

RESEARCH ARTICLE

WILEY

Hedging and utility valuation of a defaultable claim driven by Hawkes processes

Puneet Pasricha¹  | Dharmaraja Selvamuthu¹  | Paola Tardelli² 

¹Department of Mathematics, Indian Institute of Technology Delhi, New Delhi, India

²Department of Industrial and Information Engineering and Economics, University of L'Aquila, Monteluco di Roio, Italy

Correspondence

Paola Tardelli, DIIE, University of L'Aquila, Piazzale Pontieri 1, Monteluco di Roio, 67100, Italy.
Email: paola.tardelli@univaq.it

Abstract

This article studies the problem of hedging a defaultable claim via the maximization of the mean value of exponential utility, over a set of admissible strategies. The dynamics of the underlying asset is assumed to be governed by mutually exciting Hawkes processes, which captures the jumps clustering phenomenon observed in the market. The resulting market is incomplete and does not allow perfect replication. Hence, a dynamic programming approach is adopted to characterize the value function as the largest solution to a suitable backward stochastic differential equation (BSDE) with a non-Lipschitz generator. The value function of the optimal investment problem and of the indifference prices are represented in terms of limits of the sequence of value functions of suitable Lipschitz BSDEs and, further, a result of uniqueness is achieved. Finally, numerical experiments are performed to demonstrate the applicability of the proposed framework and to understand the impact of the jump-clustering on the values of the claims.

KEYWORDS

backward stochastic differential equations, default time, exponential utility, Hawkes processes, optimal investment

1 | INTRODUCTION

One of the central issues of debate, among practitioners and academics, after the subprime crisis, is the problem of pricing and hedging the contingent claims that may default before maturity. The main challenge in pricing and hedging contingent claims is to model the default event and the expected losses when a default occurs. There are two categories of models proposed in the literature to model the default event, namely structural models and reduced-form models. The class of structural models builds upon the pioneering work by Merton,¹ who gave a mechanism of default in terms of the relation between liabilities and assets at maturity time T . The structural models define the default as the time when the stochastic process describing the dynamics of total firm value falls below a certain threshold level, called the default barrier. Several authors have extended this basic model in various directions, for instance, by incorporating the jump-diffusion process or stochastic interest rates.^{2,3} On the other hand, the reduced-form models do not define the mechanism of default but consider it to be an event of surprise and set the default as the first jump time of a point process. For more details on credit risk models, interested readers can refer to Duffie and Singleton.⁴

The structural models in the literature are either based on the Poisson jump-diffusion models or Lévy processes in general. Despite their manageable mathematical properties, these models are considered to be inconsistent with the empirical observations⁵⁻⁷ reporting jump-cluster even over several days. More specifically, although Lévy processes form a general

class of models, due to their independent increment property, they do not allow for clustered jumps observed in the financial assets. Several articles (e.g., Lux and Marchesi,⁸ Mandelbrot⁹) noted that significantly large moves in the financial markets tend to be followed by additional large movements. Ait-Sahalia et al.,¹⁰ argued that “what makes financial crisis take place is typically not the initial jump, but the amplification that occurs subsequently over hours or days, and the fact that other markets become affected as well.” Moreover, this phenomenon is observed to be more prominent during periods of distress. These empirical observations suggest to relinquish the widely applied Lévy jump-diffusion models and motivate the development of the models that are capable of incorporating the default clustering and contagion effect.

On the other hand, the advent of intra-day information on financial asset price quotes and the rapidly growing empirical studies on market microstructure show that the asset prices are piece-wise constant and jump at irregularly spaced random times in reaction to trades or significant new informations. These empirical observations motivated several authors to believe that point processes may provide a more suitable framework for modeling the price process, Errais et al.¹¹ Further, even models that consider jump-diffusion trajectory processes, but do not account for the discreteness of the data may lead to wrong conclusions and may need ad-hoc methods for rounding and estimation.

This article takes a similar approach to that of Errais et al.,¹¹ and Tardelli.¹² The difference lies in the fact that, here, taking into account the aforementioned empirical observations and all financial relationships and the contagion effects affecting the financial markets that could lead to default, we assume that the dynamics of the model include the effect of mutual excitation. To this end, the price of the defaultable asset is described not only by pure jump processes, as in Tardelli,¹² but by mutually exciting Hawkes processes. The self-exciting Hawkes process is well known for its efficacy in reproducing jump clusters due to the dependence of the conditional event arrival rate, or intensity, on the counting process itself. These processes have only recently been used in problems related to financial markets. Embrechts et al.,¹³ discussed the multivariate Hawkes process and their applications to financial data. Liu and Zhu¹⁴ derived closed-form formula for price of variance swaps in a Hawkes jump-diffusion model. Pasricha and Goel¹⁵ addressed the problem of pricing power exchange options in a Hawkes jump-diffusion model. Recently, Nitu et al.¹⁶ considered Hawkes jump-diffusion model for the pricing of equity indexed annuities.

The contribution of this article to the literature is twofold. First, inspired by Errais et al.,¹¹ in this article we investigated the Markov property for the two-dimensional process consisting of a Hawkes process and its intensity. The Markov property is an important tool necessary to study the hedging problem and which allows us to generalize the idea of Tardelli.¹² In a market with only two types of assets, the money market account, and the risky asset, we discuss the problem of hedging a defaultable claim. The literature studying the issue of hedging a defaultable claim is vast, for example, Bielecki et al.¹⁷ and Jeanblanc et al.,¹⁸ who assumed the price dynamics to be a diffusion process. Moreover, the presence of Hawkes processes makes the market incomplete and a perfect replication is not possible. Thus, a hedging criterion under incompleteness has to be used, and the approach considered consists in maximizing the mean value of an exponential utility function from the terminal wealth.^{19,20} The resulting problem becomes a stochastic control problem and can be addressed using various techniques available in the literature. For an overview on stochastic control problems see the monographs Fleming and Soner,²¹ Oksendal and Sulem,²² Pham,²³ and Karoui et al.²⁴

Following the dynamic programming-based approach, the value function is characterized as the largest solution to a suitable backward stochastic differential equation (BSDE), as suggested by the procedure developed in Lim and Quenez²⁵ for a diffusion model. The maximization of the mean value of the utility function with and without considering the claim allows us to deal with the indifference valuation. But, due to the presence of the default time, the problem is strictly related to the enlargement of filtrations. More specifically, when there are two filtrations $\mathcal{F}$ and $\mathcal{G}$, such that $\mathcal{F} \subset \mathcal{G}$, as well known, a $\mathcal{G}$ -martingale is also an $\mathcal{F}$ -martingale, but the viceversa is not true, namely, an $\mathcal{F}$ -martingale does not remain a martingale with respect to the filtration $\mathcal{G}$. An exciting example is Azema's martingale, see section 4.3.8 in Jeanblanc et al.¹⁸ A classical assumption in this frame is the so-called immersion property or (H)-hypothesis, namely any $\mathcal{F}$ -local martingale must also be a $\mathcal{G}$ -local martingale, see Jeanblanc et al.¹⁸ and Bielecki et al.¹⁷ Consequently, the (H)-hypothesis is assumed in the present paper.

A further difficulty is that the value function is characterized only as the largest solution to a suitable BSDE, and the results of the uniqueness of solutions are not available because the generator does not have the Lipschitz property.²⁶ Thus, in Section 4, a uniqueness result is achieved, since we prove that the value function of optimal investment is the unique limit of a non-increasing sequence of value functions, where each elements of the sequence is obtained as the unique solution of a suitable BSDE with Lipschitz generator¹² and corresponds to the value function associated with a subset of bounded admissible strategies.

Second, the article is devoted to the results of some numerical experiments. The numerical section is an essential part of this article that are presented to demonstrate the applicability of the proposed modeling framework, and how

the inclusion of jump-clustering in the model affects the prices of defaultable claims compared to models without jumps clustering.

The remainder of the article is organized as follows. Section 2 presents the proposed framework and discusses the related hedging problem. Section 3 characterizes the hedging problem via BSDEs. Section 4 gives the construction of the non-increasing sequence converging to the value function followed by the uniqueness result. Section 5 presents the results of some numerical experiments. The conclusions and future works are in Section 6.

2 | PROPOSED FRAMEWORK AND HEDGING PROBLEM

2.1 | Mutually exciting Hawkes processes

Hawkes processes constitute a particular class of multivariate point processes, first introduced by Hawkes.²⁷ A D -variate Hawkes process is a counting-process $N_t = \{N_t^i\}_{i=1,\dots,D}$ such that the intensity vector, $\lambda_t = \{\lambda_t^{(i)}\}_{i=1,\dots,D}$ admits the representation

$$\lambda_t^{(i)} = \lambda^{(i)} + \sum_{j=1,\dots,D} \int_{-\infty}^t \mu_{ij}(t-u) dN_u^j,$$

where $\lambda^{(i)}$, $i = 1, \dots, D$ are exogenous intensities, and the so-called kernel matrix, $\mu(t) = \{\mu_{ij}(t)\}_{i,j=1,\dots,D}$, is such that, for each $1 \leq i, j \leq D$ and for $t \geq 0$, $\mu_{ij}(t) \geq 0$, otherwise, for $t < 0$, $\mu_{ij}(t) = 0$.

Furthermore, $\mu_{ij}(t)$ belongs to the space of L^1 -integrable functions and each component $\mu_{ij}(t)$ represents the effect of the jumps of N^j until time t on $\lambda^{(i)}$. There are various possible choices of the kernel $\mu_{ij}(t)$ resulting in different characteristics of the process, Bacry et al.²⁸

If $D = 1$, N_t reduces to a self exciting Hawkes process whose intensity is given by

$$\lambda_t = \lambda_0 + \int_{-\infty}^t \mu(t-u) dN_u,$$

where $\mu(t)$ represents the size of the impact of jumps. If $\mu(t) = 0$, then the process N_t turns to a Poisson process.

2.2 | The model: Representation

On a finite time horizon $T > 0$, assume that the filtered real-world probability space $(\Omega, \tilde{\mathcal{F}}, P)$ models the uncertainty in the economy. As usual, the probability space is complete and takes the right continuous version of all the filtrations considered. Let a two-dimensional mutually exciting Hawkes process (N^1, N^2) with interacting intensity $(\lambda^{(1)}, \lambda^{(2)})$, and let

$$\mathcal{F} := \{\mathcal{F}_t\}_{t \in [0, T]} \quad \mathcal{F}_t = \sigma\{N_s^1, N_s^2, 0 \leq s \leq t\},$$

with $\mathcal{F} \subset \tilde{\mathcal{F}}$. For $i = 1, 2$, assume that the intensity process of N^i admits the representation

$$\lambda_t^{(i)} = a_i + (\lambda_0^{(i)} - a_i)e^{-\delta_i t} + \sum_{j=1,2} \sum_{k=1}^{+\infty} Y_k^{(ij)} e^{-\delta_i(t-T_k^{(j)})} \mathbf{1}_{0 \leq T_k^{(j)} < t} \quad (1)$$

where, for $i = 1, 2$,

- $\lambda_0^{(i)} > 0$ is the initial value of i th intensity process at time $t = 0$.
- $a_i \geq 0$ is the constant mean-reverting level of i th intensity process.
- $\delta_i \geq 0$ is the constant mean-reverting rate of i th intensity process.
- $\{T_k^{(j)}\}_{k=1,2,\dots}, j = 1, 2$, are sequences of random variables that represent the jump times of the counting process N^j .

- $\{Y_k^{(ij)}\}_{k=1,2,\dots}, 1 \leq i, j \leq 2$, are sequences of independently and identically distributed (i.i.d.) non negative random variables with distribution function $G_{ij}(y)$. These random variables represent the sizes of self-excited jumps ($i = j$) and of mutually exciting jumps ($i \neq j$).

Observe that the events occurred before time t has an impact on the intensity process of each component, the impact being measured by the kernels. For example, $Y_{10}^{(12)}$ represents the size of the impact on the intensity of N^1 due to the jump at time $T_{10}^{(2)}$ of 10th jump of N^2 . Further, the randomness is driven by two sources of uncertainty: the time of events and the recovery at these events.

The process (N_t^1, N_t^2) with intensities given in Equation (1) admits a convenient martingale representation given below.¹¹ Note that, although Hawkes process produces clustered arrivals, it does not explode, and the integral $\int_0^t \lambda_s^{(i)} ds$ is finite for all t . In order to prove this fact, let $\delta_i = 0, i = 1, 2$. Then the intensity $\lambda^{(i)}$ is constant between the jumps with non-decreasing sample paths (in other words, N^i is a birth process), and is non-explosive and integrable. As a consequence, we deduce that N_t^i is also integrable for $\delta_i > 0$. Moreover, as $\int_0^t \lambda_s^{(i)} ds$ is integrable,

$$M_t^i = N_t^i - \int_0^t \lambda_s^{(i)} ds$$

is a martingale with respect to the filtration $\{\mathcal{F}_t\}$, and the jump times of N^1 and N^2 are $\mathcal{F}$ -stopping times (corollary 3 in chapter II.6 of Protter²⁹).

On the same probability space $(\Omega, \tilde{\mathcal{F}}, P)$, assume that τ is a strictly positive random variable with probability 1, that is, $P(\tau > 0) = 1$, that models the default time. The process $N_t^3 := \mathbf{1}_{\{\tau \leq t\}}$ is the “default indicating process” and let

$$\mathcal{G} := \{\mathcal{G}_t\}_{t \in [0, T]} \quad \mathcal{G}_t = \sigma\{N_s^1, N_s^2, N_s^3, 0 \leq s \leq t\},$$

with $\mathcal{G} \subset \tilde{\mathcal{F}}$. Analogously to Equation (1), assume the intensity of the process N^3 admits the representation

$$\lambda_t^{(3)} = a_3 + (\lambda_0^{(3)} - a_3)e^{-\delta_3 t} + \sum_{i=1,2} \sum_{k=1}^{+\infty} Y_k^{(3i)} e^{-\delta_3(t-T_k^{(i)})} \mathbf{1}_{0 \leq T_k^{(i)} < t}. \quad (2)$$

The random variables $\{Y_k^{3i}\}_{k=1,2,\dots}, i = 1, 2$, represent the effect of jumps of N^i on the default intensity $\lambda_t^{(3)}$. Assume that $\{Y_k^{3i}\}_{k=1,2,\dots}, i = 1, 2$ is a sequence of i.i.d. non-negative random variables with distribution functions $G_{3i}(y)$, representing the size of mutually exciting jumps. Recalling that $\mathcal{F}$ is the σ -algebra generated by N^1 and N^2 , while $\mathcal{G}$ is the σ -algebra generated by N^1, N^2 , and N^3 , observe that the process

$$M_t^3 = N_t^3 - \int_0^{t \wedge \tau} \lambda_s^{(3)} ds,$$

is a $\mathcal{G}_t$ -martingale. This means that τ is a totally inaccessible $\mathcal{G}$ -stopping time, while the jump times of N^1 and N^2 are $\mathcal{G}$ -totally inaccessible $\mathcal{F}$ -stopping time. Hence, N^1, N^2 , and N^3 do not have common jump times. This last condition introduces a difficulty. In order to overcome this difficulty, we assume the so-called immersion property or (H)-hypothesis in what follows. Note that, Remark 2 below uses this property as a main tool.

Assumption 1. The filtration $\mathcal{F}$ is said to be immersed in the filtration $\mathcal{G}$, or it is said to satisfy the (H)-hypothesis under the measure P , whenever any $(P, \mathcal{F}_t)$ -local martingale is also a $(P, \mathcal{G}_t)$ -local martingale.

Since the mutually and self exciting terms are defined by exponential kernels, one can show that the process $(\lambda^{(1)}, N^1, \lambda^{(2)}, N^2, \lambda^{(3)}, N^3)$ is a Markov process, even if the process $N = (N^1, N^2, N^3)$ itself does not have the Markov property (refer Errais et al.¹¹ for more details on Markov property of Hawkes process). Next, we present a proposition which provides a formula for the conditional expectation of a function f of the Markov process

$$\xi = (\lambda^{(1)}, N^1, \lambda^{(2)}, N^2, \lambda^{(3)}, N^3),$$

in terms of its infinitesimal generator, that describes the dynamics of the process itself. Further, the Dynkin formula, derived in this proposition, allows us to get an alternative approach to understand the Markov properties of all the processes involved in the model. The proof of Proposition 1 below is in Appendix.

Proposition 1. Setting $f_{\lambda^{(i)}} = \frac{\partial f}{\partial \lambda^{(i)}}$, $i = 1, 2, 3$. If, for $t \leq T$,

$$\mathbb{E} \left[\int_0^t \left| \mathcal{A}f(\xi_s) + \sum_{i=1}^3 \delta_i(\lambda_s^{(i)} - a_i) f_{\lambda^{(i)}}(\xi_s) \right| ds \right] < \infty.$$

Then, the following Dynkin formula holds

$$\mathbb{E} \left[f(\xi_T) \middle| \mathcal{G}_t \right] = f(\xi_t) + \mathbb{E} \left[\int_t^T \mathcal{A}f(\xi_s) \middle| \mathcal{G}_t \right]. \quad (3)$$

Under all the assumptions made so far, given a suitable function $f(\lambda^{(1)}, n_1, \lambda^{(2)}, n_2, \lambda^{(3)}, n_3, t)$ with continuous partial derivatives with respect to $\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)}$, and t , the infinitesimal generator of the process $(\lambda^{(1)}, N^1, \lambda^{(2)}, N^2, \lambda^{(3)}, N^3)$ is given by

$$\mathcal{A}f = \frac{\partial f}{\partial t} - \sum_{i=1}^3 \delta_i(\lambda^{(i)} - a_i) \frac{\partial f}{\partial \lambda^{(i)}} + \lambda^{(1)} \mathcal{A}^{(1)}f + \lambda^{(2)} \mathcal{A}^{(2)}f + \lambda^{(3)}(1 - n_3) \mathcal{A}^{(3)}f, \quad (4)$$

where, for $i = 1, 2$,

$$\mathcal{A}^{(i)}f = \int_0^\infty \int_0^\infty \int_0^\infty a_i^f(y) dG_{1i}(y^{(1i)}) dG_{2i}(y^{(2i)}) dG_{3i}(y^{(3i)}) - f(\lambda^{(1)}, n_1, \lambda^{(2)}, n_2, \lambda^{(3)}, n_3, t), \quad (5)$$

$$a_1^f(y) = f(\lambda^{(1)} + y^{(11)}, n_1 + 1, \lambda^{(2)} + y^{(21)}, n_2, \lambda^{(3)} + y^{(31)}, n_3, t),$$

$$a_2^f(y) = f(\lambda^{(1)} + y^{(12)}, n_1, \lambda^{(2)} + y^{(22)}, n_2 + 1, \lambda^{(3)} + y^{(32)}, n_3, t),$$

and

$$\mathcal{A}^{(3)}f = f(\lambda^{(1)}, n_1, \lambda^{(2)}, n_2, \lambda^{(3)}, n_3 + 1, t) - f(\lambda^{(1)}, n_1, \lambda^{(2)}, n_2, \lambda^{(3)}, n_3, t).$$

Remark 1. For simplicity, we take $Y^{(11)}$ and $Y^{(22)}$ in Equation (1) as null to eliminate the self exciting part which implies a mean-reversion type structure. Further, we consider three types of jumps: a jump following the upward momentum, a jump following a downward momentum and the default. More specifically, we assume that after an upward jump, the intensity of downward jump increases and similarly, after a downward jump, the intensity of an upward jump increases. Furthermore, to be in a realistic setting, we assume $Y^{(31)}$ equal to zero in Equation (2), since a positive jump in the price does not impact the default intensity. On the other hand, a negative jump in the price increases the intensity of default, the phenomenon known as downward momentum.

Furthermore, if we take the jump size to be constant in time and by choosing $a_i = \lambda_0^{(i)}$, for $i = 1, 2$, the intensity function of N^i is given by

$$\lambda_t^{(i)} = \lambda_0^{(i)} + \sum_{j=1}^2 \int_0^t \alpha_{ij} e^{-\delta_i(t-u)} dN_u^j. \quad (6)$$

The constants $\lambda_0^{(i)}$, $i = 1, 2$, are initial intensities, and $\mu(t) = \{\mu_{ij}(t)\}_{i,j=1,2}$, where $\mu_{ij}(t) = \alpha_{ij} e^{-\delta_i(t)} \mathbf{1}_{t \geq 0}$. Again, each component $\mu_{ij}(t)$ is non negative, belongs to the space of L^1 -integrable functions and represents the effect of the jumps of N^j , till time t , on λ^i . Note that the constants α_{ij} represents the jump sizes in the intensity process.

Setting $a_3 = \lambda_0^{(3)}$, the intensity of default indicator process is

$$\lambda_t^{(3)} = \lambda_0^{(3)} + \int_0^t \alpha_{32} e^{-\delta_3(t-u)} dN_u^2, \quad (7)$$

where $\mu_{32}(t) = \alpha_{32} e^{-\delta_3(t)} \mathbf{1}_{t \geq 0}$ are functions representing the effect of downward jumps of N^2 in the price, on the intensity $\lambda_t^{(3)}$.

In this case, the generator given in Equation (4) is simplified and $\mathcal{A}_i f$, $i = 1, 2$, given by Equation (5), are replaced by

$$\begin{aligned}\mathcal{A}_1 f &= f(\lambda^{(1)}, n_1 + 1, \lambda^{(2)} + \alpha_{(21)}, n_2, \lambda^{(3)}, n_3, t) - f(\lambda^{(1)}, n_1, \lambda^{(2)}, n_2, \lambda^{(3)}, n_3, t), \\ \mathcal{A}_2 f &= f(\lambda^{(1)} + \alpha_{(12)}, n_1, \lambda^{(2)}, n_2 + 1, \lambda^{(3)} + \alpha_{(32)}, n_3, t) - f(\lambda^{(1)}, n_1, \lambda^{(2)}, n_2, \lambda^{(3)}, n_3, t).\end{aligned}$$

2.3 | Dynamics of the risky asset

Consider a financial market with one risky asset and one risk-free asset. The price of the risk-free asset is assumed to be 1, that is, the risk-free interest rate is assumed to be zero. The price S of the risky asset, discounted with respect to the price of the bond, is modeled as a pure jump process, such that, for $S_0 \in \mathbb{R}^+$,

$$S_t = S_0 \exp\{Y_t\}. \quad (8)$$

The log return price Y_t is assumed to be a non-explosive process, taking real values, with initial condition $Y_0 = 0$ and the dynamics is given by

$$Y_t := \int_0^t \eta_u^1 dN_u^1 - \int_0^t \eta_u^2 dN_u^2 + \int_0^t \beta_u dN_u^3,$$

where N^1 , N^2 , and N^3 are Hawkes processes defined in the previous subsection. The motivation behind the choice of a pure jump process is as follows: the availability of intra-day information on the price quotes of financial assets, in which changes in price take place only at some irregularly spaced random times, leads us to believe that pure jump processes are more suitable to model the observed prices. Particularly, we consider mutually exciting Hawkes process which can take into account the contagion effects that affect the financial markets. The jump sizes η^1 , η^2 , and β are $\mathcal{G}_t$ -predictable processes, such that η^1 and η^2 are bounded from above and from below by two positive real constants $\underline{\eta}$, $\bar{\eta}$, that is, for $i = 1, 2$, $\underline{\eta} \leq \eta_t^i \leq \bar{\eta}$.

In the literature, many authors assume that $\beta_\tau > -1$. According to Jeanblanc et al.³⁰ and Lim and Quenez,²⁵ this condition is equivalent to $\beta_t > -1$ for any $0 \leq t \leq T$ a.s., which implies that the price process $S_\tau = S_{\tau-}(1 + \beta_\tau)$, is still positive after the default τ . However, since S is exponential, $S_\tau = S_{\tau-}e^{\beta_\tau}$ is positive, and the assumption on β is not necessary.

Remark 2. Setting

$$c_u := (e^{\eta_u^1} - 1)\lambda_u^{(1)} + (e^{-\eta_u^2} - 1)\lambda_u^{(2)} + (e^{\beta_u} - 1)\gamma_u(1 - N_u^3),$$

by a standard application of Itô formula, the representation of the price process as $\mathcal{G}_t$ -local semi-martingale, is given by

$$S_t = S_0 + \int_0^t S_u c_u du + M_t^S,$$

where M_t^S is a $\mathcal{G}_t$ -local martingale represented as

$$M_t^S = \int_0^t S_{u-}(e^{\eta_u^1} - 1)dM_u^1 + \int_0^t S_{u-}(e^{-\eta_u^2} - 1)dM_u^2 + \int_0^t S_{u-}(e^{\beta_u} - 1)dM_u^3. \quad (9)$$

Note that H-hypothesis is a necessary condition to obtain Equation (9). In order to have that the $\mathcal{G}_t$ -local martingales M_t^S , M^1 , and M^2 must be not only $\mathcal{F}_t$ -local martingales but also $\mathcal{G}_t$ -local martingales.

In the proposed framework, the investors are in an incomplete financial market and can trade in a finite time window $[0, T]$, by investing in risky stocks and a risk-free bond. Further, there also exists a default time in the market. We formally define a generic defaultable European contingent claim with maturity date T (refer Bielecki et al.¹⁷) as follows:

Definition 1. A defaultable contingent claim with maturity date T consists a triplet (X, Z, τ) where:

- (i) The default τ is a random variable that specifies the time of default, for $t \in [0, T]$. It is always assumed that τ is strictly positive with probability 1, that is, $P(\tau > 0) = 1$.

- (ii) The promised payoff X represents the random payoff received by the owner of the claim at maturity T , if there was no default prior to or at time T . This is a $\mathcal{G}_T$ -measurable random variable such that $0 \leq X \leq \bar{B}$, for a positive real constant $\bar{B}$. The actual payoff at T associated with X thus equals X if $\tau \geq T$.
- (iii) The $\mathcal{G}$ -adapted recovery process Z specifies the recovery payoff at time of default. This is the quantity received by the owner of a claim at the time of default (or at maturity), provided that the default occurred prior to or at maturity date T . We assume the following $0 \leq Z \leq \bar{B}$.

In practice, the hedging of a derivative after default is usually of minor interest.¹⁷ For instance, in a model with a single default, hedging after the default reduces to the replication of a non-defaultable claim and can be addressed following Tardelli.³¹

2.4 | Hedging problem and replication

In this subsection, we present the hedging problem. A $\mathcal{G}_t$ -predictable real-valued process θ_t is called a trading strategy, if it is S -integrable and self-financing. Let Θ denotes the class of such strategies. At any $t \in [0, T]$, θ_t describes the number of shares held in the risky asset. Noting that $\mathcal{G}_t$ is the σ -algebra generated by the point processes N^1 , N^2 , and N^3 until the time t , following Bremaud,³² we observe that any $\mathcal{G}_t$ -predictable process has the following structure,

$$\theta_t = \sum_{n \geq 0} \theta_n(t_1, t_2, \dots, t_n, t) \mathbf{1}_{t_n < t \leq t_{n+1}},$$

where $\{t_n\}_{n \geq 0}$ are the jump times of $N = (N^1, N^2, N^3)$ and θ_n are real-valued measurable functions. One can observe that any admissible strategy is completely characterized by the families $\{\theta_n\}_{n \geq 0}$ and $\{t_n\}_{n \geq 0}$.

For a strategy $\theta \in \Theta$ and for initial capital w_0 , a real positive constant, the associated wealth process is defined as

$$\begin{aligned} W_t^\theta &= w_0 + \int_0^t \theta_r dS_r \\ &= w_0 + \int_0^t \theta_r S_{r-} (e^{\eta_r^1} - 1) dN_r^1 + \int_0^t \theta_r S_{r-} (e^{-\eta_r^2} - 1) dN_r^2 + \int_0^t \theta_r S_{r-} (e^{\beta_r} - 1) dN_r^3. \end{aligned} \quad (10)$$

After receiving the premium, the seller of the contract has to hedge against his/her exposure to the risk. A hedging problem consists of finding an optimal investment strategy in order to reduce (or avoid) potential losses arising from having to honor the contract. A perfect replication of a defaultable claim is described by the following definition.

Definition 2. A perfect replication of a defaultable claim (X, Z, τ) is given by a strategy $\theta \in \Theta$, if the wealth process W_t^θ satisfies the following hedging conditions

$$W_T^\theta \mathbf{1}_{T < \tau} = X \mathbf{1}_{T < \tau} \quad \text{and} \quad W_\tau^\theta \mathbf{1}_{T \geq \tau} = Z_\tau \mathbf{1}_{T \geq \tau}.$$

Since, the incompleteness of market doesn't allow for perfect replication, we consider maximizing the expected utility function of the wealth on a random time interval $[0, T \wedge \tau]$. The expected utility of his/her final wealth gives him/her a measure of the quality of a self-financing strategy. At any $t \in [0, T]$, the agent invests the quantity θ_t in the risky traded assets. Thus, the investment process θ_t controls the dynamics of the wealth process, Equation (10). Mathematically, the problem can be stated as maximize, over $\theta \in \Theta$,

$$\mathbb{E}[u(W_{T \wedge \tau}^\theta - X(1 - N_T^3) - Z_\tau N_T^3)].$$

The utility function $u : \mathbb{R} \rightarrow \mathbb{R}$ is assumed to be strictly concave increasing differentiable and satisfies the Inada conditions,

$$u'(-\infty) := \lim_{x \rightarrow -\infty} u'(x) = +\infty \quad u'(+\infty) := \lim_{x \rightarrow +\infty} u'(x) = 0.$$

Note that, since $I_{t \leq \tau} = 1 - N_{t-}^3$, then

$$\begin{aligned} W_{T \wedge \tau}^\theta &= w_0 + \int_0^T (1 - N_{r-}^3) \theta_r S_{r-} (e^{\eta_r^1} - 1) dN_r^1 + \int_0^T (1 - N_{r-}^3) \theta_r S_{r-} (e^{-\eta_r^2} - 1) dN_r^2 \\ &\quad + \int_0^T (1 - N_{r-}^3) \theta_r S_{r-} (e^{\beta_r} - 1) dN_r^3. \end{aligned}$$

Let w be the amount of capital at time t and Θ_t denote the set of the admissible strategies on the interval $[t, T]$. Define

$$B_T(\tau) = X(1 - N_T^3) + Z_\tau N_T^3.$$

In this article, we consider an agent with exponential utility function $u_\alpha(x) = 1 - e^{-\alpha x}$, where $\alpha \in \mathbb{R}^+$ is the risk aversion parameter. Since $\xi = (\lambda^{(1)}, N^1, \lambda^{(2)}, N^2, \lambda^{(3)}, N^3)$ satisfies the Markov property, we represent the associated value process as follows:

$$V_t(w) = \text{ess sup}_{\theta \in \Theta_t} \mathbb{E} \left[1 - \exp \left\{ -\alpha (W_{T \wedge \tau}^\theta - W_{t \wedge \tau}^\theta + w - B_T(\tau)) \right\} \middle| \xi_t \right].$$

Assumption 2. The set of admissible trading strategies Θ consists of all $\mathcal{G}$ -predictable processes $\theta = \{\theta_t\}_{0 \leq t \leq T}$, such that for any θ fixed and any $s \in [0, T]$, there exists a real constant $K_{s,\theta}$ such that $W_{t \wedge \tau}^\theta - W_{s \wedge \tau}^\theta \geq -K_{s,\theta}$, $s \leq t \leq T$ a.s., Lim and Quenez.²⁵

As a consequence, $\mathbb{E} \left[\exp \left\{ -\alpha \int_t^T (1 - N_{r-}^3) \theta_r dS_r \right\} \right] < \infty$ and the value process reduces to

$$V_t(w) = 1 - e^{-\alpha w} V_t, \quad (11)$$

where

$$\begin{aligned} V_t &= \text{ess inf}_{\theta \in \Theta_t} \mathbb{E} \left[\exp \left\{ -\alpha (W_{T \wedge \tau}^\theta - W_{t \wedge \tau}^\theta - B_T(\tau)) \right\} \middle| \xi_t \right] \\ &= \text{ess inf}_{\theta \in \Theta_t} \mathbb{E} \left[\exp \left\{ -\alpha \left(\int_t^T (1 - N_{r-}^3) \theta_r dS_r - B_T(\tau) \right) \right\} \middle| \xi_t \right]. \end{aligned} \quad (12)$$

The wealth process is such that

$$W_{T \wedge \tau}^\theta - W_{t \wedge \tau}^\theta = \sum_{N(t) \leq k \leq N(T)-1} (1 - N_{t_k}^3) \theta_{t_k} (S_{t_{k+1}} - S_{t_k}),$$

where, for $k \geq 1$,

$$\theta_{t_k} = \theta_{k-1}(t_1, t_2, \dots, t_{k-1}, t_k),$$

then a representation of the value function is

$$V_t = \sum_{N(t) \leq k \leq N(T)-1} V_{t_k} \mathbf{1}_{[t_k, t_{k+1})}(t).$$

2.5 | Properties of the value function

The dynamic programming approach allows us to characterize the value function, and, more precisely, the process V_t defined in Equation (12), as the largest solution to a suitable BSDE.^{12,25,31}

Proposition 2. For $t \in [0, T \wedge \tau]$, the following holds true

- (i) The process V_t is strictly positive and bounded. Moreover $V_T = \mathbb{E} \left[e^{\alpha B_T(\tau)} \middle| \xi_T \right]$.

- (ii) For any $\theta \in \Theta$, the process $V_t \exp\{-\alpha W_{t \wedge \tau}^\theta\}$ is a $\mathcal{G}_t$ -submartingale.
 (iii) V_t is the largest process $\mathcal{G}_t$ -adapted verifying (i) such that

$$V_T = \mathbb{E} \left[e^{\alpha B_T(\tau)} \middle| \xi_T \right].$$

- (iv) The process $\theta^* \in \Theta$ is an optimal strategy if and only if $V_t \exp\{-\alpha W_{t \wedge \tau}^{\theta^*}\}$ is a $\mathcal{G}_t$ -martingale.
 (v) The process V_t admits an indistinguishable $\mathcal{G}$ -adapted cadlag representation.

All these properties can be proved on similar lines to the proof outlined in Tardelli.¹²

3 | BACKWARD STOCHASTIC DIFFERENTIAL EQUATION

As a consequence of Proposition 2, V_t is bounded, strictly positive and, for a vanishing strategy, it is a $\mathcal{G}_t$ -submartingale. Then, its Doob–Meyer decomposition is

$$V_t = V_0 + \int_0^t dM_r^V + \int_0^t A_r^V dr,$$

where the process M_t^V is a square integrable $\mathcal{G}_t$ -martingale and the process A_t^V is an increasing $\mathcal{G}_t$ -predictable process such that $A_0^V = 0$. Furthermore, a classical representation of M_t^V allows us to write

$$\begin{aligned} dM_r^V &= R_r^1 dM_r^1 + R_r^2 dM_r^2 + R_r^3 dM_r^3, \\ A_r^V &= \left[R_r^1 \lambda_r^{(1)} + R_r^2 \lambda_r^{(2)} + R_r^3 \gamma_r (1 - N_r^3) \right], \end{aligned} \quad (13)$$

where R_r^1 , R_r^2 , and R_r^3 are measurable $\mathcal{G}_t$ -predictable processes such that

$$\mathbb{E} \left\{ \int_0^t \left[(R_r^1)^2 \lambda_r^{(1)} + (R_r^2)^2 \lambda_r^{(2)} + (R_r^3)^2 \gamma_r (1 - N_r^3) \right] dr \right\} < +\infty.$$

Theorem 1. For $t \in [0, T \wedge \tau]$, setting

$$K_r^\theta(x) = \exp \left\{ -\alpha(1 - N_{r-}^3) \theta_r S_{r-} (e^x - 1) \right\} - 1, \quad (14)$$

let

$$K_r^{\theta,1} = K_r^\theta(\eta^1), \quad K_r^{\theta,2} = K_r^\theta(-\eta^2), \quad K_r^{\theta,3} = K_r^\theta(\beta).$$

Then, the process $(V_t, R_t^1, R_t^2, R_t^3)$ verifies the BSDE

$$V_t = \mathbb{E} \left[e^{\alpha B_T(\tau)} \middle| \xi_T \right] - \int_t^T dM_r^V + \int_t^T \operatorname{ess} \inf_{\theta \in \Theta_r} Q_r(\theta) dr \quad (15)$$

where, for $\lambda^{(3)} = \gamma(1 - N^3)$,

$$Q_r(\theta) = \sum_{i=1,2,3} (V_r + R_r^i) K_r^{\theta,i} \lambda_r^{(i)}.$$

Furthermore, V_t is the largest solution to the BSDE given in Equation (15), and (R_t^1, R_t^2, R_t^3) is uniquely determined by the martingale representation theorem.

The proof of Theorem 1 is in Appendix.

3.1 | Indifference price

In this subsection, we introduce the writer's indifference price, which is the price p of the claim such that the agent is indifferent between optimizing the expected utility with and without the derivative at hand (refer Becherer,³³ Frittelli,³⁴ Mania and Santacrose³⁵ and references therein). To define indifference price, consider the case in which the investor wants to maximize the expected utility of his terminal wealth *without* considering the claim. Thus, the value process turns to be $V_t^0(w) = 1 - e^{-\alpha w} V_t^0$ and recalling that Θ_t denotes the set of the admissible strategies on the time interval $[t, T]$,

$$V_t^0 = \operatorname{ess\,inf}_{\theta \in \Theta_t} \mathbb{E} \left[\exp \left\{ -\alpha \int_t^T (1 - N_{r-}^3) \theta_r dS_r \right\} \middle| \xi_t \right]. \quad (16)$$

All the properties deduced in Proposition 2 for the function V_t hold also for V_t^0 , taking into account that $B_T(T) = 0$. The Doob–Meyer decomposition of V_t^0 allows us to write that

$$V_t^0 = V_0^0 + \int_0^t dM_r^{V^0} + \int_0^t A_r^{V^0} dr$$

where $M_t^{V^0}$ is a square integrable $\mathcal{G}_t$ -martingale, $A_t^{V^0}$ is an increasing $\mathcal{G}_t$ -predictable process and $A_0^{V^0} = 0$. Again, a classical representation of $M_t^{V^0}$ allows us to write

$$\begin{aligned} dM_r^{V^0} &= R_r^{0,1} dM_r^1 + R_r^{0,2} dM_r^2 + R_r^{0,3} dM_r^3, \\ A_r^{V^0} &= \left[R_r^{0,1} \lambda_r^{(1)} + R_r^{0,2} \lambda_r^{(2)} + R_r^{0,3} \gamma_r (1 - N_r^3) \right], \end{aligned}$$

where $R_r^{0,1}$, $R_r^{0,2}$, and $R_r^{0,3}$ are measurable $\mathcal{G}_t$ -predictable processes such that

$$\mathbb{E} \left\{ \int_0^t \left[(R_r^{0,1})^2 \lambda_r^{(1)} + (R_r^{0,2})^2 \lambda_r^{(2)} + (R_r^{0,3})^2 \gamma_r (1 - N_r^3) \right] dr \right\} < +\infty.$$

Finally, analogously to Theorem 1, the next results are provided.

Theorem 2. *The process $(V_t^0, R_t^{0,1}, R_t^{0,2}, R_t^{0,3})$ verifies the BSDE*

$$V_t^0 = 1 - \int_t^T dM_r^{V^0} + \int_t^T \operatorname{ess\,inf}_{\theta \in \Theta_r} Q_r^0(\theta) dr$$

where, taking into account Equation (14), $Q_r^0(\theta) = \sum_{i=1,2,3} \left(V_r^0 + R_r^{0,i} \right) K_r^{\theta,i} \lambda_r^{(i)}$.

Furthermore, V_t^0 is the largest solution to Equation (15) and $(R_t^{0,1}, R_t^{0,2}, R_t^{0,3})$ is uniquely determined by the martingale representation theorem.

Therefore, the indifference price, the price at which the agent is indifferent between optimizing the expected utility without the contingent claim, or optimizing the expected utility including the payoff derivative $B_T(\tau)$, at time T , with the compensation p_t , is implicitly given by the following equation

$$V_t(w + p_t) = V_t^0(w). \quad (17)$$

$V_t(w)$ and $V_t^0(w)$ are defined in Equations (11), (12), and (16), respectively, that is,

$$V_t^0(w) = \operatorname{ess\,sup}_{\theta \in \Theta_t} \mathbb{E} \left[1 - \exp \left\{ -\alpha \left(w + \int_t^T (1 - N_{r-}^3) \theta_r dS_r \right) \right\} \middle| \xi_t \right] = 1 - e^{-\alpha w} V_t^0,$$

and

$$V_t(w) = \operatorname{ess\,sup}_{\theta \in \Theta_t} \mathbb{E} \left[1 - \exp \left\{ -\alpha \left(w + \int_t^T (1 - N_{r-}^3) \theta_r dS_r - B_T(\tau) \right) \right\} \middle| \xi_t \right] = 1 - e^{-\alpha w} V_t.$$

Thus, Equation (17) becomes

$$1 - e^{-\alpha(w+p_t)} V_t = 1 - e^{-\alpha w} V_t^0,$$

which leads to the following indifference price

$$p_t = \frac{1}{\alpha} \log \frac{V_t}{V_t^0}.$$

4 | UNIQUENESS RESULT: CONVERGING SEQUENCE

As already observed in the previous sections, the process V_t (or V_t^0) is characterized as the “largest” solution of a BSDE. Since this BSDE has a non-Lipschitz generator, the results of the uniqueness of solutions are not available (see Carbone et al.²⁶).

In order to prove a uniqueness result, in this section, we construct a non-increasing sequence $\{V_t^k\}_{k \geq 1}$ of value-functions, where, for each $k \geq 1$, V_t^k is the unique solution to a BSDE with Lipschitz generator. Then, we prove that V_t^k converges to V_t .^{12,25} Until now, in the previous sections, we have not made any assumption on β , but now a technical assumption is needed.

Assumption 3. The process β_t is uniformly bounded, for any $t \in [0, T]$.

Definition 3. For $k \geq 1$, let $\Theta^k = \{\theta \in \Theta : |\theta_u S_u| \leq k, \forall u \in [0, T], a.s.\}$ and let Θ_t^k be the set of the strategies in Θ^k on $[t, T \wedge \tau]$.

Further, for each $k \geq 1$, let V_t^k be defined as

$$V_t^k = \text{ess} \inf_{\theta \in \Theta_t^k} \mathbb{E} \left[1 - \exp \left\{ -\alpha (W_{T \wedge \tau}^\theta - W_{t \wedge \tau}^\theta - B_T(\tau)) \right\} \middle| \xi_t \right]$$

Remark 3. Let us note that Θ^k is a bounded set and that the argument of the expectation has finite mean value. This implies that Lebesgue theorem applies and, consequently, there exists an optimal strategy $\theta_k^* \in \Theta^k$.¹² As a consequence, all the results of Proposition 2 hold for V_t^k , which means that V_t^k is a bounded and strictly positive process and, for a vanishing strategy, it is a $\mathcal{G}_t$ -submartingale. Thus, its Doob–Meyer decomposition is given by

$$V_t^k = V_0^k + \int_0^t dM_r^{V^k} + \int_0^t A_r^{V^k} dr,$$

with $M_t^{V^k}$ square integrable $\mathcal{G}_t$ -martingale, $A_t^{V^k}$ increasing $\mathcal{G}_t$ -predictable process such that $A_0^{V^k} = 0$. Moreover, we have

$$dM_r^{V^k} = \sum_{i=1,2,3} R_r^{i,k} dM_r^i, \quad A_r^{V^k} = \sum_{i=1,2,3} R_r^{i,k} \lambda_r^{(i)},$$

where $R_r^{1,k}$, $R_r^{2,k}$, and $R_r^{3,k}$ are measurable $\mathcal{G}_t$ -predictable processes such that

$$\mathbb{E} \left\{ \int_0^t \sum_{i=1,2,3} (R_r^{i,k})^2 \lambda_r^{(i)} dr \right\} < +\infty.$$

Proposition 3. The process (V_t^k, R_t^k) , with $R_t^k = (R_t^{1,k}, R_t^{2,k}, R_t^{3,k})$, is the unique positive cadlag $\mathcal{G}$ -adapted solution of the following BSDE with Lipschitz continuous generator

$$V_t^k = \mathbb{E} \left[e^{\alpha B_T(\tau)} \middle| \xi_T \right] - \int_t^T dM_r^{V^k} + \int_t^T \text{ess} \inf_{\theta \in \Theta_r^k} Q_r^k(\theta) dr, \quad (18)$$

where $Q_r^k(\theta) = \sum_{i=1,2,3} (V_r^k + R_r^{i,k}) K_r^{\theta,i} \lambda_r^{(i)}$ and $K_r^\theta(x)$ is defined in Equation (14).

The proof of Proposition 3 is in Appendix.

Theorem 3. Under Assumption 3, $V_t = \lim_{k \rightarrow +\infty} V_t^k$ a.s., $\forall t \in [0, T]$.

The proof of Theorem 3 is in Appendix.

In Proposition 3, we prove that V_t^k is the unique solution of the BSDE given in Equation (18), then Theorem 3 allows us to have the required uniqueness result for the process V_t .

5 | NUMERICAL ILLUSTRATIONS

In this section, the results of numerical experiments are presented to study the significance of mutually exciting jumps in comparison to Poisson jump process. We consider a derivative with payoff

$$B_T = (S_T - K)^+ \mathbf{1}_{\tau \geq T} + \tilde{w}(S_\tau - K)^+ \mathbf{1}_{\tau < T} \quad (19)$$

where τ denotes the default time, that is, the jump time of the process N_t^3 , $\tilde{w}$ denotes the recovery rate and K is the strike price. Our aim is to price the vulnerable derivative with payoff given in Equation (19). For the calculation of the indifference price of the derivative given in Equation (19), we have utilized the recursive backward scheme developed in Tardelli,¹² when the underlying asset S_t follows the dynamics given in Equation (8), with intensity of the jump processes discussed in Equations (6) and (7). Before a meaningful discussion on the empirical significance of the Hawkes process is carried out, it is important to validate the recursive scheme of Tardelli¹² in the present framework. Furthermore, this validation helps us to verify that the results developed in the previous sections are free from any algebraic errors and are ready to implement. For this purpose, we compare the results obtained through applying the recursive scheme to the results obtained through Monte Carlo simulations to calculate the following expectation

$$\mathbb{E}(B_T) = \mathbb{E} \left((S_T - K)^+ \mathbf{1}_{\tau \geq T} + \tilde{w}(S_\tau - K)^+ \mathbf{1}_{\tau < T} \right). \quad (20)$$

In the following, unless otherwise stated, the values of the parameters (with changing one of the parameters in sensitivity analysis)¹ chosen to achieve our goals which mentioned above are listed as follows:

1. The initial intensities are chosen such that $\lambda_0^{(1)} = \lambda_0^{(2)} = 1$ and $\lambda_0^{(3)} = 0.01$,
2. the mean-reverting speeds are assumed to be $\delta_1 = \delta_2 = \delta_3 = \delta = 2$,
3. the jump-sizes in the intensities are $\alpha_{11} = \alpha_{12} = \alpha_{21} = \alpha_{22} = \alpha_{32} = \alpha_0 = 0.1$
4. the recovery rate $\tilde{w}$ is assumed to be equal to 0.4,
5. the risk-aversion coefficient $\alpha = 0.5$,
6. the maturity time $T = 1$,
7. the initial price $S_0 = 100$ and the strike price $K = 95$,
8. the positive and negative jumps in the asset price are 0.02 and -0.02 respectively, while the stock price is assumed to jump to $0.6 * S_0$ upon default, that is, $\beta_u = \beta = -0.4$.

To simulate the mutually exciting Hawkes process, we follow the simulation method developed in Ogata,³⁶ where the authors proposed a simulation method by distributing the accepted points to each dimension with probabilities proportional to their intensities. Figure 1A presents the comparison of the indifference price obtained through the recursive backward scheme and the expected value of Equation (20), obtained through Monte Carlo simulations. A total of 100,000 sample paths are generated for the experiments. From Figure 1A, note that the prices obtained through both the approaches match up perfectly across all the maturity times, which validates the accuracy of the recursive scheme. This further confirms that the results derived in this article are free from any algebraic errors.

With the confidence in our results and the recursive scheme, we now investigate the impact of introducing clustering of jumps in the modeling framework on the derivative price. For this purpose, we present a comparison of the derivative prices obtained in our modeling framework to the prices obtained the model with only self-exciting jumps

¹These particular values of parameters are selected just for illustration purposes. These values are chosen keeping in mind the following stability conditions for the Hawkes processes, (i) $0 < \alpha_{11} + \alpha_{12} < \theta_1$, $0 < \alpha_{21} + \alpha_{22} < \theta_2$, and $0 < \alpha_{32} < \theta_3$; (ii) $\lambda_0^{(1)}, \lambda_0^{(2)}, \lambda_0^{(3)} \geq 0$.

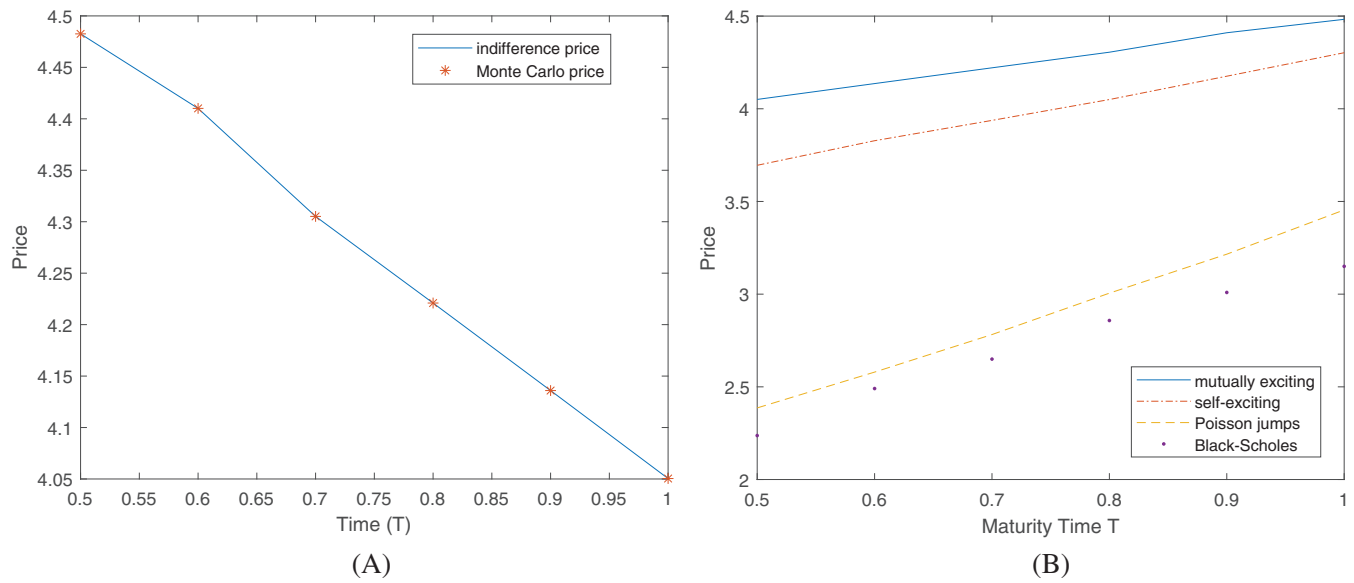


FIGURE 1 The comparison of our results with other alternatives. (A) with Monte Carlo; (B) with Poisson process

(i.e., $\alpha_{12} = \alpha_{21} = \alpha_{32} = 0$), the model with Poisson jumps (i.e., $\alpha_0 = 0$) and the standard Black-Scholes diffusion model. Note that for the illustration purposes, we have assumed that the jump sizes α_{ij} are all equal to α_0 . Note that we have the following expressions for expected number of jumps per unit time:

$$E(N_t^1) = \bar{\lambda}_1 t, \quad E(N_t^2) = \bar{\lambda}_2 t, \quad E(N_t^3) = \bar{\lambda}_3 t,$$

where

$$\bar{\lambda}_1 = \frac{\lambda_0^{(1)}}{1 - \frac{\alpha_{11} + \alpha_{12}}{\delta_1}}, \quad \bar{\lambda}_2 = \frac{\lambda_0^{(2)}}{1 - \frac{\alpha_{21} + \alpha_{22}}{\delta_1}}, \quad \bar{\lambda}_3 = \frac{\lambda_0^{(3)}}{1 - \frac{\alpha_{32}}{\delta_3}}$$

For comparison to other models, we require a consistency of the expected number of jumps in unit time, we choose the intensities of the jump processes as follows:

- For self-exciting jumps model, we have the following intensities for $i = 1, 2$,

$$\lambda_{SE,t}^{(i)} = \lambda_{SE,0}^{(i)} + \int_0^t \alpha_i^{SE} e^{-\delta_i^{SE}(t-u)} dN_{SE,u}^i. \quad (21)$$

and

$$\lambda_{SE,t}^{(3)} = \lambda_{SE,0}^{(3)}. \quad (22)$$

For self-exciting model, we have

$$E(N_{SE,t}^1) = \frac{\lambda_{SE,0}^{(1)} t}{1 - \frac{\alpha_1^{SE}}{\delta_1^{SE}}}, \quad E(N_{SE,t}^2) = \frac{\lambda_{SE,0}^{(2)} t}{1 - \frac{\alpha_2^{SE}}{\delta_2^{SE}}}, \quad E(N_{SE,t}^3) = \lambda_{SE,0}^{(3)} t.$$

Therefore, we select the parameters $\alpha_i^{SE}, \delta_i^{SE}, i = 1, 2$ and $\lambda_{SE,0}^{(i)}, i = 1, 2, 3$ such that

$$E(N_{SE,t}^1) = E(N_t^1), \quad E(N_{SE,t}^2) = E(N_t^2), \quad E(N_{SE,t}^3) = E(N_t^3),$$

which further results into the following choice

$$\delta_1^{SE} = \delta_1, \quad \delta_2^{SE} = \delta_2, \quad \alpha_1^{SE} = \alpha_{11} + \alpha_{12}, \quad \alpha_2^{SE} = \alpha_{21} + \alpha_{22}, \quad \lambda_{SE,0}^{(3)} = \frac{\lambda_0^{(3)}}{1 - \frac{\alpha_{32}}{\delta_3}}.$$

- For Poisson jumps model, we have the constant intensities for each of the three jump-processes $N_{PP,t}^i$, $i = 1, 2, 3$ and are selected such that

$$E(N_{PP,t}^1) = E(N_t^1), \quad E(N_{PP,t}^2) = E(N_t^2), \quad E(N_{PP,t}^3) = E(N_t^3),$$

which results into the following intensities

$$\lambda^{1,PP} = \bar{\lambda}_1, \quad \lambda^{2,PP} = \bar{\lambda}_2, \quad \lambda^{3,PP} = \bar{\lambda}_3.$$

Finally, we also compare our proposed model with the standard Black-Scholes (BS) model given by

$$dS_t = \sigma S_t dW_t, \quad (\because r = 0).$$

For consistency, we select σ such that $E^Q(\ln(S_t))$ is same for Black-Scholes model and our model for any time $t \in [0, T]$. For our model, we have

$$E^Q(\ln(S_t)) = \eta_1 \bar{\lambda}_1 t - \eta_2 \bar{\lambda}_2 t + \eta_3 \bar{\lambda}_3 t,$$

and for BS model, we have $E^Q(\ln(S_t)) = -\frac{1}{2}\sigma^2 t$ which implies the following choice $\sigma^2 = -2(\eta_1 \bar{\lambda}_1 - \eta_2 \bar{\lambda}_2 + \eta_3 \bar{\lambda}_3)$.

The results are presented in Figure 1B for different maturity times. First, among the four models, it can be found that the inclusion of jumps significantly impact the derivative price, which is revealed by the comparison of Black-Scholes price with the price obtained through jump models. Second, these prices under the four models are almost parallel to each other under the chosen set of parameters. Finally, the model with mutually exciting jumps leads to the highest derivative price. The contribution of clustering of jumps is revealed by comparing the results with self-exciting and Poisson jumps model. Not surprisingly, one can observe that the idea to include clustering of jumps has a significant impact on the price of the derivative, an observation that is also reported in the literature for derivative pricing Liu and Zhu¹⁴ and Pasricha and Goel.¹⁵ The inclusion of self-exciting jumps increases the price which is further increased if we also take into account the mutually-exciting jumps. Besides, the difference in the prices become more and more pronounced as the risk horizon increases, which leads to the conclusion that the effects of jump clustering become more significant at longer time horizon in comparison to shorter time horizon. Thus, the inclusion of the mutually exciting jumps can contribute to a large difference in the derivative prices, and thus it is useful to incorporate them to form a more accurate model.

Recall that for a Hawkes process, the clustering of jumps increases with increase in the value of the jump size α_0 and initial intensity λ_0 , while it decreases with increase in the value of the mean reversion speed δ . Further, from Figure 1B, we observed that clustering has a significant impact, more specifically an increasing effect, on the derivative price. It is therefore interesting to observe how each of these parameters that governs the clustering phenomenon impacts the derivative price. For these purposes, we present the sensitivity of the prices with respect to the parameters governing the clustering phenomenon, namely mean-reverting rate δ_i , initial intensity $\lambda_0^{(i)}$ and α_{ij} , in Figures 2 and 3.

As we know that the sensitivity analyses are based on varying one of the parameters while other parameters are kept unchanged from the original values assigned to them at the beginning of this section. Therefore, the range of the parameter being analyzed should be varied in an acceptable interval such that the processes are well defined. More specifically, $\lambda_0^{(i)} \geq 0$ should be non-negative and we choose the interval such that it covers the long-term means, that is, $\frac{\lambda_0^{(1)}}{1 - \frac{\alpha_{11} + \alpha_{12}}{\delta_1}} = 1.11$, $\frac{\lambda_0^{(2)}}{1 - \frac{\alpha_{21} + \alpha_{22}}{\delta_2}} = 1.11$ and $\frac{\lambda_0^{(3)}}{1 - \frac{\alpha_{32}}{\delta_3}} = 0.0105$. Similarly, the intervals of δ_i and α_{ij} are selected such that the following stability condition for each of the Hawkes processes holds true $0 < \alpha_{11} + \alpha_{12} < \theta_1$, $0 < \alpha_{21} + \alpha_{22} < \theta_2$ and $0 < \alpha_{32} < \theta_3$.

Figure 2A, allows us to observe that the price of derivative decreases with increase in the value of δ , which means that a decrease in the jump-clustering effect decreases the price, although the impact is almost linear. This observation is further confirmed by Figures 2B and 3C,D. In Figure 2B, assuming first that $\alpha_{ij} = \alpha_0$, the sensitivity analysis with respect

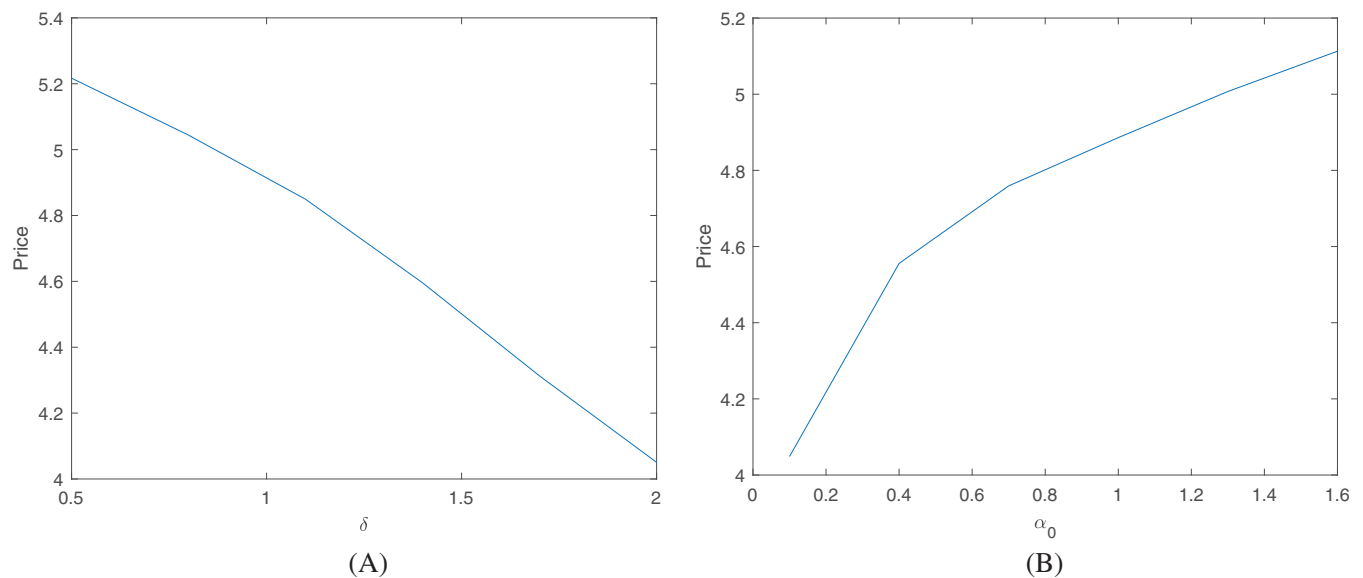


FIGURE 2 Sensitivity of prices with respect to parameters governing jump clustering namely δ the mean reverting rate and the size of jumps in the intensity α_0 . (A) Against δ ; (B) against α_0

to α_0 was performed. Then, the prices against the self-exciting jump size α_{11} and against the mutually exciting jump-size α_{12} is plotted to understand the role of each α_{ij} individually, and, because of the symmetry, the results for α_{21} and α_{22} are the same. By the observations of the figures, we get that the price increases with increasing values of α_{ij} , that is, with increasing values of jump clustering, which is as we expected. Finally, we present the results against the initial intensities of the Hawkes processes, that is, $\lambda_0^{(1)}$ and $\lambda_0^{(3)}$. Similar conclusions can be drawn for $\lambda_0^{(2)}$, due to symmetry with $\lambda_0^{(1)}$ and keeping in view that it induces negative jumps in the prices. Note that the prices increases against $\lambda_0^{(1)}$ while it decreases against $\lambda_0^{(3)}$. The reason is that a higher value of $\lambda_0^{(1)}$ implies a higher probability of having more number of positive jumps in the price of underlying asset, consequently, higher price of the derivative. On the other hand, if the value of $\lambda_0^{(3)}$ increases, this implies a higher probability of having a jump in N_t^3 and a higher probability of default. Thus, the price will be lower because no one would pay higher premium if he/she knows that there are higher chances of default in future. Thereby, the jump clustering has a significant impact on the price. This means that, not only the self-exciting processes but also those of mutual exciting should be taken into account.

6 | CONCLUSIONS AND FUTURE WORKS

In a complete friction-less market, as well known, the risks can be hedged and a unique price can be obtained. While in an incomplete market, the risks cannot be hedged away completely, and, therefore, a range of prices is obtained. In order to get a particular price for the contingent claim, some assumptions about the risk preferences of the trader have to be formulated. The utility-indifference pricing method allows us to derive such a particular price.

In this article, a pure jump model for a defaultable asset is introduced, but taking into account the mutually exciting effects of the market. Thus, the jump process follows a mutually exciting Hawkes process. A discussion on the problem of hedging a defaultable claim written on this asset is presented. The key idea is first to prove the Markov property and then to maximize the expected value of an exponential utility, and to adopt a dynamic programming approach, to characterize the value function as largest solution of a suitable BSDE. Finally, the results of some numerical experiments are presented and investigated to illustrate the significance of mutually exciting jumps and the applicability of our proposed model.

It would be interesting to study the impact of adopting a mutually-exciting Hawkes process on indifference price through experiments on the real data. We plan to implement that in future. Further, taking into account that the study of the complexity of the financial models and all financial relationships requires deep mathematical approaches, such as stochastic differential equations, an interesting open problem is to investigate the stochastic stability of the model proposed in this article. More specifically, this means studying asymptotic stability in various sense, for example, stability

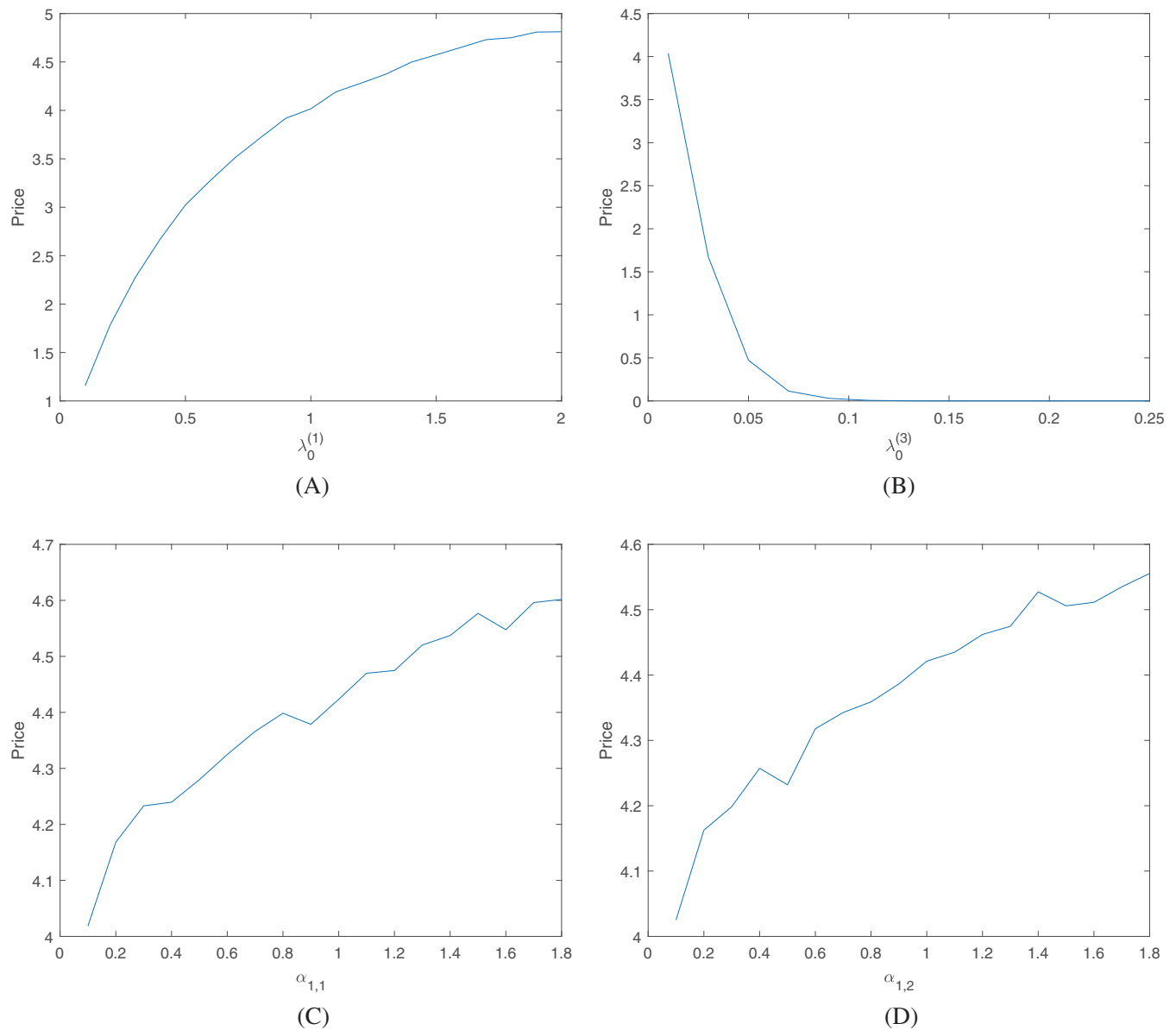


FIGURE 3 Sensitivity of prices with respect to $\lambda_0^{(1)}$, $\lambda_0^{(3)}$, α_{11} , and α_{12} . (A) Against $\lambda_0^{(1)}$; (B) against $\lambda_0^{(3)}$; (C) against α_{11} ; (D) against α_{12}

in mean, asymptotic mean-square stability, asymptotic stability almost surely and exponential p-stability. Furthermore, an interesting task could be to investigate stability for vector of stochastic differential delay equations with jumps and Markov switchings for financial models when the underlying dynamics are driven by Hawkes processes (see Swishchuk and Kazmerchuk³⁷ and Swishchuk et al.³⁸ for similar studies based on Poisson jumps).

ACKNOWLEDGMENT

We thank the editor and the anonymous reviewers for their insightful comments, which helped in improving the article. This research work is supported by the Department of Science and Technology, India.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

ORCID

Puneet Pasricha  <https://orcid.org/0000-0001-8195-3329>

Dharmaraja Selvamuthu  <https://orcid.org/0000-0003-2892-0864>

Paola Tardelli  <https://orcid.org/0000-0003-0611-7536>

REFERENCES

1. Merton RC. On the pricing of corporate debt: the risk structure of interest rates. *J Financ.* 1974;29:449-470.
2. Leland HE. Corporate debt value, bond covenants, and optimal capital structure. *J Financ.* 1994;49(4):1213-1252.
3. Zhou C. The term structure of credit spreads with jump risk. *J Bank Financ.* 2001;25(11):2015-2040. doi:10.1016/S0378-4266(00)00168-0
4. Duffie D, Singleton KJ. *Credit Risk: Pricing, Measurement, and Management*. Princeton Series in Finance. Princeton University Press; 2012.
5. Yu J. Empirical characteristic function estimation and its applications. *Econ Rev.* 2004;23(2):93-123. doi:10.1081/ETC-120039605
6. Carr P, Wu L. Time-changed levy processes and option pricing. *J Financ Econ.* 2004;71:113-141.
7. Maheu JM, McCurdy TH. News arrival, jump dynamics, and volatility components for individual stock returns. *J Financ.* 2004;59(2):755-793.
8. Lux T, Marchesi M. Volatility clustering in financial markets: a microsimulation of interacting agents. *Int J Theor Appl Financ.* 2000;3:675-702.
9. Mandelbrot B. The variation of certain speculative prices. *J Bus.* 1963;36:394-419.
10. Ait-Sahalia Y, Cacho-Diaz J, Laeven RJ. Modeling financial contagion using mutually exciting jump processes. *J Financ Econ.* 2015;117(3):585-606.
11. Errais E, Giesecke K, Goldberg LR. Affine point processes and portfolio credit risk. *SIAM J Financ Math.* 2010;1(1):642-665. doi:10.1137/090771272
12. Tardelli P. Recursive backward scheme for the solution of a BSDE with a non-lipschitz generator. *Probab Eng Inf Sci.* 2017;31(2):1-19.
13. Embrechts P, Liniger T, Lin L. Multivariate hawkes processes: an application to financial data. *J Appl Probab.* 2011;48(A):367-378.
14. Liu W, Zhu SP. Pricing variance swaps under the hawkes jump-diffusion process. *J Futur Mark.* 2019;39(6):635-655.
15. Pasricha P, Goel A. Pricing power exchange options with Hawkes jump-diffusion processes. *J Ind Manag Optim.* 2021;17(1):133-149.
16. Sharma N, Pasricha P, Dharmaraja S. Valuation of equity-indexed annuities under correlated jump-diffusion processes. *J Comput Appl Math.* 2021;395:113575. doi:10.1016/j.cam.2021.113575
17. Bielecki TR, Jeanblanc M, Rutkowski M. Hedging of credit derivatives in models with totally unexpected default. *Stochastic Processes and Applications to Mathematical Finance*. World Scientific; 2006:35-100. doi:10.1142/9789812774637_0002
18. Jeanblanc M, Yor M, Chesney M. *Mathematical Methods for Financial Markets*. Springer Finance. Springer-Verlag; 2009.
19. Hu Y, Imkeller P, Muller M. Utility maximization in incomplete markets. *Ann Appl Probab.* 2005;15(3):1691-1712.
20. Mania M, Schweizer M. Dynamic exponential utility indifference valuation. *Ann Appl Probab.* 2005;15(3):2113-2143.
21. Fleming W, Soner M. *Controlled Markov Processes and Viscosity Solutions*. Springer Science & Business Media 25. 2nd ed. Springer Verlag; 2006.
22. Oksendal B, Sulem A. *Applied Stochastic Control of Jump-diffusion Processes*. Universitext Berlin Heidelberg: Springer Verlag; 2007. doi:10.1007/978-3-540-69826-5
23. Pham H. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Stochastic Modelling and Applied Probability. Berlin Heidelberg: Springer-Verlag; 2009. doi:10.1007/978-3-540-89500-8
24. El Karoui N, Hamadene S, Matoussi A. BSDEs and applications. In: Carmona R, ed. *Indifference Pricing: Theory and Applications*. Princeton University Press; 2008:267-320.
25. Lim T, Quenez MC. Exponential utility maximization in an incomplete market with defaults. *Electron J Probab.* 2011;16(53):1434-1464. doi:10.1214/EJP.v16-918
26. Carbone R, Ferrario B, Santacroce M. Backward stochastic differential equations driven by Cadlag Martingales. *Theory Probab Appl.* 2008;52(2):304-314. doi:10.1137/S0040585X97983055
27. Hawkes A. Spectra of some self-exciting and mutually exciting point processes. *Biometrika.* 1971;58(1):83-90. doi:10.1093/biomet/58.1.83
28. Bacry E, Mastromatteo I, Muzy JF. Hawkes processes in finance. *Market Microstruct Liq.* 2015;1(01):1550005. doi:10.1142/S2382626615500057
29. Protter P. *Stochastic Integration and Differential Equations*. Stochastic Modelling and Applied Probability 2. Berlin Heidelberg: Springer, Berlin, Heidelberg; 2005. doi:10.1007/978-3-662-10061-5
30. Jeanblanc M, Matoussi A, Ngoupeyou A. Robust utility maximization in a discontinuous filtration. Preprint, University of Evry-Val d'Essonne, Evry, ArXiv e-prints; 2012.
31. Tardelli P. Utility maximization in a pure jump model with partial observation. *Probab Eng Inf Sci.* 2011;25(1):29-54.
32. Bremaud P. *Point Processes and Queues. Martingale Dynamics*. Springer Series in Statistics. Springer-Verlag; 1981.
33. Becherer D. Bounded solutions to backward SDE's with jumps for utility optimization and indifference hedging. *Ann Appl Probab.* 2006;16(4):2027-2054.
34. Frittelli M. Introduction to a theory of value Coherent with no-arbitrage principle. *Finance Stoch.* 2000;4(3):275-297. doi:10.1007/s007800050074
35. Mania M, Santacroce M. Exponential utility maximization under partial information. *Finance Stoch.* 2010;14(3):419-448. doi:10.1007/s00780-009-0114-z
36. Ogata Y. On Lewis' simulation method for point processes. *IEEE Trans Inf Theory.* 1981;27(1):23-31. doi:10.1109/TIT.1981.1056305

37. Swishchuk AV, Kazmerchuk YI. Stability of stochastic differential delay equations of Ito form with jumps and Markovian switchings, and their application in finance. *Theory Probab Math Stat.* 2002;64:167-178.
38. Swishchuk AV, Swishchuk M, Limnios N. Stochastic stability of vector SDE with applications to a stochastic epidemic model. *Int J Pure Appl Math.* 2016;106(3):801-839. doi:10.12732/ijpam.v106i3.9
39. Schachermayer W. Optimal investment in incomplete markets when wealth may become negative. *Ann Appl Probab.* 2001;11(3):694-734.

How to cite this article: Pasricha P, Selvamuthu D, Tardelli P. Hedging and utility valuation of a defaultable claim driven by Hawkes processes. *Appl Stochastic Models Bus Ind.* 2022;38(2):334-352. doi: 10.1002/asmb.2663

APPENDIX

Proof. Proof of Proposition 1.

The process $\xi = (\lambda^{(1)}, N^1, \lambda^{(2)}, N^2, \lambda^{(3)}, N^3)$ has right-continuous paths of finite variation. Hence,

$$f(\xi_t) - f(\xi_0) = \int_0^t \sum_{i=1}^3 \delta_i(\lambda_s^{(i)} - a_i) f_{\lambda^{(i)}}(\xi_s) ds + \sum_{0 < s \leq t} [f(\xi_s) - f(\xi_{s-})].$$

The last sum describes the jumps of ξ and it can be rewritten as

$$\sum_{i=1,2} \int_0^t \mathcal{A}_{s-}^{(i)} f \left(dM_s^i + \lambda_s^{(i)} ds \right) + \int_0^t \mathcal{A}_{s-}^{(3)} f \left(dM_s^3 + (1 - N_{s-}^3) \lambda_s^{(3)} ds \right).$$

The integrability condition on the predictable integrand guarantees that the quantity $\sum_{i=1,2,3} \int_0^t \mathcal{A}_{s-}^{(i)} f dM_s^i$ is a martingale (theorem 8, chapter II of Bremaud³²). As a conclusion, we get that $f(\xi) = f(\lambda^{(1)}, N^1, \lambda^{(2)}, N^2, \lambda^{(3)}, N^3)$ is a semi-martingale having a unique decomposition given by a sum of a predictable process with finite variation and a martingale as follows:

$$f(\xi_t) - f(\xi_0) = \int_0^t \mathcal{A} f(\xi_s) ds + \sum_{i=1,2,3} \int_0^t \mathcal{A}_{s-}^{(i)} f dM_s^i.$$

Since $\sum_{i=1,2,3} \int_0^t \mathcal{A}_{s-}^{(i)} f dM_s^i$ is a martingale, so is the process defined by

$$f(\xi_t) - f(\xi_0) - \int_0^t \mathcal{A} f(\xi_s) ds$$

which results to Equation (3). ■

Proof. Proof of Theorem 1.

Setting $C_t^\theta := e^{-\alpha(W_{t \wedge \tau}^\theta - W_0)}$, Itô formula provides

$$C_t^\theta = 1 + \sum_{i=1,2,3} \left\{ \int_0^t C_{r-}^\theta K_r^{\theta,i} dM_r^i + \int_0^t C_r^\theta K_r^{\theta,i} \lambda_r^{(i)} dr \right\}.$$

Taking into account that

$$\begin{aligned} \int_0^t C_{s-}^\theta dV_s &= \sum_{i=1,2,3} \left\{ \int_0^t C_{s-}^\theta R_s^i dM_s^i + \int_0^t C_s^\theta R_s^i \lambda_s^{(i)} ds \right\}, \\ \int_0^t V_{s-} dC_s^\theta &= \sum_{i=1,2,3} \left\{ \int_0^t V_{s-} C_{s-}^\theta K_s^{\theta,i} dM_s^i + \int_0^t V_s C_s^\theta K_s^{\theta,i} \lambda_s^{(i)} ds \right\}, \\ [C^\theta, V]_t &= \sum_{i=1,2,3} \left[\int_0^t C_{s-}^\theta K_s^{\theta,i} R_s^i dM_s^i + \int_0^t C_s^\theta K_s^{\theta,i} R_s^i \lambda_s^{(i)} ds \right]. \end{aligned}$$

For the product formula with V_t , by Equation (13) and since $C_0^\theta V_0 = 1 \cdot V_0 = V_0$, we have

$$\begin{aligned} C_t^\theta V_t &= C_0^\theta V_0 + \int_0^t C_{s-}^\theta dV_s + \int_0^t V_{s-} dC_s^\theta + [C^\theta, V]_t \\ &= V_0 + \sum_{i=1,2,3} \int_0^t C_{s-}^\theta \left[R_s^i + (V_{s-} + R_s^i) K_s^{\theta,i} \right] dM_s^i + \int_0^t C_s^\theta [Q_s(\theta) + A_s^V] ds. \end{aligned}$$

Since $C_t^\theta V_t = V_t e^{-\alpha(W_{t \wedge \tau}^\theta - w_0)}$ is a $\mathcal{G}_t$ -submartingale, the bounded variation term has to be increasing for any strategy and it has to vanish for the optimal strategy, which implies that

$$A_s^V = -\text{ess} \inf_{\theta \in \Theta_s} Q_s(\theta),$$

where the existence of the optimal strategy is guaranteed by theorem 2.2 given in Schachermayer.³⁹ Finally, the last assertion is a consequence of Proposition 2 by observing that

$$V_t = V_T - \int_t^T dM_r^V - \int_t^T A_r^V dr. \quad \blacksquare$$

Proof. Proof of Proposition 3.

By Remark 3, following the same line of Theorem 1, the process (V_t^k, R_t^k) is solution of Equation (18). To get the uniqueness of the solution of BSDE given in Equation (18), for $r = (r_1, r_2, r_3)$, let

$$f(t, v, r) := \text{ess} \inf_{\theta \in \Theta_t^k} \sum_{i=1,2,3} (v + r_i) K_t^{\theta,i} \lambda_t^{(i)}.$$

Note that Θ^k is bounded set and the coefficients $\lambda^{(1)}$, $\lambda^{(2)}$, γ , and K_t^θ are bounded by Assumption 3. This implies that $f(t, v, r)$ is an infimum of linear terms with respect to v and r , with uniformly bounded coefficients. In other words, f is a Lipschitz function with respect to v and r . Thus, the uniqueness follows using the results of Carbone et al.²⁶ $\blacksquare$

Proof. Proof of Theorem 3.

The proof of Theorem 3 follows from the next proposition. $\blacksquare$

Proposition 4. For $t \in [0, T]$

(i) There exists a cadlag $\mathcal{G}$ -submartingale $\bar{V}_t$, such that

$$\lim_{k \rightarrow +\infty} V_t^k = \bar{V}_t \geq V_t \quad \text{a.s.}$$

(ii) For each $\theta \in \bar{\Theta}$, where $\bar{\Theta}$ is the set of the essentially bounded admissible strategies, $e^{-\alpha(W_{t \wedge \tau}^\theta - w_0)} \bar{V}_t$ is a $\mathcal{G}$ -submartingale.

(iii) For each $t \in [0, T]$, then $V_t \geq \bar{V}_t$ a.s.

The proof of Proposition 4 can be achieved working on similar lines to the results presented in Tardelli.¹²