

TIME-CONSISTENT EVALUATION OF CREDIT RISK WITH CONTAGION

John-John Ketelbuters, Donatien Hainaut

REPRINT | 2021 / 45

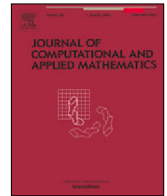
ISBA

Voie du Roman Pays 20 - L1.04.01

B-1348 Louvain-la-Neuve

Email : lidam-library@uclouvain.be

<https://uclouvain.be/en/research-institutes/lidam/isba/publication.html>



Time-consistent evaluation of credit risk with contagion

John-John Ketelbuters*, Donatien Hainaut

LIDAM, UCLouvain, Belgium

ARTICLE INFO

Article history:

Received 30 October 2020

Received in revised form 16 September 2021

Keywords:

Credit risk

Self-exciting processes

Time-consistency

ABSTRACT

A time-consistent evaluation is a dynamic pricing method according to which a risk that will be almost surely cheaper than another one at a future date should already be cheaper today. Common actuarial pricing approaches are usually not time-consistent. Pelsser and Ghalehjooghi (2016) derived time-consistent valuation principles from time-inconsistent ones. The aim of this paper is twofold. Firstly, we propose a model for credit insurance portfolios taking into account the contagion risk via self-exciting jump processes. Secondly, we extend the approach of Pelsser and Ghalehjooghi to credit insurance in this framework. Starting from classical time-inconsistent actuarial pricing methods, we derive partial integro-differential equations (PIDE) for their time-consistent counterparts. We discuss numerical methods for solving these PIDEs and their results. We draw two conclusions from these results. On the one hand, we show that time-consistent evaluations tend to give higher prices, compared to time-inconsistent evaluations. On the other hand, our results show that the time-consistency of evaluations allows to better take into account the risk of contagion in credit insurance, if such a risk exists. Finally, we propose a method to calibrate our model and use it in practice.

© 2021 Elsevier B.V. All rights reserved.

0. Introduction

Insurance markets are incomplete by nature since most of actuarial risks are not hedgeable with financial instruments. Choosing an appropriate valuation method is therefore a challenging task. This problem was initially tackled in a static way with the concepts of evaluations or risk measures. We refer the reader to e.g. [1] for an introduction about coherent risk measures and [2] for examples. Unlike insurance pricing, financial pricing is treated in a dynamic way, focused on hedging and replicating portfolios, see for example [3]. Delbaen and Schachermayer [4] show that this intrinsically dynamic approach can be applied with various securities, as long as the asset prices are bounded semimartingales. Dynamic evaluation methods extend this dynamic nature of financial pricing to insurance pricing. This type of evaluations, as well as the concept of time-consistency (or dynamic-consistency), was introduced by Riedel [5] in discrete time. Rosazza Gianin [6] uses g -expectations to extend the concept of time-consistency to continuous time. Pelsser and Stadje [7] introduce the concept of two step market evaluations, that allows to apply time-consistent evaluations on products that incorporate both a financial risk and an actuarial risk. Pelsser and Ghalehjooghi [8] use the definition of time-consistency in order to turn time-inconsistent classical evaluation principles into time-consistent ones. They derive partial differential equations (PDE) that characterize time-consistent evaluations.

Credit risk for transactions involving large quoted corporations can be hedged with derivatives like credit default swaps. Nevertheless financial markets do not propose similar covers for small or medium enterprises (SME). For this category

* Correspondence to: Voie du Roman Pays 20, 1348 Louvain-la-Neuve, Belgium.

E-mail address: john-john.ketelbuters@uclouvain.be (J.-J. Ketelbuters).

of companies, the only solution to limit credit exposure is provided by trade credit insurances. Insurers proposing these contracts rely on diversification to hedge themselves but are exposed to contagion between defaults. This motivates us to propose a time-consistent valuation framework managing both default and contagion risks.

The two main paradigms for credit risk valuation are the structural and reduced approaches. In the first one, the balance sheet of the firm value is explicitly modeled and the default occurs when the asset value falls below liabilities. For details, we refer the reader e.g. to [9–12]. The second approach is called the reduced-form or intensity-based credit model. The default is in this case triggered by the first jump of a point process. The reduced-form was initiated by the work of Duffie and Singleton [13]. Textbooks concerning this second kind of models are [14–18]. In this article, we opt for a reduced-form approach.

An endogenous way to model the contagion between defaults of companies consists in using self-exciting point processes. In this approach, any jump of this process triggers an increase of the instantaneous expected number of future jumps. [19–21] were among the first to introduce such point processes. Since then, Hawkes processes have been used to model earthquake occurrences, see e.g. [22] or [23], or terrorist activity, see [24]. They are also used for modeling financial transactions, see e.g. [25–29], or for credit risk in [30]. For a survey of properties and other possible applications of Hawkes processes, the reader is referred to [31].

The contributions of this article are organized as follows. Firstly, we introduce a new model based on self-exciting processes in which the default intensity of each company is an affine function of a common market intensity. We provide closed-form expressions for the moments of this intensity process. Secondly we find semi-closed form expressions for the Laplace transform of the intensity process. We use next this Laplace transform to compute probabilities of default. Thirdly, we extend the approach of Pelsser and Ghalehjooghi [8] to our setting in two directions. On the one hand, we evaluate in a time-consistent manner credit insurance portfolio. On the other hand, we take into account the risk of contagion. We obtain partial integro-differential equations (PIDE) for the time-consistent values and solve them numerically. Fourthly, we provide a numerical method for solving the PIDE. This method solves PIDEs for which we miss some of the boundary conditions. With the help of this method, we present a numerical illustration and analyze the impacts of contagion risk and time-consistency on the value of a credit insurance product. Finally, we propose a method to calibrate our model.

1. The model

In the intensity approach to credit risk, defaults of companies are modeled by jumps of point processes. A jump of a point process corresponds to the default of a company. In this paper, we are interested in the defaults observed in a credit insurance portfolio. The defaults correspond to the jumps of a point process $(P_t)_{t \geq 0}$. The losses induced by the defaults are modeled by the sizes of the jumps of $(P_t)_{t \geq 0}$. We start by describing precisely how the process $(P_t)_{t \geq 0}$ is built.

Let $(N_t)_{t \geq 0}$ be a point process on a filtered probability space $(\Omega, \mathbb{F}, P)$ where $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ is the natural filtration of all the processes to be described in this section. We denote the intensity of $(N_t)_{t \geq 0}$ by $(\lambda_t)_{t \geq 0}$, another stochastic process. The intensity represents the instantaneous probability to observe a jump in the process $(N_t)_{t \geq 0}$. Let $\xi_1, \xi_2, \dots$ be independent identically distributed positive square integrable random variables distributed according to the probability measure ν . We will often drop the subscript and write simply ξ for a ν -distributed random variable. We assume that the Laplace transform of ν exists and we denote it by $\tilde{\nu}$. That is, the integral $\tilde{\nu}(z) := \int_{\mathbb{R}^+} e^{-zx} \nu(dx)$ exists at least for all z in a neighborhood of 0. We also assume that the values of the Laplace transform $\tilde{\nu}$, as well as the probabilities $P(a \leq \xi < b) = \nu([a, b))$ for any real numbers $a \leq b$, can be obtained, at least numerically. The same assumption holds for the first and second moments of ν . Define the process $(P_t)_{t \geq 0}$ as

$$P_t = \sum_{k=1}^{N_t} \xi_k.$$

It is a pure positive jump process whose jump intensity is the stochastic process $(\lambda_t)_{t \geq 0}$. The dynamics of this intensity process is assumed to satisfy the stochastic differential equation (SDE)

$$d\lambda_t = \kappa(\theta - \lambda_t)dt + \eta dP_t + \sqrt{\lambda_t} \sigma dW_t \quad (1.1)$$

where $(W_t)_{t \geq 0}$ is a standard Brownian motion, $\kappa, \theta, \eta > 0$, $\sigma \geq 0$ and $\lambda_0 \in [\theta, +\infty]$. As a consequence of SDE (1.1), the intensity depends upon the history of the process $(P_t)_{t \geq 0}$. This allows the process $(P_t)_{t \geq 0}$ to exhibit a self-exciting behavior: a jump of the process $(P_t)_{t \geq 0}$ induces a jump in the intensity $(\lambda_t)_{t \geq 0}$, which in turn implies an increase of the probability to observe another jump in $(P_t)_{t \geq 0}$. The presence of $\sqrt{\lambda_t}$ in the term $\sqrt{\lambda_t} \sigma dW_t$ at SDE (1.1) ensures that $\lambda_t \geq 0$ almost surely for any $t > 0$. As early as λ_t gets close to 0, the noise term $\sqrt{\lambda_t} \sigma dW_t$ gets close to zero, so that λ_t cannot fall below 0. The parameter θ represents a level towards which the intensity tends to revert. The parameter κ in Eq. (1.1) represents the speed of the reversion towards θ and σ is a volatility parameter.

Since the jumps of $(P_t)_{t \geq 0}$ correspond to defaults, our model introduces a phenomenon of contagion between defaults: a default of a company in the portfolio triggers an increase of the intensity of defaults, thereby increasing the chances to observe other defaults in the portfolio. The self-exciting behavior of $(\lambda_t, P_t)_{t \geq 0}$ thus aims to model the clustering of defaults observed in practice, see e.g. [32]. It is important to note that the SDE (1.1) implies that the impact of past defaults dissipates at an exponential speed through time, with a rate κ . This is more visible at Eq. (2.3), where $(\lambda_t)_{t \geq 0}$ is

expressed in integral form. We observe at the last term of this equation that a default of cost dP_s at time $s \geq 0$ results in the term $\eta e^{-\kappa(t-s)} dP_s$ at time $t \geq s$ so that defaults are forgotten at an exponential speed. This property is actually crucial in the use of stochastic calculus, and therefore in the mathematical developments throughout this paper.

We close this section with a discussion about what differentiates our model from the one of Pelsser and Ghalehjooghi [8]. The main difference is the fact that the jump process used in [8] is a constant intensity compound Poisson process that is completely exogenous. More precisely, the jump diffusion process is

$$dy_t = a(t, y_t)dt + b(t, y_t)dW_t + dJ_t$$

with $J_t = \sum_{i=1}^{N_t} \xi_i$ and $(N_t)_{t \geq 0}$ a Poisson process with constant intensity $\lambda > 0$. The process $(J_t)_{t \geq 0}$ is thus a compound Poisson process with stochastic jumps $\xi_1, \xi_2, \dots$. This setting could be adapted to the modeling of a credit portfolio by taking $(y_t)_{t \geq 0}$ as the intensity of a Cox process $(D_t)_{t \geq 0}$ (i.e. a doubly stochastic Poisson process, it is a Poisson process with a stochastic intensity) whose jumps represent the defaults in the credit portfolio. In such a setting, the jumps of $(N_t)_{t \geq 0}$ represent defaults that are exogenous to the credit portfolio, whereas the jumps of $(D_t)_{t \geq 0}$ represent defaults that occur inside the credit portfolio (endogenous defaults). The difference with the model we propose in this paper is the fact that the jumps of $(D_t)_{t \geq 0}$ would have no impact on its stochastic intensity $(y_t)_{t \geq 0}$, so that no phenomenon of contagion can exist inside the credit portfolio. In other words, in the jump model of Pelsser and Ghalehjooghi [8], the occurrence of defaults in the credit portfolio can only be affected by exogenous defaults, but not by endogenous ones.

2. Expectation and variance of the λ process

This section is concerned with the computation of the expectation and variance of the intensity process described in Eq. (1.1). These computations will allow us to derive stability conditions for our model, that is, $\lim_{t \rightarrow +\infty} (\mathbb{E}[\lambda_t] \vee \mathbb{V}\text{ar}(\lambda_t)) < +\infty$.

Proposition 2.1. Assume that $\kappa > \frac{\eta}{\rho}$. For any $t \geq 0$, we have

$$\mathbb{E}[\lambda_t] = e^{a_3 t} \lambda_0 + \frac{\kappa \theta}{a_3} (e^{a_3 t} - 1) \quad (2.1)$$

and

$$\mathbb{V}\text{ar}(\lambda_t) = \frac{(e^{a_3 t} - 1)(2a_1 \lambda_0 e^{a_3 t} + a_2 (e^{a_3 t} - 1))}{2a_3} \quad (2.2)$$

where $a_1 = \sigma^2 + 2\eta^2 \mathbb{V}\text{ar}(\xi)$, $a_2 = \frac{\kappa \theta (\sigma^2 + 2\eta^2 \mathbb{V}\text{ar}(\xi))}{\eta \mathbb{E}[\xi] - \kappa}$ and $a_3 = \eta \mathbb{E}[\xi] - \kappa$.

Proof. First note that SDE (1.1) implies that

$$\lambda_t = \theta + e^{-\kappa t} (\lambda_0 - \theta) + \sigma \int_0^t e^{-\kappa(t-s)} \sqrt{\lambda_s} dW_s + \eta \int_0^t e^{-\kappa(t-s)} dP_s. \quad (2.3)$$

By taking the expectation and differentiating with respect to t , we obtain the ordinary differential equation (ODE)

$$\frac{\partial \mathbb{E}[\lambda_t]}{\partial t} = \kappa(\theta - \mathbb{E}[\lambda_t]) + \eta \mathbb{E}[\xi] \mathbb{E}[\lambda_t]$$

whose solution is well given by Eq. (2.1). For the variance, using Ito's lemma yields

$$\frac{\partial \mathbb{E}[\lambda_t^2]}{\partial t} = \mathbb{E}[\lambda_t] (2\kappa\theta + \sigma^2 + 2\eta^2 \mathbb{V}\text{ar}(\xi)) + 2\mathbb{E}[\lambda_t^2] (\eta \mathbb{E}[\xi] - \kappa). \quad (2.4)$$

By the chain rule, we find

$$\frac{\partial}{\partial t} (\mathbb{E}[\lambda_t])^2 = 2\kappa\theta \mathbb{E}[\lambda_t] + 2(\mathbb{E}[\lambda_t])^2 (-\kappa + \eta \mathbb{E}[\xi]). \quad (2.5)$$

Putting together Eqs. (2.4) and (2.5), we obtain $\frac{\partial}{\partial t} \mathbb{V}\text{ar}(\lambda_t)$. Then, replacing $\mathbb{E}[\lambda_t]$ by its expression at Eq. (2.1) and performing some simplifications leads to the following ODE

$$\frac{\partial \mathbb{V}\text{ar}(\lambda_t)}{\partial t} = e^{a_3 t} (a_1 \lambda_0 + a_2) - a_2 + 2a_3 \mathbb{V}\text{ar}(\lambda_t)$$

whose solution is given by Eq. (2.2). $\square$

From this proposition, we infer that the stability condition $\lim_{t \rightarrow +\infty} (\mathbb{E}[\lambda_t] \vee \mathbb{V}\text{ar}(\lambda_t)) < +\infty$ is equivalent to $a_3 < 0$, that is, $\eta \mathbb{E}[\xi] < \kappa$. This condition means that the speed of the mean reversion κ must be sufficiently large.

In the next section, we derive semi-closed form expressions for the Laplace transform of the intensity.

3. Laplace Transform

Let r be a real number that represents the constant interest rate and define the process $(P_t^r)_{t \geq 0}$ as the discounted version of $(P_t)_{t \geq 0}$, that is,

$$P_t^r := \int_0^t e^{-rs} dP_s.$$

From this definition, $\mathbb{E}[P_t^r]$ represents the expected discounted cost of the credit portfolio up to time t . Moreover, we will denote by $(\Lambda_t)_{t \geq 0}$ the integral of $(\lambda_t)_{t \geq 0}$, i.e. $\Lambda_t := \int_0^t \lambda_u du$. Note that $(\Lambda_t)_{t \geq 0}$ is the compensator of $(N_t)_{t \geq 0}$. The next proposition will allow us to compute the Laplace transform of this process numerically, and thus to obtain its moments.

Proposition 3.1. *There exist functions A, B such that for all $0 \leq t \leq T$,*

$$\mathbb{E}[\exp\{-v_1(N_T - N_t) - v_2(P_T^r - P_t^r) - v_3(\Lambda_T - \Lambda_t)\} | \mathcal{F}_t] = \exp\{A(v, t, T) + \lambda_t B(v, t, T)\}$$

where $v = (v_1, v_2, v_3)^t$ and A, B satisfy the system of ODEs

$$\begin{aligned} \frac{\partial A}{\partial t}(v, t, T) &= -\kappa \theta B(v, t, T) \\ \frac{\partial B}{\partial t}(v, t, T) &= \kappa B(v, t, T) - \frac{\sigma^2}{2} B^2(v, t, T) + v_3 - (e^{-v_1} \tilde{v}(-\eta B(v, t, T) - u e^{-rt}) - 1). \end{aligned}$$

Moreover A and B satisfy the terminal condition $A(v, T, T) = B(v, T, T) = 0$.

Proof. It suffices to show that there exist functions A, B such that the stochastic process

$$(M_t)_{t \in [0, T]} := \left(\exp \{A(v, t, T) + \lambda_t B(v, t, T) - v_1 N_t - v_2 P_t^r - v_3 \Lambda_t\} \right)_{t \in [0, T]}$$

is a martingale. Noting that $M_t := h(t, \lambda_t, N_t, P_t^r, \Lambda_t)$ where h is defined as

$$h(t, y, x_1, x_2, x_3) = \exp \{A(v, t, T) + y B(v, t, T) - v_1 x_1 - v_2 x_2 - v_3 x_3\}$$

and applying Ito's Lemma allows to obtain

$$\begin{aligned} dM_t &= M_t \left[\left(\frac{\partial A}{\partial t}(v, t, T) + \lambda_t \frac{\partial B}{\partial t}(v, t, T) \right) dt + B(v, t, T) \left(\kappa(\theta - \lambda_t) dt + \sigma \sqrt{\lambda_t} dW_t \right) \right. \\ &\quad \left. + \frac{\sigma^2}{2} B^2(v, t, T) \lambda_t dt - v_3 \lambda_t dt + (\exp\{-v_1 + \eta B(v, t, T)\} \xi - v_2 e^{-rt} \xi) - 1 \right] dN_t \\ &= M_t \left[\left(\frac{\partial A}{\partial t}(v, t, T) + \kappa \theta B(v, t, T) \right) + \right. \\ &\quad \left. \lambda_t \left(\frac{\partial B}{\partial t}(v, t, T) - \kappa B(v, t, T) - v_3 + \frac{\sigma^2}{2} B^2(v, t, T) + \exp\{\xi(\eta B(v, t, T) - v_2 e^{-rt}) - v_1\} - 1 \right) \right] dt \\ &\quad + M_t [\exp\{\xi(\eta B(v, t, T) - v_2 e^{-rt}) - v_1\} - 1] (dN_t - \lambda_t dt) + M_t [\sigma B(v, t, T) \sqrt{\lambda_t}] dW_t, \end{aligned}$$

where ξ is a random variable distributed according to v . Since both $(W_t)_{t \geq 0}$ and $(N_t - \int_0^t \lambda_s ds)_{t \geq 0}$ are martingales, a sufficient condition for $(M_t)_{t \geq 0}$ to be a martingale is

$$0 = \frac{\partial B}{\partial t}(v, t, T) - \kappa B(v, t, T) - v_3 + \frac{\sigma^2}{2} B^2(v, t, T) + e^{-v_1} \tilde{v}(-\eta B(v, t, T) + v_2 e^{-rt}) - 1$$

and

$$0 = \frac{\partial A}{\partial t}(v, t, T) + \kappa \theta B(v, t, T).$$

The conclusion follows. $\square$

It is now easy to characterize the Laplace transform of $(P_t^r)_{t \geq 0}$, which is the most interesting for our purpose.

Corollary 3.2. *There exist functions A, B such that for all $0 \leq t \leq T$,*

$$\mathbb{E}[\exp\{-u(P_T^r - P_t^r)\} | \mathcal{F}_t] = \exp\{A(u, t, T) + \lambda_t B(u, t, T)\}$$

where A, B satisfy the system of ODEs

$$\begin{aligned}\frac{\partial A}{\partial t}(u, t, T) &= -\kappa\theta B(u, t, T) \\ \frac{\partial B}{\partial t}(u, t, T) &= \kappa B(u, t, T) - \frac{\sigma^2}{2}B^2(u, t, T) - \tilde{v}(-\eta B(u, t, T) + ue^{-rt}) + 1\end{aligned}$$

and A, B satisfy the terminal condition $A(u, T, T) = B(u, T, T) = 0$.

Proof. Take $v_1 = v_3 = 0$ and $v_2 = u$ in Proposition 3.1. $\square$

Note that in Proposition 3.1 as well as in Corollary 3.2, it holds that $A(0, t, T) = B(0, t, T) = 0$ for all t, T . The Laplace transform of $(P_t^r)_{t \geq 0}$ will be useful in the numerical application section to compute moments.

The next section is dedicated to the derivation of the time-consistent counterparts of well known actuarial evaluations, namely the variance principle, the standard-deviation principle and the mean value principle. The reader is referred to [2] for an introduction to these actuarial evaluation principles.

4. Time-consistent evaluations

This section extends the approach of Pelsser and Ghalehjooghi to the framework we introduced. A time-consistent evaluation principle is an evaluation principle that has the following property: if it is known today that, at a future date, a risk X will be more expensive than another risk Y , then the risk X should already be more expensive today. An equivalent characterization of time-consistency is the following: let $(\Pi_t)_{t \geq 0}$ be a dynamic evaluation. Such an evaluation is time-consistent if and only if

$$\Pi_t(X) = \Pi_t(\Pi_{t+\Delta}(X)) \quad (4.1)$$

for any risk X and any $t, \Delta \geq 0$. This equivalent characterization is the one we will use in the sequel.

We assume that the companies of the credit portfolio are small or medium companies (SME) for which no credit derivatives are available. These risks being not hedgeable by financial instruments, the market is incomplete and a unique risk-neutral measure does not exist. To manage this uncertainty about the risk neutral measure, we use actuarial valuation methods and work under the real probability measure $\mathbb{P}$.

4.1. Variance principle

For a risk X expressed in a certain monetary unit, the static variance principle is computed as

$$\mathbb{E}[X] + \alpha \text{Var}(X). \quad (4.2)$$

The variance principle is derived in [33]. More precisely, it is shown that the risk premium asked by an economic agent for a small risk is approximately proportional to the variance of the risk. The parameter α is then a measure of risk aversion. Note that in Eq. (4.2), the expectation $\mathbb{E}[X]$ is expressed in the monetary unit, whereas the term $\text{Var}(X)$ is expressed in squared monetary units. Therefore the parameter α should be replaced by α/L , where $L > 0$ is a reference amount expressed in the monetary unit. Equivalently, this means that α is a quantity expressed $1/m_u$, where m_u is the monetary unit.

Let Π^v be a function that is the time-consistent valuation function derived from the variance principle of the trade credit insurance portfolio. This function depends on (t, λ_t) . In the case of the variance principle, the time-consistency requirement (4.1) writes

$$\Pi^v(t, \lambda_t) = \mathbb{E}[e^{-r\varepsilon} \Pi^v(t + \varepsilon, \lambda_{t+\varepsilon}) | \mathcal{F}_t] + \alpha \text{Var}(e^{-r\varepsilon} \Pi^v(t + \varepsilon, \lambda_{t+\varepsilon}) | \mathcal{F}_t). \quad (4.3)$$

This means that we should have

$$\lim_{\varepsilon \downarrow 0} \frac{\mathbb{E}[e^{-r\varepsilon} \Pi^v(t + \varepsilon, \lambda_{t+\varepsilon}) | \mathcal{F}_t] + \alpha \text{Var}(e^{-r\varepsilon} \Pi^v(t + \varepsilon, \lambda_{t+\varepsilon}) | \mathcal{F}_t) - \Pi^v(t, \lambda_t)}{\varepsilon} = 0, \quad (4.4)$$

for any $t, \varepsilon \geq 0, t + \varepsilon \in [0, T]$, where $T > 0$ is the maturity of the contracts. We will separately look at the expectation term and the variance term at the right-hand side of Eq. (4.3). For the expectation part, using Ito's lemma for semimartingales (see e.g. chapter 2 of Protter [34] or chapter 6 Section 7 of Rogers and Williams [35]) yields

$$\begin{aligned}\lim_{\varepsilon \downarrow 0} \mathbb{E} \left[\frac{\Pi^v(t + \varepsilon, \lambda_{t+\varepsilon}) - \Pi^v(t, \lambda_t)}{\varepsilon} | \mathcal{F}_t \right] &= \Pi_t^v(t, \lambda_t) + \kappa(\theta - \lambda_t) \Pi_{\lambda}^v(t, \lambda_t) \\ &+ \frac{\sigma^2 \lambda_t}{2} \Pi_{\lambda\lambda}(t, \lambda_t) + \mathbb{E}[\xi + \Pi^v(t, \lambda_t + \eta\xi) - \Pi^v(t, \lambda_t)] \lambda_t\end{aligned} \quad (4.5)$$

where Π_t^v , Π_λ^v and $\Pi_{\lambda\lambda}^v$ are shorthands for $\partial\Pi^v/\partial t$, $\partial\Pi^v/\partial\lambda$ and $\partial^2\Pi^v/\partial\lambda^2$ respectively and $\xi \sim v$. Note that the result of Eq. (4.5) is the infinitesimal generator of Π^v , that we will denote by $\mathcal{A}\Pi^v$. For the variance part, we will use infinitesimal generators. First, let us recall the following elementary identity for variance

$$\mathbb{V}\text{ar}(Y_{t+\varepsilon}|\mathcal{F}_t) = \mathbb{E}[Y_{t+\varepsilon}^2|\mathcal{F}_t] - Y_t^2 - (\mathbb{E}[Y_{t+\varepsilon}|\mathcal{F}_t] - Y_t)^2 - 2Y_t(\mathbb{E}[Y_{t+\varepsilon}|\mathcal{F}_t] - Y_t) \quad (4.6)$$

where $(Y_t)_{t \geq 0}$ is used instead of $(\Pi^v(t, \lambda_t))_{t \geq 0}$ for shorter notations. Dividing Eq. (4.6) by ε on both sides and letting $\varepsilon \downarrow 0$ leads to

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathbb{V}\text{ar}(\Pi^v(t + \varepsilon, \lambda_{t+\varepsilon})|\mathcal{F}_t)}{\varepsilon} = \mathcal{A}(\Pi^{v2}) - 2\Pi^v(\mathcal{A}\Pi^v) \quad (4.7)$$

where Π^{v2} means $(\Pi^v)^2$ and is used to avoid too many parentheses. We also remind that Π^v is of course a shorthand for $\Pi^v(t, \lambda_t)$. The use of Ito's lemma for semimartingales and the chain rule allows us to get the following expression for the infinitesimal generator $\mathcal{A}(\Pi^{v2})$:

$$\begin{aligned} \mathcal{A}(\Pi^{v2}) &= 2\Pi^v\Pi_t^v + 2\Pi^v\Pi_\lambda^v\kappa(\theta - \lambda_t) + \frac{1}{2}\sigma^2\lambda_t \left(2(\Pi_\lambda^v)^2 + 2\Pi^v\Pi_{\lambda\lambda}^v \right) \\ &\quad + (\mathbb{E}[(\xi + \Pi^v(t, \lambda_t + \eta\xi))^2] - \Pi^{v2})\lambda_t \end{aligned} \quad (4.8)$$

where, again, $\xi \sim v$. Then, starting from Eq. (4.4), using that $e^{-r\varepsilon} = 1 - r\varepsilon + o(\varepsilon)$ and using our computations regarding the expectation and the variance, we obtain a PIDE for the time-consistent variance price:

$$\begin{aligned} r\Pi^v &= \Pi_t^v + \kappa(\theta - \lambda_t)\Pi_\lambda^v + \frac{1}{2}\sigma^2\lambda_t[\Pi_{\lambda\lambda}^v + 2\alpha(\Pi_\lambda^v)^2] \\ &\quad + \mathbb{E}[\xi + \Pi^v(t, \lambda_t + \eta\xi) - \Pi^v(t, \lambda_t)]\lambda_t + \alpha\mathbb{E}[(\xi + \Pi^v(t, \lambda_t + \eta\xi) - \Pi^v(t, \lambda_t))^2]\lambda_t \end{aligned} \quad (4.9)$$

Given the computations we already performed, deriving a PIDE for the time-consistent standard-deviation principle is an easy matter.

4.2. Standard-deviation principle

For a risk X expressed in a certain monetary unit, the static standard-deviation principle is computed as

$$\mathbb{E}[X] + \beta\sqrt{\mathbb{V}\text{ar}(X)}. \quad (4.10)$$

Let Π^s be a function that is the time-consistent valuation derived from the standard-deviation principle. Again, this function depends on (t, λ_t) . For any $t + \varepsilon \in [0, T]$, we should have

$$\Pi^s(t, \lambda_t) = \mathbb{E}[e^{-r\varepsilon}\Pi^s(t + \varepsilon, \lambda_{t+\varepsilon})|\mathcal{F}_t] + \beta\sqrt{\mathbb{V}\text{ar}(e^{-r\varepsilon}\Pi^s(t + \varepsilon, \lambda_{t+\varepsilon})|\mathcal{F}_t)} \quad (4.11)$$

for a fixed constant $\beta > 0$. As noted in [8], the terms $\mathbb{E}[X]$ and $\sqrt{\mathbb{V}\text{ar}(X)}$ have different time scales. That is, the expectation term scales linearly with the time whereas the standard-deviation term scales with the square root of the time. Indeed, for a small $\Delta_t > 0$, Eq. (4.5) shows that

$$\mathbb{E}[\Pi(t + \Delta_t, \lambda_{t+\Delta_t}) - \Pi(t, \lambda_t)|\mathcal{F}_t] \approx (\mathcal{A}\Pi)\Delta_t$$

whereas Eq. (4.7) shows that

$$\mathbb{V}\text{ar}(\Pi(t + \Delta_t, \lambda_{t+\Delta_t})|\mathcal{F}_t) \approx (\mathcal{A}(\Pi^2) - 2\Pi\mathcal{A}\Pi)\Delta_t$$

so that both the expectation and the variance scales linearly with time. It follows that the standard-deviation scales linearly with the square root of the time. Therefore, the β parameter must be replaced by $\sqrt{\varepsilon}\beta$ in Eq. (4.11), which yields

$$\Pi^s(t, \lambda_t) = \mathbb{E}[e^{-r\varepsilon}\Pi^s(t + \varepsilon, \lambda_{t+\varepsilon})|\mathcal{F}_t] + \beta\sqrt{\varepsilon}\sqrt{\mathbb{V}\text{ar}(e^{-r\varepsilon}\Pi^s(t + \varepsilon, \lambda_{t+\varepsilon})|\mathcal{F}_t)}. \quad (4.12)$$

From an economic point of view, it means that if we want to compare a standard-deviation measured over a year with a standard-deviation measured over a month, the annual outcome should be scaled with $\sqrt{1/12}$ to become fairly comparable. From Eq. (4.12), we proceed as for the time-consistent variance principle: we divide by ε and let ε tend to 0. Using the results we obtained for the variance principle, we have

$$\begin{aligned} r\Pi^s(t, \lambda_t) &= \Pi_t^s(t, \lambda_t) + \kappa(\theta - \lambda_t)\Pi_\lambda^s(t, \lambda_t) + \frac{\sigma^2\lambda_t}{2}\Pi_{\lambda\lambda}^s(t, \lambda_t) \\ &\quad + \mathbb{E}[\xi + \Pi^s(t, \lambda_t + \eta\xi) - \Pi^s(t, \lambda_t)]\lambda_t + \beta\lim_{\varepsilon \downarrow 0} \sqrt{\frac{\mathbb{V}\text{ar}(\Pi^s(t + \varepsilon, \lambda_{t+\varepsilon})|\mathcal{F}_t)}{\varepsilon}}. \end{aligned} \quad (4.13)$$

Eq. (4.13) also provides a more mathematical reason to the multiplication of the standard-deviation by $\sqrt{\varepsilon}$: it is exactly what we need to obtain the convergence of the last term when $\varepsilon \rightarrow 0$. With the help of Eqs. (4.7) and (4.8) we can compute the limit of the standard-deviation term and arrive to the following PIDE

$$\begin{aligned} r\Pi^s(t, \lambda_t) &= \Pi_t^s(t, \lambda_t) + \kappa(\theta - \lambda_t)\Pi_\lambda^s(t, \lambda_t) + \frac{\sigma^2\lambda_t}{2}\Pi_{\lambda\lambda}^s(t, \lambda_t) \\ &\quad + \mathbb{E}[\xi + \Pi^s(t, \lambda_{t-} + \eta\xi) - \Pi^s(t, \lambda_{t-})]\lambda_t \\ &\quad + \beta \left[\mathbb{E}[(\xi + \Pi^s(t, \lambda_t + \eta\xi) - \Pi^s)^2]\lambda_t + \sigma^2\lambda_t(\Pi_\lambda^s)^2 \right]^{1/2}. \end{aligned} \quad (4.14)$$

In the next subsection, we derive the time-consistent counterpart of an evaluation known as the mean value principle.

4.3. Mean value principle

The static mean value principle is given by

$$\Pi(X) = v^{-1}\left(\mathbb{E}[v(X)]\right) \quad (4.15)$$

where v can be any convex increasing function. The time-consistent evaluation operator Π^m derived from the static Mean Value principle should satisfy

$$\Pi^m(t, \lambda_t) = v^{-1}(\mathbb{E}[v(e^{-r\varepsilon}\Pi^m(t + \varepsilon, \lambda_{t+\varepsilon}))|\mathcal{F}_t])$$

for any $0 \leq t + \varepsilon \leq T$, with $t, \varepsilon \geq 0$. It implies that

$$\lim_{\varepsilon \downarrow 0} \frac{\mathbb{E}\left[v\left(\frac{\Pi^m(t+\varepsilon, \lambda_{t+\varepsilon})}{e^{r\varepsilon}}\right) | \mathcal{F}_t\right] - v(\Pi^m(t, \lambda_t))}{\varepsilon} = 0$$

for all $t \geq 0$. Using this condition and Ito's Lemma on the process $\left(v\left(\frac{\Pi^m(t+\varepsilon, \lambda_{t+\varepsilon})}{e^{r\varepsilon}}\right)\right)_{\varepsilon \geq 0}$, we will derive a partial differential equation for the time-consistent price. In order to apply Ito's Lemma, we compute the partial derivatives of the function $(\varepsilon, \lambda) \mapsto f(\varepsilon, \lambda) := v\left(\frac{\Pi^m(t+\varepsilon, \lambda)}{e^{r\varepsilon}}\right)$ by the use of the chain rule. This leads to

$$\begin{aligned} \frac{\partial f}{\partial \varepsilon}(s, \lambda_{t+s}) &= \\ v' \left(\frac{\Pi^m(t+s, \lambda_{t+s})}{e^{rs}} \right) \frac{\Pi_t^m(t+s, \lambda_{t+s}) - r\Pi^m(t+s, \lambda_{t+s})}{e^{rs}}, \\ \frac{\partial f}{\partial \lambda}(s, \lambda_{t+s}) &= v' \left(\frac{\Pi^m(t+s, \lambda_{t+s})}{e^{rs}} \right) \frac{\Pi_\lambda^m(t+s, \lambda_{t+s})}{e^{rs}}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 f}{\partial \lambda^2}(s, \lambda_{t+s}) &= \left(\frac{\Pi_\lambda^m(t+s, \lambda_{t+s})}{e^{rs}} \right)^2 v'' \left(\frac{\Pi^m(t+s, \lambda_{t+s})}{e^{rs}} \right) \\ &\quad + v \left(\frac{\Pi^m(t+s, \lambda_{t+s})}{e^{rs}} \right) \frac{\Pi_{\lambda\lambda}^m(t+s, \lambda_{t+s})}{e^{rs}}. \end{aligned}$$

As before, from Ito's Lemma, we find

$$\begin{aligned} 0 &= \frac{\sigma^2}{2} \lambda_t (\Pi_\lambda^m(t, \lambda_t))^2 v''(\Pi^m(t, \lambda_t)) + \mathbb{E}[v(\xi + \Pi^m(t, \lambda_t + \eta\xi)) - v(\Pi^m(t, \lambda_t))]\lambda_t \\ &\quad + v'(\Pi^m(t, \lambda_t)) \left[\Pi_t^m(t, \lambda_t) - r\Pi^m(t, \lambda_t) + \Pi_\lambda^m(t, \lambda_t)\kappa(\theta - \lambda_t) + \frac{\sigma^2}{2} \lambda_t \Pi_{\lambda\lambda}^m(t, \lambda_t) \right]. \end{aligned} \quad (4.16)$$

which can be used as a PIDE for the time-consistent price.

4.4. The exponential principle

The exponential principle is a particular case of the mean value principle. It is obtained by defining the function v by $v(x) = e^{\gamma^{-1}x}$, $\gamma > 0$. A PIDE for the time-consistent exponential principle is obtained by replacing v , $v'(x) = \gamma^{-1}e^{\gamma^{-1}x}$ and

$v''(x) = \gamma^{-2} e^{\gamma^{-1}x}$ in Eq. (4.16). We obtain

$$\begin{aligned} 0 &= \frac{\sigma^2}{2} \lambda_t (\Pi_\lambda^m(t, \lambda_t))^2 \gamma^{-2} e^{\gamma^{-1} \Pi^m(t, \lambda_t)} \\ &+ \frac{\Pi_t^m(t, \lambda_t) - r \Pi^m(t, \lambda_t) + \Pi_\lambda^m(t, \lambda_t) \kappa(\theta - \lambda_t) + \frac{\sigma^2}{2} \lambda_t \Pi_{\lambda\lambda}^m(t, \lambda_t)}{\gamma e^{-\gamma^{-1} \Pi^m(t, \lambda_t)}} \\ &+ \mathbb{E} \left[e^{\gamma^{-1}(\xi + \Pi^m(t, \lambda_t + \eta\xi))} - e^{\gamma^{-1} \Pi^m(t, \lambda_t)} \right] \lambda_t. \end{aligned} \quad (4.17)$$

Since $e^{\gamma^{-1} \Pi^m(t, \lambda_t)} > 0$, we can divide both sides of Eq. (4.17) by $e^{\gamma^{-1} \Pi^m(t, \lambda_t)}$, so that we get

$$\begin{aligned} 0 &= \frac{\Pi_t^m(t, \lambda_t) - r \Pi^m(t, \lambda_t) + \Pi_\lambda^m(t, \lambda_t) \kappa(\theta - \lambda_t) + \frac{\sigma^2}{2} \lambda_t \Pi_{\lambda\lambda}^m(t, \lambda_t)}{\gamma} \\ &+ \mathbb{E} \left[e^{\gamma^{-1}(\xi + \Pi^m(t, \lambda_t + \eta\xi) - \Pi^m(t, \lambda_t))} - 1 \right] \lambda_t + \frac{\sigma^2}{2} \lambda_t (\Pi_\lambda^m(t, \lambda_t))^2 \gamma^{-2}. \end{aligned} \quad (4.18)$$

This plays a role in the numerical methods part. Indeed, the exponential of the evaluation $e^{\gamma^{-1} \Pi^m(t, \lambda_t)}$ can lead to huge numbers, causing problems in the numerical solving of the PDE. It is therefore important to get rid of it.

5. Numerical application

The goal of this numerical application section is to explore numerically the impact of the risk of contagion between defaults and the impact of the time-consistency of evaluations. We are also interested in observing possible interactions between time-consistency and contagion. In order to do so, we start by explaining how to compute actual evaluations on the basis of the results presented in this paper. This is done in Section 5.1 for the time-inconsistent evaluations and in Section 5.2 for their time-consistent counterparts. The numerical results are presented and analyzed in Section 5.3. Finally, we propose a calibration method for our model in Section 5.4.

5.1. Time-inconsistent setting

The time-inconsistent evaluations represented in the upcoming graphs correspond to the functions

$$[0, T] \ni t \mapsto \mathbb{E}[e^{rt}(P_t^r - P_t^r)|\mathcal{F}_t] + \alpha \text{Var}(e^{rt}(P_t^r - P_t^r)|\mathcal{F}_t), \quad (5.1)$$

$$[0, T] \ni t \mapsto \mathbb{E}[e^{rt}(P_t^r - P_t^r)|\mathcal{F}_t] + \beta \sqrt{T-t} \sqrt{\text{Var}(e^{rt}(P_t^r - P_t^r)|\mathcal{F}_t)} \quad (5.2)$$

and

$$[0, T] \ni t \mapsto \gamma \ln \left(\mathbb{E} \left[\exp \left\{ \frac{e^{rt}(P_t^r - P_t^r)}{\gamma} \right\} | \mathcal{F}_t \right] \right) \quad (5.3)$$

where $e^{rt}(P_t^r - P_t^r) = \int_t^T e^{-r(s-t)} dP_s$ is the remaining discounted cost of the portfolio seen from time $t \in [0, T]$. The expectation is computed with

$$\begin{aligned} \mathbb{E}[(P_t^r - P_t^r)|\mathcal{F}_t] &= (-1) \frac{\partial}{\partial u} \mathbb{E}[\exp\{-u(P_t^r - P_t^r)\}|\mathcal{F}_t] \Big|_{u=0} \\ &= - \left(\frac{\partial A}{\partial u}(u, t, T) + \lambda_t \frac{\partial B}{\partial u}(u, t, T) \right) \Big|_{u=0} \\ &\approx - \frac{A(\Delta_u, t, T) - A(-\Delta_u, t, T)}{2\Delta_u} - \lambda_t \frac{B(\Delta_u, t, T) - B(-\Delta_u, t, T)}{2\Delta_u} \end{aligned} \quad (5.4)$$

that holds for a small $\Delta_u > 0$. The functions A and B come from Corollary 3.2. Similarly, we have for the second moment

$$\begin{aligned} \mathbb{E}[(P_t^r - P_t^r)^2|\mathcal{F}_t] &= \frac{\partial^2}{\partial u^2} \mathbb{E}[\exp\{-u(P_t^r - P_t^r)\}|\mathcal{F}_t] \Big|_{u=0} \\ &= \frac{\partial}{\partial u} \left[(-1) \left(\frac{\partial A}{\partial u}(u, t, T) + \lambda_t \frac{\partial B}{\partial u}(u, t, T) \right) \exp\{A(u, t, T) + \lambda_t B(u, t, T)\} \right] \Big|_{u=0} \\ &= \left[\left(\frac{\partial^2 A}{\partial u^2}(u, t, T) + \lambda_t \frac{\partial^2 B}{\partial u^2}(u, t, T) \right) + \left(\frac{\partial A}{\partial u}(u, t, T) + \lambda_t \frac{\partial B}{\partial u}(u, t, T) \right)^2 \right] \Big|_{u=0} \\ &\approx \left(\frac{A(\Delta_u, t, T) + A(-\Delta_u, t, T)}{\Delta_u^2} + \lambda_t \frac{B(\Delta_u, t, T) + B(-\Delta_u, t, T)}{\Delta_u^2} \right) \\ &+ \left(\frac{A(\Delta_u, t, T) - A(-\Delta_u, t, T)}{2\Delta_u} + \lambda_t \frac{B(\Delta_u, t, T) - B(-\Delta_u, t, T)}{2\Delta_u} \right)^2 \end{aligned} \quad (5.5)$$

where we used the fact that $A(0, t, T) = B(0, t, T) = 0$ within the usual second order difference approximation $\frac{d^2 f}{dx^2}(x) \approx (f(x + \Delta) - 2f(x) + f(x - \Delta))/\Delta^2$. By combining Eqs. (5.4) and (5.5), we get

$$\text{Var}((P_T^r - P_t^r)|\mathcal{F}_t) \approx \frac{A(\Delta_u, t, T) + A(-\Delta_u, t, T)}{\Delta_u^2} + \lambda_t \frac{B(\Delta_u, t, T) + B(-\Delta_u, t, T)}{\Delta_u^2}. \quad (5.6)$$

We now explain how we approximate the values of the functions A and B with the help of Corollary 3.2. We proceed recursively starting from $A(u, T, T) = B(u, T, T) = 0$. We use a forward difference approximation, which gives for $k = 0, 1, \dots$

$$\begin{aligned} \frac{\partial B}{\partial t}(u, T - (k+1)\Delta_t, T) &\approx \frac{B(u, T - k\Delta_t, T) - B(u, T - (k+1)\Delta_t, T)}{\Delta_t} \\ &= \kappa B(u, T - (k+1)\Delta_t, T) - \frac{\sigma^2}{2} B^2(u, T - (k+1)\Delta_t, T) \\ &\quad - \tilde{v}(-\eta B(u, T - (k+1)\Delta_t, T) + ue^{-r(T-(k+1))}) + 1 \end{aligned} \quad (5.7)$$

and we solve numerically for $B(u, T - (k+1)\Delta_t, T)$. Similarly, the values of A are found using

$$A(u, T - (k+1)\Delta_t, T) = A(u, T - k\Delta_t, T) + \Delta_t \kappa \theta B(u, T - (k+1)\Delta_t, T). \quad (5.8)$$

Note that the difference steps Δ_u and Δ_t were fixed to 10^{-6} and 2×10^{-5} respectively in all our computations of time-inconsistent evaluations.

5.2. Time-consistent setting

In this section, we explain the numerical method that allows us to solve the 3 PIDEs derived in Section 4. To this end, we start by explaining the general setting of the method. Afterwards we address the problem of solving each PIDE specifically.

We work with a finite difference method. There are two variables to discretize: the intensity λ and the time t . The finite difference method is performed on a finite grid for the two variables. The limits for the time will be taken to be 0 and T for the time, T being the maturity of the contract whereas the limits for λ are taken to be 0 and $\lambda_{\max} > 0$. The boundary $\lambda_{\max}$ will be fixed so that it is large enough to be attained by $(\lambda_t)_{t \geq 0}$ with a very low probability. We define n_λ and n_t , strictly positive integers, to be the number of steps in the grid for λ and t respectively. These integers imply mesh sizes for both discretized variables, we will denote those mesh sizes by $\Delta_\lambda := \lambda_{\max}/n_\lambda$ and $\Delta_t := T/n_t$. The grid for the time is $(t_0, t_1, \dots, t_{n_t})$, where $t_i = i\Delta_t$ for each i . The grid for the intensity λ_t is $(\lambda_0, \lambda_1, \dots, \lambda_{n_\lambda})$ where $\lambda_j = j\Delta_\lambda$ for each j .

The approximated values of the function Π on the grid will be denoted by $p_{i,j} \approx \Pi(t_i, \lambda_j)$. The matrix that contains all approximated values $p_{i,j}$ will be denoted by $P = (p_{i,j})_{\substack{i=0, \dots, n_t \\ j=0, \dots, n_\lambda}}$. The approximations of partial derivatives will be denoted as

$$\partial_t p_{i,j} \approx \frac{\partial \Pi}{\partial t}(t_i, \lambda_j) \quad \partial_\lambda p_{i,j} \approx \frac{\partial \Pi}{\partial \lambda}(t_i, \lambda_j) \quad \partial_{\lambda\lambda} p_{i,j} \approx \frac{\partial^2 \Pi}{\partial \lambda^2}(t_i, \lambda_j) \quad (5.9)$$

so that $\partial_t P = (\partial_t p_{i,j})_{\substack{i=1, \dots, n_t+1 \\ j=1, \dots, n_\lambda+1}}$, $\partial_\lambda P = (\partial_\lambda p_{i,j})_{\substack{i=1, \dots, n_t+1 \\ j=1, \dots, n_\lambda+1}}$ and $\partial_{\lambda\lambda} P = (\partial_{\lambda\lambda} p_{i,j})_{\substack{i=1, \dots, n_t+1 \\ j=1, \dots, n_\lambda+1}}$ are matrices that approximate the needed partial derivatives of Π .

The only boundary condition in this problem is $\Pi(T, \lambda) = 0$ for any possible intensity λ . Indeed, at time T the contract is over and is thus worth 0. Consequently, we set $p_{n_t, j} = 0$ for all $j \in \{0, 1, \dots, n_\lambda\}$. All the other $p_{i,j}$ will be found with a finite difference method based on the PIDEs. The approximations of the partial derivatives with respect to time are, for all $i \in \{0, 1, \dots, n_t - 1\}$ and $j \in \{0, 1, \dots, n_\lambda\}$, given by

$$\partial_t p_{i,j} = \frac{p_{i+1,j} - p_{i,j}}{\Delta_t}.$$

Concerning the approximations of the partial derivatives with respect to λ , we use the following formulas

$$\partial_\lambda p_{i,j} = \begin{cases} \theta_{\text{or}} \frac{p_{i,j+2} - p_{i,j}}{2\Delta_\lambda} + (1 - \theta_{\text{or}}) \frac{p_{i+1,j+2} - p_{i+1,j}}{2\Delta_\lambda} & \text{if } j = 0 \\ \theta_{\text{or}} \frac{p_{i,j} - p_{i,j-2}}{2\Delta_\lambda} + (1 - \theta_{\text{or}}) \frac{p_{i+1,j} - p_{i+1,j-2}}{2\Delta_\lambda} & \text{if } j = n_\lambda \\ \theta_{\text{or}} \frac{p_{i,j+1} - p_{i,j-1}}{2\Delta_\lambda} + (1 - \theta_{\text{or}}) \frac{p_{i+1,j+1} - p_{i+1,j-1}}{2\Delta_\lambda} & \text{if } j \notin \{0, n_\lambda\} \end{cases} \quad (5.10)$$

$$\partial_{\lambda\lambda} p_{i,j} = \begin{cases} \theta_{\text{or}} \frac{p_{i,j+2} - 2p_{i,j+1} + p_{i,j}}{(\Delta_\lambda)^2} + (1 - \theta_{\text{or}}) \frac{p_{i+1,j+2} - 2p_{i+1,j+1} + p_{i+1,j}}{(\Delta_\lambda)^2} & \text{if } j = 0 \\ \theta_{\text{or}} \frac{p_{i,j} - 2p_{i,j-1} + p_{i,j-2}}{(\Delta_\lambda)^2} + (1 - \theta_{\text{or}}) \frac{p_{i+1,j} - 2p_{i+1,j-1} + p_{i+1,j-2}}{(\Delta_\lambda)^2} & \text{if } j = n_\lambda \\ \theta_{\text{or}} \frac{p_{i,j+1} - 2p_{i,j} + p_{i,j-1}}{(\Delta_\lambda)^2} + (1 - \theta_{\text{or}}) \frac{p_{i+1,j+1} - 2p_{i+1,j} + p_{i+1,j-1}}{(\Delta_\lambda)^2} & \text{if } j \notin \{0, n_\lambda\} \end{cases} \quad (5.11)$$

where $\theta_{or} \in [0, 1]$ is an over relaxation parameter (the subscript “or” stands for over relaxation). This parameter was fixed to 0.5 in our computations. Our approximations of partial derivatives with respect to λ present a special feature. At Eqs. (5.10) and (5.11) in the cases $j \in \{0, n_\lambda\}$, the approximations are modified so that we do not fall outside of the grid. This trick allows us to solve the PIDEs in spite of the lack of boundary conditions.

5.2.1. Variance principle

In this subsection, we focus on solving PIDE (4.9) numerically. The partial derivatives will be approached according to Eqs. (5.9), (5.10) and (5.11). In addition to partial derivatives with respect to t and λ , this equation also contains two terms that correspond to the expected impact of a default on the price, that is, $\mathbb{E}[\Pi^v(t, \lambda_t + \eta\xi) - \Pi^v(t, \lambda_t)]$ and $\mathbb{E}[(\xi + \Pi^v(t, \lambda_t + \eta\xi) - \Pi^v(t, \lambda_t))^2]$. These terms will be approximated by discretization of the random variable ξ , which is equivalent to Riemann sum approximations. More precisely the first term is approximated by

$$\begin{aligned} & \mathbb{E}[\xi + \Pi^v(t_i, \lambda_j + \eta\xi) - \Pi^v(t_i, \lambda_j)] \\ & \approx \mathbb{E}[\xi] + \sum_{k=j+1}^{n_\lambda} p_{i,k}^v \left(\left[\frac{(k-j-1)}{\eta} \Delta_\lambda, \frac{(k-j)}{\eta} \Delta_\lambda \right) \right) \\ & + p_{i,n_\lambda}^v \left(\left[\frac{n_\lambda-j}{\eta} \Delta_\lambda, +\infty \right) \right) - p_{i,j}^v. \end{aligned}$$

Before approaching the second term, it is rewritten as

$$\begin{aligned} & \mathbb{E}[(\xi + \Pi^v(t_i, \lambda_j + \eta\xi) - \Pi^v(t_i, \lambda_j))^2] \\ & \approx \mathbb{E}[\xi^2] + \mathbb{E}[\Pi^{v2}(t_i, \lambda_j + \eta\xi)] + 2\mathbb{E}[\xi \Pi^v(t_i, \lambda_j + \eta\xi)] \\ & + \Pi^{v2}(t_i, \lambda_j) - 2\Pi^v(t_i, \lambda_j)\mathbb{E}[\xi] - 2\Pi^v(t_i, \lambda_j)\mathbb{E}[\Pi^v(t_i, \lambda_j + \eta\xi)]. \end{aligned}$$

The term $\mathbb{E}[\Pi^{v2}(t_i, \lambda_j + \eta\xi)]$ is approximated by

$$\begin{aligned} & \mathbb{E}[\Pi^{v2}(t_i, \lambda_j + \eta\xi)] \\ & \approx \sum_{k=j+1}^{n_\lambda+1} (p_{i,k})^2 \left(\left[\frac{(k-j-1)}{\eta} \Delta_\lambda, \frac{(k-j)}{\eta} \Delta_\lambda \right) \right) \\ & + (p_{i,n_\lambda+1})^2 \left(\left[\frac{n_\lambda+1-j}{\eta} \Delta_\lambda, +\infty \right) \right), \end{aligned}$$

whereas $\mathbb{E}[\xi \Pi^v(t_i, \lambda_j + \eta\xi)]$ is approximated by

$$\begin{aligned} & \mathbb{E}[\xi \Pi^v(t_i, \lambda_j + \eta\xi)] \\ & \approx \sum_{k=j+1}^{n_\lambda+1} \frac{(k-j)}{\eta} \Delta_\lambda p_{i,k}^v \left(\left[\frac{(k-j-1)}{\eta} \Delta_\lambda, \frac{(k-j)}{\eta} \Delta_\lambda \right) \right) \\ & + \mathbb{E}[\xi] p_{i,n_\lambda+1}^v \left(\left[\frac{n_\lambda+1-j}{\eta} \Delta_\lambda, +\infty \right) \right). \end{aligned}$$

Let $p_{i,j}^v$ be the approximation of $\Pi^v(t_i, \lambda_j)$ and P^v be the matrix that contains all the approximations $p_{i,j}^v$. As already pointed out, the last column of P^v , denoted by¹ $P_{*n_t}^v$, is filled with zeros. We can now replace the terms of PIDE (4.9) by their approximations described above. For each $k \in \{0, 1, \dots, n_t - 1\}$, we can find P_{*k}^v from $P_{*(k+1)}^v$ by solving numerically a system of equations.

Solving numerically PIDE (4.14), i.e. the one that describes the time-consistent standard-deviation principle, is done in the same way. That is, the method and the approximations are the same, but the system of equations that allows to find P_{*k}^s from $P_{*(k+1)}^s$ is modified so that it corresponds to PIDE (4.14).

5.2.2. Exponential principle

We focus on the specificity of the numerical resolution of PIDE (4.18). Apart from the partial derivatives that are again approached by the formulas (5.9), (5.10) and (5.11), we need to approximate $\mathbb{E}\left[e^{\gamma^{-1}(\xi + \Pi^m(t_i, \lambda_j + \eta\xi))}\right]$. This is done with a

¹ That is actually the column number $n_t + 1$ of P^v , since there are $n_t + 1$ points in the time grid.

Table 5.1

Parameters used for the computations.

	θ	λ_0	κ	σ	η	ρ	r	T	β	α	γ
SE = yes	0.1	0.1	0.24	0.05	0.0001	0.001	0.02	10	0.1	0.1ρ	10000
SE = no	0.152598	0.1	0.24	0.05	0	0.001	0.02	10	0.1	0.1ρ	10000

discretization of the random variable ξ :

$$\begin{aligned} & \mathbb{E} \left[e^{\gamma^{-1}(\xi + \Pi^m(t_i, \lambda_j + \eta\xi))} \right] \\ & \approx \sum_{k=j+1}^{n_\lambda} \exp \left\{ \gamma^{-1} \left(\frac{(k-j)}{\eta} \Delta_\lambda + p_{i,k}^m \right) \right\} \nu \left(\left[\frac{(k-j-1)}{\eta} \Delta_\lambda, \frac{(k-j)}{\eta} \Delta_\lambda \right) \right) \\ & + \exp \left\{ \gamma^{-1} \left(\frac{n_\lambda-j}{\eta} \Delta_\lambda + p_{i,n_\lambda}^m \right) \right\} \nu \left(\left[\frac{n_\lambda-j}{\eta} \Delta_\lambda, +\infty \right) \right). \end{aligned}$$

We can now proceed inductively as before: P_{*k}^m is deduced from $P_{*(k+1)}^m$ by solving numerically a system of equations.

5.3. Results

This section studies numerically the convergence and stability of the numerical procedure described in Section 5.2 for the computation of the time-consistent evaluations. More precisely, the time-consistent evaluations we represent on the graphs are given by the functions $[0, T] \ni t \mapsto \Pi^v(t, 0.1)$, $[0, T] \ni t \mapsto \Pi^s(t, 0.1)$ and $[0, T] \ni t \mapsto \Pi^m(t, 0.1)$ respectively for the variance, standard-deviation and exponential principles. Note that the price at time t is computed with the assumption that λ_t equals the target level 0.1. In other words, the evaluation at time t is computed with the assumption that the process $(\lambda_t)_{t \geq 0}$ satisfies SDE (1.1) but that its value at time t happens to be the target level θ (which is set to 0.1 in the computations).

For the computations, it is also assumed that the jump measure ν satisfies

$$\nu(dx) = \rho \exp\{-x\rho\}dx, \quad x \geq 0$$

so that the jump size ξ is exponentially distributed with mean ρ^{-1} . This is equivalent to assuming that the loss given default has the memoryless property, i.e. $P(\xi > t + s | \xi > s) = P(\xi > t)$ for all $s, t \geq 0$. For example, when a default occurs, the probability of losing more than 2000 monetary units given that the loss exceeds 1000 monetary units is the same as the probability of losing more than 1000 monetary units. The values of the parameters we fixed for our numerical experiments are given in Table 5.1. In the abbreviation “SE = yes”, SE stands for “Self-Exciting”, meaning that the model exhibits a phenomenon of contagion between defaults (that is, $\eta > 0$). The second line, “SE = no”, corresponds to a model without contagion ($\eta = 0$). Recall that the parameters α , β and γ are specific to the variance, standard-deviation and exponential principle respectively. Note that in the section devoted to the variance principle, we pointed out that the variance term should be divided by a reference amount L . This reference amount L is here chosen to be the average default cost, ρ^{-1} , as indicated in Table 5.1 (parameter α). The maturity is fixed to $T = 10$ years.

5.3.1. Convergence

We present and discuss the numerical results obtained by solving the PIDEs with the method described in Section 5.2 and the parameters of Table 5.1 (SE = yes). We start by checking the convergence and stability of the results. The convergence with respect to Δ_t is fast and the results are not modified by the choice of $\lambda_{\max}$, as we can observe on Fig. 5.1. In this Figure, we represent the evolution of the time-consistent price for the 3 principles with an intensity kept constant and equal to $\lambda_0 = \theta = 0.1$. Note that in order to obtain Fig. 5.1, Δ_λ was fixed to 0.0001. Fig. 5.2 shows the convergence with respect to Δ_λ , which is visibly slower than the convergence with respect to Δ_t .

Now that we have checked numerically the convergence of our method, we will use it to assess the impact of time-consistency and self-excitation on the price.

5.3.2. Joint impact of time-consistency and self-excitation

In this section, the time-consistent evaluations are compared with their time-inconsistent equivalents. The time-inconsistent evaluations are given by the functions of Eqs. (5.1), (5.2) and (5.3). These functions are computed with the help of the method described in detail in Section 5.1. Note that, as for the time-consistent evaluations, the value of λ_t is needed to compute the evaluation at time t . It is why we always assume that λ_t equals the target level 0.1 in the computations.

Computing prices that ignore contagion requires some comments. Canceling the phenomenon of contagion in the portfolio can be simply done in our model. Indeed, it suffices to set $\eta = 0$. However, if we do so, the new model is not anymore fairly comparable to the previous one. As a matter of fact, setting $\eta = 0$ will lead to a lower expected number

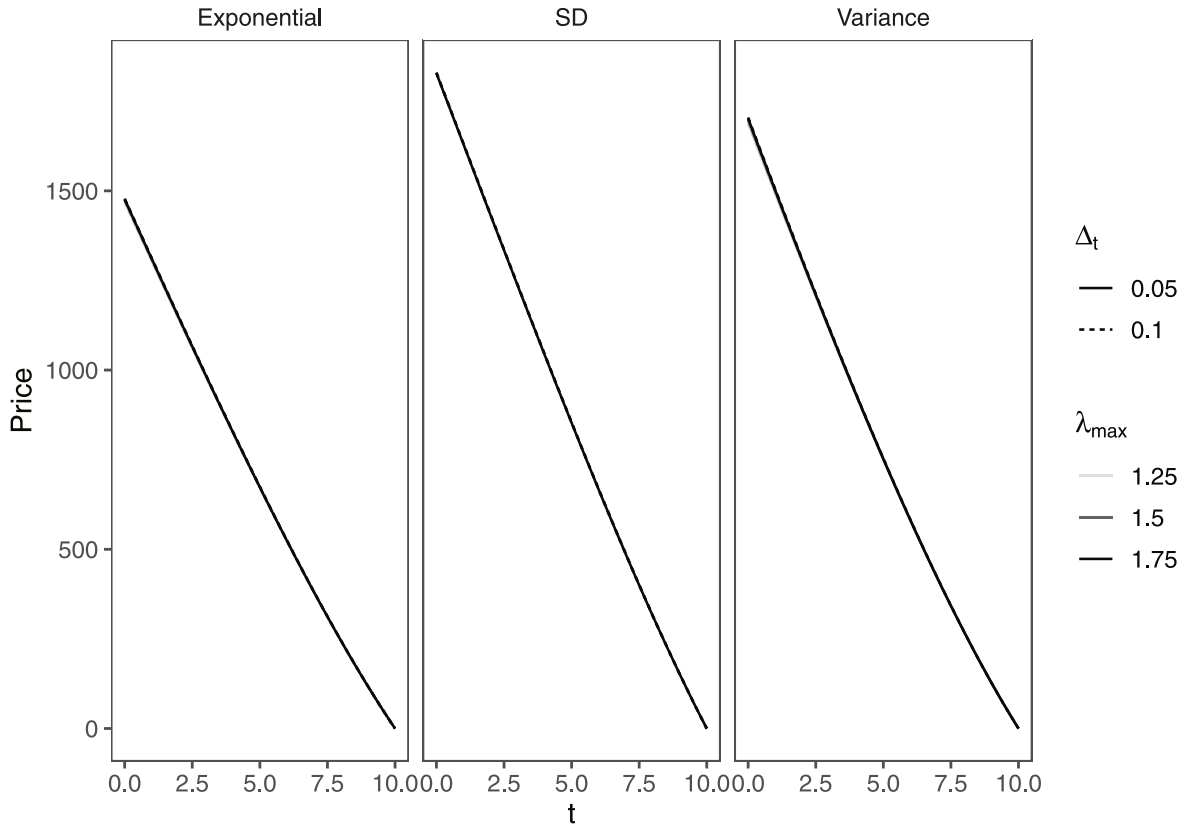


Fig. 5.1. Evolution of the time-consistent variance, Standard-Deviation (SD) and exponential principles.

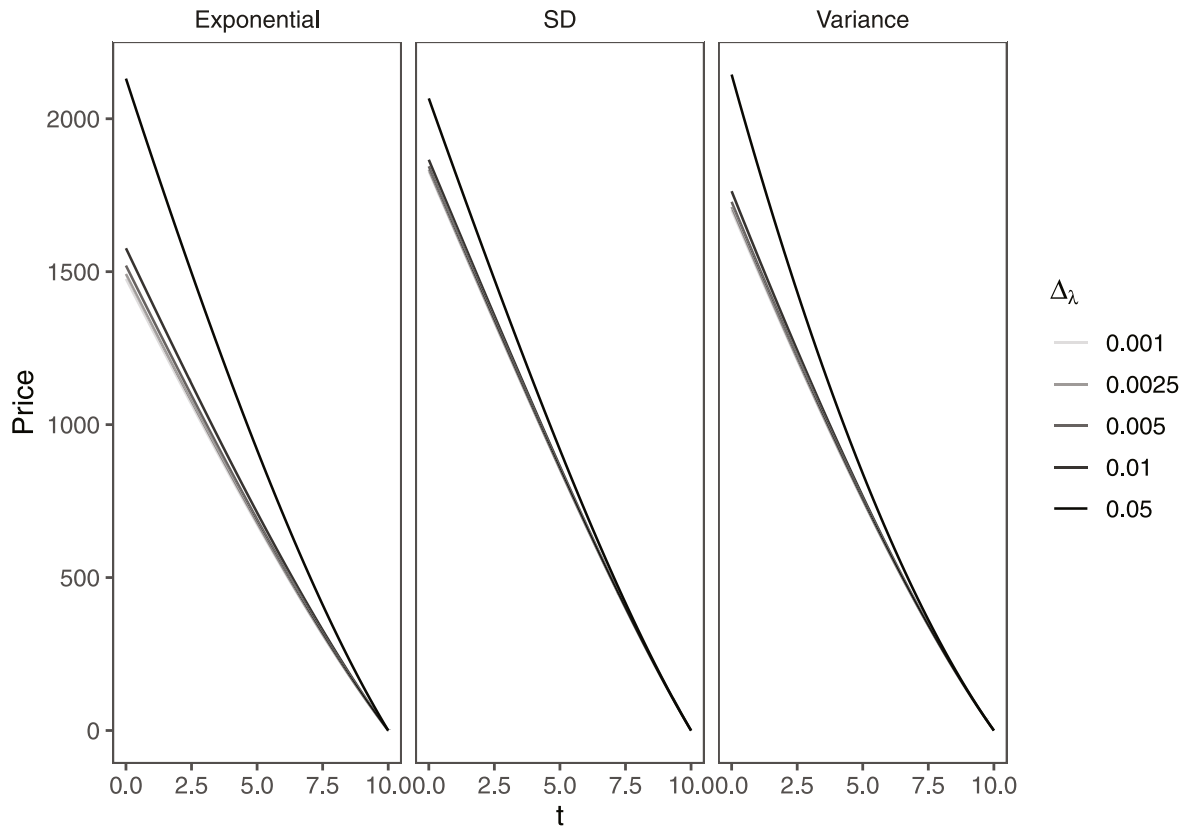
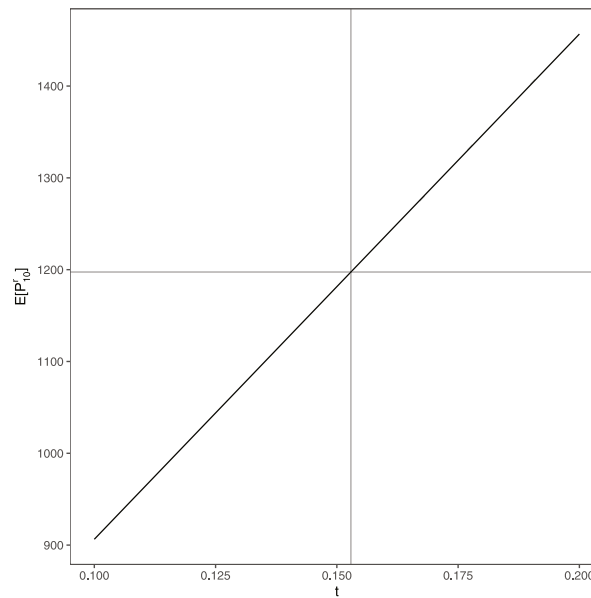
of defaults and therefore to lower prices. This is why we will adjust the parameter θ in order to keep the expected discounted cost $\mathbb{E}[P_T^r]$ unchanged. This is illustrated at Fig. 5.3 and the new “corrected” value for θ is raised to 0.152898. More precisely, with the parameters of the first line of Table 5.1, we computed that $\mathbb{E}[P_T^r] = 1197.441$ with the method explained in Section 5.1. This corresponds to the horizontal line of Fig. 5.3. This Figure also displays the expected value of P_T^r with the same parameters, with the exception of $\eta = 0$, as a function of θ . Therefore, it illustrates that θ should be raised to 0.152598 (vertical line) to keep the expected value unchanged, as indicated in Table 5.1.

Fig. 5.4 compares the prices in four different settings. The first setting combines contagion and time-consistency (TC = yes, SE = yes) and is represented by a dashed dark curve. The second setting is time-consistent but ignores contagion (TC = yes, SE = no). It corresponds to the continuous dark curve. The third and fourth settings are not time-consistent. The third takes the phenomenon of contagion into account and is represented by a clear dashed curve whereas the fourth, represented by a clear continuous curve, does not.

Note that in all three cases, with the exception of the standard-deviation principle where the two continuous lines are indistinguishable on the graph, the dark curves are always above the clear curves. This means that using time-consistent evaluations leads to higher prices in most cases, compared to the corresponding time-inconsistent evaluations. Fig. 5.5 makes it easier to see by showing the difference between the time-consistent evaluations and the time-inconsistent ones. A positive difference means that the time-consistent evaluation is more expensive. The standard-deviation principle exhibits the smallest differences, so that it is the least time-inconsistent in our setting. It also distinguishes itself because it is the only one that leads to higher (but very close) prices in the time-inconsistent setting, which happens in the absence of a risk of contagion only. In all other cases, the positive differences indicate that time-inconsistent evaluations tend to underestimate the cost of the credit portfolio. The fact that the dashed curves always lie above the continuous ones means that this underestimation problem is much more severe when there is a risk of contagion. As a conclusion, the existence of a risk of contagion in a credit portfolio can be seen as an additional reason to work with time-consistent evaluations.

5.4. Calibration to real data

A common practice for calibrating a model in finance is to choose the set of parameters that minimizes the sum of the squared differences between the market prices and the prices implied by the model. However, this approach does not

Fig. 5.2. Convergence with respect to Δ_λ .Fig. 5.3. Expected number of defaults $\mathbb{E}[N_{10}]$ in function of θ with $\eta = 0$.

make sense in our case since no market prices are available. We must therefore calibrate the model on past losses that we assume to be representative of the future. Assume that we observed such data on the period of time $[0, T]$. The data consists of pairs $D := \{(t_i, m_i) \in [0, T] \times \mathbb{R}^+ : 1 \leq i \leq n\}$ that describe the times $0 \leq t_1 < \dots < t_n \leq T$ of the defaults that occurred between 0 and T and the corresponding lost amounts $m_1, \dots, m_n > 0$ in the $N_T = n$ case. Then, assuming that

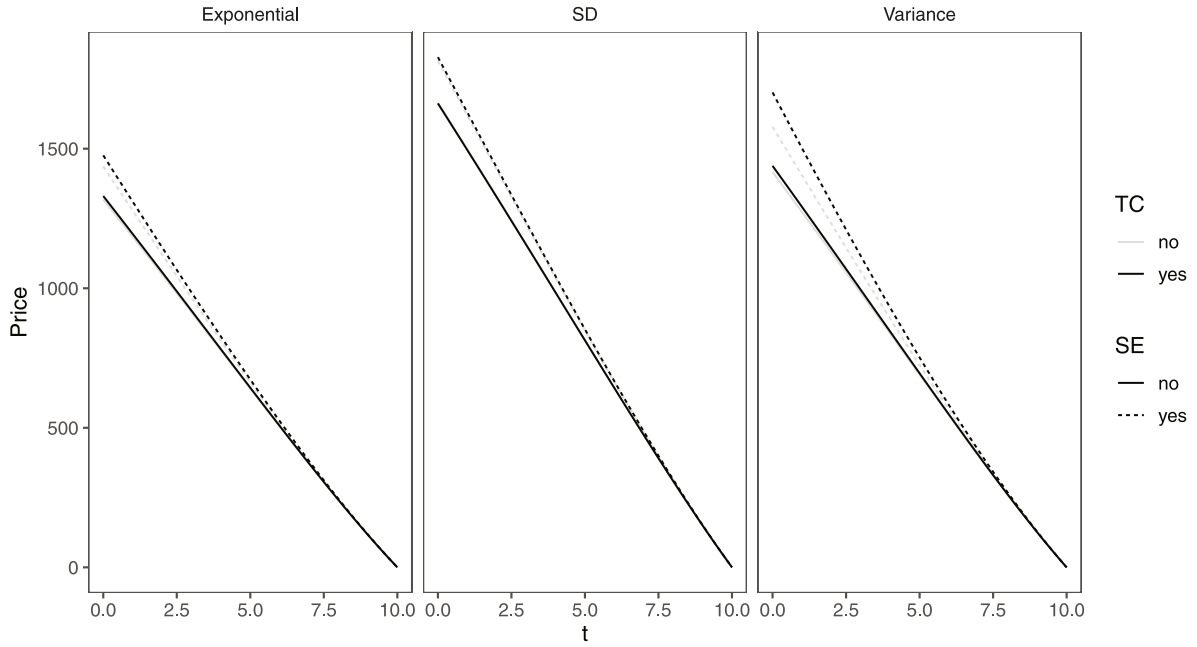


Fig. 5.4. Evolution of the time-consistent and time-inconsistent prices, with and without contagion.

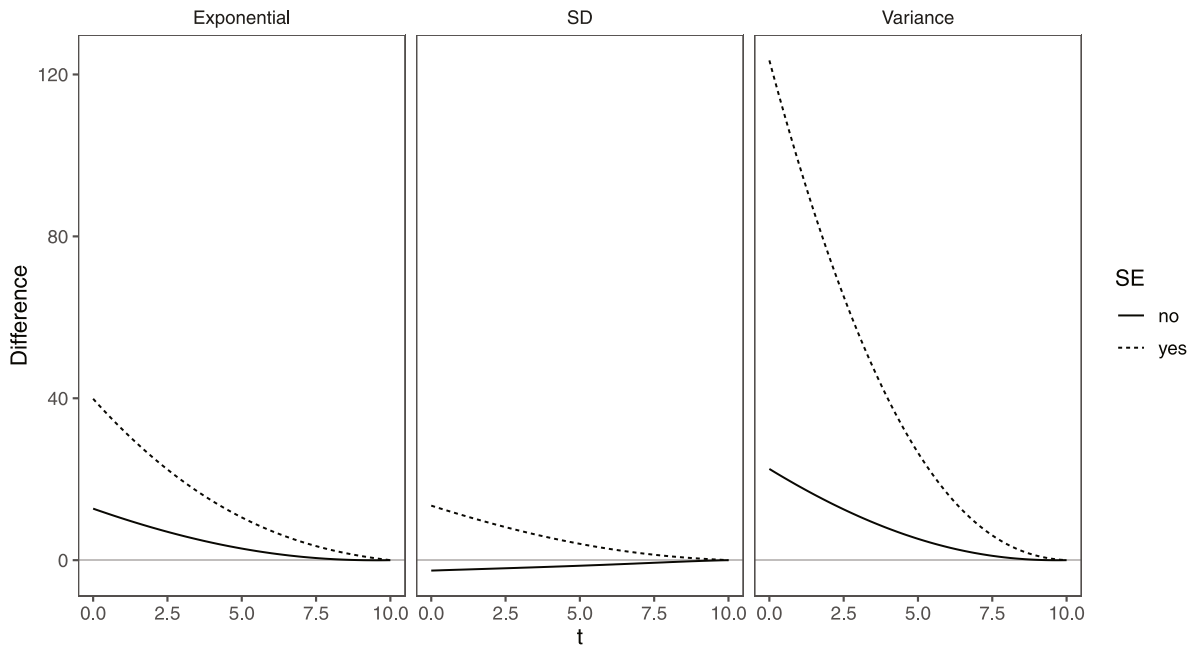


Fig. 5.5. Time-consistent evaluation minus their time-inconsistent counterparts.

$\sigma = 0$, we can numerically obtain the parameters by maximizing the likelihood. The reason for setting $\sigma = 0$ is the fact that the brownian term in the intensity $(\lambda_t)_{t \geq 0}$ cannot be observed in the data D . Proposition 7.2.III in [36] states that the log-likelihood function L associated to $(N_t)_{t \geq 0}$ is given by

$$L(D) = \sum_{j=1}^n \ln(\lambda_{t_j-}) - \int_0^T \lambda_t dt + T,$$

where λ_{t-} denotes the left limit $\lim_{s \uparrow t} \lambda_s$. In particular, note that $\lambda_{t_j-} = \lambda_{t_j} - m_j$. Following Eq. (2.3), for any $t \in [0, T]$, the observed past intensity satisfies

$$\lambda_t = \theta + e^{-\kappa t}(\lambda_0 - \theta) + \eta \sum_{t_i \leq t} e^{-\kappa(t-t_i)} m_i \quad (5.12)$$

so that

$$\sum_{j=1}^n \ln(\lambda_{t_j-}) = \sum_{j=1}^n \ln \left(\theta + e^{-\kappa t_j}(\lambda_0 - \theta) + \eta \sum_{i=1}^{j-1} e^{-\kappa(t_j-t_i)} m_i \right).$$

Moreover,

$$\int_0^T \lambda_t dt = \theta T + (\lambda_0 - \theta) \frac{1 - e^{-\kappa T}}{\kappa} + \eta \int_0^T \sum_{t_i \leq t} e^{-\kappa(t-t_i)} m_i dt.$$

The last term can be computed as

$$\begin{aligned} & \int_0^T \sum_{t_i \leq t} e^{-\kappa(t-t_i)} m_i dt \\ &= \underbrace{\int_0^{t_1} \sum_{t_i \leq t} e^{-\kappa(t-t_i)} m_i dt}_{=0} + \sum_{j=1}^{n-1} \int_{t_j}^{t_{j+1}} \sum_{t_i \leq t} e^{-\kappa(t-t_i)} m_i dt + \int_{t_n}^T \sum_{t_i \leq t} e^{-\kappa(t-t_i)} m_i dt \\ &= \sum_{j=1}^{n-1} \sum_{i=1}^j m_i \int_{t_j}^{t_{j+1}} e^{-\kappa(t-t_i)} dt + \sum_{i=1}^n m_i \int_{t_n}^T e^{-\kappa(t-t_i)} dt \\ &= \sum_{j=1}^{n-1} \sum_{i=1}^j \frac{m_i e^{\kappa t_i}}{\kappa} (e^{-\kappa t_j} - e^{-\kappa t_{j+1}}) + \sum_{i=1}^n \frac{m_i e^{\kappa t_i}}{\kappa} (e^{-\kappa t_n} - e^{-\kappa T}). \end{aligned}$$

Hence, we conclude that the log-likelihood L can be expressed as

$$\begin{aligned} L(D) &= \sum_{j=1}^n \ln \left(\theta + e^{-\kappa t_j}(\lambda_0 - \theta) + \eta \sum_{i=1}^{j-1} e^{-\kappa(t_j-t_i)} m_i \right) + (1 - \theta)T - (\lambda_0 - \theta) \frac{1 - e^{-\kappa T}}{\kappa} \\ &\quad - \frac{\eta}{\kappa} \left(\sum_{j=1}^{n-1} \sum_{i=1}^j m_i (e^{-\kappa(t_j-t_i)} - e^{-\kappa(t_{j+1}-t_i)}) \right) - \frac{\eta}{\kappa} \sum_{i=1}^n m_i (e^{-\kappa(t_n-t_i)} - e^{-\kappa(T-t_i)}). \end{aligned}$$

Considering L as a function of the parameters, the maximum likelihood estimators can be found by a numerical maximization. Then we can use the default costs $m_1, m_2, \dots, m_n$ to determine the jump size distribution ν . For example, with our assumption that the jump sizes are exponentially distributed, the parameter $\rho > 0$ of the exponential distribution can be approximated by $\hat{\rho} = (\frac{1}{n} \sum_{i=1}^n m_i)^{-1}$, which is also the maximum likelihood estimator.

We assess the performance of the proposed calibration procedure with the help of simulations. In order to do so, we generated 10,000 paths of the processes $(\lambda_t)_{t \geq 0}$ and $(P_t)_{t \geq 0}$ via an Euler discretization scheme. More precisely, the discretized path $(\lambda_0, \lambda_\Delta, \lambda_{2\Delta}, \dots, \lambda_{d\Delta} = \lambda_T)$ of $(\lambda_t)_{t \geq 0}$ is generated through the following procedure: we set $\lambda_0 \in \mathbb{R}$ and recursively compute

$$\lambda_{(j+1)\Delta} = \lambda_{j\Delta} + (\theta - \lambda_{j\Delta})\Delta + \eta(X_{j+1} \wedge 1)Y_{j+1}$$

where X_{j+1} is Poisson distributed with parameter $\Delta\lambda_{j\Delta}$ and Y_{j+1} is exponentially distributed with parameter ρ . All the Poisson and exponential random variables are of course independent. The discretized path $(P_0 = 0, P_\Delta, P_{2\Delta}, \dots, P_{d\Delta} = P_T)$ of $(P_t)_{t \geq 0}$ is then simply

$$P_{j\Delta} = \sum_{i=1}^j (X_i \wedge 1)Y_i.$$

The observation period was set to $T = 10$ years and the number of discretization steps was set to $d = 3650$, giving a daily discretization. To simplify the calibrations, we assumed that $\lambda_0 = \theta$. This is not a strong assumption because Eq. (2.3) implies that the impact of the gap $(\lambda_0 - \theta)$ between λ_0 and θ on the intensity $(\lambda_t)_{t \geq 0}$ is $e^{-\kappa t}(\lambda_0 - \theta)$ at time t , i.e. it decreases exponentially with time. Our 10,000 simulations gave us 10,000 values of the maximum likelihood estimator $(\hat{\eta}, \hat{\kappa}, \hat{\rho}, \hat{\theta})$ of $(\eta, \kappa, \rho, \theta)$. The average values of the 10,000 estimations are compared to the real parameters used for the

Table 5.2
Real parameters and average of 10,000 maximum likelihood estimations.

	η	κ	ρ	θ
Real value	0.002	5	0.0005	1
Average estimation	0.002485	5.149	0.0005077	1.067

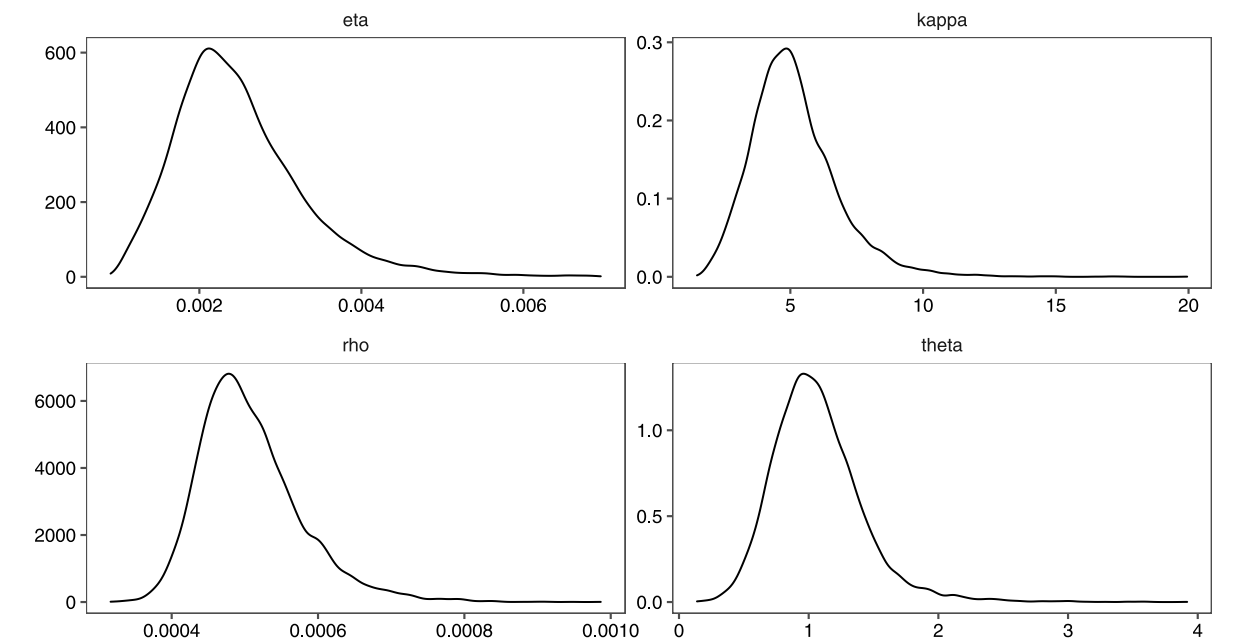


Fig. 5.6. Density plots of the 10,000 maximum likelihood estimations.

simulations in Table 5.2. Fig. 5.6 displays density plots based on the 10,000 values of the maximum likelihood estimators. We observe that all the estimators are reasonably close to the true values on average.

We close this section by pointing out the interest of Eq. (5.12) for a practitioner. Once the parameters are estimated, this equation allows to retrieve the value of the intensity at each time t on the basis of the history of past defaults. As a consequence, one is always able to compute the time-consistent price at all time so that our results are usable in practice.

6. Conclusions

In this paper we propose a model for credit insurance portfolio allowing to take the risk of contagion into account. This is done with the help of self-exciting processes. We start by providing closed-form formula for the two first moments of such processes. Then we derive semi-closed form expressions for the Laplace transform of these self-exciting processes. These semi-closed form expressions are then used to derive the distribution of the time of default in a credit insurance model relying on self-exciting processes. We derive PIDEs for time-consistent evaluations of credit insurance portfolio, extending the method of Pelsser and Ghalehjooghi [8] to models which incorporate self-excitation. In order to numerically solve these PIDEs, we describe a method allowing to find the solution, in spite of the lack of boundary conditions of our problem. By solving numerically these PIDEs, we are able to bring to light the effect of time-consistency on the prices and possible interactions between time-consistency of evaluations and the existence of a risk of contagion. More precisely, we showed that the prices tend to be higher when they are computed with time-consistent evaluations. We also showed that the difference is greater when a risk of contagion exists, motivating the use of time-consistent evaluations in such contexts. Finally, we propose a calibration of the model via likelihood maximization, which qualifies our results for practical use.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was also supported by the Fonds de la Recherche Scientifique - FNRS under Grant n° 33658713. The authors thank the reviewer for helpful comments.

References

- [1] P. Artzner, F. Delbaen, J. Eber, D. Heath, Coherent measures of risk, *Math. Finance* 9 (3) (1999) 203–228.
- [2] R. Kaas, M. Goovaerts, J. Dhaene, M. Denuit, *Modern Actuarial Risk Theory Using R*, Springer Verlag, 2008.
- [3] F. Black, M. Scholes, The pricing of options and corporate liabilities, *J. Polit. Econ.* 81 (1973) 637–659.
- [4] F. Delbaen, W. Schachermayer, A general version of the fundamental theorem of asset pricing, *Math. Ann.* 300 (3) (1994) 463–520.
- [5] F. Riedel, Dynamic coherent risk measures, *Stochastic Process. Appl.* 112 (2) (2004) 185–200.
- [6] E. Rosazza Gianin, Risk measures via g-expectations, *Insurance Math. Econom.* 39 (1) (2006) 19–34.
- [7] A. Pelsner, M. Stadje, Time-consistent and market-consistent evaluations, *Math. Finance* 24 (1) (2014) 25–65.
- [8] A. Pelsner, A. Ghalehjooghi, Time-consistent actuarial valuations, *Insurance Math. Econom.* 66 (2016) 97–112.
- [9] R. Merton, On the pricing of corporate debt : The risk structure of interest rates, *J. Finance* 29 (2) (1974) 449–470.
- [10] F. Black, J.C. Cox, Valuing corporate securities : Some effects of bonds indenture provisions, *J. Finance* 31 (1976) 351–367.
- [11] R. Geske, The valuation of corporate liabilities as compound options, *J. Financ. Quant. Anal.* 12 (14) (1977) 541–552.
- [12] F.A. Longstaff, E.S. Schwartz, A simple approach for valuing risky fixed and floating rate debt, *J. Finance* 50 (1995) 789–819.
- [13] D. Duffie, K. Singleton, Modeling term structures of defaultable bonds, *Rev. Financ. Stud.* 12 (1999) 687–720.
- [14] T. Bielecki, Rutkowski, *Credit Risk : Modeling, Valuation and Hedging*, in: Springer Finance, Springer, 2002.
- [15] D. Duffie, K. Singleton, *Credit Risk : Pricing, Measurement, and Management*, Princeton University Press, 2003.
- [16] D. Lando, *Credit Risk Modeling : Theory and Applications*, in: Princeton Series in Finance, Princeton University Press, 2004.
- [17] M. Jeanblanc, M. Yor, M. Chesney, *Mathematical Methods for Financial Markets*, Springer, 2009.
- [18] W. Schoutens, J. Cariboni, *Lévy Processes in Credit Risk*, A John Wiley and Sons Ltd, 2009.
- [19] A. Hawkes, Point spectra of some mutually exciting point processes, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 33 (1971) 438–443.
- [20] A. Hawkes, Spectra of some self-exciting and mutually exciting point processes, *Biometrika* 58 (1971) 83–90.
- [21] A. Hawkes, D. Oakes, A cluster representation of a self-exciting process, *J. Appl. Probab.* 11 (1974) 493–503.
- [22] F. Musmeci, D. Vere-Jones, A space-time clustering model for historical earthquakes, *Ann. Inst. Statist. Math.* 44 (1992) 1–11.
- [23] Y. Ogata, Space-time point process models for earthquake occurrences, *Ann. Inst. Statist. Math.* 50 (2) (1998) 397–402.
- [24] M.D. Porter, G. White, Self-exciting hurdle models for terrorist activity, *Ann. Appl. Stat.* 6 (1) (2012) 106–124.
- [25] L. Bauwens, N. Hautsch, Modelling financial high frequency data using point processes, in: *Handbook of Financial Time Series*, Springer Nature, 2009, pp. 953–979.
- [26] E. Bacry, I. Mastromatteo, J. Muzy, Hawkes processes in Finance, *Mark. Microstruct. Liq.* 1 (1) (2015) 1550005.
- [27] D. Hainaut, Clustered Lévy processes and their financial applications, *J. Comput. Appl. Math.* 319 (2017) 117–140.
- [28] D. Hainaut, F. Moraux, A switching self-exciting jump diffusion process for stock prices, *Ann. Financ.* 15 (2) (2019) 267–306.
- [29] D. Hainaut, S. Goutte, A switching microstructure model for stock prices, *Math. Final. Econom.* 13 (2019) 459–490.
- [30] Y. Ma, W. Xu, Structural credit risk modelling with Hawkes jump diffusion processes, *J. Comput. Appl. Math.* 303 (2016) 69–80.
- [31] A. Reinhard, A review of self-exciting spacio-temporal point processes and their applications, *Statist. Sci.* 33 (3) (2018) 299–318.
- [32] Y. Ait-Sahalia, J. Cacho-Diaz, R. Laeven, Modeling financial contagion using mutually exciting jump processes, *J. Finan. Econ.* 117 (3) (2015) 585–606.
- [33] J. Pratt, Risk aversion in the small and in the large, *Econometrica* 32 (1) (1964) 122–136.
- [34] P. Protter, *Stochastic Integration and Differential Equations. Second Edition, Version 2.1*, Springer, 2003.
- [35] C. Rogers, D. Williams, *Diffusions, Markov Processes and Martingales, Vol. 2*, Cambridge Mathematical Library, 2000.
- [36] D.J. Daley, D. Vere-Jones, *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*, second ed., Springer, 2002.