Optimal dividends for a NatCat insurer in the presence of a climate tipping point

Autoría no ditelliana: Albrecher, Hansjörg

Autoría ditelliana: Azcue, Pablo; Muler, Nora

Fecha de publicación: 24/11/2025

Publicado originalmente en: Canadian Journal of Statistics (e-ISSN: 1708-945X)

The Canadian Journal of Statistics | La revue canadienne de statistique
Vol. 000, No. 00, 2025, e70029

Optimal dividends for a NatCat insurer in the presence of a climate tipping point

Hansjörg ALBRECHER @
Department of Actuarial Science, Faculty of Business and Economics, Swiss Finance Institute and Expertise Center for Climate Extremes (ECCE), University of Lausanne, Lausanne, Switzerland

Pablo AZCUE and **Nora MULER**
Departamento de Matematicas, Universidad Torcuato Di Tella, Ciudad de Buenos Aires, Argentina

Abstract We study optimal dividend strategies for an insurance company facing natural catastrophe claims, anticipating the arrival of a climate tipping point after which the claim intensity and/or the claim size distribution of the underlying risks deteriorates irreversibly. For shot-noise Cox claim arrivals, we show that the nonstationary feature of such a tipping point can be an advantage for shareholders. Assuming the tipping point arrives according to an Erlang distribution, we demonstrate that the corresponding system of two-dimensional stochastic control problems admits a viscosity solution, which can be numerically approximated in an efficient way. The results are illustrated through several numerical examples. Also, the sensitivity of the optimal dividend strategies to the presence of a climate tipping point is analyzed.

Résumé Nous étudions des stratégies de dividendes optimales pour une compagnie d'assurance confrontée à des sinistres liés àux catastrophes naturelles, en anticipant l'arrivée d'un point de bascule climatique après lequel l'intensité des sinistres et/ou la distribution des montants de sinistres des risques sous-jacents se détériore de manière irréversible. Pour des arrivées de sinistres modélisées par un processus de Cox à bruit de choc, nous montrons que la caractéristique non stationnaire induite par un tel point de bascule peut constituer un avantage pour les actionnaires. En supposant que le point de bascule survienne selon une loi d'Erlang, nous démontrons que le système correspondant de problèmes de contrôle stochastique bidimensionnels admet une solution de viscosité, qui peut être approchée numériquement de manière efficace. Les résultats sont illustrés par plusieurs exemples numériques, et la sensibilité des stratégies de dividendes optimales à la présence d'un point de bascule climatique est analysée.

Keywords Climate change; Hamilton–Jacobi–Bellman equation; insurance; optimal dividends; shot-noise Cox process; viscosity solution.
MSC2020 Primary 91G50; Secondary 49L25; 93E20.
@ Corresponding author hansjoerg.albrecher@unil.ch

1 Introduction

In recent years, it has become amply clear that climate change leads to more frequent and more severe weather extremes (see e.g. Seneviratne et al. (2021)), thus leading to an increase of natural catastrophes, cf. e.g. Lloyd and Shepherd (2020). Correspondingly, the insurance of losses due to natural catastrophes (NatCat) is becoming increasingly challenging, see e.g. Schneider (2023). Traditionally, the reinsurance market was able to provide protection for respective lines of business to a large extent (see Albrecher et al. (2017) for a survey on reinsurance from an actuarial perspective), but due to these recent developments, reinsurance premiums for several NatCat perils are rising, causing increasing concern in the insurance industry. In this context, it is of interest to see whether the insurance of NatCat lines remains feasible and/or attractive for insurance companies. Among several different angles from which one might want to study such a question, in Albrecher et al. (2024) we recently looked into the insurance of NatCat risks from a shareholder perspective. In particular, we studied whether the variability of claim intensities over time, which naturally appears in NatCat lines, can also have upward potential for shareholders when performance is measured in terms of the expected value of the aggregate discounted future dividends from the insurance portfolio, a criterion that goes back to De Finetti (1957) and has been studied in terms of optimal respective dividend strategies in risk theory ever since, see e.g. Albrecher and Thonhauser (2009) and Avanzi (2009) for surveys. In Albrecher et al. (2024), it was recently shown that if the Poisson claim intensity in a NatCat portfolio is modelled by a (stochastic) shot-noise process and this process is observable, then the variability of the claim intensity can indeed enhance the value of the insurance portfolio, provided the dividend strategy is suitably adapted to the evolution of the observed claim intensity over time. For the solution of the respective two-dimensional optimal control problem, it was necessary to assume a certain degree of stationarity of this shot-noise process, restricting the considerations to NatCat portfolios for which neither the intensity nor the claim severity distribution could systematically increase over time.

In this article, we go one step further and allow for such a systematic change (in the sense of a deterioration) in the analysis. To ensure that the typical stochastic control techniques from previous studies remain applicable, we embed this extension within a Markovian framework by assuming that a "change-point" in the conditions occurs at a stochastic time, driven by an independent renewal process with specified characteristics. Concretely, we consider a Cox process for the claim arrivals due to a shot-noise process as in Albrecher et al. (2024), but at an exponentially (or Erlang) distributed time, the dynamics of the process changes to a more severe situation. While the developed technique will, in principle, allow for several such changes over time, we restrict the considerations to a single occurrence in this article, simplifying the presentation and numerical implementation when deriving the optimal dividend strategy. One can view such a sudden change as a discrete and mathematically tractable approximation of an increasing trend. On the other hand, such a "change-point" also has a natural interpretation in terms of a *climate tipping point*, which is a critical threshold that, once crossed, leads to an irreversible change in climatic conditions. This can happen in response to a gradual warming of air and ocean surface temperature. In recent years, an increasing number of climate scientists predicted such a tipping point to occur with a large probability (see e.g. Lenton et al. (2019)), and we choose to focus on the tipping point interpretation of our model in the rest of this article. Here we assume that, after the occurrence of this tipping point, we will have a higher intensity of claims and/or a change in the claim severity distribution. Our goal is to see that if one is able to actually observe the occurrence of such a tipping point and has realistic assumptions on the implications for the risk characteristics of a NatCat portfolio, it is possible to derive an optimal dividend strategy for an insurance portfolio exposed to this risk. We assume independent and identically distributed (i.i.d.) claims, but with a potentially different distribution before and after the tipping point. The premiums for the policyholders are calculated according to an expected value principle (with a safety loading), which stays constant before the tipping point (so that the risk of changed claim intensities due to the arrival of catastrophes is a priori borne by the insurance company and diversified over time), but at the time of the tipping point, this premium is adapted to its new level because of the systematically increased risk, and then kept constant again at that higher level.

The setup will allow us to solve the optimal dividend problem as an iteration of two-dimensional optimal control problems, where one variable is the current surplus of the portfolio and the other is the current claim intensity level. We will adapt the results of Albrecher et al. (2024) to solve an auxiliary optimization problem that consists of determining the optimal dividend strategy for a compound Cox process with shot-noise intensity for an exponential time horizon. The solution of this problem can then be concatenated to provide the general solution for the optimal dividend problem in the presence of the tipping point. As it will turn out, also in this situation, the expected discounted dividends for the shareholders increase in the case where the processes are observable and the premium adaptation is done instantaneously.

On the conceptual level, there are some similarities between the present article and dividend problems in Markov-modulated risk models, see e.g. Wei et al. (2010), Jiang and Pistorius (2012), and Szölgyenyi (2015), but here we switch only in one direction. On the other hand, our underlying compound Cox risk model makes the analysis of the stochastic control problem much more complicated, and the exposition is necessarily relatively technical. We will also deal with a situation where the tipping point arrives at an Erlang(n) time, which introduces memory into the arrival process of the tipping point, but is still computationally tractable. Technically,

DOI: 10.1002/cjs.70029

this can be handled as n consecutive switches at independent exponential times, an idea that goes back, in mathematical finance, to Carr (1998) and in actuarial science to Asmussen et al. (2002), and which was already used in another context of dividend optimization in Albrecher and Thonhauser (2012).

Note that the method we employ here can also be useful in other settings where one needs to solve iterative optimal control problems and the solution of the previous problem serves as a boundary value of the next one. While the computational complexity might increase when one wants to apply multiple iterations of such a procedure, it is still feasible methodologically. In that vein, one could, for instance, approximate a deterministic time horizon for a change of the model dynamics by an Erlang(n) distribution with higher n at the expense of n iterated stochastic problems, and one could also introduce several (instead of one) "deterioration checkpoints" of environmental conditions in the insurance activity. More generally, the method proposed in this article might be useful in other contexts where, for computational reasons, one would like to approximate some nonstationary phenomenon by iterative applications of Markovian sections that pass on their result as a boundary condition to the next level.

The rest of the article is organized as follows. Section 2 formulates an auxiliary optimal control problem that will be used later as a building block for the general model. This auxiliary problem setting with a predefined exit value at a stopping time may also be of interest in its own right for other modelling circumstances. Some basic properties of the respective value function are derived. Section 3 then derives the Hamilton–Jacobi–Bellman (HJB) equation of this auxiliary problem and shows that the optimal value function is the smallest viscosity supersolution of the HJB equation. Section 4 provides an important boundary result as well as an asymptotic property of this value function. In Section 5, a numerical scheme is introduced that approximates the optimal value function on a finite grid in two-dimensional space, and it is shown that this approximation converges uniformly to the optimal value function. In Section 6, the auxiliary model studied in the previous sections is applied to solve a dividend problem for a NatCat insurer like the one in Albrecher et al. (2024), but now with the presence of a climate tipping point. In Section 7, we work out several concrete numerical illustrations that extend those in Albrecher et al. (2024) by the presence of such a climate tipping point and we interpret the results. Finally, Section 8 concludes. All proofs are in the Appendix.

2 An auxiliary optimization problem

Let us introduce an auxiliary optimization problem. It will be a crucial ingredient to solve inductively the problem of optimizing expected cumulative discounted dividend payouts in the presence of a climate tipping point, as described in Section 6.

We assume that the uncontrolled surplus of the company follows a compound Cox process with drift, generated by a shot-noise intensity until a stopping time given by a random exponential time $\overline{\tau} \sim Exp(\xi)$ independent of the compound Cox process. That is, for $t < \overline{\tau}$,

$$X_t = x + pt - \sum_{j=1}^{N_t} U_j, \tag{1}$$

where x is the initial surplus of the portfolio and p is the premium rate. Moreover, let us denote by τ_j the arrival time of claim j. We assume that the counting process

$$N_t = \#\{j : \tau_j \leq t\}$$

is an inhomogeneous Poisson process with (stochastic) intensity λ_t. The process N_t and the claim sequence $\{U_i\}_{i\geq 1}$ are independent of each other. The intensity $\boldsymbol{\lambda} = (\lambda_t)_{t\geq 0}$ is a shot-noise process given by

$$\lambda_t = \lambda_t^c + \sum_{k=1}^{\widetilde{N}_t} Y_k \, e^{-d(t-T_k)}, \tag{2}$$

where T_k is the arrival time of catastrophe k, which generates an upward jump in the intensity process of size Y_k. The positive random variables $(Y_k)_{k\geq 1}$ are assumed to be i.i.d. (and independent of $\widetilde{N}_t$) with distribution function F_Y and finite expectation. The arrival process of catastrophes

$$\widetilde{N}_t = \#\{k : T_k \leq t\}$$

is a Poisson process of constant intensity β, independent of the current intensity value. After the arrival of a catastrophe, the induced intensity decays deterministically and exponentially with rate $d > 0$. $\underline{\lambda} > 0$ is a baseline intensity and

$$\lambda_t^c = \underline{\lambda} + e^{-dt}(\lambda - \underline{\lambda}) \tag{3}$$

is the sum of that baseline intensity and the deterministic decay function of the initial intensity level $\lambda = \lambda_0 \geq \underline{\lambda}$ (see for instance Albrecher and Asmussen (2006) for such a model assumption in risk theory). Note that $\lambda_t \geq \underline{\lambda}$ for all $t \geq 0$ and if the initial intensity satisfies $\lambda > \underline{\lambda}$, then $\lambda_t > \underline{\lambda}$ for all $t \geq 0$. The baseline $\underline{\lambda}$ can be interpreted as the claim intensity of "regular" noncatastrophic claims in the portfolio. All quantities are considered in a suitable compound Cox filtered space $\left(\Omega^{\lambda}, \mathcal{F}^{\lambda}, \left(\mathcal{F}_t^{\lambda}\right)_{t\geq 0}, \mathbb{P}^{\lambda}\right)$ conditional on the process $\boldsymbol{\lambda}$; for details see Albrecher et al. (2024, Section 2) and Dassios and Jang (2003).

The company pays dividends to the shareholders up to $\overline{\tau}$ or the ruin time (whichever comes first). We denote the dividend strategy as $L = (L_t)_{t\geq 0}$, where L_t is the amount of cumulative dividends paid up to time t. To extend the definition of L_t for all $t \geq 0$, let us consider L_t constant after the minimum between $\overline{\tau}$ and the ruin time. The dividend payment strategy is admissible if it is nondecreasing, càdlàg (right-continuous with left limits), adapted to $\left(\mathcal{F}_t^{\lambda}\right)_{t\geq 0}$, and satisfies $L_0 \geq 0$ and $L_t \leq X_t$ for $t < \tau^L$, where the ruin time τ^L is defined as

$$\tau^L = \inf\{t \geq 0 : X_t - L_{t^-} < 0\}. \tag{4}$$

This last condition means that ruin can happen only at the arrival of a claim and that no lump sum dividend payment can be made at the time of ruin. We define the controlled surplus process as

$$X_t^L = X_t - L_t. \tag{5}$$

Denote by $\Pi_{x,\lambda}$ the set of admissible dividend strategies starting with initial surplus level $x \geq 0$ and initial intensity $\lambda \geq \underline{\lambda}$. Let us define for a function $u : [0, \infty) \times [\underline{\lambda}, \infty) \to [0, \infty)$, the growth condition

$$u(x, \lambda) \leq K + x \text{ for all } (x, \lambda) \in (0, \infty) \times (\underline{\lambda}, \infty) \text{ and some } K > 0. \tag{6}$$

We consider a function $g : [0, \infty) \times [\underline{\lambda}, \infty) \to [0, \infty)$, which represents in the auxiliary model the exit value of the company at time $\overline{\tau}$, and we study the problem of maximizing the expected cumulative discounted dividend payouts. We assume the following conditions on g:

Assumption (A1). g is uniformly continuous in $[0, \infty) \times [\underline{\lambda}, \infty)$, locally Lipschitz in $(0, \infty) \times (\underline{\lambda}, \infty)$, nondecreasing in x, nonincreasing in λ, such that

$$\lim_{\lambda\to\infty} \sup_x (g(x, \lambda) - x) = 0,$$

$g_x \geq 1$ a.e. and $g(x, \lambda) = (x - \hat{x}) + g(\hat{x}, \lambda)$ for some $\hat{x} \geq 0$.

Assumption (A1) for the function g collects, in fact, the properties of the optimal value function that maximizes the expected cumulative discounted dividend payments in the case where the uncontrolled surplus of the company follows a compound Cox process with drift generated by a shot-noise intensity until ruin without a switching time (see Albrecher et al. (2024)) and these are the natural assumptions for the climate tipping point problem (see Section 6).

For any initial surplus level $x \geq 0$ and initial intensity $\lambda \geq \underline{\lambda}$, we define the optimal value function as

$$V(x, \lambda) = \sup_{L\in\Pi_{x,\lambda}} J(L; x, \lambda), \tag{7}$$

where

$$J(L; x, \lambda) = \mathbb{E}\left(\int_{0^-}^{\overline{\tau}\wedge\tau^L} e^{-qs} dL_s + e^{-q(\overline{\tau}\wedge\tau^L)} I_{\{\overline{\tau}\wedge\tau^L<\tau^L\}} g(X^L_{\overline{\tau}\wedge\tau^L}, \lambda_{\overline{\tau}\wedge\tau^L})\right). \tag{8}$$

Here, $q > 0$ is a constant discount factor. We assume that the premium rate p is obtained using the expected value principle of the asymptotic distribution of λ_t. That is

$$p = (1 + \eta)\mathbb{E}(U_1)\lambda_{\text{av}}, \tag{9}$$

where $\eta > 0$ is the safety loading,

$$\Lambda_t = \int_0^t \lambda_s ds, \tag{10}$$

and

$$\lambda_{\text{av}} = \lim_{t\to\infty} \mathbb{E}\left(\frac{\Lambda_t}{t}\right) = \underline{\lambda} + \frac{\beta\, \mathbb{E}(Y_1)}{d}. \tag{11}$$

The optimal value function V will inherit the basic properties from g, as we will show in the next proposition and in Section 4.

In the remainder of this section, we prove some elementary properties of the optimal value function. The proofs of the propositions in this section are given in Appendix A1.

Proposition 1. *V is nonincreasing in λ, uniformly continuous in $[0,\infty)\times[\underline{\lambda},\infty)$, locally Lipschitz in $[0,\infty)\times(\underline{\lambda},\infty)$, nondecreasing in x, and satisfies the growth condition (6).*

We now state the so-called dynamic programming principle (DPP), but skip its proof as it is similar to that of Lemma 3.6 of Albrecher et al. (2024) and Lemma 1.2 of Azcue and Muler (2014).

Lemma 1. *For any initial surplus $x \geq 0$ and any finite stopping time τ, we can write*

$$V(x,\lambda) = \sup_{L\in\Pi_{x,\lambda}} \left(\mathbb{E}(\int_{0^-}^{\tau\wedge\overline{\tau}\wedge\tau^L} e^{-qs} dL_s + I_{\{\overline{\tau}\wedge\tau\wedge\tau^L=\overline{\tau}\}} g(X^L_{\overline{\tau}}, \lambda_{\overline{\tau}}) + I_{\{\overline{\tau}\wedge\tau\wedge\tau^L=\tau\}} e^{-q\tau} V(X^L_{\tau}, \lambda_\tau) \right).$$

In the next proposition, we show the asymptotic behaviour of V as λ goes to infinity.

Proposition 2. *For all $x \geq 0$, it holds that $\lim_{\lambda\to\infty} V(x,\lambda) = x$.*

3 HJB equation and viscosity solutions for the auxiliary problem

In this section we obtain the HJB equation associated with the auxiliary problem and state that the optimal value function V defined in (7) is a viscosity solution. We also show that the value function of any admissible strategy is smaller than any supersolution of the HJB equation. Finally, we characterize the optimal value function V as the smallest viscosity supersolution of the HJB equation. These results are a generalization of those given in Albrecher et al. (2024), Section 4, in the absence of a switching time. The proofs of the results of this section are deferred to Appendix A2.

The HJB equation of the optimization problem (7) is given by

$$\max\{\mathcal{L}(V)(x,\lambda), 1 - V_x(x,\lambda)\} = 0, \tag{12}$$

where

$$\begin{aligned} \mathcal{L}(V)(x,\lambda) &= pV_x(x,\lambda) - d(\lambda-\underline{\lambda})V_\lambda(x,\lambda) - (q+\lambda+\beta)V(x,\lambda) \\ &\quad + \lambda\int_0^x V(x-\alpha,\lambda)dF_U(\alpha) + \beta\int_0^\infty V(x,\lambda+\gamma)dF_Y(\gamma) + \xi(g(x,\lambda) - V(x,\lambda)). \end{aligned} \tag{13}$$

While the derivation of equation (12) follows standard procedures, it may be instructive to interpret it concretely. At each point in time, one has to decide, based on the current surplus level x and intensity level λ, whether paying dividends is preferable to not paying them. The left-hand side in the max operator refers to the case of no dividend payment. In that scenario, the process evolves according to the infinitesimal generator given in (13). As the premium rate is p, the infinitesimal change in the x-direction in the absence of a claim is given by $pV_x(x,\lambda)$. A claim arrives with intensity λ, so we might interpret this as the surplus process being "killed" at that rate (term $-\lambda V(x,\lambda)$) and immediately "revived" with a modified surplus level that depends on the claim size with cdf F_U (integral term $\lambda\int_0^x V(x-\alpha,\lambda)dF_U(\alpha)$). In the λ-direction, in the absence of a jump in the shot-noise process, the infinitesimal change in the claim intensity level is described by the term $-d(\lambda-\underline{\lambda})V_\lambda(x,\lambda)$ (which, due to the Markov property of the intensity process, also comprises the decay of all previous intensity shots), and since an intensity shot arrives with intensity β, we might interpret this as the process being "killed" ($-\beta V(x,\lambda)$) and "revived" again with the respective up-jump with cdf F_Y (which is captured in the integral term $\beta\int_0^\infty V(x,\lambda+\gamma)dF_Y(\gamma)$). Finally, the term $-qV(x,\lambda)$ accounts for the time value due to the discounting of future payments, and the term $\xi(g(x,\lambda)-V(x,\lambda))$ keeps track of switching events, which occur with intensity ξ and replace $V(x,\lambda)$ with its boundary value $g(x,\lambda)$ upon switching. Note that $\xi=0$ leads to the integro-differential operator considered in Albrecher et al. (2024) without switching time $\overline{\tau}$. Moreover, one can interpret that it is not optimal to pay out dividends if $V_x(x,\lambda)$ is larger than 1, because paying out dividends at that moment will realize a slope of 1 in the direction of the surplus x. This means that one benefits from the dynamics of the surplus process more than when paying dividends, translating into the right-hand side of (12) being negative in that case.

The structure of equation (12) suggests that the state space $[0,\infty)\times[\underline{\lambda},\infty)$ is partitioned into a no-action region $\mathcal{NA}$ (in which no dividends are paid) and an action region $\mathcal{A}$ (in which dividends are paid). The points in the $\mathcal{NA}$ region satisfy $\mathcal{L}(V)=0$ and $1-V_x<0$, whereas the points in $\mathcal{A}$ satisfy $\mathcal{L}(V)\le 0$ and $1-V_x=0$.

Because of Proposition 1, the optimal value function V is locally Lipschitz. However, it might not be differentiable at some points, so we cannot claim that V is a solution of the HJB equation in the classical sense. We prove instead that V is a viscosity solution of the HJB equation. Let us define this standard notion.

Definition 1. A uniformly continuous function $\overline{u}:[0,\infty)\times\left[\underline{\lambda},\infty\right)\to\mathbb{R}$, that is locally Lipschitz in $(0,\infty)\times\left[\underline{\lambda},\infty\right)$, is a *viscosity supersolution* of (12) at $(x,\lambda)\in(0,\infty)\times(\underline{\lambda},\infty)$ if any continuously differentiable function $\varphi:(0,\infty)\times\left[\underline{\lambda},\infty\right)\to\mathbb{R}$ such that $\varphi(y,\cdot)$ is bounded for any $y\ge 0$, $\varphi(x,\lambda)=\overline{u}(x,\lambda)$, and such that $\overline{u}-\varphi$ reaches its minimum at (x,λ), satisfies

$$\max\{\mathcal{L}(\varphi)(x,\lambda),1-\varphi_x(x,\lambda)\}\le 0.$$

A uniformly continuous function $\underline{u}:[0,\infty)\times\left[\underline{\lambda},\infty\right)\to\mathbb{R}$, that is locally Lipschitz in $(0,\infty)\times\left[\underline{\lambda},\infty\right)$, is a *viscosity subsolution* of (12) at $(x,\lambda)\in(0,\infty)\times(\underline{\lambda},\infty)$ if any continuously differentiable function $\psi:(0,\infty)\times(\underline{\lambda},\infty)\to\mathbb{R}$ such that $\psi(y,\cdot)$ is bounded for any $y\ge 0$, $\psi(x,\lambda)=\underline{u}(x,\lambda)$, and such that $\underline{u}-\psi$ reaches its maximum at (x,λ), satisfies

$$\max\{\mathcal{L}(\psi)(x,\lambda),1-\psi_x(x,\lambda)\}\ge 0.$$

A function $u:(0,\infty)\times(\underline{\lambda},\infty)\to\mathbb{R}$ which is both a supersolution and subsolution at $(x,\lambda)\in(0,\infty)\times(\underline{\lambda},\infty)$ is called a *viscosity solution* of (12) at (x,λ).

Proposition 3. *V is a viscosity solution of the HJB equation (12) at any $(x,\lambda)\in(0,\infty)\times(\underline{\lambda},\infty)$.*

The optimal value function can now be characterized in the following way:

Proposition 4. *The optimal value function V is the smallest viscosity supersolution of (12) among those that are nonincreasing in λ and satisfy the growth condition (6).*

Next, the usual viscosity verification result can be deduced.

Corollary 1. *Consider a family of admissible strategies $\{L^{x,\lambda}\in\Pi_{x,\lambda}:(x,\lambda)\in(0,\infty)\times(\underline{\lambda},\infty)\}$. If the function $W(x,\lambda):=J(L^{x,\lambda};x,\lambda)$ is a viscosity supersolution of (12), then $W(x,\lambda)$ is the optimal value function. Also, if for each $k\ge 1$ there exists a family of strategies $\{L_k^{x,\lambda}\in\Pi_{x,\lambda}:(x,\lambda)\in(0,\infty)\times(\underline{\lambda},\infty)\}$ such that $W(x,\lambda):=\lim_{k\to\infty}J(L_k^{x,\lambda};x,\lambda)$ is a viscosity supersolution of the HJB equation (12) in $(0,\infty)\times(\underline{\lambda},\infty)$, then W is the optimal value function V.*

4 Asymptotic properties of the optimal value function for the auxiliary problem

From Corollary 1 we get the next proposition, in which we prove that there exists an explicit threshold $\overline{x}$, such that if the surplus is above this explicit threshold, then one should pay at least the exceeding surplus as dividends, regardless of the current intensity of the shot-noise process; so, $[\overline{x},\infty)\times[\underline{\lambda},\infty)\subset\mathcal{A}$. However, as we will show in the examples, the boundaries between the no-action region $\mathcal{NA}$ and the action region $\mathcal{A}$ depend on λ, so it is a complex problem and the value $\overline{x}$ is just an upper bound. The proofs of the results of this section and some necessary auxiliary results are given in Appendix A3.

Proposition 5. *$V(x,\lambda) = x - \overline{x} + V(\overline{x},\lambda)$ for $x \geq \overline{x}$, where*

$$\overline{x} = \max\left\{\frac{p + \xi(g(\hat{x},\lambda) - \hat{x})}{q}, \hat{x}\right\}$$

and $\hat{x}$ is as defined in Assumption (A1).

We have the following uniform convergence as the current intensity of the shot-noise process goes to infinity.

Proposition 6. *It holds that* $\lim_{\lambda\to\infty}\left(\sup_{x\geq 0} V(x,\lambda) - x\right) = 0.$

5 Numerical scheme for the auxiliary problem

We describe here the numerical method, which is based on admissible strategies restricted to a finite grid. The numerical scheme is derived from the one in Azcue and Muler (2021) and it is a generalization of the one described in Sections 6–8 in Albrecher et al. (2024) when there is no-switching time $\overline{\tau}$ (i.e., $\xi = 0$). Proposition 6 suggests that we can approximate the optimal value function by one over a bounded grid on λ. Also, from Proposition 5, we can assume that $(x,\lambda) \in \mathcal{A}$ for $x > \overline{x}$. This allows us to make a numerical computation staying in a finite grid. For some $\delta > 0$, $\Delta > 0$, and $m_1 \in \mathbb{N}$, we define the finite grid

$$\overline{\mathcal{G}}_\delta \times \overline{\mathcal{H}}_\Delta,$$

where

$$\overline{\mathcal{G}}_\delta = \{x_n^\delta = np\delta : x_n^\delta \leq \overline{x}\} \text{ and } \overline{\mathcal{H}}_\Delta = \{\lambda_m^\Delta = \underline{\lambda} + m\Delta : 0 \leq m \leq m_1\}.$$

Let $n_1 = \max\left\{n : x_n^\delta \in \overline{\mathcal{G}}_\delta\right\}$. In order to discretize on λ, we consider the auxiliary process

$$\widehat{\lambda}_t = \sigma^\Delta(\lambda_t) \wedge \lambda_{m_1}^\Delta \in \overline{\mathcal{H}}_\Delta$$

instead of the original shot-noise process $\boldsymbol{\lambda} = (\lambda_t)_{t\geq 0}$, where

$$\sigma^\Delta(\lambda) = \min\{\lambda_m^\Delta \in \mathcal{H}_\Delta : \lambda_m^\Delta \geq \lambda\} \in \mathcal{H}_\Delta. \tag{14}$$

So, if $\lambda_t \leq \lambda_{m_1}^\Delta$, we are discretizing λ_t considering the closest largest point in $\mathcal{H}_\Delta$, and if $\lambda_t \geq \lambda_{m_1}^\Delta$, then $\widehat{\lambda}_t = \lambda_{m_1}^\Delta$. We also introduce the auxiliary function

$$\rho^\delta(x) = \max\{x_n^\delta \in \overline{\mathcal{G}} : x_n^\delta \leq x\}, \tag{15}$$

which gives the closest point of the grid $\overline{\mathcal{G}}_\delta$ below x.

Define now the following local strategies: either the company pays no dividends or pays immediately a lump sum $p\delta$ as dividends; moreover, the company can finish the insurance contract at any time. We modify these local strategies in such a way that the controlled surplus and the intensity always lie in $\overline{\mathcal{G}}_\delta \times \overline{\mathcal{H}}_\Delta$ immediately after the arrival of a claim. Let τ^U and U be the arrival time and the size of the next claim, and T and Y be the arrival time and the size of the next intensity upward jump. We first define the three possible control actions at any point of $(x_n^\delta, \lambda_m^\Delta) \in \overline{\mathcal{G}}_\delta \times \overline{\mathcal{H}}_\Delta$ as follows:

- Control action $\mathbf{E}_0$: Pay no dividends up to the time $\delta \wedge \tau^U \wedge T \wedge \overline{\tau}$. The control action $\mathbf{E}_0$ can only be applied for current surplus with $n < n_1$.
    1. In the case that $\delta < \tau^U \wedge T \wedge \overline{\tau}$, the surplus at time δ is $x_{n+1}^\delta \in \overline{\mathcal{G}}_\delta$, and the intensity is $\widehat{\lambda}_\delta = \sigma^\Delta(\underline{\lambda} + e^{-d\delta}(\lambda_m^\Delta - \underline{\lambda}))$.
    2. If $\tau^U < \delta \wedge T \wedge \overline{\tau}$, the uncontrolled surplus at time τ^U is $x_n^\delta + \tau^U p - U$ and the intensity is $\widehat{\lambda}_{\tau^U} = \sigma^\Delta\left(\underline{\lambda} + e^{-d\tau^U}(\lambda_m^\Delta - \underline{\lambda})\right)$; if $x_n^\delta + \tau^U p - U$ is positive, the company pays immediately the minimum amount of dividends in such a way that the end surplus is $\rho^\delta(x_n^\delta + \tau^U p - U)$ and the amount paid as dividends is equal to $x_n^\delta + \tau^U p - U - \rho^\delta(x_n^\delta + \tau^U p - U)$. Ruin occurs in the case where the surplus $x_n^\delta + \tau^U p - U < 0$ at time $\tau^U \leq \delta \wedge T \wedge \overline{\tau}$.

DOI: 10.1002/cjs.70029

3. If $T < \delta \wedge \tau^U \wedge \overline{\tau}$, the end surplus is $\rho^\delta(x_n^\delta + Tp)$, the amount paid as dividends is $x_n^\delta + Tp - \rho^\delta(x_n^\delta + Tp)$, and the intensity is $\widehat{\lambda}_T = \sigma^\Delta(\underline{\lambda} + e^{-dT}(\lambda_m^\Delta - \underline{\lambda}) + Y)$.

4. If $\overline{\tau} < T \wedge \delta \wedge \tau^U$, the compound Cox process switches and the exit value of the company becomes $g(x_n^\delta + \overline{\tau}p, \widehat{\lambda}_{\overline{\tau}})$, where $\widehat{\lambda}_{\overline{\tau}} = \sigma^\Delta\left(\underline{\lambda} + e^{-d\overline{\tau}}(\lambda_m^\Delta - \underline{\lambda})\right)$.

- Control actions $\mathbf{E}_1$: The company pays immediately $p\delta$ as dividends, so the controlled surplus becomes $x_{n-1}^\delta \in \mathcal{G}_\delta$. The control action $\mathbf{E}_1$ can be applied only for current surplus $x_n^\delta > 0$.

- Control action $\mathbf{E}_F$: The company opts to pay the current surplus x_n^δ as dividends and immediately terminates the insurance policies.

We denote the space of control actions as

$$\mathcal{E} = \{\mathbf{E}_F, \mathbf{E}_1, \mathbf{E}_0\}. \tag{16}$$

The numerical scheme will choose inductively which of the possible local actions in each of the points in the grid is "better": $\mathbf{E}_0$ or $\mathbf{E}_1$ or $\mathbf{E}_F$. Roughly speaking, the idea of this construction is to find, for each point in $\overline{\mathcal{G}}_\delta \times \overline{\mathcal{H}}_\Delta$, the best local strategy (in the sense that maximizes expected discounted dividends until the switching random exponential time $\overline{\tau}$, after which the terminal value is g) restricted to the family of control actions defined above. For that purpose, let us introduce the operators related to the control actions in $\mathcal{E}$. Consider the operators $\widehat{\mathcal{T}}_0$, $\mathcal{T}_1$, and $\mathcal{T}_F$ in the set of functions

$$\mathcal{W}^{\delta,\Delta} = \{w : \mathcal{G}_\delta \times \mathcal{H}_\Delta \to [0,\infty)\}$$

defined as follows:

$$\widehat{\mathcal{T}}_0(w)(x_n^\delta, \lambda) := \mathbb{P}(\delta < T \wedge \tau^U \wedge \overline{\tau}) e^{-q\delta} w(x_{n+1}^\delta, \widehat{\lambda}_\delta) + \widehat{\mathcal{I}}(w)(x_n^\delta, \lambda_m^\Delta), \tag{17}$$

where

$$\begin{aligned}\widehat{\mathcal{I}}(w)(x_n^\delta, \lambda_m^\Delta) &= \mathbb{E}(I_{\delta\wedge T\wedge\tau^U\wedge\overline{\tau}=\tau^U} I_{\{x+p\tau^U-U\geq 0\}} e^{-q\tau^U} w(\rho^\delta(x_n^\delta + p\tau^U - U), \widehat{\lambda}_{\tau^U}))\\ &+ \mathbb{E}(I_{\delta\wedge T\wedge\tau^U\wedge\overline{\tau}=T}\ e^{-qT} w(x_n^\delta, \widehat{\lambda}_T)) + \mathbb{E}(I_{\delta\wedge T\wedge\tau^U\wedge\overline{\tau}=T} e^{-qT} pT)\\ &+ \mathbb{E}(I_{\delta\wedge T\wedge\tau^U\wedge\overline{\tau}=\tau^U} I_{\{x+p\tau^U-U\geq 0\}} e^{-q\tau^U}(x_n^\delta + p\tau^U - U - \rho^\delta(x_n^\delta + p\tau^U - U)))\\ &+ \mathbb{E}(I_{\delta\wedge T\wedge\tau^U\wedge\overline{\tau}=\overline{\tau}}\ e^{-q\overline{\tau}} g(x_n^\delta + p\overline{\tau}, \widehat{\lambda}_{\overline{\tau}}))\end{aligned}$$

and

$$\mathcal{T}_1(w)(x_n^\delta, \lambda) = w(x_{n-1}^\delta, \lambda) + \delta p, \qquad \mathcal{T}_F(w)(x_n^\delta, \lambda) = x_n^\delta. \tag{18}$$

Here, τ^U and U are the arrival time and the size of the next claim, and T and Y are the arrival time and the size of the next intensity upward jump. We also introduce the operator $\widehat{\mathcal{T}}$ in $\mathcal{W}^\delta$ as

$$\widehat{\mathcal{T}} = \max\{\widehat{\mathcal{T}}_0, \mathcal{T}_1, \mathcal{T}_F\}. \tag{19}$$

The operator $\widehat{\mathcal{T}}_0$ is associated with the local action $\mathbf{E}_0$, the operator $\mathcal{T}_1$ with the local action $\mathbf{E}_1$, and the operator $\mathcal{T}_F$ with the local action $\mathbf{E}_F$. The control action $\mathbf{E}_F$ is never optimal but facilitates the numerical scheme. It can be proved, much like in Albrecher et al. (2024), that there exists a unique fixed point of $\widehat{\mathcal{T}}$ which approximates uniformly the optimal value V taking δ and Δ small enough and m_1 large enough. Let us denote this unique fixed point of $\widehat{\mathcal{T}}$ as $\widehat{V}^{\delta,\Delta}$, so $\widehat{\mathcal{T}}(\widehat{V}^{\delta,\Delta}) = \widehat{V}^{\delta,\Delta}$. We can then state the following result, the proofs of which are totally analogous to those of Proposition 7.2 and Theorem 7.4 in Albrecher et al. (2024), so we omit them here.

Proposition 7. *For any $\varepsilon > 0$, there exist δ and Δ small enough so that $0 \leq V - \widehat{V}^{\delta,\Delta} \leq \varepsilon$ in $[0,\infty) \times [\underline{\lambda}, \infty)$.*

In order to obtain an approximation of $\widehat{V}^{\delta,\Delta}$, we proceed inductively in the following way:

1. We define as the initial function $\widehat{W}_1(x_n^\delta, \lambda_m^\Delta) = x_n^\delta$. Note that $\widehat{W}_1$ corresponds to applying the local action $\mathbf{E}_F$ in all the points of the grid.

2. We calculate $\widehat{W}_{l+1} = \widehat{\mathcal{T}}(\widehat{W}_l)$. The action regions $\widehat{\mathcal{A}}_{l+1}$ (where dividends are paid) and no-action regions $\widehat{\mathcal{N}\mathcal{A}}_{l+1}$(where no dividends are paid) associated to $\widehat{W}_{l+1}$ are obtained as

$$\widehat{\mathcal{A}}_{l+1} = \{(x_n^\delta, \lambda_m^\Delta) \in \overline{\mathcal{G}}_\delta \times \overline{\mathcal{H}}_\Delta \,:\, \widehat{\mathcal{T}}(\widehat{W}_l)(x_n^\delta, \lambda_m^\Delta) = \widehat{\mathcal{T}}_1(\widehat{W}_l)(x_n^\delta, \lambda_m^\Delta)\},$$

$$\widehat{\mathcal{N}\mathcal{A}}_{l+1} = \{(x_n^\delta, \lambda_m^\Delta) \in \overline{\mathcal{G}}_\delta \times \overline{\mathcal{H}}_\Delta - \widehat{\mathcal{A}}_{l+1} \,:\, \widehat{\mathcal{T}}(\widehat{W}_l)(x_n^\delta, \lambda_m^\Delta) = \widehat{\mathcal{T}}_0(\widehat{W}_l)(x_n^\delta, \lambda_m^\Delta)\}.$$

Note that $\mathbf{E}_F$ is never optimal; we have just used $\mathbf{E}_F$ to construct the initial function $\widehat{W}_1$.

The proof that $\lim_{l\to\infty} \widehat{W}_l = \widehat{V}^{\delta,\Delta}$ is as in Section 8 of Albrecher et al. (2024).

6 A compound shot-noise risk model with a tipping point

The auxiliary risk model studied in the previous sections can now be used to study a full model with a tipping point ω, where the surplus process of an insurance company evolves according to a compound Cox shot-noise risk model as defined in Section 2 with a certain set of parameters and distributional assumptions up to a random time horizon ω (not necessarily exponentially distributed). After time ω, the compound Cox process switches to another compound Cox process. Concretely, assume that we have a random tipping point $\omega \sim$ Erlang(k, ξ) for a given k and ξ. Then

$$\omega = \omega^k + \omega^{k-1} + \dots . + \omega^1,$$

where ω^i are i.i.d. exponential random variables with parameter ξ.

A main motivation for the choice of an Erlang distribution for ω is tractability, as it exploits the memoryless property iteratively, and yet allows us to deviate from an exponential shape of the respective density of the waiting time to much more peaked scenarios, as k grows (and ξ is chosen to, for instance, match a certain expected value). An even more general choice would be the class of matrix-exponential distributions, which is dense in the class of all distributions on the positive half-line (see e.g. Bladt and Nielsen (2017)). However, for the purposes of the present article, there is no added value in going to that considerably higher level of complexity in the calculations.

In terms of concrete parameter estimation, if one had sufficiently many historical instances for tipping point occurrences in the past, one could then develop statistical procedures to estimate k and ξ for $\omega \sim$ Erlang(k, ξ) with respect to some optimization criteria. However, historical data for such events are scarce, and for the present purpose we would suggest estimating the expected waiting time k/ξ (for instance, according to some concrete climate projections, cf. Lenton et al. (2019)) and the degree of uncertainty around that value being represented by the value of k. Furthermore, we assume here that all ω^i are observable. That may be a realistic assumption in situations where the approaching of the tipping point is a gradual process with some climate indicators reaching intermediate thresholds, and the arrival of the tipping point ω itself may also be considered as a major observable instance. In case ω^i is not observable, one faces a filtering problem for the state i, which is an additional complication, but may still lead to solvable cases see e.g. Leobacher et al. (2014) for another filtering context.

We will now illustrate how to use the auxiliary problem (iteratively and in a backward direction) to solve the optimal dividend problem for this case.

We call State m, where $1 \le m \le k$, the state in which the time until the tipping point will happen is given by $\omega^m + \omega^{m-1} + \dots . + \omega^1$. Let us call State 0 the final compound Cox process, after the tipping point happened. Hence, starting at State m, at a random exponential time ω^m, the process switches to State $m-1$. To solve the problem of maximizing dividends, we therefore need to apply an iterative backward procedure: from the final State 0 to State 1 to State 2, and so on.

Let us first consider State 0 (that is, after the tipping point). Given an initial surplus x and initial intensity λ, the expected discounted dividends until ruin for an admissible strategy $L = (L_t)_{t\ge 0}$ are given by

$$J^0(L) = \mathbb{E}\left[\int_0^\tau e^{-qs} dL_s\right],$$

where $q > 0$ is the discount rate and τ the ruin time. The optimal value function is given by

$$V^0(x, \lambda) = \sup_L J^0(L).$$

This problem is exactly the one considered in Albrecher et al. (2024), and it can be solved with the same numerical procedure described in the previous section, but putting $\xi = 0$. Consider now State $m \geq 1$. Let us define the optimal value function V^m in State m, assuming that we have already obtained V^{m-1}. This problem can be solved using the auxiliary problem defined in (7) and (8) with $V := V^m$ and $g := V^{m-1}$. Before and after the climate tipping point, the compound Cox process can be completely different; the unique restriction is that the baseline intensities fulfil $\underline{\lambda}_m \geq \underline{\lambda}_{m-1}$.

The same technique can in fact be used for more than one climate tipping point. Consider, for instance, two consecutive tipping points: the first arriving at an Erlang(k_1, ξ_1) time, and the second at an Erlang(k_2, ξ_2) time later. One can then obtain numerically the value functions $V^0, V^1, \ldots, V^{k_2}$ for the last tipping point as described before, then define the auxiliary problem with $g := V^{k_2}$, and start again solving the problem iteratively and backwardly for the tipping point Erlang(k_1, ξ_1) as before. In that case, one would hence need to solve $k_1 + k_2 + 1$ numerical auxiliary problems. The shortcoming of the latter procedure is that, despite the fact that in each step the numerical approximation of the value function converges to the optimal value function, numerical errors may accumulate as the number of steps increases.

7 Numerical results

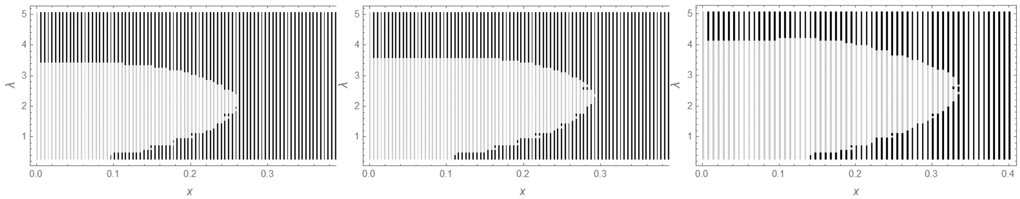
In this section, we show some numerical results. Because of the uniform convergence result, when δ and Δ are sufficiently small, the optimal value function can be obtained with an arbitrarily small approximation error. In our numerical scheme, we ensure that δ and Δ are sufficiently small by comparing the value functions for (δ, Δ) with those of $(\delta/2, \Delta/2)$. We consider examples with varying claim size distributions and upward jump distributions in the shot-noise intensity process.

We will first consider an example with exponential claim sizes, where at an Erlang(2)-distributed tipping point the intensity of catastrophe arrivals increases. It turns out that in this situation, a barrier strategy is optimal where the barrier level depends on both the current surplus level x and the intensity level λ. In Section 7.2, we consider an example with Erlang(2) claim sizes for a situation where, under constant intensity, a two-band strategy is known to be optimal, and we show that for certain λ-values, such a two-band remains optimal, whereas for small and large values of λ, it collapses to a barrier strategy, where the barrier level again depends on x and λ. In Section 7.3, we consider a situation where at the tipping point not only the catastrophe arrival dynamics but also the claim size distribution changes. Finally, in Section 7.4 we consider a variant of the example in Section 7.3 with truncated claims due to a policy limit.

7.1 Example 1: Exponential claims and increasing number of catastrophies

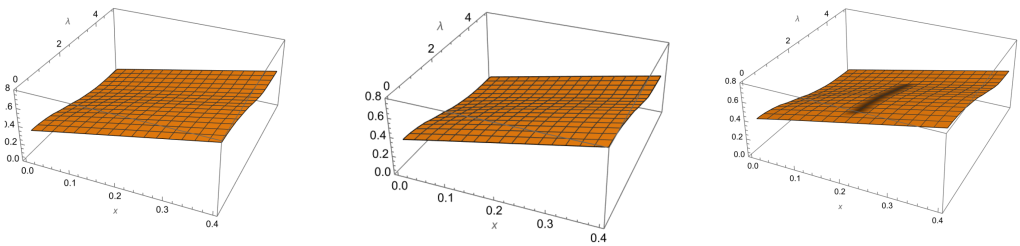
Consider a climate tipping point at an Erlang$(2, 1/3)$ time (so the expected time until its arrival is 6 years) with the frequency of catastrophes increasing after the tipping point. The base intensity equals $\underline{\lambda} = 1/4$, and we assume an exponential intensity jump distribution with cdf $F_Y(x) = 1 - e^{-x/2}$, $d = 7/10$ (so that the additional claim arrival intensity due to a catastrophe is halved after 1 year), as well as an exponential claim size distribution with cdf $F_U(x) = 1 - e^{-10x}$. The discount rate is chosen to be $q = 1/5$, and the insurance safety loading applied in the policies equals $\eta = 1/5$. The concrete values chosen are of realistic magnitudes, and they are also chosen in such a way that the action and non-action regions in the plots are nicely visible. Before the tipping point, we consider the intensity of jump arrivals to be $\beta = 1/3$ (so we expect a catastrophe to happen every 3 years) and after the tipping point we assume $\beta = 1/2$ (i.e., afterward, we expect a catastrophe to happen every 2 years). We then obtain from (9) that the premium rate equals $p = 101/700$ before the tipping point and $p = 141/700$ after the tipping point. Note that we have designed the example in such a way that the setting after the tipping point is exactly that of Example 1 of Albrecher et al. (2024), Section 8. For the grid parameters, it turns out that to consider a finite grid of 60×60 points in the compact set $(x, \lambda) \in [0, 0.4] \times [1/4, 5]$ is sufficient here. All the points outside this compact set are defined to be in the action region.

Figure 1 depicts the action (dark) and no-action (light) regions of the optimal dividend strategy in the different states as a function of current surplus level x and current intensity value λ. In Figure 1a, the regions are plotted for State 2, which is the situation in the beginning ($t = 0$). Figure 1b then shows the regions in State 1, when the first of the two exponential components of the Erlang(2) waiting time has already materialized. One sees that if (for some reason) State 1 is observable (i.e., one knows it to be already close to the tipping point, due to environmental variables having deteriorated considerably, temperature having risen further, etc.), then the dividend strategy should already be adapted, namely the dividend barrier (as a function of current intensity λ) should be raised. Finally, Figure 1c depicts the optimal action and no-action regions after the tipping point



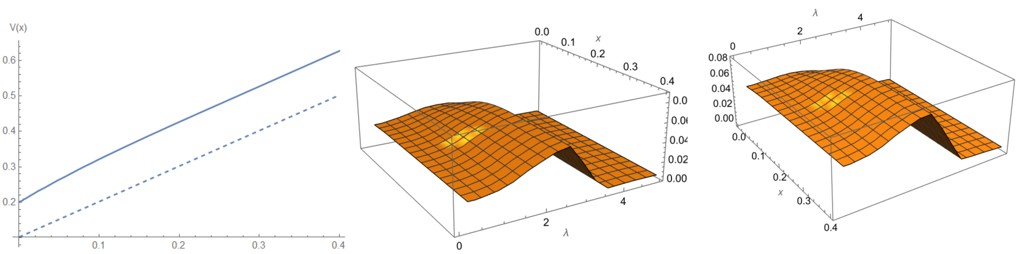
(a) State 2 **(b)** State 1 **(c)** State 0

Figure 1: Action regions (dark) and no-action regions (light) of the optimal strategy as a function of λ and x for each of the states in Example 1.



(a) $V^2(x,\lambda)$ **(b)** $V^1(x,\lambda)$ **(c)** $V^0(x,\lambda)$

Figure 2: Value functions $V^i(x,\lambda)$ as a function of x and λ for each State i in Example 1.



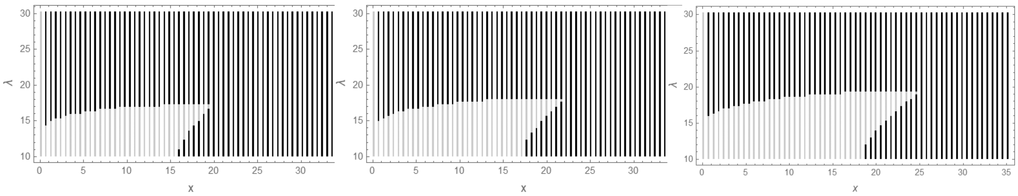
(a) $V^2(x,\lambda_{\text{av}})$ vs. $V_{\text{CL}}(x,\lambda_{\text{av}})$ **(b)** $V^2(x,\lambda)-V^{\text{NT}}(x,\lambda)$ **(c)** $V^1(x,\lambda)-V^{\text{NT}}(x,\lambda)$

Figure 3: Comparison of the value function $V^2(x,\lambda_{\text{av}})$ (solid) and $V_{\text{CL}}(x,\lambda_{\text{av}})$ (dashed) as a function of x (left), $V^2(x,\lambda)-V^{\text{NT}}(x,\lambda)$ (middle), and $V^1(x,\lambda)-V^{\text{NT}}(x,\lambda)$ (right) in Example 1.

(State 0), which leads to a strategy where lowers dividends are paid; this plot is identical to the one in Albrecher et al. (2024), Figure 8.1(A).

The optimal value functions V^2, V^1, and V^0 in State 2, State 1, and State 0, respectively, are depicted in Figure 2 as a function of initial surplus x and initial claim intensity level λ. Figure 2c again corresponds to Albrecher et al. (2024), Figure 8.1(B).

Note that the premium rate is kept constant before the tipping point despite the changing value of momentary claim intensity λ (so the insurance company is assuming and diversifying that risk over time), but it is raised at the moment of the tipping point to account for the average increase of the intensity at that point. Therefore, when we want to compare the attractiveness of a natural catastrophe element in the portfolio to the one of a "regular" compound Poisson portfolio without catastrophe covers, a fair comparison seems to be the one where in the compound Poisson model we also raise the premium at the tipping point in the same fashion. In Figure 3a, $V^2(x,\lambda_{\text{av}})$ is compared with the corresponding value function $V_{\text{CL}}(x,\lambda_{\text{av}})$ of the classical risk model analogue with the same constant intensity λ_{av} as the long-term average of the shot-noise process before and (with an increased value) also after the tipping point. Therefore, the two receive the same premium amounts, both before and after the tipping point; for better visual comparison, we start the shot-noise intensity in that same value $\lambda_0=\lambda_{\text{av}}$. It was shown by Albrecher et al. (2024) that the time variability of the claim intensity in the shot-noise scenario, given that its value can be observed at every point in time and that the dynamics are known, is in fact preferable for the shareholders in terms of the value function when compared to the situation with a constant claim intensity (the safety loading for the policyholders staying at the same



(a) State 2 **(b)** State 1 **(c)** State 0

Figure 4: Action regions (dark) and no-action regions (light) of the optimal strategy as a function of λ and x for each of the states in Example 2.

level). An interpretation is that this advantage comes from the additional degree of freedom of choosing the appropriate dividend payment strategy as a function of not only current surplus x but also current intensity λ. Figure 3a shows that this stays true also when a tipping point is introduced. Note that the dashed line in Figure 3a is not identical to the one in Albrecher et al. (2024, Figure 8.1(D)) (here the premium rate changes at the tipping point), although the two are very close.

Finally, we also look into the effects of the tipping point itself on a NatCat insurance portfolio. Figure 3b plots the difference between the value function $V^2(x,\lambda)$ and the value function $V^{\mathrm{NT}}(x,\lambda)$ of a process with the same initial dynamics, but without a tipping point coming later, across all relevant values of x and λ. One sees that the difference is always positive. Intuitively, the additional variability of the claim intensity after the tipping point (due to the increased value of β) is an advantage, as was the shot-noise variation an advantage over a homogeneous Poisson portfolio. That is, from that point of view, the presence of a tipping point is an upward potential for shareholders, as long as the increased claim intensity is charged to the policyholders according to an expected value principle with the same risk loading. The difference is most pronounced for initial claim intensity values slightly above the long-term average λ_{av} (which for this example is 1.20 before and 1.67 after the tipping point). Figure 3c plots the difference $V^1(x,\lambda) - V^{\mathrm{NT}}(x,\lambda)$, which leads to larger (also always positive) values. The reason is that the value function V^1 is already larger, given that the transition to State 0 with its upward potential is already closer. One can also interpret Figures 1b, 2b, and 3c as the situation of Example 1 with an exp(1/3) instead of an Erlang (2,1/3) time horizon until the tipping point.

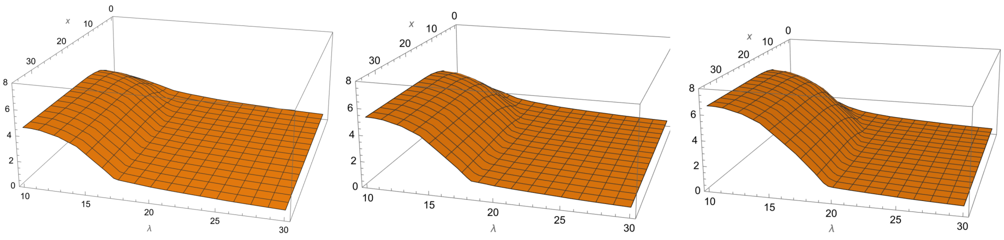
7.2 Example 2: Erlang claims and increasing number of catastrophes

Consider a climate tipping point Erlang$(2,1/3)$ with the frequency of catastrophes increasing from $\beta = 1/6$ before the tipping point to $\beta = 1/5$ afterwards. Here we consider an Erlang$(2,1)$ claim size distribution (i.e., $F_U(x) = 1-(1+x)e^{-x}$, $x>0$) and we switch to parameters for which we know from Azcue and Muler (2014) that a two-band strategy maximizes the expected discounted dividend payments in the absence of shot-noise jumps in the Poisson intensity. Concretely, we consider $\underline{\lambda} = 10$, a catastrophe shot-noise component with exponential intensity jumps with cdf $F_Y(x) = 1-e^{-x/2}$, $d = 1/5$, a discount rate $q = 1/10$, and $\eta = 7/100$. We obtain, from (9), that the premium rate equals $p = 749/30$ before the tipping point and $p = 642/25$ after the tipping point. So, this example is the setting of Example 2 of Albrecher et al. (2024, Section 8) after the tipping point, but we have a reduced shot-noise jump intensity (and correspondingly reduced premium rate) before the tipping point.

For the grid parameters, it suffices to consider a finite grid of 60×60 points in the compact set $(x,\lambda) \in [0,35] \times [10,30]$. Again, all points outside this compact set are defined to be in the action region.

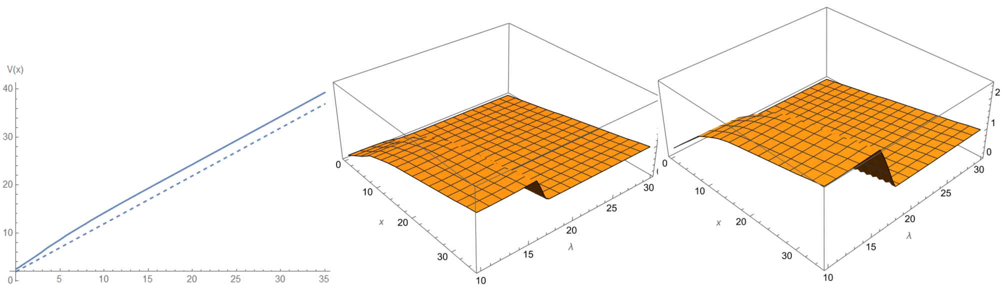
Figure 4 depicts the optimal action (dark) and no-action (light) regions in the different states with respect to x and λ. Figure 4c is identical to the respective regions in the plot of Albrecher et al. (2024), Figure 8.1(A), and one observes that also before the tipping point a similar pattern holds, with only mild adaptations that are only of scaling type. Indeed, in each of the states the optimality of the two-band regime is retained for a small strip of values of λ. For smaller values of λ, the two-band strategy collapses to a barrier strategy, and for large values of λ, the optimal strategy is to pay out all the surplus immediately as dividends. Figure 5 depicts the respective optimal value function for each combination of x and λ in each of the three states (plotted in excess of x). In this case, the differences of the value function between the different states are relatively small.

Figure 6a compares $V^2(x,\lambda_{\mathrm{av}})$ with the corresponding value function $V_{\mathrm{CL}}(x,\lambda_{\mathrm{av}})$ of the respective classical risk model analogue with constant intensity λ_{av} and the same premium income. The additional variability due to the shot-noise structure before and after the tipping point is again an advantage for the shareholders, although the difference is smaller in this case. In Figure 6b, we see that the presence of the tipping point itself in



(a) $V^2(x,\lambda) - x$ **(b)** $V^1(x,\lambda) - x$ **(c)** $V^0(x,\lambda) - x$

Figure 5: Value functions $V^i(x,\lambda)$ minus x as a function of x and λ for each State i in Example 2.



(a) $V^2(x,\lambda_{\text{av}})$ vs. $V_{\text{CL}}(x,\lambda_{\text{av}})$ **(b)** $V^2(x,\lambda) - V^{\text{NT}}(x,\lambda)$ **(c)** $V^1(x,\lambda) - V^{\text{NT}}(x,\lambda)$

Figure 6: Comparison of the value function $V^2(x,\lambda_{\text{av}})$ (solid) and $V_{\text{CL}}(x,\lambda_{\text{av}})$ (dashed) as a function of x (left), $V^2(x,\lambda) - V^{\text{NT}}(x,\lambda)$ (middle), and $V^1(x,\lambda) - V^{\text{NT}}(x,\lambda)$ (right) in Example 2.

a compound shot-noise Cox process is also advantageous for the shareholders, and Figure 6c depicts the same situation for an exp(1/3) time horizon. As in the previous example, this difference is larger than in Figure 6b.

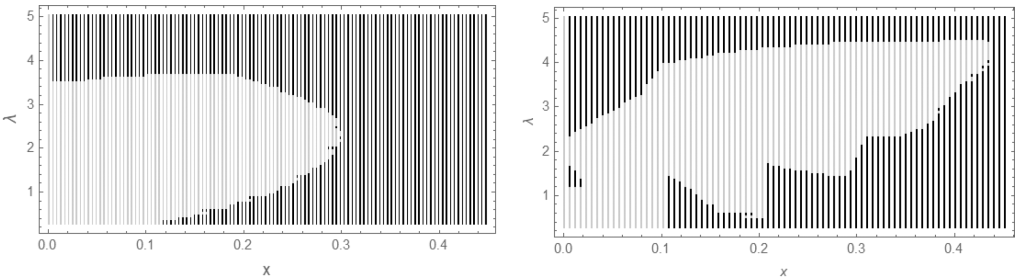
7.3 Example 3: Claim changes and increasing number of catastrophes

Next, let us consider a situation where before the climate tipping time the claim sizes are exponential with the same parameters as in Example 1 and afterwards they are deterministic at size 1/10. The latter corresponds to the situation of Example 3 of Albrecher et al. (2024), Section 8, where no tipping point was considered. As observed in that paper, the deterministic claim size assumption leads to an intricate action/no-action region pattern, and we are interested to see how the anticipation of the latter before the tipping point influences the shape of the optimal action regions. We assume here that the climate tipping point occurs at an exponentially distributed time exp(1/3), to avoid the additional complication of an Erlang(2) time horizon, so here we have only two states. Assume $\underline{\lambda} = 1/4$, an exponential intensity jump distribution with cdf $F_Y(x) = 1 - e^{-x/2}$, $d = 7/10$, a discount rate $q = 1/5$, and the insurance safety loading applied in the policies equal to $\eta = 1/5$. As in Example 1, the tipping point increases the intensity of jump arrivals from $\beta = 1/3$ before the tipping point to $\beta = 1/2$ afterwards. The expected claim size before and after the tipping point remains the same. We then obtain, from (9), that the premium rate equals $p = 101/700$ before the tipping point and $p = 141/700$ afterwards. For the grid parameters, the finite grid of 80×80 points in the compact set $[0, 0.45] \times [1/4, 5]$ is appropriate here (and all other points are defined to be in the action region).

Figure 7 depicts the action (dark) and no-action (light) optimal regions in the different states with respect to x and λ. We see that the anticipation of a deterministic rather than exponential claim size distribution after the tipping point does not change the optimal strategy in the initial State 0 (Figures 7a and 2b coincide). Furthermore, Figure 7b is identical to Albrecher et al. (2024), Figure 8.6(A). Figure 8 depicts the value functions, Figure 9a shows again that the additional variability of the intensity is an advantage for shareholders, and Figure 9b depicts the advantage of the presence of a tipping point over the situation where the dynamics before the tipping point does not change in the future.

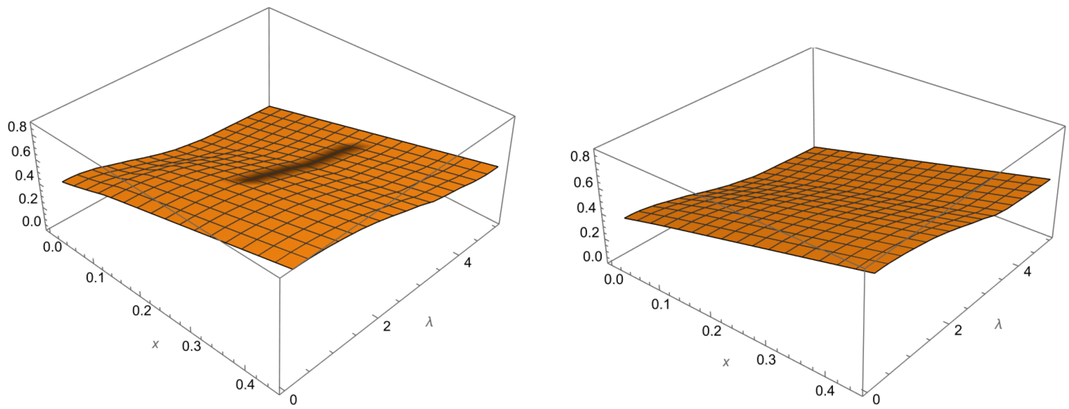
7.4 Example 4: A variant of Example 3 with truncated claims

One may question how the deterministic claims after the tipping point in Example 3 might be justified in practice. One such situation could be where the insurance policies contain, in fact, an upper coverage limit,



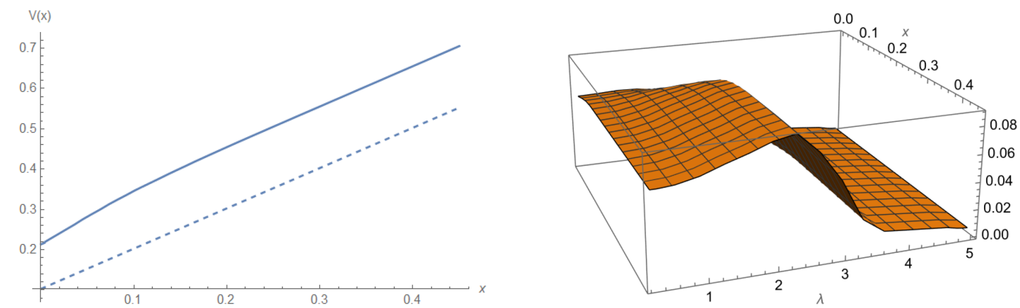
(a) State 1 **(b)** State 0

Figure 7: Action regions (dark) and no-action regions (light) of the optimal strategy as a function of λ and x for each of the states in Example 3.



(a) $V^1(x,\lambda)$ **(b)** $V^0(x,\lambda)$

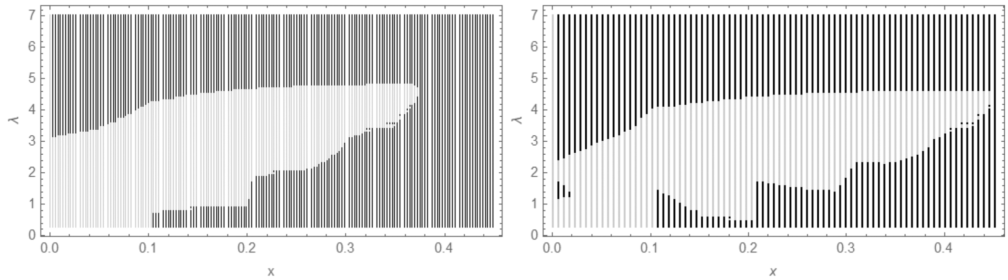
Figure 8: Value functions $V^i(x,\lambda)$ as a function of x and λ for each State i in Example 3.



(a) $V^1(x,\lambda_{\mathrm{av}})$ vs. $V_{\mathrm{CL}}(x,\lambda_{\mathrm{av}})$ **(b)** $V^1(x,\lambda) - V^{\mathrm{NT}}(x,\lambda)$

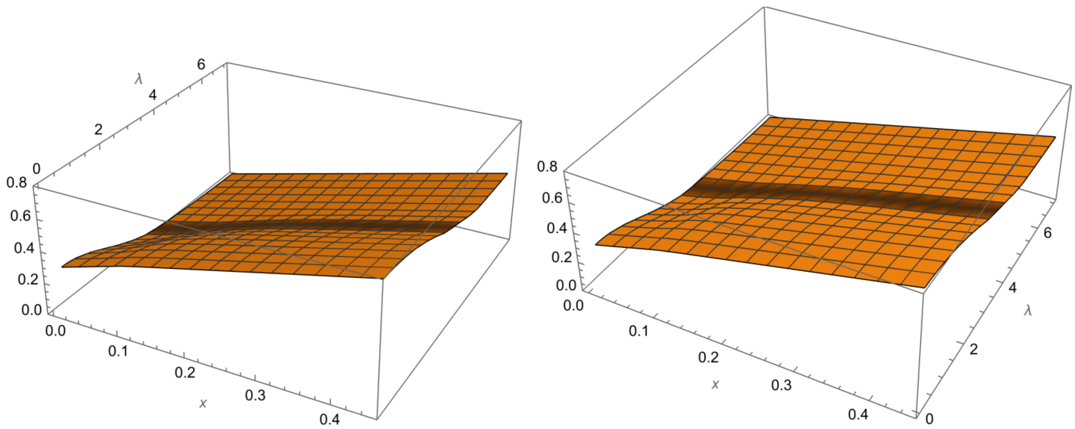
Figure 9: Comparison of the value function $V^1(x,\lambda_{\mathrm{av}})$ (solid) and $V_{\mathrm{CL}}(x,\lambda_{\mathrm{av}})$ (dashed) as a function of x (left) and $V^2(x,\lambda) - V^{\mathrm{NT}}(x,\lambda)$ in Example 3.

and where after the tipping point the claims are all so large that the upper limit is in fact always reached. Note that in that case, the uncertainty for which policyholders buy insurance (and pay a safety loading) is then solely the frequency and not the size of the claims. In the setting of Example 3, one would then have to assume that the claim size distribution before the tipping point is exponential, but truncated at some threshold, which in this example we choose to be 1/10. That is, $F_U(x) = \left(1 - e^{-10x}\right) I_{x<1/10} + I_{x\geq 1/10}$ before and $F_U(x) = I_{x\geq 1/10}$ after the tipping point. All other assumptions and parameter values of Example 3 remain. We then obtain from (9) that the premium rate equals $p = 101(-1+e)/(700\,e) \approx 0.09$ before the tipping point and $p = 141/700 \approx 0.20$ afterwards. For the grid parameters, the finite grid of 80×110 points in the compact set $[0, 0.45] \times [1/4, 7]$ is appropriate here (with all remaining points being in the action region).



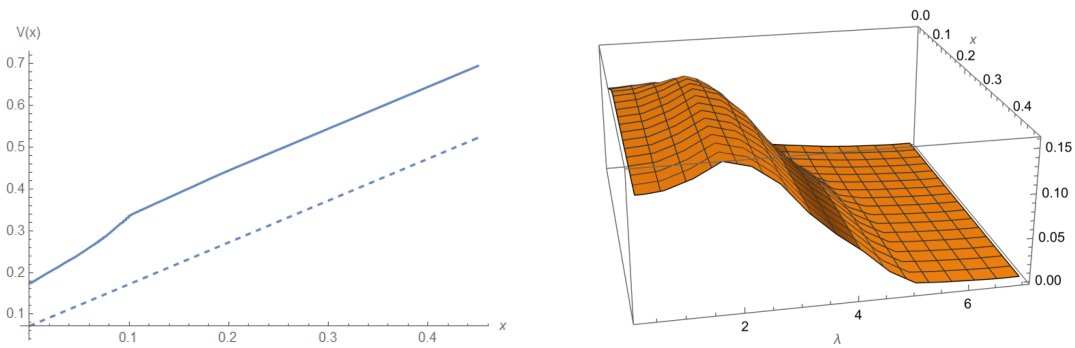
(a) State 1 **(b)** State 0

Figure 10: Action regions (dark) and no-action regions (light) of the optimal strategy as a function of λ and x for each of the states in Example 4.



(a) $V^1(x,\lambda)$ **(b)** $V^0(x,\lambda)$

Figure 11: Value functions $V^i(x,\lambda)$ as a function of x and λ for each State i in Example 4.



(a) $V^1(x,\lambda_{\text{av}})$ vs. $V_{\text{CL}}(x,\lambda_{\text{av}})$ **(b)** $V^1(x,\lambda) - V^{\text{NT}}(x,\lambda)$

Figure 12: Comparison of the value function $V^1(x,\lambda_{\text{av}})$ (solid) and $V_{\text{CL}}(x,\lambda_{\text{av}})$ (dashed) as a function of x (left) and $V^2(x,\lambda) - V^{\text{NT}}(x,\lambda)$ in Example 4.

Figure 10 depicts the action and no-action optimal regions in the two states with respect to x and λ, and Figure 11 gives the respective value functions, which show behaviour similar to those of Example 3. Figure 12a compares the situation of the tipping point again with the one of a compound Poisson setup with constant intensity before and after the tipping point. Note here the kink at $x = 1/10$, where (due to the deterministic claim amount of that size) the value function is continuous but not differentiable, so the notion of a viscosity solution is essential. Figure 12b again assesses the quantitative advantage of having a tipping point for the shareholders for Example 4.

8 Conclusion

In this article, we extended the analysis of Albrecher et al. (2024) to the situation where an insurance company holding a NatCat portfolio faces a deterioration in the future, which we refer to as a tipping point. Under some assumptions, we solved the resulting optimal control problem to identify the optimal dividend payment strategy that maximizes the expected aggregate discounted dividend payments for shareholders over the lifetime of the portfolio, using viscosity solutions of the related HJB equation. We proved uniform convergence of solutions on a finite grid of the parameter space and illustrated the optimal solutions and their properties for a number of concrete cases for which the number of catastrophes increases after the tipping point. It turns out that when the premiums are raised in a fair manner at that point (according to the same premium principle before and after the tipping point), the additional variability of the claim intensity can turn out to be an advantage for the shareholders. That is, under the assumption that all states and parameters are observable and that the model accurately describes reality, increased frequency and severity of catastrophes do not necessarily lead to more unattractive conditions for shareholders, and hence the insurance industry. It will be interesting to study in future research how certain model assumptions can be relaxed, and whether such a conclusion also holds in more general settings.

Data sharing

There is no data analysis in this article.

Acknowledgements

We would like to thank the referees for helpful remarks that improved the presentation of the manuscript. Pablo Azcue contributed to the present manuscript before he unfortunately passed away in April 2024; we dedicate this work to him. Open access publishing facilitated by University of Lausanne, as part of the Wiley - Université de Lausanne agreement via the Consortium of Swiss Academic Libraries.

ORCID

Hansjörg Albrecher: https://orcid.org/0000-0002-5434-9270
Pablo Azcue: https://orcid.org/0000-0003-2624-1841
Nora Muler: https://orcid.org/0000-0002-8618-719X

A Appendix

A1 Proofs of Section 2

For technical reasons, let us define by $v^{\lambda}(x)$ the optimal value function that maximizes the expected cumulative discounted dividend payouts with constant intensity λ and premium rate p (i.e., a Cramér–Lundberg process) with switching value (at an exponential time $\overline{\tau} \sim Exp(\xi)$ independent of the Cramér–Lundberg process) $h^{\lambda}(x) = g(x, \lambda)$. It is straightforward to see the following result (see for instance Wei et al. (2011) for a similar problem):

Lemma 2. *v^{λ} is nondecreasing with $v^{\lambda}(x) - x$ positive, nondecreasing, and bounded. Moreover, $\lim_{\lambda\to\infty} v^{\lambda}(x) = x$ and $v^{\lambda}(x) \leq x + K$ for some $K > 0$.*

Proof of Proposition 1. We skip the proof that V is nonincreasing in λ, uniformly continuous in λ, and that $V(x, \lambda) \leq v^{\underline{\lambda}}(x)$ because the proofs are rather cumbersome and similar to those of Proposition 3.1 in Albrecher et al. (2024). Then, from Lemma 2, V satisfies (6).

Let us show that $V(\cdot, \lambda) : [0, \infty) \to (0, \infty)$ is locally Lipschitz. It is straightforward to show that $V(\cdot, \lambda)$ is non-decreasing because

$$V(x_2, \lambda) - V(x_1, \lambda) \geq x_2 - x_1.$$

Let us prove the other inequality. Take $t_0 = (x_2 - x_1)/p < 1$ and consider $\lambda_{t_0} = \underline{\lambda} + e^{-dt_0}(\lambda - \underline{\lambda}) < \lambda$. Given an initial surplus $x \geq 0$ and $\varepsilon > 0$, consider an admissible strategy $L = (L_t)_{t\geq 0} \in \Pi_{x_2,\lambda_{t_0}}$ such that

$$J(L; x_2, \lambda_{t_0}) \geq V(x_2, \lambda_{t_0}) - \varepsilon$$

for any $x_2 > x_1$. Let us define the strategy $\widetilde{L} \in \Pi_{x_1,\lambda}$ that, starting with surplus x_1, pays no dividends if $X_t^{\widetilde{L}} < x_2$ and follows strategy L after the current reserve reaches (x_2, λ_{t_0}); hence

$$\widetilde{L}_t = \begin{cases} 0 & \text{if} \quad t \le t_0 \ \wedge \tau_1 \wedge T_1 \wedge \overline{\tau} \\ L_{t-t_0} & \text{if} \quad t \ge t_0 \text{ and } \tau_1 \wedge T_1 \wedge \overline{\tau} > t_0 \\ 0 & \text{if} \quad \tau_1 \wedge T_1 < t \text{ and } \tau_1 \wedge T_1 \le t_0 < \overline{\tau}. \end{cases}$$

Therefore,

$$\begin{aligned} V(x_1,\lambda) &\ge J(\widetilde{L}; x_1, \lambda) \\ &\ge \mathbb{P}(\tau_1 \wedge T_1 \wedge \overline{\tau} > t_0) J(L; x_2, \lambda_{t_0}) e^{-qt_0} \\ &\ge \left(V(x_2, \lambda_{t_0}) - \varepsilon\right) e^{-qt_0} \mathbb{P}(\tau_1 \wedge T_1 \wedge \overline{\tau} > t_0), \end{aligned}$$

and

$$V(x_2,\lambda) - V(x_1,\lambda) \le V(x_2,\lambda) - \left(V(x_2,\lambda_{t_0}) - \varepsilon\right)\mathbb{P}(\tau_1 \wedge T_1 \wedge \overline{\tau} > t_0) e^{-qt_0}.$$

So, using that V is nonincreasing on λ and $t_0 < 1$,

$$\begin{aligned} V(x_2,\lambda) - V(x_1,\lambda) &\le V(x_2,\lambda) - V(x_2,\lambda_{t_0})\mathbb{P}(\tau_1 \wedge T_1 \wedge \overline{\tau} > t_0) e^{-qt_0} \\ &\le V(x_2,\lambda) - V(x_2,\lambda)\mathbb{P}(\tau_1 \wedge T_1 \wedge \overline{\tau} > t_0) e^{-qt_0} \\ &= V(x_2,\lambda)\left(1 - e^{-(q+\beta+\lambda+\xi)t_0}\right) \\ &\le V(x_2,\underline{\lambda})(q+\beta+\lambda+\xi)t_0 \\ &= V(x_2,\underline{\lambda})\frac{q+\beta+\lambda+\xi}{p}(x_2 - x_1), \end{aligned}$$

which establishes the result.

Let us prove now that $V(x,\cdot)$ is locally Lipschitz in the open set $(\underline{\lambda}, \infty)$. Since V is nonincreasing on λ, we have

$$0 \le V(x,\lambda_1) - V(x,\lambda_2).$$

Let us prove the other inequality. Take $L \in \Pi_{x,\lambda_1}$ such that $V(x,\lambda_1) \le J(L; x, \lambda_1) + \varepsilon$, and the associated controlled process (X_t^L, λ_t) starting at (x, λ_1). Consider $\delta_0 > 0$ such that $\underline{\lambda} + e^{-d\delta_0}(\lambda_2 - \underline{\lambda}) = \lambda_1$. Hence,

$$\delta_0 = \frac{1}{d}\log\left(\frac{\lambda_2 - \underline{\lambda}}{\lambda_1 - \underline{\lambda}}\right).$$

Since $\log(a) < a$ for $a > 0$, we also have that

$$\delta_0 \le \frac{1}{d(\lambda_1 - \underline{\lambda})}(\lambda_2 - \lambda_1).$$

Define $\widetilde{L} \in \Pi_{x,\lambda_2}$ as

$$\widetilde{L}_t = \begin{cases} pt & \text{if} \quad t \le \delta_0 \ \wedge \tau_1 \wedge T_1 \wedge \overline{\tau} \\ L_{t-\delta_0} & \text{if} \quad t \ge \delta_0 \text{ and } \tau_1 \wedge T_1 \wedge \overline{\tau} > \delta_0 \\ p(\tau_1 \wedge T_1) & \text{if} \quad t > \tau_1 \wedge T_1, t < \overline{\tau} \text{ and } \delta_0 \ge \tau_1 \wedge T_1. \end{cases}$$

Since

$$\int_0^{\delta_0} \left(\underline{\lambda} + e^{-ds}(\lambda_2 - \underline{\lambda})\right) ds \le \lambda_2 \delta_0,$$

we obtain

$$J(\widetilde{L}; x, \lambda_2) \ge J(L; x, \lambda_1) e^{-q\delta_0} \mathbb{P}(\tau_1 \wedge T_1 \wedge \overline{\tau} > \delta_0)$$

$$\begin{aligned}&\geq J(L;x,\lambda_1)e^{-(q+\beta+\lambda+\xi)\delta_0}\\&\geq J(L;x,\lambda_1)(1-(q+\beta+\lambda+\xi)\delta_0).\end{aligned}$$

As a result, taking λ_2 and λ_1 close enough so that $\delta_0 < 1$,

$$\begin{aligned}V(x,\lambda_1)-V(x,\lambda_2) &\leq J(L;x,\lambda_1)-J(\widetilde{L},x,\lambda_2)+\varepsilon\\&\leq J(L;x,\lambda_1)(q+\beta+\lambda+\xi)\delta_0+\varepsilon\\&\leq V(x,\lambda_1)\frac{q+\beta+\lambda+\xi}{d(\lambda_1-\underline{\lambda})}(\lambda_2-\lambda_1)+\varepsilon,\end{aligned}$$

and the result follows. At the lower boundary $\underline{\lambda}$, we only have the uniform continuity result because the Lipschitz constant blows up as $\lambda \searrow \underline{\lambda}$. □

Proof of Proposition 2. For λ large enough, consider $t(\lambda)$, so $\underline{\lambda}+e^{-dt(\lambda)}(\lambda-\underline{\lambda}) = \ln(\lambda)$. Hence,

$$t(\lambda)=\frac{1}{d}\ln\left(\frac{\lambda-\underline{\lambda}}{\ln(\lambda)-\underline{\lambda}}\right).$$

Take a near-optimal strategy $L=(L_t)_{t\geq 0}\in\Pi_{x,\lambda}$ such that $V(x,\lambda)\leq J(L;x,\lambda)+\varepsilon$. By definition, $\lambda_t \geq \lambda_t^c = \underline{\lambda}+e^{-dt}(\lambda-\underline{\lambda})$, and so $\lambda_t \geq \ln(\lambda)$ for $t \leq t(\lambda)$. Then, using Lemma 1, Lemma 2 and Proposition 1,

$$\begin{aligned}V(x,\lambda)-\varepsilon &\leq J(L;x,\lambda)\\&=\mathbb{E}\left(\int_{0^-}^{t(\lambda)\wedge\overline{\tau}\wedge\tau^L} e^{-qs}dL_s + I_{\{\overline{\tau}\wedge t(\lambda)\wedge\tau^L=\overline{\tau}\}}g(X^L_{\overline{\tau}},\lambda_{\overline{\tau}})\right)\\&\quad+e^{-qt(\lambda)}\mathbb{E}\Big(I_{\{\overline{\tau}\wedge\tau\wedge\tau^L=t(\lambda)\}}V(X^L_{t(\lambda)},\lambda_{t(\lambda)})\Big)\\&\leq v^{\ln(\lambda)}(x)+e^{-qt(\lambda)}v^{\underline{\lambda}}(x+pt(\lambda)).\end{aligned}$$

Since $t(\lambda)\to\infty$ as $\lambda\to\infty$ and by Lemma 2, we have that $\lim_{\lambda\to\infty} v^{\ln(\lambda)}(x)=x$, and we obtain for some $K>0$,

$$\lim_{\lambda\to\infty} e^{-qt(\lambda)}v^{\underline{\lambda}}(x+pt(\lambda)) \leq \lim_{\lambda\to\infty} e^{-qt(\lambda)}(x+pt(\lambda)+K)=0.$$

Hence we have the result. □

A2 Proofs of Section 3

Proof of Proposition 3. Let us show that V is a viscosity supersolution. Given initial values $(x,\lambda)\in(0,\infty)\times(\underline{\lambda},\infty)$ and any $l\geq 0$, we consider the admissible strategy $L\in\Pi_{x,\lambda}$, where the company pays dividends with constant rate l until $\tau^L\wedge\overline{\tau}$. Let φ be a test function for supersolution of (12) at (x,λ); the associated controlled surplus is given by $X_t^L = x+(p-l)t$ and the intensity process is $\lambda_t=\underline{\lambda}+e^{-dt}(\lambda-\underline{\lambda})$ for $t<\tau_1\wedge T_1\wedge\overline{\tau}$. Applying Lemma 1 with stopping time $\tau_1\wedge T_1\wedge\overline{\tau}\wedge h$, we have

$$\begin{aligned}0&=V(x,\lambda)-\varphi(x,\lambda)\\&\geq\mathbb{E}\left(\int_0^{\tau_1\wedge T_1\wedge\overline{\tau}\wedge h} e^{-qt}l dt\right)\\&\quad+\mathbb{E}\Big(I_{\{\tau_1<T_1\wedge\overline{\tau}\wedge h\}}e^{-q\tau_1}\varphi(x+(p-l)\tau_1-U_1,\underline{\lambda}+e^{-d\tau_1}(\lambda-\underline{\lambda}))\Big)\\&\quad+\mathbb{E}\Big(I_{\{T_1<\tau_1\wedge\overline{\tau}\wedge h\}}e^{-qT_1}\varphi(x+(p-l)T_1,\underline{\lambda}+e^{-dT_1}(\lambda-\underline{\lambda})+Y_1)\Big)\\&\quad+\mathbb{E}\Big(I_{\{h<\tau_1\wedge\overline{\tau}\wedge T_1\}}e^{-qh}\varphi(x+(p-l)h,\underline{\lambda}+e^{-dh}(\lambda-\underline{\lambda}))\Big)\end{aligned}$$

$$+\mathbb{E}\Big(I_{\{\overline{\tau}<\tau_1\wedge T_1\wedge h\}}e^{-q\overline{\tau}}g(x+(p-l)\overline{\tau},\underline{\lambda}+e^{-d\overline{\tau}}(\lambda-\underline{\lambda}))\Big)-\varphi(x,\lambda).$$

So, dividing by h and taking $h\to 0^+$, and taking into account that τ_1, T_1 and $\overline{\tau}$ are independent exponential random variables, we obtain

$$\mathcal{L}(\varphi)(x,\lambda)+l(1-\varphi_x(x,\lambda))\leq 0.$$

Hence, taking $l\to\infty$,we get

$$\max\{\mathcal{L}(\varphi)(x,\lambda),1-\varphi_x(x,\lambda)\}\leq 0.$$

This proves that V is a viscosity supersolution at (x,λ).

We omit the proof that V is a viscosity subsolution. This result is similar to the one in Proposition 4.1 in Albrecher et al. (2024). □

The next lemma will be used to prove Proposition 4, and the proof is similar to the one of Lemma A.3 in Albrecher et al. (2024) in which $\xi=0$.

Lemma 3. *Let $\overline{u}$ be a non-negative supersolution of (12) satisfying the growth condition (6) and nonincreasing on λ. Then, the sequence of positive convolution functions $\overline{u}^m : (0,\infty)\times(\underline{\lambda},\infty)\to\mathbb{R}$ defined as*

$$\overline{u}^m(x,\lambda)=m^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\overline{u}(x+s,\lambda+t)\phi(ms)\phi(mt)\,ds\,dt,$$

where ϕ is a non-negative continuously differentiable function with support included in $(0,1)$ such that $\int_0^1\phi(x)dx=1$, satisfies

(a) $\overline{u}^m$ is continuously differentiable;

(b) $\overline{u}^m(x,\lambda)\leq K+x$ and $\overline{u}^m$ is nonincreasing on λ;

(c) $1\leq\overline{u}^m_x\leq((q+\lambda+\beta+\xi)/p)\overline{u}(x+1/m,\lambda)$ in $(0,\infty)\times(\underline{\lambda},\infty)$;

(d) $\overline{u}^m\to\overline{u}$ uniformly on compact sets in $(0,\infty)\times(\underline{\lambda},\infty)$ and $\nabla\overline{u}^m$ converges to $\nabla\overline{u}$ a.e. in $(0,\infty)\times(\underline{\lambda},\infty)$;

(e) There exists a sequence c_m with $\lim_{m\to\infty}c_m=0$ such that

$$\sup_{(x,\lambda)\in A_0}\mathcal{L}(\overline{u}^m)(x,\lambda)\leq c_m,\text{where } A_0=[0,x_0]\times[\lambda_0,\lambda_1],$$

where $x_0>0$ and $\lambda_0,\lambda_1\in(\underline{\lambda},\infty)$.

Proof of Proposition 4. Consider $\overline{u}$ a non-negative supersolution of (12) satisfying the growth condition (6). Take any admissible strategy $L\in\Pi_{x,\lambda}$, and (X_t^L,λ_t) as the corresponding controlled risk process starting at (x,λ). Take $\overline{u}^m$ as defined in Lemma 3 in $(0,\infty)\times(\underline{\lambda},\infty)$ and let us extend this function as $\overline{u}^m(x,\lambda)=0$. As in the proof of Proposition 4.2 in Albrecher et al. (2024), but considering the switching time $\overline{\tau}$, we have (using $\overline{u}^m_x\geq 1$),

$$\begin{aligned}&\mathbb{E}\Big(I_{t\wedge\tau^L<\overline{\tau}}\overline{u}^m(X^L_{t\wedge\tau^L},\lambda_{t\wedge\tau^L})e^{-q(t\wedge\tau^L)}+I_{\overline{\tau}<t\wedge\tau^L}e^{-q\overline{\tau}}g(X_{\overline{\tau}},\lambda_{\overline{\tau}})\Big)-\overline{u}^m(x,\lambda)\\&\leq\mathbb{E}\left(\int_0^{\overline{\tau}\wedge t\wedge\tau^L}\mathcal{L}(\overline{u}^m)(X^L_{s-},\lambda_{s-})e^{-qs}ds\right)-\mathbb{E}\left(\int_{0^-}^{\overline{\tau}\wedge t\wedge\tau^L}e^{-qs}dL_s\right).\end{aligned}\tag{A1}$$

By the monotone convergence theorem, and since L_t is a nondecreasing process, we obtain

$$\lim_{t\to\infty}\mathbb{E}\left(\int_{0^-}^{\overline{\tau}\wedge t\wedge\tau^L}e^{-qs}dL_s\right)=\mathbb{E}\left(\int_{0^-}^{\overline{\tau}\wedge\tau^L}e^{-qs}dL_s\right).$$

We get from Lemma 3-(c)

$$-(q+\lambda+\beta+\xi)\overline{u}(x+\frac{1}{m},\lambda)\leq \mathcal{L}(\overline{u}^{m})(x,\lambda)\leq (q+\lambda+\beta+\xi)\overline{u}\left(x+\frac{1}{m},\lambda\right)+\beta\overline{u}(x,\lambda)+\xi g(x,\lambda). \tag{A2}$$

From Lemma 3-(b), the growth condition (6) on $\overline{u}$ and g, and since $X_s^L \leq x+ps$, $\lambda_s > \underline{\lambda}$, we obtain

$$\overline{u}(X_s^L,\lambda_s)\leq K+x+ps,\ g(X_s,\lambda_s)\leq K+x+ps \tag{A3}$$

for some $K>0$. Therefore, by the bounded convergence theorem, we have

$$\begin{aligned}&\lim_{t\to\infty}\mathbb{E}\left(\int_0^{\overline{\tau}\wedge t\wedge \tau^L}\mathcal{L}(\overline{u}^{m})(X_{s-}^L,\lambda_{s-})e^{-qs}ds+I_{\overline{\tau}\wedge t<\tau^L}\left(e^{-q\overline{\tau}}g(X_{\overline{\tau}},\lambda_{\overline{\tau}})\right)\right)\\&=\mathbb{E}\left(\int_0^{\overline{\tau}\wedge \tau^L}\mathcal{L}(\overline{u}^{m})(X_{s-}^L,\lambda_{s-})e^{-qs}ds+I_{\overline{\tau}<\tau^L}\left(e^{-q\overline{\tau}}g(X_{\overline{\tau}},\lambda_{\overline{\tau}})\right)\right).\end{aligned} \tag{A4}$$

From (A1) and (A4), we get

$$\begin{aligned}&\lim_{t\to\infty}\mathbb{E}\left(I_{t\wedge\tau^L<\overline{\tau}}\overline{u}^{m}(X_{t\wedge\tau^L}^L,\lambda_{t\wedge\tau^L})e^{-q(t\wedge\tau^L)}\right)-\overline{u}^{m}(x,\lambda)\\&=-\overline{u}^{m}(x,\lambda)\\&\leq \mathbb{E}\left(\int_0^{\tau^L}\mathcal{L}(\overline{u}^{m})(X_{s-}^L,\lambda_{s-})e^{-qs}ds\right)-\mathbb{E}\left(\int_{0^-}^{\overline{\tau}\wedge\tau^L}e^{-qs}dL_s+I_{\overline{\tau}<\tau^L}e^{-q\overline{\tau}}g(X_{\overline{\tau}},\lambda_{\overline{\tau}})\right)\\&=\mathbb{E}\left(\int_0^{\tau^L}\mathcal{L}(\overline{u}^{m})(X_{s-}^L,\lambda_{s-})e^{-qs}ds\right)-J(L;x,\lambda).\end{aligned} \tag{A5}$$

We now show that

$$\limsup_{m\to\infty}\mathbb{E}\left(\int_0^{\tau^L}\mathcal{L}(\overline{u}^{m})(X_{s-}^L,\lambda_{s-})e^{-qs}ds\right)\leq 0. \tag{A6}$$

Since $\overline{u}^{m}(x,\lambda)\leq K+x$, we get from Lemma 3-(b) and (c),

$$\begin{aligned}\left|\mathcal{L}(\overline{u}^{m})(X_{s-}^L,\lambda_{s-})\right| &\leq (q+\lambda+\beta+\xi)\overline{u}\left(X_{s-}^L+\frac{1}{m},\lambda_{s-}\right)+\beta\overline{u}(X_{s-}^L,\lambda_{s-})\\&\leq (q+\lambda+\beta+\xi)\overline{u}\left(X_{s-}^L+\frac{1}{m},\underline{\lambda}\right)+\beta\overline{u}(X_{s-}^L,\underline{\lambda})\\&\leq (q+\lambda+2\beta+\xi)(K+x+1+ps).\end{aligned}$$

As a consequence, given any $\varepsilon>0$, there exists T large enough such that

$$\int_T^{\infty}\mathcal{L}(\overline{u}^{m})(X_{s-}^L,\lambda_{s-})e^{-qs}ds<\frac{\varepsilon}{4} \tag{A7}$$

for any $m\geq 1$. For $s\leq T$, the controlled process $X_{s-}^L\leq x_0:=x+pT$, $\lambda_s\geq\lambda_T:=\underline{\lambda}+\left(\lambda-\underline{\lambda}\right)e^{-dT}$. Since $\boldsymbol{\lambda}$ is a piecewise decreasing process that has upward jumps at exponential times with jumps of finite mean, it is possible to find a constant l large enough such that

$$\mathbb{P}(\sup_{s\leq T}\lambda_s>l)<\frac{\varepsilon}{4(q+\lambda+2\beta+\xi)\left(x+\frac{1}{m}+ps\right)}.$$

Now, take the compact set $[0,x_0]\times[\lambda_T,l]$ and the set

$$C=\left\{(\lambda_s)_{s\leq T}\ :\ \sup_{s\leq T}\lambda_s\leq M\right\}.$$

From Lemma 3-(e), there exists m_0 large enough such that for any $m \geq m_0$,

$$I_C \int_0^T \mathcal{L}(\overline{u}^m)(X^L_{s^-}, \lambda_{s^-}) e^{-qs} ds \leq c_m \int_0^T e^{-qs} ds \leq \frac{c_m}{q} \leq \frac{\varepsilon}{2},$$

and so taking the expectation, we get (A6). Then, from (A5) and using Lemma 3-(d), we obtain

$$\overline{u}(x,\lambda) = \lim_{m\to\infty} \overline{u}^m(x,\lambda) \geq J(L;x,\lambda). \tag{A8}$$

As a consequence, since V is a viscosity solution of (12), the result follows. □

A3 Proofs of Section 4

Proof of Proposition 5. Let us define for $M \geq \hat{x} > 0$, where $\hat{x}$ is defined in Assumption (A1),

$$V_M(x,\lambda) = \begin{cases} V(x,\lambda) & \text{if} \quad x \leq M, \\ V(M,\lambda) + x - M & \text{if} \quad x > M. \end{cases}$$

Note that $V_M(x,\lambda)$ is a limit of value functions of admissible strategies in $\Pi_{x,\lambda}$. V_M is locally Lipschitz, increasing on x and nonincreasing on λ (see Proposition 1). Also, V_M is a viscosity supersolution of (12) for $x < M$. In addition, $\partial_x V_M(x,\lambda) = 1$ for $x > M$. Let us show that V_M is a viscosity supersolution of (12) for $x > M$ taking M large enough. It is enough to show that $\mathcal{L}(V_M)(x,\lambda) \leq 0$ for $x > M$. We have

$$\begin{aligned}
\mathcal{L}(V_M)(x,\lambda) &\leq p - (q+\lambda)V_M(x,\lambda) + \lambda \int_0^x V_M(x-\alpha,\lambda) dF_U(\alpha) + \xi(g(\hat{x},\lambda) + M - \hat{x} - V(M,\lambda)) \\
&\leq p - (q+\lambda)(V(M,\lambda) + x - M) + \lambda \int_0^x (V(M,\lambda) + x - \alpha - M) dF_U(\alpha) \\
&\quad + \xi(g(\hat{x},\lambda) + M - \hat{x} - V(M,\lambda)) \\
&\leq p - qV(M,\lambda) + \xi(g(\hat{x},\lambda) + M - \hat{x} - V(M,\lambda)) \\
&\leq p - qV(M,\lambda) + \xi(g(\hat{x},\lambda) - \hat{x}).
\end{aligned}$$

And so, since $V(x,\lambda) \geq x$,

$$\mathcal{L}(V_M)(x,\lambda) \leq p - qM + \xi(g(\hat{x},\lambda) - \hat{x}).$$

Hence, taking $M \geq \overline{x}$, we obtain $\mathcal{L}(V_M)(x,\lambda) \leq 0$ for $x > M$.

Let us define

$$B = \{(x,\lambda) \in [0,\infty) \times [\underline{\lambda},\infty) : V \text{ is differentiable}\}.$$

B is a dense set in $[0,\infty) \times [\underline{\lambda},\infty)$, and for any point $(M,\lambda) \in B$ with $M \geq \overline{x}$, we have that $\partial_x V_M(M,\lambda) = 1$. Consequently, by Corollary 1, we obtain that V_M coincides with V. □

Proof of Proposition 6. We have that $V(x,\lambda) = x + a(\lambda)$ for $x \geq \overline{x}$ as a consequence of Proposition 5, taking

$$a(\lambda) := V(\overline{x},\lambda) - \overline{x}.$$

Also, from Proposition 2, $\lim_{\lambda\to\infty} V(\overline{x},\lambda) - \overline{x} = 0$ and $V(\overline{x},\lambda)$ is nonincreasing on λ. This implies $\lim_{\lambda\to\infty} a(\lambda) \searrow 0$ and so

$$\lim_{\lambda\to\infty} \left(\sup_{x\geq\overline{x}} V(x,\lambda) - x \right) = 0.$$

In addition, since for all $x \geq 0$ such that $V_x(x,\lambda)$ exists, $V_x(x,\lambda) \geq 1$, we get that $h(x,\lambda) := V(x,\lambda) - x$ is nondecreasing on x. As a consequence,

$$V(x,\lambda) - x \leq V(\overline{x},\lambda) - \overline{x} = a(\lambda),$$

so that we get the result. □

References

Albrecher, H. and Asmussen, S. (2006). Ruin probabilities and aggregate claims distributions for shot noise Cox processes. *Scandinavian Actuarial Journal*, 2006(2):86–110.

Albrecher, H., Azcue, P., and Muler, N. (2024). Optimal dividend strategies for a catastrophe insurer. *Frontiers of Mathematical Finance*, 3(2):304–344.

Albrecher, H., Beirlant, J., and Teugels, J. (2017). *Reinsurance: Actuarial and Statistical Aspects*, John Wiley and Sons, Chichester.

Albrecher, H. and Thonhauser, S. (2009). Optimality results for dividend problems in insurance. *Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas*, 103:295–320.

Albrecher, H. and Thonhauser, S. (2012). On optimal dividend strategies in insurance with a random time horizon. In Cohen, S. N., Madan, D. B., Siu, T. K., & Yang, H. (Eds.), *Stochastic Processes, Finance and Control: A Festschrift in Honor of Robert J Elliott*, World Scientific, Hackensack, NJ, 157–179.

Asmussen, S., Avram, F., and Usabel, M. (2002). Erlangian approximations for finite-horizon ruin probabilities. *ASTIN Bulletin*, 32(2):267–281.

Avanzi, B. (2009). Strategies for dividend distribution: A review. *North American Actuarial Journal*, 13(2):217–251.

Azcue, P. and Muler, N. (2014). *Stochastic Optimization in Insurance: A Dynamic Programming Approach*, Springer Briefs in Quantitative Finance, Springer, New York.

Azcue, P. and Muler, N. (2021). A multidimensional problem of optimal dividends with irreversible switching: A convergent numerical scheme. *Applied Mathematics and Optimization*, 83:1613–1649.

Bladt, M. and Nielsen, B. F. (2017). *Matrix-Exponential Distributions in Applied Probability*, Vol. 81, Springer, New York.

Carr, P. (1998). Randomization and the American put. *The Review of Financial Studies*, 11(3):597–626.

Dassios, A. and Jang, J. W. (2003). Pricing of catastrophe reinsurance and derivatives using the Cox process with shot noise intensity. *Finance and Stochastics*, 7:73–95.

De Finetti, B. (1957). Su un'impostazione alternativa della teoria collettiva del rischio. In *Transactions of the 15th International Congress of Actuaries*, Vol. 2, 433–443.

Jiang, Z. and Pistorius, M. (2012). Optimal dividend distribution under Markov regime switching. *Finance and Stochastics*, 16:449–476.

Lenton, T. M., Rockström, M. J., Gaffney, O., Rahmstorf, S., Richardson, K., Steffen, W., and Schellnhuber, H. J. (2019). Climate tipping points – Too risky to bet against. *Nature*, 575(7784):592–595.

Leobacher, G., Szölgyenyi, M., and Thonhauser, S. (2014). Bayesian dividend optimization and finite time ruin probabilities. *Stochastic Models*, 30(2):216–249.

Lloyd, E. and Shepherd, T. (2020). Environmental catastrophes, climate change, and attribution. *Annals of the New York Academy of Sciences*, 1469(1):105–124.

Schneider, R. (2023). Extreme events require new forms of financial collaboration to become more resilient. *Environment Systems and Decisions*, 43(4):544–554.

Seneviratne, S., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S. et al. (2021). Weather and climate extreme events in a changing climate. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, 1513–1766.

Szölgyenyi, M. (2015). Dividend maximization in a hidden Markov switching model. *Statistics and Risk Modeling*, 32(3–4):143–158.

Wei, J., Yang, H., and Wang, R. (2010). Optimal reinsurance and dividend strategies under the Markov-modulated insurance risk model. *Stochastic Analysis and Applications*, 28(6):1078–1105.

Wei, J., Yang, H., and Wang, R. (2011). Optimal threshold dividend strategies under the compound Poisson model with regime switching. In Kohatsu-Higa, A., Privault, N., & Sheu, S. J. (Eds.), *Stochastic Analysis with Financial Applications. Progress in Probability*, Vol. 65, Springer, Basel.

Received 11 December 2024 — Accepted 21 July 2025