Catastrophe Bond Pricing Under Renewal Process

Saeid Safarveisi*

Actuarial Research Group, AFI, Faculty of Economics and Business, KU Leuven, Leuven, Belgium

saeid.safarveisi@kuleuven.be

Dixon Domfeh

Department of Computer Science, Georgia Institute of Technology, Atlanta, United States

dnkwantabisa3@gatech.edu

Arpita Chatterjee

Department of Mathematical Sciences, Georgia Southern University, Georgia, United States

achatterjee@georgiasouthern.edu

Version: July 29, 2023

Abstract

This paper introduces a framework for catastrophe bond pricing. The renewal process and Cox-Ingersoll-Ross (CIR) process are considered to model uncertainty sources associated with insurance and financial risks, respectively, which behave independently from one another. We perform a two-step valuation method, in which a class of all equivalent measures corresponding to the insurance risk is selected such that a renewal compound process (used as the loss index in the pay-off function) preserves its structure when changing the probability measure. Applying the Bayesian inference together with the information extracted from historical data and capital market enables us to fully calibrate the pricing model. In addition, the proposed framework allows us to assess the effect of different inter-arrival time distributions on the CAT bond price, which has not been discussed in the literature at the time of writing this paper. Coupled with the aforesaid advantage, the derived pricing formula under some special cases of the renewal process can account for the market prices

*Corresponding author

of claim frequency and claim severity separately.

JEL classification: C11; C15; Q54; G12; G22

1 Introduction

As catastrophe risk stemming from climate change increases (re)insurers are turning to financial markets for capital adequacy.(Re)insurers use insurance-linked securities (ILS) such as catastrophe (CAT) bonds to raise additional capital to increase their capacity to handle the increasing risk of natural disasters. For a financial investor, these bonds are particularly attractive due to their so-called *zero-beta* (or low correlation to the broad financial market) and relatively high yield. A CAT bond is an exotic derivative whose cash-flows are contingent on the occurrence of natural disasters. CAT bonds have become popular as an alternative risk transfer mechanism since their inception in 1994 with an estimated outstanding market capitalization of $37 billion as of 2023[1].

The increase in CAT bonds popularity as a viable alternative investment has drawn the attention of researchers and investors to understand how these instruments are priced. CAT bonds are particularly unique, since their underlying risk cover two distinct sources of risk (i.e., interest rate and catastrophe risk) which makes their valuation challenging. The pricing of CAT bond contracts is of great importance as it determines a fair value at which the insurer and the investor are willing to enter a deal. However, from a technical point of view, the approach of finding the price is not straightforward in light of the existence of both insurance and financial risks. To this end, many researchers have proposed different methods of pricing.

Some authors (see for example, Cox and Pedersen (2000), Lee and Yu (2002), Ma and Ma (2013)) have extended the risk-neutral approach used in the valuation of primitive financial derivatives to price CAT bonds. Traditional risk-neutral valuation principles are not directly applicable to CAT bonds. Since CAT bonds have an actuarial dimension, the use of pure financial valuation principles may underprice the inherent catastrophe risk (Domfeh et al. 2022). The catastrophe risk component of CAT bonds lies outside the financial market which naturally leads to an incomplete market setting for CAT bonds. Under an incomplete market, there is no unique price of a financial instrument. In other words, there is no unique equivalent martingale under risk-neutral valuation. Regardless, some researchers have taken inspiration from other markets such as credit derivatives in the face of incomplete markets. Under these market settings, risk-neutral is achieved by using a pricing kernel implied by the market. This means that the models used in valuation are calibrated to observed market prices. For instance, Beer and Braun (2022) developed a market-consistent pricing model for CAT bonds using observed secondary CAT bond prices and implied catastrophe intensity rate. Domfeh et al. (2022) utilized

[1] https://www.artemis.bm/dashboard/

an entropy risk-neutral measure under a Bayesian framework to derive market-consistent risk premia for CAT bonds in the primary market.

The intractability of the catastrophe risk component in CAT bond pricing is often due to the inability to replicate payoffs or cash flows linked to catastrophes. Earlier researchers have made several strong assumptions about the CAT bond markets, allowing them to price these instruments as they would for traditional financial derivatives. However, these assumptions were quite restrictive and did not capture of the inherent risk in pricing. For instance, Cox and Pedersen (2000), Lee and Yu (2002), Ma and Ma (2013) treat CAT bonds as zero-beta security. Under this assumption, investors should earn a zero-risk premium for holding CAT bonds in their portfolio because catastrophe is uncorrelated with the overall market. In other words, CAT bonds have idiosyncratic risk (Merton, 1976). This assumption has an immediate consequence on how the catastrophe risk component of CAT bonds is treated in valuation. Specifically, the aggregate loss process (i.e., loss intensity and severity) retains its physical distribution under the risk-neutral measure without any transformation or adjustment. In this way, the only risk-neutral measure is the interest rate risk. However, other researchers have attempted to relax this strong assumption by deriving an independent risk-neutral measure for the catastrophe risk component of CAT bonds. For example, Tang and Zhongyi (2019) apply a distortion measure known as the Wang Transform to the catastrophe risk component in their two-step valuation approach. The Wang Transform amplifies the catastrophe risk under the risk-neutral measure (see e.g, Hamada and Sherris (2003), Kijima and Murimachi (2008), Li et al. (2013)).

The main contribution of this paper is to generalize a family of equivalent risk-neutral measures for catastrophe risk and apply them to CAT bond valuation. To this end, we utilize the two-step valuation framework proposed by Pelsser and Stadje (2014) and later applied it to CAT bond pricing by Tang and Zhongyi (2019). Under the two-step valuation, CAT bonds' sources of risk (financial and catastrophe risks) are independent. Therefore, the resultant price is a product risk-neutral measure, $\mathbb{Q}_1$ and $\mathbb{Q}_2$ respectively. We extend and generalize the characterization of equivalent measures inspired by the works of Macheras and Tzaninis (2020) in the context of CAT bond pricing. The authors introduced a class of equivalent measures such that a compound renewal process under the physical measure remains a compound renewal process under its corresponding risk-neutral measure. It is critical to note that the characterization under the corresponding risk-neutral measure has been transformed (adjusted) under the Random-Nikodym derivative to preserve its structure. This is entirely different from the treatment of catastrophe risk under the strong assumption described earlier. The idea of equivalent martingale for catastrophe risk component of CAT bonds is intimately linked to the actuarial consistent principle which can be traced back to Delbaen and Haezendonck (1989).

In the CAT bond literature, many researchers treat the aggregate claim process as a compound Poisson process where the counting process follows an exponentially distributed inter-arrival time. Burnecki et al. (2019) propose a time-inhomogeneous compound Poisson process to formalize a design for pricing contingent convertible catastrophe bonds (CocoCat). Burnecki and Giuricich (2019) in their attempt to address the heavy-tailed nature of catastrophe losses still maintains a non-homogeneous Poisson process with a deterministic intensity function of time. However, due to the un-anticipatory nature of catastrophic events, the assumptions that claims occur as (non)homogeneous Poisson process may be unsatisfactory. To address this issue, Ma et al. (2017) propose a doubly stochastic Poisson process (also known as a Cox process) where the Poisson intensity is stochastic, to model the claim arrival process.

While we see nothing wrong in their approaches, we argue that the aggregate loss process need not necessarily be a compound Poisson process, but rather a flexible compound process depending on the data generation process itself. For example, the inter-arrival process could follow a gamma distribution. The renewal process is a general process which includes a Poisson process. In this paper, we formalize the generalization of the compound renewal process to include other distributions inspired by the recent work of Macheras and Tzaninis (2020). We apply our framework to the earthquake dataset and calibrate the model under several scenarios and test the sensitivity of CAT bond prices where the inter-arrival time follows either Exponential, Gamma, Weibull, or Mixed-exponential distributions. We model the interest rate risk component of catastrophe bonds as another stochastic process that follows Cox-Ingersoll-Ross (CIR) model.

There are other econometric and machine learning-based approaches that have been applied to CAT bonds. While these methodologies are not valuation methods in the strict sense, they nonetheless reveal interesting patterns in the price formation on CAT bonds [2]. Of important notice is the earlier work of Lane (2000) who models the expected excess return (EER) on CAT bonds as a function of the probability of first loss (PFL) and conditional expected loss (CEL) with Cobb-Douglas function to capture the asymmetrical nature of the catastrophe losses and PFL. This two-factor model would serve as the seminal work to characterize the behavior of the CAT bond market. Lane and Oliver (2008) find that CAT bonds could be priced based on existing bond contracts characteristics such as the underlying peril, the expected loss, the wider

[2] Typically, econometric and machine learning approaches are applied to already existing CAT bonds either in the primary or secondary market. Such models examine the historical relationship between several covariates which might help predict future prices.

capital market cycle, and the risk profile of the transaction. Papachristou (2011) utilizes a generalized additive model to examine the factors that can affect the CAT bond premia and concludes that factors such as underwriting cycles, rating class, issuer, territory covered, catastrophe risk modeler, and trigger type are relevant in the primary market. Braun (2016) and Gomez and Carcamo (2014) come to a similar conclusion when they use econometric models to analyze already issued CAT bonds in the primary market. We will also note in passing that more recently Götze et al. (2020) and Makariou et al. (2021) have applied machine learning methods such as random forest on primary market data to predict CAT bond prices.

The remainder of the paper is organized as follows: Section 2 discusses the model assumptions and valuation framework. Section 3 provides details on the dynamics of interest rate risk under both physical and risk-neutral measures. Section 4 elaborates on the distribution of the aggregate loss process under both measures. In Section 5, the CAT bond prices associated with two scenarios and model assumptions thereof are derived. Section 6 presents a numerical illustration that includes parameter estimation, model calibration, and sensitivity analysis. Finally, Section 7 concludes the paper.

2 Model assumption and valuation framework

In this section, we build up our valuation framework for CAT bond pricing based on the two-step valuation approach established by Pelsser and Stadje (2014), which was developed later by Tang and Zhongyi (2019).

The probability space on which all random variables will be defined is described as follows: We model the uncertainty in the market by using a filtered probability space $(\Omega, \mathcal{F}, \underline{\mathcal{F}}, \mathbb{P})$ where the flow of information is modeled by an increasing family of sub sigma-fields of $\mathcal{F}$, $\underline{\mathcal{F}} = (\mathcal{F}_t)_{(t\geq 0)}$, satisfying all usual conditions. We also assume the existence of the probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}, \underline{\mathcal{F}})$. As it is postulated that financial risks and actuarial risks are independent under both measures, a complete configuration of the probability space can then be represented by $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$, $\mathbb{P}(\omega) = \mathbb{P}_1(\omega_1) \times \mathbb{P}_2(\omega_2)$, and $\mathbb{Q}(\omega) = \mathbb{Q}_1(\omega_1) \times \mathbb{Q}_2(\omega_2)$, for $\omega = (\omega_1, \omega_2), \omega_1 \in \Omega_1, \omega_2 \in \Omega_2$, where the first component refers to the financial risk and the second component refers to the insurance risk. A detailed study of the exact form of the above-mentioned filtrations associated with insurance and financial risks under the independence assumption can be found in Safarveisi and Hirbod (2021).

The classical pricing formula for asset pricing mentions that under arbitrage-free assumption, there exist an equivalent martingale measure $\mathbb{Q}$ (in the sense of absolute continuity) for the reference measure $\mathbb{P}$ such that any price process $\{V(t) : 0 \leq t \leq T\}$ discounted at a risk-free rate is a martingale. Hence, the present value of a claim at time t can be written as follows:

$$V^{\mathbb{Q}}(t) = \mathbb{E}^{\mathbb{Q}}[D(t,T)V(T)|\mathcal{F}_t] \tag{2.1}$$

where $D(t,T) = exp\{\int_t^T r_s ds\}$ is called the discount factor in which r_s is a risk-free interest rate. Furthermore, $V(T)$ specifies the claim's value at time T, which is a random variable. It is well-known that a complete market leads to a unique price while an incomplete market results in an infinite number of prices, in which case $V^{\mathbb{Q}}(t)$ is supposed to be in the interval $(\inf_{\mathbb{Q}\in\mathcal{M}} V^{\mathbb{Q}}(t), \sup_{\mathbb{Q}\in\mathcal{M}} V^{\mathbb{Q}}(t))$, where $\mathcal{M}$ is the class of all possible equivalent measures under which the price process becomes a martingale. In the sequel, we assume that the claim's payoff at maturity time T represented by $V(T)$ only consists of insurance risk. As the discount factor $D(t,T)$ is defined as a function of interest rate that itself is a macroeconomics element, it can be seen as a part connected to the financial risk. Therefore, using the independence assumption, we can rewrite the relation (2.1) in this way:

$$V^{\mathbb{Q}}(t) = \mathbb{E}^{\mathbb{Q}_1}[D(t,T)|\mathcal{F}_t^1] \times \mathbb{E}^{\mathbb{Q}_2}[V(T)|\mathcal{F}_t^2] \tag{2.2}$$

where $\mathcal{F}_t^1$ and $\mathcal{F}_t^2$ are natural filtrations associated with financial and insurance risks, respectively.

3 Equivalent measure for financial risk

We suppose that the dynamics of interest rate process r_t under measure $\mathbb{P}_1$ is governed by the Cox-Ingersoll-Ross (CIR) model (Cox et al. 1985), of which stochastic differential equation (SDE)is given by

$$dr_t^{\mathbb{P}_1} = \theta(m - r_t)dt + \sigma\sqrt{r_t}dW_t^{\mathbb{P}_1}, \tag{3.1}$$

where $\mathbb{P}_1$ is defined on measurable space $(\Omega_1, \mathcal{F}_1)$, $W_t^{\mathbb{P}_1}$ is a Brownian motion process, $\theta > 0$ is the mean-reverting force measurement (i.e., the speed of mean-reverting), $m > 0$ is a long-run interest rate mean, and $\sigma > 0$ is the volatility parameter for the interest rate. The Feller condition $2m\theta > \sigma^2$ guarantees that r_t is almost surely strictly positive (Feller 1951). Note here that we may want to use interchangeably, for instance, $\mathbb{P}_1$ as a superscript or with a dash for a process in order to stress that the given process is described under measure $\mathbb{P}_1$.

To price the financial risk contained in relation (2.2), the dynamics of interest rate r_t under a risk-neutral measure $\mathbb{Q}_1$, namely $r_t^{\mathbb{Q}_1}$, should be derived. The mean idea is to employ the well-known Girsanov's Theorem which states that if $W_t^{\mathbb{P}_1}$ is a Brownian motion process, and we shift the process by $\int_0^t \frac{\lambda}{\sigma}\sqrt{r_s^{\mathbb{P}_1}}ds$, then the shifted process is again a Brownian motion process under risk-neutral probability measure $\mathbb{Q}_1$, where $\mathbb{P}_1$ and $\mathbb{Q}_1$ are linked together via a likelihood ratio, also known as Radon-Nikodym derivative formula, defined as below

$$\frac{d\mathbb{Q}_1}{d\mathbb{P}_1}\bigg|_{\mathcal{F}_t^1} = \exp\left\{\int_0^t \frac{\lambda}{\sigma}\sqrt{r_s^{\mathbb{P}_1}}dW_s^{\mathbb{P}_1} - \frac{1}{2}\int_0^t \frac{\lambda^2}{\sigma^2}r_s^{\mathbb{P}_1}ds\right\} \tag{3.2}$$

where λ is a constant which determines the market price of risk. In the CIR model, the market price of risk has the form $\lambda_t = \frac{-\lambda\sqrt{r_t}}{\sigma}$[3], which is a popular choice used as a kernel function when applying Girsanov´Theorem. It is easy to show that the dynamics of r_t under risk-neutral measure $\mathbb{Q}_1$ is given by

$$dr_t^{\mathbb{Q}_1} = \theta^*(m^* - r_t^{\mathbb{Q}_1})dt + \sigma\sqrt{r_t^{\mathbb{Q}_1}}dW_t^{\mathbb{Q}_1} \tag{3.3}$$

with new parameters θ^* and m^*, which are defined by

$$\theta^* = \theta + \lambda, \ m^* = \frac{\theta m}{\theta + \lambda} \tag{3.4}$$

4 Equivalent measure for insurance risk

In order to price the insurance risk contained in the payoff function $V(T)$ presented in (2.2), we first need to identify the underlying risk process that captures the respective insurance risk. For this purpose, we start with typical definitions of a CAT bond's payoff function.

4.1 CAT bond's structure

In this paper, we consider a one-period framework for the CAT bond whose conditional payment takes place only once during the length of the contract, as opposed to a multi-period framework where multiple conditional payments are made periodically up to and including the maturity time T. For more information about the latter type, readers can refer to, for example, Safarveisi

[3]The market price of risk here represents the market´expectation about the rate evolution - a positive value of λ is interpreted as the situation where a significant decrease of rates may happen, while a negative value of which shows an significant increase of the rates.

and Hirbod (2021). We follow a stylized structure (i.e., the former type) provided by Ma and Ma (2013).

A zero-coupon CAT bond delivers a face value Z at maturity time T if the condition $L_T \leq D$ happens, where D stands for a contractually defined threshold level and L_T is a loss index reflecting losses from natural catastrophes over the time interval $[0, T]$, and a proportion $q \in (0,1)$ of the face value is paid to the bondholder if the trigger event $L_T > D$ occurs, that is,

$$P_{CAT}(T) = \begin{cases} Z & \text{if } L_T \leq D \\ qZ & \text{if } L_T > D \end{cases} \tag{4.1}$$

According to the foregoing structure selected for the CAT bond contract[4], the payment is contingent on the behaviour of the index-linked catastrophe loss L_T, which is modeled by a counting process and a sequence of positive independent and identically random variables. The counting process defines the number of claims over time interval $[0, T]$, and random variables denote the losses resulting from natural disasters. Such a CAT bond contract whose trigger mechanism is characterized by the underlying loss index L_T is called a non-indemnity-type CAT bond. In many applications, the loss index L_T is chosen to be an aggregate loss process. This model-based loss index plays a fundamental role when calculating the insurance premium in actuarial science, and one can make different assumptions for modeling purposes. Moreover, from (4.1), it is obvious that L_T is the main information to determine the price of a CAT bond contract. In subsequence sections, we discuss further the assumptions that can be made to model L_T.

4.2 Underlying risk process model

Throughout this subsection, the probability space under consideration is $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$. We denote by $X = \{X_n\}_{n\in\mathbb{N}}$ a sequence of $\mathbb{P}_2$-i.i.d positive real-valued random variables named claim size process, $N = \{N_t\}_{t\in\mathbb{R}_+}$ a counting process or claim number, $T = \{T_n\}_{n\in\mathbb{N}_0}$ a claim arrival time process, and $W = \{W_n\}_{n\in\mathbb{N}}$ a claim interarrival process. In fact, W is said to be inter-arrival times between consecutively arriving events, i.e., $W_n = T_n - T_{n-1}$. The counting process N is called a renewal counting process with parameter δ if the interarrival process W is P_2-i.i.d with distribution $\mathbf{K}^\delta$. In particular, when $\delta > 0$ and $\mathbf{K}^\delta = \mathbf{EXP}(\delta)$, namely Exponential distribution with parameter δ, the counting process N becomes a homogeneous Poisson process[5] with intensity δ where δ can be interpreted as the rate of arrivals. So, in a homogeneous

[4] Also known as index-based binary CAT bond contract

[5] Other well-known examples are negative binomial process and general inverse Gaussian process, of which $\mathbf{K}^\delta$ is distributed as Gamma and general inverse Gaussian, respectively.

Poisson process, W_n's are independent and exponentially distributed with constant parameter δ. In the situation where parameter δ varies with time, i.e., it is a deterministic function of time, the counting process N is a non-homogeneous Poisson process for which W_n's are not independent and exponentially distributed anymore. In case δ itself is a random variable, N turns to a mixed renewal counting process. This process can be seen as a special case of a stochastic renewal counting process where δ is itself a stochastic process (also called Cox process). In this paper, we restrict our work to the renewal counting process with fixed parameter δ.

The underlying risk process $L = \{L_t\}_{t\in\mathbb{R}_+}$ is said to be the aggregate claims process induced by (N, X), which is defined as $L_t = \sum_{n=1}^{N_t} X_n$ for any $t \geq 0$. Accordingly, if N is a renewal counting process with parameter δ, and independent of X, the aggregate claims process is called a $\mathbb{P}_2$-compound renewal process (CRP for short) specified by $\mathbf{K}^\delta$ and $\mathbb{P}_{X_1}$ (henceforth written as $\mathbb{P}_2$-CRP($\mathbf{K}^\delta$, $\mathbb{P}^2_{X_1}$)[6]). An example of the CRP is called the compound Poisson process (CPP for short) in which the counting process N is supposed to be a Poisson process, i.e., L_t is a $\mathbb{P}_2$-CPP(δ, $\mathbb{P}^2_{X_1}$)[7].

4.3 Equivalent measure under renewal risk model

The fact that the underlying risk process L_t follows a compound renewal process induced by (N, X), implies that we need to change the distribution of (N, X) as a result of a change of measures. The new obtained pricing measure is characterized by market prices of insurance risks associated with the claim number and claim size, which reflect the risk averseness of an economic agent in the market. Characterization of such an equivalent measure can be conducted based on the information that L_t provides for us. Macheras and Tzaninis (2020) introduced a class of all equivalent measures such that a compound renewal process under reference measure remains a compound renewal process under its corresponding equivalent measure. More formally, let $\mathcal{F}^2_t = \mathcal{F}^L_t$ be the natural filtration generated by random process L_t, in symbol we write $\mathcal{F}^L_t = \sigma(L_s, s \leq t)$ which means sigma-algebra generated by process L_s. Denote by $\mathbf{\Lambda}^{\rho(\delta)}$ the distribution function of interarrival process W with parameter $\rho(\delta)$ under measure $\mathbb{Q}_2$, where ρ is an arbitrary function. We already know that the distribution function of W under measure $\mathbb{P}_2$ is denoted by $\mathbf{K}^\delta$. It can be shown that the Radon-Nikodym derivative satisfying the property of preserving the structure of a renewal compound process under both measures is

[6]For simplicity in notation, we use $\mathbb{P}^2_{X_1}$ to denote the distribution of X_1 under measure $\mathbb{P}_2$.
[7]Note here that a $\mathbb{P}_2$-CPP(δ, $\mathbb{P}^2_{X_1}$) is in fact a $\mathbb{P}_2$-CRP($\mathbf{K}^\delta = \mathbf{EXP}(\delta)$, $\mathbb{P}^2_{X_1}$)

of the following form

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_t^2} = \Big[\prod_{j=1}^{N_t} h^{-1}(\gamma(X_j)) \times \frac{d\mathbb{Q}^2_{W_1}}{d\mathbb{P}^2_{W_1}}(W_j)\Big] \times \frac{1-\mathbf{\Lambda}^{\rho(\delta)}(t-T_{N_t})}{1-\mathbf{K}^{\delta}(t-T_{N_t})} \tag{4.2}$$

where h and γ are real-valued Borel measurable mapping from $(0,\infty)$ to $\mathbb{R}$ such that $\mathbb{E}^{\mathbb{P}_2}[h^{-1}(\gamma(X_j))] = 1$, $\mathbb{E}^{\mathbb{P}_2}[X_1^l h^{-1}(\gamma(X_j))] < \infty$ (for $l = 1,2$), and $\frac{d\mathbb{Q}^2_{W_1}}{d\mathbb{P}^2_{W_1}}$ is the Radon-Nikodym derivative of distribution W_1 under measure $\mathbb{Q}_2$ with respect to distribution W_1 under measure $\mathbb{P}_2$. In practice, we set $h = \ln$ (natural logarithm) and $\gamma = h(f)$ with f being a Radon-Nikodym derivative of $\mathbb{Q}^2_{X_1}$ with respect to $\mathbb{P}^2_{X_1}$, in symbol we write $f = \frac{d\mathbb{Q}^2_{X_1}}{d\mathbb{P}^2_{X_1}}$ where $\mathbb{Q}^2_{X_1}$ and $\mathbb{P}^2_{X_1}$ are assumed to be equivalent measures. To make the current paper self-contained, we provide a rough proof of (4.2) that is originally inspired by Macheras and Tzaninis (2020), see Appendix C. In the following examples, we explain more on the derivation of the Radon-Nikodym derivative using relation (4.2).

Example 4.3.1 *Let $\mathbb{P}_2$ and $\mathbb{Q}_2$ be probability measures such that $\{L_t\}_{t\in\mathbb{R}_+}$ is a $\mathbb{P}_2$-CPP(δ, $\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CPP($\rho(\delta)$, $\mathbb{Q}^2_{X_1}$). This implies that the counting process N is a Poisson process, and hence the inter-arrival process W, which is independent of N, contains a sequence of independent and exponentially distributed random variables under both measures. Therefore, we can write*

$$\begin{aligned}
\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_t^2} &= \Big[\prod_{j=1}^{N_t} e^{\gamma(X_j)} \times \frac{\rho(\delta)e^{-\rho(\delta)W_j}}{\delta e^{-\delta W_j}}\Big] \times \frac{e^{-\rho(\delta)(t-T_{N_t})}}{e^{\delta(t-T_{N_t})}} \\
&= e^{\sum_{j=1}^{N_t}\gamma(X_j)} \times (\frac{\rho(\delta)}{\delta})^{N_t} \times e^{-(\rho(\delta)-\delta)\sum_{j=1}^{N_t} W_j} \times e^{-t(\rho(\delta)-\delta)} \times e^{T_{N_t}(\rho(\delta)-\delta)} \\
&= e^{\sum_{j=1}^{N_t}\gamma(X_j)} \times (\frac{\rho(\delta)}{\delta})^{N_t} \times e^{-t(\rho(\delta)-\delta)} \qquad (4.3)
\end{aligned}$$

where the last line is due to the fact that $T_{N_t} = \sum_{j=1}^{N_t} W_j$.

Relation (4.3) can be reformulated through the following notation which was introduced by Macheras and Tzaninis (2020): Define a real-valued Borel measurable function $\beta_\delta(x) = \gamma(x) + \alpha_\delta$ such that $\alpha_\delta = \ln\rho(\delta) + \ln \mathbb{E}^{\mathbb{P}_2}[W_1]$. Knowing the fact that W_1 is $\mathbb{P}_2$-$\mathbf{EXP}(\delta)$ and putting $\rho(\delta) = \frac{e^{\alpha_\delta}}{\mathbb{E}^{\mathbb{P}_2}[W_1]}$ which itself leads to $\alpha_\delta = \ln(\frac{\rho(\delta)}{\delta})$ as $\mathbb{E}^{\mathbb{P}_2}[W_1] = 1/\delta$, relation (4.3) turns into

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_t^2} = \exp\left\{\sum_{j=1}^{N_t}\beta(X_j) - \delta t\mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)} - 1]\right\}, \tag{4.4}$$

Relation (4.4) is the main result that was proved by Delbaen and Haezendonck (1989). This is however not surprising since a compound Poisson process is a special case of a compound renewal process, and one can generate subclasses by means of the general framework proposed by Macheras and Tzaninis (2020). We now intend to find the distribution of claim number and claim size inducing aggregate process $\{L_t\}_{t\in\mathbb{R}_+}$ under new measure $\mathbb{Q}_2$. First, $\mathbb{E}^{\mathbb{P}_2}[\exp\{\gamma(X_1)\}] = 1$ together with $\beta_\delta(x) - \alpha_\delta = \gamma(x)$ yield

$$\mathbb{E}^{\mathbb{P}_2}[\exp\{\beta(X_1)\}] = \exp\{\alpha_\delta\} = \frac{\rho(\delta)}{\delta} \tag{4.5}$$

and so we have that

$$\rho(\delta) = \delta\mathbb{E}^{\mathbb{P}_2}[\exp\{\beta(X_1)\}] \tag{4.6}$$

Second, for all $A \in \mathcal{B}(\mathbb{R}_+)$ where $\mathcal{B}(\mathbb{R}_+)$ is defined to be the Borel sets on $\mathbb{R}_+$, we can write

$$Q^2_{X_1}(A) = \mathbb{E}^{\mathbb{P}_2}\left[I_A\frac{d\mathbb{Q}^2_{X_1}}{d\mathbb{P}^2_{X_1}}\right] = \mathbb{E}^{\mathbb{P}_2}\left[I_A e^{\gamma(X_1)}\right] = \mathbb{E}^{\mathbb{P}_2}\left[I_A\frac{e^{\beta(X_1)}}{\mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)}]}\right] = \int_A \frac{e^{\beta(x_1)}}{\mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)}]}d\mathbb{P}^2_{X_1} \tag{4.7}$$

From (4.6) and (4.7), we conclude that process L_t is a $\mathbb{Q}_2$-CPP$\big(\delta\mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)}],\ \frac{e^{\beta(x_1)}}{\mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)}]}\mathbb{P}^2_{X_1}\big)$. As proposed by Muermann (2003), an alternative representation of relations (4.4) and (4.7) can be expressed in this way: Define $\kappa = \mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)}]$ and $\nu(x_1) = \frac{e^{\beta(x_1)}}{\mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)}]}$. Then, we have that

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\bigg|_{\mathcal{F}^2_t} = \exp\left\{\sum_{j=1}^{N_t}\ln(\kappa\nu(X_j)) + \delta t(1-\kappa)\right\}, \tag{4.8}$$

where κ and $v(.)$ can be interpreted as the market prices of claim number risk and claim size risk, respectively. Hence, the characterization of Radon-Nikody derivative (4.8) yields that under measure $\mathbb{P}_2$, L_t is a $\mathbb{P}_2$-CPP(δ, $\mathbb{P}^2_{X_1}$) while under measure $\mathbb{Q}_2$ is a $\mathbb{Q}_2$-CPP$\big(\delta\kappa,\ \nu(x_1)\mathbb{P}^2_{X_1}\big)$.

Example 4.3.2 *Let $\mathbb{P}_2$ and $\mathbb{Q}_2$ be probability measures such that $\{L_t\}_{t\in\mathbb{R}_+}$ is a $\mathbb{P}_2$-CRP($\mathbf{K}^\delta$, $\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CPP($\rho(\delta)$, $\mathbb{Q}^2_{X_1}$) with $\mathbf{K}^\delta = \boldsymbol{Ga}(\delta)$[8], where $\delta = (\eta_1, \eta_2)$ and η_1 is assumed to be a positive integer. Relation (4.2) gives us that*

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\bigg|_{\mathcal{F}^2_t} = e^{\sum_{j=1}^{N_t}\gamma(X_j)} \times \left(\frac{\rho(\delta)\Gamma(\eta_1)}{\eta_2^{\eta_1}}\right)^{N_t} \times \left(\prod_{j=1}^{N_t}\frac{1}{W_j^{\eta_1-1}}\right) \times \frac{e^{-t(\rho(\delta)-\eta_2)}}{\sum_{i=0}^{\eta_1-1}\frac{(\eta_2(t-T_{N_t}))^i}{i!}} \tag{4.9}$$

[8]**Ga**(δ) represents the distribution function of a Gamma distribution with the density given by

$$f(x) = \frac{\eta_2^{\eta_1}}{\Gamma(\eta_1)}x^{\eta_1-1}e^{-\eta_2 x} \quad (x \geq 0)$$

Analogously to the previous example, one can show that

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_t^2} = \exp\left\{\sum_{j=1}^{N_t} ln(\kappa\nu(X_j)) + N_t ln\left(\frac{\eta_2\Gamma(\eta_1)}{\eta_2^{\eta_1}\eta_1}\right) - (\eta_1 - 1)\sum_{j=1}^{N_t} ln(W_j) + \frac{t\eta_2}{\eta_1}(\eta_1 - \kappa) - ln\left(\sum_{i=0}^{\eta_1 - 1}\frac{\big(\eta_2(t - T_{N_t})\big)^i}{i!}\right)\right\} \tag{4.10}$$

where κ and $\nu(.)$ are as before. In this example, L_t is $\mathbb{Q}_2$-CPP$\big(\frac{\eta_2}{\eta_1}\kappa,\, \nu(x_1)\mathbb{P}^2_{X_1}\big)$.

Example 4.3.3 *Let $\mathbb{P}_2$ and $\mathbb{Q}_2$ be probability measures such that $\{L_t\}_{t\in\mathbb{R}_+}$ is a $\mathbb{P}_2$-CRP($\mathbf{K}^\delta$, $\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CPP($\rho(\delta)$, $\mathbb{Q}^2_{X_1}$) with $\mathbf{K}^\delta = $ **WE**(δ)[9], where $\delta = (\eta_3, \eta_4)$. Then, we have that*

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_t^2} = e^{\sum_{j=1}^{N_t}\gamma(X_j)} \times \left(\frac{\eta_4^{\eta_3}\rho(\delta)}{\eta_3}\right)^{N_t} \times \left(\prod_{j=1}^{N_t}\frac{1}{W_j^{\eta_3 - 1}}\right) \times \left(\prod_{j=1}^{N_t} e^{-\rho(\delta)W_j + \frac{W_j^{\eta_3}}{\eta_4^{\eta_3}}}\right) \times \frac{e^{\rho(\delta)(t - T_{N_t})}}{e^{\frac{-(t-T_{N_t})^{\eta_3}}{\eta_4^{\eta_3}}}} \tag{4.11}$$

Similarly, it can be shown that

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_t^2} = \exp\left\{\sum_{j=1}^{N_t} ln(\kappa\nu(X_j)) + N_t ln\left(\frac{\eta_4^{\eta_3}}{\eta_4\Gamma(1 + \frac{1}{\eta_3})\eta_3}\right) - (\eta_3 - 1)\sum_{j=1}^{N_t} ln(W_j) + \frac{1}{\eta_4^{\eta_3}}\sum_{j=1}^{N_t} W_j^{\eta_3} - \frac{t\kappa}{\eta_4\Gamma(1 + \frac{1}{\eta_3})} + \frac{(t - T_{N_t})^{\eta_3}}{\eta_4^{\eta_3}}\right\} \tag{4.12}$$

for which L_t is $\mathbb{Q}_2$-CPP$\Big(\Big[\eta_4\Gamma(1 + \frac{1}{\eta_3})\Big]^{-1}\kappa,\, \nu(x_1)\mathbb{P}^2_{X_1}\Big)$.

Example 4.3.4 *Let $\mathbb{P}_2$ and $\mathbb{Q}_2$ be probability measures such that $\{L_t\}_{t\in\mathbb{R}_+}$ is a $\mathbb{P}_2$-CRP($\mathbf{K}^\delta$, $\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CPP($\rho(\delta)$, $\mathbb{Q}^2_{X_1}$) with $\mathbf{K}^\delta = $ **MIX-EXP**(δ)[10], where $\delta = (\phi, \eta_5, \eta_6)$. Then, we have*

[9] **WE**(δ) represents the distribution function of a Weibull distribution with the density given by

$$f(x) = \frac{\eta_3}{\eta_4^{\eta_3}} x^{\eta_3 - 1} e^{-(\frac{x}{\eta_4})^{\eta_3}} \quad (x \geq 0)$$

[10] **MIX-EXP**(δ) represents the distribution function of a mixture of two exponential distributions with the density given by

$$f(x) = \phi\eta_5 e^{-\eta_5 x} + (1 - \phi)\eta_6 e^{-\eta_6 x} \quad (x \geq 0),$$

with $\phi \in (0, 1)$.

that

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_t^2} = \exp\left\{\sum_{j=1}^{N_t}\gamma(X_j)\right\} \times \left(\prod_{j=1}^{N_t}\frac{\rho(\delta)e^{-\rho(\delta)W_j}}{\phi\eta_5 e^{-\eta_5 W_j}+(1-\phi)\eta_6 e^{-\eta_6 W_j}}\right) \times \frac{e^{\rho(\delta)(t-T_{N_t})}}{\Delta_1} \quad (4.13)$$

where $\Delta_1 = 1-[\phi(1-e^{-\eta_5(t-T_{N_t})})+(1-\phi)(1-e^{-\eta_6(t-T_{N_t})})]$. *Following the same procedure discussed earlier, we have that*

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_t^2} = \exp\left\{\sum_{j=1}^{N_t} ln(\kappa\nu(X_j)) - N_t ln\left(\frac{\phi}{\eta_5}+\frac{(1-\phi)}{\eta_6}\right) - \frac{\kappa t}{\frac{\phi}{\eta_5}+\frac{(1-\phi)}{\eta_6}} - ln(\Delta_1)\right\} \quad (4.14)$$

for which L_t *is* $\mathbb{Q}_2$*-CPP*$\big(\big(\frac{\phi}{\eta_5}+\frac{(1-\phi)}{\eta_6}\big)^{-1}\kappa,\ \nu(x_1)\mathbb{P}^2_{X_1}\big)$.

In the examples mentioned above, we have considered a CPP for L_t under measure $\mathbb{Q}_2$, in which case we observed how, with the help of notations defined by Macheras and Tzaninis (2020), the Radon-Nikodym derivative could be represented in terms of parameters contained in the distributions that were assumed under the real-world measure. In fact, when L_t is a CPP under $\mathbb{Q}_2$, by means of these notations, it is possible to find an alternative representation based on the real-world parameters for the intensity parameter contained in the Poisson process under measure $\mathbb{Q}$, which itself helps to express the Radon-Nikodym solidly based on the real-world parameters. However, finding such a representation appears to be more complicated when we assume a CRP for L_t under measure $\mathbb{Q}_2$. One reason is that the definition of α_δ can be only used in the case of one-dimensional space for parameter $\rho(\delta)$. In the next example, we will show this hurdle.

Example 4.3.5 *Let* $\mathbb{P}_2$ *and* $\mathbb{Q}_2$ *be probability measures such that* $\{L_t\}_{t\in\mathbb{R}_+}$ *is a* $\mathbb{P}_2$*-CRP(*$\mathbf{K}^\delta$*,* $\mathbb{P}^2_{X_1}$*) and* $\mathbb{Q}_2$*-CRP(*$\mathbf{\Lambda}^{\rho(\delta)}$*,* $\mathbb{Q}^2_{X_1}$*) with* $\mathbf{K}^\delta = \boldsymbol{Ga}(\delta)$ *and* $\mathbf{\Lambda}^{\rho(\delta)} = \boldsymbol{Ga}(\rho(\delta))$*, where* $\delta = (\xi_1,\xi_2)$ *and* $\rho(\delta) = (\epsilon_1,\epsilon_2)$*. Applying relation (4.2), we get that*

$$\begin{aligned}\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_t^2} &= \left[\prod_{j=1}^{N_t} e^{\gamma(X_j)} \times \frac{\frac{\epsilon_2^{\epsilon_1}}{\Gamma(\epsilon_1)}W_j^{\epsilon_1-1}e^{-\epsilon_2 W_j}}{\frac{\xi_2^{\xi_1}}{\Gamma(\xi_1)}W_j^{\xi_1-1}e^{-\xi_2 W_j}}\right] \times \frac{\sum_{i=0}^{\epsilon_1-1}\frac{(\epsilon_2(t-T_{N_t}))^i}{i!}e^{-\epsilon_2(t-T_{N_t})}}{\sum_{i=0}^{\xi_1-1}\frac{(\xi_2(t-T_{N_t}))^i}{i!}e^{-\xi_2(t-T_{N_t})}} \\ &= \exp\left\{\sum_{j=1}^{N_t}\gamma(X_j)\right\} \times \left(\frac{\epsilon_2^{\epsilon_1}\Gamma(\xi_1)}{\xi_2^{\xi_1}\Gamma(\epsilon_1)}\right)^{N_t} \times \left[\prod_{j=1}^{N_t}W_j^{\epsilon_1-\xi_1}\right] \times \frac{\sum_{i=0}^{\epsilon_1-1}\frac{(\epsilon_2(t-T_{N_t}))^i}{i!}}{\sum_{i=0}^{\xi_1-1}\frac{(\xi_2(t-T_{N_t}))^i}{i!}}\end{aligned} \quad (4.15)$$

where ξ_1 *and* ϵ_1 *are said to be positive integers.*

In this example, we observe that $\rho(\delta)$ is of two-dimension, which does not allow us to find an alternative representation for the Radon-Nikodym derivative in terms of only real-world parameters. Another interesting case is when we assume a mixture of exponential for the inter-arrival distribution under both measures.

Example 4.3.6 *Let L_t be a $\mathbb{P}_2$-CRP($\mathbf{K}^\delta$, $\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CRP($\mathbf{\Lambda}^{\rho(\delta)}$, $\mathbb{Q}^2_{X_1}$) with $\mathbf{K}^\delta = \textbf{\textit{MIX-EXP}}(\delta)$ and $\mathbf{\Lambda}^{\rho(\delta)} = \textbf{\textit{MIX-EXP}}(\rho(\delta))$, where $\delta = (\phi, \eta_5, \eta_6)$ and $\rho(\delta) = (\phi^*, \eta_7, \eta_8)$.*

5 CAT bond price

5.1 Scenario I

According to (2.2), the price of a zero-coupon CAT bond with maturity time T, pay-off function (4.1), and model assumptions stated in examples 4.3.1, 4.3.2, 4.3.3, and 4.3.4 is given by

$$\begin{aligned} V_1^{\mathbb{Q}}(t) &= \mathbb{E}^{\mathbb{Q}_1}[D(t,T)|\mathcal{F}_t^1] \times \mathbb{E}^{\mathbb{Q}_2}[P_{CAT}(T)|\mathcal{F}_t^2] \\ &= P(t,T,r(t),\theta^*,m^*,\sigma)\Big[Z - (Z - qZ)\mathbb{Q}_2(L_T > D)\Big] \end{aligned} \tag{5.1}$$

where the first term in (5.1) denotes the zero-coupon bond price expressed by

$$P(t,T,r(t),\theta^*,m^*,\sigma) = A(t,T)e^{-B(t,T)r(t)}, \tag{5.2}$$

with $A(t,T)$ and $B(t,T)$ being defined as below:

$$\begin{aligned} A(t,T) &= \left[\frac{2\vartheta e^{(\vartheta+\theta^*)(T-t)/2}}{(\vartheta+\theta^*)(e^{\vartheta(T-t)}-1)+2\vartheta}\right]^{\frac{2m^*\theta^*}{\sigma^2}}, \\ B(t,T) &= \frac{2(e^{\vartheta(T-t)}-1)}{(\vartheta+\theta^*)(e^{\vartheta(T-t)}-1)+2\vartheta}, \\ \vartheta &= \sqrt{(\theta^*)^2+2\sigma^2}. \end{aligned}$$

(see, e.g., Brigo and Mercurio **?**), and that

$$\mathbb{Q}_2(L_T > D) = e^{-\delta\kappa T}\sum_{k=0}^{\infty}\frac{(\delta\kappa T)^k}{k!}\mathbb{Q}_2\left(\sum_{i=1}^{k} X_i > D\right) \tag{5.3}$$

Determining κ and $\nu(x)$ is based on the specification of function $\beta(.)$, which itself characterizes the distribution of claim frequency and claim severity under the risk-neutral measure. Following Delbaen and Haezendonck (1989), different choices are possible:

- $\beta_1(x) = \alpha$, where α is a constant (known as the expected value principle).
- $\beta_2(x) = ln(a + bx)$ with $b > 0$ and $a = 1 - b\mathbb{E}^{\mathbb{P}_2}[X_1] > 0$ (known as the variance principle).
- $\beta_3(x) = cx - ln(\mathbb{E}^{\mathbb{P}_2}[e^{cX_1}])$ with $c > 0$ (known as the Esscher principle).

For example, if we choose $\beta_1(x) = \alpha$, then we get $\kappa = e^{\alpha}$ and $\nu(x) = 1$, which means that the distribution of claim severity remains unchanged. In contrast, the intensity parameter of the claim frequency distribution under the risk-neutral measure is scaled by e^{α}.

5.2 Scenario II

We now give our attention to examples 4.3.5 and 4.3.6 where L_t is assumed to be CRP under both measures. For such a situation, as we discussed earlier, it is not possible to acquire a representation based on κ and $\nu(.)$ and the CAT bond price will also rely on risk-neutral parameters contained in inter-arrival time distribution, which can not be estimated easily from information available in the capital market. More precisely, the probability of aggregate claim process L_t can be written as follows under the assumptions of examples 4.3.5 and 4.3.6, respectively; the proofs can be found in the Appendix:

$$\begin{aligned}\mathbb{Q}_2(L_T > D) &= \sum_{k=0}^{\infty} \mathbb{Q}_2(N_T = k)\mathbb{Q}_2\left(\sum_{i=1}^{k} X_i > D\right) \\ &= \sum_{k=0}^{\infty} \sum_{s=k\epsilon_1}^{k\epsilon_1+\epsilon_1-1} \frac{e^{-\epsilon_2 T}(\epsilon_2 T)^s}{s!} \mathbb{Q}_2\left(\sum_{i=1}^{k} X_i > D\right) \end{aligned} \tag{5.4}$$

and

$$\begin{aligned}\mathbb{Q}_2(L_T > D) = \sum_{n=0}^{\infty} \Bigg\{ &\sum_{s=0}^{n} \binom{n}{s} \pi_1^s \pi_2^{n-s} \text{CGDC}(s, \lambda_1, n-s, \lambda_2; T) \\ &- \sum_{s=0}^{n+1} \binom{n+1}{s} \pi_1^s \pi_2^{(n+1)-s} \text{CGDC}(s, \lambda_1, (n+1)-s, \lambda_2; T) \Bigg\} \mathbb{Q}_2\left(\sum_{i=1}^{k} X_i > D\right)\end{aligned}$$

where the notation CGDC denotes the cumulative distribution function of Gamma Distribution Convolution (GDC), $\pi_1 = \phi^*$, $\pi_2 = 1 - \phi^*$, $\lambda_1 = \eta_7$, and $\lambda_2 = \eta_8$. In this paper, we find the estimation of inter-arrival parameters under the risk-neutral measure by using the Bayesian framework.

5.2.1 Bayesian framework for risk-neutral parameters

The foundation of the Bayesian approach considers the specification of a sampling model from which our data is drawn and a marginal distribution of unknown parameters contained in the target distribution, called a prior distribution. Using Bayes'rule, the information about unknown parameters is updated through the conditional model of observed data ~~w~~ and the prior model of unknown parameters to achieve a new distribution called a posteriori. This transition from a prior distribution to a posterior distribution can be seen as a situation where we move from the physical measure to a risk-neutral measure. Then, the inverse probability of unknown parameters under the physical measure (i.e., posterior distribution) can be employed to find a reasonable estimation for corresponding risk-neutral parameters. The Bayesian estimation is one that minimizes the posteriori expected of a given loss function. The commonly used loss functions are the squared error loss function, absolute value error loss function, and weighted squared error loss function. Here, we consider a loss function based on the Kullback-Leibler divergence (KLD). By its standard definition, KLD defines the relative distance between two absolutely continuous measures $\mathbb{P}$ and $\mathbb{Q}$ in the entropy sense[11]. Let $f(x)$ be a density function for a continuous random variable X, characterized by the parameter Θ. From the Bayesian perspective, a necessary form for a proper loss function under KLD is then given by

$$\mathrm{KL}(\Theta \parallel \hat{\Theta}) = \mathrm{KL}\big(f(x;\Theta) \parallel f(x;\hat{\Theta})\big) = \int_{\mathcal{A}} \log \frac{f(x;\Theta)}{f(x;\hat{\Theta})} f(x;\Theta)dx, \tag{5.5}$$

where we call $\mathrm{KL}(\Theta \parallel \hat{\Theta})$ as the Kullback error loss function (KEL) corresponding to the density function f. The Bayesian estimation is then defined as the estimator Θ^B that minimizes KLD between the actual parameter of interest Θ and its possible estimation $\hat{\Theta}$, that is,

$$\Theta^B = \arg\min_{\hat{\Theta}} \mathbb{E}^{\mathbb{P}}[\mathrm{KL}(\Theta \parallel \hat{\Theta}) \,|\, \mathbf{x}] \tag{5.6}$$

We consider the problem of example 4.3.5 in which under physical measure the inter-arrival time distribution $\mathbf{W}$ is Gamma distribution with shape and rate parameters ξ_1 and ξ_2, while under risk-neutral measure is Gamma distribution with shape and rate parameters ϵ_1 and ϵ_2. Let $\mathbf{W}_m = (W_1, W_2, \cdots, W_m)$ be a complete sample from $\mathbf{W}$. We specify the likelihood and prior of the model as follows:

$$\begin{aligned} \mathbf{W}_m|\xi_1, \xi_2 &\sim \mathrm{Ga}(\xi_1, \xi_2), \quad \xi_1 > 0, \quad \xi_2 > 0, \\ (\xi_1, \xi_2) &\sim \pi(\delta) \end{aligned}$$

[11] In an incomplete market where the class of all possible equivalent measures is not unique, the desired measure is selected to minimize KLD, called the minimal entropy martingale measure.

where

$$
\begin{aligned}
f(\mathbf{w}_m|\xi_1,\xi_2) &= \frac{\xi_2^{m\xi_1}}{\Gamma^m(\xi_1)}\Big(\prod_{i=1}^{m} w_i^{\xi_1-1}\Big)\exp\Big\{-\xi_2\sum_{i=1}^{m} w_i\Big\} &(5.7)\\
\pi(\delta) &\propto \frac{\xi_2}{\Gamma(\xi_1)}\exp\left\{(\xi_1-1)\frac{\psi(\xi_1)}{\Gamma(\xi_1)}-\xi_1\right\} &(5.8)
\end{aligned}
$$

with $\psi(\xi_1)=\frac{\partial}{\partial\xi_1}log\Gamma(\xi_1)=\frac{\Gamma'(\xi_1)}{\Gamma(\xi_1)}$ being the diGamma function. The selected non-informative prior distribution above is derived by the justified maximal data information prior (JMDIP) which maximizes the prior average information in the data density minus the information in the prior density. A simulation study conducted by Moala et al. (2013) has shown that for Gamma distribution JMDIP outperforms in a class of non-informative priors such as Jeffrey's Prior, Reference Prior, and Tibishirani's Prior. Assuming the above likelihood and prior distribution, the joint posterior distribution for parameters ξ_1 and ξ_2 is given by,

$$
\begin{aligned}
f(\xi_1,\xi_2|\mathbf{w}_m) &\propto f(\mathbf{w}_m|\xi_1,\xi_2)\times\pi(\delta)\\
&\propto \frac{\xi_2^{m\xi_1}}{\Gamma^m(\xi_1)}\Big(\prod_{i=1}^{m} w_i^{\xi_1-1}\Big)\exp\Big\{-\xi_2\sum_{i=1}^{m} w_i\Big\}\times\frac{\xi_2}{\Gamma(\xi_1)}\exp\left\{(\xi_1-1)\frac{\psi(\xi_1)}{\Gamma(\xi_1)}-\xi_1\right\}
\end{aligned}
\tag{5.9}
$$

Using relation (5.5), the KLD corresponding to Ga(ξ_1,ξ_2) with density $f(x;\xi_1,\xi_2)$ is given by:

$$
\begin{aligned}
\mathrm{KL}(\Theta\parallel\hat{\Theta}) &= \int_0^\infty \log\frac{f(x;\xi_1,\xi_2)}{f(x;\hat{\xi_1},\hat{\xi_2})}f(x;\xi_1,\xi_2)dx,\\
&= \log\frac{\xi_2^{\xi_1}}{\Gamma(\xi_1)}-\log\frac{\hat{\xi_2}^{\hat{\xi_1}}}{\Gamma(\hat{\xi_1})}+(\hat{\xi_2}-\xi_2)\frac{\xi_1}{\xi_2}-(\hat{\xi_1}-\xi_1)\big(\psi(\xi_1)-\log\xi_2\big)
\end{aligned}
\tag{5.10}
$$

According to relation (5.6), the Bayes estimate of ξ_1 and ξ_2 can be found by differentiating the following equation with respect to $\hat{\xi_1}$ and $\hat{\xi_2}$ and setting equal to zero,

$$
\int_0^\infty\int_0^\infty\left[\log\frac{\xi_2^{\xi_1}}{\Gamma(\xi_1)}-\log\frac{\hat{\xi_2}^{\hat{\xi_1}}}{\Gamma(\hat{\xi_1})}+(\hat{\xi_2}-\xi_2)\frac{\xi_1}{\xi_2}-(\hat{\xi_1}-\xi_1)\big(\psi(\xi_1)-\log\xi_2\big)\right]f(\xi_1,\xi_2|\mathbf{w}_m)d\xi_1d\xi_2,
\tag{5.11}
$$

which results in

$$\hat{\xi_2} = \frac{\hat{\xi_1}}{\mathbb{E}^{\mathbb{P}}[\frac{\xi_1}{\xi_2}|\mathbf{w}]}, \quad \psi(\hat{\xi_1}) = \mathbb{E}^{\mathbb{P}}\big[\psi(\xi_1) - \log \xi_2 |\mathbf{w}\big] + \log \hat{\xi_2}. \tag{5.12}$$

Applying the following difference equation,

$$\psi(x+h) - \psi(x) = \sum_{i=0}^{h-1} \frac{1}{x+i}, \tag{5.13}$$

with $h = 1$ follows that

$$\hat{\xi_2} = \frac{\hat{\xi_1}}{\mathbb{E}^{\mathbb{P}}[\frac{\xi_1}{\xi_2}|\mathbf{w}]}, \quad \hat{\xi_1} = \frac{1}{\mathbb{E}^{\mathbb{P}}[\psi(\xi_1+1) - \psi(\xi_1)|\mathbf{w}]} \tag{5.14}$$

A Markov Chain Monte Carlo (MCMC) simulation according to the Gibbs sampling scheme can be used to generate samples from the posterior distribution, but as the exact analytical form of posterior joint and marginal distributions are not available, it is more convenient to apply Metropolis-Hastings (MH) algorithm as prescribed by Moala et al. (2013). To generate samples from ξ_1 and ξ_2, we run Algorithm 1 where $f(\xi_1^{(t)}, \xi_2^{(t-1)}|\mathbf{w})$ is given by (5.9), and $\mathbf{Ga}(\frac{\xi_1^{(t-1)}}{c}, c)$ and $\mathbf{Ga}(\frac{\xi_2^{(t-1)}}{d}, d)$ are proposal distribution of which hype-parameters c and d are selected such that a good mixing of the chains and the convergence of the MCMC samples of parameters are obtained. The generated sample hereby is applied to find Bayes estimators of ξ_1 and ξ_2, denoted by ξ_1^B and ξ_2^B, through (5.12). These Bayesian estimations of real-world parameters are then regarded as estimations of their corresponding risk-neutral parameters, that is, we set $\hat{\epsilon_1} = \xi_1^B$ and $\hat{\epsilon_2} = \xi_2^B$.

Let us now consider the problem of example 4.3.6 in which under physical measure the inter-arrival time distribution $\mathbf{W}$ is a mixture of two exponential distributions characterized by parameters (ϕ, η_5, η_6), while under risk-neutral measure is a mixture of two exponential distributions characterized by parameters (ϕ^*, η_7, η_8). There are two ways to specify the likelihood function of a mixture model. To illustrate this clearly, we start with the basic definition. Generally, a random variable Y with a g-component mixture density can be generated in this way. Suppose that $\boldsymbol{Z}$ be the set of g latent variables where the k-th element of $\boldsymbol{Z}$ is defined to be zero or one, according to whether the component of origin of Y in the mixture is equal to k or not. More precisely, consider $\boldsymbol{Z}_j$ as the latent variable vector for the j-th sample point. Therefore, $\boldsymbol{Z}_j$ can

Algorithm 1 Metropolis-Hasting sampling for Gamma distribution

Initialize: $\xi_1^{(0)}$ and $\xi_2^{(0)}$

for $t = 1, 2, \cdots$ **do**

Generate new value $\xi_1^{(t)}$ from $\mathbf{Ga}(\frac{\xi_1^{(t-1)}}{c}, c)$, and accept it with the following probability known as the MH ratio:

$$u(\xi_1^{(t-1)}, \xi_1^{(t)}) = \min\left\{1, \frac{\mathbf{Ga}(\frac{\xi_1^{(t-1)}}{c}, c) f(\xi_1^{(t)}, \xi_2^{(t-1)}|\mathbf{w})}{\mathbf{Ga}(\frac{\xi_1^{(t)}}{c}, c) f(\xi_1^{(t-1)}, \xi_2^{(t-1)}|\mathbf{w})}\right\}$$

Generate new value $\xi_2^{(t)}$ from $\mathbf{Ga}(\frac{\xi_2^{(t-1)}}{d}, d)$, and accept it with the following probability:

$$u(\xi_2^{(t-1)}, \xi_2^{(t)}) = \min\left\{1, \frac{\mathbf{Ga}(\frac{\xi_2^{(t-1)}}{d}, d) f(\xi_1^{(t)}, \xi_2^{(t)}|\mathbf{w})}{\mathbf{Ga}(\frac{\xi_2^{(t)}}{d}, d) f(\xi_1^{(t)}, \xi_2^{(t-1)}|\mathbf{w})}\right\}$$

end for

be a vector like this:

$$\boldsymbol{Z}_j = (Z_{1j}, Z_{2j}, \cdots, Z_{gj}) \quad \text{or} \quad \boldsymbol{Z}_j = (Z_{kj}); \quad k = 1, 2, \cdots, g \tag{5.15}$$

where $Z_{kj} = 1$ if the j-th sample point belongs to the k-th component (or cluster), otherwise $Z_{kj} = 0$. We also know that each element of vector $\boldsymbol{Z}_j$ occurs independently with probabilities $\pi_1, \pi_2, \cdots, \pi_g$. So, it is easy to see that $\boldsymbol{Z}_j$ is distributed according to a multinomial distribution consisting of one draw on g categories with probabilities $\pi_1, \pi_2, \cdots, \pi_g$, i.e., $\boldsymbol{Z}_j \sim Multi(1, \boldsymbol{\pi})$, where $\boldsymbol{\pi} = (\pi_1, \pi_2, \cdots, \pi_g)$. The probability density function is then given by:

$$p(\boldsymbol{z}_j) = \mathbb{P}(\boldsymbol{Z}_j = \boldsymbol{z}_j) = \mathbb{P}(Z_{1j} = 1)^{z_{1j}} \times \mathbb{P}(Z_{2j} = 1)^{z_{2j}} \times \cdots \times \mathbb{P}(Z_{gj} = 1)^{z_{gj}} = \prod_{k=1}^{g} \pi_k^{z_{kj}} \tag{5.16}$$

For example, if we know that the j-th sample point belongs to the second component, then for the vector $\boldsymbol{Z}_j$ we observe the vector $\boldsymbol{z}_j = (z_{1j}, z_{2j}, \cdots, z_{jg}) = (0, 1, 0, \cdots, 0)$ with the probability function:

$$p(\boldsymbol{z}_j) = \mathbb{P}\big(\boldsymbol{Z}_j = (0, 1, 0, \cdots, 0)\big) = \pi_2 \tag{5.17}$$

Let $\boldsymbol{z}_j = k$ denotes the fact that the j-th sample point fall in the k-th component, i.e., the k-th element of vector $\boldsymbol{z}_j$ is equal to one and the others are equal to zero. Based on this definition, we make another assumption which says that the conditional density of y_j given $\boldsymbol{z}_j = k$ is $f_k(y_j)$. Then, following the same logic used for the $p(\boldsymbol{z}_j)$, we can conclude that

$$f(y_j|\boldsymbol{z}_j) = \prod_{k=1}^{g} f_k(y_j)^{z_{kj}} \tag{5.18}$$

In order to obtain the mixture model of $f(y)$, we just need to apply the Bayes rule by summing up the terms on $\boldsymbol{z}$ to get the probability density function of $f(y)$. That is,

$$f(y) = \sum_{k=1}^{g} f(y, \boldsymbol{z}) = \sum_{k=1}^{g} f(y|\boldsymbol{z})p(\boldsymbol{z}) = \sum_{k=1}^{g} \pi_k f_k(y) \tag{5.19}$$

where $f_k(y)$ are called the component densities of the mixture model and the π_k are mixing weights that satisfy:

$$\sum_{k=1}^{g} \pi_k = 1 \quad \text{and} \quad 0 \leqslant \pi_k \leqslant 1 \quad (k = 1, 2, \cdots, g) \tag{5.20}$$

Suppose that $\mathcal{Y} = (y_1, y_2, ..., y_n)$ is the vector of n observed data points. Using (5.19), the likelihood function is then given by

$$f(\mathcal{Y}; \Theta) = \prod_{i=1}^{n} \left\{ \sum_{k=1}^{g} \pi_k f_k(y_i; \Theta_k) \right\} \tag{5.21}$$

Another useful decomposition of (5.21) is based on latent variables. We consider the vector $\mathcal{X} = (\mathcal{Y}, \mathcal{C})$ involving latent variables Z_i defined in (5.15), for which $\mathcal{C} = (Z_1, Z_2, ..., Z_n)$, where $Z_i = (Z_{ki})$ with $\mathbb{P}(Z_{ki} = 1) = \pi_k$, for $i = 1, 2, ..., n$ and $k = 1, 2, ..., g$. Therefore, the complete version of the likelihood function becomes

$$\begin{aligned} f(\mathcal{X}; \Theta) = f(\mathcal{Y}, \mathcal{C}; \boldsymbol{\Theta}) &= \prod_{i=1}^{n} f(y_i, Z_i; \boldsymbol{\Theta}) = \prod_{i=1}^{n} f(y_i|Z_i; \boldsymbol{\Theta})p(Z_i) \\ &= \prod_{i=1}^{n} \left[\prod_{k=1}^{g} f_k(y_i; \Theta_k)^{Z_{ki}} \right] \times \prod_{k=1}^{g} \pi_k^{Z_{ki}} \\ &= \prod_{i=1}^{n} \prod_{k=1}^{g} [\pi_k f_k(y_i; \Theta_k)]^{Z_{ki}} \end{aligned} \tag{5.22}$$

We adopt the above-mentioned setting to construct our conditional distribution for Bayesian framework under prior distributions selected for unknown parameters contained in the mixture model. For simplicity in notation and avoidance of confusion, suppose that $\mathbf{z} = (z_1, z_2, \cdots, z_n)$ is a vector of latent variables where z_i, corresponding to the i-th sample point, takes value $k \in \{1, 2, \cdots, g\}$ with probability $p(z_i = k) = w_k$. The prior of mixing weights $\mathbf{w} = (w_1, w_2, \cdots, w_g)$ is chosen to be a standard Dirichlet prior, that is,

$$\mathbf{w} \sim \text{Dirichlet}_g(\alpha_1, \alpha_2, \cdots, \alpha_g) = \frac{\Gamma(\alpha)}{\prod_{k=1}^{g} \Gamma(\alpha_k)} \prod_{k=1}^{g} w_k^{\alpha_k - 1} \tag{5.23}$$

where $\alpha_1, \alpha_2, \cdots, \alpha_g$ are non-negative hyper-parameters, and $\alpha = \sum_{k=1}^{g} \alpha_k$. Inspired by (5.16), one can write the conditional distribution for the cluster allocations in this way:

$$p(\mathbf{z}|\mathbf{w}) = \prod_{k=1}^{g} w_k^{n_k} \tag{5.24}$$

where $n_k = \sum_{i=1}^{n} I_{\{k\}}(z_i)$ is the number of samples attributed to the k-th cluster. By marginalising (5.24) with respect to $\mathbf{w}$, we get the marginal distribution of $\mathbf{z}$,

$$p(\mathbf{z}) = \frac{\Gamma(\alpha)}{\Gamma(\alpha + n)} \prod_{k=1}^{g} \frac{\Gamma(\alpha_k + n_k)}{\Gamma(\alpha_k)} \tag{5.25}$$

Motivated by (5.22), the joint distribution for $\mathbf{y} = \mathcal{Y}$ and $\mathbf{z}$ is conveniently proportional with

$$p(\mathbf{y}, \mathbf{z}|\Theta) \propto \prod_{k=1}^{g} \left\{ \Gamma(\alpha_k + n_k) \prod_{i:z_i = k} f_k(y_i|\Theta_k) \right\} \tag{5.26}$$

It is assumed that the vector $\mathbf{w}$ is independent of $\mathbf{y}$, which can be considered reasonable in a mixture model. Therefore, from (5.23), (5.24), (5.25), we are able to find the conditional distribution $p(\mathbf{w}|\mathbf{z})$, i.e., the posterior distribution of $\mathbf{w}$,

$$p(\mathbf{w}|\mathbf{z}) = \frac{p(\mathbf{z}|\mathbf{w})p(\mathbf{w})}{p(\mathbf{z})} = \text{Dirichlet}_g(\alpha_1 + n_1, \alpha_2 + n_2, \cdots, \alpha_g + n_g) \tag{5.27}$$

In addition, it is possible to find the posterior distribution of Θ by associating a conjugate prior with each parameter Θ_k in (5.26). In other words,

$$p(\Theta|\mathbf{y}, \mathbf{z}) \propto \prod_{k=1}^{g} \left\{ \prod_{i:z_i = k} f_k(y_i|\Theta_k) \right\} \pi(\Theta_k) \tag{5.28}$$

We are now in a position to specify our non-parametric Bayesian framework for the mixture of two exponential distributions as follows:

$$\begin{aligned} y_i|z_i = k &\sim \mathbf{EXP}(\lambda_k) \\ \mathbf{w} = (w_1, w_2) &\sim \text{Dirichlet}_2(\alpha_1, \alpha_2) \\ \lambda_k &\sim \mathbf{Ga}(\tau_k, \Psi_k); \quad i = 1, 2, \cdots, n, \quad k = 1, 2. \end{aligned}$$

where $(\alpha_k, \tau_k, \Psi_k)$ are known hyper-parameters. The corresponding Gibbs samplers with states $t = 1, 2, \cdots$ can be written as follows:

Algorithm 2 Gibbs sampling for an exponential mixture

Initialize: $\mathbf{w}^{(0)} = (w_1^{(0)}, w_2^{(2)})$ and $\Theta^{(0)} = (\lambda_1^{(0)}, \lambda_2^{(0)})$
for $t = 1, 2, 3, \cdots$ **do**
Generate $z_i^{(t)}$ $(i = 1, \cdots, n, k = 1, 2)$ from

$$\mathbb{P}(z_i^{(t)} = k|w_k^{(t-1)}, \lambda_k^{(t-1)}, y_i) \propto w_k^{(t-1)} f(y_i|\lambda_k^{(t-1)}) = w_k^{(t-1)} \lambda_k^{(t-1)} \exp\{-\lambda_k^{(t-1)} y_i\}$$

Compute $n_k^{(t)} = \sum_{i=1}^n I_{\{k\}}(z_i)$ and $(s_k^y)^{(t)} = \sum_{i=1}^n I_{\{k\}}(z_i)\, y_i$
Generate $\mathbf{w}^{(t)}$ from $\text{Dirichlet}_2(\alpha_1 + n_1^{(t)}, \alpha_2 + n_2^{(t)})$
Generate $\lambda_k^{(t)}$ from $\mathbf{Ga}(\tau_k + n_k^{(t)}, \Psi_k + (s_k^y)^{(t)})$
end for

To be able to find the Bayes estimate of mixture parameters $\Theta = (w_1, w_2, \lambda_1, \lambda_2)$, we need to solve the minimization problem in (5.6), which can be easily shown to be equivalent to the following maximization problem:

$$\begin{aligned} \Theta^B &= \arg\min_{\hat{\Theta}} \mathbb{E}^{\mathbb{P}}[\text{KL}(\Theta \parallel \hat{\Theta}) \,|\, \boldsymbol{y}] \\ &= \arg\min_{\hat{\Theta}} \mathbb{E}_{\Theta}^{\mathbb{P}}\Big[\mathbb{E}_{\boldsymbol{Y}}^{\mathbb{P}}\big[\log\big(\frac{f(\boldsymbol{Y};\Theta)}{f(\boldsymbol{Y};\hat{\Theta})}\big)\big]\Big|\boldsymbol{y}\Big] \\ &= \arg\max_{\hat{\Theta}} \mathbb{E}_{\Theta}^{\mathbb{P}}\Big[\mathbb{E}_{\boldsymbol{Y}}^{\mathbb{P}}\big[\log f(\boldsymbol{Y};\hat{\Theta})\big]\Big|\boldsymbol{y}\Big] \end{aligned} \tag{5.29}$$

where the inner expectation is taken with respect to $\boldsymbol{Y}$, while the outer expectation is computed under the posterior distribution $\Theta|\boldsymbol{y}$. Due to the complexity with which the objective function in (5.29) can be optimized, we apply the idea of the EM method for which a complete version of the likelihood function is considered, that is, we replace $\log f(\boldsymbol{Y}; \hat{\Theta})$ with

$$\log f(\boldsymbol{Y}, \boldsymbol{Z}; \hat{\Theta}) = \sum_{j=1}^{n} \sum_{k=1}^{g} Z_{kj} \left[\log \hat{w}_k + \log f_k(y_j|\hat{\lambda}_k)\right] \tag{5.30}$$

The E-step suggests substituting the latent variable Z_{kj} with an expectation that is taken under probability distribution of the latent variable conditionally to the observed data and current value of the parameter, that is, we replace Z_{kj} with

$$\begin{aligned} \mathbb{E}[Z_{kj}|y_j, \hat{\Theta}_k] &= \mathbb{P}(Z_{kj} = 1|y_j, \hat{\Theta}_k) = \frac{f(y_j|Z_{kj} = 1, \hat{\lambda}_k)\mathbb{P}(Z_{kj} = 1)}{\sum_{l=1}^{g} f(y_j|Z_{lj} = 1, \hat{\lambda}_k)\mathbb{P}(Z_{lj} = 1)} \\ &= \frac{f_k(y_j|\hat{\lambda}_k)\hat{w}_k}{\sum_{l=1}^{g} f_k(y_j|\hat{\lambda}_k)\hat{w}_l} = H_k(y_j) \end{aligned} \tag{5.31}$$

Given the value $\widehat{\Theta^{(t-1)}}$ at the $(t-1)$-th iteration, the E-step calculates

$$\mathcal{Q}(\hat{\Theta}, \widehat{\Theta^{(t-1)}}) = \sum_{j=1}^{n} \sum_{k=1}^{g} (\log \hat{w}_k)\, \mathbb{E}^{\mathbb{P}}_{\Theta}\left[\mathbb{E}^{\mathbb{P}}_{\boldsymbol{Y}}\left[H_k^{(t-1)}(Y_j)\right]\middle|\boldsymbol{y}\right] + \mathbb{E}^{\mathbb{P}}_{\Theta}\left[\mathbb{E}^{\mathbb{P}}_{\boldsymbol{Y}}\left[H_k^{(t-1)}(Y_j)\log f_k(Y_j|\hat{\lambda}_k)\right]\middle|\boldsymbol{y}\right] \tag{5.32}$$

After that the M-step finds the revised parameter $\widehat{\Theta^{(t)}}$ according to the following problem

$$\widehat{\Theta^{(t)}} = \arg\max_{\hat{\Theta}} \mathcal{Q}(\hat{\Theta}, \widehat{\Theta^{(t-1)}}) \tag{5.33}$$

The iteration between the E-step and M-step continues until a convergence criterion is satisfied. Solving (5.33) yields

$$\widehat{w_k^{(t)}} = \frac{\sum_{j=1}^{n} \mathbb{E}^{\mathbb{P}}_{\Theta}\left[\mathbb{E}^{\mathbb{P}}_{\boldsymbol{Y}}\left[H_k^{(t-1)}(Y_j)\right]\middle|\boldsymbol{y}\right]}{n} \tag{5.34}$$

$$\widehat{\lambda_k^{(t)}} = \arg\max_{\hat{\lambda}_k} \sum_{j=1}^{n} \mathbb{E}^{\mathbb{P}}_{\Theta}\left[\mathbb{E}^{\mathbb{P}}_{\boldsymbol{Y}}\left[H_k^{(t-1)}(Y_j)\log f_k(Y_j|\hat{\lambda}_k)\right]\middle|\boldsymbol{y}\right] \tag{5.35}$$

6 Numerical illustration

6.1 Parameter calibration of the CIR model

To calibrate the parameters of the CIR model (3.1), we take the daily historical yields on the 3-month US treasury bills from January 2, 1990, to May 25, 2022[12] as depicted in Figure 2.

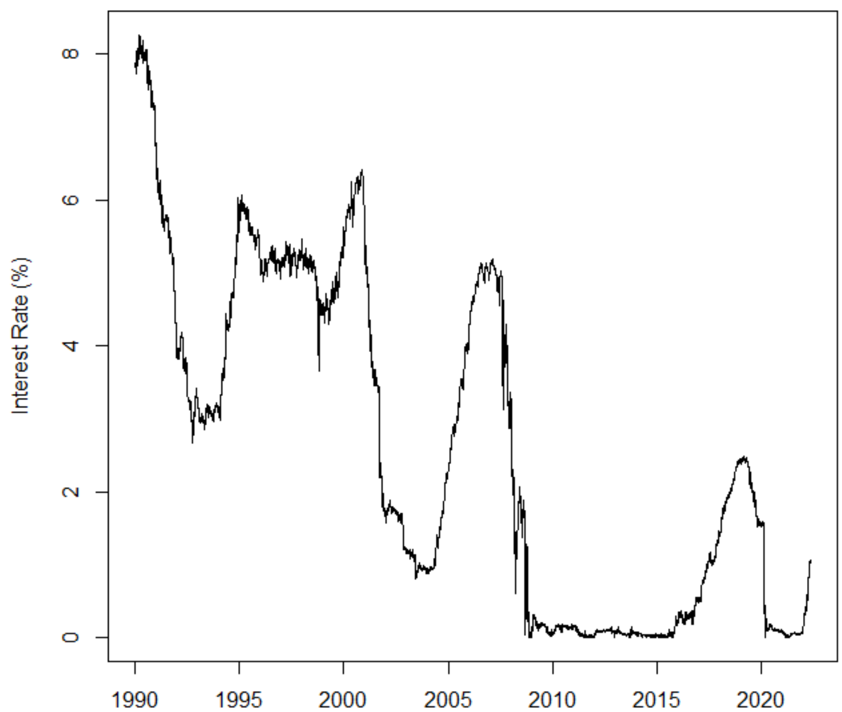


Figure 1: Historical data on interest rates observed from January 2, 1990, to May 25, 2022.

In this section, we use the maximum likelihood estimation (MLE) method for parameters of the CIR model (3.1), of which the transition probability density is given by

$$f(r_t|r_u) = c^* p_{\chi^2_{(d_1,d_2)}}(c^* r_t), \quad t > u, \tag{6.1}$$

where $c^* = \frac{4\theta}{\sigma^2(1-e^{-\theta(t-u)})}$ and $p_{\chi^2_{(d_1,d_2)}}(.)$ denotes the probability density of a non-central chi-square distribution with the degree of freedom $d_1 = \frac{4\theta m}{\sigma^2}$ and the non-central parameter $d_2 =$

[12]Note here that it is not essential to use the same time period for the catastrophe loss observations, as we already assumed that insurance risk and financial risks are independent.

$\frac{4\theta e^{-\theta(t-u)}}{\sigma^2(1-e^{-\theta(t-u)})} r_u$ (see, e.g., Fergusson 2019). If the data $r^* = \{r_1, r_2, \cdots, r_n\}$ is given according to the historical data, then the log-likelihood function can be written as:

$$l(\theta, m, \sigma; r^*) = \sum_{i=1}^{n} ln(c^*) + \sum_{i=1}^{n} ln(p_{\chi^2_{(d_1,d_2)}}(c^* r_i | r_{i-1})) \tag{6.2}$$

For the purpose of finding the MLE of (6.2), a numerical optimization technique can be applied. For that, initial values for parameters are essential to be specified to start the iteration process embedded in the algorithm. In this paper, the initial values are obtained using the ordinary least square estimation (OLSE) method. The main idea to achieve approximate estimates of parameters contained in the CIR model on the basis of OLSE method is to find a regression version that can describe a discretized version of the CIR model derived by the Euler discretization technique. According to the Euler scheme, the SDE associated with the CIR model (3.1) can be represented as follows:

$$r_{t_{i+1}} - r_{t_i} = \theta(m - r_{t_i})\Delta t_i + \sigma\sqrt{r_{t_i}}\Delta W_i \tag{6.3}$$

where $\Delta t_i = (t_{i+1} - t_i)/250$[13] and $\Delta W_i = W_{t_{i+1}} - W_{t_i}$, for $i = 0, 1, 2, \cdots, n-1$. After performing some simple algebraic manipulations, one can realize that the matrix representation of the regression model corresponding to the discretized version (6.3) is of the form:

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{bmatrix}}_{\boldsymbol{Y}} = \underbrace{\begin{bmatrix} \sqrt{\frac{\Delta t_1}{r_1}} & \sqrt{\Delta t_1 r_1} \\ \sqrt{\frac{\Delta t_2}{r_2}} & \sqrt{\Delta t_2 r_2} \\ \vdots & \vdots \\ \sqrt{\frac{\Delta t_{n-1}}{r_{n-1}}} & \sqrt{\Delta t_{n-1} r_{n-1}} \end{bmatrix}}_{\boldsymbol{Z}} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} \sigma\mathcal{N}_1(0,1) \\ \sigma\mathcal{N}_2(0,1) \\ \vdots \\ \sigma\mathcal{N}_{n-1}(0,1) \end{bmatrix}}_{\boldsymbol{\epsilon}} \tag{6.4}$$

where $y_i = \frac{r_{t_{i+1}} - r_{t_i}}{\sqrt{\Delta t_i r_{t_i}}}$, $\beta_1 = \theta m$, $\beta_2 = -\theta$, and $\mathcal{N}_i(0,1)$ are independent random variables, each having standard normal distribution. Then, the OLSE of $\boldsymbol{\beta}$ and σ become

$$\hat{\boldsymbol{\beta}}_{OLS} = (\boldsymbol{Z}^T\boldsymbol{Z})^{-1}\boldsymbol{Z}^T\boldsymbol{Y}, \quad \widehat{\sigma^2} = \frac{1}{n-2}||\boldsymbol{Y} - \boldsymbol{Z}\hat{\boldsymbol{\beta}}_{OLS}||^2 \tag{6.5}$$

where $||.||$ denotes the Euclidean distance. We use these estimates as initial values for a numerical optimization of the likelihood function (6.2). Table 1 represents the final results for the OLS and MLE of the CIR parameters in (3.1).

[13] Since a daily base observations are used for the interest rate, the time step of the Euler scheme is computed as the time difference between consecutive points divided by the number of working days per year, which as a

Method (Under physical measure)	$\hat{\theta}$	$\hat{m}$	$\hat{\sigma}$
OLSE (Initial values)	0.277	0.018	0.077
MLE (Optimal values)	0.254	0.011	0.074

Table 1: Estimated parameters of the CIR model under the physical measure.

The risk-neutralized parameters represented in (3.4) can be found by knowing the estimated physical parameters as well as the constant parameter λ which determines the market price of interest rate risk. Following the same idea in Ahmad and Wilmott (2006), we find an estimation for parameter λ empirically. We summarize our steps as follows:

For simplicity in notation, suppose that $P(t, T, r_t)$ denotes the zero-coupon bond price at time t with maturity time T. In the first step, we consider the partial differential equation (PDE) associated with the CIR model, which is given by

$$\frac{\partial P}{\partial t}(t,T,r_t) + [\theta(m - r_t) + \lambda r_t]\frac{\partial P}{\partial r_t}(t,T,r_t) + \frac{1}{2}\sigma^2 r_t \frac{\partial^2 P}{\partial r_t^2}(t,T,r_t) - r_t P(t,T,r_t) = 0 \tag{6.6}$$

In the second step, the Taylor series expansion of the zero-coupon bond price around $t = T$ with final condition $P(T, T, r_T) = 1$ is derived by the following polynomial representation:

$$P(t,T,r_t) = \sum_{j=0}^{\infty} c_j(r_t)(T-t)^j \tag{6.7}$$

Substituting (6.7) into (6.6) gives the following recursive relation for $c_j(r_t)$:

$$c_{j+1}(r_t) = \frac{1}{j+1}\left\{[\theta(m - r_t) - \lambda r_t]c_j'(r_t) + \frac{1}{2}\sigma^2 r_t c_j''(r_t) - r_t c_j(r_t)\right\}, \quad j = 0, 1, \cdots \tag{6.8}$$

where $c_j'(r_t)$ and $c_j''(r_t)$ are the first and second derivatives of function $c_j(r_t)$ with respect to r_t. Using (6.8), an approximation for the zero-coupon bond price is given by

$$P(t,T,r_t) \approx 1 - r_t(T-t) + \frac{1}{2}\left\{(\lambda + \theta)r_t - \theta m - r_t^2\right\}(T-t)^2 + \cdots, \quad \text{as} \quad t \longrightarrow T \tag{6.9}$$

From this we have

$$-\frac{\log(P(t,T,r_t))}{(T-t)} \approx -r_t + \frac{1}{2}\left\{\theta m - (\lambda+\theta)r_t\right\}(T-t) + \cdots, \quad \text{as} \quad t \longrightarrow T \tag{6.10}$$

standard convention is said to be 250 days.

Relation (6.10) shows that the slope of the yield curve at the short end is equal to $\frac{1}{2}\Big\{\theta m-(\lambda+\theta)r_t\Big\}$. Observing the slope of the yield curve and short rate each day can be useful for deriving a time series for the parameter λ. By comparing the empirical slope with its corresponding analytical slope, we can generate this time series. We report the value of λ by taking an average of this time series. To end this, we consider a dataset consisting of daily rates on US treasury bills with maturity times of 3 months, 6 months, and 1 year. The 3-month maturity rate is considered the short rate, while the 6-month and 1-year maturity rates are used to calculate the yield curve's slope per day. Using this approach, we have determined that $\lambda = -0.033$. This value is then applied to find the risk-neutral parameters, see Table (2).

Risk neutral parameters	θ^*	m^*	σ^*
Estimation	0.221	0.013	0.074

Table 2: Estimates of the CIR model parameters under risk-neutral measure.

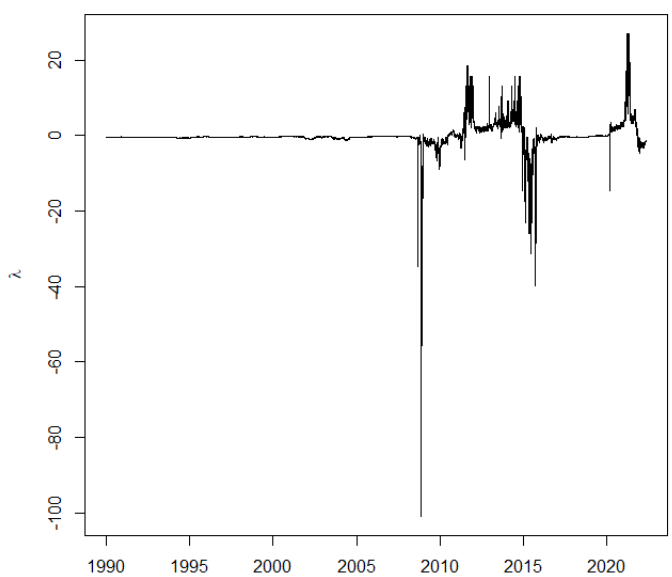


Figure 2: λ time series

6.2 Parameters estimation of aggregate claim process

6.2.1 Loss data description

The data employed in this paper is the property damages (in dollars) resulting from earthquakes in the US, spanning the period from 1906 to 2005; more information can be found in Vranes and Pielke (2009). To take the inflation into account, we adjust the data to 2020 value using the Consumer Price Index[14] (CPI). As the yearly observation for the CPI covers years between 1947 to 2021, a generalized additive model is used to predict CPI values from 1906 to 1946. Figure 3 shows the annual frequency of earthquake occurrence and the adjusted losses per year.

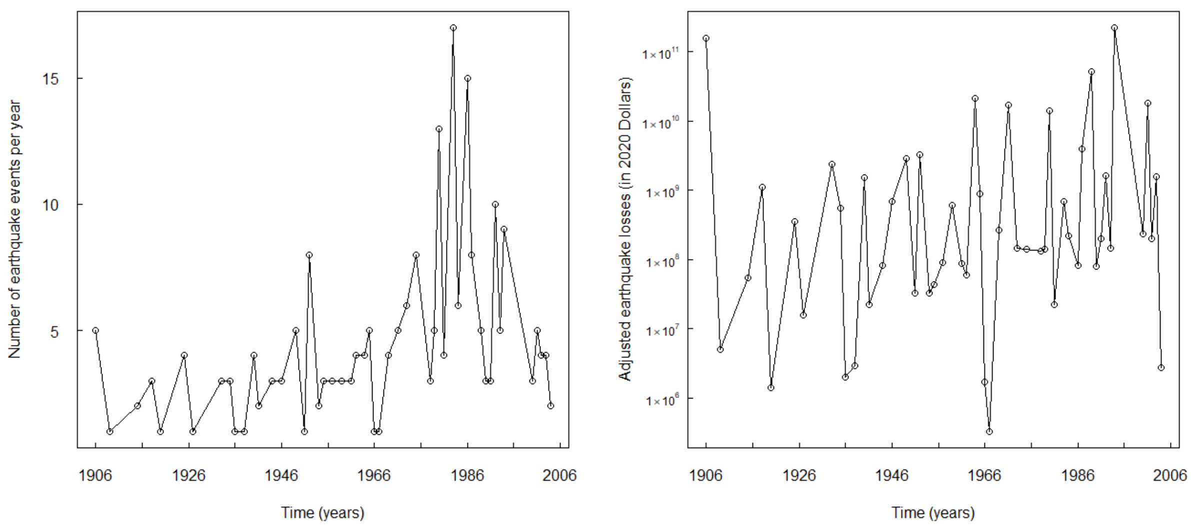


Figure 3: Left plot illustrates the number of claims, per year, from April 1906 to September 2004, right plot illustrates the adjusted loss severity (on log-scale), per year, from April 1906 to September 2004.

6.2.2 Loss distribution

In the CAT bond pricing literature, a wide range of heavy-tailed distributions can be considered for the loss data. The following loss severity distributions, as discussed in Giuricich and Burnecki (2019), are selected to fit the data at our disposal:

[14] Available at https://fred.stlouisfed.org/series/CPIAUCSL

A Burr type XII distribution with shape parameters $a > 0$ and $b > 0$, and scale parameter $c > 0$ is given by:

$$\text{Burr}(a, b, c) = \frac{\frac{ab}{c}(\frac{x}{c})^{a-1}}{(1 + (\frac{x}{c})^a)^{b+1}} \tag{6.11}$$

A generalized Pareto distribution with shape parameter $a \in \mathbb{R}$ and scale parameter $c > 0$ is given by:

$$\text{GP}(a, c) = \frac{1}{c}(1 + \frac{ax}{c})^{-(1+\frac{1}{a})} \tag{6.12}$$

A generalized Extreme Value distribution with shape parameter $a \in \mathbb{R} \setminus \{0\}$, location parameter $b \in \mathbb{R}$, and scale parameter $c > 0$ is given by:

$$\text{GEV}(a, b, c) = \frac{1}{c} exp\left(-(1 + \frac{a(x-b)}{c})^{-\frac{1}{a}}\right) \times \left(1 + \frac{a(x-b)}{c}\right)^{-1-\frac{1}{a}} \tag{6.13}$$

Using the MLE method, we estimate all parameters contained in the above distributions, and then we perform the goodness-of-fit test for fitted distributions by means of well-known non-parametric tests, namely, the Kolmogorov-Smirnov (KS), Anderson-Darling (AD), and Cramér Von Mises (CVM) tests, each of which measures the discrepancy between empirical distribution and theoretical distribution in its own way (see, e.g., Ma and Ma 2013). Following Giuricich and Burnecki (2019), we approximate the p-values corresponding to each test via the bootstrap method as non-parametric tests basically require distributions to be fully specified while in our case we applied fitted distributions.

Distribution	Parameters	MLE
Burr type XII	$\widehat{a, b, c}$	$0.82, 0.70, 2.85 \times 10^7$
Generalized Pareto	$\widehat{a, c}$	$2.2, 2.6 \times 10^7$
Generalized extreme value	$\widehat{a, b, c}$	$2.6, 1.05 \times 10^8, 2.72 \times 10^8$

Table 3: Estimated parameters of loss severity distributions using MLE

The MLE of parameters and goodness-of-fit test results at the confidence level 0.05 have been provided in Tables 3 and 4, respectively. The results show that GEV cannot be fitted to our data. To check the performance of fitted distributions, we use Akaik information criteria (AIC) and Bayesian information criteria (BIC). The results of Table 5 show that GPD fits the best among others.

Distribution	KS	AD	CVM
Burr type XII	$0.08(0.94)$	$1.76(0.94)$	$0.15(0.95)$
Generalized Pareto	$0.04(0.60)$	$0.56(0.30)$	$0.04(0.70)$
Generalized extreme value	$0.30(< 0.05^{*})$	$51.80(< 0.05)$	$9.20(< 0.05)$

Table 4: Test statistic values of in-sample goodness-of-fit tests for fitted distributions with their corresponding p-values in the parenthesises, each being estimated by 1000 Monte Carlo simulation. (*): the significance level is equal to $\%5$.

Criteria	Burr type XII	Generalized Pareto	Generalized Extreme Value
AIC	9217.523	9207.845	9310.124
BIC	9227.798	9214.695	9320.398

Table 5: Goodness-of-fit criterion for checking the performance of different fitted loss distributions.

6.2.3 Frequency distribution

In regard to the distribution of the counting process N_t which models the number of claims, we already considered different distributions for inter-arrival times, namely, Exponential, Gamma, Weibull, and a mixture of two Exponential, each resulting in different models for N_t. Our goal here is to fit these distributions to arriving time observations (in days) extracted from the recorded dates of earthquake occurrence and then assess their fit adequacy by implementing goodness-of-fit tests as elaborated in the previous subsection.

Non-parametric tests provided in Table 6 suggest that all distributions are suited for the arriving time observation. We observe from Table 8 that the Gamma distribution possesses the lowest AIC and BIC values, which means it fits the best among others. This provides evidence supporting the fact that in real applications, there exist cases where the exponential assumption for the inter-arrival times of catastrophe events may not be an optimal choice, and so one needs to work under a more general framework than the Poisson process.

Distribution	Parameters	MLE
Exponential	$\widehat{\delta}$	0.002
Gamma	$\widehat{\eta_1, \eta_2}$	0.76, 0.002
Wiebull	$\widehat{\eta_3, \eta_4}$	0.84, 398.70
Mix-Exponential	$\widehat{\phi, \eta_5, \eta_6}$	0.87, 0.002, 0.02

Table 6: Estimated parameters of inter-arrival time distributions (under physical measure) using MLE, where the Expectation-Maximization algorithm is adopted for the case of the Mix-Exponential distribution.

Distribution	KS	AD	CVM
Exponential	$0.09(0.13)$	$1.18(0.08)$	$0.15(0.16)$
Gamma	$0.07(0.26)$	$0.25(0.74)$	$0.04(0.70)$
Wiebull	$0.07(0.13)$	$0.31(0.53)$	$0.05(0.46)$
Mix-Exponential	$0.06(0.42)$	$0.27(0.61)$	$0.03(0.65)$

Table 7: Test statistic values of in-sample goodness-of-fit tests for fitted distributions with their corresponding p-values in the parenthesises, each being estimated by performing 1000 Monte Carlo simulations.

Criteria	Exponential	Gamma	Wiebull	Mix-Exponential
AIC	1175.830	1173.201	1173.815	1177.676
BIC	1178.249	1178.039	1178.653	1187.351

Table 8: Goodness-of-fit criterion for checking the performance of different fitted inter-arrival time distributions.

6.3 Sensitivity analysis

Let the CAT bond contract of our interest cover earthquake losses with face value $Z = 1$ USD and $q = 0.5$. The distribution for severity variables is considered to be a GPD as defined in (6.12) with estimated parameters depicted in Table 3. The threshold level D is selected as a point beyond which the mean excess plot is linearly increasing according to the Extreme Value Theorem. We start by looking at scenario I, where under the physical measure L_t is a CRP with inter-arrival distribution being Exponential, Gamma, Weibull, and Mixture of exponential, while under risk-neutral measure L_t is a CPP. We aim to investigate how the CAT bond price changes with respect to the risk premium c and maturity time T under the aforesaid distributions. The function $\beta_3(x)$ is chosen for transforming the severity distribution function once switching from the physical measure to the risk-neutral measure. It is easy to show that by choosing $\beta_3(x)$, we will end up with $\kappa = 1$ and $\nu(x) = \frac{\exp\{cx\}}{\mathbb{E}^{\mathbb{P}_2}[\exp\{cX_1\}]}$, where the parameter c can be seen as an indicator of the earthquake risk premium, and $\nu(x)$ is the well-known Esscher transform, which is commonly employed as a kernel function for pricing in incomplete markets. Figure 4 represents the change of CAT bond price with respect to premium and maturity

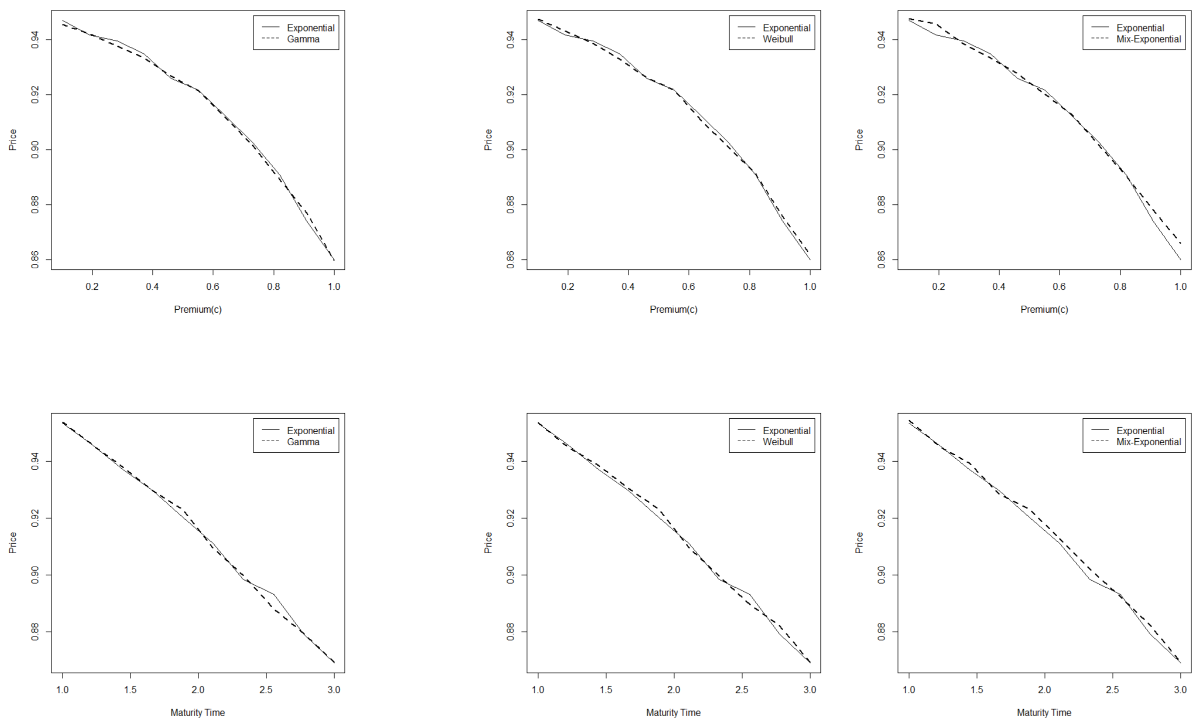


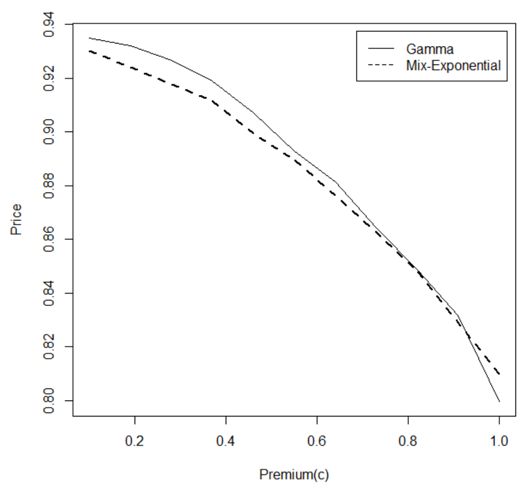
Figure 4: Plots on the top depict the change of CAT bond price with respect to the risk premium c, assuming different inter-arrival time distributions with maturity $T = 2$, while those on the bottom show the change of CAT price with respect to maturity T with $c = 0.6$.

time, indicating the fact that a higher-risk premium (required by the bond investor) and a longer maturity time (i.e., the time at which the contract is expired) result in a lower bond price. While CAT bond prices with Gamma and Weibull distributions are deemed to behave similarly to that with exponential distribution (considered as a benchmark), the bond price with the mixture of exponential distribution tends to be higher for very small and very large values of premium c. As far as the change of maturity is concerned, the CAT bond price with the mixture of exponential tends to be consistently higher than that with exponential distribution, in contrast to the case of Gamma and Weibull distributions in which the CAT bond price is either higher or lower.

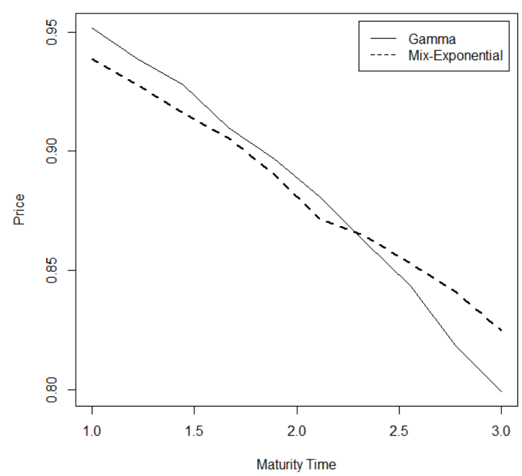
In scenario II, we assumed that the inter-arrival time is distributed as a Gamma or a mixture of exponential under both measures, in that, L_t is a compound renewal process under $\mathbb{P}_2$ and $\mathbb{Q}_2$, in which case we applied Bayesian inference to find parameter estimates of inter-arrival time distributions under $\mathbb{Q}_2$. Table 9 summarizes the results based on relations (5.14), (5.34), and (5.35):

Distribution	Parameters	Bayes estimate
Gamma	$\widehat{\epsilon_1, \epsilon_2}$	2.474, 0.01
Mix-Exponential	$\widehat{\phi^*, \eta_7, \eta_8}$	0.85, 1.063, 0.063

Table 9: Estimate of parameters contained in Gamma and mixture of two exponential distributions using Bayesian framework when scenario II is under consideration



(a)



(b)

Figure 5: Plot (a) and (b) show the change of CAT bond price with respect to premium c and maturity time T, respectively, by considering scenario II for which the inter-arrival time distribution under both measures is assumed to be Gamma distribution (solid line) and a mixture of two exponential distribution (dash line).

We observe from Figure 5 that the CAT bond price with Gamma distribution is higher than the mixture of exponential for a small value of premium c while their price difference becomes negligible as the premium increases. On the other hand, the CAT bond price with Gamma distribution is higher compared with the mixture of exponential before maturity $T = 2.3$ (in the year), and as we move beyond this point the price for the mixture of exponential becomes higher. Based on our analysis, it is obvious that the choice of inter-arrival distribution impacts more significantly on the CAT bond price in scenario II.

7 Conclusion

We have developed a pricing model for a zero-coupon CAT bond, in which the aggregate loss process used in the CAT bond payoff function is allowed to follow a general compound renewal process, where the inter-arrival time distribution can be a different distribution, such as Gamma, Weibull, or a mixture of exponential, than the exponential distribution. Our pricing framework ensures that a compound renewal process under the physical measure remains a compound renewal process under a risk-neutral measure. We derived the pricing formula under two general scenarios, where the aggregate loss process either follows a compound renewal process under the physical measure but a compound Poisson process under the risk-neutral measure (scenario I), or follows a compound renewal process under both measures (scenario II).

Throughout the paper, we made the assumption that financial risk and insurance risk behave independently. Therefore, we modeled interest rates separately using the Cox-Ingersol-Ross model, and calibrated the market price of risk using market data. For the loss distribution, we fitted different heavy-tailed distributions to earthquake data, and selected the Generalized Pareto distribution based on the AIC and BIC criteria. Under scenario I, we used Maximum Likelihood Estimation and Expectation Maximization methods to estimate the parameters contained in the inter-arrival time distribution and severity distribution. In contrast, under scenario II, we solved a Bayes minimization problem based on the Kulback-Laibler divergence loss function. Finally, we conducted a sensitivity analysis to examine the impact of premium and maturity time on the CAT bond price.

The numerical experiments demonstrated that the choice of inter-arrival time distribution has a significant impact on the price of the CAT bond under both scenarios, with scenario II showing a more pronounced effect. Furthermore, the results indicate that an increase in the premium and maturity time of the CAT bond leads to a decrease in its price. To establish a benchmark price, we assumed that the aggregate loss process L_t follows a Compound Poisson process under both measures, implying that the counting process follows a Poisson process under both measures. In scenario I, we observed that the deviation of the bond price from our benchmark price was greater for the mixture of exponential distribution than for the Gamma and Weibull distributions at very large and small values of premium.

References

Feller, W. 1951. Two singular diffusion problems. Ann. of Math. (2), 54:173–182. (Cited on pages 290 and 551).

Cox, J. C., J. J. E. Ingersoll, and S. A. Ross. 1985. A theory of the term structure of interest rates. Econometrica, 385–407.

Macheras, N.D. and Tzaninis, S.M., 2020. A characterization of equivalent martingale measures in a renewal risk model with applications to premium calculation principles. Modern Stochastics: Theory and Applications, 7(1), pp.43-60

Delbaen, F. and Haezendonck, J., 1989. A martingale approach to premium calculation principles in an arbitrage-free market. Insurance: Mathematics and Economics, 8(4), pp.269-277.

Muermann, A., 2003. Actuarially consistent valuation of catastrophe derivatives. The Wharton Financial Institution Center Working Paper Series, pp.03-18

Foss, S., Konstantopoulos, T. and Zachary, S., 2007. Discrete and continuous time modulated random walks with heavy-tailed increments. Journal of Theoretical Probability, 20(3), pp.581-612

ANTONIO MOALA, F.E.R.N.A.N.D.O., LUIZ RAMOS, P.E.D.R.O. and ALBERTO ACHCAR, J.O.R.G.E., 2013. Bayesian inference for two-parameter gamma distribution assuming different noninformative priors. Revista Colombiana de Estadística, 36(2), pp.319-336.

Godin, F., VS. Lai, and DA. Trottier. 2019. A general class of distortion operators for pricing contingent claims with applications to CAT bonds. Scandinavian Actuarial Journal 2019(7):558-84.

Wesolowski, C.A., Wanasundara, S.N., Wesolowski, M.J., 2016. A gamma-distribution convolution model of 99mTc-MIBI thyroid time-activity curves. EJNMMI Phys 3, 31 (2016).

Di Salvo, F., 2006. The exact distribution of the weighted convolution of two gamma distributions. In Atti della XLIII Riunione Scientifica SIS (Acts of the 43 Annual Meeting of the Italian Scientific Society). Università di Torino: Società Italiana di Scientifica (pp. 511-4).

Fergusson, K., 2019. Asymptotics of bond yields and volatilities for extended CIR models under the real-world measure. Scandinavian Actuarial Journal, 1-36.

Ahmad, R. and Wilmott, P., 2006. The market price of interest-rate risk: Measuring and modelling fear and greed in the fixed-income markets. Wilmott magazine, pp.64-70.

Vranes, K. and Pielke, R., 2009. Normalized earthquake damage and fatalities in the United States: 1900-2005. Natural Hazards Review, 10(3), p.84.

Giuricich, M.N. and Burnecki, K., 2019. Modelling of left-truncated heavy-tailed data with application to catastrophe bond pricing. Physica A: Statistical Mechanics and Its Applications, 525, pp.498-513.

Ma, Z.G. and Ma, C.Q., 2013. Pricing catastrophe risk bonds: A mixed approximation method. Insurance: Mathematics and Economics, 52(2), pp.243-254.

Beer, S. and Braun, A., 2022. Market-consistent valuation of natural catastrophe risk. Journal of Banking & Finance, 134, p.106350.

Cox, S. H. and Pedersen, H. W. (2000). Catastrophe risk bonds. North American Actuarial Journal, 4, pp.56-82.

Lee, J., and Yu, M. (2002). Pricing default-risky CAT bonds with moral hazard and basis risk. Journal of Risk and Insurance, 69, pp. 25-44.

Merton, R. C. (1976). Option pricing when the underlying stock returns are discontinuous. Journal of Financial Economica, pp. 125-144.

Tang, Q., and Zhongyi, Y. (2019). CAT bond pricing under a product probability with POT risk characterization. ASTIN Bulletin: The of the IAA, 49, pp. 457-490.

Pelsser, A., and Stadje, M. (2014). Time-consistent evaluations. Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics, 24, pp. 25-65.

Burnecki, K., Giuricich, M. N., and Palmowski, Z. (2019). Valuation of contingent convertible catastrophe bonds—The case for equity conversion. Insurance: Mathematics and Economics, 88, 238-254.

Ma, Z., Ma, C., and Xiao, S. (2017). Pricing zero-coupon catastrophe bonds using EVT with doubly stochastic Poisson arrivals. Discrete Dynamics in Nature and Society, 2017.

Lane, M. N. (2000). Pricing risk transfer transactions. ASTIN Bulletin, 30, 259-293.

Lane, M. N., & Olivier, M. (2008). Catastrophe Risk Pricing: An Empirical Analysis. Policy Research Working Paper, 4765. World Bank, Washington, DC.

Papachristou, D. (2011). Statistical analysis of the spread of catastrophe bonds at the time of issue. ASTIN Bulletin: The Journal of the International Actuarial Association, 41, pp. 251–277.

Braun, A. (2016). Pricing in the primary market for CAT bonds: new empirical evidence. Journal of Risk and Insurance, 83, pp. 811–847.

Gomez, L., and Carcamo, U. (2014). A multi-factor pricing model for CAT bonds in the secondary market. Journal of Business, Economics and Finance, 3, pp. 247-258.

Götze, T., Gürtler,M., and Witowski, E. (2020). Improving CAT bond pricing models via machine learning. Journal of Asset Management, 21, pp. 428–446.

Makariou, D., Barrieu, P., and Chen, Y. (2021). A random forest based approach for predicting spreads in the primary catastrophe bond market. Insurance: Mathematics and Economics, 101, Part B, pp. 140–162.

Kijima, M., and Murimachi, Y. (2008). An extension of the Wang transform derived from Buhlmann's economic premium principle for insurance risk. Insurance: Mathematics and Economics, 42, pp. 887–896.

Hamada, M., and Sherris, M. (2003). Contingent claim pricing using probability distortion operator: methods from insurance risk pricing and their relationship to financial theory. Applied Mathematical Finance, 10, pp. 19–47.

Li, B., Wang, T., and Tian, W. (2013). Risk measures and asset pricing models with new versions of Wang transform. In Uncertainty Analysis in Econometrics with Applications (pp. 155-167). Berlin, Heidelberg: Springer.

Domfeh, D., Chatterjee, A., and Dixon, M. (2022). A Bayesian valuation framework for catastrophe bonds. http://dx.doi.org/10.2139/ssrn.3998313

Safarveisi, S., Hirbod, A. (2021). Financial and Actuarial Valuations of CAT bonds, under revision.

A Appendix

Proof. Recall that the distribution of renewal process N_t under measure $\mathbb{Q}_2$ is given by

$$\begin{aligned}
\mathbb{Q}_2(N_t = n) &= \mathbb{Q}_2(\{N_t \geq n\} \cap \{N_t \geq n+1\}^c) \\
&= \mathbb{Q}_2(\{N_t \geq n\}) - \mathbb{Q}_2(\{N_t \geq n+1\}) \\
&= \mathbb{Q}_2(T_n \leq t) - \mathbb{Q}_2(T_{n+1} \leq t), \qquad \text{(A.1)}
\end{aligned}$$

where $T_n = \sum_{i=1}^{n} W_i$ (i.e., the arrival time of n-th event) can be written as the sum of n interarrival times W_i's, each of which follows a gamma distribution with parameter $n\epsilon_1$ and ϵ_2. Similarly, T_{n+1} is a Gamma with parameters $(n+1)\epsilon_1$ and ϵ_2. Hence, we have that

$$\mathbb{Q}_2(N_t = n) = \sum_{s=n\epsilon_1}^{+\infty} \frac{e^{-\epsilon_2 t}(\epsilon_2 t)^s}{s!} - \sum_{s=(n+1)\epsilon_1}^{+\infty} \frac{e^{-\epsilon_2 t}(\epsilon_2 t)^s}{s!} = \sum_{s=n\epsilon_1}^{n\epsilon_1+\epsilon_1-1} \frac{e^{-\epsilon_2 t}(\epsilon_2 t)^s}{s!}, \qquad \text{(A.2)}$$

which means that the event that exactly n events have occurred in the renewal process may be thought of as the event that between $n\epsilon_1$ and $n\epsilon_1 + \epsilon_1 - 1$ have occurred in the underlying Poisson process with intensity parameter ϵ_2. Accordingly, in example 4.3.5, we can say that the problem of renewal process can be turned into the problem of Poisson process where the counting process N_t takes its values between $n\epsilon_1$ and $n\epsilon_1 + \epsilon_1 - 1$ for $n = 0, 1, 2, \cdots$. ∎

B Appendix

Proof. To find the distribution function associated with L_t under the assumptions presented in example 4.3.6, we first obtain the probability function of claim number N_t, which itself follows a counting renewal process. To this purpose, it is then essential, according to relation (A.1), to derive the distribution function associated with the n-th arrival time T_n, which is a sum of i.i.d random variables following a mixture of two exponential distributions **MIX-EXP**$(\rho(\delta))$. We express our problem in the following with a general format.

Let random variable Y follow a mixture of two exponential distributions that is

$$Y \sim \pi_1 \mathbf{EXP}(\lambda_1) + \pi_2 \mathbf{EXP}(\lambda_2), \qquad \text{(B.1)}$$

where π_i are mixing weights with the property that $\pi_1 + \pi_2 = 1$. Furthermore, assume that $\{Y_1, Y_2, \cdots, Y_n\}$ is a sample of i.i.d random variables drawn from $f_Y(y)$. Then, the probability density function of $Z = \sum_{j=1}^{n} Y_j$ is given by

$$f_Z(z) = \sum_{k=0}^{n} \binom{n}{k} \pi_1^k \pi_2^{n-k} \frac{\lambda_1^k \lambda_2^{n-k}}{\Gamma(n)} e^{-\lambda_1 z} z^{n-1} {}_1F(n-k, n, (\lambda_1 - \lambda_2) z) \tag{B.2}$$

with ${}_1F(a, b, c)$ being a confluent hypergeometric function of the first kind defined as below:

$${}_1F(a; c; Z) = \begin{cases} \sum_{i=1}^{\infty} \frac{(a)_i}{(b)_i} \frac{c^i}{i!} & \text{hypergeometric series representation} \\ \frac{\Gamma(b)}{\Gamma(b-1)\Gamma(a)} \int_0^1 e^{ct} t^{a-1} (1-t)^{b-a-1} dt & \text{Integral representation} \end{cases} \tag{B.3}$$

where the Pochhammer symbol $(a)_i = \frac{\Gamma(a+i)}{\Gamma(a)} = a(a+1) \cdots (a+i-1)$.

Define random vector $\boldsymbol{\Upsilon} = (\Upsilon_1, \Upsilon_2, \cdots, \Upsilon_n)$ consisting of n i.i.d Bernoulli variable $\Upsilon_i \sim$ **Ber**(π_1) which are associated with each Y_i in this way:

$$Y_i | \Upsilon_i = k \sim \textbf{EXP}(\lambda_{2-k}); \quad \text{for} \quad k = 0, 1. \tag{B.4}$$

We define $\Upsilon^* = \sum_{i=1}^{n} \Upsilon_i \sim \textbf{Bin}(n, \pi_1)$. Conditioning on $\boldsymbol{\Upsilon}^*$, we can write

$$Z = \sum_{j=1}^{n} Y_j | \Upsilon^* = \left(\sum_{j:\Upsilon_j = 0} Y_j + \sum_{j:\Upsilon_j = 1} Y_j \right) \Big| \Upsilon^*, \tag{B.5}$$

from which one can conclude that $Z | \boldsymbol{\Upsilon}^* \sim Z_1 + Z_2$, where

$$Z_1 \sim \textbf{Ga}(\Upsilon^*, \lambda_1) \quad \text{and} \quad Z_2 \sim \textbf{Ga}(n - \Upsilon^*, \lambda_2) \tag{B.6}$$

This shows that the distribution of Z given Υ^* can be written as the sum of two independent random variables distributed as gamma with different shape and rate parameters. Denote by GDC(a, b, c, d; x) the gamma distribution convolution consisting of **Ga**(a, b) and **Ga**(c, d), which is given by:

$$\text{GDC}(a, b, c, d; x) = \begin{cases} \frac{b^a d^c}{\Gamma(a+c)} e^{-bx} x^{a+c-1} {}_1F(c, a+c, (b-d)x), & x > 0 \\ 0, & x \leq 0 \end{cases}, \tag{B.7}$$

(see, e.g., Wesolowski et al. (2016) and Di Salvo (2006)). Finally, marginalizing in Z leads to

$$\begin{aligned} f_Z(z) &= \sum_{k=0}^{n} f(z|\Upsilon^*)p(\Upsilon^*=k) \\ &= \sum_{k=0}^{n} \binom{n}{k} \pi_1^k \pi_2^{n-k} \text{GDC}(k,\lambda_1,n-k,\lambda_2;z), \end{aligned}$$

which can be viewed as a mixture of $(n+1)$ components with mixing weights equal to $\binom{n}{k}\pi_1^k\pi_2^{n-k}$. Using (A.1) and (B.2), we derive the probability function of the renewal process N_t in example 4.3.6,

$$\begin{aligned} \mathbb{Q}_2(N_t=n) &= \mathbb{Q}_2(T_n \le t) - \mathbb{Q}_2(T_{n+1} \le t) \\ &= \sum_{k=0}^{n} \binom{n}{k} \pi_1^k \pi_2^{n-k} \text{CGDC}(k,\lambda_1,n-k,\lambda_2;t) \\ &\quad - \sum_{k=0}^{n+1} \binom{n+1}{k} \pi_1^k \pi_2^{(n+1)-k} \text{CGDC}(k,\lambda_1,(n+1)-k,\lambda_2;t) \end{aligned}$$

where the notation CGDC denotes the cumulative distribution function of GDC, $\pi_1 = \phi^*$, $\pi_2 = 1-\phi^*$, $\lambda_1 = \eta_7$, and $\lambda_2 = \eta_8$. ■

C Appendix

Proof. Assume that $\mathcal{F}^W = \{\mathcal{F}_n^W\}_{n\in\mathbb{N}_0}$ and $\mathcal{F}^X = \{\mathcal{F}_n^X\}_{n\in\mathbb{N}_0}$ are the natural filtration of W and X, respectively. It can be shown that the following holds true (for unexplained details, interested readers can refer to Macheras and Tzaninis (2020)):

$$\forall A \in \mathcal{F}_t^L \quad \exists B_k \in \sigma(\mathcal{F}_t^W \cup \mathcal{F}_t^X)(\text{for every} k \in \mathbb{N}_0) \ s.t \ A \cap \{N_t=k\} = B_k \cap \{N_t=k\} \tag{C.1}$$

Using (C.1) we yield

$$\mathbb{Q}_2(A) = \sum_{k=0}^{\infty} \mathbb{Q}_2\big(B_k \cap \{N_t=k\}\big) = \sum_{k=0}^{\infty} \mathbb{Q}_2\big(B_k \cap \{T_k \le t\} \cap \{W_{k+1} > t-T_k\}\big) \tag{C.2}$$

For a fixed but arbitrary $n \in \mathbb{N}_0$, we define

$$G_n = \bigcap_{j=1}^{n} (W_j^{-1}(E_j) \cap X_j^{-1}(F_j)) \cap \{W_{n+1} > t - T_n\} \tag{C.3}$$

where $E_j, F_j \in \mathcal{B}(\mathbb{R}_+)$ for any $j \in \{1, 2, \cdots, n\}$. We then have that

$$\begin{aligned}
\mathbb{Q}_2(G_n) &= \mathbb{E}^{\mathbb{Q}_2}\left[I_{E_1}(W_1)I_{F_1}(X_1)\cdots I_{E_n}(W_n)I_{F_n}(X_n)I_{\{W_{n+1}>t-T_n\}}\right] \\
&= \left[\prod_{j=1}^{n} \mathbb{E}^{\mathbb{Q}_2}[I_{E_j}(W_j)]\right] \times \left[\prod_{j=1}^{n} \mathbb{E}^{\mathbb{Q}_2}[I_{F_j}(X_j)]\right] \times \mathbb{E}^{\mathbb{Q}_2}[I_{\{W_{n+1}>t-T_n\}}] \\
&= \left[\prod_{j=1}^{n} \mathbb{E}^{\mathbb{P}_2}\left[\frac{d\mathbb{Q}^2_{W_1}}{d\mathbb{P}^2_{W_1}}(W_j)I_{E_j}(W_j)\right]\right] \times \left[\prod_{j=1}^{n} \mathbb{E}^{\mathbb{P}_2}\left[\frac{d\mathbb{Q}^2_{X_1}}{d\mathbb{P}^2_{X_1}}(X_j)I_{F_j}(X_j)\right]\right] \\
&\quad \times \frac{\mathbb{E}^{\mathbb{Q}_2}[I_{\{W_{n+1}>t-T_n\}}]}{\mathbb{E}^{\mathbb{P}_2}[I_{\{W_{n+1}>t-T_n\}}]} \times \mathbb{E}^{\mathbb{P}_2}[I_{\{W_{n+1}>t-T_n\}}] \\
&= \mathbb{E}^{\mathbb{P}_2}\left[I_{G_n} \prod_{j=1}^{n} \left(\frac{d\mathbb{Q}^2_{W_1}}{d\mathbb{P}^2_{W_1}}(W_j)\frac{d\mathbb{Q}^2_{X_1}}{d\mathbb{P}^2_{X_1}}(X_j)\right) \frac{\mathbb{E}^{\mathbb{Q}_2}[I_{\{W_{n+1}>t-T_n\}}]}{\mathbb{E}^{\mathbb{P}_2}[I_{\{W_{n+1}>t-T_n\}}]}\right]
\end{aligned} \tag{C.4}$$

A monotone class argument allows one to employ (C.4) in (C.2) for any $k \in \mathbb{N}_0$, which leads to the following Radon-Nikodym derivative.

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}^2_t} = \left[\prod_{j=1}^{N_t} \left(\frac{d\mathbb{Q}^2_{W_1}}{d\mathbb{P}^2_{W_1}}(W_j)\frac{d\mathbb{Q}^2_{X_1}}{d\mathbb{P}^2_{X_1}}(X_j)\right)\right] \times \frac{\mathbb{Q}_2(W_{n+1} > t - T_{N_t})}{\mathbb{P}_2(W_{n+1} > t - T_{N_t})} \tag{C.5}$$

■