Pricing Index-linked Catastrophe Bonds via Monte Carlo Simulation

Justin van der Merwe

A dissertation submitted to the Faculty of Commerce, University of Cape Town, in partial fulfilment of the requirements for the degree of Master of Philosophy.

May 30, 2016

MPhil in Mathematical Finance,
University of Cape Town.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Philosophy in the University of the Cape Town. It has not been submitted before for any degree or examination in any other University.

Justin van der Merwe

May 30, 2016

Abstract

The pricing framework used in this dissertation allows for the specification of catastrophe risk under the real-world measure. This gives the user a great deal of freedom in the assumptions made about the underlying catastrophe risk process (referred to in this dissertation as the aggregate loss process). Therefore, this dissertation aims to shed light on the effect of various assumptions and considerations on index-linked CAT bond prices based on the Property Claims Services (PCS) index. Also, given the lack of a closed-form solution to the pricing formulae used and the lack of a liquidly-traded secondary market, this dissertation compares two approximation methods to evaluate expressions involving the aggregate loss process: Monte Carlo simulation and a mixed-approximation method. The two price-approximation methods are largely consistent and seem to agree particularly in the upper quantiles of the distribution of the aggregate loss process. Another key consideration is that the third-party estimating the catastrophe losses in North America, PCS, only records catastrophe losses above $25 million. This dissertation therefore also explores the issue of left-truncated data and its effect when estimating the parameters of the aggregate loss process. For this purpose, it introduces a non-parametric approach to compare, in sample, the results of ignoring the threshold and taking it into account. In both these exercises, it becomes apparent that very heavy-tailed distributions need to be used with caution. In the former case, the use of very heavy-tailed distributions places restrictions on the distributions that can be used for the mixed-approximation method. Finally, as a more realistic avenue this dissertation proposes a simple stochastic intensity model to compare with the deterministic intensity model and found that, by parsimony, the deterministic intensity seems to provide a reasonable model for the upper quantiles of the aggregate loss process. The key results of this dissertation are that the pricing of CAT bonds depends on the quantiles of the aggregate loss process, as in evident both when comparing the approximation methods and the deterministic and stochastic intensity functions, and that left-truncation should be taken into account when valuing index-linked CAT bonds using data from PCS.

Acknowledgements

I would like to thank my supervisors, Mario Giuricich and A/Prof. Peter Ouwehand. I am very grateful for their support and accessibility when tackling any problems I encountered. Also, I would like to thank Manuel and Luby Washkansky for the scholarship that assisted with my Master's studies. Finally, I owe all of this to my parents for their unwavering support throughout my academic career.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Problems arise for insurers when facing catastrophe risks which cannot be reduced by diversifying across policies and regions. Catastrophe risks can be categorised as high-severity, low-probability risks and do not typically conform to the law of large numbers (Anderson *et al.*, 2000). Traditionally, insurers make use of the reinsurance market to mitigate this risk (Anderson *et al.*, 2000). However, the occurrence of Hurricane Andrew in 1992 and the Northridge Earthquake in 1994 called into question the reinsurance market's ability to bear this risk and motivated the need to transfer some of it to the financial markets (Cummins and Weiss, 2009). Therefore, catastrophe (CAT) bonds were introduced as a means of transferring this risk to capital markets where the potential losses of insurers from these catastrophes pale in comparison to the size of these markets (Anderson *et al.*, 2000).

A CAT bond has payoffs that are linked to the occurrence of predefined catastrophic events and forms part of a broader category of assets called insurance-linked securities. CAT bonds often focus mainly on catastrophic property risks. By design, they are challenging to price as there are few tradable assets with payoffs linked to the occurrence of catastrophes (Cox *et al.*, 2000) and therefore replication becomes difficult. However, this is why investors are attracted by these types of investments, since they are uncorrelated with financial markets and are therefore seen as zero beta investments (Cox *et al.*, 2000; Cummins and Weiss, 2009). The high yields of these instruments also make them attractive for potential investors (Anderson *et al.*, 2000). On the supply side, Cummins and Weiss (2009) attribute the rise in CAT bond issuance to the growth in property values in catastrophe-prone areas, which further compromises reinsurers' abilities to bear the risk of these events. One can also point to increased urbanisation in these catastrophe-prone areas as contributing to the severity of natural disasters and therefore the prevalence of CAT bonds.

1.1 CAT Bond Structure

The typical structure of a CAT bond issue, as shown in Figure 1.1 below, is described as follows.

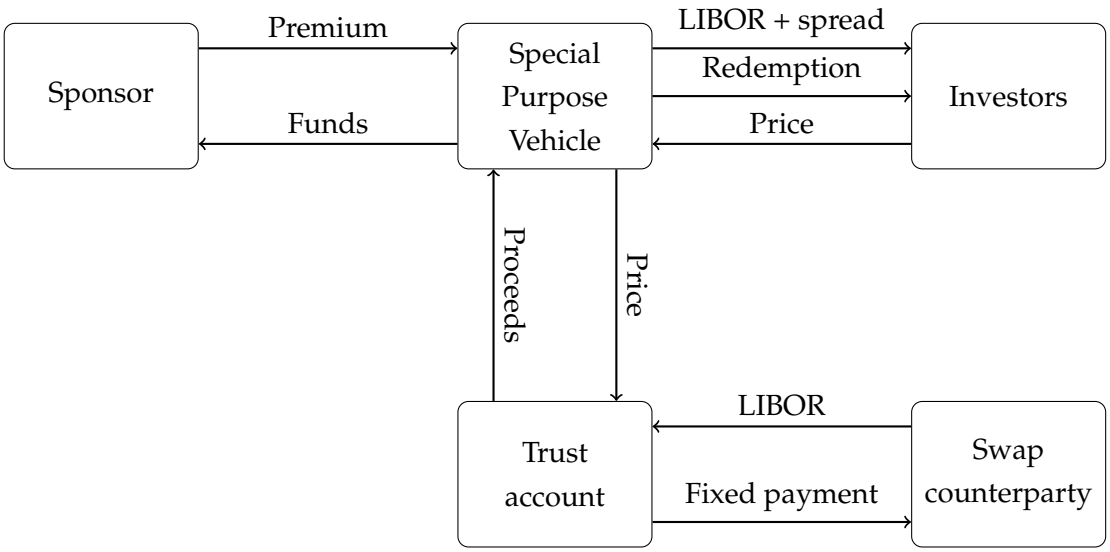


Fig. 1.1: Typical structure of a simplified CAT bond.

Issuing a CAT bond requires the creation of a special-purpose vehicle (SPV), which receives the funds from the CAT bond investors and deposits them in a trust account (Cummins and Weiss, 2009). The entity issuing the CAT bond is known as the sponsor and is typically an insurance company. Note that a CAT bond will, in general, 'trigger' when the relevant recorded losses from the prespecified catastrophic events exceeds some predefined level or if the recorded characteristics from a prespecified catastrophe meet some criteria. The three main types of CAT bond triggers are defined in Section 1.2. If the CAT bond triggers, the funds are released to the sponsor by the SPV and the amount paid is usually proportional to the losses from the event (Cummins and Weiss, 2009). As can be easily noted, there is effectively a reinsurance contract between the sponsor and the SPV, where the sponsor pays premiums to the SPV (Anderson *et al.*, 2000). In return for providing this protection, investors receive coupon payments based on the London Interbank Offer Rate (LIBOR) plus a spread (Cummins and Weiss, 2009). In addition, an interest rate swap protects the sponsor from interest rate movements (Cummins and Weiss, 2009).

The structure of CAT bonds ensures that they are fully-collateralised, significantly reducing credit risk (Cummins and Weiss, 2009). This makes them more

favourable for the sponsor than typical reinsurance, as insurers are not exposed to the credit risk of a particular reinsurer (Anderson *et al.*, 2000). Furthermore, from an investor's point of view, the assets held by the trust provide collateral against the principal amount of the CAT bond. Also, the interest income received from these assets covers part of the coupon payments. Indeed, the low levels of credit risk may contribute to their popularity compared to other insurance-linked securities. The 2008 financial crisis, however, showed that these instruments are not completely free of credit risk, as was present both for the interest rate swap counterparty and the instruments invested in by the trust (Cummins and Weiss, 2009).

1.2 Types of CAT Bond Triggers

According to Anderson *et al.* (2000), there are three types of CAT bond triggers: indemnity, index-linked and parametric. Indemnity triggers are defined in terms of losses experienced by the sponsor, index-linked triggers are defined in terms of industry-wide losses and parametric triggers are defined in terms of a physical measurement from a catastrophic event.

There is clearly moral hazard associated with using indemnity CAT bonds (Cox *et al.*, 2000; Cummins and Weiss, 2009). Anderson *et al.* (2000) noted that, in this case, insurers may be incentivised to relax their underwriting procedures. A further problem with indemnity triggers is that in order for investors to understand the underlying risk of their investment, they need to understand the risks of the particular insurer to which they are exposed (Cummins and Weiss, 2009). CAT bonds with this type of trigger can also take longer to settle as more verification is required in determining the particular insurer's losses (Cummins and Weiss, 2009). Index-linked and parametric triggers, on the other hand, expose CAT bond issuers to basis risk, with parametric triggers presenting the most basis risk (Anderson *et al.*, 2000).

The benefits of investing in index-linked CAT bonds are reduced moral hazard as well as no need to have specific information about the sponsor when compared to indemnity CAT bonds, and also reduced basis risk when compared to parametric triggers (Anderson *et al.*, 2000). Often, index-linked CAT bond triggers are based on data provided by third-party organisations such as the Insurance Services Office's Property Claims Services in the USA.

1.3 PCS Catastrophe Loss Data

Property Claims Services (PCS) is a division of the Insurance Services Office, a subsidiary of Verisk Analytics (Kerney, 2013). PCS is responsible for estimating insurers' losses in North America, and its associated territories, arising from natural disasters expected to cause more than $25 million in damage (Kerney, 2013). The time series of these PCS-recorded losses, above the threshold of $25 million, forms the PCS index. Here an underlying issue of left-truncated data arises, as only losses exceeding the threshold are recorded in the index. Note that this threshold was originally set at $5 million before it was increased to $25 million in 1997 (Chernobai *et al.*, 2006).

PCS provides an industry-wide estimate for total vehicle, personal property and commercial property claims in each state affected by a relevant recorded catastrophe (Kerney, 2013). PCS data is often used when specifying index-linked CAT bond triggers and is also sometimes used in indemnity triggers (where only PCS-declared catastrophes are used) (Kerney, 2013). Insurers and reinsurers can also make use of this data in setting catastrophe reserves (Kerney, 2013).

Artemis (2016) and a brief review of the literature highlights the growing use of PCS data when issuing CAT bonds: PCS data has been used in 35% of the total CAT bond issuance reported by PCS in 2015. A greater shift towards the use of index-linked triggers motivates the need for research into how the underlying PCS index should be modelled and, ultimately, how index-linked CAT bonds based on such an index should be priced.

1.4 Research and Motivation

Both sponsors of and investors in CAT bonds require a method for calculating CAT bond prices. Sponsors, or insurers in particular, need to understand the effects of different threshold levels, terms-to-maturity and other CAT-bond design-related factors on the amount of capital they receive from the issue. Pension funds, who make up a large proportion of CAT bond investors, need to value their assets periodically and therefore may require such an understanding. There are, however, many different approaches to value CAT bonds based on vastly different assumptions (Cummins and Geman, 1995; Baryshnikov *et al.*, 1998; Cox *et al.*, 2000) and therefore there is no consensus on how these values should be calculated. Given the inability to calculate an explicit price, there are a number of possible assumptions that one could make about the underlying catastrophe risk process, each having a different impact on the price that is calculated. This dissertation serves to

shed some light on the effect of using various assumptions on the calculated prices of CAT bonds, under a specified pricing framework. This dissertation is organised as follows.

Chapter 2 discusses previous literature relating to CAT bond pricing. Based on this discussion, Chapter 3 then defines the pricing framework to be used in this dissertation and derives CAT bond pricing formulae for zero-coupon and coupon-paying CAT bonds. The estimation of the parameters for the catastrophe risk process is then explored in Chapter 4, for both the loss severity (amount) and intensity (frequency). Given the lack of a liquid, secondary market, the comparison of any pricing formula to observable market-based CAT bond prices is difficult. Therefore, Chapter 5 considers two approximation methods, namely Monte Carlo simulation and a mixed-approximation method to examine the extent to which they agree. The issue of left-truncation is a key limitation in CAT bond pricing and is therefore considered in Chapter 6. Chapter 7 then looks at the effect on CAT bond prices when using a stochastic intensity, as opposed to a deterministic intensity. Finally, Chapter 8 provides a summary of the key conclusions. Overall, this dissertation attempts to contribute to existing literature by analysing the effects of various assumptions that could be made as well as highlights areas where the user needs to be particularly cautious.

Chapter 2

Literature Review

There has been much literature devoted to the pricing of CAT bonds, however, because of the incomplete pricing setting, there is no consensus on the appropriate approach. Incompleteness here refers to the inability to replicate the payoff of the underlying instrument and it must be noted that the CAT bond market is incomplete, since it is largely uncorrelated with financial markets (Cox and Pedersen, 2000).

Merely imposing that a market is arbitrage-free will only limit the price of an instrument to a range of possible arbitrage-free values (Cox and Pedersen, 2000). Indeed, Jarrow (2009) noted that assuming the market is arbitrage-free but not complete means that there is more than one equivalent martingale measure and therefore more than one price. So, Cox and Pedersen (2000) state that in order to calculate explicit prices for CAT bonds an assumption needs to be made about the distribution of the catastrophe risk process, in addition to assumptions about distributions for financial market variables such as interest rates. With this statement in mind, Cox and Pedersen (2000) developed a discrete-time pricing framework for catastrophe bonds, based on the theory of a representative agent. They assume that, in the real world, variables dependent only on financial events are independent of variables that depend only on catastrophic events. This assumption is consistent with previous studies. In this framework, they then show that: (1) these variables will be independent under the risk-neutral measure and (2) the expectation of a variable which depends only on catastrophe risk is the same under the real-world and risk-neutral measures. This effectively allows for the specification of catastrophe risk under the real-world measure and that of financial risk under the risk-neutral measure. Zimbidis *et al.* (2007) and Shao *et al.* (2015) used their framework for pricing earthquake CAT bonds. Ma and Ma (2013) applied it in a continuous setting and developed a mixed-approximation method for pricing zero-coupon and coupon-paying CAT bonds, since it is difficult to derive closed-form solutions under their framework.

Cummins and Geman (1995), on the other hand, used an arbitrage-free approach to price catastrophe insurance futures and call spreads. Baryshnikov *et al.* (1998) made the assumption of continuous trading and developed an arbitrage-free pricing framework for pricing CAT bonds. The work of Baryshnikov *et al.* (1998) was corrected, expounded and extended by Burnecki and Kukla (2003). Cox and Pedersen (2000) explained that to get explicit prices in an incomplete setting, real-world probabilities need to be used for the risk variable/s that cannot be hedged. The advantage of the approach of Cox and Pedersen (2000) is that, as previously mentioned, it allows for working under the real-world measure for the catastrophe risk, and the risk-neutral measure for financial risk. The disadvantage of the approach by Baryshnikov *et al.* (1998) is that, in contrast to Cox and Pedersen (2000), to use real-world catastrophe data they assume that the real-world and risk-neutral measures are the same. This is seen as a drawback of their approach (Nowak and Romaniuk, 2013).

Also, by design, catastrophe bonds are fairly similar to defaultable bonds. Therefore, Jarrow (2010) developed a pricing model for CAT bonds based on reduced-form, credit risk models. This pricing methodology has been implemented to value multiple-trigger CAT bonds based on drought risks in Kenya (Sun *et al.*, 2015). This approach, however, has only been implemented for single-event catastrophes.

There are other features which have been considered in the pricing of CAT bonds. The addition of credit risk, moral hazard and basis risk to the pricing of CAT bonds was considered by Lee and Yu (2002). The issue of credit risk in CAT bonds was also investigated by Liu *et al.* (2014). Unger (2010) developed a numerical partial differential equation approach for pricing callable CAT bonds. Nowak and Romaniuk (2013) considered CAT bonds with stepwise and piecewise-linear payoff functions.

An important component in the pricing of CAT bonds is the catastrophe risk process. For indemnity CAT bonds, this process is defined as the accumulated losses of the individual insurer and for index-linked CAT bonds, it is defined as the underlying loss index. Therefore, a key assumption is the manner in which this catastrophe risk process is modelled. Vaugirard (2003a,b) modelled the catastrophe risk process using a Poisson jump diffusion process. However, the catastrophe risk process is more typically modelled using a compound Poisson process (CPP) (Embrechts and Meister, 1997; Aase, 1999; Chernobai *et al.*, 2006; Ma and Ma, 2013). In the latter case, the catastrophe risk process is commonly referred to as the aggregate loss process, where it should be noted that losses are restricted to insured losses. Moreover, using a CPP is consistent with the way in which the PCS index is reported, with jumps occurring at the times when the catastrophe losses are re-

ported.

Chernobai *et al.* (2006) considered the issue of the left-truncated PCS loss data in calculating ruin probabilities using various continuous distributions. They noted that failing to account for the left-truncation, present in the PCS loss data, in the distribution fitting will result in an overestimated mean, underestimated variance and an underestimation of the upper quantiles of the distribution in the context of the PCS loss data. They therefore found that ignoring the threshold led to an underestimation of the ruin probabilities. The PCS loss data is of course a key input in estimating the intensity and severity of the aggregate loss process on which the CAT bond prices are based. However, as yet, this issue has not been addressed in the context of CAT bond pricing.

A potential criticism of the use of the non-homogeneous compound Poisson process, which is often used, is the specification of a deterministic intensity function for the arrival of catastrophes. Dassios and Jang (2003) suggested that any realistic model for the aggregate loss process must have a stochastic intensity component. Along with Albrecher *et al.* (2004) and Schmidt (2014), they used a shot-noise process to incorporate a stochastic intensity into the modelling of the loss process. The shot-noise process captures the exponential decay observed in the individual claims after a catastrophe occurs. This process, however, is suitable in the context of modelling individual claims and not aggregate losses, as one would not expect to observe an exponential decay of the aggregate losses after a catastrophe. Lin *et al.* (2009) considered a log-normal intensity model to price contingent capital and motivated this by the apparent exponential trend in the number of catastrophes observed over the 50 year period from 1949 to 1999. More recently, Ma *et al.* (2015) used the Black-Derman-Toy intensity model to price CAT bonds using PCS data from 2000 to 2010. A criticism of these two papers, however, is the manner in which they estimated their parameters for their proposed log-normal intensity. This process involved setting the volatility term to zero and fitting the resultant deterministic function using non-linear least squares. Ma *et al.* (2015) then calculated CAT bond prices based on varying levels of the volatility term. The criticism here is that all of the parameters are not estimated from the data, in particular the volatility.

Overall, the clear theme from the literature is the lack of consensus on the pricing approach needed to value CAT bonds. There is therefore a great deal of flexibility in the assumptions that can be made about the underlying catastrophe risk process and the estimation of the parameters.

Chapter 3

Pricing Framework

The pricing approach of this dissertation will follow that of Ma and Ma (2013), which is loosely based on the pricing framework of Cox and Pedersen (2000). This chapter serves to demonstrate how the main assumptions made in the discrete-time pricing framework of Cox and Pedersen (2000) ultimately simplify the pricing of CAT bonds in a continuous-time setting. This is done in a similar fashion to Ma and Ma (2013). Note that the pricing framework of Ma and Ma (2013), and therefore Cox and Pedersen (2000), is an equilibrium pricing model, rather than an arbitrage-free model, and is based on the theory of a representative agent. A detailed derivation of this discrete-time framework of Cox and Pedersen (2000) is given in Appendix A.

3.1 Probabilistic Setting

Firstly, a formal specification of the continuous-time pricing framework is given. This specification follows that of Ma and Ma (2013) very closely.

For $0 < T < \infty$, let $\mathbb{T} = [0, T]$ be a continuous trading interval. Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in\mathbb{T}})$ be a filtered probability space. Ω represents the states of the world, $\mathcal{F}$ is a σ-algebra of subsets of Ω and $\mathbb{P} : \mathcal{F} \to [0, 1]$ represents the real-world probability measure. $(\mathcal{F}_t)_{t\in\mathbb{T}}$ denotes an increasing filtration with $\mathcal{F}_t \subset \mathcal{F}\ \forall\, t \in \mathbb{T}$.

Let $\{V_t : t \in \mathbb{T}\}$ denote the CAT bond price process. The value of this process, particularly at $t = 0$, is the focus of this dissertation. Let $L = \{L_t : t \in \mathbb{T}\}$ denote the catastrophe risk process, which is the cumulative PCS loss index in this case. Here $\{L_t : t \in \mathbb{T}\}$ is assumed to follow a (non)-homogeneous compound Poisson process (and is consequently referred to as the aggregate loss process). Therefore, it is assumed that there is a (non)-homogeneous Poisson point process $\{N_t : t \in \mathbb{T}\}$ with a predictable bounded intensity $\lambda(t)$ which generates the occurrences of catastrophic events. Let $0 \leq t_1 \leq ... \leq t_i \leq ... \leq T$ denote the times of these catastrophic events over $\mathbb{T}$.

Note that L_t is predictable as well as left-continuous, and is defined by

$$L_t = \sum_{i=1}^{N_t} X_i \tag{3.1}$$

where $L_0 = 0$, $X = \{X_i : i \geq 1\}$ denotes the losses reported by the PCS index over time and $\{N_t : t \in \mathbb{T}\}$ is defined as above. Furthermore, these losses (X_i) are assumed to be independent and identically distributed, each having the cumulative distribution function F. The processes $\{N_t : t \in \mathbb{T}\}$ and $\{X_i : i \geq 1\}$ are assumed to be independent. Finally, $\{r_t : t \in \mathbb{T}\}$ denotes the risk-free rate process.

3.2 Zero-coupon and Coupon-paying CAT Bond Structures

This dissertation will consider two hypothetical CAT bond structures: a zero-coupon (ZC) CAT bond and a coupon-paying (CP) CAT bond. There is a tendency in the literature to consider these simpler CAT bond structures (Ma and Ma, 2013; Ma *et al.*, 2015; Nowak and Romaniuk, 2013) given the non-standardised nature of CAT bond issues in reality. These simpler structures can of course be extended to more realistic structures, which would depend on the term sheet of the particular CAT bond issue in question.

Let Z denote the redemption amount of the CAT bond, C denote the quarterly coupon payment of the CP CAT bond and D denote the threshold level.[1] ρ is defined to be the recovery rate, where $0 \leq \rho < 1$, and will reduce either the principal and/or coupon payments if the predefined catastrophic event occurs.

The payoff function of the ZC CAT bond at time T is defined by

$$\mathrm{P}_{ZC}(T) = \begin{cases} Z & \text{if } L_T \leq D \\ \rho Z & \text{if } L_T > D. \end{cases}$$

This structure is the same as that considered in Ma and Ma (2013).

Similarly, the payoff function of the CP CAT bond at times t_k for $t_1 = 0.25, ..., t_m = T$, where m is the number of quarters in $[0, T]$, is defined by:

$$\mathrm{P}_{CP}(t_k) = \begin{cases} C + \mathbb{I}_{\{t_k = T\}} Z & \text{if } L_{t_k} \leq D \\ \rho(C + \mathbb{I}_{\{t_k = T\}} Z) & \text{if } L_{t_k} > D. \end{cases}$$

Note also that the above CP CAT bond structure is more realistic than that considered by Ma and Ma (2013), who only included a coupon payment at the maturity

[1] In this dissertation, CP CAT bonds with quarterly coupon payments are considered; this is in line with a majority of CAT bond issues.

date. However, typically, the coupon payment C will be made up of LIBOR plus a spread, which is more in line with the typical structure of defaultable corporate bonds.

3.3 Pricing Formulae

Now that the instruments being valued have been given, the underlying market is specified. Firstly, the CAT bond market is assumed to be arbitrage-free and therefore there exists a risk-neutral measure $\mathbb{Q}$, which is in line with Burnecki *et al.* (2005). Moreover, the underlying aggregate loss process has jumps and the CAT bond market is therefore incomplete. Much research (Lee and Yu, 2002; Lin *et al.*, 2009; Ma and Ma, 2013) make use of the assumption that catastrophe risk is non-systematic and diversifiable, and therefore has a zero risk premium; this follows from the work of Merton (1976). This assumption is used in this dissertation.

As mentioned in Chapter 2, Cox and Pedersen (2000) make the assumption that variables dependent only on financial events are independent of variables that depend only on catastrophic events, under the real-world measure. Furthermore, the key results from their framework are that: the expectation of catastrophe risk variables is the same under the real-world and risk-free probability measures, and that catastrophe risk variables and financial market variables are independent under the risk-neutral measure. See Appendix A for further details. Importantly, these two assumptions allow for the use of real-world data for modelling the catastrophe risk and a risk-neutral model for calculating the discount factor. Therefore, there is no need to employ CAT bond market prices, which are difficult to obtain. In a similar vein as Ma and Ma (2013), given these two assumptions, the pricing formulae for both the ZC CAT bond and CP CAT bond are easily derived and it is these pricing formulae which illustrate, quite clearly, the necessity for real-world data.

Theorem 3.3.1. In an arbitrage-free market, the price of the ZC CAT bond is given by

$$V_t = B(t,T)Z[\rho + (1-\rho)\mathbb{P}(L_T \leq D)] \tag{3.2}$$

Proof.

$$\begin{aligned} V_t &= \mathbb{E}^{\mathbb{Q}}[e^{-\int_t^T r_s ds} \mathrm{P}_{ZC}(T)|\mathcal{F}_t] \\ &= \mathbb{E}^{\mathbb{Q}}[e^{-\int_t^T r_s ds}|\mathcal{F}_t]\ \mathbb{E}^{\mathbb{Q}}[\rho Z \mathbb{I}_{\{L_T > D\}} + Z\mathbb{I}_{\{L_T \leq D\}}|\mathcal{F}_t] \qquad (3.3) \\ &= B(t,T)\mathbb{E}^{\mathbb{P}}[\rho Z \mathbb{I}_{\{L_T > D\}} + Z\mathbb{I}_{\{L_T \leq D\}}|\mathcal{F}_t] \qquad (3.4) \\ &= B(t,T)Z[\rho + (1-\rho)\mathbb{P}(L_T \leq D)] \end{aligned}$$

□

where $B(t,T) = \mathrm{E}^{\mathbb{Q}}[e^{-\int_t^T r_s ds}|\mathcal{F}_t]$ and denotes the price of a riskless zero-coupon bond at time t with maturity at T. Equation (3.3) follows from the independence assumption and Equation (3.4) from the equivalence of the real-world and risk-neutral expectations for catastrophe risk variables.

Theorem 3.3.2. Similarly, the price of a CP CAT bond is given by

$$V_t = \sum_{k=1}^{m-1} B(t,t_k)C[\rho + (1-\rho)\mathbb{P}(L_{t_k} \leq D)] + B(t,T)(Z+C)[\rho + (1-\rho)\mathbb{P}(L_T \leq D)] \tag{3.5}$$

Proof.

$$\begin{aligned}
V_t &= \mathbb{E}^{\mathbb{Q}}\left[\sum_{k=1}^{m} e^{-\int_t^{t_k} r_s ds} \mathrm{P}_{CP}(t_k)\middle|\mathcal{F}_t\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\sum_{k=1}^{m-1} e^{-\int_t^{t_k} r_s ds} C(\rho \mathbb{I}_{\{L_{t_k}>D\}} + \mathbb{I}_{\{L_{t_k}\leq D\}}) + e^{-\int_t^{T} r_s ds}(Z+C)(\rho \mathbb{I}_{\{L_T>D\}} + \mathbb{I}_{\{L_T\leq D\}})\middle|\mathcal{F}_t\right] \\
&= \sum_{k=1}^{m-1} \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{t_k} r_s ds}\middle|\mathcal{F}_t\right] \mathbb{E}^{\mathbb{Q}}\left[C(\rho \mathbb{I}_{\{L_{t_k}>D\}} + \mathbb{I}_{\{L_{t_k}\leq D\}})\middle|\mathcal{F}_t\right] \\
&\quad + \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{T} r_s ds}\middle|\mathcal{F}_t\right] \mathbb{E}^{\mathbb{Q}}\left[(Z+C)(\rho \mathbb{I}_{\{L_T>D\}} + \mathbb{I}_{\{L_T\leq D\}}\middle|\mathcal{F}_t\right] \\
&= \sum_{k=1}^{m-1} B(t,t_k)\mathbb{E}^{\mathbb{P}}\left[C(\rho \mathbb{I}_{\{L_{t_k}>D\}} + \mathbb{I}_{\{L_{t_k}\leq D\}})\middle|\mathcal{F}_t\right] \\
&\quad + B(t,T)\mathbb{E}^{\mathbb{P}}\left[(Z+C)(\rho \mathbb{I}_{\{L_T>D\}} + \mathbb{I}_{\{L_T\leq D\}}\middle|\mathcal{F}_t\right] \\
&= \sum_{k=1}^{m-1} B(t,t_k)C[\rho + (1-\rho)\mathbb{P}(L_{t_k} \leq D)] + B(t,T)(Z+C)[\rho + (1-\rho)\mathbb{P}(L_T \leq D)].
\end{aligned}$$

□

Note that, for this dissertation, the financial risk analysis is limited to constant interest rates. Therefore,

$$B(t,s) = e^{-\int_t^s r ds} = e^{-r(s-t)}$$

for $s = t_1, \ldots, t_m = T$. A constant interest rate assumption is selected because the focus will be on the effects of varying assumptions of the catastrophe risk variables on the prices of the considered CAT bonds, as well as varying some of the features of the CAT bond issue. More specifically, most of the results of this dissertation focus on the differences in pricing surfaces under various assumptions. Therefore, it can possibly be argued that interest rate uncertainty will have similar effects on the pricing surfaces being compared. Given the simplified nature of the

CAT bond structures considered above, incorporating the current term structure into the CAT bond valuation simply amounts to obtaining the market-observable bond prices at the relevant dates and using these as discount factors. Of course, CAT bonds typically have coupon payments that are defined as a certain spread above an observable market interest rate such as LIBOR. In this case, the pricing formulae would require an appropriate interest rate model. However, this dissertation aims to study the effects of changing the catastrophe risk assumptions only. The constant interest rate assumption prevents any confounding effects of interest rate assumptions with those made about the catastrophe risk process.

It is clear from Formulae (3.2) and (3.5) above that the only random component in the pricing formulae is L_T or L_{t_k} for $k = 1, \ldots, m$. The remaining chapters describe the considerations in evaluating the probabilistic expressions involving the aggregate loss process at the required dates (e.g. $\mathbb{P}(L_T \leq D)$). The constant parameters that will be used for all CAT bond prices in this dissertation are given by $Z = 1, C = 0.05, \rho = 0.5$ and $r = 0.06$, which are in line with current literature.

Chapter 4

Distribution and Intensity Fitting

4.1 Data Description

The focus of this dissertation is on the pricing of PCS-linked CAT bonds and, in order to price such bonds, relevant data was required. The data set was based on the PCS loss index, as described in Section 1.3, and contained PCS catastrophe claims data from January 1985 to July 2011. The data was adjusted to January 2012 using the Consumer Price Index (CPI) obtained from the US Department of Labour. This adjustment using CPI is in line with Chernobai *et al.* (2006) and Ma and Ma (2013), though it may not be the best approach. An alternative adjustment would be to use the Case-Shiller Index. However, a disadvantage of using the index is that it does not consider moveable property. Furthermore, CPI was used so that results would be comparable to those of Chernobai *et al.* (2006) and Ma and Ma (2013). As in Chernobai *et al.* (2006), all claims exceeding $25 million or its historical equivalent are used for parameter estimation. A possible justification for this is that these are the only claims that affect the PCS loss index, given the left-truncation issue discussed in Chapter 2.

The purpose of this chapter is to estimate the parameters of the aggregate loss process in Equation (3.1). As mentioned before, typically, the catastrophe risk process is modelled as a non-homogeneous, compound Poisson process (Chernobai *et al.*, 2006; Ma and Ma, 2013). This requires the specification of the loss-severity and frequency components of the process.

4.2 Distribution Fitting

To model the losses X_i, continuous distributions are considered, where the support of the distribution is the positive real line, in order to prevent negative losses (Chernobai *et al.*, 2006). The exception is the generalised extreme value (GEV) distribution where the support depends on the parameters of the distribution. The

probability density functions (pdfs) as well as any restrictions on the parameters are given in Table 4.1.

Tab. 4.1: Continuous distribution functions to be considered for loss severity modelling, their probability density functions and any restrictions on the parameters.

Distribution	Probability Density Function (pdf), parameters and restrictions
Exponential	$f_X(x;\mu) = \frac{1}{\mu}e^{\frac{1}{\mu}x}$, $\mu > 0$
Log-normal	$f_X(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}x}e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}$, $\mu,\sigma > 0$
Gamma	$f_X(x;a,b) = \frac{1}{b^a\Gamma(a)}x^{a-1}e^{-\frac{x}{b}}$, $a,b > 0$
Weibull	$f_X(x;a,b) = \frac{b}{a}(\frac{x}{a})^{b-1}e^{-(\frac{x}{a})^b}$, $a,b > 0$
Burr type XII	$f_X(x;\zeta,c,k) = \frac{kc(\frac{x}{\zeta})^{c-1}}{\zeta(1+(\frac{x}{\zeta})^c)^{k+1}}$, $\zeta,c,k > 0$
Generalised Pareto (GP)	$f_X(x;k,\sigma) = \frac{1}{\sigma}(1+\frac{kx}{\sigma})^{-(1+\frac{1}{k})}$, $\sigma > 0, k \in \mathbb{R}$
Inverse Gaussian (IG)	$f_X(x;\mu,\lambda) = \sqrt{\frac{\lambda}{2\pi x^3}}\exp(-\frac{\lambda}{2\mu^2 x}(x-\mu)^2)$, $\mu,\lambda > 0$
Generalised Extreme Value (GEV)	$f_X(x;\mu,\sigma) = \begin{cases} \frac{1}{\sigma}\exp\left(-\left(1+\frac{k(x-\mu)}{\sigma}\right)^{-\frac{1}{k}}\right)\left(1+\frac{k(x-\mu)}{\sigma}\right)^{-1-\frac{1}{k}} & \text{if } k \neq 0, 1+k\left(\frac{x-\mu}{\sigma}\right) > 0 \\ \frac{1}{\sigma}\exp\left(-\exp\left(-\frac{(x-\mu)}{\sigma}\right)-\frac{(x-\mu)}{\sigma}\right) & \text{if } k = 0 \end{cases}$

Maximum likelihood estimation (MLE) is used to estimate the parameters with the exception of the generalised Pareto and GEV distributions. These two distributions are fitted using the method of Maximum Products of Spacing (MPS). Wong and Li (2006) noted that the MLE procedure may not converge for these distributions due to a potentially unbounded likelihood function (the unboundedness ultimately depending on the parameter values) and advocated the use of this alternative procedure in these two cases. Instead of finding the parameter(s) γ that maximise the likelihood function, this method aims to maximise the following statistic:

$$\mathcal{M}(\gamma) = \prod_{k=1}^{n+1}(F_\gamma(x_{(k)}) - F_\gamma(x_{(k-1)}))^{n+1}$$

which is, naturally, bounded by the definition of the cumulative distribution function F_γ. Note that $x_{(i)}$ refers to the ith order statistic.

For both the MLE and MPS estimation procedures, the losses are required to be independent and identically distributed. To a certain extent, in this data set, the independence assumption is justified by the fact that the data includes estimated losses from a number of different catastrophes, such as hurricanes, windstorms, fires and earthquakes.

Once the distributions are fitted using the relevant estimation procedure described above (either MLE or MPS), one needs to assess the goodness of fit. As in Chernobai *et al.* (2006) the following goodness of fit tests are considered: the Kolmogorov-Smirnov (D), Anderson-Darling (A^2), Cramer von-Mises (W^2) and Kuiper (V) tests. It should be noted that, as the parameters are estimated from the

data, Monte Carlo simulation is needed to calculate the p-values and critical values (Ross, 2002). So Monte Carlo simulation is used in the testing and 1000 simulations are used for each test.

Before proceeding, a caveat to consider is the fact that PCS only record losses that are expected to cause damages of $25 million or more. This is an issue which has not been addressed in previous literature in the context of CAT bond pricing, and is a novel feature of this dissertation. As mentioned in Chapter 2, Chernobai *et al.* (2006) addressed this issue when calculating ruin probabilities in the context of left-truncated data and noted that ignoring this left-truncation would lead to an underestimation of such probabilities. To account for this non-randomly missing data, Chernobai *et al.* (2006) recommended the use of the following procedure for estimating the parameters. This procedure is described in detail below.

Once again, let γ denote the vector of parameter(s) of the distribution being fitted. Now, there is a threshold $H > 0$ such that data is only observed in the interval $[H, \infty)$. Let $\mathcal{X}$ denote the sample space and suppose there is a sample $\mathbf{x} = (x_1, \ldots, x_n)$ from $[H, \infty)$ of size n, observed in $[T_1, T_2]$. The aim is to estimate γ whilst accounting for the left-truncation from H. Note that, in this context, H is equal to $25 million.

As specified in Chernobai *et al.* (2006), given that the total or 'complete' number of observations is unknown, the joint density (with respect to the product of the counting and Lebesgue measures) on $\mathcal{X}$ is given by:

$$g_{\mathbf{x}}(\mathbf{x}, \lambda^o(t), \gamma) = \frac{1}{n!} \left(\int_{T_1}^{T_2} \lambda^o(t) \mathrm{dt} \right)^n e^{-\int_{T_1}^{T_2} \lambda^o(t) \mathrm{dt}} \prod_{k=1}^{n} \frac{f_\gamma(x_k)}{1 - F_\gamma(H)}$$

where $F_\gamma(H) = \mathbb{P}(X \leq H)$, $\lambda^o(t)$ is the intensity function for the data in excess of the threshold H and f_γ as well as F_γ are the probability density function and cumulative density function respectively of the loss distribution with unknown parameter(s) γ.

Note that the independence of N_t and X_i is evident in the above formula. Also, this joint density implies that the parameters of the loss distribution can be estimated independently of the intensity (Chernobai *et al.*, 2006) - the estimation of the intensity is considered in Section 4.3.

The parameters of the distribution are estimated by calculating

$$\hat{\gamma}^c_{\text{MLE}} = \arg\max_{\gamma} \log \left(\prod_{k=1}^{n} \frac{f_\gamma(x_k)}{1 - F_\gamma(H)} \right), \tag{4.1}$$

which corresponds to maximising the log-likelihood, and therefore likelihood, of the observations $\mathbf{x}$ occurring.

Note that the Maximum Product of Spacings expression can similarly be adjusted to account for the left-truncation in the following manner

$$\mathcal{M}(\gamma^c) = \prod_{k=1}^{n+1} \left(\frac{F_{\gamma^c}(x_{(k)}) - F_{\gamma^c}(x_{(k-1)})}{1 - F_{\gamma^c}(H)} \right)^{n+1}.$$

The parameters are then estimated by calculating

$$\hat{\gamma}^c_{\text{MPS}} = \arg\max_{\gamma} \log \left(\prod_{k=1}^{n+1} \left(\frac{F_{\gamma}(x_{(k)}) - F_{\gamma}(x_{(k-1)})}{1 - F_{\gamma}(H)} \right)^{n+1} \right). \tag{4.2}$$

At this point, it is important to distinguish between the following two estimation procedures. The naïve estimation procedure refers to estimating the parameters of the distributions without taking into account the threshold level of \$25 million. The conditional estimation procedure refers to that proposed by Chernobai *et al.* (2006) and described above; a method that attempts to account for the left-truncation present in the data set.[2] The results of the naïve and conditional estimation procedures are given in Table 4.2 below.

Before proceeding to the goodness of fit tests for the fitted distributions in Table 4.2, some important observations can be made. Firstly, there is a clear change from the parameters fitted under the naïve approach to those under the conditional approach.[3] Also, as one would expect, the values of $F_{\hat{\gamma}^c}(H)$ are higher under the conditional approach. This is consistent with the results of Chernobai *et al.* (2006). There is a marked change in the values of $F_{\hat{\gamma}^c}(H)$ in the case of the gamma and the Weibull distributions. The gamma parameters fitted under the conditional approach show that almost 100% of the data lies below the threshold H. However, this feature is clearly unrealistic. Similarly, for the Weibull distribution, under the conditional approach $F_{\hat{\gamma}^c}(H)$ is approximately 52% which is also considerably high given the relatively low threshold level of \$25 million. Therefore, on this basis both the gamma and Weibull distributions are excluded from further analysis. Ma and Ma (2013) considered the GEV distribution for the loss process for final analysis. However, a potential issue with the use of this distribution is that the support of the distribution depends on the parameters. Indeed, for the conditional case, the

[2] Another possible method, in accounting for the left-truncation, would be to re-insert the sponsor's losses into the data set (i.e. those that are less than \$25 million). The problem with this approach is that, firstly, it may be extremely difficult to obtain such data and, secondly, it will not follow the same distribution as the PCS loss data.

[3] Note also that the MLE and MPS parameter estimates for the exponential, log-normal, gamma, Weibull, Burr and inverse Gaussian distributions were similar, showing the consistency of these two estimation procedures.

Tab. 4.2: Fitted parameters of the continuous distributions under the naïve and conditional fitting approaches, as well as calculated values of $F_{\widehat{\gamma}^c}(H)$.

Distribution	$\widehat{\gamma}, F_{\widehat{\gamma}^c}(H)$	Naïve approach	Conditional approach
Exponential	μ	5.63×10^8	5.38×10^8
	$F_{\widehat{\gamma}^c}(H)$	4.35%	4.54%
Log-normal	μ, σ	$18.59, 1.19$	$18.58, 1.49$
	$F_{\widehat{\gamma}^c}(H)$	5.06%	14.86%
Gamma	a, b	$0.54, 1.04 \times 10^9$	$6.79 \times 10^{-8}, 2.23 \times 10^9$
	$F_{\widehat{\gamma}^c}(H)$	14.84%	$\approx 100\%$
Weibull	a, b	$3.37 \times 10^8, 0.66$	$5.56 \times 10^7, 0.38$
	$F_{\widehat{\gamma}^c}(H)$	16.5%	52.32%
Burr type XII	ζ, c, k	$7.21 \times 10^7, 2.59, 0.36$	$9.53 \times 10^7, 1.57, 0.70$
	$F_{\widehat{\gamma}^c}(H)$	2.23%	7.78%
Generalised Pareto	k, σ	$0.59, 1.96 \times 10^8$	$0.73, 1.26 \times 10^8$
(GP)	$F_{\widehat{\gamma}^c}(H)$	11.56%	16.85%
Inverse Gaussian	μ, λ	$5.63 \times 10^8, 1.30 \times 10^8$	$5.28 \times 10^8, 9.38 \times 10^8$
(IG)	$F_{\widehat{\gamma}^c}(H)$	2.84%	6.28%
Generalised Extreme Value	k, μ, σ	$1.25, 1.47 \times 10^8, 1.26 \times 10^8$	$0.88, 1.03 \times 10^8, 9.48 \times 10^7$
(GEV)	$F_{\widehat{\gamma}^c}(H)$	0.78%	6.01%

support of the fitted distribution is $[-\$4.72 \text{ million}, \infty)$. The parameters estimated by Ma and Ma (2013) moreover suggest that the distribution can take on negative values - this was indeed checked. However, this is not pointed out in their paper. This feature, in that the support can take on negative values, is not suitable for the modelling of losses arising from catastrophes. Therefore, the GEV distribution is excluded from further analysis.

Chernobai *et al.* (2015) advocated the use of adjusted goodness of fit tests when using left-truncated data. These adjusted goodness of fit tests are given in Appendix B. Once again, the p-values and critical values are calculated using 1000 Monte Carlo simulations. The results of the standard and adjusted tests are given in Table 4.3 below.

Tab. 4.3: Critical values (above) and p-values (below) for the Kolmogorov-Smirnov (D), Kuiper (V), Anderson-Darling (A^2) and Cramer von-Mises (W^2) tests.

	Naïve approach				**Conditional approach**			
	D	V	A^2	W^2	D	V	A^2	W^2
Exponential	9.23	10.43	157.06	32.76	9.19	9.74	200.40	38.56
	< 0.005	< 0.005	< 0.005	< 0.005	< 0.005	< 0.005	< 0.005	< 0.005
Log-normal	2.19	3.64	7.76	1.19	1.39	2.15	2.94	0.49
	< 0.005	< 0.005	< 0.005	< 0.005	0.07	0.005	0.22	0.12
Burr	0.92	1.78	1.39	0.19	0.68	1.09	0.33	0.06
	0.01	0.02	< 0.005	< 0.005	0.50	0.36	0.83	0.72
Generalized Pareto	3.30	5.38	16.64	2.21	1.03	1.73	1.74	0.29
	< 0.005	< 0.005	< 0.005	< 0.005	0.03	0.03	0.13	0.04
Inverse Gaussian	3.29	4.13	14.40	2.78	2.37	3.04	7.95	1.42
	< 0.005	< 0.005	< 0.005	< 0.005	< 0.005	< 0.005	< 0.005	< 0.005

Firstly, none of the distributions considered fit the data well using the naïve approach, similar to the results of Chernobai *et al.* (2006). This remains the case for the exponential and inverse Gaussion distributions using the conditional approach and therefore they are excluded at this point. The Burr distribution appears to fit the loss data particularly well under the conditional fitting approach. The log-normal distribution passes the Kolmogorov-Smirnov, Anderson-Darling and Cramer-von-Mises tests at the 5% significance level while the generalised Pareto distribution passes only the Anderson Darling test, however in this case the p-values are close to 0.05 for the other three tests. Therefore the Burr, generalised Pareto and log-normal distributions are used for further analysis.

4.3 Intensity Fitting

In order to model the frequency, the time-dependent deterministic function $\lambda_d^o(t)$ proposed by Ma and Ma (2013) is considered together with a constant intensity λ_c^o. The deterministic function proposed is given by

$$\lambda_d^o(t) = a + b\sin^2(t+c) + d\exp\left(\cos\left(\frac{2\pi t}{w}\right)\right),$$

where $a > 0, d > 0$ and $w > 0$.

This intensity function was proposed based on the annual number of catastrophes. The sinusoidal term is used to capture the cyclical trend observed in the data while the exponential term accounts for a large increase in the intensity. The Non-Linear Least Squares procedure is performed on the annual number of catastrophes as recorded by the PCS index, to estimate the coefficients of $\lambda_d^o(t)$.

On the other hand, the constant intensity is estimated as follows

$$\lambda_c^o = \frac{\text{Number of events in } [0, n]}{n}$$

where n is the number of years over which the PCS data is observed.

As in Ma and Ma (2013), the following measures are used to assess the fit of the deterministic intensity: mean absolute error (MAE), root mean squared error (RMSE), Theil's coefficient (U), the Nash-Sutcliffe coefficient of efficiency (E) and the index of agreement (D). Their equations are given by

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|O_i - P_i|,$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - P_i)^2},$$

$$\text{U} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - P_i)^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}O_i^2} + \sqrt{\frac{1}{N}\sum_{i=1}^{N}P_i^2}},$$

$$\text{E} = 1 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O})^2} \text{ and}$$

$$\text{D} = 1 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(|P_i - \bar{O}| + |O_i - \bar{O}|)},$$

where $\bar{O} = \frac{1}{N}\sum_{i=1}^{N}O_i$, O_i are the observed number of catastrophes at time i, P_i are the predicted number of catastrophes at time i and N is the total number of observations.

A better fitting intensity will give a lower value for MAE, RMSE and U and a higher value for E and D (Ma and Ma, 2013). The estimated parameters for the deterministic and constant intensities are given in Table 4.4 and the calculated measures of fit are given in Table 4.5 below.

Tab. 4.4: Estimated parameters for constant and deterministic intensities.

Intensity	**Parameters**				
$\lambda_d^o(t)$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{d}$	$\hat{\omega}$
	24.45136	2.27945	-2.60845	1.44533	3.54381
λ_c^o	27.7383				

Tab. 4.5: Measures of Fit: the mean absolute error (MAE), the root mean square error (RMSE), Theil's coefficient (U), the index of agreement (d) and the Nash-Sutcliffe coefficient of efficiency (E).

Intensity	MAE	RMSE	U	D	E
$\lambda_d^o(t)$	4.45761	5.27878	0.09504	0.37080	0.069703
λ_c^o	4.65697	5.47899	0.09830	0.07367	-0.002202

The period of the cosine term suggests that a large increase in the intensity is expected about every 3.54 years, which does not seem unreasonable and is also sufficiently conservative.

The MAE, RMSE and U statistics are higher for the constant intensity while the D and E statistics are lower. Therefore, the measures of fit suggest that the deterministic time-dependent intensity provides a better fit than the constant intensity considered, as would be expected for a reasonably specified intensity function.

When taking into account the left-truncation, it is important at this point to distinguish between the observed and the complete number of events denoted with a superscript o and c respectively. Given that losses are only observed above H, the observed intensity $\lambda^o(t)$ is given by $(1 - F_{\gamma^c}(H))\lambda^c(t)$, where $\lambda^o(t)$ can be either deterministic or constant. Therefore, the estimated parameters $\hat{\gamma}^c$ from Section 4.2 can be used to estimate the complete intensity function to account for the left-truncation in the following manner (Chernobai *et al.*, 2006),

$$\hat{\lambda}^c(t) = \frac{\hat{\lambda}^o(t)}{1 - F_{\hat{\gamma}^c}(H)}. \tag{4.3}$$

where $\hat{\gamma}^c$ is estimated by MLE or MPS where applicable.

It is noted by Chernobai *et al.* (2006) that the fraction of missing data below the threshold H is "added back" by the above conditional fitting approach, which modifies both the fitted parameters and the intensity, and makes an important assumption in that the distribution fitted is the true, underlying distribution.

Going forward, the 'conditional approach' in the context of CAT bond pricing will refer to using the conditionally fitted parameters for the loss distribution (by Equation (4.1) and (4.2)) and scaling up the intensity (see Equation (4.3)) in the manner described above.

Chapter 5

Comparison of Two Approximation Methods

The purpose of this chapter is to evaluate two methods for approximating the real-world probabilistic expressions involving L_T or L_{t_k} for $k = 1, \ldots, m$, to ultimately calculate CAT bond prices. In this chapter, the pricing formulae of the ZC and CP CAT bonds are firstly implemented by Monte Carlo simulation (MCS) for a range of threshold levels and maturity dates. These results are then compared with the mixed-approximation method of Ma and Ma (2013). The same range of threshold levels and terms-to-maturity are considered as in Ma and Ma (2013) and are given by $T = [0.25, 0.5, ..., 2.5]$ and $D \in [\$3.74\,\text{billion}, \$44.84\,\text{billion}]$: the threshold level ranges from the average quarterly loss to three times the average annual loss (Ma and Ma, 2013). Given that there are no closed-form solutions for the CAT bond pricing formulae in Chapter 3, these comparisons will be used to determine whether the Monte Carlo approach gives similar results to the mixed-approximation method, where both approaches assume that the catastrophe risk process follows a compound Poisson process. These comparisons will be made both in the cases of constant and deterministic intensities.

5.1 Monte Carlo Simulation

Assume that the ZC or CP CAT bond is to be valued at $t = 0$. For the ZC CAT bond, the only random variable of interest is the aggregate loss process's value at maturity, i.e. L_T. For the CP CAT bond, simulated values of the aggregate loss process are required on each coupon-paying date, i.e. L_{t_k} for $t_1 = 0.25, \ldots, t_m = T$. The method for simulating the aggregate loss process over the interval $[0, T]$ is described as follows. Firstly, simulate a random number of events N in the specified time interval $[0, T]$. Then, given $N = n$, simulate n losses from the specified distribution for X. This relies on the independence assumption of the loss process X

and the arrival process N. Finally, summing the values of X_i that exceed \$25 million provides a realisation of the aggregate loss process L at time T for the ZC CAT bond or at time t_k for $k = 1, \ldots, m$ for the CP CAT bond.

Where the aggregate loss process follows a homogeneous compound Poisson process with constant intensity λ_c, simulating N simply involves generating a Poisson random variable with intensity $\lambda_c T$. Note that for the CP CAT bond, it is necessary to simulate N_i for $i = 1, \ldots, m_T$, where m_t is the number of quarters in $[0, t]$ and one has that

$$\sum_{i=1}^{m_T} N_i = N.$$

The values of L at each coupon date are given by

$$L_t = \sum_{i=1}^{M_t} X_i$$

where $M_t = \sum_{i=1}^{m_t} N_i$. Where the aggregate loss process follows a non-homogeneous compound Poisson process with deterministic intensity $\lambda_d(t)$, for both the ZC and CP CAT bond, the thinning algorithm is used. This is implemented using the algorithm described in Burnecki and Weron (2005). Once a sufficient number of realisations of L_{t_k} are simulated for the required term-to-maturity and/or coupon paying dates, the probabilistic expressions given in Equations (3.2) and (3.5) can be evaluated for any threshold level. For example,

$$\mathbb{P}(L_t \leq D) \approx \frac{\text{Number of simulated values of } L_t \leq D}{\text{Total number of simulations}}.$$

5.2 Mixed-Approximation Method

The probabilistic expression in Equation (3.2), for the ZC CAT bond, can be expressed as follows (as given in Ma and Ma (2013)):

$$\begin{aligned} V_t &= B(t,T)Z[\rho + (1-\rho)\mathbb{P}(L_T \leq D)] \\ &= B(t,T)Z[\rho + (1-\rho)\sum_{n=0}^{\infty} p_{nT} F_n^*(D)] \end{aligned} \tag{5.1}$$

where $p_{ns} = e^{-\int_t^s \lambda_u du} \frac{(\int_t^s \lambda_u du)^n}{n!}$ and $F_n^*(D) = \mathbb{P}(X_1 + X_2 + \ldots + X_n \leq D)$, the latter being the n-fold convolution of F.

Similarly, the price of the CP CAT bond can be expressed as follows:

$$V_t = \sum_{k=1}^{m-1} B(t,t_k)C[\rho + (1-\rho)\mathbb{P}(L_{t_k} \leq D)] + B(t,T)(Z+C)[\rho + (1-\rho)\mathbb{P}(L_T \leq D)]$$
$$= \sum_{k=1}^{m-1} B(t,t_k)C[\rho + (1-\rho)\sum_{n=0}^{\infty} p_{nt_k} F_n^*(D)] + B(t,T)(Z+C)[\rho + (1-\rho)\sum_{n=0}^{\infty} p_{nT} F_n^*(D)]. \tag{5.2}$$

Given that the aggregate loss process is commonly used in actuarial science, there has been much research into approximating the above convolution using an approximating distribution (Ma and Ma, 2013). These approximation methods include, amongst others, the Edgeworth method and the Esscher approximation (Reijnen *et al.*, 2005).

The basic idea is to use a well-known distributional form to approximate expressions of the form given in Equations (5.1) and (5.2). The parameters are defined in terms of the mean, variance, skewness and/or kurtosis of the loss process X and/or aggregate loss process L (Reijnen *et al.*, 2005).

Chaubey *et al.* (1998) used a mixed-approximation method which includes both the gamma and inverse Gaussian (IG) distributions. This was implemented in the context of CAT bond pricing by Ma and Ma (2013).

Let $\mu_L, \sigma_L, \kappa_{3L}$ and κ_{4L} denote the mean, standard deviation, skewness and excess kurtosis of the aggregate loss process and let κ_{3X} denote the skewness of the loss distribution. The parameterisation of the gamma approximating distribution is given by

$$f_L(l) \approx f_{\text{gamma}}(l) = \frac{\beta^\alpha (l-x_0)^{\alpha-1} e^{-\beta(l-x_0)}}{\Gamma(\alpha)},$$

for $l > x_0$, where $\alpha = \left(\frac{2}{\kappa_{3L}}\right)^2, \beta = \frac{2}{\kappa_{3L}\sigma_L}, x_0 = \mu_L - \frac{2\sigma_L}{\kappa_{3L}}$.

Similarly, the parameterisation of the inverse Gaussian approximating distribution is given by

$$f_L(l) \approx f_{\text{IG}}(l) = \frac{\alpha}{\sqrt{2\pi\beta(l-x_0)^3}} \exp\left(-\frac{(\alpha - \beta(l-x_0))^2}{2\beta(l-x_0)}\right),$$

for $l > x_0$, where $\alpha = \left(\frac{3}{\kappa_{3L}}\right)^2, \beta = \frac{3}{\kappa_{3L}\sigma_L}, x_0 = \mu_L - \frac{3\sigma_L}{\kappa_{3L}}$,

and that of the the gamma-IG approximating distribution is given by

$$f_L(l) \approx f_{\text{gamma-IG}}(l) = w f_{\text{gamma}}(l) + (1-w) f_{\text{IG}}(l),$$

where $w = \frac{\kappa_{4L} - \kappa_{4\text{IG}}}{\kappa_{4\text{gamma}} - \kappa_{4\text{IG}}} = \frac{\kappa_{4L} - 5\kappa_{3L}^2/3}{-\kappa_{3L}^2/6} = \frac{10\kappa_{3L}^2 - 6\kappa_{4L}}{\kappa_{3L}^2}$.

The following rule of thumb was used by Reijnen *et al.* (2005),

$$\text{Approximation method} = \begin{cases} \text{gamma-IG approximation} & \text{if } 0 \leq \kappa_{3X} \leq 5 \text{ and } 0 \leq \kappa_{4L} \leq 1.5 \\ \text{IG approximation} & \text{if } 5 < \kappa_{3X} < 15 \text{ or } 1.5 < \kappa_{4L} < 50. \end{cases}$$

Reijnen *et al.* (2005) found that this rule of thumb performed better than using only a single approximating approach. Therefore, this rule of thumb was used by Ma and Ma (2013) and is consequently used in this dissertation. Therefore, the probabilistic expressions can be approximated using

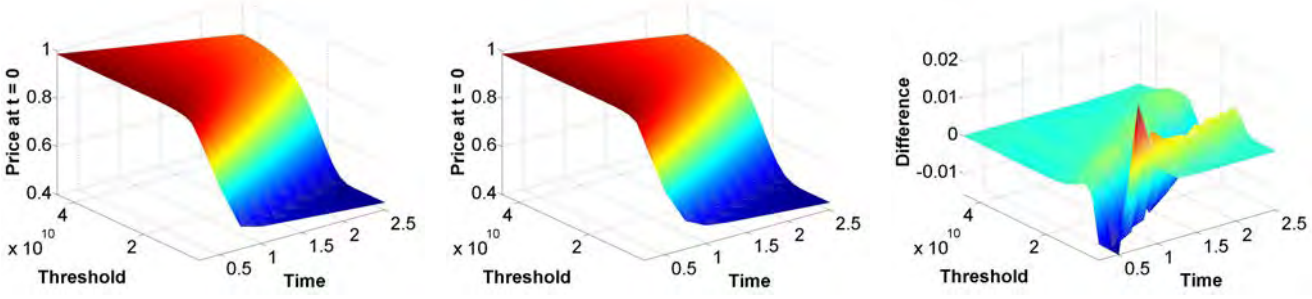
$$\mathbb{P}(L_t \leq D) \approx \begin{cases} F_{\text{gamma-IG}}(D) = wF_{\text{gamma}}(D) + (1-w)F_{\text{IG}}(D) & \text{if } 0 \leq \kappa_{3X} \leq 5 \text{ and } 0 \leq \kappa_{4L} \leq 1.5 \\ F_{\text{IG}}(D) & \text{if } 5 < \kappa_{3X} < 15 \text{ or } 1.5 < \kappa_{4L} < 50. \end{cases}$$

5.3 Results of Comparison

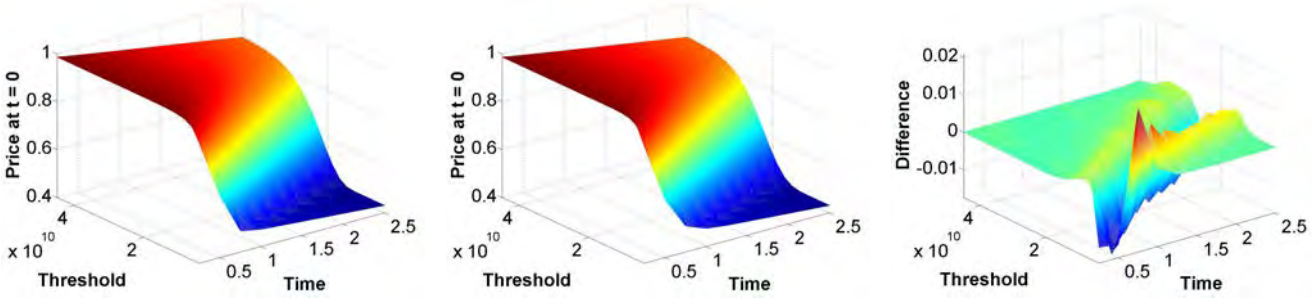
It is important to note that the mixed-approximation method, derived by Chaubey *et al.* (1998), requires the calculation of the kurtosis of the aggregate loss process L. This in turn requires the existence of the first four moments for the loss distribution of X. The use of heavy-tailed distributions, however, places certain restrictions on the parameters for higher moments such as the kurtosis to exist. This is problematic in the case of the generalised Pareto and Burr distributions.

For the generalised Pareto distribution, the skewness is finite when $k < \frac{1}{3}$ and the kurtosis is finite when $k < \frac{1}{2}$. For the Burr distribution, $\mathbb{E}[X^q]$ is defined only for $q < ck$ (Klugman *et al.*, 2004).

Upon inspection of the parameters estimated in Section 4.2., it is clear that the parameters do not satisfy the conditions required for these higher moments to exist. Therefore, the mixed-approximation method can only be implemented in the case of the log-normal distribution. Consequently, the results from only this distribution are used to test whether the Monte Carlo approach gives similar results to the mixed-approximation method, both approaches using the assumption that the catastrophe risk process follows a compound Poisson process. The pricing surfaces for both the Monte Carlo and the mixed-approximation approaches are given in Figures 5.1 and 5.2 below.

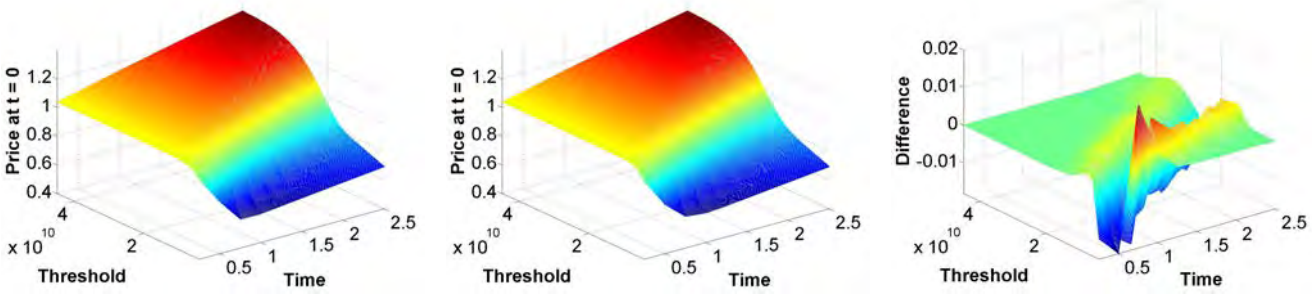


(a) Mixed-approx., Const. intensity. **(b)** Monte Carlo, Const. intensity. **(c)** Difference.

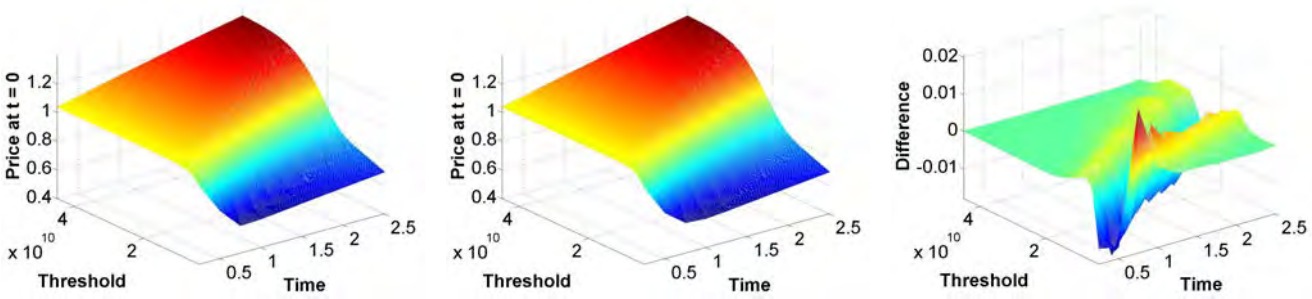


(d) Mixed-approx., Det. intensity. **(e)** Monte Carlo, Det. intensity. **(f)** Difference.

Fig. 5.1: Comparison of mixed-approximation method and Monte Carlo simulation for ZC CAT bonds. (Det. ≡ deterministic; Const. ≡ constant).



(a) Mixed-approx., Const. intensity. **(b)** Monte Carlo, Const. intensity. **(c)** Difference.



(d) Mixed-approx., Det. intensity. **(e)** Monte Carlo, Det. intensity. **(f)** Difference.

Fig. 5.2: Comparison of mixed-approximation method and Monte Carlo simulation for CP CAT bonds. (Det. ≡ deterministic; Const. ≡ constant).

The results above seem to suggest that the two approximation methods are largely consistent. In both cases, the CAT bond pricing surfaces behave as expected. For the ZC CAT bond, the price increases with threshold level and decreases with term-to-maturity, as one would expect. The highest price occurs where the threshold level is at its highest and the term-to-maturity is at its lowest and has a value approximately equal to $e^{-0.06(0.25)} = 0.98511$, since there is a very low probability of CAT bond triggering. Similarly, the lowest price occurs where the threshold level is at its lowest and the term-to-maturity at its highest. For the CP CAT bond, there is now an additional benefit of holding the CAT bond for longer which partially offsets the higher probability of the CAT bond triggering for higher terms-to-maturity. Therefore, the CP CAT bond pricing surface tends to increase with term-to-maturity. Moreover, using the Monte Carlo approach as the benchmark, the mixed-approximation method appears to perform well where the CAT bond pricing surface is relatively flat, and poorly where there is a larger change in the gradient. This is evident for both CAT bond structures.

Investigating further, in the case of the ZC CAT bond the approximation methods are mainly consistent for the upper quantiles of the distribution of L_T. In the pricing surfaces above, the "upper quantiles of the distribution of L_T" refers to the triangular region specified by the following points: the point corresponding to the lowest threshold level and lowest term-to-maturity, the point corresponding to the highest threshold level and lowest term-to-maturity and the point corresponding to the highest threshold level and highest term-to-maturity. The differences are close to zero in this region of the graphs above for the ZC CAT bond, both for the constant and deterministic intensities. For the CB CAT bond, the interpretation is less obvious. However, the event that $\{L_T > D\} \subset \{L_{t_k} > D\}$ for $k = 1, \ldots, m-1$, as the aggregate loss process is monotonically increasing. Therefore, it is reasonable to interpret the CP CAT bond in terms of the quantiles of L_T as well. Similarly, for the CP CAT bond, the differences are close to zero in this region. The results above suggest that the mixed-approximation method produces quite consistent results with the MCS approach in the upper quantiles of L_T, at least in the case of the log-normal distribution. The mixed-approximation method has the obvious advantage of being computationally efficient. The major disadvantage of this method, however, is the non-existence of higher moments for very heavy-tailed distributions, including the Burr and GP distributions. Therefore, MCS is used for the remainder of this dissertation.

Chapter 6

Naïve vs. Conditional Fitting Approach

Chernobai *et al.* (2006) noted that failing to account for the left-truncation in the distribution fitting will result in an overestimated mean, underestimated variance and an underestimation of the upper quantiles of the distribution in the context of the PCS loss data. Therefore, it is expected that CAT bond prices will differ when accounting for the left-truncation in the distribution fitting methodology. In this section, pricing surfaces are produced based on the naïve and conditional fitting approaches, for the log-normal, GP and Burr distributions using the deterministic intensity function from Chapter 4 in each case.

Now, in order to compare these two approaches, an alternative pricing approach referred to as the "non-parametric" approach is introduced. Under this approach realisations of the aggregate loss process, and therefore the pricing surfaces, are obtained based on the empirical distribution function (EDF) of both the loss estimates and the waiting times between catastrophes. This method is described in Section 6.1.

6.1 Non-parametric Approach

The EDF of the loss amounts X is defined as follows:

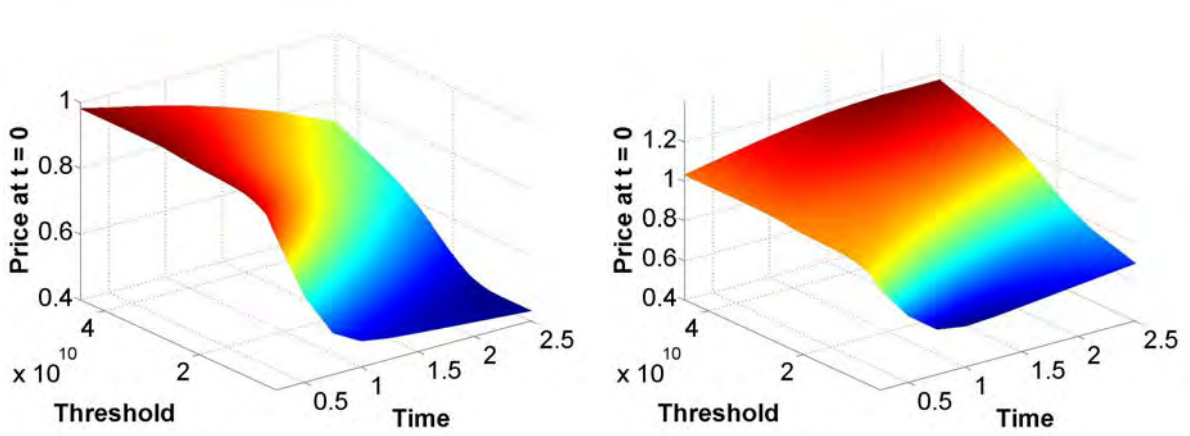
$$F_X(x) = \begin{cases} 0 & \text{if } x \leq x_{(1)} \\ \frac{1}{n} & \text{if } x_{(1)} < x \leq x_{(2)} \\ \frac{2}{n} & \text{if } x_{(2)} < x \leq x_{(3)} \\ \vdots & \\ 1 & \text{if } x > x_{(n)}. \end{cases}$$

This is similarly defined for the waiting times W between catastrophes;

$$F_W(w) = \begin{cases} 0 & \text{if } w \leq w_{(1)} \\ \frac{1}{n} & \text{if } w_{(1)} < w \leq w_{(2)} \\ \frac{2}{n} & \text{if } w_{(2)} < w \leq w_{(3)} \\ \vdots & \\ 1 & \text{if } w > w_{(n)}, \end{cases}$$

where $x_{(1)}, x_{(2)}, ..., x_{(n)}$ and $w_{(1)}, w_{(2)}, ..., w_{(n)}$ are the order statistics of the n observed losses and waiting times.

Generating a sample Y from the EDF of either X or W is relatively straightforward. Simulate $V \sim U(0,1)$ and set $Y = x_{(i)}$ or $Y = w_{(i)}$ respectively if $\frac{i-1}{n} < V \leq \frac{i}{n}$. The aim is to produce CAT bond pricing surfaces using these EDFs to compare the naïve and conditional approaches. The implementation of the non-parametric simulation is similar to the previously described Monte Carlo simulation methods, where firstly waiting times are simulated until their sum is less than or equal to the term-to-maturity and, thereafter, the resulting number of events is recorded. Then, given this number of events, the corresponding losses are simulated. In this case, all the simulated losses will be in excess of $25 million. The motivation for using this method is that no distributional assumptions are made, both in the case of loss severity and intensity. However, a downfall of this fitting methodology is that the losses are very specific to the data observed and an implicit assumption is made in that the loss amounts in the future will not differ from those observed in the sample. The results for 20 000 simulations are given in Figure 6.1 below.



(a) ZC CAT bond. **(b)** CP CAT bond.

Fig. 6.1: Pricing surfaces based on the non-parametric approach.

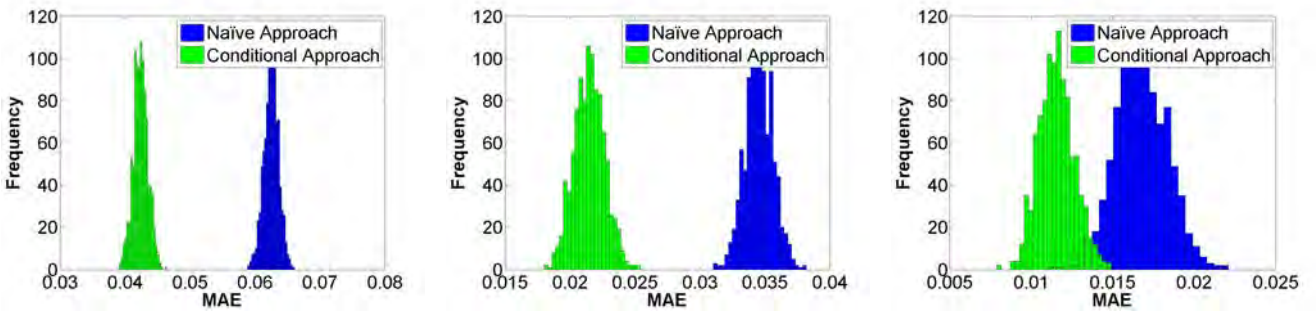
6.2 Comparison with Non-parametric Approach

A comparison can now be made between the resulting 'empirical pricing surface' and those produced using the parameter estimates from the naïve and conditional approaches. Note that this approach does not displace the formal tests in determining the appropriate distribution. Also, it must be noted that the non-parametric approach is more appropriate when working with large datasets which is not the case for extreme events (Burnecki *et al.*, 2010). This comment applies to the PCS data set, since this data set is small. In this context, the non-parametric approach is used simply to compare the naïve and conditional approaches for the log-normal, GP and Burr distributions. It serves purely as a 'robustness' check. Intuitively, one would expect a pricing surface simulated with parameters fitted using a more correct estimation procedure to lie closer the 'empirical pricing surface' within an in-sample context. To measure the absolute differences in these pricing surfaces, a MAE is used where the summation is taken over all the prices that make up the grid used to construct the pricing surfaces in the figures. The MAE is calculated by using the following formula:

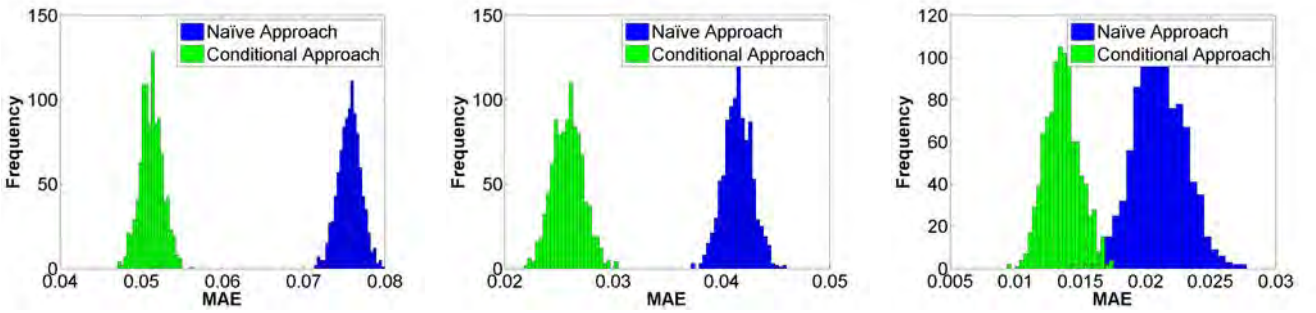
$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|V_{0i}^{MCS} - V_{0i}^{NPS}|$$

where V_{0i}^{MCS} is the i^{th} CAT bond price simulated by the normal Monte Carlo approach, for either the naïve or conditional approaches, V_{0i}^{NPS} is the i^{th} CAT bond price simulated by the non-parametric approach and the summation is taken over all CAT bond prices on the grid of threshold levels and terms-to-maturity.

Of course, the pricing surfaces in all these cases are produced by Monte Carlo simulation and it would not be fair to draw conclusions based on a single observation of the MAE. Therefore, 1000 pricing surfaces are simulated with 1000 observations each for the non-parametric approach and for the naïve and conditional approaches, for all three distributions. Based on these surfaces, 1000 MAEs are calculated for the naïve and conditional approaches. The histograms of the naïve and conditional MAEs are superimposed and compared for the three distributions, as shown in Figure 6.2 below.



(a) ZC CAT bond, Log-normal. **(b)** ZC CAT bond, GP. **(c)** ZC CAT bond, Burr.



(d) CP CAT bond, Log-normal. **(e)** CP CAT bond, GP. **(f)** CP CAT bond, Burr.

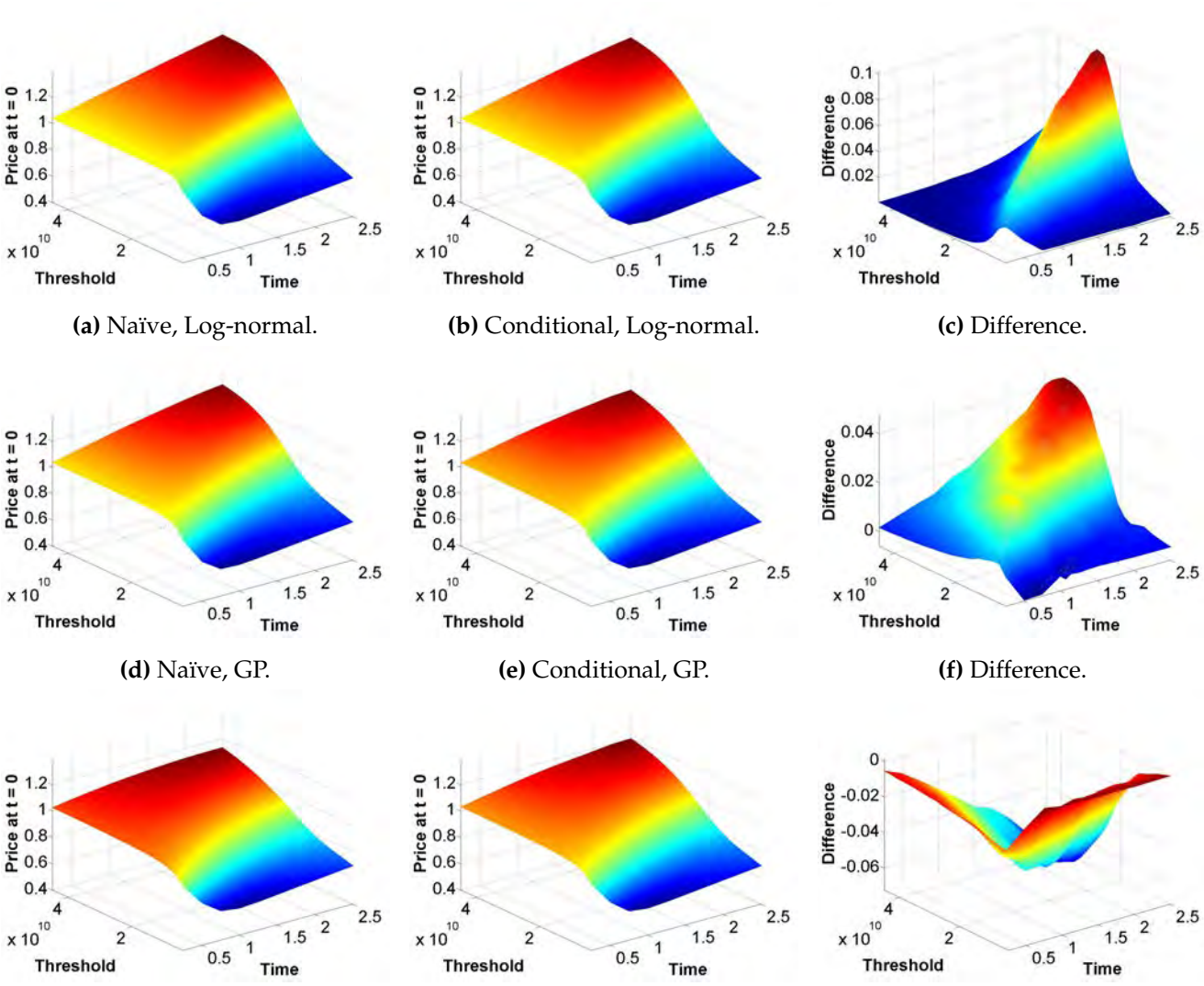
Fig. 6.2: MAE's for the naïve and conditional approaches.

In all three cases, the MAE is lower for the conditional approach compared to the naïve approach. Therefore, the results seem to suggest that the conditional approach improves the fit of the model in sample.

Focusing only on the intensity, intuitively the naïve approach is expected to underestimate the number of catastrophic events. The intensity parameters are estimated using the observed data, which only includes catastrophic events in the interval $[H, \infty)$. However, under the naïve approach, the fitted intensity is used to simulate the number of catastrophic events over the entire positive real line. There is therefore a reduction in the number of events simulated in $[H, \infty)$ for the naïve approach; the naïve approach does not take into account the threshold. The non-parametric approach only simulates losses in $[H, \infty)$, as the observed losses are only in this interval. Therefore, the simulated number of catastrophes under the naïve approach will typically be lower than those simulated using the non-parametric approach. The conditional approach accounts for this by scaling up the observed intensity.

Now, ignoring the threshold is expected to underestimate the upper quantiles of the distribution. Therefore, given that CAT bond prices are typically based on the upper quantiles of the aggregate loss process, it is expected that in general the

CAT bond prices generated using the naïve approach will be greater than those generated using the conditional approach. Comparisons between naïve and conditional CP CAT bond pricing surfaces for the three distributions are shown in Figure 6.3 below, where each pricing surface is produced using 20 000 simulations. Similar comparisons for the ZC CAT bond are given in Appendix D. The log-normal and GP cases seem to behave as expected; the conditional pricing surface tends to lie below the naïve pricing surface. However, for the Burr distribution the opposite is observed.



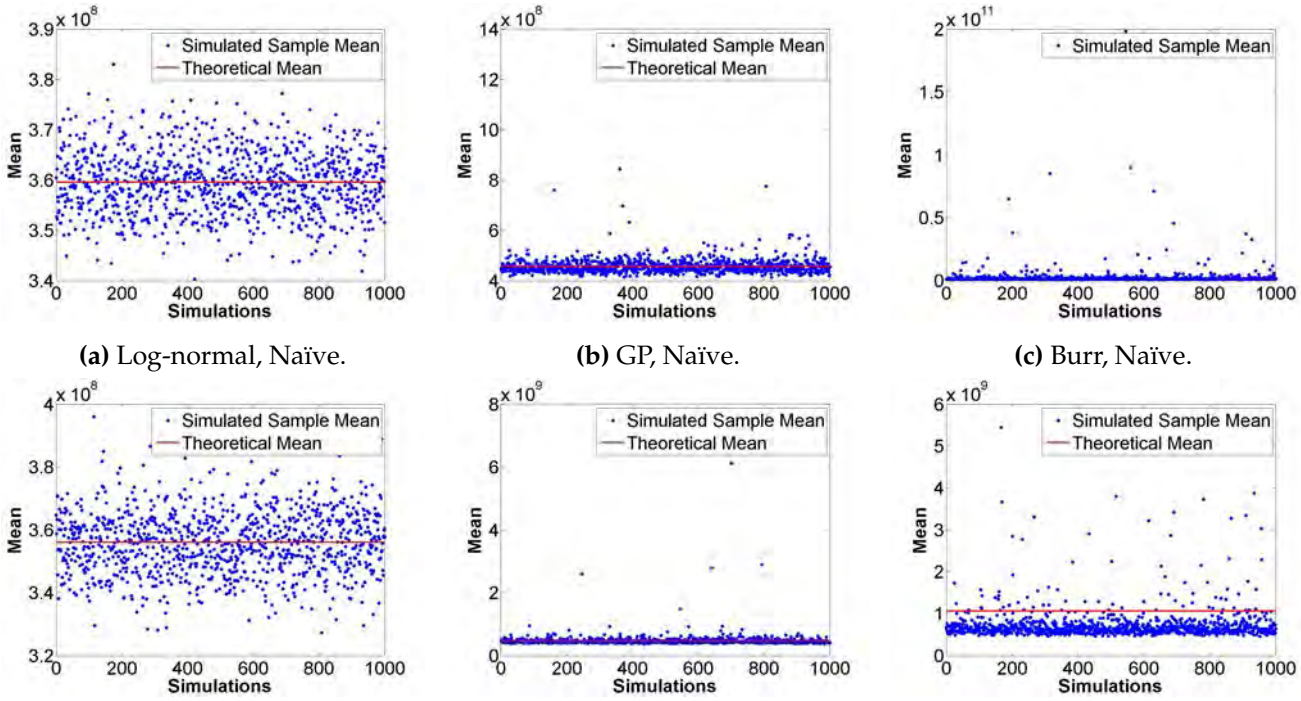
(a) Naïve, Log-normal. **(b)** Conditional, Log-normal. **(c)** Difference.

(d) Naïve, GP. **(e)** Conditional, GP. **(f)** Difference.

(g) Naïve, Burr. **(h)** Conditional, Burr. **(i)** Difference.

Fig. 6.3: CP CAT bond pricing surfaces for the naïve and conditional approaches and their difference, for the three fitted distributions.

This anomaly involving the Burr distribution highlights a very important issue: one needs to be particularly careful when dealing with very heavy-tailed distributions. To further illustrate this point, 1000 samples of size 10 000 are simulated from the three distributions and the sample means are calculated and plotted. In addition, the theoretical means are calculated and included in the plots. Note that for the Burr distribution fitted using the naïve approach, the theoretical mean is infinite and only the sample means are plotted. Indeed, Chernobai *et al.* (2006) excluded the Burr distribution from further analysis as the infinite first moment is inappropriate for calculating ruin probabilities.



(a) Log-normal, Naïve. **(b)** GP, Naïve. **(c)** Burr, Naïve.

(d) Log-normal, Conditional. **(e)** GP, Conditional. **(f)** Burr, Conditional.

Fig. 6.4: Theoretical means and simulated sample means for the three distributions, for both the naïve and conditionally fitted distributions.

As seen in Figure 6.4 above, for the log-normal distribution, and to a certain extent the generalised Pareto distribution, the sample means centre around the theoretical mean as one would expect. However, for the Burr distribution the sample mean tends to be markedly lower than the theoretical mean, due to the very heavy right tail. Burnecki *et al.* (2010) pointed out that this can occur when the product of the Burr parameters ck is greater than but very close to one. In this case, ck is approximately 1.09.

A possible explanation for the sample mean problem is the difficulty in sam-

pling from an extremely heavy-tailed distribution. The very heavy right tail of the fitted Burr distribution may require a very fine random number generator to adequately sample from the extreme right-hand tail of the distribution.

Note that, in an attempt to account for this issue, an adjustment was made to the simulated Burr random variables, in the conditional case. The difference between the the theoretical mean and the sample mean, based on 1 000 000 simulated Burr observations, was added to each simulated Burr random variable. This produced the expected behaviour, with the unconditional pricing surface lying above the conditional one as in shown in Figure 6.5 below.

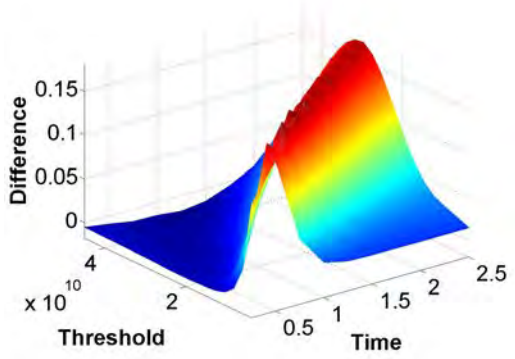


Fig. 6.5: Difference between naïve and conditional CP CAT bond pricing surfaces after the adjustment.

However, this would be expected given that this effectively shifts down the conditional pricing surface. An alternative method is loosely based by the method of stratified sampling. This alternative method is described as follows. Let C denote an arbitrary cut-off level. Firstly, simulate $U \sim U(0,1)$ and $V \sim U(0,1)$, independent of U. If U lies between 0 and C, set U_{new} to be UV so that U_{new} lies between 0 and U. If U lies between C and 1, set U_{new} to be $U + (1-U)V$ so that U_{new} lies between U and 1. Finally, perform the probability integral transform, using the inverse cumulative distribution function of the estimated Burr distribution, on U_{new}. The method attempts to more accurately sample from the tails of the estimated Burr distribution, particularly focusing on the right tail. Therefore, the cut-off level is typically chosen to be very close to 1 (e.g. $C = 0.9$). The problem here is, firstly, that the choice of cut-off level is rather subjective. A second point is that the method is very sensitive to the choice of cut-off level and only produces consistent results for certain choices of cut-off level (e.g. $C = 0.9$) and not for others (e.g. $C = 0.999$). Therefore, the method does not appear to be robust.

In sum the above issue with the Burr distribution raises a broader problem of the simulation from very heavy-tailed distributions, and would make room for further research.

Overall, the results of this chapter seem to suggest that the conditional approach should be adopted to account for the left-truncation in the PCS data set. Note that this also applies when implementing the mixed-approximation method, by using the parameters estimated using the conditional fitting approach and scaling up the moments of the (non)-homogeneous Poisson process. Furthermore, the results bring to light the problems associated with simulating from very heavy-tailed distributions which have not been given much attention in previous literature.

Chapter 7

Pricing Using a Stochastic Intensity

Another potential issue arising in this dissertation is the use of a deterministic intensity function in the modelling of the aggregate loss process. Dassios and Jang (2003) suggested that any realistic model for the aggregate loss process must have a stochastic intensity component. This chapter considers three important issues regarding the use of a stochastic intensity model, and these issues have not been fully addressed in previous literature on the matter. Firstly, it attempts to address the estimation issue discussed in Chapter 2 and therefore looks at how the parameters of this model can be estimated from the data available. Secondly, given the estimated parameters, it proposes a method for simulating the number of catastrophic events using specified stochastic intensity model. These two steps are implemented in the context of the log-normal intensity proposed by Lin *et al.* (2009). Finally, it proposes an alternative model for the stochastic intensity to be used for the aggregate loss process.

As mentioned in Chapter 2, Lin *et al.* (2009) advocated the use of a log-normal model to describe the intensity. The stochastic differential equation (SDE) for this model, commonly referred to as Geometric Brownian Motion, is given by

$$d\lambda_t = \lambda_t(\mu dt + \sigma dW_t) \tag{7.1}$$

where $\{W_t : t \in \mathbb{T}\}$ is a standard Brownian motion. A possible justification for the use of a log-normal intensity model is the fact that CAT bonds are typically issued over a short period. Hence, although the log-normal model would be expected to explode, this would be less likely over shorter time periods.

7.1 Estimation of Stochastic Intensity Parameters

In this chapter, the stochastic models considered are estimated using maximum likelihood, based on the transition density of the relevant model. The quarterly number of catastrophic events are used, where the estimates of the intensity are calculated as the (annualised) number of events in each quarter. By calculating the estimated intensities in this manner, the implicit assumption is that the intensity in constant over each quarter.

Applying Ito's Formula to Equation (7.1) gives

$$\log\left(\frac{\lambda_t}{\lambda_{t-\frac{1}{4}}}\right) = \frac{1}{4}\left(\mu - \frac{1}{2}\sigma^2\right) + \sigma\sqrt{\frac{1}{4}}Z$$

where $Z \sim N(0,1)$.

In this case, the maximum-likelihood estimates of the log-normal model are simply calculated as the mean and standard deviation of the log-changes of observed quarterly intensities (Phillips and Yu, 2009).

The results of this exercise are given in Table 7.1 below.

Tab. 7.1: Maximum-likelihood estimates for the log-normal intensity model.

Model	$\hat{\mu}$	$\hat{\sigma}$
Log-normal	1.08938	1.47606

The appropriateness of these estimated parameters will be discussed in the next section.

7.2 Simulation

Theoretically speaking, given the definition of a doubly stochastic Poisson process, the number of catastrophes could be simulated using the thinning algorithm based on a realised path of the intensity process. However, to be consistent with the manner in which the intensity estimates are obtained, the simulation procedure for the number of catastrophes during each period is performed as described below. As before, m_T denotes the number of quarters in $[0, T]$. Now, let $\hat{\lambda}_j$ denote the simulated intensity for quarter j for $j = 1, \ldots, m_T$, $\hat{N}_j$ denote the simulated number of catastrophes in quarter j for $j = 1, \ldots, m_T$, $P(\lambda)$ denote a Poisson distribution with parameter λ and λ_0 denote the most recently observed estimate for the intensity. Firstly, simulate a realisation of the intensity $\hat{\lambda}_j$ for each quarter until the term-to-maturity based on the estimated stochastic model, starting from the

most recently observed intensity λ_0. Then, given the realisation of the intensity in each quarter, the number of events in that quarter $\hat{N}_j$ is simulated by generating a Poisson random variable $P(\hat{\lambda}_j)$ with the realised intensity for that quarter. This is consistent with the implicit assumption made in the estimation procedure. The interval of estimation, and therefore simulation, could be reduced in size to better approximate the underlying stochastic process for the intensity. However, a smaller interval length would result in a larger number of observations of intervals with no catastrophes, which would undermine the use of a non-negative intensity model. Figure 7.1 illustrates the simulation procedure.

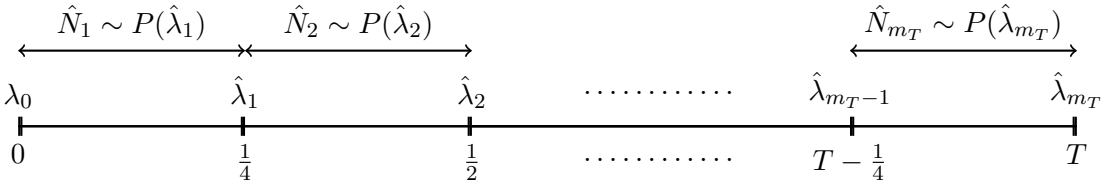


Fig. 7.1: Diagram illustrating the stochastic intensity simulation procedure.

As a reasonability check, catastrophes are simulated over a 2.5 year period based on realisations of the log-normal intensity process. However, as seen in Figure 7.2 below, even over a short period the number of catastrophes seems unreasonably high. Therefore, this model is deemed to be unsatisfactory. Note that the Black-Derman-Toy model also leads to log-normal dynamics for the intensity, so one would expect it to produce similar behaviour.

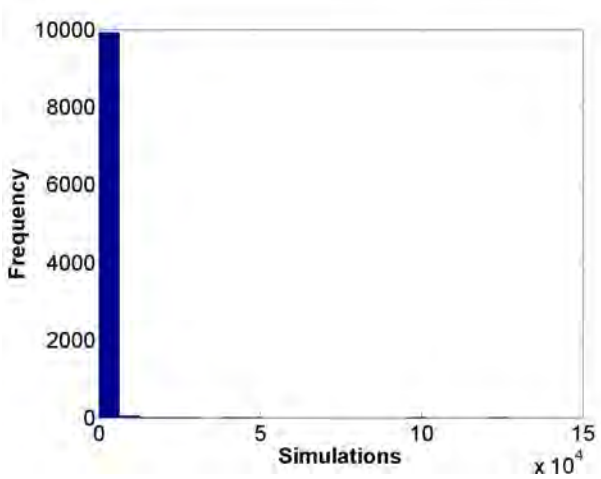


Fig. 7.2: Simulated number of catastrophic events over 2.5 years using the log-normal intensity model.

7.3 Cox-Ingersoll-Ross process

An examination of the quarterly number of events shows that the data appears to exhibit mean-reverting behaviour.

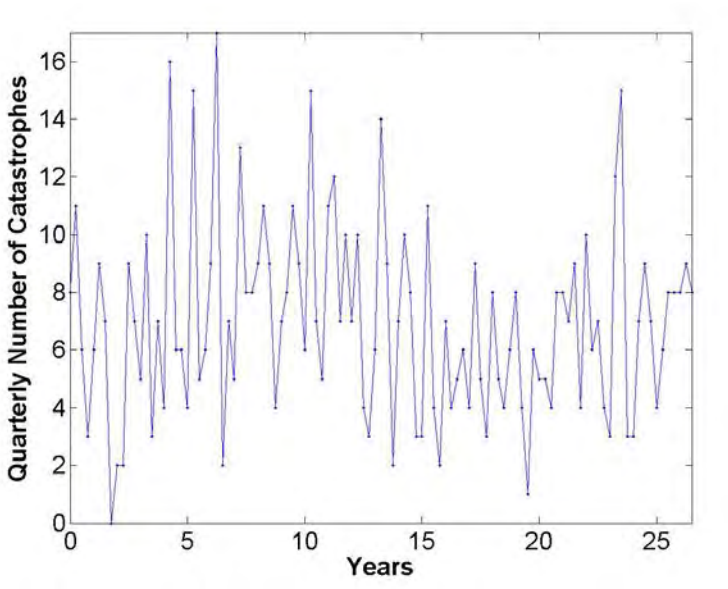


Fig. 7.3: Quarterly number of catastrophic events from the PCS data set.

Therefore, on consideration of Figures 7.2 and 7.3, the Cox-Ingersoll-Ross (CIR) model (Cox *et al.*, 1985) is proposed to capture the apparent mean-reverting behaviour, whilst maintaining non-negativity of the intensity process. The SDE for the Cox-Ingersoll-Ross process is given by

$$\lambda_t = \alpha(\mu - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t \tag{7.2}$$

where $\alpha, \mu, \sigma > 0$, $2\alpha\mu \geq \sigma^2$ and W_t is a standard Brownian motion.

A possible criticism of the use of the Cox-Ingersoll-Ross model is the fact that it does not incorporate an increasing trend. An argument against this is the fact that CAT bonds' terms-to-maturity are typically very short and the observed increasing trend is very gradual over long periods of time. Maximum likelihood estimation is used to estimate the parameters of the Cox-Ingersoll-Ross model using the method

of and code provided by Kladıvko (2007)[4]. This method was developed in the context of estimating the Cox-Ingersoll-Ross model parameters by maximum likelihood estimation based on an observed time series of interest rates. It is applied in this context to an observed time series of 'estimated intensities'. For further details, the estimation procedure of Kladıvko (2007) is outlined in Appendix E. The estimated parameters are given in Table 7.2.

Tab. 7.2: Maximum-Likelihood estimates for the CIR intensity model.

Model	$\hat{\alpha}$	$\hat{\mu}$	$\hat{\sigma}$
CIR	51.48610	27.96281	26.23422

Firstly, note that $2\hat{\alpha}\hat{\mu} = 2879.39206 \geq \hat{\sigma}^2 = 688.23430$ so the model is well-defined. Secondly, the estimated mean-reverting level $\hat{\mu}$ is very close to the constant intensity calculated in Chapter 4. Furthermore, simulating the number of catastrophes over a 2.5 year period, as shown in Figure 7.4, shows that the estimation procedure of Kladıvko (2007) appears to give reasonable results for the model. As a comparison, based on the observed data, the average number of catastrophes per 2.5 years was 69.3458. The simulations were performed using the procedure given by Andersen *et al.* (2010), which is also briefly illustrated in Appendix E.

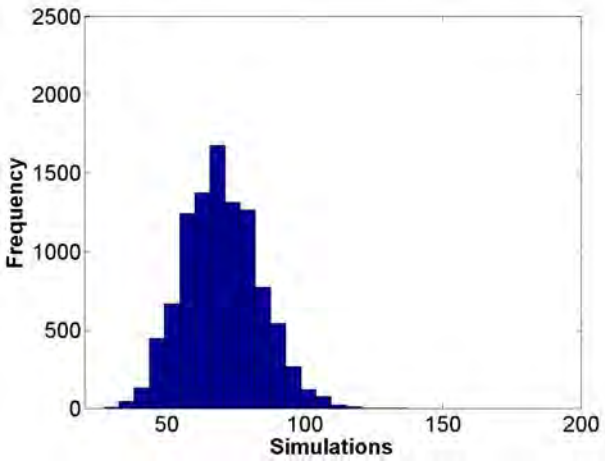
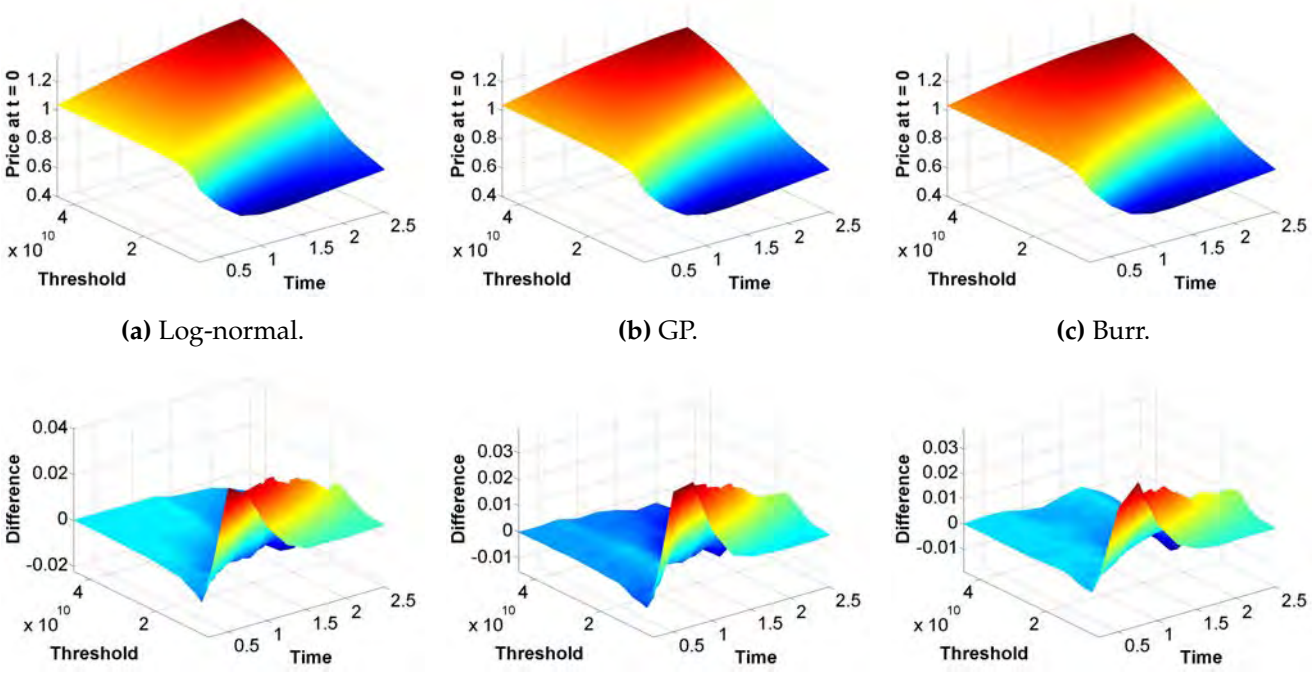


Fig. 7.4: Simulated number of catastrophic events over 2.5 years using the CIR intensity model.

[4] Available at: http://www.mathworks.com/matlabcentral/fileexchange/37297-maximum-likelihood-estimation -of-the-cox-ingersoll-ross-process–the-matlab-implementation

7.4 Comparing the Deterministic and Stochastic Intensities

Below in Figure 7.5 are graphs of the CP CAT bond pricing surfaces using the CIR intensity for 20 000 simulations, as well as the difference between these surfaces and those generated using the deterministic intensity in Chapter 4. Given the results of Chapter 6, the conditional approach is used; i.e. the conditional loss distribution parameters are used for both the deterministic and stochastic intensities, where the realised intensity is scaled by $\frac{1}{1-F_{\hat{\gamma}^c}(H)}$ (as before). The results for the ZC CAT bond can be found in Appendix F.



(a) Log-normal. **(b)** GP. **(c)** Burr.

(d) Difference: Log-normal. **(e)** Difference: GP. **(f)** Difference: Burr.

Fig. 7.5: CP CAT bond pricing surfaces using the CIR intensity for the three distributions and the difference between the deterministic and CIR intensity.

The results seem to suggest that the use of a stochastic intensity model does lead to differences in the CAT bond prices produced compared to the deterministic intensity. However, upon closer examination, by focusing on the upper quantiles of the distribution of L_T as in Chapter 5, the prices produced for the deterministic and stochastic case appear to be fairly similar. Furthermore, these results are consistent across the three distributions considered. This suggests that the use of a more complicated, and arguably more realistic, stochastic intensity model may depend on the region in which the CAT bond in being priced, as specified in terms of

the quantiles of the aggregate loss process. For the upper quantiles, by parsimony, the simpler deterministic intensity could be used as it appears to produce similar results.

Overall, the examination of a stochastic intensity has led to the following findings. Firstly, the results illustrate that a log-normal stochastic intensity model may not be appropriate to model the occurrence of catastrophic events. Secondly, the use of the Cox-Ingersoll-Ross process leads to more reasonable results. There are, however, limitations when working with extreme data such the downward limit on the size of the estimation window for the intensity. Finally, when compared with a deterministic intensity function, the results do not differ greatly for the upper quantiles of the distribution of L_T.

Chapter 8

Conclusions

This dissertation aimed to contribute to the literature in three ways. Firstly, it examined the effect of two approximation methods, namely the mixed-approximation method and Monte Carlo simulation. Secondly, it investigated the issue of left-truncation and its effect on CAT bond prices. Finally, it looked at incorporating a stochastic intensity model into CAT bond pricing.

The mixed-approximation and Monte Carlo simulation approaches seemed to agree particularly for the regions where the price is determined by the upper quantiles of the aggregate loss process at maturity L_T. However, the heavy-tailed nature of the data set meant that only the log-normal model could be used for this mixed approximation method (i.e. the required higher moments did not exist for the estimated generalised Pareto and Burr distributions). If computing speed is a concern, the log-normal distribution is deemed acceptable and if the threshold level is relatively high then the mixed-approximation method should provide reasonably consistent CAT bond prices, using Monte Carlo simulation as a benchmark. It was also found that taking into account the left-truncation led to more satisfactory pricing surfaces: a non-parametric approach was introduced to illustrate this point, with prices under the conditional approach lying closer to the 'empirical' CAT bond prices than those under the naïve approach. Once again, issues regarding the use of very-heavy-tailed distributions were brought to light, particularly in the case of the Burr distribution. Therefore, any user of catastrophe claims data should be cautious and bare these issues in mind when choosing the appropriate loss distribution. Finally, it was noted that the use of an arguably more realistic stochastic intensity depends on the quantiles of the distribution of the aggregate loss process at maturity. For the upper quantiles, by parsimony, the deterministic intensity may be more suitable.

As stated previously, there are a number of possible assumptions that can affect the pricing of CAT bonds. Overall, the results of this dissertation served to inform CAT bond pricing as well as illustrated various major considerations.

Bibliography

Aase, K. (1999). An equilibrium model of catastrophe insurance futures and spreads, *The Geneva Papers on Risk and Insurance Theory* **24**(1): 69–96.

Albrecher, H., Hartinger, J. and Tichy, R. F. (2004). QMC techniques for CAT bond pricing, *Monte Carlo Methods and Applications* **10**(3-4): 197–211.

Andersen, L. B., Jäckel, P. and Kahl, C. (2010). Simulation of square-root processes, *Encyclopedia of Quantitative Finance* .

Anderson, R. R., Bendimerad, F., Canabarro, E. and Finkemeier, M. (2000). Analyzing insurance-linked securities, *The Journal of Risk Finance* **1**(2): 49–75.

Artemis (2016). PCS index triggers widely utilised in 2015 CAT bond issuance.
URL: *http://www.artemis.bm/blog/2016/01/06/pcs-index-triggers-widely-utilised-in-2015-cat-bond-issuance/.*

Baryshnikov, Y., Mayo, A. and Taylor, D. (1998). Pricing of CAT bonds, *Working Paper* .

Burnecki, K. and Kukla, G. (2003). Pricing of zero-coupon and coupon CAT bonds, *Applicationes Mathematicae* **30**(3): 315–324.

Burnecki, K., Kukla, G. and Taylor, D. (2005). *Statistical Tools for Finance and Insurance*, Springer Berlin Heidelberg, Berlin, Heidelberg, chapter Pricing of Catastrophe Bonds, pp. 93–114.

Burnecki, K., Misiorek, A. and Weron, R. (2010). Loss distributions, *MPRA paper*, University Library of Munich, Germany.
URL: *http://EconPapers.repec.org/RePEc:pra:mprapa:22163*

Burnecki, K. and Weron, R. (2005). Modeling of the risk process, *Statistical Tools for Finance and Insurance*, Springer, pp. 319–339.

Chaubey, Y. P., Garrido, J. and Trudeau, S. (1998). On the computation of aggregate claims distributions: some new approximations, *Insurance: Mathematics and Economics* **23**(3): 215–230.

Chernobai, A., Burnecki, K., Rachev, S., Trück, S. and Weron, R. (2006). Modelling catastrophe claims with left-truncated severity distributions, *Computational Statistics* **21**(3-4): 537–555.

Chernobai, A., Rachev, S. T. and Fabozzi, F. J. (2015). Composite goodness-of-fit tests for left-truncated loss samples, *Handbook of Financial Econometrics and Statistics* pp. 575–596.

Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985). A theory of the term structure of interest rates, *Econometrica* **53**(2): 385–407.

Cox, S. H., Fairchild, J. R. and Pedersen, H. W. (2000). Actuarial and economic aspects of securitization of risk, *ASTIN Bulletin* **20**(1): 157 – 193.

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

Cummins, J. D. and Weiss, M. A. (2009). Convergence of insurance and financial markets: Hybrid and securitized risk-transfer solutions, *Journal of Risk and Insurance* **76**(3): 493–545.

Cummins, J. and Geman, H. (1995). Pricing catastrophe insurance futures and call spreads: An arbitrage approach, *The Journal of Fixed Income* **4**(4): 46–57.

Dassios, A. and Jang, J.-W. (2003). Pricing of catastrophe reinsurance and derivatives using the Cox process with shot noise intensity, *Finance and Stochastics* **7**(1): 73–95.

Embrechts, P. and Meister, S. (1997). Pricing insurance derivatives, the case of CAT-futures, *Proceedings of the 1995 Bowles Symposium on Securitization of Risk, Georgia State University Atlanta, Society of Actuaries, Monograph M-FI97-1*, pp. 15–26.

Jarrow, R. A. (2009). Credit risk models, *Annual Review of Financial Economics* **1**(1): 37–68.

Jarrow, R. A. (2010). A simple robust model for CAT bond valuation, *Finance Research Letters* **7**(2): 72 – 79.

Kerney, G. (2013). Everything you need to know about the PCS catastrophe loss index, *Property Claims Services* .
URL: *http://www.scribd.com/doc/127385433/Everything-You-Need-to-Know-about-the-PCS-Catastrophe-Loss-Index#scribd*

Kladıvko, K. (2007). Maximum likelihood estimation of the cox-ingersoll-ross process: the matlab implementation, *Technical Computing Prague* .
URL: *http://citeseerx.ist.psu.edu/viewdoc/downloaddoi=10.1.1.456.4683&rep=rep1&type=pdf*

Klugman, S. A., Panjer, H. H. and Willmot, G. E. (2004). *Loss models: from data to decisions*, Vol. 715, John Wiley & Sons, Hoboken, New Jersey.

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

Lin, S.-K., Chang, C.-C. and Powers, M. (2009). The valuation of contingent capital with catastrophe risks, *Insurance: Mathematics and Economics* **45**(1): 65–73.

Liu, J., Xiao, J., Yan, L. and Wen, F. (2014). Valuing catastrophe bonds involving credit risks, *Mathematical Problems in Engineering* **2014**.

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

Ma, Z.-G., Zou, X.-Y., Xiao, S.-S. and Ma, C.-Q. (2015). The valuation of catastrophe bonds with doubly stochastic compound poisson losses, *preprint* .

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous, *Journal of Financial Economics* **3**(1-2): 125–144.

Nowak, P. and Romaniuk, M. (2013). Pricing and simulations of catastrophe bonds, *Insurance: Mathematics and Economics* **52**(1): 18 – 28.

Phillips, P. C. and Yu, J. (2009). Maximum likelihood and gaussian estimation of continuous time models in finance, *Handbook of financial time series*, Springer, pp. 497–530.

Reijnen, R., Albers, W. and Kallenberg, W. C. (2005). Approximations for stop-loss reinsurance premiums, *Insurance: Mathematics and Economics* **36**(3): 237–250.

Ross, S. (2002). *Simulation*, Academic Press, San Diego, USA.

Schmidt, T. (2014). Catastrophe insurance modeled by shot-noise processes, *Risks* **2**(1): 3–24.

Shao, J., Pantelous, A. and Papaioannou, A. (2015). Catastrophe risk bonds with applications to earthquakes, *European Actuarial Journal* **5**(1): 113–138.

Sun, L., Turvey, C. G. and Jarrow, R. A. (2015). Designing catastrophic bonds for catastrophic risks in agriculture, *Agricultural Finance Review* **75**(1): 47–62.

Unger, A. J. (2010). Pricing index-based catastrophe bonds: Part 1: Formulation and discretization issues using a numerical PDE approach, *Computers & Geosciences* **36**(2): 139–149.

Vaugirard, V. E. (2003a). Pricing catastrophe bonds by an arbitrage approach, *The Quarterly Review of Economics and Finance* **43**(1): 119–132.

Vaugirard, V. E. (2003b). Valuing catastrophe bonds by Monte Carlo simulations, *Applied Mathematical Finance* **10**(1): 75–90.

Wong, T. and Li, W. (2006). A note on the estimation of extreme value distributions using maximum product of spacings, *Time Series and Related Topics*, Institute of Mathematical Statistics, pp. 272–283.

Zimbidis, A. A., Frangos, N. E. and Pantelous, A. A. (2007). Modeling earthquake risk via extreme value theory and pricing the respective catastrophe bonds, *ASTIN bulletin* **37**(01): 163–183.

Appendix A

Cox & Pedersen's Valuation Framework

A.1 Probabilistic Structure

Let $T < \infty$ and $\mathbb{T}' = \{0, 1, ..., k, ..., T\}$ be the discrete-time trading interval. The discrete-time model developed by Cox and Pedersen (2000) consists of primary financial market variables and catastrophe risk variables. These variables are described by the filtered probability spaces

$$(\Omega^{(1)}, \mathcal{F}^{(1)}, \mathbb{P}^{(1)}) \text{ and } (\Omega^{(2)}, \mathcal{F}^{(2)}, \mathbb{P}^{(2)})$$

respectively, where $\mathcal{F}^{(1)} = (\mathcal{F}_k^{(1)})_{k\in\mathbb{T}'}$ and $\mathcal{F}^{(2)} = (\mathcal{F}_k^{(2)})_{k\in\mathbb{T}'}$ are defined to be an increasing sequences of σ-algebras. Both $\Omega^{(1)}$ and $\Omega^{(2)}$ are assumed to be finite. Note also that $\mathbb{P}^{(1)}$ and $\mathbb{P}^{(2)}$ are defined by $\mathbb{P}^{(1)} : \mathcal{F}_T^{(1)} \to [0, 1]$ and $\mathbb{P}^{(2)} : \mathcal{F}_T^{(2)} \to [0, 1]$ respectively.

In order to define the overall probability space

$$(\Omega, \mathcal{F}, \mathbb{P}),$$

Ω is set to be $\Omega = \Omega^{(1)} \times \Omega^{(2)}$ where $\omega = (\omega^{(1)}, \omega^{(2)})$. Therefore, $\omega^{(1)}$ represents a state of the financial risk variables, $\omega^{(2)}$ represents a state of the catastrophe risk variables and ω represents their joint occurrence. Furthermore, $\mathcal{F}$ is defined by $\mathcal{F} = (\mathcal{F}_k)_{k\in\mathbb{T}'}$ where $\mathcal{F}_k = \mathcal{F}_k^{(1)} \times \mathcal{F}_k^{(2)}$ for $k \in \mathbb{T}'$. Finally, the probability measure $\mathbb{P} : \mathcal{F}_T \to [0, 1]$ is defined by the product measure

$$\mathbb{P}(\omega) = \mathbb{P}^{(1)}(\omega^{(1)})\mathbb{P}^{(2)}(\omega^{(2)}).$$

Note that this specification of $\mathbb{P}$ implies that the financial risk and the catastrophe risk variables are independent under the real-world measure, which provides the model with its desirable features.

To formalise this notion of independence, two new filtrations are defined,

$$\mathcal{A}^{(1)} = (\mathcal{A}_k^{(1)})_{k\in\mathbb{T}'} \text{ and } \mathcal{A}^{(2)} = (\mathcal{A}_k^{(2)})_{k\in\mathbb{T}'}$$

where $\mathcal{A}_k^{(1)} = \mathcal{F}_k^{(1)} \times \{\emptyset, \Omega^{(2)}\}$ and $\mathcal{A}_k^{(2)} = \{\emptyset, \Omega^{(1)}\} \times \mathcal{F}_k^{(2)}$ for $k \in \mathbb{T}'$.

Lemma A.1.1. $\mathcal{A}_T^{(1)}$ and $\mathcal{A}_T^{(2)}$ are independent under $\mathbb{P}$.

Proof. Let $\alpha_1 \in \mathcal{A}_T^{(1)}$ and $\alpha_2 \in \mathcal{A}_T^{(2)}$. Then $\alpha_1 = \{A, \Omega^{(2)}\}$ for some $A \in \mathcal{F}_T^{(1)}$ and $\alpha_2 = \{\Omega^{(1)}, B\}$ for some $B \in \mathcal{F}_T^{(2)}$.

$$\begin{aligned}
\mathbb{P}(\alpha_1 \cap \alpha_2) &= \mathbb{P}(\{A, \Omega^{(2)}\} \cap \{\Omega^{(1)}, B\}) \\
&= \mathbb{P}(\{A, B\}) \\
&= \mathbb{P}^{(1)}(A)\, \mathbb{P}^{(2)}(B) \\
&= (\mathbb{P}^{(1)}(A) \times 1)\, (1 \times \mathbb{P}^{(2)}(B)) \\
&= (\mathbb{P}^{(1)}(A)\mathbb{P}^{(2)}(\Omega^{(2)}))\, (\mathbb{P}^{(1)}(\Omega^{(1)})\mathbb{P}^{(2)}(B)) \\
&= \mathbb{P}(A \times \Omega^{(2)})\, \mathbb{P}(\Omega^{(1)} \times B) \\
&= \mathbb{P}(\alpha_1)\, \mathbb{P}(\alpha_2)
\end{aligned}$$

□

A.2 Valuation Framework

The pricing framework of Cox and Pedersen (2000) is an equilibrium pricing model, rather than an arbitrage-free model, and is based on the theory of a representative agent. In the representative agent technique, the two important quantities are the aggregate consumption process $\{C^*(k) : k = 0, 1, ..., T\}$ and the representative agent's utility function $\{u_k(.) : k = 0, 1, ..., T\}$.

By the theory of the representative agent, the price $V_0(d)$ at time 0 of a generic future cashflow process $d = \{d(k)|k = 1, 2, ..., T\}$ is given by:

$$V_0(d) = \mathrm{E}^{\mathbb{P}}\left[\sum_{k=1}^{T} \frac{u_k'(C^*(k))}{u_0'(C^*(0))} d(k)\right]. \tag{A.1}$$

More generally, the price $V_n(d)$ at time n of a generic future cashflow process $d = \{d(k)|k = n+1, n+2, ..., T\}$ is given by:

$$V_n(d) = \mathrm{E}^{\mathbb{P}}\left[\sum_{k=n+1}^{T} \frac{u_k'(C^*(k))}{u_n'(C^*(n))} d(k) \middle| \mathcal{F}_n\right]. \tag{A.2}$$

These are equivalent to Equation (5.5) and (5.6) in Cox and Pedersen (2000).

The following variables are defined to illustrate the equivalence of the above pricing formula with risk-neutral valuation. This notion of equivalence is formalised in Theorem 1.4.

The one-period interest rates are defined by

$$r(k) := \frac{1}{\mathrm{E}^{\mathbb{P}}\left[\frac{u'_{k+1}(C^*(k+1))}{u'_k(C^*(k))}\middle|\mathcal{F}_k\right]} - 1 \tag{A.3}$$

for $k = 0, 1, ..., T-1$.
Given the valuation formulae above, it is clear that $r(k)$ is defined as the rate of return earned over one time period. Equation (A.3) can be expressed equivalently as

$$\frac{1}{1+r(k)} := \mathrm{E}^{\mathbb{P}}\left[\frac{u'_{k+1}(C^*(k+1))}{u'_k(C^*(k))}\middle|\mathcal{F}_k\right]$$

for $k = 0, 1, ..., T-1$.

The Radon-Nikodym Derivative is defined by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} := \prod_{i=0}^{T-1}[1+r(i)]\frac{u'_T(C^*(T))}{u'_0(C^*(0))},$$

and the Radon-Nikodym Process is defined by

$$Z(k) := \mathrm{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_k\right]$$

for $k = 0, 1, ..., T$.

The money-market account is defined by

$$B(k) := \begin{cases} \prod_{i=0}^{k-1}[1+r(i)] & \text{if } k = 1, 2, ..., T \\ 1 & \text{if } k = 0. \end{cases}$$

Additionally, the stochastic process $\zeta(k)$ is defined as:

$$\zeta(k) := \begin{cases} \prod_{i=0}^{k-1}[1+r(i)]\frac{u'_k(C^*(k))}{u'_0(C^*(0))} & \text{if } k = 1, 2, ..., T \\ 1 & \text{if } k = 0. \end{cases}$$

The importance of the above stochastic process is shown in lemma A.2.1 and in the proof of Theorem A.2.3.

Lemma A.2.1. The process $\{\zeta(k) : k = 0, 1, ..., T\}$ is a $\mathbb{P}$-martingale and $\zeta(k) = Z(k)$ for $k = 0, 1, ..., T$.

Proof. Given that $\mathbb{T}'$ is discrete, it suffices to show that $\mathrm{E}^{\mathbb{P}}[\zeta(k+1)] = \zeta(k)$ for $k = 1, 2, ..., T$.

$$\begin{aligned}
\mathrm{E}^{\mathbb{P}}[\zeta(k+1)] &= \mathrm{E}^{\mathbb{P}}\left[\zeta(k)[1+r(k)]\frac{u'_{k+1}(C^*(k+1))}{u'_k(C^*(k))}\middle|\mathcal{F}_k\right] \\
&= \zeta(k)[1+r(k)]\mathrm{E}^{\mathbb{P}}\left[\frac{u'_{k+1}(C^*(k+1))}{u'_k(C^*(k))}\middle|\mathcal{F}_k\right] && \text{(because } \zeta(k)[1+r(k)] \text{ is } \mathcal{F}_k\text{-measurable)} \\
&= \zeta(k) && \text{(by definition of } r(k)).
\end{aligned}$$

Moreover,

$$\begin{aligned}\zeta(k) &= \mathrm{E}^{\mathbb{P}}[\zeta(T)|\mathcal{F}_k] && \text{(because } \{\zeta(k) : k = 0, 1, ..., T\} \text{ is a } \mathbb{P}\text{-martingale)} \\ &= \mathrm{E}^{\mathbb{P}}[\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_k] && \text{(since } \zeta(T) = \frac{d\mathbb{Q}}{d\mathbb{P}}\text{)} \\ &= Z(k).\end{aligned}$$

□

Using Lemma 1.2 and the fact that $\zeta(T) = \frac{d\mathbb{Q}}{d\mathbb{P}}$, it is clear that

$$\mathbb{Q}(\Omega) = \int_{\Omega} \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} = \mathrm{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\right] = \mathrm{E}^{\mathbb{P}}\left[\zeta(T)\right] = \mathrm{E}^{\mathbb{P}}\left[\zeta(0)\right] = 1.$$

Finally, the following well-known result is necessary before proceeding.

Theorem A.2.2 (Bayes Theorem)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with an equivalent martingale measure $\mathbb{Q}$ and let X be an $\mathcal{F}_{T^-}$-measurable random variable. Then,

$$\mathrm{E}^{\mathbb{Q}}[X|\mathcal{F}_n] = \frac{1}{Z(n)}\mathrm{E}^{\mathbb{P}}[XZ(T)|\mathcal{F}_n].$$

Proof. For proof, see Cox and Pedersen (2000). □

Now, the equivalence between the Representative Agent Technique and Risk-neutral Valuation is shown in Theorem A.2.3.

Theorem A.2.3. The price $V_0(d)$ at time 0 of a generic future cashflow process $d = \{d(k)|k = 1, 2, ..., T\}$, given by Equation (A.1), can equivalently be calculated as

$$V_0(d) = \mathrm{E}^{\mathbb{Q}}\left[\sum_{k=1}^{T} \frac{1}{B(k)} d(k)\right].$$

More generally, the price $V_n(d)$ at time n of a generic future cashflow process $d = \{d(k)|k = n + 1, n + 2, ..., T\}$, given by Equation (A.2), can equivalently be calculated as

$$V_n(d) = \mathrm{E}^{\mathbb{Q}}\left[\sum_{k=n+1}^{T} \frac{B(n)}{B(k)} d(k) \middle| \mathcal{F}_n\right].$$

Proof. For $k = 1, 2, ..., T$

$$\begin{aligned}
\mathrm{E}^{\mathbb{Q}}\left[\frac{1}{B(k)}d(k)\right] &= \mathrm{E}^{\mathbb{P}}\left[\frac{1}{B(k)}d(k)\frac{d\mathbb{Q}}{d\mathbb{P}}\right] && \text{(change of measure)}\\
&= \mathrm{E}^{\mathbb{P}}\left[\mathrm{E}^{\mathbb{P}}\left[\frac{1}{B(k)}d(k)\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_k\right]\right] && \text{(Tower property)}\\
&= \mathrm{E}^{\mathbb{P}}\left[\frac{1}{B(k)}d(k)\mathrm{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_k\right]\right] && \text{(because}\frac{1}{B(k)}d(k)\text{ is }\mathcal{F}_k\text{-measurable)}\\
&= \mathrm{E}^{\mathbb{P}}\left[\frac{1}{B(k)}d(k)Z(k)\right] && \text{(by definition of } Z(k))\\
&= \mathrm{E}^{\mathbb{P}}\left[\frac{1}{B(k)}d(k)\zeta(k)\right] && \text{(since } Z(k)=\zeta(k))\\
&= \mathrm{E}^{\mathbb{P}}\left[\frac{u_k'(C^*(k))}{u_0'(C^*(0))}d(k)\right] && \text{(by definition of } \zeta(k))\\
&= V_0(d).
\end{aligned}$$

More generally, for $k = n+1, ..., T$,

$$\begin{aligned}
\mathrm{E}^{\mathbb{Q}}\left[\frac{B(n)}{B(k)}d(k)\middle|\mathcal{F}_n\right] &= \frac{1}{Z(n)}\mathrm{E}^{\mathbb{P}}\left[\frac{B(n)}{B(k)}d(k)Z(T)\middle|\mathcal{F}_n\right] && \text{(Bayes Theorem)}\\
&= \frac{1}{Z(n)}\mathrm{E}^{\mathbb{P}}\left[\mathrm{E}^{\mathbb{P}}\left[\frac{B(n)}{B(k)}d(k)Z(T)\middle|\mathcal{F}_k\right]\middle|\mathcal{F}_n\right] && \text{(by Tower property)}\\
&= \frac{1}{Z(n)}\mathrm{E}^{\mathbb{P}}\left[\mathrm{E}^{\mathbb{P}}\left[\frac{B(n)}{B(k)}d(k)\zeta(T)\middle|\mathcal{F}_k\right]\middle|\mathcal{F}_n\right] && \text{(since } Z(T)=\zeta(T))\\
&= \frac{1}{Z(n)}\mathrm{E}^{\mathbb{P}}\left[\frac{B(n)}{B(k)}d(k)\mathrm{E}^{\mathbb{P}}\left[\zeta(T)|\mathcal{F}_k\right]\middle|\mathcal{F}_n\right] && \text{(because } \frac{1}{Z(n)} \text{ is } \mathcal{F}_n\text{-measurable)}\\
&= \frac{1}{Z(n)}\mathrm{E}^{\mathbb{P}}\left[\frac{B(n)}{B(k)}d(k)\zeta(k)\middle|\mathcal{F}_n\right] && (\zeta \text{ is a } \mathbb{P}\text{-martingale})\\
&= \mathrm{E}^{\mathbb{P}}\left[\frac{B(n)}{\zeta(n)}\frac{\zeta(k)}{B(k)}d(k)\middle|\mathcal{F}_n\right] && \text{(since } Z(n)=\zeta(n))\\
&= \mathrm{E}^{\mathbb{P}}\left[\frac{u_k'(C^*(k))}{u_n'(C^*(n))}d(k)\middle|\mathcal{F}_n\right] && \text{(by definition of } \zeta(k) \text{ and } \zeta(n))\\
&= V_n(d).
\end{aligned}$$

□

Assumption 1. *Aggregate consumption depends only on the financial market variables.*

This assumption leads to the following two important lemmas, because the derivation of the CAT bond pricing formula of Ma and Ma (2013) makes use of them.

Lemma A.2.4. The expectation of a random variable that depends only on catastrophe risk is the same under the real-world and risk-neutral measures.

Proof. Firstly note that, by Assumption 1., the one-period interest rates only depend on the financial risk variables. By definition, the same is true for the Radon-Nikodym derivative. Therefore, $\frac{d\mathbb{Q}}{d\mathbb{P}} \in \mathcal{A}_T^{(1)}$.

Now, let $\alpha_{\text{CAT}} \in \mathcal{A}_T^{(2)}$.

$$\begin{aligned}
\mathbb{Q}(\alpha_{\text{CAT}}) &= \mathrm{E}^{\mathbb{Q}}\left[\mathbb{I}_{\alpha_{\text{CAT}}}\right] \\
&= \mathrm{E}^{\mathbb{P}}\left[\mathbb{I}_{\alpha_{\text{CAT}}}\frac{d\mathbb{Q}}{d\mathbb{P}}\right] \\
&= \mathrm{E}^{\mathbb{P}}\left[\mathbb{I}_{\alpha_{\text{CAT}}}\right]\mathrm{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\right] && \text{(by definition of } \mathbb{P}) \\
&= \mathrm{E}^{\mathbb{P}}\left[\mathbb{I}_{\alpha_{\text{CAT}}}\right] && \text{(by definition of } \frac{d\mathbb{Q}}{d\mathbb{P}}) \\
&= \mathbb{P}(\alpha_{\text{CAT}}).
\end{aligned}$$

□

Lemma A.2.5. Financial risk and catastrophe risk variables are independent under the risk-neutral measure.

Proof. Let $\alpha_{\text{FIN}} \in \mathcal{A}_T^{(1)}$ and $\alpha_{\text{CAT}} \in \mathcal{A}_T^{(2)}$.

$$\begin{aligned}
\mathbb{Q}(\alpha_{\text{FIN}} \cap \alpha_{\text{CAT}}) &= \mathrm{E}^{\mathbb{Q}}\left[\mathbb{I}_{\alpha_{\text{FIN}}}\mathbb{I}_{\alpha_{\text{CAT}}}\right] \\
&= \mathrm{E}^{\mathbb{P}}\left[\mathbb{I}_{\alpha_{\text{FIN}}}\mathbb{I}_{\alpha_{\text{CAT}}}\frac{d\mathbb{Q}}{d\mathbb{P}}\right] \\
&= \mathrm{E}^{\mathbb{P}}\left[\left(\mathbb{I}_{\alpha_{\text{FIN}}}\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\mathbb{I}_{\alpha_{\text{CAT}}}\right] \\
&= \mathrm{E}^{\mathbb{P}}\left[\mathbb{I}_{\alpha_{\text{FIN}}}\frac{d\mathbb{Q}}{d\mathbb{P}}\right]\mathrm{E}^{\mathbb{P}}\left[\mathbb{I}_{\alpha_{\text{CAT}}}\right] && \left(\text{since } \mathbb{I}_{\alpha_{\text{FIN}}}\frac{d\mathbb{Q}}{d\mathbb{P}} \in \mathcal{A}_T^{(1)} \text{ and } \mathbb{I}_{\alpha_{\text{CAT}}} \in \mathcal{A}_T^{(2)}\right) \\
&= \mathrm{E}^{\mathbb{Q}}\left[\mathbb{I}_{\alpha_{\text{FIN}}}\right]\mathrm{E}^{\mathbb{P}}\left[\mathbb{I}_{\alpha_{\text{CAT}}}\right] \\
&= \mathbb{Q}(\alpha_{\text{FIN}})\mathbb{P}(\alpha_{\text{CAT}}) \\
&= \mathbb{Q}(\alpha_{\text{FIN}})\mathbb{Q}(\alpha_{\text{CAT}}) && \text{(by Lemma A.2.4.)}.
\end{aligned}$$

□

Both Lemma A.2.4 and A.2.5 are key results necessary to simplify the CAT bond pricing formula of Ma and Ma (2013).

Appendix B

Adjusted Goodness of Fit Tests

Tab. B.1: Comparison of goodness of fit test statistics for complete and left-truncated data samples. Note that $F_n(x)$ refers to the EDF of the data sample, $z_j = F_{\hat{\gamma}}(x_{(j)})$ and $z_H = F_{\hat{\gamma}^c}(H)$.

Test	Sample	Statistic	Computing formula (Chernobai *et al.*, 2006)
D	Complete	$\sqrt{n} \cdot \sup_x \lvert F_n(x) - \widehat{F}(x) \rvert$	$\sqrt{n} \cdot \max\left\{ \sup_j \left\{ \frac{j}{n} - z_j \right\}, \sup_j \left\{ z_j - \frac{j-1}{n} \right\} \right\}$
	Left-truncated	$\sqrt{n} \cdot \sup_x \lvert F_n(x) - \widehat{F^*}(x) \rvert$	$\frac{\sqrt{n}}{1-z_H} \cdot \max\left\{ \sup_j \left\{ z_H + \frac{j}{n}(1-z_H) - z_j \right\}, \sup_j \left\{ z_j - \left(\frac{j-1}{n}(1-z_H) \right) \right\} \right\}$
V	Complete	$\sqrt{n}\left(\sup_x \left\{ F_n(x) - \widehat{F}(x) \right\} + \sup_x \left\{ \widehat{F}(x) - F_n(x) \right\} \right)$	$\sqrt{n}\left(\sup_j \left\{ \frac{j}{n} - z_j \right\} + \sup_j \left\{ z_j - \frac{j-1}{n} \right\} \right)$
	Left-truncated	$\sqrt{n}\left(\sup_x \left\{ F_n(x) - \widehat{F^*}(x) \right\} + \sup_x \left\{ \widehat{F^*}(x) - F_n(x) \right\} \right)$	$\frac{\sqrt{n}}{1-z_H}\left(\sup_j \left\{ z_H + \frac{j}{n}(1-z_H) - z_j \right\} + \sup_j \left\{ z_j - \left(\frac{j-1}{n}(1-z_H) \right) \right\} \right)$
AD^2	Complete	$n \cdot \int_{-\infty}^{\infty} \frac{\left(F_n(x) - \widehat{F}(x) \right)^2}{\widehat{F}(x)\left(1 - \widehat{F}(x) \right)} d\widehat{F}(x)$	$-n + \frac{1}{n}\sum_{j=1}^{n}(1-2j)\log(z_j) - \frac{1}{n}\sum_{j=1}^{n}(1+2(n-j))\log(1-z_j)$
	Left-truncated	$n \cdot \int_{H}^{\infty} \frac{\left(F_n(x) - \widehat{F^*}(x) \right)^2}{\widehat{F^*}(x)\left(1 - \widehat{F^*}(x) \right)} d\widehat{F^*}(x)$	$-n + 2n\log(1-z_H) - \frac{1}{n}\sum_{j=1}^{n}(1+2(n-j))\log(1-z_j) + \frac{1}{n}\sum_{j=1}^{n}(1-2j)\log\lvert z_j - z_H \rvert$
W^2	Complete	$n\int_{-\infty}^{\infty}\left(F_n(x) - \widehat{F}(x) \right)^2 d\widehat{F}(x)$	$\frac{n}{3} + \frac{1}{n}\sum_{j=1}^{n}(1-2j)\log(z_j) + \sum_{j=1}^{n} z_j^2$
	Left-truncated	$n\int_{-\infty}^{\infty}\left(F_n(x) - \widehat{F^*}(x) \right)^2 d\widehat{F^*}(x)$	$\frac{n}{3} + \frac{nz_h}{1-z_H} + \frac{1}{n(1-z_H)}\sum_{j=1}^{n}(1-2j)z_j + \frac{1}{(1-z_H)^2}\sum_{j=1}^{n}(z_j - z_H)^2$

Appendix C

Mixed-Approximation Method

C.1 Useful Formulae

For a compound Poisson process $L = \sum_{i=1}^{N} X_i$, the raw moments are given by the following formulae (Reijnin, 2005):

$\mu_{1L} = \mu_{1N}\mu_{1X}$,
$\mu_{2L} = \mu_{2N}\mu_{1X}^2 + \mu_{1N}\mu_{2X} - \mu_{1N}\mu_{1X}^2$,
$\mu_{3L} = \mu_{3N}\mu_{1X}^3 + 3\mu_{2N}\mu_{1X}\mu_{2X} - 3\mu_{2N}\mu_{1X}^3 + \mu_{1N}\mu_{3X} - 3\mu_{1N}\mu_{2X}\mu_{1X} + 2\mu_{1N}\mu_{1X}^3$,
$\mu_{4L} = \mu_{4N}\mu_{1X}^4 + 6\mu_{3N}\mu_{1X}^2\mu_{2X} - 6\mu_{3N}\mu_{1X}^4 + 3\mu_{2N}\mu_{2X}^2 - 18\mu_{2N}\mu_{1X}^2\mu_{2X} + 4\mu_{2N}\mu_{1X}\mu_{3X}$
$+ 11\mu_{2N}\mu_{1X}^4 + \mu_{1N}\mu_{4X} - 4\mu_{1N}\mu_{3X}\mu_{1X} + 12\mu_{1N}\mu_{2X}\mu_{1X}^2 - 3\mu_{1N}\mu_{2X}^2 - 6\mu_{1N}\mu_{1X}^4$,

where $\mu_{iY} = \mathrm{E}[Y^i]$.

The raw moments of a (non)-homogeneous Poisson process N with intensity λ_t on $[0, T]$ are given by

$\mu_{1N} = v$
$\mu_{2N} = v(1 + v)$
$\mu_{3N} = v(1 + 3v + v^2)$
$\mu_{4N} = v(1 + 7v + 6v^2 + v^3)$

where $v = \int_t^T \lambda_s ds$.

C.2 Parameterisation

Note that the parameterisations given in the specification of the mixed-approximation method differ from those used in $Matlab$. In $Matlab$, the gamma distribution is specified in terms of x, a and b where

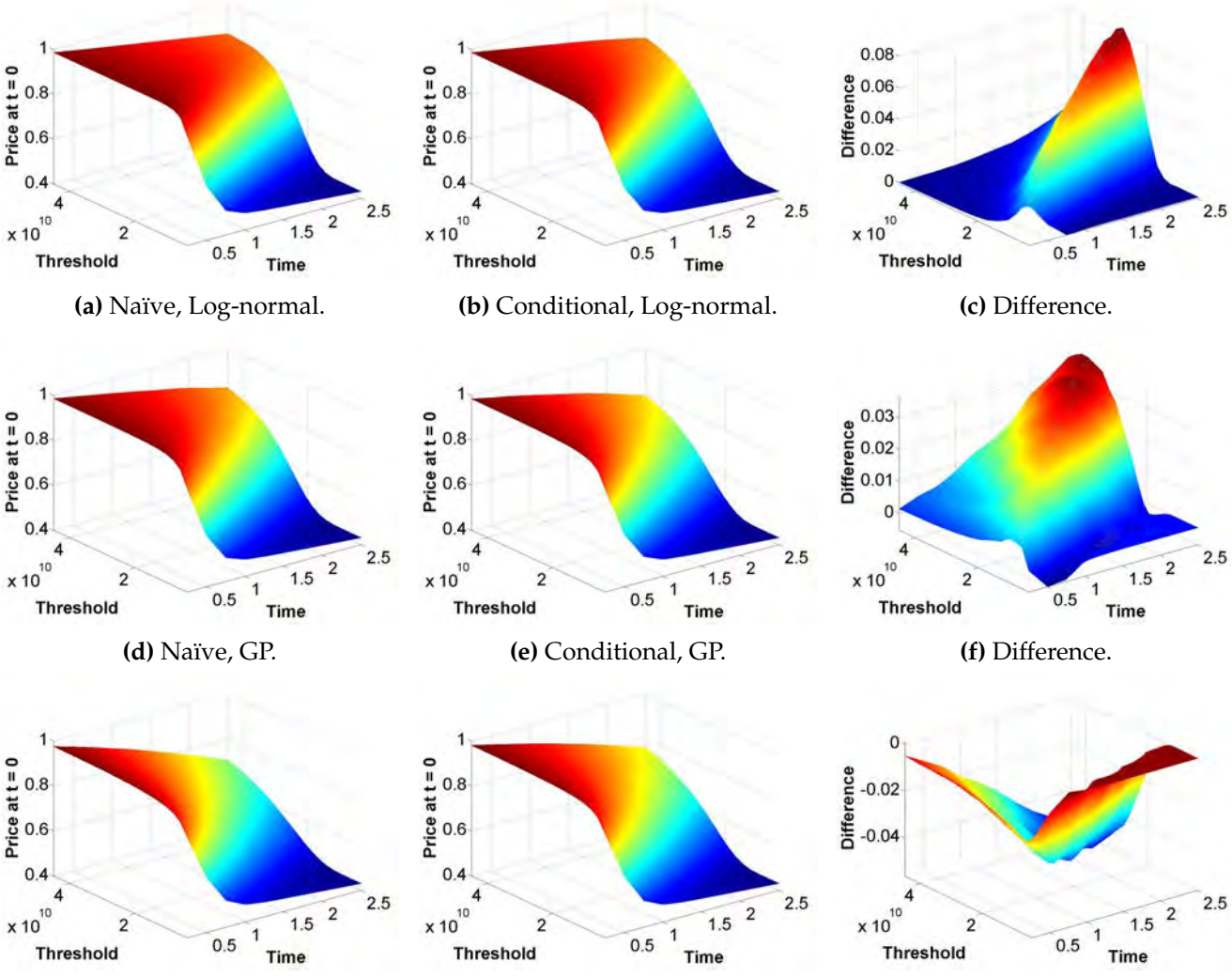
$$x = l - x_0,\ a = \alpha \text{ and } b = \frac{1}{\beta}.$$

Similarly, the inverse Gaussian distribution is specified in terms of x, λ and μ where

$$x = l - x_0,\ \lambda = \frac{\alpha^2}{\beta} \text{ and } \mu = \frac{\alpha}{\beta}.$$

Appendix D

ZC CAT Bonds I



(a) Naïve, Log-normal. **(b)** Conditional, Log-normal. **(c)** Difference.

(d) Naïve, GP. **(e)** Conditional, GP. **(f)** Difference.

(g) Naïve, Burr. **(h)** Conditional, Burr. **(i)** Difference.

Fig. D.1: ZC CAT bond pricing surfaces for the naïve and conditional approaches and their difference, for the three fitted distributions. .

Appendix E

Cox-Ingersoll-Ross process

E.1 Maximum Likelihood Estimation

The following description of the maximum likelihood estimation procedure for the Cox-Ingersoll-Ross process follows that of Kladıvko (2007) very closely.

Let $\{r_{t_i} : i = 1, 2, \ldots, N\}$ be an observed time series of interest rates, let $\theta = (\alpha, \mu, \sigma)$ be the parameters of the Cox-Ingersoll-Ross model that need to be estimated. The estimation procedure is based on the transition density of the Cox-Ingersoll-Ross process which is shown in Equation (E.1) below.

As in Kladıvko (2007), given r_t, for a change in time Δt the transition density of $r_{t+\Delta t}$ is given by

$$p(r_{t+\Delta t}|r_t, \theta) = ce^{-u-v}\left(\frac{v}{u}\right)^{\frac{q}{2}} I_q(2\sqrt{uv}) \tag{E.1}$$

where

$$c = \frac{2\alpha}{\sigma^2(1 - e^{\alpha\Delta t})},$$

$$u = cr_t e^{-\alpha\Delta t},$$

$$v = cr_{t+\Delta t},$$

$$q = \frac{2\alpha\mu}{\sigma^2} - 1,$$

and $I_q(2\sqrt{uv})$ is the modified Bessel function of the first kind of order q. The aim is to estimate θ via maximum likelihood estimation and the likelihood function is given, in term of the transition densities, by

$$L(\theta) = \prod_{i=1}^{N-1} p(r_{t_{i+1}}|r_{t_i}, \theta).$$

The log-likelihood, which is typically used in estimation, can be expressed as follows

$$\begin{aligned}\log L(\theta) &= \sum_{i=1}^{N-1} \log(p(r_{t_{i+1}}|r_{t_i},\theta)) \\ &= (N-1)\log c + \sum_{i=1}^{N-1}\left(-u_{t_i} - v_{t_{i+1}} + \frac{1}{2}q\log\left(\frac{v_{t_{i+1}}}{u_{t_i}}\right) + \log(I_q(2\sqrt{u_{t_i}v_{t_{i+1}}}))\right).\end{aligned}$$

Kladıvko (2007) notes that there is an issue of rapid divergence for the $Matlab$ Bessel function, which hinders the estimation process. Therefore, he suggests using the scaled version of the Bessel function $I_q^1(2\sqrt{uv}) = I_q(2\sqrt{uv})\exp(-2\sqrt{uv})$, which deals with this issue.

$$\log L(\theta) = (N-1)\log c + \sum_{i=1}^{N-1}\left(-u_{t_i} - v_{t_{i+1}} + \frac{1}{2}q\log\left(\frac{v_{t_{i+1}}}{u_{t_i}}\right) + \log(I_q^1(2\sqrt{u_{t_i}v_{t_{i+1}}} + 2\sqrt{u_{t_i}v_{t_{i+1}}}))\right).$$

Therefore, the parameters of the Cox-Ingersoll-Ross model can be estimated using

$$\hat{\theta} = \arg\max_{\theta}\ \log L(\theta).$$

To specify the starting values, Kladıvko (2007) advocates the use of the following Ordinary Least Squares estimates, based on a discretised version of Equation (7.2). The initial estimates for α and θ, given by α^*, and μ^* respectively are obtained as follows

$$(\alpha^*,\mu^*) = \arg\max_{\alpha,\mu}\sum_{i=1}^{N-1}\left(\frac{r_{t_{i+1}} - r_{t_i}}{\sqrt{r_{t_i}}} - \frac{\alpha\mu\Delta t}{\sqrt{r_{t_i}}} + \alpha\sqrt{r_{t_i}}\right)^2$$

and σ^* is given by the standard deviation of the residuals.

$Matlab$ code for the implementation of the procedure described above is provided by Kladıvko (2007)[5].

E.2 Simulation

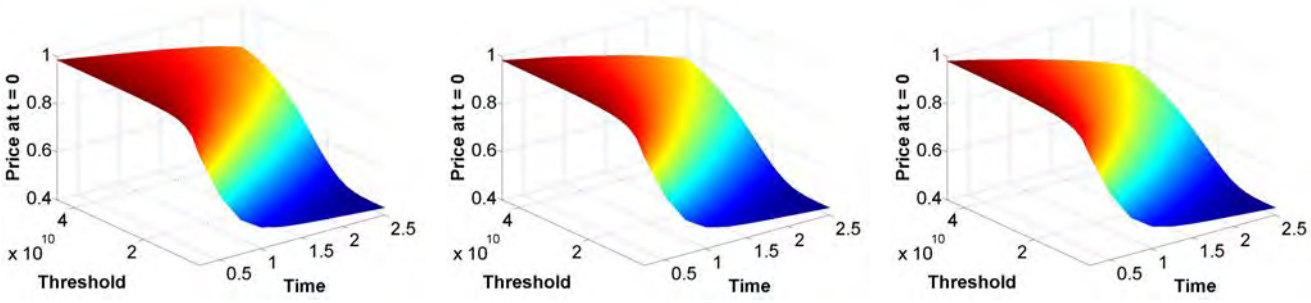
The paths for the CIR intensity process were simulated using the following algorithm, as given in Andersen *et al.* (2010) where it is assumed $\lambda_{t+\Delta t}$ is simulated given λ_t.

Firstly, generate $N \sim P(q)$ where $q = \frac{1}{2}\lambda_t n(t,t+\Delta t)$ and $n(t,T) = \frac{4\alpha e^{-\alpha(T-t)}}{\sigma^2(1-e^{-\alpha(T-t)})}$ for $T > t$. Then, given N, generate $H \sim \chi^2_v$ where $v = d + 2N$, where $d = \frac{4\alpha\mu}{\sigma^2}$. Finally, set $\lambda_{t+\Delta t} = H\left(\frac{\exp(-\alpha\Delta t)}{n(t,t+\Delta t)}\right)$.

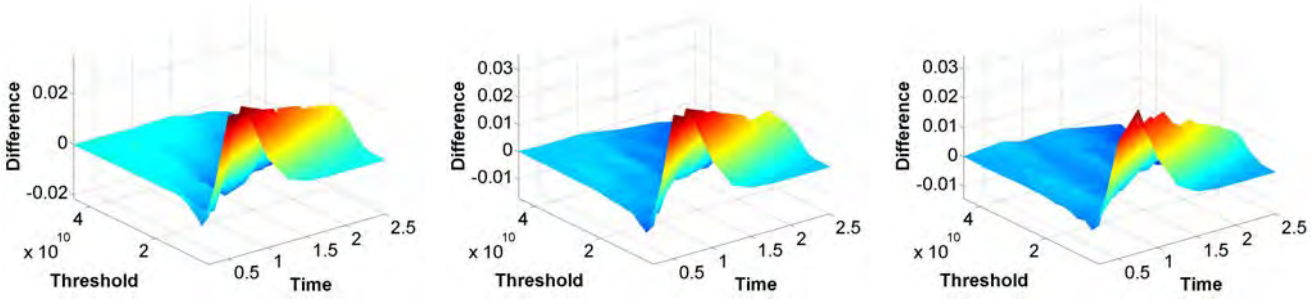
[5] Available at: http://www.mathworks.com/matlabcentral/fileexchange/37297-maximum-likelihood-estimation -of-the-cox-ingersoll-ross-process–the-matlab-implementation

Appendix F

ZC CAT Bonds II



(a) Log-normal. **(b)** GP. **(c)** Burr.



(d) Difference: Log-normal. **(e)** Difference: GP. **(f)** Difference: Burr.

Fig. F.1: ZC CAT bond pricing surfaces using the CIR intensity for the three distributions and the difference between the deterministic and CIR intensity.