KONINKLIJKE VLAAMSE ACADEMIE VAN BELGIE
VOOR WETENSCHAPPEN EN KUNSTEN

ACTUARIAL AND FINANCIAL MATHEMATICS CONFERENCE

Interplay between Finance and Insurance

February 5-6, 2015

Michèle Vanmaele, Griselda Deelstra, Ann De Schepper,
Jan Dhaene, Wim Schoutens, Steven Vanduffel & David Vyncke (Eds.)

CONTACTFORUM

KONINKLIJKE VLAAMSE ACADEMIE VAN BELGIE
VOOR WETENSCHAPPEN EN KUNSTEN

ACTUARIAL AND FINANCIAL MATHEMATICS CONFERENCE

Interplay between Finance and Insurance

February 5-6, 2015

**Michèle Vanmaele, Griselda Deelstra, Ann De Schepper,
Jan Dhaene, Wim Schoutens, Steven Vanduffel & David Vyncke (Eds.)**

CONTACTFORUM

Handelingen van het contactforum "Actuarial and Financial Mathematics Conference. Interplay between Finance and Insurance" (5-6 februari 2015, hoofdaanvrager: Prof. M. Vanmaele, UGent) gesteund door de Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten.

Afgezien van het afstemmen van het lettertype en de alinea's op de richtlijnen voor de publicatie van de handelingen heeft de Academie geen andere wijzigingen in de tekst aangebracht. De inhoud, de volgorde en de opbouw van de teksten zijn de verantwoordelijkheid van de hoofdaanvrager (of editors) van het contactforum.

KONINKLIJKE VLAAMSE ACADEMIE VAN BELGIE
VOOR WETENSCHAPPEN EN KUNSTEN

Paleis der Academiën
Hertogsstraat 1
1000 Brussel

D/2015/0455/06
ISBN 978 90 6569 150 7

KONINKLIJKE VLAAMSE ACADEMIE VAN BELGIE
VOOR WETENSCHAPPEN EN KUNSTEN

Actuarial and Financial Mathematics Conference
Interplay between finance and insurance

CONTENTS

KONINKLIJKE VLAAMSE ACADEMIE VAN BELGIE
VOOR WETENSCHAPPEN EN KUNSTEN

Actuarial and Financial Mathematics Conference
Interplay between finance and insurance

PREFACE

It was for the 8th time that our two-day international "Actuarial and Financial Mathematics Conference" was organized in Brussels, on February 5-6 2015. The organizing committee consisted of colleagues from 6 Belgian universities, i.e. Ghent University, the University of Antwerp, the KU Leuven and the Vrije Universiteit Brussel on the one hand, and the Université Libre de Bruxelles and the Université Catholique de Louvain on the other hand. As for the previous editions, we could use the facilities of the Royal Flemish Academy of Belgium for Science and Arts. Next to 8 invited lectures, there were 8 selected contributions as well as a poster session with 10 posters. We felt honoured by the presence of renowned international speakers, both from academia and from practice, and by the participation of leading international researchers in the scientific committee.

There were 131 registrations in total, with 89 participants from Belgium, and 42 participants from 15 other countries from all continents. The population was mixed, with 59% of the participants associated with a university (PhD students, post doc researchers and professors), and with 41% working in the banking and insurance industry.

On the first day, February 5, we had 8 speakers, among them 4 international and eminent invited speakers, alternated with 4 contributions selected by the scientific committee.
In the morning, the first speaker was *Prof.dr. Andrew Cairns,* from Heriot Watt University Edinburgh (UK), on "Securitization and Hedging of Longetivity Risk"; later, we could listen to *Prof.dr. Guillen Montserra*, University of Barcelona (Spain) about "Uplift predictive modeling in pricing, retention and cross selling of insurance policies". These two lectures were alternated by 2 presentations with researchers from the Netherlands and Belgium.
In the afternoon, we welcomed *Prof.dr. Véronique Maume-Deschamps,* University Claude Bernard 1 Lyon (France); she delivered a lecture entitled "On the estimation of aggregated VaR with marginal and/or dependence information". Afterwards, *Prof.dr. Fabio Bellini,* University of Milano Bicocca (Italy) presented his research, with a paper "Elicitable risk measures and expectiles". Next to these two invited speakers, there were two more selected contributions by young Swiss and Belgian researchers.

During the lunch break, we organized a poster session, preceded by a poster storm session, where the 10 different posters were presented concisely by the researchers. The posters remained available in the central hall during the whole conference, so that they could be consulted and discussed during the coffee breaks. The posters attracted a great deal of interest, judging by the lively interaction between the participants and the posters' authors.

Also on the second day, February 6, we had 8 lectures, with 4 keynote speakers and 4 selected contributions again. The first speaker was *Prof.dr. Bernt Øksendal*, University of Oslo (Norway), with a lecture on "Optimal control of stochastic Volterra equations and applications to financial markets with memory". Afterwards *Prof.dr. Wim Schoutens*, KU Leuven (Belgium) informed us on his current research outcomes in a lecture "Conic Finance Explained and Applied". In the afternoon, we could first listen to *Prof.dr. Rudi Zagst*, Technische Universität München (Germany) with a well-received lecture on "Closed-form solutions for Guaranteed Minimum Accumulation Benefits". The closing lecture was delivered by *Prof.dr. Uwe Wystup*, MathFinance AG (Germany) and University of Antwerp (Belgium); the title of this last presentation was "Volatility as investment – Crash Protection with Calendar Spreads of Variance Swaps". The other 4 presentations were again selected from a large number of submissions by the scientific committee; the speakers came from the Netherlands, Belgium and the United Kingdom.

The proceedings contain six papers and extended abstracts, giving an overview of the topics and activities at the conference.

We are much indebted to the members of the scientific committee, *Hansjörg Albrecher* (University of Lausanne, Switzerland), *Carole Bernard* (Grenoble Ecole de Management, France), *Tahir Choulli* (University of Alberta, Canada), *Michel Denuit* (Université Catholique de Louvain, Belgium), *Jan Dhaene* (Katholieke Universiteit Leuven, Belgium), *Ernst Eberlein* (University of Freiburg, Germany), *Monique Jeanblanc* (Université d'Evry Val d'Essonne, France), *Ragnar Norberg* (SAF, Université Lyon 1, France), *Ludger Rüschendorf* (University of Freiburg, Germany), *Steven Vanduffel* (Vrije Universiteit Brussel, Belgium), *Michel Vellekoop* (University of Amsterdam, The Netherlands), and the chair *Griselda Deelstra* (Université Libre de Bruxelles, Belgium). We appreciate their excellent scientific support, their presence at the meeting and their chairing of sessions. We also thank Wouter Dewolf (Ghent University, Belgium), for the administrative work.
We are very grateful to our sponsors, namely the Royal Flemish Academy of Belgium for Science and Arts, the Research Foundation — Flanders (FWO), the Scientific Research Network (WOG) "Stochastic modelling with applications in finance", le Fonds de la Recherche Scientifique (FNRS), Advanced Mathematical Methods for Finance (AMAMEF), the AG Health Insurance Chair at the KU Leuven, the BNP Paribas Fortis Chair in Banking at the Vrije Universiteit Brussel and Université Libre de Bruxelles, and Generali, and the exhibitors NAG and Springer. Without them it would not have been possible to organize this event in this very enjoyable and inspiring environment.

The continuing success of the meeting encourages us to go on with the organization of this contact-forum, in order to create future opportunities for exchanging ideas and results in this fascinating research field of actuarial and financial mathematics.

The editors:
Griselda Deelstra, Ann De Schepper, Jan Dhaene, Wim Schoutens, Steven Vanduffel,
Michèle Vanmaele, David Vyncke

The other members of the organizing committee:
Pierre Devolder, Karel In't Hout

INVITED TALK

VALUE AT RISK ESTIMATION OF AGGREGATED RISKS USING MARGINAL LAWS AND SOME DEPENDENCE INFORMATION

Andrés Cuberos[†], Esterina Masiello[§] and Véronique Maume-Deschamps[§]

[†] *Université de Lyon, Université Lyon 1, Laboratoire SAF EA 2429, SCOR SE, France*
[§] *Université de Lyon, Université Lyon 1, Institut Camille Jordan ICJ UMR 5208 CNRS, France*
Email: acuberos@scor.com, esterina.masiello@univ-lyon1.fr, veronique.maume@univ-lyon1.fr

Abstract

Estimating high level quantiles of aggregated variables (mainly sums or weighted sums) is crucial in risk management for many application fields such as finance, insurance, environment .... This question has been widely treated but new efficient methods are always welcome; especially if they apply in (relatively) high dimension. We propose an estimation procedure based on the *checkerboard copula*. It allows to get good estimations from a (quite) small sample of the multivariate law and a full knowledge of the marginal laws. This situation is realistic for many applications, mainly in insurance. Moreover, we may also improve the estimations by including in the checkerboard copula some additional information (on the law of a sub-vector or on extreme probabilities).

1. INTRODUCTION

Consider a vector of random variables $\mathbf{X} = (X_1, \ldots, X_d)$ and a measurable function $\Psi : \mathbb{R}^d \to \mathbb{R}^+$. In the context of quantitative risk management $\mathbf{X}$ is known as a risk vector and generally represents the profit-losses of a portfolio at a given future date. $\Psi(\mathbf{X})$, the aggregated risk, represents its total future position. The main examples of aggregation functions are: the sum, max, weighted sums or a slightly more complex function that may include stop-loss reinsurance type function on each of the marginals. In this paper, we will be essentially concerned with $\Psi = \sum$. We are interested here in the estimation of the Value at Risk of $\Psi(\mathbf{X})$. To this purpose, we will assume that the distributions $F_1, \ldots, F_d$ of the marginal variables $X_1, \ldots, X_d$ are known and that some information on the dependence between them is given. Usually this information is available via some observations of the joint distribution and also via expert opinion.

In general neither the marginals nor the dependence of the risk vector $\mathbf{X}$ will be known. However in many cases the knowledge on the marginal distributions will be much more important than the knowledge on the dependence. For example when some observations of the vector $\mathbf{X}$ are available, inferences one can do on the marginal distributions give better results than inferences one can do on the multivariate distribution. Moreover, on each marginal risk, one can also have extra information that is more unusual in the joint distribution, as for example expert opinion or prior information. So, even if the assumption of the knowledge of marginal distribution may seem exaggerate, there is, however, in practice much more knowledge on the marginal distributions than on the dependence structure of the random vector.

When the marginals are known but the dependence is unknown, the re-arrangement algorithm (introduced in special cases in Rüschendorf (1983) and Rüschendorf (1982)) allows to obtain bounds on the distribution of $\Psi(\mathbf{X})$ (Puccetti (2012)) working well for $d \geq 30$. By improving the re-arrangement algorithm, bounds on the VaR are obtained in Embrechts et al. (2013) in high dimensional ($d \geq 1000$) inhomogeneous portfolio. Cases in which some kind of dependence information is available lead to narrower bounds (Bernard et al. (2013), Bernard and Vanduffel (2015)) for the risk measure at hand. Bounds are also derived in Cossette et al. (2014) for dependence structures described by different copula models. A general mathematical framework which interpolates between marginal knowledge and full knowledge of the distribution function of $\mathbf{X}$ is considered in Embrechts and Puccetti (2010).

In this paper, we propose to use the checkerboard copula (introduced in Mikusinski and Taylor (2010)) to merge the information given by a small sample of the distribution of $\mathbf{X}$ with the known marginal distributions. Then, we introduce the checkerboard copula with information on the tail and with information on a sub-vector, to take into account some additional information which may improve the VaR estimation (see Section 3.1). Some simulations are provided in Section 3.2.
We begin (see Section 2) with a brief discussion on the admissible multivariate distribution with fixed marginal and aggregated laws: given marginal laws and a distribution for $\Psi(\mathbf{X})$, what are the possible multivariate distributions for $\mathbf{X}$?

2. THE INVARIANT AGGREGATION COPULA CLASS

Let $\mathbf{F}$ be the distribution function of $\mathbf{X} = (X_1, \ldots, X_d)$. By Sklars Theorem, there exists a copula distribution C in $[0, 1]^d$ such that

$$\mathbf{F}(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)),$$

and moreover when the marginal random variables of $\mathbf{X}$ are absolutely continuous this copula C is unique. We shall assume that the marginals of $\mathbf{X}$ are absolutely continuous. The aggregation function $\Psi : \mathbb{R}^d \to \mathbb{R}$ is considered to be measurable and non-decreasing on each variable. Let us denote by $F_\Psi(x) = \mathbb{P}(\Psi(\mathbf{X}) \leq x)$. Of course, the copula of the vector $\mathbf{X}$ determines the distribution of $\Psi(\mathbf{X})$. Nevertheless, the copula specification may be redundant, as for any copula C there may exist infinite set of copulas $\mathcal{C}_{\Psi,C}$ such that $\Psi(\mathbf{X}^C) =^d \Psi(\mathbf{X}^{C'})$ for any $C' \in \mathcal{C}_{\Psi,C}$, where $\mathbf{X}^C$ denotes a random vector with same marginals as $\mathbf{X}$ with copula C.

The Fréchet class of the marginal distributions $F_1, \ldots, F_d$, denoted by $\mathfrak{F}_d(F_1, F_2, \ldots, F_d)$ consists of all d-multivariate distributions with $F_1, \ldots, F_d$ as marginals. This class is completely determined by the class of all d-copulas, i.e.:

$$\mathfrak{F}_d(F_1, \ldots, F_d) = \{F : F(x_1, \ldots, x_d) = C\left(F_1(x_1), \ldots, F_d(x_d)\right)\}.$$

Moreover, when the marginals are absolutely continuous there is a bijective correspondence between both classes.

In the context of risk aggregation the following more useful class has been introduced in Bernard et al. (2014).

Definition 2.1 *An aggregate risk S is called an admissible risk of marginal distributions $F_1, \ldots, F_d$ if it can be written as $S = \Psi(X_1, \ldots, X_d)$ where $X_i \sim F_i$ for $i = 1, \ldots, d$. The admissible risk class is defined by the set of admissible risks of given marginal distributions:*

$$\mathfrak{S}_d(F_1, \ldots, F_d, \Psi) = \{\Psi(X_1, \ldots, X_d) : X_i \sim F_i, i = 1, \ldots, d\}.$$

Some interesting properties of this class have been presented in Bernard et al. (2014) when Ψ is the sum. Here we present a related class from the copula point of view and with more general aggregation functions when possible.

Definition 2.2 *Let $\boldsymbol{X}$ be a random vector and Ψ an aggregation function. The class of copulas*

$$\mathcal{C}(\boldsymbol{X}, \Psi) = \{C \in \mathcal{C} : \Psi(\boldsymbol{X}^C) =^d \Psi(\boldsymbol{X})\}$$

is the invariant aggregation copula class of $\boldsymbol{X}$ and Ψ.

The invariant aggregation class is related to the set of admissible risk, in a similar way as the copulas are related to the Fréchet class:

$$F \in \mathfrak{F}_d(F_1, \ldots, F_d) \Leftrightarrow \exists C : F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d))$$
$$S \in \mathfrak{S}_d(X_1, \ldots, X_d, \Psi) \Leftrightarrow \exists \mathcal{C}_{X,\Psi} : \forall C \in \mathcal{C}_{X,\Psi} \quad S =^d \Psi(X^C),$$

We present some examples and results that show explicitly that this class is not trivial.

Example 1 *We construct explicitly two different random vectors (X, Y) and (X', Y') such that $X =^d X'$, $Y =^d Y'$ and $X + Y =^d X' + Y$. Let (X, Y) be any random vector in $\mathbb{R}^2$ with density f. Suppose that for some $\epsilon > 0$ and some $a < b$ and $c < d$ with $b - a = d - c$ we have that $f(x, y) > \epsilon$ for any $(x, y) \in [a, b] \times [c, d]$. The equality $\int_0^{2\pi} \int_0^{2\pi} \sin(x - y)\, dxdy = 0$ implies that*

$$g(x, y) = f(x, y) + \epsilon \sin\left(2\pi\frac{x - a}{b - a} - 2\pi\frac{y - c}{d - c}\right) I_{[a,b]\times[c,d]}(x, y)$$

is a density function. Moreover, as for any t the following equations hold,

$$\int_0^t \int_0^{2\pi} \sin(x - y)\, dxdy = 0 \quad \textit{and} \quad \int_0^{2\pi} \int_0^t \sin(x - y)\, dxdy = 0,$$

then the marginal densities of f and g are identical. Furthermore, it can also be checked easily that $\int_0^{2\pi} \int_0^{2\pi} \sin(x-y) I_{\{0\leq x+y\leq t\}}(x, y)\, dxdy = 0$ for any $t > 0$, thus if (X', Y') is a random vector with density g, it satisfies that $X' + Y' =^d X + Y$.

The example above may be generalized in any dimension.

Proposition 2.1 *If X admits a density then $\mathcal{C}_{X,+}$ has infinite elements.*

By definition, any element of the class $\mathcal{C}_{X,\Psi}$ characterizes $\Psi(\mathbf{X})$. The following result shows that in some cases we can always find a symmetrical copula in $\mathcal{C}_{X,\Psi}$.

Proposition 2.2 *If X admits a density, its marginals are identical and Ψ is a symmetrical aggregation function then there exists a symmetrical copula C such that $\Psi(\boldsymbol{X}) =^d \Psi(\boldsymbol{X}^C)$.*

Proof. Let $f(x_1, \ldots, x_d)$ be the density of $\mathbf{X}$. Define $g(x_1, \ldots, x_d)$ as

$$g(x_1, \ldots, x_d) = \frac{1}{d!} \sum_{\sigma \in S_d} f(x_{\sigma(1)}, \ldots, x_{\sigma(d)}),$$

where S_d is the set of all the permutations of $\{1 \ldots, d\}$. Let $\mathbf{X}'$ be a random vector with density g. Then it is easy to check that the marginals of $\mathbf{X}'$ are distributed as the marginals of $\mathbf{X}$. It follows equally, from the symmetry of Ψ, that $\Psi(\mathbf{X}) =^d \Psi(\mathbf{X}')$. As the density of $\mathbf{X}'$ is completely symmetrical so is its copula. ■

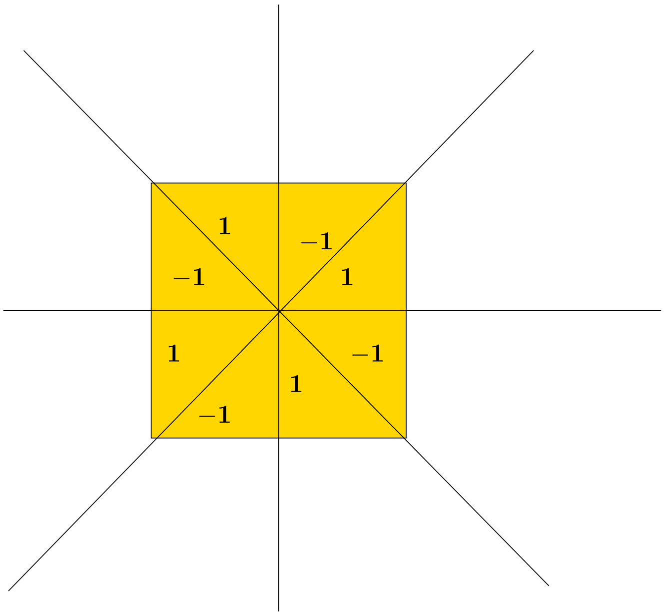
Remark 2.1 *In the case of d dimensional Archimedean copulas, it is known that the copula C is uniquely determined by its diagonal δ, $\delta(t) = C(t, \ldots, t)$ if $\delta'(1-) = d$ (see Frank (1996), Sungur and Yang (1996) in dimension 2 and Erdely et al. (2014) in higher dimension). This means that if $\Psi = \max$ or $\Psi = \min$, given a fixed common law for $X_1, \ldots, X_d$ and a fixed law for $\Psi(\boldsymbol{X})$, then there is only one Archimedean copula which leaves $\Psi(\boldsymbol{X})$ and the marginal laws invariant. Nevertheless, using constructions in Nelsen et al. (2008), infinitely many copulas with a fixed diagonal may be constructed.*

Below we provide a construction of infinitely many laws of random vectors with a fixed law for their $\max$ and fixed marginal laws (remark that if the marginal laws are not the same, then the law of the max is not determined by the diagonal of the copula).

Proposition 2.3 *Assume that X is absolutely continuous with density f such that $\inf_K f > 0$ with $K = \prod_{i=1}^{d} [a_i, b_i]$. If K is symmetric with respect to the diagonal, then there exists φ a density function such that $f \equiv \varphi$ outside K, $f \neq \varphi$ on K and the random vector $\widetilde{\boldsymbol{X}}$ whose density function is φ is such that*

- *for $i = 1, \ldots, d$, $\widetilde{X}_i \overset{\mathcal{L}}{=} X_i$,*
- $\Psi(\boldsymbol{X}) \overset{\mathcal{L}}{=} \Psi(\widetilde{\boldsymbol{X}})$.

Proof. We sketch the proof in dimension 2. Let $f > \varepsilon$ on K and $\varphi = f + \varepsilon\gamma$ where γ has its support in K as shown below:



It is easy to verify that the random vector $\widetilde{\mathbf{X}}$ whose density is φ has the same marginal laws as $\mathbf{X}$ and $\max(\widetilde{\mathbf{X}}) \overset{\mathcal{L}}{=} \max(\mathbf{X})$. ■

Even if the example above seems trivial it shows how unnecessary is the full knowledge of the copula distribution when studying an aggregation.

3. NON-PARAMETRIC ESTIMATION OF THE AGGREGATION DISTRIBUTION, WHEN THE MARGINALS ARE KNOWNS

We have seen in the above section that the exact copula estimation can be considered as a redundant exercise when estimating the distribution of an aggregation of $\mathbf{X}$. The information given by the copula of $\mathbf{X}$ is redundant and it really suffices to pick any copula from the class $\mathcal{C}_{X,\Psi}$. This may be seen as a justification of the fact that when the marginals are known, there is some flexibility in the copula estimation in order to estimate the aggregated distribution. Here we propose a non-parametric approach, when choosing the copula from data when the marginals are known, and we compare it with the classical empirical estimation that does not integrate the marginal knowledge.

3.1. Checkerboard and Empirical Copulas

In this section we propose a non-parametric estimator of the distribution of $\Psi(\mathbf{X})$ when marginals $F_1, \ldots, F_d$ are known and an independent and identically distributed (i.i.d.) sample $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)}$ is given. We compare its behaviour with the classical empirical cumulative distribution function

$$\widehat{F_\Psi}(t) = \frac{1}{n} \sum_{i=1}^{n} 1\{\Psi(\mathbf{X}^{(i)}) \leq t\},$$

where no marginal information is used.

Let F be the cumulative distribution function (c.d.f.) of $\mathbf{X}$. Then $F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d))$ where C is a copula function and F_i is the c.d.f of X_i, $i = 1 \ldots, d$. Let μ_C be the probability measure associated with C, i.e such that $\mu_C([0, x]) = C(x)$ for any $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$. By a μ-decomposition of a set $A \subset \mathbb{R}^d$ we mean a finite family of subsets $\{A_i \subset A\}$ such that

1. $\mu(A_i \cap A_j) = 0$ whenever $i \neq j$;
2. $\sum_i \mu(A_i) = \mu(A)$.

Definition 3.1 *A copula C^* is a checkerboard copula if there exists a λ-decomposition $\mathcal{A} = \{(a_i, b_i)\}$ of I^d, the d-dimensional unit cube, made out of d-intervals such that for all i*

1. $\mu_{C^*}((a_i, b_i)) = \mu_C((a_i, b_i))$*;*
2. μ_{C^*} *is uniform on* (a_i, b_i)*.*

If C is any copula such that $\mu_{C^*}(A) = \mu_C(A)$ for any $A \in \mathcal{A}$, then we say that C^* is a checkerboard approximation of C (Mikusinski, 2010 (Mikusinski and Taylor (2010))).
For $m \in \mathbb{N}$, let us consider the partition (modulo a 0 measure set) of $[0,1]^d$ given by the m^d squares:

$$I_{i,m} = \prod_{j=1}^{d} \left[\frac{i_j - 1}{m}, \frac{i_j}{m}\right], \; i = (i_1, \ldots, i_d), \; i_j \in \{1, \ldots, m\}.$$

We shall denote by C^*_m the checkerboard copula associated with this partition. The definition of the checkerboard copula may be rewritten as:

$$C^*_m(x) = \sum_i m^d \mu(I_{i,m}) \lambda([0, x] \cap I_{i,m})$$

where $[0, x] = \prod_{i=1}^d [0, x_i]$, for $x = (x_1, \ldots, x_d) \in [0,1]^d$ and λ is the d-dimensional Lebesgue measure.

The checkerboard copula is defined throughout a probability measure; in order to verify that it is a copula, it is sufficient to verify that for $x = (x_1, \ldots, x_d) \in [0,1]^d$ with $x_j = 1$ if $j \neq k$, $C^*_m(x) = x_k$. This is a simple computation (see Mikusinski and Taylor (2010) for more details and for various results of convergence of C^*_m to C). The following proposition is a simple computation.

Proposition 3.1 *Let C^*_m be the checkerboard copula defined above. We have:*

$$\sup_{x \in [0,1]^d} |C^*_m(x) - C(x)| \leq \frac{d}{2m}.$$

We also define two kind of *checkerboard copula with additional information*. First of all, we consider the case where the distribution of a subvector $\mathbf{X}^J = (X_i)_{i \in J}$, $J \subset \{1, \ldots, d\}$, is known, $|J| = k < d$. Denote C^J the copula of $\mathbf{X}^J$. Let μ^J be the probability measure on $[0,1]^k$ associated with C^J. For $i = (i_1, \ldots, i_d)$, let $x = (x_1, \ldots, x_d) \in [0,1]^d$, $x^J = (x_j)_{j \in J}$, $x^{-J} = (x_j)_{j \notin J}$ and

$$I^J_{i,m} = \left\{ x \in [0,1]^d \;/\; x_j \in \left[\frac{i_j - 1}{m}, \frac{i_j}{m}\right], \; j \in J \right\},$$

$$I_{i,m}^{-J} = \left\{ x \in [0,1]^d \;/\; x_j \in \left[\frac{i_j - 1}{m}, \frac{i_j}{m} \right], \; j \notin J \right\}.$$

The checkerboard copula with information on $\mathbf{X}^J$ is defined below.

Definition 3.2 *Consider the probability measure on $[0,1]^d$ defined by*

$$\mu_m^J([0,x]) = \sum_{i \subset \{1,\ldots,d\}} \frac{m^{d-k}}{\mu^J(I_{i,m}^J)} \mu(I_{i,m}) \lambda([0,x^{-J}] \cap I_{i,m}^{-J}) \mu^J([0,x^J] \cap I_{i,m}^J).$$

Let C_m^J, the checkerboard copula with additional information on $\mathbf{X}^J$, be defined by $C_m^J(x) = \mu_m^J([0,x])$.

Proposition 3.2 *C_m^J is a copula, it approximates C:*

$$\sup_{x \in [0,1]^d} |C_m^J(x) - C(x)| \le \frac{d}{2m}.$$

If $\mathbf{X}^J$ and $\mathbf{X}^{-J}$ are independent then,

$$\sup_{x \in [0,1]^d} |C_m^J(x) - C(x)| \le \frac{d-k}{2m}.$$

We may also add information on the tail and so define the following particular checkerboard copula.

Definition 3.3 *For $t \in]0,1[$, let $E = \left(\prod_{i=1}^d [0,t]^d \right)^c$. We assume that $\mu_C(E)$ is known. Consider the partition of the cube given by $J_{i,m} = t \cdot I_{n,i}^d$ for d-tuple $i = (i_1, \ldots, i_d)$ in $\{0, 1/n, \ldots, (n-1)/n\}^d$. Define $C_m^{\mathcal{E}}$ as the checkerboard copula associated with the partition $(J_{i,m}), E$, that is*

$$C_m^{\mathcal{E}}(x) = \mu_C(E^c) \mu_m^*([0,x]/t) \mathbf{1}_{E^c}(x) + \frac{\mu_C(E)}{\lambda(E)} \lambda([0,x] \cap E).$$

This is the checkerboard copula with extra information on the tail.

The empirical copula, introduced by Deheuvels (1979), may be used to estimate non parametrically the copula.

Definition 3.4 *Let $\mathbf{X}^{(1)}, \ldots \mathbf{X}^{(n)}$ be n independent copies of $\mathbf{X}$ and $R_i^{(1)}, \ldots, R_i^{(n)}$, $i = 1, \ldots, d$ their marginals ranks, i.e.,*

$$R_i^{(j)} = \sum_{k=1}^n 1\{X_i^{(j)} \ge X_i^{(k)}\}, \; i = 1, \ldots, d, \; j = 1, \ldots, n.$$

The empirical copula $\widehat{C}_n$ of $\mathbf{X}^{(1)}, \ldots \mathbf{X}^{(n)}$ is defined as

$$\widehat{C}_n(u_1, \ldots, u_d) = \frac{1}{n} \sum_{k=1}^n 1 \left\{ \frac{1}{n} R_1^{(k)} \le u_1, \ldots, \frac{1}{n} R_d^{(k)} \le u_d \right\}.$$

Using the empirical copula $\widehat{C}_n$ and the empirical probability measure $\widehat{\mu}$ associated with $\widehat{C}_n$, we may define empirical versions of the checkerboard copulas defined above.

Definition 3.5 *Let $\boldsymbol{X}^{(1)}, \dots \boldsymbol{X}^{(n)}$ be n independent copies of $\boldsymbol{X}$.*

- *The empirical checkerboard copula $\widehat{C}^*_m$ is defined by*

$$\widehat{C}^*_m(x) = \sum_i m^d \widehat{\mu}(I_{i,m})\lambda([0,x] \cap I_{i,m}).$$

- *The empirical checkerboard copula with information on a sub-vector $\boldsymbol{X}^J$ is defined by*

$$\widehat{C}^J_m(x) = \sum_{i \subset \{1,\dots,d\}} \frac{m^{d-k}}{\mu^J(I^J_{i,m})} \widehat{\mu}(I_{i,m})\lambda([0,x^{-J}] \cap I^{-J}_{i,m})\mu^J([0,x^J] \cap I^J_{i,m}).$$

- *The empirical checkerboard copula with information on the tail is defined by:*

$$\widehat{C}^{\mathcal{E}}_m(x) = \mu_C(E^c)\widehat{C}^*_m(x/t)\mathbf{1}_{E^c}(x) + \frac{\mu_C(E)}{\lambda(E)}\lambda([0,x] \cap E).$$

In what follows $\widehat{C}^{cb}_m$ denotes any of the three empirical checkerboard copulas defined above. We propose the following estimation procedure.
Assume the marginal laws are known and a (quite small) sample of size n of $\mathbf{X}$ is available.

1. Estimate μ by $\widehat{\mu}$ using the empirical copula.
2. Obtain the empirical checkerboard copula $\widehat{C}^{cb}_m$ (depending if some additional information is known).
3. Simulate a sample of size N from the copula $\widehat{C}^{cb}_m$ for N big:

$$(u_1^{(1)}, \dots, u_d^{(1)}), \dots, (u_1^{(N)}, \dots, u_d^{(N)})$$

4. Get a sample of $\Psi(\mathbf{X})$ using the marginals to transform the above sample:

$$\Psi\left(F_1^{\leftarrow}(u_1^{(1)}), \dots, F_d^{\leftarrow}(u_d^{(1)})\right), \dots, \Psi\left(F_1^{\leftarrow}(u_1^{(N)}), \dots, F_d^{\leftarrow}(u_d^{(N)})\right)$$

5. Estimate the distribution function F_Ψ of $\Psi(\mathbf{X})$ empirically using the above sample: $\widehat{F}(\Psi)$ is the empirical distribution function from the sample above.

Proposition 3.3 *For some $A > 0$, let $A\sqrt{n} \leq m \leq n$. Assume that $\Psi(\boldsymbol{X})$ is absolutely continuous and C has continuous partial derivatives*

$$\|F_\Psi - \widehat{F}(\Psi)\|_\infty = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right).$$

Proof. Use the convergence result by Fermanian et al. (2004). ∎

3.2. Numerical Application

In this section we use the estimator of the distribution function G_Ψ to estimate the $\mathrm{VaR}_p(S)$ for $S = X_1 + \cdots + X_d$ at different confidence levels $0 < p < 1$. We will consider the Pareto–Clayton model because, in that case, the exact value of $\mathrm{VaR}_p(S)$ is known.

3.2.1. The Pareto–Clayton model

We consider $\mathbf{X} = (X_1, \dots, X_d)$ such that:

$$\mathbb{P}(X_1 > x_1, \dots, X_d > x_d \,|\, \Lambda = \lambda) = \prod_{i=1}^{d} e^{-\lambda x_i},$$

that is, conditionally on the value of Λ the marginals of $\mathbf{X}$ are independent and exponentially distributed.
If Λ is Gamma distributed, then theX_i's are Pareto distributed with dependence given by a survival Clayton copula.
If Λ is Levy distributed, then the X_i's are Weibull distributed with a Gumbel survival copula.
These models have been studied by Oakes (1989), Yeh (2007). In the context of multivariate risk theory, they have been used e.g. in Maume-Deschamps et al. (2014) and Dacorogna et al. (2014).
In what follows, we consider that $\Lambda \rightsquigarrow \Gamma(\alpha, \beta)$, so that the X_i's are Pareto (α, β) distributed and the dependence structure is described by a survival Clayton copula with parameter $1/\alpha$. In Dubey (1970), it is shown that S follows the so-called Beta prime distribution:

$$F_S(x) = F_\beta\left(\frac{x}{1+x}\right)$$

where F_β is the distribution function of the Beta$(d\beta, \alpha)$ distribution. The inverse of F_S (or VaR function of S) can also be expressed in function of the inverse of the Beta distribution

$$F_S^{\leftarrow}(p) = \frac{F_\beta^{\leftarrow}(p)}{1 - F_\beta^{\leftarrow}(p)}.$$

From these results (see also Cuberos et al. (2015)), we may compute $\mathrm{VaR}_\alpha(S)$.

3.2.2. Simulation study in dimension 2

We consider a Pareto–Clayton model in dimension 2, with $\beta = 1$ and $\alpha = 1$. The multivariate sample is of size $n = 30$, we perform $N = 1000$ runs. Table 1 below compares the method using the checkerboard copula with the direct estimation from the empirical distribution of S. We give the real value, the mean estimation and the relative standard deviation (in $\%$ age). On this example, the checkerboard method seems performant, especially, it is much more stable than the empirical one for high level quantiles.

In Table 2 below, we have performed the checkerboard estimation with information in the tail, for various values of t. We remark that injecting information in the tail improves the estimation (mean and stability), especially for VaR levels higher than t.

	VaR 80%	VaR 90%	VaR 95%	VaR 99%	VaR 99.5%	VaR 99.9%
Exact value	8.5	18.5	38.5	198.5	398.5	1998.5
Empirical	9	19.5	45	293.5	341.4	379.7
	(40%)	(55%)	(105%)	(547%)	(566%)	(717%)
Checkerboard	8.8	19.4	41.8	224.5	432.6	2049.7
	(19%)	(23%)	(26%)	(10%)	(11%)	(11%)

Table 1: Mean and relative deviation for the Pareto–Clayton sum in dimension 2.

	VaR 80%	VaR 90%	VaR 95%	VaR 99%	VaR 99.5%	VaR 99.9%
Exact value	8.5	18.5	38.5	198.5	398.5	1998.5
$t = 1$ (no information)	8.8	19.4	41.8	224.5	432.6	2049.7
	(19%)	(23%)	(26%)	(10%)	(11%)	(11%)
$t = 0.995$	8.9	20.1	39.9	185.6	401.9	2212
	(18%)	(22%)	(23%)	(12%)	(4%)	(9%)
$t = 0.99$	8.9	19.8	38.5	202.6	456.3	2134.1
	(18%)	(21%)	(21%)	(3%)	(3%)	(9%)
$t = 0.95$	8.8	21.4	41.1	220.8	427.2	2047.1
	(15%)	(9%)	(1%)	(3%)	(4%)	(10%)

Table 2: Mean and relative deviation for the Pareto–Clayton sum in dimension 2, using information in the tail.

3.2.3. Simulation study in higher dimension

We conclude with a simulation in dimension 10. We consider a Clayton–Pareto model with $\beta = 1$, $\alpha = \frac{1}{2}$. The multivariate sample size is $n = 75$, then $n = 175$. We perform $N = 1000$ runs for the checkerboard method, see Table 3. As in dimension 2, we remark that the checkerboard method performs well.

Finally, some technical simulation issues have to be treated in order to perform simulations with informations in the tail in dimension higher than 2 as well as simulations including informations on a sub-vector.

	VaR 80%	VaR 90%	VaR 95%	VaR 99%	VaR 99.5%	VaR 99.9%
Exact value	12.2	19.2	29	70.1	100.8	230.5
Empirical, $n = 75$	12.6	20	29.9	62.2	75.8	86.7
	(12%)	(15%)	(19%)	(39%)	(58%)	(71%)
Checkerboard, $n = 75$	12.5	20.1	31.2	74.8	92.4	152.6
	(10%)	(13%)	(14%)	(20%)	(20%)	(16%)
Empirical, $n = 150$	12.4	19.6	30.3	67.3	89.9	121
	(8%)	(11%)	(14%)	(27%)	(38%)	(59%)
Checkerboard, $n = 150$	12.4	19.6	29.8	75.4	107.6	173.9
	(7%)	(9%)	(12%)	(16%)	(21%)	(19%)

Table 3: Mean and relative deviation for the Pareto–Clayton sum in dimension 10.

References

C. Bernard and S. Vanduffel. A new approach to assessing model risk in high dimensions. *Journal of Banking and Finance*, 58:166–178, 2015.

C. Bernard, L. Rüschendorf, and S. Vanduffel. Value-at-Risk bounds with variance constraints. preprint, 2013.

C. Bernard, X. Jiang, and R. Wang. Risk aggregation with dependence uncertainty. *Insurance: Mathematics and Economics*, 54(1):93–108, 2014.

H. Cossette, M.-P. Côté, M. Mailhot, and E. Marceau. A note on the computation of sharp numerical bounds for the distribution of the sum, product or ratio of dependent risks. *Journal of Multivariate Analysis*, 130:1–20, 2014.

A. Cuberos, E. Masiello, and V. Maume-Deschamps. High level quantile approximations of sums of risks. hal-01020597v2, 2015.

M. Dacorogna, L. El Bahtouri, and M. Kratz. Explicit diversification benefit for dependent risks. preprint, 2014.

P. Deheuvels. La fonction de dépendance empirique et ses propriétés. *Acad. Roy. Belg. Bull. Cl. Sci.*, 65(5):274–292, 1979.

S.D. Dubey. Compound gamma, beta and F distributions. *Metrika*, 16(1):27–31, 1970.

P. Embrechts and G. Puccetti. Risk aggregation. In P. Jaworski, F. Durante, W. Haerdle, and T. Rychlik, editors, *Copula Theory and Its Applications*, Lecture Notes in Statistics - Proceedings 198, pages 111–126. Springer Berlin/Heidelberg, 2010.

P. Embrechts, G. Puccetti, and L. Rüschendorf. Model uncertainty and VaR aggregation. *Journal of Banking & Finance*, 37(8):2750–2764, 2013.

A. Erdely, J.M. González-Barrios, and M.M. Hernández-Cedillo. Frank's condition for multivariate Archimedean copulas. *Fuzzy Sets and Systems*, 240:131–136, 2014.

J.-D. Fermanian, D. Radulovic, and M. Wegkamp. Weak convergence of empirical copula processes. *Bernoulli*, 10(5):847–960, 2004.

M.J. Frank. Diagonals of copulas and Schröder's equation. *Aequationes Mathematicae*, 51:150, 1996.

V. Maume-Deschamps, D. Rullière, and K. Saïd. On the optimal capital allocation by minimizing multivariate risk indicators. hal-01082559, 2014.

P. Mikusinski and M.D. Taylor. Some approximations of n-copulas. *Metrika*, 72(3):385–414, 2010.

R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, and M. Úbeda-Flores. On the construction of copulas and quasi-copulas with given diagonal sections. *Insurance: Mathematics and Economics*, 42(2):473–483, 2008.

D. Oakes. Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, 84(406):487–493, 1989.

L. Puccetti, G.and Rüschendorf. Computation of sharp bounds on the distribution of a function of dependent risks. *Journal of Computational and Applied Mathematics*, 236(7):1833–1840, 2012.

L. Rüschendorf. Random variables with maximum sums. *Advances in Applied Probability*, 14: 623–632, 1982.

L. Rüschendorf. Solution of a statistical optimization problem by rearrangement methods. *Metrika*, 30(1):55–61, 1983.

E.A. Sungur and Y. Yang. Diagonal copulas of Archimedean class. *Communications in Statistics-Theory and Methods*, 25(7):1659–1676, 1996.

H.-C. Yeh. The frailty and the Archimedean structure of the general multivariate Pareto distributions. *Bulletin Institute of Mathematics Acedemia Sinica*, 2(3):713–729, 2007.

CONTRIBUTED TALK

LOSS COVERAGE IN INSURANCE MARKETS: WHY ADVERSE SELECTION IS NOT ALWAYS A BAD THING

MingJie Hao, Pradip Tapadar and R. Guy Thomas

School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, CT2 7NF, UK
Email: mh586@kent.ac.uk, P.Tapadar@kent.ac.uk, R.G.Thomas@kent.ac.uk

Abstract

This paper investigates equilibrium in an insurance market where risk classification is restricted. Insurance demand is characterised by an iso-elastic function with a single elasticity parameter. We characterise the equilibrium by three quantities: equilibrium premium; level of adverse selection; and "loss coverage", defined as the expected population losses compensated by insurance. We find that equilibrium premium and adverse selection increase monotonically with demand elasticity, but loss coverage first increases and then decreases. We argue that loss coverage represents the efficacy of insurance for the whole population; and therefore, if demand elasticity is sufficiently low, adverse selection is not always a bad thing.

1. INTRODUCTION

Restrictions on insurance risk classification can lead to troublesome adverse selection. A simple version of the usual argument is as follows. If insurers cannot charge risk-differentiated premiums, more insurance is bought by higher risks and less insurance is bought by lower risks. This raises the equilibrium pooled price of insurance above a population-weighted average of true risk premiums. Also, since the number of higher risks is usually smaller than the number of lower risks, the total number of risks insured usually falls. This combination of a rise in price and fall in demand is usually portrayed as a bad outcome, both for insurers and for society.

However, it can be argued that from a social perspective, higher risks are those more in need of insurance. Also, the compensation provided by insurance for a wide class of risks appears to be widely regarded as a desirable objective, which public policymakers often seek to promote (for example by tax relief on premiums). Insurance of one higher risk contributes more in expectation to this objective than insurance of one lower risk. This suggests that public policymakers might welcome increased purchasing by higher risks, except for the usual story about adverse selection.

The usual story about adverse selection overlooks one point: with adverse selection, expected losses compensated by insurance can be *higher* than with no adverse selection. The rise in equilibrium price with adverse selection reflects a shift in coverage towards higher risks. From a public policymaker's viewpoint, this means that more of the "right" risks, i.e. those more likely to suffer loss, buy insurance. If the shift in coverage is large enough, it can more than outweigh the fall in numbers insured. This result of higher expected losses compensated by insurance, i.e. higher "loss coverage", might be regarded by a public policymaker as a better outcome for society than that obtained with no adverse selection.

Another way of putting this is that a public policymaker designing risk classification policies in the context of adverse selection normally faces a trade-off between insurance of the "right" risks (those more likely to suffer loss), and insurance of a larger number of risks. The optimal trade-off depends on demand elasticities in higher and lower risk-groups, and will normally involve at least some adverse selection. The concept of loss coverage quantifies this trade-off, and provides a metric for comparing the effects of different risk classification schemes.

2. MOTIVATING EXAMPLES

We now give three heuristic examples of insurance market equilibria to illustrate the concept of loss coverage and the possibility that loss coverage may be increased by some adverse selection.

Suppose that in a population of 1,000 risks, 16 losses are expected every year. There are two risk-groups. The high risk-group of 200 individuals has a probability of loss 4 times higher than those in the low risk-group. This is summarised in Table 1.

	Low risk-group	High risk-group	Aggregate
Risk	0.01	0.04	0.016
Total population	800	200	1000
Expected population losses	8	8	16
Break-even premiums (differentiated)	0.01	0.04	0.016
Numbers insured	400	100	500
Insured losses	4	4	8
Loss coverage			0.5

Table 1: Full risk classification with no adverse selection.

We further assume that the probability of loss is not altered by the purchase of insurance, i.e. there is no moral hazard. An individual's risk-group is fully observable to insurers and all insurers are required to use the same risk classification regime. The equilibrium, or "break-even", price of insurance is determined as the price at which insurers make zero profit.

Under our first risk classification regime, insurers operate full risk classification, charging actuarially fair premiums to members of each risk-group. We assume that the proportion of each risk-group which buys insurance under these conditions, i.e. the "fair-premium proportional demand", is 50%. Table 1 shows the outcome, which can be summarised as follows:

(a) There is no adverse selection, as premiums are actuarially fair and the demand is at the fair-premium proportional demand.

(b) Half the losses in the population are compensated by insurance. We heuristically characterise this as a "loss coverage" of 0.5.

Now suppose that a new risk classification regime is introduced, where insurers have to charge a single "pooled" price to members of both the low and high risk-groups. One possible outcome is shown in Table 2, which can be summarised as follows:

(a) The pooled premium of 0.02 at which insurers make zero profits is calculated as the demand-weighted average of the risk premiums: $(300 \times 0.01 + 150 \times 0.04)/450 = 0.02$.

(b) The pooled premium is expensive for low risks, so fewer of them buy insurance (300, compared with 400 before). The pooled premium is cheap for high risks, so more of them buy insurance (150, compared with 100 before). Because there are 4 times as many low risks as high risks in the population, the total number of policies sold falls (450, compared with 500 before).

(c) There is moderate adverse selection, as the break-even pooled premium exceeds population-weighted average risk and the aggregate demand has fallen.

(d) The resulting loss coverage is 0.5625. The shift in coverage towards high risks more than outweighs the fall in number of policies sold: 9 of the 16 losses (56%) in the population as a whole are now compensated by insurance (compared with 8 of 16 before).

	Low risk-group	High risk-group	Aggregate
Risk	0.01	0.04	0.016
Total population	800	200	1000
Expected population losses	8	8	16
Break-even premiums (pooled)	0.02	0.02	0.02
Numbers insured	300	150	450
Insured losses	3	6	9
Loss coverage			0.5625

Table 2: No risk classification leading to moderate adverse selection but higher loss coverage.

Another possible outcome under the restricted risk classification scheme, this time with more severe adverse selection, is shown in Table 3, which can be summarised as follows:

(a) The pooled premium of 0.02154 at which insurers make zero profits is calculated as the demand-weighted average of the risk premiums: $(200 \times 0.01 + 125 \times 0.04)/325 = 0.02154$.

(b) There is severe adverse selection, with further increase in pooled premium and significant fall in demand.

(c) The loss coverage is 0.4375. The shift in coverage towards high risks is insufficient to outweigh the fall in number of policies sold: 7 of the 16 losses (43.75%) in the population as a whole are now compensated by insurance (compared with 8 of 16 in Table 1, and 9 out of 16 in Table 2).

	Low risk-group	High risk-group	Aggregate
Risk	0.01	0.04	0.016
Total population	800	200	1000
Expected population losses	8	8	16
Break-even premiums (pooled)	0.02154	0.02154	0.02154
Numbers insured	200	125	325
Insured losses	2	5	7
Loss coverage			0.4375

Table 3: No risk classification leading to severe adverse selection and lower loss coverage.

Taking the three tables together, we can summarise by saying that compared with an initial position of no adverse selection (Table 1), moderate adverse selection leads to a higher fraction of the population's losses compensated by insurance (higher loss coverage) in Table 2; but too much adverse selection leads to a lower fraction of the population's losses compensated by insurance (lower loss coverage) in Table 3.

3. MODEL

Based on the motivation in the previous section, we now develop a model to analyse the impact of restricted risk classification on equilibrium premium, adverse selection and loss coverage. We first outline the model assumptions and define the underlying concepts.

3.1. Population Parameters

We assume that a population of risks can be divided into a high risk-group and a low risk-group, based on information which is fully observable by insurers. Let μ_1 and μ_2 be the underlying risks (probabilities of loss), i.e. $\mu_1 = E[\text{loss}|\text{low-risk group}], \mu_2 = E[\text{loss}|\text{high-risk group}]$. Let p_1 and p_2 be the respective population fractions, i.e. a risk chosen at random from the entire population has a probability of p_i of belonging to the risk-group $i = 1, 2$. For simplicity, we assume that all losses are of unit size. All quantities defined below are for a single risk sampled at random from the population (unless the context requires otherwise).

The expected loss is given by:

$$E[L] = \mu_1 p_1 + \mu_2 p_2, \tag{1}$$

where L is the loss for a risk chosen at random from the entire population.

Information on risk being freely available, insurers can distinguish between the two risk-groups accurately and charge premiums π_1 and π_2 for risks in risk-groups 1 and 2 respectively.

The expected insurance coverage is given by:

$$E[Q] = d(\mu_1, \pi_1)p_1 + d(\mu_2, \pi_2)p_2, \tag{2}$$

where $d(\mu_i, \pi_i)$ denotes the proportional demand for insurance for risk-group i at premium π_i, i.e. the probability that an individual selected at random from risk-group i buys insurance.

The expected premium is given by:

$$E[Q\pi] = d(\mu_1, \pi_1)p_1\pi_1 + d(\mu_2, \pi_2)p_2\pi_2, \tag{3}$$

where π is π_1 or π_2 with probability p_1 or p_2 respectively.

The expected insurance claim, i.e. the loss coverage, is given by:

$$\text{Loss coverage: } E[QL] = d(\mu_1, \pi_1)p_1\mu_1 + d(\mu_2, \pi_2)p_2\mu_2, \tag{4}$$

where we assume no moral hazard, i.e. purchase of insurance has no bearing on the risk. Loss coverage can also be thought of as *risk-weighted* insurance demand.

Finally, dividing Equation (4) by Equation (2) we obtain the expected claim per policy, say $\rho(\pi_1, \pi_2)$, which is given by:

$$\text{Expected claim per policy: } \rho(\pi_1, \pi_2) = \frac{E[QL]}{E[Q]} \tag{5}$$

3.2. Demand for insurance

In the previous section, we have introduced the concept of proportional demand for insurance, $d(\mu_i, \pi_i)$, when a premium π_i is charged for risk-group with true risk μ_i. In this section, we specify a functional form for $d(\mu_i, \pi_i)$ and its relevant properties.

De Jong and Ferris (2006) suggested axioms for an insurance demand function:

(a) $d(\mu_i, \pi_i)$ is a decreasing function of premium π_i for all risk-groups i;

(b) $d(\mu_1, \pi) < d(\mu_2, \pi)$, i.e. at a given premium π, the proportional demand is greater for the higher risk-group;

(c) $d(\mu_i, \pi_i)$ is a decreasing function of the premium loading π_i/μ_i; and

(d) for our model, where all losses are of unit size, we need to add $d(\mu_i, \pi_i) < 1$, i.e. the highest possible demand is when all members of the risk-group buy insurance.

These authors suggested a "flexible but practical" exponential-power demand function, and this approach was also followed by Thomas (2008, 2009). However the exponential-power function, whilst very flexible, is also rather intractable. In the present paper, we use a more tractable function which satisfies the axioms above and for which the price elasticity of demand λ, is a constant for all risk-groups, i.e.:

$$-\frac{\pi_i}{d(\mu_i,\pi_i)}\frac{\partial d(\mu_i,\pi_i)}{\partial \pi_i} = \lambda. \tag{6}$$

Solving Equation (6) leads to the "iso-elastic" demand function:

$$d(\mu_i,\pi_i) = \tau_i \left(\frac{\pi_i}{\mu_i}\right)^{-\lambda}, \tag{7}$$

where $\tau_i = d(\mu_i,\mu_i)$ is the "fair-premium demand" for insurance for risk-group i, that is the probability that a member sampled randomly from risk-group i buys insurance when premiums are actuarially fair.

The formula specifies demand as a function of the "premium loading" (π_i/μ_i). When the premium loading is high (insurance is expensive), demand is low, and vice versa. The λ parameter controls the shape of the demand curve. The "iso-elastic" terminology reflects that the price elasticity of demand is the same everywhere along the demand curve.

Clearly, iso-elastic demand functions satisfy axioms (a) and (c) of De Jong and Ferris (2006). Axioms (b) and (d) appear superficially to require conditions on the fair-premium demands τ_1 and τ_2. However, if we define *fair-premium demand-shares* α_1 and α_2 as:

$$\text{Fair-premium demand-share: } \alpha_i = \frac{\tau_i p_i}{\tau_1 p_1 + \tau_2 p_2}, \quad i = 1,2 \tag{8}$$

then it turns out that that α_1 and α_2 fully specify the population structure in the form required for our model. Since increasing the τ_i is mathematically equivalent to decreasing the p_i and vice versa, we do not need to specify any particular values for them. We can analyse the model for the full range of fair-premium demand-shares $0 < \alpha_i < 1$, knowing that for every α_i there must exist some corresponding combination of p_i and τ_i which satisfies the axioms (b) and (d) above.

4. EQUILIBRIUM

In the model in Section 3, an insurance market equilibrium is reached when the premiums charged (π_1,π_2) ensure that the expected profit, $f(\pi_1,\pi_2) = 0$, where:

$$\begin{aligned} f(\pi_1,\pi_2) &= E[Q\pi] - E[QL] && (9)\\ &= d(\mu_1,\pi_1)(\pi_1-\mu_1)p_1 + d(\mu_2,\pi_2)(\pi_2-\mu_2)p_2. && (10)\end{aligned}$$

4.1. Risk-differentiated Premiums

An obvious solution to the profit equation $f(\pi_1,\pi_2) = 0$ is to set $(\pi_1,\pi_2) = (\mu_1,\mu_2)$, i.e. setting premiums equal to the respective risks results in an expected profit of zero for each risk group and also in aggregate. We shall refer to this case as *risk-differentiated premiums*.

Following the notations introduced in Section 3, the expected insurance coverage is given by:

$$E[Q] = \tau_1 p_1 + \tau_2 p_2. \tag{11}$$

Also, $(\pi_1, \pi_2) = (\mu_1, \mu_2)$, i.e. the expected premium and expected claim are equal and given by:

$$E[Q\pi] = E[QL] = \tau_1 p_1 \mu_1 + \tau_2 p_2 \mu_2. \tag{12}$$

So, the expected claim per policy is:

$$\rho(\mu_1, \mu_2) = \frac{E[QL]}{E[Q]} = \frac{\tau_1 p_1 \mu_1 + \tau_2 p_2 \mu_2}{\tau_1 p_1 + \tau_2 p_2} = \alpha_1 \mu_1 + \alpha_2 \mu_2. \tag{13}$$

4.2. Pooled Premium

Next we consider the case of *pooled premium*. This is where risk classification is banned, so that insurers have to charge the same premium π_0 for both risk-groups, i.e. $\pi_1 = \pi_2 = \pi_0$, leading to $f(\pi_0, \pi_0) = 0$. For convenience, we omit one argument for all bivariate functions if both arguments are equal, e.g. we write $f(\pi)$ for $f(\pi, \pi)$.

Equation (9) leads to the following relationship for the equilibrium premium π_0:

$$\pi_0 = \frac{E[QL]}{E[Q]}. \tag{14}$$

The existence of a solution for $f(\pi) = 0$ within the interval $(\mu_1, \mu_2]$ is obvious, because $f(\pi)$ is a continuous function with $f(\mu_1) < 0$ and $f(\mu_2) \geq 0$. Assuming an iso-elastic demand function with constant elasticity of demand, λ, Equation (14) provides a unique solution:

$$\pi_0 = \frac{\alpha_1 \mu_1^{\lambda+1} + \alpha_2 \mu_2^{\lambda+1}}{\alpha_1 \mu_1^{\lambda} + \alpha_2 \mu_2^{\lambda}}. \tag{15}$$

This can be written as a weighted average of the true risks μ_1 and μ_2:

$$\pi_0 = v\mu_1 + (1 - v)\mu_2, \quad \text{where} \quad v = \frac{\alpha_1}{\alpha_1 + \alpha_2 \left(\frac{\mu_2}{\mu_1}\right)^{\lambda}}. \tag{16}$$

Note that π_0 does not depend directly on the individual values of the population fractions (p_1, p_2) and fair-premium demands (τ_1, τ_2), but only indirectly on these parameters through the demand-shares (α_1, α_2). So, populations with the same true risks (μ_1, μ_2) and demand-shares (α_1, α_2) have the same equilibrium premium, even if the underlying (p_1, p_2) and (τ_1, τ_2) are different.

Figure 1 shows the plots of pooled equilibrium premium against demand elasticity, λ, for two different population structures with the same true risks $(\mu_1, \mu_2) = (0.01, 0.04)$ but different fair-premium demand-shares (α_1, α_2). The following observations can be derived from Equations (15) and (16), and are illustrated by Figure 1:

(a) $\lim_{\lambda \to 0} \pi_0 = \alpha_1 \mu_1 + \alpha_2 \mu_2 = \rho(\mu_1, \mu_2)$. If demand is inelastic, changing premium has no impact on demand, and so the equilibrium premium will be the same as the expected claim per policy if risk-differentiated premiums were charged.

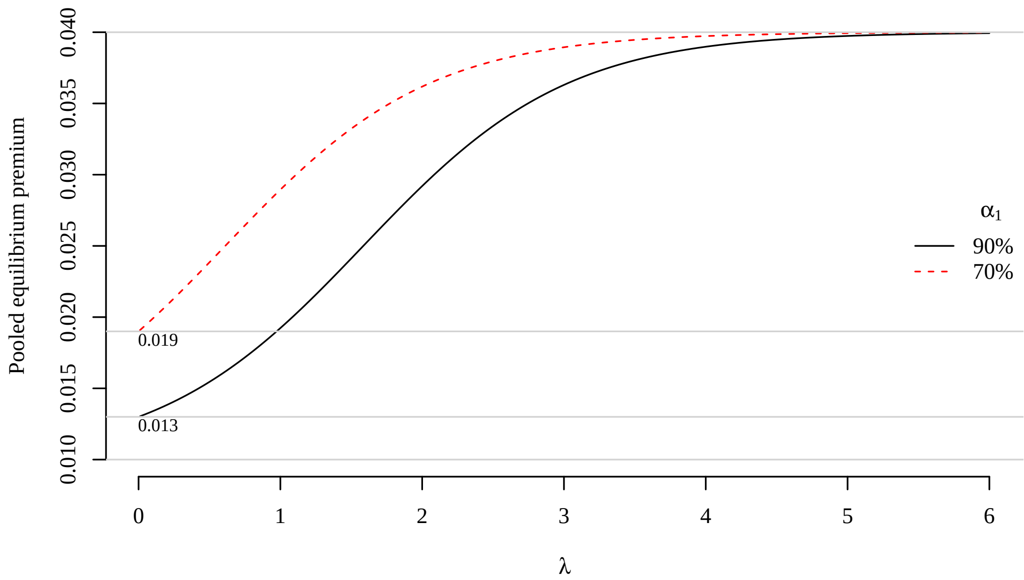


Figure 1: Pooled equilibrium premium as a function of λ for two populations with the same $(\mu_1, \mu_2) = (0.01, 0.04)$ but different values of α_1.

(b) π_0 is an increasing function of λ. In Equation (16), increasing λ reduces the weight v on low-risk, resulting in an increase in the equilibrium premium π_0.

(c) $\lim_{\lambda\to\infty} \pi_0 = \mu_2$. If demand elasticity is very high, demand from the low risk-group falls to zero for any premium above their true risk μ_1. The only remaining insureds are then all high risks, so the equilibrium premium must move to $\pi_0 = \mu_2$.

(d) π_0 is a decreasing function of α_1. If the fair-premium demand-share α_1 of the lower risk-group increases, we would expect the equilibrium premium to fall.

5. ADVERSE SELECTION

Adverse selection is typically defined in the economics literature as a positive correlation (or equivalently, covariance) of coverage and losses (e.g. for a survey see Cohen and Siegelman (2010)). Using the notations developed in Section 3, this can be quantified by the covariance between the random variables Q and L, i.e. $E[QL] - E[Q]E[L]$. We prefer to use the ratio rather than the difference, so our definition is:

$$\text{Adverse selection} = \frac{E[QL]}{E[Q]E[L]} = \frac{\rho(\pi_1, \pi_2)}{E[L]}, \tag{17}$$

where $\rho(\pi_1, \pi_2)$ is the expected claim per policy as defined in Equation (5). This metric for adverse selection is intuitively appealing: it is the the ratio of the *expected claim per policy* to the *expected* *loss per risk*, where the risk is drawn at random from the whole population.

To compare the severity of adverse selection under different risk classification regimes, we need to define a reference level of adverse selection. We use the level under risk-differentiated premiums, $\rho(\mu_1, \mu_2)/E[L]$, and so define the *adverse selection ratio* as:

$$\text{Adverse selection ratio: } S(\pi_1, \pi_2) = \frac{\rho(\pi_1, \pi_2)}{\rho(\mu_1, \mu_2)} = \frac{\rho(\pi_1, \pi_2)}{\alpha_1\mu_1 + \alpha_2\mu_2}. \tag{18}$$

Note that as the same underlying population is used in both the numerator and the denominator of the ratio, the population expected loss $E[L]$ gets cancelled and does not play any further role.

An interesting case arises for pooled equilibrium premium, where by Equation (14), an equilibrium premium π_0 satisfies the condition:

$$\pi_0 = \frac{E[QL]}{E[Q]} = \rho(\pi_0, \pi_0). \tag{19}$$

So in the particular case of pooled equilibrium premium, π_0, we have:

$$\text{Adverse selection ratio: } S(\pi_0) = \frac{\pi_0}{\alpha_1\mu_1 + \alpha_2\mu_2}. \tag{20}$$

In essence, the pooled equilibrium premium itself (scaled by the expected claim per policy under risk-differentiated premiums) provides a measure of the adverse selection.

Figure 2 shows the adverse selection ratio under pooling for two populations with the same underlying risks $(\mu_1, \mu_2) = (0.01, 0.04)$ and demand elasticity λ, but different values of the fair-premium demand-share α_1.

The following properties of the adverse selection ratio, $S(\pi_0)$, follow directly from the observations in Section 4:

(a) $S(\pi_0) \geq 1$, as the pooled equilibrium premium, π_0, is always higher than the expected claim per policy for risk-differentiated premiums.

(b) $S(\pi_0)$ is an increasing function of the underlying demand elasticity.

(c) $\lim_{\lambda\to\infty} S(\pi_0) = \frac{\mu_2}{\alpha_1\mu_1+\alpha_2\mu_2}$, i.e. when demand is very elastic, the adverse selection ratio tends towards a limit where only higher risks are insured.

The adverse selection ratio is always higher under pooling than under risk-differentiated premiums. It also increases monotonically with the underlying demand elasticity. Therefore this metric is unable to distinguish between cases where pooling gives a better outcome for society as a whole (Table 2 in the motivating examples in Section 2) and cases where pooling gives a worse outcome for society as a whole (Table 3 in the motivating examples in Section 2). This leads us to the concept of loss coverage ratio discussed in the next section.

6. LOSS COVERAGE

The motivating examples in Section 2 suggested loss coverage, heuristically characterised as the proportion of the population's losses compensated by insurance, as a measure of the social effi-

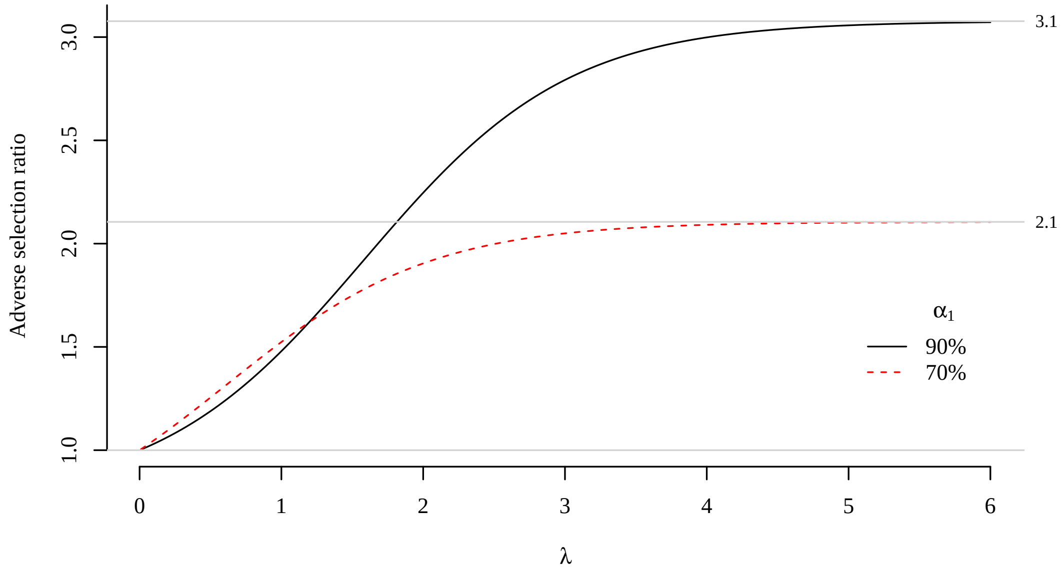


Figure 2: Adverse selection ratio as a function of λ for two populations with the same $(\mu_1, \mu_2) = (0.01, 0.04)$ but different values of α_1.

cacy of insurance. This can be formally quantified in our model by the expected insurance claim, $E[QL]$, as defined in Section 3 as:

$$\text{Loss coverage: } LC(\pi_1, \pi_2) = E[QL]. \tag{21}$$

To compare the relative merits of different risk classification regimes, we define a reference level of loss coverage. We use the level under risk-differentiated premiums (i.e. the same approach as for adverse selection in Equation (18)), and so define the *loss coverage ratio*, as follows:

$$\text{Loss coverage ratio: } C = \frac{LC(\pi_1, \pi_2)}{LC(\mu_1, \mu_2)}. \tag{22}$$

Considering loss coverage ratio for pooled premium, i.e $\pi_1 = \pi_2 = \pi_0$, and using the iso-elastic demand function with demand elasticity λ, in Equation (4), gives:

$$C(\lambda) = \frac{1}{\pi_0^\lambda} \frac{\alpha_1 \mu_1^{\lambda+1} + \alpha_2 \mu_2^{\lambda+1}}{\alpha_1 \mu_1 + \alpha_2 \mu_2}, \tag{23}$$

where π_0 is the pooled equilibrium premium given in Equation (15). The above can also be conveniently re-expressed as:

$$C(\lambda) = \frac{\left[w\beta^{1-\lambda} + (1-w)\right]^\lambda \left[w + (1-w)\beta^\lambda\right]^{1-\lambda}}{\beta^{\lambda(1-\lambda)}}, \text{ where} \tag{24}$$

$$w = \frac{\alpha_1 \mu_1}{\alpha_1 \mu_1 + \alpha_2 \mu_2} \quad \text{and} \quad \beta = \frac{\mu_2}{\mu_1} > 1.$$

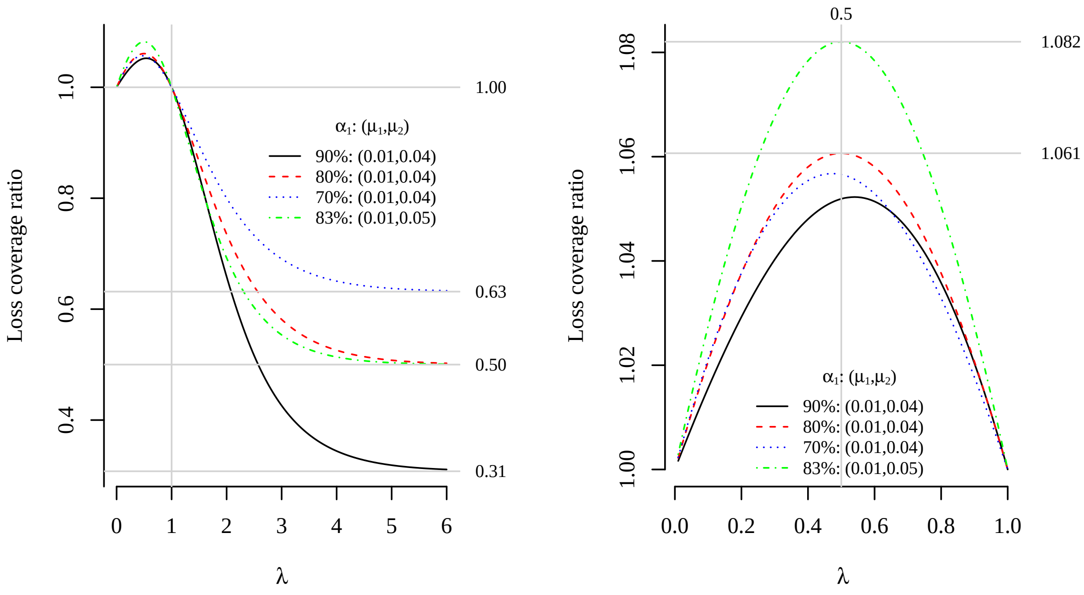


Figure 3: Loss coverage ratio as a function of λ for four population structures.

Figure 3 shows loss coverage ratio for four population structures. Both plots in Figure 3 show the same example, with the right-hand plot zooming over the range $0 < \lambda < 1$.

We make the following observations:

(a) $\lim_{\lambda \to 0} C(\lambda) = 1$. This follows directly from Equation (23). Intuitively, if demand is inelastic then pooling must give the same loss coverage as fair premiums.

(b) $\lim_{\lambda \to \infty} C(\lambda) = 1 - w = \frac{\alpha_2 \mu_2}{\alpha_1 \mu_1 + \alpha_2 \mu_2}$. This follows from Equation (24). Recall that for highly elastic demand, equilibrium is achieved when only high risks buy insurance at the equilibrium premium $\pi_0 = \mu_2$, which explains the above result.

(c) For $\lambda > 0$,

$$\lambda \lesseqgtr 1 \Rightarrow C(\lambda) \gtreqless 1. \tag{25}$$

The result implies that pooling produces higher loss coverage than fair premiums if demand elasticity is less than 1. (The proof of this result is beyond the scope of this paper.)

(d)

$$\max_{w,\lambda} C = \frac{1}{2}\left(\sqrt[4]{\frac{\mu_2}{\mu_1}} + \sqrt[4]{\frac{\mu_1}{\mu_2}}\right) = \frac{1}{2}\left(\sqrt[4]{\beta} + \frac{1}{\sqrt[4]{\beta}}\right). \tag{26}$$

As can be seen from the right-hand plot of Figure 3, for a given value of relative risk, β, loss coverage ratio attains its maximum when $\lambda = 0.5$ and $w = 0.5$. Moreover, the maximum loss coverage ratio increases with increasing relative risk. This implies that a pooled premium might be highly beneficial in the presence of a small group with very high risk exposure. (The proof of this result is beyond the scope of this paper.)

7. CONCLUSIONS

This paper has investigated insurance market equilibrium under restricted risk classification with iso-elastic demand. The equilibrium was characterised by three quantities: equilibrium premium, adverse selection, and loss coverage, defined as the expected losses compensated by insurance. We investigated how these quantities varied depending on the elasticity of demand for insurance, which was assumed to be equal for high and low risk-groups.

The equilibrium premium (and adverse selection) increases monotonically with demand elasticity. However, loss coverage ratio increases from 1, to a maximum for demand elasticity of around 0.5 and then decreases, falling back to 1 for demand elasticity of 1. So, restricting risk classification increases loss coverage if demand elasticity is less than 1. This is despite the fact that restricting risk classification will always increase adverse selection. In other words, the concept of loss coverage suggests that adverse selection is not always a bad thing.

References

A. Cohen and P. Siegelman. Testing for adverse selection in insurance markets. *The Journal of Risk and Insurance*, 77(1):39–84, 2010.

P. De Jong and S. Ferris. Adverse selection spirals. *ASTIN Bulletin*, 36:589–628, 2006.

R.G. Thomas. Loss coverage as a public policy objective for risk classification schemes. *Journal of Risk and Insurance*, 75:997–1018, 2008.

R.G. Thomas. Demand elasticity, risk classification and loss coverage: when can community rating work? *ASTIN Bulletin*, 39:403–428, 2009.

POSTER SESSION

MODELING LARGE CLAIMS WITH PARETO ARCTAN DISTRIBUTION

Enrique Calderín–Ojeda[†] **and Emilio Gómez–Déniz**[§]

[†]*Department of Economics, University of Melbourne, Australia,*
Level 4, FBE Building 105, 111 Barry Street, Melbourne, VIC 3010, Australia
[§] *Department of Quantitative Methods and TiDES Institute, ULPGC, Spain,*
Facultad de Economía, Empresa y Turismo, Campus Universitatio de Tafira, 35017 Las Palmas de Gran Canaria, Spain
Email: ecalderin@unimelb.edu.au, emilio.gomez-deniz@ulpgc.es

In this paper a new methodology to derive probabilistic models based on the circular $\tan^{-1}$ function is presented. This procedure generates a new survival function after incorporating an extra scale parameter α to a given parent survival function. The latter survival function is determined as limiting case when α tends to zero. By choosing as parent the classical Pareto survival function, the Pareto ArcTan (PAT) distribution is obtained. After providing a comprehensive analysis of its statistical properties, theoretical results with reference to insurance are illustrated. Its performance is compared by means of the well–known Norwegian fire insurance data with other existing heavy-tailed distributions in the literature such as Pareto, Stoppa and shifted lognormal distributions.

1. INTRODUCTION

Most of the methods for generalizing probability density functions in the statistical literature are based on the idea of incorporating a new parameter to a classical distribution (see Marshall and Olkin (1997)). The probabilistic models derived from this methodology usually exhibit more flexibility and they include classical distributions for particular values of the new parameter attached to the initial family.

In this work a new method to add a parameter to a family of distributions is proposed after making a change of variable in the truncated Cauchy distribution. As a result, a class of probabilistic models is obtained. In particular, if this methodology is applied to the Pareto distribution, the Pareto ArcTan (PAT) distribution is derived.

In general insurance, only a few large claims arising in the portfolio represent the largest part of the payments made by the insurer. The PAT distribution provides a more accurate description of large losses than other heavy-tailed distributions do. In this particular, the performance of this model, by means of the well–known Norwegian fire insurance data is compared to Pareto, shifted lognormal and Stoppa (see Kleiber and Kotz (2003)) distributions.

2. GENESIS AND PROPERTIES

The half–Cauchy distribution truncated at $\alpha > 0$ has pdf given by

$$f(y) = \frac{1}{\tan^{-1}\alpha}\frac{1}{1+y^2}, \quad 0 < y < \alpha. \tag{1}$$

In the latter expression, $\tan^{-1}$ is the inverse of the circular tangent function. Let us consider now the transformation $y = \alpha\bar{F}(x)$, where $\bar{F}(x)$ is the survival function of a random variable X with support in $[a, b]$ and where a and b can be finite or non–finite. Then, the corresponding pdf of the random variable X obtained from (1) results

$$f(x;\alpha) = \frac{1}{\tan^{-1}\alpha}\frac{\alpha f(x)}{1+[\alpha\bar{F}(x)]^2}, \tag{2}$$

for $a \le x \le b$ and $\alpha > 0$. The survival function of X, which is obtained from (2) by integration, is given by

$$\bar{F}(x;\alpha) = \frac{\tan^{-1}(\alpha\bar{F}(x))}{\tan^{-1}\alpha}. \tag{3}$$

Observe that when (3) is applied to the classical Pareto distribution with survival function given by $\bar{F}(x) = \left(\frac{\sigma}{x}\right)^{\theta}$, $x \ge \sigma$, $\sigma > 0$, the new survival function is

$$\bar{F}(x;\alpha) = \frac{\tan^{-1}(\alpha(\sigma/x)^{\theta})}{\tan^{-1}\alpha}, \quad x \ge \sigma. \tag{4}$$

Then, the corresponding pdf is

$$f(x;\alpha) = \frac{1}{\tan^{-1}\alpha}\frac{\alpha\sigma^{\theta}\theta x^{\theta-1}}{(\alpha\sigma^{\theta})^2 + x^{2\theta}}, \quad x \ge \sigma. \tag{5}$$

The expression (5) includes as particular case the Pareto distribution as limiting case when $\alpha \to 0$.

The r–th moment of the new distribution about zero. This is given by

$$E(X^r) = \frac{\alpha\theta\sigma^r}{(\theta - r)\tan^{-1}\alpha}\,{}_2F_1\left(1, \frac{\theta - r}{2\theta}; \frac{3\theta - r}{2\theta}; -\alpha^2\right), \quad \theta > r.$$

From (4) the quantile function x_γ is simply derived

$$x_\gamma = \sigma\left[\frac{1}{\alpha}\tan(\gamma\tan^{-1}\alpha)\right]^{-1/\theta}, \quad \text{with } 0 < \gamma < 1,$$

from which the median can be easily obtained.

Besides, the mode, which can be obtained by differentiating (5) with respect to the variable x, is given by

$$x_{Mo} = \left[\frac{(\theta - 1)(\sigma^{\theta}\,\alpha)^2}{1+\theta}\right]^{1/2\,\theta}.$$

The value at risk (VaR) is defined as the amount of capital required to ensure that the insurer does not became insolvent with a high degree of certainty. The VaR of a random variable X which follows the PAT distribution is the q quantile and it is given by

$$\text{VaR}(X;q) = \sigma\left(\frac{\alpha}{\tan((1-q)\tan^{-1}\alpha))}\right)^{1/\theta}.$$

The use of the VaR is questionable due to the lack of subadditivity, for that reason the expected loss given that the loss exceeds the q quantile of the distribution of X, the tail value at risk (TVaR), is considered. Then, if X follows a PAT distribution, for any quantile q the tail value at risk is given by

$$\begin{aligned}\text{TVaR}(X;q) =& \frac{1}{1-q}\int_q^1 \text{VaR}(x;q)\,dq \\ =& -\frac{\sigma\theta\alpha^{1/\theta}\tan^{\frac{\theta-1}{\theta}}\left(\tan^{-1}\alpha - q\tan^{-1}\alpha\right)}{(\theta-1)(q-1)\tan^{-1}(\alpha)} \\ & \times {}_2F_1\left(1, \frac{\theta-1}{2\theta}; \frac{3}{2}-\frac{1}{2\theta}; -\tan^2\left(\tan^{-1}\alpha - q\tan^{-1}\alpha\right)\right).\end{aligned}$$

3. NUMERICAL APPLICATIONS

The versatility of (5), as compared with different heavy–tailed distributions such as to Pareto, shifted lognormal and Stoppa, is proven by analyzing real actuarial loss data. This set of data describes a Norwegian fire insurance portfolio from 1989 to 1992 (see Beirlant et al. (2004)). This data set includes the claim value on 2,585 fire insurance losses in Norwegian Krone ($\times 1000$ NOK). A priority of 500 units was in force, thus no claims below this limit were recorded.

3.1. Estimation and Model Assessment

Table 1 provides different measures of goodness–of–fit based on information–criterion approach. The negative of the maximum of the log–likelihood (NLL) is exhibited in the second column of this Table.

Distribution	**NLL**	**AIC**	**BIC**	**CAIC**
Pareto	21058.50	42119.00	42124.85	42125.85
Sh. Logn	20993.60	41993.19	42010.76	42013.76
Stoppa	20984.56	41975.13	41992.70	41995.70
PAT	20935.66	41785.31	41887.03	41889.03

Table 1: Different measures of model assessment for Norwegian fire loss data.

Next the Akaike's Information Criteria (AIC, which is calculated by twice NLL plus twice the number of parameters), evaluated at the maximum likelihood estimates is shown in the third column; finally in the last two columns of this Table we give the Bayesian information criterion (BIC, which is obtained as twice the NLL at the estimates plus $k \ln(n)$, where k is the number of free parameters and n is the sample size) and the Consistent Akaike's Information Criteria (CAIC), a correction factor based on the sample size is used to compensate for the overestimating nature of AIC. It is defined as twice NLL plus $k(1 + \ln(n))$, again k is the number of free parameters and n refers to the sample size. A lower value of these measures is desirable. These results show that the PAT distribution provide a better fit than do the classical Pareto, shifted Lognormal and Stoppa distributions, even when some of these distributions use a larger number of parameters.

3.2. Point Estimation of High Quantiles

High quantiles of the distribution of the claim amounts has been traditionally considered as a measure that provides useful information for practitioners. Empirical and fitted models quantiles for the four models considered are displayed in Figure 1. The PAT distribution provides the best fit to data. The Pareto and Stoppa distribution tend to overestimate the extreme values where shifted lognormal model underestimates them.

3.3. Tail Value at Risk

In Table 2, the tail value at risk (TVaR), for empirical values and different security levels has been calculated for the models considered. This risk measure describes the expected loss given that the loss exceeds the security level (quantile).

Security Level	**Empirical**	**Pareto**	**Shifted lognormal**	**Stoppa**	**PAT**
0.90	9936.60	136861.19	7997.18	14796.88	10814.99
0.95	15635.29	267267.40	11331.77	24728.08	17231.80
0.99	42475.20	1266000.34	23237.94	81290.92	50793.41

Table 2: TVar for the different models considered.

4. CONCLUSIONS AND FORTHCOMING RESEARCH

In this research proposal a new mechanism to derive probability distributions by adding a parameter to a parent distribution function have been introduced. When the parent distribution is Pareto,

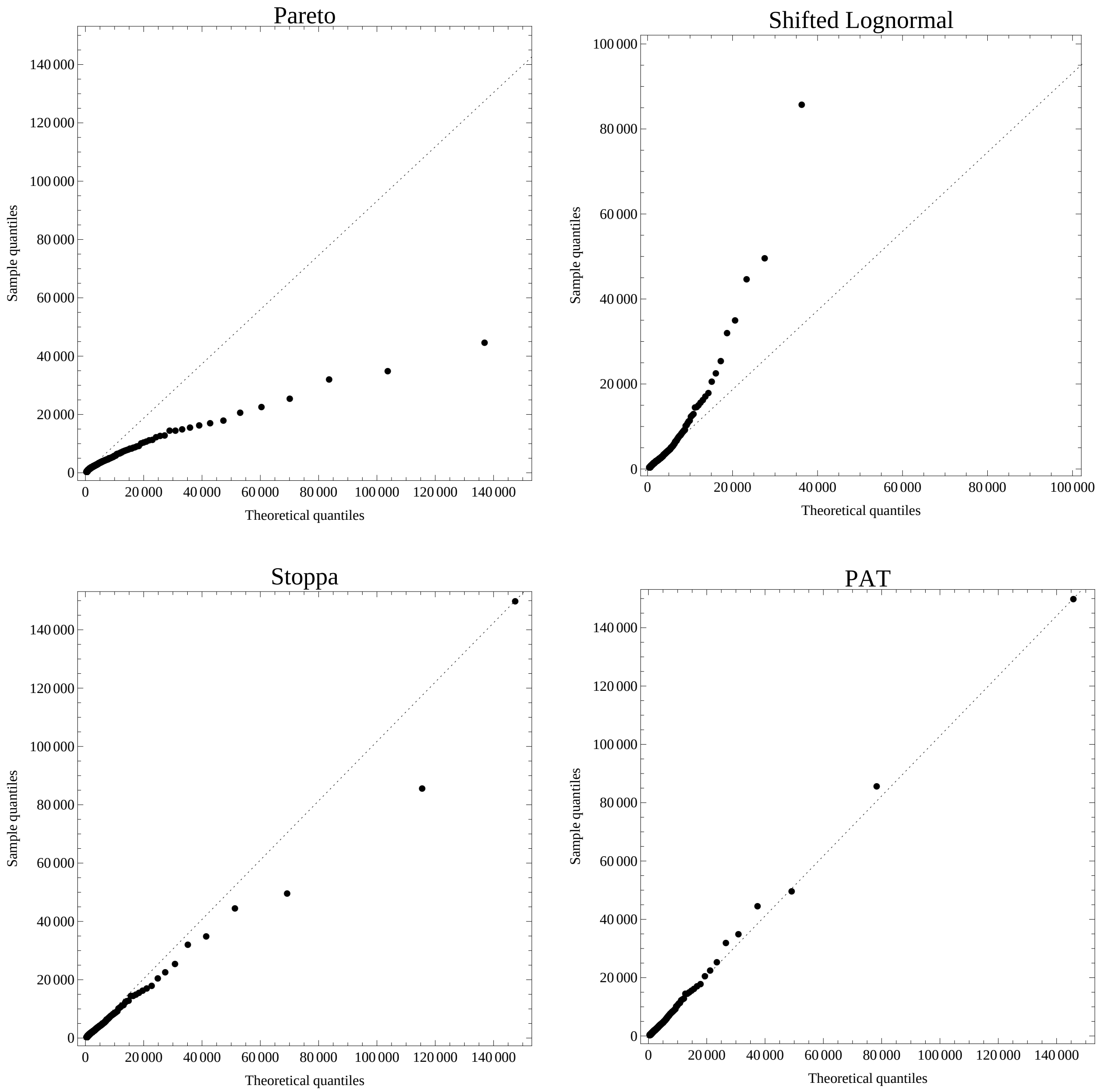


Figure 1: Q–Q plots for Norwegian fire loss data.

the Pareto ArcTan (PAT) is derived. The PAT model seems suitable for modeling payments that include a positive priority, with no claims below that threshold, and losses that combine data with high frequencies near the lower limit together with large upper tail derived from massive losses with low frequencies. Its performance has been proven using the well–known Norwegian insurance fire claim data. Numerical results illustrate the PAT distribution outperforms other existing heavy–tailed distributions for this set of data. Although the method has only been applied to the classical Pareto distribution, certainly this procedure can be extended by allowing the choice of other probabilistic families as parent distribution. Surely, more flexible distributions will be obtained.

Acknowledgements

Authors have been partially funded by grant ECO2013-47092 (Ministerio de Economía y Competitividad, Spain). EGD also acknowledges the Faculty of Business and Economics and the Centre of Actuarial Studies at the University of Melbourne for their special support, since part of this paper was written while EGD was visiting the University of Melbourne in July 2014.

References

J. Beirlant, Y. Goegebeur, J. Segers, and J. Teugels. *Statistics of Extremes: Theory and Applications*. Wiley, 2004.

C. Kleiber and S. Kotz. *Statistical Size Distributions in Economics and Actuarial Science*. John Wiley & Sons, 3rd edition, 2003.

A. W. Marshall and I. Olkin. A new method for adding a parameter to a family of distributions with application to the Exponential and Weibull families. *Biometrika*, 84(3):641–652, 1997.

DECOMPOSITION OF THE CONDITIONAL ASSET RETURN DISTRIBUTION: A BAYESIAN APPROACH

Evangelia N. Mitrodima, Jim E. Griffin and Jaideep S. Oberoi

School of Mathematics, Statistics & Actuarial Science, University of Kent,
Cornwallis Building, CT2 7NF, Canterbury, Kent, UK
Email: em260@kent.ac.uk, J.E.Griffin-28@kent.ac.uk, J.S.Oberoi@kent.ac.uk

We estimate the conditional asset return distribution by modelling a finite number of quantiles using Bayesian analysis. The motivation for this is to jointly incorporate time-varying dynamics of shape and scale of the asset return distribution in a robust manner. We also want to avoid any violations of the quantile orderings. Additionally, we want to address the challenges with reliably estimating such models. Thus, Bayesian analysis using Adaptive Markov Chain Monte Carlo (MCMC) methods is adopted.

1. INTRODUCTION

In this paper, we estimate jointly the scale and the shape of the conditional return distribution in a robust way. Our aim is to estimate the conditional distribution by using quantile regression, in particular a CAViaR model, see Engle and Manganelli (2004), for each of a set of quantiles.

The approach of modelling the quantiles of the distribution directly has been shown to be a robust approach in cases where non-normality holds, as in the case of asset returns, or in cases where the aim is to fit the tails of the distribution. By extending the single quantile and combining quantile estimates at different probability levels we are able to use valuable and different information provided at different sides of the distribution.

Engle and Manganelli (2004) state that, "... But the CAViaR specifications are more general than these GARCH models. Various forms of non-iid error distributions can be modelled in this way. In fact, these models can be used for situations with constant volatilities but changing error distributions, or situations in which both error densities and volatilities are changing." This is our starting point as we seek a robust way to model both the scale and the shape of the conditional asset return distribution.

Interquartile Range (IQR) is the difference between the upper 75% and the lower 25% quartile. It is a robust measure of scale and therefore useful for modelling asset returns which are found to be skewed with heavy tails relative to a Normal distribution. By standardising the quantiles by the estimated time-varying IQR, fat tails should be reduced and the dynamics of the shape are separated from the scale dynamics.

An important issue is the crossing problem when modelling quantiles. By using a single quantile model one might end up with an estimate for the 1% that is higher than the 5% quantile for example. This violates the correct ordering of the quantiles, see Chernozhukov et al. (2008), and Chernozhukov et al. (2010).

In this work we perform a Bayesian analysis using Adaptive MCMC methods. By this methodology, constraints are imposed by construction and the crossing problem is addressed. The MCMC method sidesteps the well-known estimation issues of CAViaR-type models. In particular, quantile estimation requires optimization of a non-linear, not everywhere convex objective function. Thus, the quality of the parameter estimates is very sensitive to the implementation of the optimization algorithm. We choose to use Adaptive MCMC methods because they offer solutions in cases where the target distribution is not tractable, as here. In particular they allow adaptation at every step of the algorithm and they converge to the correct distribution, see Andrieu and Thoms (2008) for a review.

2. THE MODEL

We decompose the asset return distribution by separating the dynamics of the shape (quantiles) and scale (IQR). By this, we are able to model the asset return distribution semiparametrically and identify the type of departures during different periods such as those of high volatility, skewness or low volatility.

Our empirical results suggest that this framework is superior to traditional approaches for single quantile in terms of forecasting and better explaining the evolution of the tails of the distribution. We base quantile estimation on a finite sample of quantiles of the left and the right side of the distribution that are estimated jointly with IQR and standardised by the time-varying IQR. Let the θ-quantile $q_{\theta,t}$ at time t be modelled as

$$q_{\theta,t} = IQR_t \left(u_\theta + \sum_{i=1}^{p} \beta_{\theta,i} \frac{q_{\theta,t-i}}{IQR_{t-i}} + \sum_{i=1}^{p} \frac{\ell(F_{t-i}; \gamma_{\theta,i}, \dots, \gamma_{\theta,p})}{IQR_{t-i}} \right)$$

for $\theta = 0.99, 0.95, 0.25, 0.05, 0.01$, where F_{t-i} is the information set up to and including time $t-i$, and $\ell(.)$ is a possibly non-linear function, u_θ is the intercept of the quantile, $\beta_{\theta,i}$ is the autoregressive parameter and $\gamma_{\theta,i}$ is the parameter on the lags of returns, quantiles etc. Let the time t-IQR be modelled as

$$IQR_t = u + \sum_{i=1}^{p} \beta_i IQR_{t-i} + \sum_{i=1}^{p} \ell(F_{t-i}; \gamma_i, \dots, \gamma_p),$$

and $q_{0.75,t} = IQR_t + q_{0.25,t}$.

3. BAYESIAN ESTIMATION METHODOLOGY

Classical statistical inference on parameter estimates in the joint quantile model setup is feasible, but there are some issues that are not addressed in a robust way, such as tuning the right starting point of the optimisation. This happens because quantile estimation requires optimisation of a non-linear, not everywhere convex objective function. The problem of incorrectly ordered quantiles remains in this framework. Also, calculating the standard errors of the parameter estimates is challenging.

In order to address these issues, we perform a Bayesian analysis using an Adaptive MCMC algorithm. By deriving the Laplace-type likelihood, as in Yu and Moyeed (2001), we are able to obtain the parameters by the Metropolis-Hastings algorithm (MH), see Hastings (1970), and Metropolis et al. (1953).

Let the likelihood (full conditional) be

$$L(u_\theta, \beta_\theta, \gamma_\theta; y, q) \propto \exp\{\sum_{t=1}^{T} \sum_{\theta} [y_t - q_\theta(y_t|F_{t-1})][\theta - I(y_t < q_\theta(y_t|F_{t-1}))]\},$$

where $I(.)$ denotes an indicator function, and T is the sample size.

Here we discuss the steps of the MCMC algorithm to fit the joint quantile model with IQR. We choose the prior to be informative over the possible region following a Normal distribution. For the intercept of the quantiles we choose $u_\theta \sim N(t^{-1}(\theta, 5), 10^6)$, where $t^{-1}(\theta, 5)$ is the inverse of the Student t-distribution at θ with 5 degrees of freedom. For the autoregressive parameters we choose $\beta_\theta \sim N(0.5, 10^6)$, and finally for the parameters on past returns we choose $\gamma_\theta \sim N(0, 10^6)$.

The quantiles and the IQR can be updated simultaneously. The parameters are updated in blocks for each quantile, where $\varphi_\theta = (u_\theta, \beta_\theta, \gamma_\theta)$ are the parameters $u_\theta, \beta_\theta, \gamma_\theta$ of the joint quantile model with IQR. Thus, we update each dimension consecutively using a MH step.

1. Perform the following steps for $g = 1, 2, \ldots, N$
   - φ'_θ is drawn from the conditional density $P({\varphi_\theta}', {\varphi_\theta}^{g-1}|y)$.
   - Propose φ'_θ and accept with probability

$$\alpha = \min\left\{1, \frac{f(y|{\varphi_\theta}')f({\varphi_\theta}')P({\varphi_\theta}', {\varphi_\theta}^{g-1}|y)}{f(y|{\varphi_\theta}^{g-1})f({\varphi_\theta}^{g-1})P({\varphi_\theta}^{g-1}, {\varphi_\theta}'|y)}\right\}.$$

2. Perform the following steps for $g = N+1, N+2, \ldots, M$
   - φ'_θ is drawn from the full conditional density $f(\varphi_\theta|y)$.
   - Propose φ'_θ and accept with probability

$$\alpha = \min\left\{1, \frac{f(y|{\varphi_\theta}')f({\varphi_\theta}')f({\varphi_\theta}', {\varphi_\theta}^{g-1}|y)}{f(y|{\varphi_\theta}^{g-1})f({\varphi_\theta}^{g-1})f({\varphi_\theta}^{g-1}, {\varphi_\theta}'|y)}\right\}.$$

3. The set of accepted values of φ_θ is saved and represents the sample from the target distribution.

We use a combined Adaptive Random Walk (RW) for the burn-in period and an Adaptive Independent Kernel (IK) Metropolis Hastings (MH) algorithm for the remaining iterations, where the sample mean and covariance matrix of the burn-in iterates is employed for each parameter grouping.

We adjust the volatility of the Student-t proposed distribution with 5 degrees of freedom whenever the acceptance rate falls out of bounds around 0.234, as suggested by Gelman et al. (1997). By this we are able to update the kernels given the performance of the algorithm which is tuned so that we are not accepting or rejecting too much.

4. EMPIRICAL ANALYSIS AND RESULTS

We use a sample of daily returns of 4 international indices and 5 stocks trading on the London, NASDAQ, and New York Stock Exchanges. These are the FTSE trading on the London Stock Exchange, NASDAQ trading on the NASDAQ Stock Market, Standard and Poor's 500 (SP500), International Business Machines (IBM), Xerox Corporation (XEROX), Walt Disney company (DISNEY), Caterpillar Inc. (CAT), Dow Chemical company (DOW) and Boeing company (Boeing) trading on the New York Stock Exchange.

Here we present results for FTSE. The data set ranges from November 1984 to November 2014. In table 1 and figure 1 we present results of the Adaptive MCMC described above for 40000 iterations and a burn-in period of 15000.

In table 1 we provide a summary of the Bayesian posteriors for one of the joint quantile specifications, the Joint Component Asymmetric Slope model (J-C-AS) — where we set the number of lags of past information at one. This specification uses a two component process to model IQR in order to account for a slow moving component. We replace u with a time-varying process that induces a long memory property to the IQR, and allows for smooth adjustments to the level of the IQR under different market conditions. The deviation $IQR_{t-1} - u_{t-1}$ is the component that represents an adjusted distance from the unconditional mean. The dynamics of u_t capture the dependence in $\overline{IQR_t}$, albeit with an adjusted mean level. Overall, we introduce a long memory feature in the IQR process similar to that in the component GARCH models, see Engle and Lee (1999).

$$
\begin{aligned}
IQR_t = u_t + \beta_1(IQR_{t-1} - u_{t-1}) + \gamma_1 y^+_{t-1} + \delta_1 y^-_{t-1}, \\
u_t = \alpha + \beta_2 u_{t-1} + \gamma_2 y_{t-1}.
\end{aligned}
$$

The quantiles are given by

$$
q_{\theta,t} = IQR_t \left(u_\theta + \beta_\theta \frac{q_{\theta,t-1}}{IQR_{t-1}} + \gamma_\theta \frac{|y_{t-1}|}{IQR_{t-1}} \right),
$$

for $\theta = 0.99, 0.95, 0.25, 0.05, 0.01$, and $q_{0.75,t} = IQR_t + q_{0.25,t}$.

Table 1 gives sensible parameter estimates. In particular, the intercept of the quantiles is negative or positive corresponding to the left or the right side of the distribution. The autoregressive parameters ensure stationarity for all the quantiles, IQR is highly persistent while others are not.

Parameter	Estimate	StDev	LowerCI	UppperCI	Parameter	Estimate	StDev	LowerCI	UppperCI
$u_{0.99}$	1.2571	0.6091	1.2495	1.2646	$\gamma_{0.05}$	-0.0218	0.0468	-0.0224	-0.0212
$\beta_{0.99}$	0.2905	0.3311	0.2864	0.2946	$u_{0.01}$	-1.3578	0.6142	-1.3654	-1.3501
$\gamma_{0.99}$	0.0388	0.0736	0.0379	0.0397	$\beta_{0.01}$	0.2627	0.3113	0.2588	0.2665
$u_{0.95}$	0.8874	0.4446	0.8819	0.8929	$\gamma_{0.01}$	-0.2130	0.0809	-0.2140	-0.2120
$\beta_{0.95}$	0.2932	0.3465	0.2889	0.2975	β_1	0.9194	0.0196	0.9192	0.9197
$\gamma_{0.95}$	0.0141	0.0379	0.0137	0.0146	γ_1	0.0163	0.0138	0.0161	0.0165
$u_{0.25}$	-0.2726	0.1114	-0.2739	-0.2712	δ_1	-0.1947	0.0237	-0.1950	-0.1944
$\beta_{0.25}$	0.3947	0.2321	0.3918	0.3975	α	0.0002	0.0003	0.0002	0.0002
$\gamma_{0.25}$	-0.0142	0.0190	-0.0144	-0.0139	β_2	0.9409	0.0238	0.9407	0.9412
$u_{0.05}$	-0.8945	0.4590	-0.9002	-0.8888	γ_2	0.0270	0.0156	0.0268	0.0272
$\beta_{0.05}$	0.3189	0.3380	0.3147	0.3230					

Table 1: Summary statistics for the posteriors of the parameters of J-C-AS for FTSE.

Parameters on the returns of positive quantiles are positive, increasing the quantile each time while parameters on the returns of negative quantiles are negative, decreasing it.

Figure 1 gives the plots for the $q_{\theta,t}$ quantiles, and the corresponding standardised quantiles by IQR, $\hat{q}_{\theta,t} = \frac{q_{\theta,t}}{IQR_t}$. We show that by modelling IQR jointly with the quantiles (standardised by IQR) we are able to capture the time-varying scale, leaving few outliers.We are also able to address the crossing problem. We are currently analysing if there are indeed any dynamics or other time-varying features in the shape of the distribution after adjusting for scale.

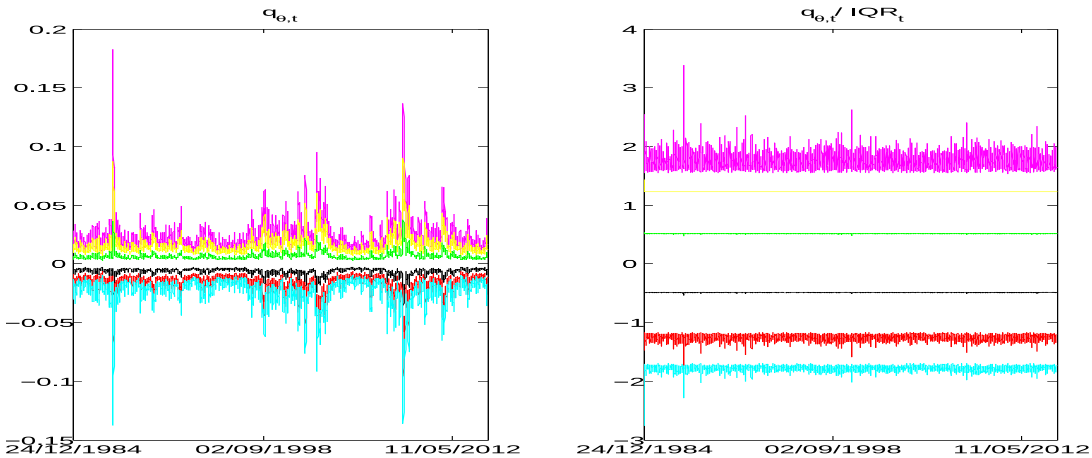


Figure 1: Jointly estimated $q_{\theta,t}$ quantiles with IQR, and standardised $\hat{q}_{\theta,t}$ quantiles for FTSE.

5. CONCLUSIONS

This paper presents a method for Bayesian semiparametric inference in joint quantile models. Joint quantile models estimated with IQR provide evidence for being able to capture dynamics that are consistent with the concept of time-varying risk.

Posteriors concentrate on sensible values for quantiles at different probability levels. Having checked the autocorrelation function for sampled values from each parameter, we find that this decays quickly for all parameters. This suggests that the algorithm mixes quickly.

Also extreme quantiles at the tails are more volatile than the quantiles of the main body, as expected. In line with prior evidence asymmetry parameters suggest that negative returns are more likely to cause higher increases in the left tail of the distribution than positive returns, this is stronger for more extreme quantiles.

We are able to analyse the dynamics of both the scale and the shape which would be difficult to capture using traditional models.

Future work will consider estimating the conditional asset return distribution without using the link between the likelihood and the Laplace-type distribution in order to account for the potential interdependence between individual quantiles.

References

C. Andrieu and J. Thoms. A tutorial on adaptive MCMC. *Statistics and Computing*, 18(4):343–373, 2008.

Victor Chernozhukov, Iván Fernández-Val, and Alfred Galichon. Improving point and interval estimates of monotone functions by rearrangement. *Biometrica*, 96:559–575, 2008.

Victor Chernozhukov, Iván Fernández-Val, and Alfred Galichon. Quantile and probability curves without crossing. *Econometrica*, 78:1093–1125, 2010.

Robert. F. Engle and G. Lee. A permanent and transitory component model of stock return volatility. In *Cointegration, Causality and Forecasting: A festschrift in honor of Clive W.J. Granger*, pages 475–497. Oxford University Press, 1999.

Robert F. Engle and Simone Manganelli. CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *American Statistical Journal of Business and Economics*, 4:367–381, 2004.

A. Gelman, W. R. Gilks, and G. O. Roberts. Weak convergence and optimal scaling of random walk Metropolis algorithms. *Ann. Appl. Probab.*, 7:110–120, 1997.

W.K. Hastings. Monte-Carlo sampling methods using Markov chains and their applications. *Biometrica*, 57:97–109, 1970.

Roger Koenker and Gilbert Bassett. Regression quantiles. *Econometrica*, 46:33–50, 1978.

N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21:1087–1092, 1953.

K. Yu and R.A Moyeed. Bayesian quantile regression. *Statistics and Probability Letters*, 54: 437–447, 2001.

AFFINE LIBOR MODELS WITH REAL-VALUED PROCESSES

Wolfgang Müller and Stefan Waldenberger

Graz University of Technology, Institute of Statistics, NAWI Graz
Kopernikusgasse 24/III, 8010 Graz, Austria
Email: w.mueller@tugraz.at, stefan.waldenberger@tugraz.at

1. INTRODUCTION

Market models, the most famous example being the LIBOR market model, are very popular in the area of interest rate modeling. If these models generate nonnegative interest rates they usually do not give semi-analytic formulas for both basic interest rate derivatives, caps and swaptions. One exception is the class of affine LIBOR models proposed by Keller-Ressel et al. (2013). Using nonnegative affine processes as driving processes affine LIBOR models guarantee nonnegative forward interest rates and lead to semi-analytical formulas for caps and swaptions, so that calibration to interest rate market data is possible. This paper modifies the setup of Keller-Ressel et al. (2013) to allow for not necessarily nonnegative affine processes. This modification still leads to semi-analytical formulas for caps and swaptions and guarantees nonnegative forward interest rates, but allows for a wider class of driving affine processes and hence is more flexible in producing interest rate skews and smiles. This paper summarizes the results of the paper Müller and Waldenberger (2015), further information as well as proofs can be found in the original paper.

Affine processes

Let $X = (X_t)_{0\le t\le T}$ be a homogeneous Markov process with values in $D = \mathbb{R}^m_{\ge 0} \times \mathbb{R}^n$ realized on a measurable space $(\Omega, \mathcal{A})$ with filtration $(\mathcal{F}_t)_{0\le t\le T}$, with regards to which X is adapted. Denote by $\mathbb{P}^x[\cdot]$ and $\mathbb{E}^x[\cdot]$ the corresponding probability and expectation when $X_0 = x$. X is said to be an affine process, if its characteristic function has the form

$$\mathbb{E}^x\left[\mathrm{e}^{u\cdot X_t}\right] = \exp\left(\phi_t(u) + \psi_t(u)\cdot x\right), \quad u \in \mathrm{i}\,\mathbb{R}^d, x \in D, \tag{1}$$

where $\phi : [0,T]\times \mathrm{i}\,\mathbb{R}^d \to \mathbb{C}$ and $\psi : [0,T]\times \mathrm{i}\,\mathbb{R}^d \to \mathbb{C}^d$ with $\mathrm{i}\,\mathbb{R}^d = \{u \in \mathbb{C}^d : \mathrm{Re}(u) = 0\}$. By homogeneity and the Markov property the conditional characteristic function satisfies

$$\mathbb{E}^x\left[e^{u\cdot X_t}\middle|\, \mathcal{F}_s\right] = \exp\left(\phi_{t-s}(u) + \psi_{t-s}(u)\cdot X_s\right).$$

An affine process X is called analytic (see Keller-Ressel (2008)), if X is stochastically continuous and the interior of the set

$$\mathcal{V} := \left\{ u \in \mathbb{C}^d : \sup_{0 \le s \le T} \mathbb{E}^x \left[\mathrm{e}^{\mathrm{Re}(u) \cdot X_s} \right] < \infty \quad \forall x \in D \right\}, \tag{2}$$

contains 0. In this case the functions ϕ and ψ have continuous extensions to $\mathcal{V}$, which are analytic in the interior, such that (1) holds for all $u \in \mathcal{V}$.

Interest rate market models

Consider a tenor structure $0 < T_1 < \cdots < T_N < T_{N+1} =: T$ and a market consisting of zero coupon bonds with maturities $T_1, \ldots, T_{N+1}$. Their price processes $(P(t, T_k))_{0 \le t \le T_k}$ are assumed to be nonnegative semimartingales on a filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{0 \le t \le T}, \mathbb{P})$, which satisfy $P(T_k, T_k) = 1$ almost surely. If there exists an equivalent probability measure $\mathbb{Q}^T$ such that the normalized bond price processes $P(\cdot, T_k)/P(\cdot, T)$ are martingales, the market is arbitrage-free. In this case we can define equivalent martingale measures $\mathbb{Q}^{T_k}$ for the numeraires $P(t, T_k)$ instead of $P(t, T)$ by

$$\frac{\mathrm{d}\,\mathbb{Q}^{T_k}}{\mathrm{d}\,\mathbb{Q}^T} = \frac{1}{P(T_k, T)} \frac{P(0, T)}{P(0, T_k)}. \tag{3}$$

2. THE MODIFIED AFFINE LIBOR MODEL

On the filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{0 \le t \le T}, \mathbb{Q}^T)$ consider an analytic one-dimensional affine process X with a fixed starting value x_0. For $u \in \mathcal{V}$ with $-u \in \mathcal{V}$ consider the martingales M^u,

$$M_t^u := \mathbb{E}^{\mathbb{Q}^T} \left[\cosh(u X_T) \middle| \mathcal{F}_t \right] = \frac{1}{2} \left(\mathrm{e}^{\phi_{T-t}(u) + \psi_{T-t}(u) X_t} + \mathrm{e}^{\phi_{T-t}(-u) + \psi_{T-t}(-u) X_t} \right). \tag{4}$$

By the symmetry of the cosinus hyperbolicus $M^u = M^{-u}$, hence one may restrict u to be nonnegative. For the given tenor structure $0 < T_1 < \cdots < T_N \le T_{N+1} = T$ and the market setup from before define the normalized bond prices for $k = 1, \ldots, N$ and $t \le T_k$ as

$$\frac{P(t, T_k)}{P(t, T)} := M_t^{u_k}, \qquad u_k \in \{ v \in \mathcal{V} : v \ge 0, -v \in \mathcal{V} \}.$$

With $M_t^{u_k}$ being a $\mathbb{Q}^T$-martingale the model is arbitrage-free. For every $x \in \mathbb{R}$ the function $u \mapsto \cosh(ux)$ is increasing in $u \in \mathbb{R}_{\ge 0}$ and satisfies $\cosh(ux) \ge 1$ so that if

$$u_1 \ge u_2 \ge \cdots \ge u_N \ge 0,$$

it holds that

$$\frac{P(t, T_1)}{P(t, T)} \ge .. \ge \frac{P(t, T_N)}{P(t, T)} \ge 1 \tag{5}$$

and forward interest rates

$$F^k(t) = \frac{1}{\Delta_k}\left(\frac{P(t,T_{k-1})}{P(t,T_k)} - 1\right) = \frac{1}{\Delta_k}\left(\frac{M_t^{u_{k-1}}}{M_t^{u_k}} - 1\right), \qquad \Delta_k = T_k - T_{k-1},$$

are nonnegative for all $0 \le t \le T_k$. To fit initial market data one has to choose the sequence (u_k) so that $M_0^{u_k} = P(0,T_k)/P(0,T)$.

Remark 2.1 *As in the affine LIBOR model for a monotonically decreasing sequence (u_k) forward interest rates are not only nonnegative, but bounded below by strictly positive time-dependent constants (the bounds can be calculated numerically). This is not a big issue if these bounds are close to zero, but has to be checked during the calibration process.*

In the modified affine LIBOR model the change of measure to the T_k-forward measure $\mathbb{Q}^{T_k}$ is given by

$$\frac{\mathrm{d}\,\mathbb{Q}^{T_k}}{\mathrm{d}\,\mathbb{Q}^{T}} = \frac{P(0,T)}{P(0,T_k)} M_{T_k}^{u_k} = \frac{M_{T_k}^{u_k}}{M_0^{u_k}}. \tag{6}$$

Here $M_t^{u_k}$ is a sum of exponentials of X_t, while in the affine LIBOR model the corresponding term is a single exponential. This means that contrary to the affine LIBOR model the process X is not an inhomogeneous affine process under $\mathbb{Q}^{T_k}$ and it is not possible to calculate the $\mathbb{Q}^{T_k}$-moment generating function of $P(t,T_{k-1})/P(t,T_k)$. Nevertheless it is possible to get analytical formulas for the prices of caplets and swaptions.

Option pricing in the modified affine LIBOR model

The derivation of the pricing formulas for caplets and swaptions is based on a method first applied in Jamshidian (1989). Here we will consider only swaptions since caplets can be treated similarly[1]. Note that if $u_k = u_{k-1}$ the corresponding forward interest rate F^k always stays zero. To exclude such pathological examples assume that the sequence (u_k) is strictly decreasing. In this section random variables are often viewed as functions of the value of the driving process X. Specifically consider the functions $M_t^u : \mathbb{R} \to \mathbb{R}$,

$$x \mapsto M_t^u(x) := \frac{1}{2}\left(\mathrm{e}^{\phi_{T-t}(u)+\psi_{T-t}(u)x} + \mathrm{e}^{\phi_{T-t}(-u)+\psi_{T-t}(-u)x}\right). \tag{7}$$

The time t value of martingale M^u in (4) is then $M_t^u = M_t^u(X_t)$. In the rest of the paper M_t^u will denote both, the function and the value of the stochastic processes, where the correct interpretation should be clear from context.

Consider a swap which is part of the tenor structure. That is, consider $1 \le \alpha < \beta \le N$ and the according interest rate swap with forward swap rate

$$S_{\alpha,\beta}(t) = \frac{P(t,T_\alpha) - P(t,T_\beta)}{\sum_{k=\alpha+1}^{\beta} \Delta_k P(t,T_k)}, \qquad \Delta_k = T_k - T_{k-1}.$$

[1]Actually caplet prices then coincide with prices of swaptions with only one underlying period. The difference between those two derivatives is the payoff time.

The payoff of a put swaption on the above swap with strike K is then

$$\sum_{k=\alpha+1}^{\beta} P(T_\alpha, T_k)\Delta_k \left(K - S_{\alpha,\beta}(T_\alpha)\right)_+ = \left(P(T_\alpha, T_\beta) + K \sum_{k=\alpha+1}^{\beta} \Delta_k P(T_\alpha, T_k) - 1\right)_+$$
$$= \left(\frac{M_{T_\alpha}^{u_\beta}}{M_{T_\alpha}^{u_\alpha}} + \sum_{k=\alpha+1}^{\beta} K\Delta_k \frac{M_{T_\alpha}^{u_k}}{M_{T_\alpha}^{u_\alpha}} - 1\right)_+ .$$

The following lemma can be used to simplify this payoff function.

Lemma 2.1 *For $i = 1, \ldots, n$ let $u_0 \geq u_i \geq 0$, where for at least one i $u_0 > u_i$. Let $c_i > 0$ be positive constants. Define a function $g : \mathbb{R} \to \mathbb{R}$ by*

$$g(x) := \sum_{i=1}^{n} c_i \frac{M_t^{u_i}(x)}{M_t^{u_0}(x)}. \tag{8}$$

Then g has a unique maximum at some point $\xi \in \mathbb{R}$ and is strictly monotonically decreasing to 0 on the left and right side of ξ.

Since the function $M_{T_\alpha}^{u_\beta}(x)/M_{T_\alpha}^{u_\alpha}(x) + \sum_{k=\alpha+1}^{\beta} K\Delta_k M_{T_\alpha}^{u_k}(x)/M_{T_\alpha}^{u_\alpha}(x)$ is of the form of Lemma 2.1, it has a unique maximum ξ and one can find constants $\kappa_1 \leq \xi \leq \kappa_2$ such that after a change of measure using (6) the value of a put swaption is

$$\mathrm{PutSwaption}(t, T_\alpha, T_\beta, K) = P(t,T)\mathbb{E}^{\mathbb{Q}^T}\left[f_{\alpha,\beta}^K(X_{T_\alpha}) \middle| \mathcal{F}_t \right],$$

where

$$f_{\alpha,\beta}^K(x) = \left(M_{T_\alpha}^{u_\beta}(x) - M_{T_\alpha}^{u_\alpha}(x) + \sum_{k=\alpha+1}^{\beta} K\Delta_k M_{T_\alpha}^{u_k}(x) \right) \mathbb{I}\left\{\kappa_1 < x < \kappa_2\right\}. \tag{9}$$

Here the expectation is under $\mathbb{Q}^T$, where the moment generating function of X_{T_α} is known in closed form. One can use Fourier inversion methods to arrive at the following formula.

Theorem 2.2 *Let $R \in \mathcal{V} \cap \mathbb{R}$. In the modified affine LIBOR model the price of a put swaption is*

$$\mathrm{PutSwaption}(t, T_\alpha, T_\beta, K) = \frac{P(t,T)}{\pi} \int_0^\infty \mathrm{Re}\left(\mathcal{M}_{X_{T_\alpha}|X_t}(R + \mathrm{i}u) \hat{f}_{\alpha,\beta}^K(u - \mathrm{i}R) \right) \mathrm{d}u. \tag{10}$$

The Fourier transform $\hat{f}_{\alpha,\beta}^K$ is for $R \notin \{0, u_\alpha, \ldots, u_\beta\}$

$$\hat{f}_{\alpha,\beta}^K(z) = \frac{1}{\mathrm{i}z}\Big(h_{\kappa_1,\kappa_2}^{T_\alpha}(-\mathrm{i}z, u_\beta) - h_{\kappa_1,\kappa_2}^{T_\alpha}(-\mathrm{i}z, u_\alpha) + K \sum_{k=\alpha+1}^{\beta} \Delta_k h_{\kappa_1,\kappa_2}^{T_\alpha}(-\mathrm{i}z, u_k) \Big), \tag{11}$$

where $h_{\kappa_1,\kappa_2}^t(z,u)$

$$\begin{aligned} h_{\kappa_1,\kappa_2}^t(z,u) := {} & \mathrm{e}^{\phi_{T-t}(u)} \frac{\psi_{T-t}(u)}{2(z + \psi_{T-t}(u))} \left(\mathrm{e}^{(z+\psi_{T-t}(u))\kappa_2} - \mathrm{e}^{(z+\psi_{T-t}(u))\kappa_1} \right) \\ & + \mathrm{e}^{\phi_{T-t}(-u)} \frac{\psi_{T-t}(-u)}{2(z + \psi_{T-t}(-u))} \left(\mathrm{e}^{(z+\psi_{T-t}(-u))\kappa_2} - \mathrm{e}^{(z+\psi_{T-t}(-u))\kappa_1} \right). \end{aligned} \tag{12}$$

In order to calculate $\hat{f}^K_{\alpha,\beta}$ one has to find the roots κ_1, κ_2 of the function

$$g^K_{\alpha,\beta}(x) := \frac{M^{u_\beta}_{T_\alpha}(x)}{M^{u_\alpha}_{T_\alpha}(x)} + \sum_{k=\alpha+1}^{\beta} K\Delta_k \frac{M^{u_k}_{T_\alpha}(x)}{M^{u_\alpha}_{T_\alpha}(x)} - 1. \tag{13}$$

By Lemma 2.1 this amounts to finding the roots of a function which has a single optimum and is monotonic when moving away from this optimum. Numerical determination of the roots of such well-behaved one-dimensional functions poses no problem. Having determined those bounds valuation reduces to a one-dimensional integration of a function that is falling at least like $1/x^2$ (depending on the moment generating function of the affine process), so also numerical integration is feasible. Caps, floors and options like digital options or Asset-or-Nothing options can be calculated in a similar manner.

Volatility surfaces

To apply the proposed model one needs to fix an affine process. Here we consider an Ornstein-Uhlenbeck process (parameters θ, λ) which is generated by a Levy process which consists of a Brownian motion (parameter σ^2) and double exponentially distributed jumps (parameters β^+, β^-, α^+, α^-). Details can be found in Müller and Waldenberger (2015). The functions ϕ and ψ in this case read

$$\begin{aligned}
\phi_t(u) &= \frac{\sigma^2 u^2}{4\lambda}(1-\mathrm{e}^{-2\lambda t}) + \theta u(1-\mathrm{e}^{-\lambda t}) \\
&+ \frac{\beta^+ + \beta^-}{2} \ln\left(\frac{(\alpha^+ - \mathrm{e}^{-\lambda t}u)(\alpha^- + \mathrm{e}^{-\lambda t}u)}{(\alpha^+ - u)(\alpha^- + u)}\right) \\
&+ \frac{\beta^+ - \beta^-}{2} \ln\left(\frac{(\alpha^+ - \mathrm{e}^{-\lambda t}u)(\alpha^- + u)}{(\alpha^+ - u)(\alpha^- + \mathrm{e}^{-\lambda t}u)}\right), \\
\psi_t(u) &= \mathrm{e}^{-\lambda t}u
\end{aligned}$$

and for this process $\mathcal{V} = \{u \in \mathbb{C} : -\alpha^- < \mathrm{Re}(u) < \alpha^+\}$.

With this affine process it is possible to generate volatility smiles as well as volatility skews. For illustration of a possible volatility smile we consider a term structure with constant interest rates of 3.5%. The tenor structure and therefore the forward interest rates are based on half year intervals. Implied volatilities are then calculated for caplets with maturities over a 5-year period and strikes ranging from 0.02 to 0.07. The resulting volatility surface displayed in figure 1 shows a very pronounced smile. As mentioned in the previous chapters forward interest rates in this type of model will be bounded from below. The bounds in these examples are at 1% for the forward interest rate expiring after half a year and decrease to basically 0% for the forward interest rate which expires in 5 years. Hence they are well within reasonable boundaries.

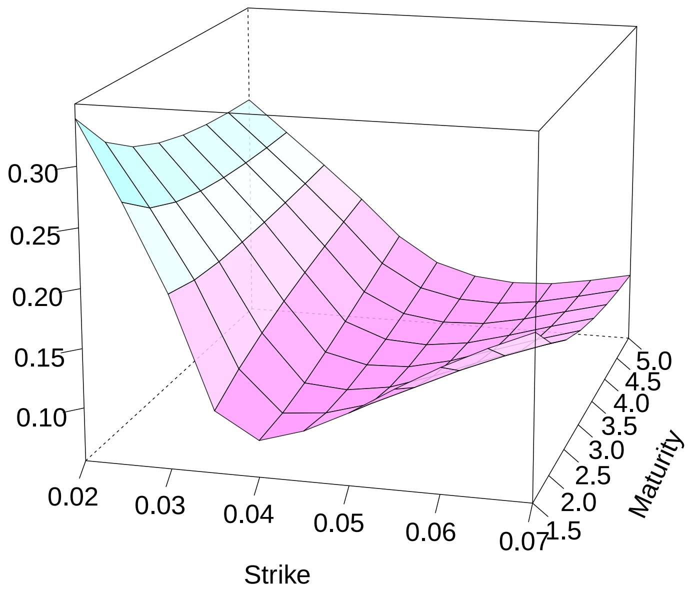


Figure 1: Implied volatility smile of caplets generated by an OU process with parameters $\lambda = 0.02, \alpha^+ = 50, \alpha^- = 5, \beta^+ = 50, \beta^- = 10, \sigma = 0, \theta = 0, x = 1$ and $T = 10$.

3. CONCLUSION

Classical interest rate market models are not capable of simultaneously allowing for semi-analytical pricing formulas for caplets and swaptions and guaranteeing nonnegative forward interest rates. One exception are the affine LIBOR models presented in Keller-Ressel et al. (2013). This paper modifies their approach to also allow for driving processes which are not necessarily nonnegative. Caplet and swaption valuation is possible via one-dimensional numerical integration. This allows for a fast calculation of implied volatilities for these types of interest rate derivatives. With the additional flexibility of real-valued affine processes this type of model is capable of producing skewed implied volatility surfaces as well as implied volatility surfaces with pronounced smiles.

References

F. Jamshidian. An exact bond option formula. *The Journal of Finance*, 44(1):205–209, 1989.

M. Keller-Ressel. *Affine processes - Theory and application in finance*. PhD-thesis, Wien University of Technology, 2008.

M. Keller-Ressel, A. Papapantoleon, and J. Teichmann. The affine LIBOR models. *Mathematical Finance*, 23(4):627–658, 2013.

W. Müller and S. Waldenberger. Affine LIBOR models driven by real-valued affine processes. *arXiv:1503.00864*, 2015.

CONTINUITY AND DIFFERENTIABILITY PROPERTIES OF THE SURVIVAL PROBABILITIES IN RISK MODELS WITH INVESTMENTS AND THEIR APPLICATIONS

Olena Ragulina

Department of Probability Theory, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kyiv, Volodymyrska Str. 64, 01601 Kyiv, Ukraine
Email: ragulina.olena@gmail.com

We consider a generalisation of the classical risk model where an insurance company invests all surplus in risk-free and risky assets proportionally. We investigate continuity and differentiability of the infinite-horizon and finite-horizon survival probabilities and discuss some applications of these results. In particular, we apply these results to find analytic expressions for the infinite-horizon survival probabilities in the classical risk model where the claim sizes are exponentially distributed and the insurance company applies a franchise.

1. INTRODUCTION

We deal with the classical risk model (see, e.g. Asmussen (2000), Grandell (1991), Rolski et al. (1999)) where all surplus of an insurance company is invested in risk-free and risky assets proportionally. In the classical risk model claim sizes form a sequence $(Y_i)_{i\ge 1}$ of nonnegative i.i.d. random variables with c.d.f. $F(y) = \mathbb{P}[Y_i \le y]$ and finite expectation μ. Let τ_i be the time when the ith claim arrives. The number of claims on the time interval $[0, t]$ is a Poisson process $(N_t)_{t\ge 0}$ with constant intensity $\lambda > 0$. Thus, the total claims on $[0, t]$ equal $\sum_{i=1}^{N_t} Y_i$. We set $\sum_{i=1}^{0} Y_i = 0$ if $N_t = 0$. The insurance company has a nonnegative initial surplus x and receives premiums with constant intensity $c > 0$.

In addition, we assume that all surplus is invested in risk-free and risky assets. The price of the risk-free asset equals $B_t = B_0 e^{rt}$ at time t, where B_0 is the price of the risk-free asset at the time $t = 0$, and $r > 0$ is a risk-free interest rate. The price of the risky asset equals $S_t = S_0 \exp\left(\tilde{r}t + \sum_{i=1}^{\tilde{N}_t} \tilde{Y}_i\right)$ at time t, where $S_0 > 0$ is the price of risky asset at the time $t = 0$, $\tilde{r} > 0$ is a constant, $(\tilde{Y}_i)_{i\ge 1}$ is a sequence of i.i.d. random variables with c.d.f. $\tilde{F}(y) = \mathbb{P}[\tilde{Y}_i \le y]$ such that $0 < \tilde{F}(0) < 1$, and $(\tilde{N}_t)_{t\ge 0}$ is a Poisson process with constant intensity $\tilde{\lambda} > 0$. We denote by $\tilde{\tau}_i$ the time of ith jump of $(\tilde{N}_t)_{t\ge 0}$ and set $\sum_{i=1}^{0} \tilde{Y}_i = 0$ if $\tilde{N}_t = 0$. All the random variables and processes in this model are mutually independent.

We suppose that all surplus is invested in the assets proportionally. That is at any time $t \geq 0$ the part α is invested in the risky asset and the part $1-\alpha$ is invested in the risk-free asset, where $0 < \alpha \leq 1$. Set $\bar{r} = \alpha\tilde{r} + (1-\alpha)r$.

Let $X_t(x)$ be the surplus of the insurance company at time t provided that its initial surplus is x. Then the surplus process $\big(X_t(x)\big)_{t\geq 0}$ follows the equation

$$X_t(x) = x + \int_0^t (\bar{r}X_s(x) + c)\,ds + \alpha \sum\nolimits_{i=1}^{\tilde{N}_t} X_{\tilde{\tau}_i -}(x)\cdot(e^{\tilde{Y}_i} - 1) - \sum\nolimits_{i=1}^{N_t} Y_i, \quad t \geq 0. \quad (1)$$

The infinite-horizon ruin probability is given by $\psi(x) = \mathbb{P}\big[\inf_{s\geq 0}\ X_s(x) < 0\big]$. The finite-horizon ruin probability is given by $\psi(x,t) = \mathbb{P}\big[\inf_{0\leq s\leq t}\ X_s(x) < 0\big]$. The infinite-horizon and finite-horizon survival probabilities equal $\varphi(x) = 1-\psi(x)$ and $\varphi(x,t) = 1-\psi(x,t)$, respectively.

The remainder of this paper is organized as follows. In Section 2, Theorems 2.1 and 2.2 describe continuity and differentiability of the infinite-horizon and finite-horizon survival probabilities. The proofs of these theorems are given in Bondarev and Ragulina (2012). We also discuss some applications of these results in Section 2. In Section 3, we concentrate in particular on the case where a franchise is applied in the classical risk model without any investments.

2. CONTINUITY AND DIFFERENTIABILITY OF THE SURVIVAL PROBABILITIES

We define the set Z^* as follows:

$$Z^* = \Big\{z\colon \mathbb{E}\big[\big(1-\alpha+\alpha e^{\tilde{Y}_1}\big)^{-z}\big] < \infty, \quad 0 < z \leq 1\Big\}.$$

Theorem 2.1 *1. Let the surplus process $\big(X_t(x)\big)_{t\geq 0}$ follow* (1) *under the above assumptions. Then the function $\varphi(x)$ is continuous on $\mathbb{R}_+$.*
2. Moreover, let there be $z_0 \in Z^*$ such that

$$\tilde{\lambda}\big(\mathbb{E}\big[\big(1-\alpha+\alpha e^{\tilde{Y}_1}\big)^{-z_0}\big] - 1\big) - \bar{r}z_0 < 0.$$

Then the following assertions hold.

(i) The function $\varphi(x)$ is continuously differentiable on $\mathbb{R}_+$, except at positive points of discontinuity of $F(y)$.

(ii) If $x > 0$ is a point of discontinuity of $F(y)$ and $F(x) - F(x_-) = p$, then $\varphi(x)$ has the left and right derivatives $\varphi'_-(x)$ and $\varphi'_+(x)$, respectively, and

$$\varphi'_-(x) - \varphi'_+(x) = \frac{\lambda p \varphi(0)}{\bar{r}x + c} > 0.$$

(iii) The function $\varphi(x)$ satisfies the integro-differential equation

$$(\bar{r}x + c)\varphi'_+(x) = (\lambda + \tilde{\lambda})\varphi(x) - \lambda\int_0^x \varphi(x-y)\,dF(y) - \tilde{\lambda}\int_{-\infty}^{+\infty} \varphi\big((1-\alpha+\alpha e^y)x\big)\,d\tilde{F}(y)$$

on $\mathbb{R}_+$ with the boundary condition $\lim_{x\to+\infty}\varphi(x) = 1$.

(iv) For all $x_0 \geq 0$*, we have*

$$\sup_{x \in [x_0, +\infty)} |\varphi'_+(x)| \leq \frac{\lambda + \tilde{\lambda}}{\bar{r} x_0 + c}.$$

Theorem 2.2 *Let the surplus process* $\big(X_t(x)\big)_{t\geq 0}$ *follow* (1) *under the above assumptions.*
1. If Y_i*,* $i \geq 1$*, have a p.d.f.* $f(y)$*, which is continuous on* $\mathbb{R}_+$*, and* $\tilde{Y}_i$*,* $i \geq 1$*, have a p.d.f.* $\tilde{f}(y)$*, which is continuous on* $\mathbb{R}$*, then* $\varphi(x,t)$ *is continuous on* $(0,+\infty) \times [0,+\infty)$ *as a function of two variables.*
2. If the p.d.f. $f(y)$ *of* Y_i*,* $i \geq 1$*, has the derivative* $f'(y)$ *on* $\mathbb{R}_+$*, such that* $|f'(y)|$ *is integrable and bounded on* $\mathbb{R}_+$*, and the p.d.f.* $\tilde{f}(y)$ *of* $\tilde{Y}_i$*,* $i \geq 1$*, has the derivative* $\tilde{f}'(y)$ *on* $\mathbb{R}$*, such that* $|\tilde{f}'(y)|$*,* $\tilde{f}(y)e^{-y}$*, and* $|\tilde{f}'(y)|e^{-y}$ *are integrable and bounded on* $\mathbb{R}$*, then*

(i) $\varphi(x,t)$ *has partial derivatives w.r.t.* x *and* t *on* $(0,+\infty) \times [0,+\infty)$*, which are continuous as functions of two variables;*

(ii) $\varphi(x,t)$ *satisfies the partial integro-differential equation*

$$\frac{\partial \varphi(x,t)}{\partial t} - (\bar{r}x + c)\frac{\partial \varphi(x,t)}{\partial x} + (\lambda + \tilde{\lambda})\varphi(x,t) - \lambda \int_0^x \varphi(x-y,t)\, dF(y) - \tilde{\lambda} \int_{-\infty}^{+\infty} \varphi\big((1 - \alpha + \alpha e^y)x, t\big)\, d\tilde{F}(y) = 0$$

on $(0,+\infty) \times [0,+\infty)$ *with the boundary conditions* $\varphi(x,0) = 1$ *and* $\lim_{x\to+\infty} \varphi(x,t) = 1$*;*

(iii) for all $x_0 > 0$ *and* $T > 0$*, we have*

$$\sup_{\substack{x \in [x_0,+\infty),\\ t \in [0,T]}} \left| \frac{\partial \varphi(x,t)}{\partial x} \right| \leq C_0 \bigg(C_1 + \frac{C_2}{x_0} \bigg),$$

where

$$C_0 = \begin{cases} \frac{1 - e^{(\bar{r} - \lambda - \tilde{\lambda})T}}{\lambda + \tilde{\lambda} - \bar{r}} & \text{if } \lambda + \tilde{\lambda} \neq \bar{r}, \\ T & \text{if } \lambda + \tilde{\lambda} = \bar{r}, \end{cases} \qquad C_1 = \lambda \bigg(f(0) + \int_0^{+\infty} |f'(y)|\, dy \bigg),$$

$$C_2 = \frac{\tilde{\lambda}}{\alpha} \bigg((1-\alpha) \int_{-\infty}^{+\infty} e^{-y} \tilde{f}(y)\, dy + \int_{-\infty}^{+\infty} \big((1-\alpha)e^{-y} + \alpha\big) |\tilde{f}'(y)|\, dy \bigg).$$

We consider applications of these results in three directions. First, we use the bounds for the derivatives of the survival probability w.r.t. the initial surplus to get formulas connecting the accuracy and reliability of the uniform approximations of the survival probabilities by their statistical estimates (see Bondarev and Ragulina (2012)). Next, we use analogues of these results for the classical risk model without investments to solve problems of optimal control by franchise and deductible amounts from viewpoint of the infinite-horizon survival probability maximization (see Ragulina (2014)). Finally, we also use analogues of these results for the classical risk model without investments to find analytic expressions for the infinite-horizon survival probabilities in a few cases when the c.d.f. of claim sizes is a sum of absolutely continuous and discrete components (see Ragulina (2011)).

3. SURVIVAL PROBABILITY IN THE CLASSICAL RISK MODEL WITH A FRANCHISE

We now consider the classical risk model without any investments where the insurance company applies a franchise. A franchise is a provision in an insurance policy whereby an insurer does not pay unless damage exceeds the franchise amount.

Let the net profit condition hold, i.e. $c > \lambda\mu$. Moreover, we assume that the insurance company uses the expected value principle for premium calculation, which means that $c = \lambda\mu(1+\theta)$, where $\theta > 0$ is a safety loading.

In the classical risk model the surplus process $\big(X_t(x)\big)_{t\ge 0}$ is defined as

$$X_t(x) = x + ct - \sum\nolimits_{i=1}^{N_t} Y_i, \quad t \ge 0. \tag{2}$$

Let d be a franchise amount. We choose it at the initial time and do not change it later. We make the following natural assumption concerning this amount: $0 \le d < +\infty$. In particular, if $d = 0$, then a franchise is not used. Let $Y_i^{(d)}$, $i \ge 1$, denote an insurance compensation for the ith claim provided that the franchise amount is d. We let $F^{(d)}(y)$ stand for the c.d.f. of $Y_i^{(d)}$.

Normally, a franchise also implies reduction of insurance premiums. We suppose that the safety loading $\theta > 0$ is constant. Thus, the premium intensity is given by $c^{(d)} = \lambda(1+\theta)\,\mathbb{E}\big[Y_i^{(d)}\big]$ provided that the insurance company uses the expected value principle for premium calculation.

Let $X_t^{(d)}(x)$ be the surplus of the insurance company at time t provided that its initial surplus is x, and the franchise amount is d. Then (2) for the surplus process $\big(X_t^{(d)}(x)\big)_{t\ge 0}$ can be rewritten as follows

$$X_t^{(d)}(x) = x + c^{(d)}t - \sum_{i=1}^{N_t} Y_i^{(d)}, \quad t \ge 0. \tag{3}$$

Let $\varphi^{(d)}(x)$ denote the corresponding infinite-horizon survival probability. In what follows, we deal with exponentially distributed claim sizes only. In this case we have

$$F^{(d)}(y) = \begin{cases} 0 & \text{if} \quad y < 0, \\ 1 - e^{-d/\mu} & \text{if} \quad 0 \le y < d, \\ 1 - e^{-y/\mu} & \text{if} \quad y \ge d. \end{cases}$$

The next theorem gives analytic expressions for $\varphi^{(d)}(x)$. It is easily seen that in this case the c.d.f. of the insurance compensation is a sum of absolutely continuous and discrete components. That is why analytic expressions for the survival probability turn out different on certain intervals.

To formulate the next theorem, introduce the constants

$$\gamma = (1+\theta)(\mu+d), \quad C_{1,1} = \frac{\theta}{1+\theta}, \quad A_{2,0} = -\frac{\theta}{(1+\theta)(\gamma+\mu)}\,e^{-d/\gamma},$$

$$C_{2,1} = \frac{\theta}{1+\theta}\left(1 + \frac{\gamma\mu + d(\gamma+\mu)}{(\gamma+\mu)^2}\,e^{-d/\gamma}\right), \quad C_{2,2} = -\frac{\theta\gamma\mu}{(1+\theta)(\gamma+\mu)^2}\,e^{d/\mu}.$$

Moreover, let the constants $A_{n+1,\,i}$, $0 \le i \le n-1$, be given in a recurrent way by formulas

$$A_{n+1,\,n-1} = -\frac{A_{n,\,n-2}}{n(\gamma+\mu)}\,e^{-d/\gamma}, \quad n \ge 2,$$

$$A_{n+1,\,j} = -\frac{1}{\gamma+\mu}\left[(j+2)\gamma\mu A_{n+1,\,j+1} + \frac{1}{j+1}\left(\sum_{i=j-1}^{n-2} A_{n,\,i}\,(-d)^{i-j+1}\binom{i+1}{j}\right)e^{-d/\gamma}\right],$$
$$1 \le j \le n-2, \quad n \ge 3,$$

$$A_{n+1,\,0} = -\frac{1}{\gamma+\mu}\left[2\gamma\mu A_{n+1,\,1} + \left(C_{n,\,1} + \sum_{i=0}^{n-2} A_{n,\,i}\,(-d)^{i+1}\right)e^{-d/\gamma}\right], \quad n \ge 2.$$

Next, let the constants $B_{n+1,\,i}$, $0 \le i \le n-2$, be given in a recurrent way by formulas

$$B_{3,\,0} = \frac{C_{2,\,2}}{\gamma+\mu}\,e^{d/\mu},$$

$$B_{n+1,\,n-2} = \frac{B_{n,\,n-3}}{(n-1)(\gamma+\mu)}\,e^{d/\mu}, \quad n \ge 3,$$

$$B_{n+1,\,j} = \frac{1}{\gamma+\mu}\left[(j+2)\gamma\mu B_{n+1,\,j+1} + \frac{1}{j+1}\left(\sum_{i=j-1}^{n-3} B_{n,\,i}\,(-d)^{i-j+1}\binom{i+1}{j}\right)e^{d/\mu}\right],$$
$$1 \le j \le n-3, \quad n \ge 4,$$

$$B_{n+1,\,0} = \frac{1}{\gamma+\mu}\left[2\gamma\mu B_{n+1,\,1} + \left(C_{n,\,2} + \sum_{i=0}^{n-3} B_{n,\,i}\,(-d)^{i+1}\right)e^{d/\mu}\right], \quad n \ge 3.$$

Finally, let the constants $C_{n+1,\,1}$ and $C_{n+1,\,2}$ be given by formulas

$$\begin{aligned}
C_{n+1,\,1} = C_{n,\,1} &+ \frac{\gamma\mu(A_{n,\,0} - A_{n+1,\,0})}{\gamma+\mu} \\
&+ \sum_{i=0}^{n-3}\left(A_{n,\,i} - A_{n+1,\,i} + \frac{(i+2)\gamma\mu(A_{n,\,i+1} - A_{n+1,\,i+1})}{\gamma+\mu}\right)(nd)^{i+1} \\
&+ \left(A_{n,\,n-2} - A_{n+1,\,n-2} - \frac{n\gamma\mu A_{n+1,\,n-1}}{\gamma+\mu}\right)(nd)^{n-1} - A_{n+1,\,n-1}\,(nd)^{n} \\
&+ \frac{\gamma\mu}{\gamma+\mu}\left(\sum_{i=0}^{n-3}(i+1)(B_{n,\,i} - B_{n+1,\,i})(nd)^{i} - (n-1)B_{n+1,\,n-2}\,(nd)^{n-2}\right) \\
&\times \exp\left(-nd\,\frac{\gamma+\mu}{\gamma\mu}\right), \quad n \ge 2,
\end{aligned}$$

$$C_{3,\,2} = C_{2,\,2} + \frac{\gamma\mu B_{3,\,0}}{\gamma+\mu} - 2dB_{3,\,0} + \frac{\gamma\mu(A_{3,\,0} - A_{2,\,0} + 4dA_{3,\,1})}{\gamma+\mu}\,\exp\left(2d\,\frac{\gamma+\mu}{\gamma\mu}\right),$$

$$\begin{aligned}
C_{n+1,\,2} = C_{n,\,2} &+ \frac{\gamma\mu(B_{n+1,\,0} - B_{n,\,0})}{\gamma+\mu} \\
&+ \sum_{i=0}^{n-4}\left(B_{n,\,i} - B_{n+1,\,i} + \frac{(i+2)\gamma\mu(B_{n+1,\,i+1} - B_{n,\,i+1})}{\gamma+\mu}\right)(nd)^{i+1} \\
&+ \left(B_{n,\,n-3} - B_{n+1,\,n-3} + \frac{(n-1)\gamma\mu B_{n+1,\,n-2}}{\gamma+\mu}\right)(nd)^{n-2} - B_{n+1,\,n-2}\,(nd)^{n-1} \\
&+ \frac{\gamma\mu}{\gamma+\mu}\left(\sum_{i=0}^{n-2}(i+1)(A_{n+1,\,i} - A_{n,\,i})(nd)^{i} + nA_{n+1,\,n-1}\,(nd)^{n-1}\right) \\
&\times \exp\left(nd\,\frac{\gamma+\mu}{\gamma\mu}\right), \quad n \ge 3.
\end{aligned}$$

Theorem 3.1 *Let the surplus process* $\left(X_t^{(d)}(x)\right)_{t\ge 0}$ *follow* (3) *under the above assumptions with* $0 < d < +\infty$*, and the claim sizes be exponentially distributed with mean* μ*. Then*

$$\varphi^{(d)}(x) = \varphi^{(d)}_{n+1}(x) \quad \textit{for all} \quad x \in [nd, (n+1)d), \quad n \ge 0,$$

where

$$\varphi^{(d)}_1(x) = C_{1,\,1}\,e^{x/\gamma},$$

$$\varphi^{(d)}_2(x) = \left(C_{2,\,1} + A_{2,\,0}\,x\right)e^{x/\gamma} + C_{2,\,2}\,e^{-x/\mu},$$

$$\varphi^{(d)}_{n+1}(x) = \left(C_{n+1,\,1} + \sum_{i=0}^{n-1} A_{n+1,\,i}\,x^{i+1}\right)e^{x/\gamma} + \left(C_{n+1,\,2} + \sum_{i=0}^{n-2} B_{n+1,\,i}\,x^{i+1}\right)e^{-x/\mu}, \quad n \ge 2.$$

The proof of Theorem 3.1 is given in Ragulina (2011).

References

S. Asmussen. *Ruin Probabilities*. World Scientific, Singapore, 2000.

B.V. Bondarev and E.Yu. Ragulina. On the finite-time nonruin probability of an insurance company with investments in the financial (B, S)-market. *Cybernetics and Systems Analysis*, 48(5):736–748, 2012.

J. Grandell. *Aspects of Risk Theory*. Springer, New York, 1991.

O. Ragulina. On the survival probability of an insurance company in the classical risk model with use of a conditional deductible and a liability limit. *Applied Statistics. Actuarial and Financial Mathematics*, 2011(1):27–46, 2011. In Ukrainian.

O. Ragulina. Maximization of the survival probability by franchise and deductible amounts in the classical risk model. *Springer Optimization and Its Applications*, 90:287–300, 2014.

T. Rolski, H. Schmidli, V. Schmidt, and J. Teugels. *Stochastic Processes for Insurance and Finance*. John Wiley & Sons, Chichester, 1999.

De Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten coördineert jaarlijks tot 25 wetenschappelijke bijeenkomsten, ook contactfora genoemd, in de domeinen van de natuurwetenschappen (inclusief de biomedische wetenschappen), menswetenschappen en kunsten. De contactfora hebben tot doel Vlaamse wetenschappers of kunstenaars te verenigen rond specifieke thema's.

De handelingen van deze contactfora vormen een aparte publicatiereeks van de Academie.

Contactforum "Actuarial and Financial Mathematics Conference" (5-6 February 2015, Prof. M. Vanmaele)

Deze handelingen van de "Actuarial and Financial Mathematics Conference 2015" geven een inkijk in een aantal onderwerpen die in de editie van 2015 van dit contactforum aan bod kwamen. Zoals de vorige jaren handelden de voordrachten over zowel actuariële als financiële onderwerpen en technieken met speciale aandacht voor de wisselwerking tussen beide. Deze internationale conferentie biedt een forum aan zowel experten als jonge onderzoekers om hun onderzoeksresultaten ofwel in een voordracht ofwel via een poster aan een ruim publiek voor te stellen bestaande uit academici uit binnen- en buitenland alsook collega's uit de bank- en verzekeringswereld.