

Novel tools in quantitative risk management

Weihao Choo

Department of Applied Finance and Actuarial Studies
Macquarie University

PhD dissertation

2015

Abstract

This thesis proposes novel methods to analyze risk and dependence across a joint probability distribution. It is well known in finance and insurance that risk and dependence are vastly different in the tails compared to the rest of the distribution. Tails characterise events such as market crisis and natural catastrophes, and contribute to a significant portion of overall risk and dependence. However typical measures of risk and dependence capture the overall result and mask variations across the probability distribution.

Random quantities are partitioned into infinitesimal layers capturing outcomes of various magnitude and likelihood. Risk and dependence are then measured across layers using established methods such as distortion and correlation. Layers are standard constructs representing (re)insurance coverage, capital consumption and shortfall, derivative payouts, and tranches of collateralised debt obligations. This thesis expresses layer endpoints using percentiles or more commonly known as Values-at-Risk (VaRs), hence each layer occupies a relative position in the probability distribution.

This thesis also extends distortion risk measurement by capturing upside risk in addition to downside risk. In financial and insurance markets with strong competition and limited availability of capital, an explicit view of upside risk is required to reflect opportunity costs.

Developments in this thesis formalise existing, and reveal new, insights to risk and diversification. For example the framework explains weak diversification in financial and insurance markets despite moderate correlations overall. The framework also deals with problems such as setting capital buffers, reinsurance purchase and assessing the credit quality of debt tranches. These insights arise from a deeper understanding of how risk and dependence varies across a probability distribution.

Proposed methods apply consistent concepts such as VaRs, distortion and layers, and hence form a coherent analytical framework. These concepts are well established and hence the resulting framework integrates and expands current disparate approaches. The proposed framework is a complete tool to quantitative risk management, by first analysing risk and dependence when imperfectly dependent random quantities are aggregated, and then guiding strategies to optimally manage and reduce risk.

Acknowledgements

I would like to express my gratitude to my supervisor, Professor Piet de Jong, for his critique of my research ideas which strongly guided the development of this thesis. Piet also taught me that research ideas are most convincingly presented when expressed in the simplest form.

My thanks also go to my wife, Qinyi Phua, who strongly supported my research work over the years despite our busy work schedules.

Finally I thank the examiners for their valuable feedback.

Contents

1	Thesis contributions and the literature	4
2	Layer dependence	26
3	Mean and risk densities	47
4	Analyzing systematic risk and diversification	72
5	Tradeoff premiums	89
6	Integrating proposed tools and future research areas	111
	Complete bibliography	115

Chapter 1

Thesis contributions and the literature

1.1 Overview of this thesis

This thesis presents papers discussing novel solutions to four critical problems in quantitative risk management. These problems relate to analyzing risk, dependence and diversification in a joint probability distribution, and forming balanced risk measures capturing both upside and downside.

Section 1.2 provides an overview of quantitative risk management. Section 1.3 identifies the four problems addressed in this thesis. Section 1.4 introduces well known concepts of layers, Value-at-Risk and distortion risk which are critical to remaining discussion in this chapter. Subsequent sections delve into each problem, explaining its role and importance in quantitative risk management, summarising and critiquing current approaches, and outlining proposed solutions.

Proposed solutions modify, enhance and combine established approaches in the literature. Proposed solutions to the four problems also integrate to form a coherent quantitative risk management framework.

Subsequent chapters of this thesis are structured as follows:

- Chapter 2 presents the paper “*Layer dependence as a measure of local dependence*,” which measures local dependence across percentiles of a joint probability distribution.
- Chapter 3 presents the paper “*Mean and risk densities and their applications to risk management*,” which analyzes how mean and risk varies

across layers of a random loss.

- Chapter 4 presents the paper “*Insights to systematic risk and diversification across a joint probability distribution*,” integrating layer dependence and risk density curves to analyze systematic risk and diversification when random, imperfectly dependent losses are aggregated.
- Chapter 5 presents the paper “*The tradeoff insurance premium as a two-sided generalisation of the distortion premium*,” an extension of distortion risk measurement which explicitly reflects upside risk in addition to downside risk and is related to cumulative prospect theory.
- Chapter 6 concludes this thesis by discussing how proposed approaches form an integrated approach to quantitative risk management, and outlining future research areas.

1.2 Quantitative risk management

Risk generally refers to unpredictable outcomes having a positive or negative impact¹. Risk is inherent in financial and insurance markets. For example banks are exposed to future uncertain movements in stock markets, interest rates, credit defaults and the state of the economy. Insurers are impacted by the same factors and insurance claims volatility particularly from man-made and natural catastrophes.

Quantitative risk management quantifies and manages risk. McNeil et al. (2005) discusses the history, issues and common techniques of quantitative risk management. Risk quantification is not only driven by the probabilistic behaviour of outcomes, but also risk perception. High risk aversion magnifies risks, whereas risk neutrality dismisses risks. Quantitative risk management typically focuses on adverse rather than favourable outcomes. For instance banks and insurers calculate the risk of unexpected losses and hold capital buffers accordingly. Other examples of focusing on adverse outcomes are the purchase of reinsurance cover for catastrophe losses, using derivatives to hedge portfolio losses, putting risk discounts on securities prices, and early warning systems for adverse financial movements.

Quantitative risk management is increasingly important for financial and insurance companies. For example Solvency II (Eling et al. 2007) imposes formal risk management requirements on insurers, including the need to hold

¹A formal definition of risk is discussed later in this thesis, using distortion.

capital commensurate with market, credit, underwriting, catastrophe and operational risks. Basel II (Engelmann and Rauhmeier 2006) and more recently Basel III (King and Tarbert 2011) impose equally if not more stringent requirements on banks. Major insurers and banks typically develop stochastic models of their business and manage risks according to simulated results (Kaufmann et al. 2001). Quantitative risk management is also an important component of the broader Enterprise Risk Management (ERM) (Nocco and Stulz 2006), where quantifiable and non-quantifiable risks are managed holistically to take advantage of synergies and diversification.

Quantitative risk management is a broad area covering a wide range of topics. This thesis focusses on measuring and analyzing risk and dependence. Specific interest areas are outlined in the next section. Subsequent sections discuss the importance of these interest areas, current approaches and their limitations, and proposed solutions. The following is a brief description of dependence and risk measurement:

- Risk measurement quantifies the magnitude and likelihood of adverse outcomes. Quantified risks are critical inputs to risk management decisions. For example banks hold capital buffers which increase with risks of stock market downturns, credit defaults and other adverse financial movements (King and Tarbert 2011). Reinsurance purchases by insurers refer to the extent of catastrophe risk. Value-at-Risk is a common risk measure but has shortcomings, and refined measures satisfying coherence properties have been proposed (McNeil et al. 2005).
- Dependence measurement captures the degree of association between random outcomes, and is typically based on linear correlation (McNeil et al. 2005). Dependence is a common feature of financial and insurance markets. For example returns from various stocks are dependent, particularly when in distress (Rodriguez 2007). Natural catastrophes create significant losses across insurers and their portfolios such as property, motor and liability.

Risk and dependence measurement become intimately connected when random losses from different sources are aggregated. Imperfect dependence creates diversification benefits: a reduction in aggregate risk when favourable outcomes in a group of losses offset adverse outcomes in other losses. On the other hand strong dependence leads to catastrophic consequences when loss outcomes are simultaneously adverse. Effective quantitative risk management exploits diversification. However diversification is often treated as an anecdotal phenomenon and is insufficiently analysed in the literature.

Risk and dependence are rarely static across a joint probability distribution, and change significantly in the tails capturing extreme outcomes such as market crashes and catastrophes. Reflecting varying risk and dependence behaviour is hence critical in quantitative risk management, but often insufficiently acknowledged in current approaches.

1.3 Specific areas addressed in this thesis

This thesis proposes novel solutions to the following four specific problems in risk and dependence measurement. Remaining sections of this chapter delve into each problem. The four problems are:

- Measuring local dependence between random losses. Of interest is dependence at various parts of the joint distribution, rather than overall dependence. To illustrate the importance of measuring local dependence, consider stock returns which are highly dependent when in distress but less dependent otherwise under normal circumstances. Focussing on overall dependence understates tail dependence and leads to insufficient capital buffers against market risk.
- Decomposing the mean and risk of a random loss and understanding contributions by various parts of its probability distribution. For example, significant portions of the mean and risk of a right skewed loss distribution are concentrated in the upper tail. Understanding mean and risk contributions yields targeted risk management strategies to achieve an optimal mean–risk combination.
- Analyzing risk and diversification when imperfectly dependent losses are aggregated. Aggregation reduces risk. For example the stock market index is less volatile than its component stocks. Systematic risk remains after diversification. Of interest are key sources of systematic risk and diversification in the joint distribution so that, similar to the previous problem, optimal strategies can be formulated.
- Reflecting upside in addition to downside when measuring risk. Despite the widespread focus on downside risk in the literature, upside risk is also important and neglecting it leads to opportunity losses. For example, excessive focus on downside risk leads to an uncompetitive insurance premium, whereas acknowledging and reflecting favourable outcomes achieves a more balanced premium.

For the first three problems described above, the literature extensively discusses measures of *overall* dependence, risk and diversification, but offers less insights into how dependence, risk and diversification *vary* across the probability distribution. For example linear correlation does not indicate differing dependence between random losses at their 50th, 75th and 99th percentiles. These insights are important because risk management strategies usually target parts of loss distributions rather than their entirety. For example, reinsurance covers losses above the excess. Capital protects against losses below itself. Derivatives hedge movements in a defined region.

Problems discussed in this thesis, and their proposed solutions, are generic and apply to any situation involving risk and dependence. Quantifiable risks typically involve monetary quantities in insurance and finance, but can also be extended to quantities such as the amount of rainfall, passenger volume and population size. These non-monetary quantities are uncertain and involve risk, and each depends on a number of other uncertain factors.

Marginal behaviour of random variables is assumed known. The modeling of marginal behaviour is well covered in the literature. Parametric or non-parametric distributions are typically first fitted to data, and tests are then performed to assess the suitability of the fit. References include Feller (2008) and Hogg and Klugman (2009). Extreme value theory (Kotz and Nadarajah 2000) addresses tail behaviour representing extreme outcomes.

A single period is assumed. Therefore the time series behaviour of random variables, for example daily stock market returns over one year, are ignored. Typical time series models in quantitative risk management, such as Autogressive Moving Average (ARMA) and Generalised Autoregressive Conditional Heteroskedasticity (GARCH) models, are discussed in McNeil et al. (2005) and Lamoureux and Lastrapes (1990).

1.4 Established concepts used in this thesis

Common to proposed solutions in this thesis are concepts of layers, Values-at-Risk and distortion risks. Hence proposed solutions naturally integrate to form a coherent quantitative risk management framework. The final chapter discusses the integration. This section introduces each concept, setting the scene for the discussion of proposed solutions in subsequent sections.

Any loss can be decomposed into additive layers. The $[a, b]$ -layer of a loss is the excess over a , capped at $b - a$. In particular the $[a, \infty]$ -layer is the unbounded excess above a , and the $[0, b]$ -layer is the loss capped at b . Layers

are standard insurance and financial constructs. For example insurance and reinsurance cover a layer of the loss defined by an excess and limit. Capital buffers divide losses into two layers: capital consumed and capital shortfall. Derivatives and collateralised debt obligations also involve layers, sometimes known as tranches with attachment and detachment points. Bailouts of a distressed company cover a layer of losses suffered. High layers capture rare, extreme outcomes and low layers characterise common, attritional outcomes. Statistical properties of layers are discussed in Campana and Ferretti (2014), Wang (1998), Wang (1995) and Miccolis (1977). Lee (1988) adopts a graphical approach to explain key concepts and results. Insurance pricing of loss layers is discussed in Evans (2001) and Salzmänn (1963). Mandel et al. (2012) and Duffie and Garleanu (2001) discuss tranches in collateralised debt obligations and how they partially enhance the credit quality of the debt.

Values-at-Risk (VaRs) are percentiles of a loss distribution (McNeil et al. 2005). Writing loss outcomes as VaRs shows their relative position in the probability distribution. For example the 50% VaR is the median or middle outcome, whereas the 75% VaR is the 75th largest out of 100 outcomes. VaRs adjust to the shape and scale of the loss distribution, and are hence comparable across loss distributions. In contrast absolute dollar amounts are not probabilistic and may be a commonly exceeded outcome in a loss distribution but a rare, extreme outcome in another. Similar to layers, VaRs are standard insurance and financial constructs. For example Solvency II insurance regulation applies 90% and 99.5% VaRs (Eling et al. 2007) to capital requirements. Banking regulations Basel II and more recently Basel III also reference VaRs in risk measurement (Chernobai et al. 2008).

Lastly distortion risk (Wang 1996) is risk formed by the difference between the average outcome under a distorted, conservative probability distribution and the original, objective distribution². Choo and De Jong (2009) shows that distortion risks are proportional to loss volatility and risk aversion, and are equivalent to loss aversion risks and spectral risks (Acerbi 2002), both defined as weighted average of VaRs. Examples of distortion risks are discussed in Wang (1995), Wang (2000) and Choo and De Jong (2009), and include the proportional hazards risk, conditional-tail-expectation and expected-maximal-loss. Distortion risks are coherent (Artzner et al. 1999): positively homogenous, translation invariant, monotonic and subadditive.

²The original formulation only includes the distorted average and does not take the difference with the original average.

1.5 Local dependence and its measurement

1.5.1 Importance of local dependence

Dependence is inherent in financial, insurance, commodities and other markets. For example interest rates, unemployment, stock market returns and credit defaults are interdependent and partially rely on the state of the economy. In addition these quantities are linked to corresponding quantities in other geographies and countries. Insurance claims from motor, property and liability classes are also interdependent, due to their reliance on common economic and social factors, as well as weather patterns.

Measuring and subsequently modeling dependence is thus a critical part of quantitative risk management. To illustrate the importance of dependence, suppose a bank holds capital against the risk of stock market crashes and credit defaults by borrowers. If these two risk factors are strongly dependent, then large amounts of capital are required to cover market crashes and credit defaults occurring simultaneously. On the other hand if dependence is weak, then required capital is significantly lower since market crashes are unlikely to coincide with credit defaults – a diversification benefit.

Dependence typically varies across the joint distribution, giving rise to the need for *local* dependence measures. For example moderate returns from various stocks may be weakly dependent but extreme returns are strongly dependent (Rodriguez (2007) and Hartmann et al. (2004)). In insurance, attritional losses from various lines are weakly dependent, however a catastrophic event creates equally extreme losses across all lines. Failure to acknowledge varying dependence, and applying the average dependence across the joint distribution, leads to for example underestimation of tail events where extreme outcomes from various risk factors occur simultaneously. The 2008 global financial crisis (Kolb 2010) is a relevant case study.

1.5.2 Measures of overall and local dependence

Dependence measures lie between ± 1 , with -1 indicating perfect negative dependence (countermonotonicity) and 1 indicating perfect positive dependence, (comonotonicity) (Dhaene et al. 2002). Positive dependence implies random variables tend to increase simultaneously, and vice versa for negative dependence. Dependence is measured either between random variables in their original values or percentile rank transforms. The latter leads to “rank dependence” which is calculated from the copula (Nelson 1999). The former

is distorted by marginal distributions and only covers a subset of $[-1, 1]$, even with perfect dependence (McNeil et al. 2005). Rank dependence is free of distortion by marginal distributions, and covers $[-1, 1]$ completely.

Measures of overall dependence, including Pearson’s correlation, Spearman’s ρ and Kendall’s τ (McNeil et al. 2005), do not characterise the dependence *structure* of the joint distribution. As explained in the previous subsection, dependence typically varies across joint distributions of financial and insurance quantities, and tail dependence is common. Failure to measure and hence properly model the dependence structure results in ineffective risk management, such as holding insufficient capital.

Local dependence measures address the drawback of overall dependence measures, by measuring dependence at various points or parts of the joint distribution. The set of local dependence values calculated across the entire joint distribution characterises the dependence structure. Local dependence measures are a summary of the joint distribution or copula, but are less summarised than overall dependence measures. The following describes current local dependence measures and their shortcomings:

- Tail concentration (Durante et al. (2014), Venter (2002)) is the conditional probability of percentile ranks falling in identical tail regions, and is calculated from the diagonal section of the copula (Fredricks and Nelsen 1997). Varying the tail region yields dependence at various parts of the joint distribution. Calculating the conditional probability in extreme tails yields coefficients of tail dependence by Joe (1997).

Tail concentration, being a probability, excludes actual values of random variables. As a result tail concentration does not always vary coherently across the joint distribution. In addition negative dependence is not obviously shown from tail concentration values.

- Correlation curve (Bjerve and Doksum 1993) applies regression principles and measures dependence between a random variable and a neighbourhood of another. The measurement combines conditional variances and changes in conditional expectations across neighbourhoods. A larger change in conditional expectation or lower conditional variance implies higher local dependence and vice versa.

Despite satisfying several coherence properties, correlation curve is difficult to calculate on data, due to the reliance on conditional expectation and conditional variance in an infinitesimal neighbourhood. Calculated values are volatile even for large samples.

- Bairamov et al. (2003), Jones (1996) and Holland and Wang (1987) discuss bivariate local dependence measures. These capture dependence between different neighbourhoods of two random variables. In contrast tail concentration and correlation curves are univariate measures.

Bivariate local dependence measures maintain the dimension of the bivariate joint distribution, whereas univariate measures summarise and extract dependence information. This thesis focuses on univariate local dependence measures.

Apart from the shortcomings described above, current local dependence measures are generally disconnected from overall dependence measures such as Spearman's ρ and Kendall's τ . As a result calculated local dependence values may be consistently higher or lower than overall dependence. Layer dependence, a local dependence measure proposed in this thesis, is consistent with Spearman's ρ . Layer dependence also satisfies several other practical properties. The next subsection discusses layer dependence.

Specifying local dependence values, that is the dependence structure, is an intuitive approach to model copulas. For example Fredricks and Nelsen (1997) and Durante et al. (2006) discuss copulas constructed from a given diagonal section which is in turn implied from tail concentration values. However local dependence measurement needs to be appropriate in order to construct a copula exhibiting the desired dependence structure.

1.5.3 Layer dependence as a local dependence measure

This thesis proposes a local dependence measure called “layer dependence.” Layer dependence captures rank dependence and is hence calculated entirely from the copula. As discussed in the previous subsection, rank dependence is free from distortion by marginal distributions and is preferred over dependence calculated between original values. Chapter 2 contains a paper defining, illustrating and analyzing layer dependence.

Layer dependence accurately reflects local dependence, and satisfies important properties described below. For a copula exhibiting weak lower tail dependence and strong upper tail dependence, layer dependence increases from near 0 to near 1. Suppose observations are simulated from a copula and plotted on the unit square. Then layer dependence is higher at any point along the 45° line if observations are tightly clustered around the point, and lower if observations are dispersed.

The definition and concept of layer dependence are as follows. As the name suggests, layer dependence involves layers discussed in §1.4 and is defined from the covariance between the percentile rank of a random variable and an infinitesimal layer of another. The covariance is scaled to yield a value of one if random variables are comonotonic. An infinitesimal layer of a random variable captures its movements at a point, hence layer dependence is dependence between a random variable and local movements of another.

Layer dependence satisfies practical properties similar to Spearman's ρ : between -1 and 1 , constant and equal to -1 , 0 and 1 for countermonotonic, independent and comonotonic random variables, sign switching when ranking order reverses, and taking on higher values when dependence is stronger.

Layer dependence has a direct relationship with Spearman's ρ : taking a weighted average of layer dependence values across all layers is Spearman's ρ . This relationship is intuitive and appealing – averaging local dependence yields overall dependence. Weights attached to layer dependence values are quadratic, peaking at the median layer and approaching zero at the tails. These weights imply Spearman's ρ understates tail dependence, a critical characteristic of insurance and financial quantities. More appropriate overall dependence measures are formed by averaging layer dependence values using weights reflecting the importance of dependence at various layers. These alternate measures satisfy similar coherence properties as Spearman's ρ if weights are non-negative and integrate to 1.

Layer dependence values are broadly similar to correlation curve and tail concentration values. However layer dependence enhances and outperforms these measures in a number of ways. Layer dependence refines tail concentration by reflecting average dispersion between percentile ranks, and is hence a better measure of local dependence. In addition negative dependence is indicated by negative layer dependence values. Layer dependence calculations on data only involve conditional tail expectations and are hence simpler and more stable than correlation curves. Lastly layer dependence is coherent and is directly connected to Spearman's ρ as discussed above.

Layer dependence captures tail dependence consistently with coefficients of tail dependence by Joe (1997). Layer dependence at extreme layers and coefficients of tail dependence are both one when tail dependence is perfect. Perfect tail dependence requires random variables to simultaneously attain their maximum or minimum values.

Calculating layer dependence at the first instance from a parametric copula or data extracts essential and interpretable information: the dependence structure. This is often more informative than positing a parametric copula,

as the implication of its parametric form and parameters on the dependence structure is indirect. Similar problems apply when data is scarce or volatile and the dependence structure is masked. Computing layer dependence from data facilitates the selection and fitting of an appropriate parametric copula. Denuit et al. (2005), Genest and Rivest (1993) and Oakes (1989) discusses the fitting of parametric copulas.

1.6 Mean and risk decomposition

1.6.1 Risk measurement and risk behaviour

A key exercise in quantitative risk management is calculating risk values of random losses such as from insurance claims, credit defaults and stock market downturns. Risk values, or risks, enable random losses to be assessed and compared, and are key inputs to risk management decisions. For example insurance risk drives premium loadings, reinsurance purchases and capital buffers. Credit and market risks influence lending margins, capital buffers, investment decisions and hedging positions for derivatives.

Risk measures assign risk values to random losses based on their probability distribution, and are also known as premium principles. This thesis applies distortion risk measures described in §1.4. McNeil et al. (2005) and Young (2004) discuss common, specific risk measures such as standard deviation, Esscher premium, Value-at-Risk and conditional-tail-expectation. Risk typically increases with the volatility and skewness of the loss distribution. Risk perception also influences risk: greater risk aversion leads to greater risk and vice versa, whilst risk neutrality implies zero risk.

This thesis studies the risk *behaviour* of random losses: how risk varies across the probability distribution. Similar to dependence discussed in §1.5, random losses may have equal overall risk but different risk behaviour, and hence require different risk management strategies. Any measure of overall risk, regardless of its sensitivity to the loss distribution, does not completely characterise the risk profile of a loss. A “risk density curve” is thus required to capture risk across the loss distribution. For a skewed loss distribution, risk density is low for moderate outcomes but high for extreme outcomes such as catastrophic insurance losses or stock market crashes.

Analyzing risk behaviour is important to form optimal and targeted risk management strategies. For example, excess-of-loss reinsurance is purchased to cover losses above a threshold, and an optimal threshold balances risks of

covered and retained losses and coverage cost. Setting an appropriate capital buffer involves similar considerations. In finance, pricing collateralised debt obligations and derivatives requires an understanding of the risk of payouts in different tranches.

1.6.2 Risk measurement across loss layers

Risk behaviour is captured by measuring risk across layers of a random loss using a consistent risk measure. As described in §1.4, layers are standard insurance and financial constructs. A loss is formed by additive layers with low layers representing attritional, likely outcomes and high layers capturing extreme, rare outcomes. As layers of the same underlying loss are comonotonic, risks of individual layers add to overall risk if the risk measure is additive over comonotonic random variables³.

The following summarises current literature on risk measurement across loss layers. The literature in this area is arguably less established compared to measures of overall risk. As explained in the previous subsection, analyzing risk behaviour is important to gain insights and form targeted risk management strategies.

- Wang (1995) calls the survival function the premium layer density as it computes the mean value or premium of infinitesimal layers. The premium layer density is distorted to deliver risk-adjusted premiums of layers. Integrating the distorted premium layer density forms overall risk. Wang (1995) applies distorted premium layer densities to investigate premiums when the limit of an insurance contract is increased.
- Ladoucette and Teugels (2006) analyzes risk measures including Value-at-Risk, variance, and coefficient of variation across layers of an insurance loss. The analysis is also extended to consider layers of a random sum of insurance losses.
- Hürlimann (1998) uses a distortion risk measure involving a two-stage loss transformation and the Hardy-Littlewood pricing principle. This distortion risk measure is used to construct distribution-free layer premiums satisfying several practical properties.
- Certain measures of overall risk focus on tail layers of a loss and hence indicate risk behaviour. For example conditional-tail-expectation takes

³Distortion risk measures are additive over comonotonic random variables.

the expected value of losses beyond a threshold. Adding the scaled variance of these losses yields modified–tail–variance (Furman and Zitikis 2008a). Varying the threshold indicates risks of various tail layers.

- Salzmänn (1963) and Evans (2001) perform empirical studies of appropriate premiums rates across loss layers for property insurance. Finger (1976) performs a similar study, assuming a lognormal loss distribution.

1.6.3 Mean and risk densities over VaR layers

Similar to Wang (1995), this thesis constructs mean and risk densities indicating the mean and distortion risk of infinitesimal loss layers. A novel and critical change to loss layers is applied: layer endpoints are expressed in VaR rather than dollar terms. Hence mean and risk densities are defined over the unit interval indicating the percentile rather than over the original loss scale. As noted in §1.4, VaRs occupy relative positions in the probability distribution and adjust to its shape and scale. The layer from the 50th VaR to 75th VaR, for example, captures the top 50% of losses and the top 25% of losses are capped at the 75th percentile. VaR layers are therefore comparable between loss distributions. In contrast original dollar layers may be attritional or rare depending on the loss distribution.

Mean and risk densities are critical constructs in explaining the mean and risk behaviour of a random loss. As layers of the same loss are comonotonic, and distortion risks are additive over comonotonic random variables, integrating risk densities over any subset of the unit interval yields the risk of a larger layer. In addition the entire area under the risk density is overall distortion risk. The same result applies to mean densities, although the comonotonicity condition does not need to hold. Mean and risk densities are analogous to probability densities, representing quantities over an infinitesimal area, with integration yielding the same over a larger area.

Defining mean and risk densities over VaR layers delivers the following properties and results. They also formalise current risk insights.

- Mean and risk densities across loss distributions are graphed over the unit interval on the horizontal axis. Scale effects are isolated and captured by the vertical axis. In contrast scale effects are shown in horizontal and vertical axes if layers are defined on the original loss scale.
- The relative risk of a VaR layer (the ratio between risk and mean densities) is monotonic increasing and does not involve the loss distribution. Hence higher VaR layers are always riskier than lower VaR layers.

- The mean density characterises local skewness or volatility compared to an exponential distribution. As the mean density is flat for an exponential distribution, an increasing mean density at any layer indicates greater skewness comparatively.
- The overall distortion risk of a loss relative to its mean is an average of risk ratios across layers weighted by the mean density. As risk ratios are increasing and independent of the loss distribution, losses with high tail volatility (increasing mean density) are relatively riskier.
- Mean and risk densities defined over VaR layers, when applied to risk management problems such as capital setting, yield solutions expressed as VaRs instead of dollars. This is consistent with the standard use of VaRs in finance and insurance as described in §1.4.

Mean and risk densities provide solutions and insights to common risk management problems. These problems include assessing insurance coverage under different excess and limits, setting optimal capital buffers to reflect risks of capital shortfall and surplus, using reinsurance to alter a loss distribution, and comparing the credit quality of various debt tranches. Although these problems can be solved using current statistical approaches, mean and risk densities provide more elegant solutions and deliver additional insights.

Mean and risk densities integrate layer dependence discussed in the previous section with the analysis of systematic risk and diversification discussed in the next section. Hence defining mean and risk densities over VaR layers forms an integrated analytical framework.

1.7 Systematic risk and diversification

1.7.1 Concepts of systematic risk and diversification

Risk reduces or diversifies when imperfectly dependent random variables are aggregated. For example the return on a market index is less volatile than returns of components forming the index. Average investment returns also stabilise over time. Pooling insurance claims reduces their volatility to an acceptable level. Diversification generally arises when adverse outcomes for a random variable are offset by favourable outcomes in other random variables. Therefore diversification weakens when random variables become more dependent and likely to be simultaneously adverse.

Diversification implies the risk of any random loss is split into diversifiable and non-diversifiable. The latter is known as systematic risk in finance (Luenberger 1998). Systematic risk contributes to aggregate risk, whereas diversifiable risk is eliminated upon aggregation. Of interest is the extent of diversification in each random loss with greater diversification reducing aggregate risk. In insurance, systematic risks add up to aggregate risk and are hence an allocation of aggregate risk to component random losses. Kalkbrener (2005) and Denault (2001) propose axioms for a coherent allocation including no undercut: the systematic risk of any component loss is less than its standalone risk before aggregation.

Similar to risk and dependence discussed in the previous two sections, systematic risk and diversification varies across the loss distribution. Developments in this thesis show that the extent of diversification at any part of the distribution is inversely related to *local* dependence with the aggregate random variable such as the index return or aggregate insurance loss. Current literature, further discussed in the next subsection, makes this observation usually in an intuitive manner rather than under a formal setting.

Developing insights to systematic risk and diversification yields risk management strategies aimed at reducing aggregate risk effectively, by targeting areas of joint loss distributions with high systematic risk and low diversification. For example, excess-of-loss reinsurance is purchased to only cover tails of losses with high rather than low systematic risk. Or consider a company comprising of business units. The systematic risk of a business unit drives its risk-adjusted performance and remuneration. Hence it is critical for every business unit to understand sources of its systematic risk and ways to maximise diversification.

1.7.2 Current approaches to analyse systematic risk and diversification

Approaches to derive systematic or allocated risks are well established in the literature. An overview is shown below. Current approaches provide broad insights into drivers of systematic risk, usually the dependence between component random variables. This thesis provides further insights by analyzing systematic risk and diversification across loss distributions.

The following is an overview of current approaches to derive systematic or allocated risk and analyse diversification:

- In the capital asset pricing model (Luenberger (1998), Sharpe (1964)),

the systematic risk of a security is proportional to correlation between its return and the market return, whilst overall risk of the security is the standard deviation of its return. Hence diversification reduces with the dependence between security and market returns.

- Choo and De Jong (2010) provides similar insights to systematic risks as the capital asset pricing model. By applying the Euler allocation principle (Buch and Dorfleitner (2008), McNeil et al. (2005)) to aggregate distortion risk, the systematic risk of a component loss is its covariance with a function of the aggregate loss. Furman and Zitikis (2008b) and Tsanakas and Christofides (2006) show similar results for allocated and systematic risks.
- The Euler allocation principle satisfies coherence axioms described in Denault (2001) and Kalkbrener (2005). These axioms include no undercut (allocated risk is less than standalone risk), symmetry (equal allocation to random variables with equal risk contribution) and riskless allocation (no allocation to risk-free random variables). Applying game theory (Shapley 1974) yields consistent results. Other properties and applications of Euler allocation are discussed in Tasche (2007).
- Sherris (2006) and Myers and Read Jr (2001) derive allocations based on option values. Dhaene et al. (2012) proposes a general approach by minimising differences between allocated risk and losses. Van Gulick et al. (2012) argues against Euler allocation and performs an allocation by minimising expected shortfall in various portfolios. Cummins (2000) and Venter (2004) summarise and critique current allocations.

1.7.3 Proposed analytical framework for systematic risk and diversification

This thesis constructs a framework to analyze systematic risk and diversification across layers of component losses being aggregated. The framework expands risk densities in §1.6 and forms links with layer dependence in §1.5. Insights gained from the proposed framework are critical to managing and reducing aggregate risk.

Cornerstone to the proposed framework are systematic risk densities indicating systematic risks of infinitesimal VaR layers forming a component loss. These risk densities are akin to those discussed in §1.6 and identify risk contributions by various parts of the loss distribution. However systematic

risk densities allow for diversification and exclude diversifiable risk. Systematic risk is measured as per Choo and De Jong (2010) by applying Euler allocation to aggregate risk measured using distortion. Thus the systematic risk of a VaR layer is its covariance with a function of the aggregate loss. Since covariances are additive, integrating the systematic risk density over an interval yields systematic risks of a larger layer, and the entire area under the density is the overall systematic risk of the component loss as per Choo and De Jong (2010).

Calculating the ratio between systematic and standalone (i.e. before aggregation) risk densities reveals the lack of diversification in each VaR layer of component losses. Ratios close to 1 indicate minimal diversification at the layer, whereas ratios close to 0 or even negative indicate strong diversification at the layer. Ratios closely relate to layer dependence discussed in §1.5, between component and aggregate losses. Hence *local* dependence between component and aggregate losses drives the level of systematic risk and diversification across the probability distribution. Strong local dependence at a component loss layer increases its systematic risk and reduces diversification. In particular strong tail dependence leads to weak diversification and high systematic risk at high VaR layers, and complete dependence implies zero diversification across all layers.

The negative relationship between local dependence and diversification across layers explains large systematic risks in financial markets exhibited for example during the 2008 global financial crisis (Kolb 2010). This is despite relatively weak to moderate overall correlations observed over time. Strong tail dependence in financial markets (Rodriguez (2007) and Hartmann et al. (2004)), coupled with skewed return distributions and large tail risks before diversification (Cont (2001), Hsieh (1988)), leads to significant amounts of non-diversifiable or systematic risk. Strong diversification below the tails does not significantly reduce overall risk as risks of return distributions are concentrated in the tails where diversification is weak.

Systematic risk and diversification insights developed in this thesis are critical to advanced risk management involving multiple random losses. By plotting and comparing standalone and systematic densities of a component loss, VaR layers with high and low diversification are identified and treated appropriately. For example, reinsurance, hedging and other risk mitigation actions target tails with large systematic risk and weak diversification. Tails with low systematic risk and strong diversification are retained or even expanded, even if they have large standalone risks. These strategies reduce aggregate risk effectively.

This thesis also allocates mean and risk densities of the aggregate loss to component losses. The allocation is critical when risk management strategies such as stop-loss reinsurance and aggregate hedging target aggregate loss layers, and the cost and impact of these aggregate strategies are allocated to component losses. Allocated mean and risk densities of a component loss differ from its mean and systematic risk densities: the former refers to aggregate loss layers, whilst the latter uses component loss layers. The allocation applies conditional mean sharing where an aggregate loss is allocated to component losses based on their conditional expectations (Denuit and Dhaene 2012). Allocated mean and risk densities integrate to the mean and systematic risk of component losses, hence the allocation is unbiased.

1.8 Forming balanced views of risk

1.8.1 Two-sided versus one-sided risk measurement

As highlighted in §1.6, risks of random insurance and financial quantities form critical inputs to insurance premium loadings, lending margins, capital buffers, reinsurance purchases and derivative hedging positions. Risk measures, or premium principles, are well established in the literature and common examples are discussed in Young (2004) and McNeil et al. (2005).

Due to widely perceived negative consequences of uncertainty and volatility, risk measures are typically one-sided and quantify adverse rather than favourable outcomes. Common one-sided risk measures are discussed in the next subsection. On the other hand, two-sided risk measures capture favourable in addition to adverse outcomes. The distinction between one-sided and two-sided risk measures is well articulated in Dhaene et al. (2003):

Two-sided risk measure (TRM): A two-sided risk measure measures the “distance” between the risky situation and the corresponding risk-free situation when both favorable and unfavorable discrepancies are taken into account.

One-sided risk measure (ORM): A one-sided risk measure measures the distance between the risky situation and the corresponding risk-free situation when only unfavorable discrepancies contribute to the “risk.”

Balanced, two-sided risk measurement is crucial in modern financial and insurance markets where competition is intense and unnecessary conservatism

results in opportunity costs. For example when setting capital buffers, excessive conservatism and focussing only on larger than expected losses leads to high buffers and high holding costs, whereas taking a more balanced view and acknowledging the possibility of favourable outcomes yields potentially higher rates of returns. In insurance pricing, taking a one-sided negative view of claim costs results in a high risk loading and an uncompetitive premium, leading to loss of business to competitors with less pessimistic and more realistic views of both favourable and unfavourable claims experience. The importance of capturing upside risk, and strategies of doing so, are discussed generally in Hillson (2003) and Hillson (2002).

1.8.2 Current risk measurement approaches

The following are commonly used risk measures or premium principles. They are generally one-sided and focus on downside risk or adverse outcomes.

- Value-at-Risk (VaR) (Dowd and Blake 2006) is a high percentile in a loss distribution. VaR is commonly used in insurance and finance. For example Solvency II insurance regulations apply 90% and 99.5% VaRs (Eling et al. 2007) in capital requirements. Banking regulations Basel II and Basel III also reference VaRs (Chernobai et al. 2008).
- Conditional-tail-expectation (CTE) (Rockafellar and Uryasev 2002) is the average loss beyond a specified VaR. CTE addresses shortcomings of VaR, by reflecting the magnitude of outcomes above the VaR and satisfying subadditivity (Dowd and Blake 2006). CTE is used in Swiss solvency regulations (Embrechts and Hofert 2014).
- Distortion risk measures (Wang 1996) are expectations under increased loss survival probabilities, implying adverse outcomes are more likely. Distortion risk measures capture CTE, proportional hazard and other risk measures. Choo and De Jong (2009) shows distortion risk measures are equivalent to loss aversion reserves and spectral risk measures (Acerbi 2002), both being weighted averages of VaRs with higher VaRs given greater weight. The weighted averaging concept is important as it is the starting point to generate two-sided risk measures as discussed in the next subsection.
- The zero utility premium (Gerber 1985) is the certainty equivalent of a random loss using a risk averse utility function. The premium always exceeds the expected loss. Using the exponential utility function yields

the exponential premium, which is a weighted average of Esscher premiums (Van Heerwaarden et al. 1989). Heilpern (2003) calculates zero utility premiums using rank-dependent utility theory (Quiggin 1982), whilst Kaluszka and Krzeszowiec (2011) proposes a similar approach using cumulative prospect theory (Tversky and Kahneman 1992).

In the above examples there is always a one-sided focus on larger loss outcomes, resulting in a positive risk loading above the mean. Although it can be argued that this conservatism can be tempered with appropriate selection of risk parameters⁴, there is no explicit allowance for upside risk or smaller loss outcomes. This thesis proposes a risk measure which explicitly captures, and controls, the relative importance of upside and downside risks.

1.8.3 Tradeoff premiums as two-sided risk measures

This thesis proposes “tradeoff premiums” which are two-sided extensions of one-sided distortion risk measures. Upside and downside risks are explicitly captured and a “loss appetite” controls their relative importance. Close links are established with subjective probability in cumulative prospect theory.

Tradeoff premiums are weighted averages of loss outcomes expressed in VaRs. Weights are U-shaped. For one-sided distortion risks, loss aversion reserves or spectral risks discussed in the previous subsection, weights are increasing to reflect the importance of larger loss outcomes. With tradeoff premiums, weights decrease up to an exogenous “loss appetite,” and increase thereafter. U-shaped weights stress the importance of larger *and* smaller loss outcomes, respectively representing downside and upside risk. A low loss appetite implies weights are mostly increasing and generates a conservative premium dominated by downside risk. On the other hand a high loss appetite creates mostly decreasing weights and an aggressive premium with greater focus on upside and less on downside.

Cumulative prospect theory (Tversky and Kahneman 1992) supports a U-shaped weight function: over-weighting extreme, unlikely outcomes and under-weighting average, likely outcomes. Further S-shaped distortion operators are implied from tradeoff premiums, consistent with probability adjustment functions described in cumulative prospect theory.

Manipulating tradeoff premiums yields weighted averages of a distortion risk, capturing downside risk, and its “dual” (Wang 2000), capturing upside

⁴This for example can be the CTE or VaR at a lower threshold, or the zero utility premium using a utility function with a lower risk aversion index.

risk. This property emphasizes the two-sided nature of tradeoff premiums. An example tradeoff premium is the two-sided VaR, a weighted average of lower and upper VaRs. The two-sided CTE is a weighted average of two CTEs separately capturing lower and upper tails. In all cases the loss appetite specifies the weights placed on downside and upside risks.

Tradeoff premiums match the description of two-sided risk measures described in Dhaene et al. (2003): the distance between risky and risk-free positions where the risky position captures favourable and unfavourable outcomes. Subtracting the mean loss (risk-free position) from the tradeoff premium (risky position) yields the combination of a loading and discount. The loading and discount capture the volatility of losses above and below the loss appetite respectively. Therefore the difference between the tradeoff premium and the mean loss depends on the relative volatility of the two tails separated at the loss appetite. Higher upper tail volatility yields an overall positive difference and vice versa.

Tradeoff premiums satisfy translation invariance, positive homogeneity and monotonicity properties of coherent risk measures (Artzner et al. 1999). In addition tradeoff premiums are additive for comonotonic random losses. Only subadditivity is not generally satisfied: the tradeoff premium of a sum of random losses may exceed the sum of individual tradeoff premiums. Tradeoff premiums are not subadditive due to their two-sided nature. Downside and upside risks have opposite impact on tradeoff premiums and both reduce or “diversify” upon aggregation. For right skewed loss distributions or low loss appetite, downside risk dominates upside risk, yielding subadditive tradeoff premiums. Conversely left skewed loss distributions or high loss appetite give rise to superadditivity.

1.9 Conclusion

This thesis proposes novel solutions to critical problems around measuring risk and dependence behaviour in a joint probability distribution. These problems arise from observations that risk and dependence typically vary across the distribution and hence measures of overall risk and dependence, often the focus in the literature, provide insufficient information. In addition competition is creating greater focus on upside risk in addition to downside risk, driving the need for two-sided risk measures.

Proposed solutions, although novel, apply established concepts such as correlation, layers, VaR, distortion risk, and Euler allocation. These concepts are applied consistently across proposed solutions, naturally leading to

an integrated and coherent quantitative risk management framework. Commonalities between proposed solutions are:

- Loss outcomes are expressed in VaRs instead of absolute amounts, and modelled based on their percentile ranks.
- Random losses are decomposed into VaR layers when constructing layer dependence, mean and risk densities, as well as analyzing systematic risk and diversification.
- Distortion risk is used to form risk densities, analyse systematic risk and diversification, and measure upside and downside risks in the tradeoff premium.

Proposed solutions are explained in the next four chapters. The final chapter discusses the integration of proposed solutions, and outlines potential future research areas.

Chapter 2

Layer dependence

The following paper introduces, analyzes and illustrates layer dependence as a measure of local dependence.

Layer dependence as a measure of local dependence

Abstract

A new measure of local dependence called “layer dependence” is proposed and analysed. Layer dependence measures the dependence between two random variables at different percentiles in their joint distribution. Layer dependence satisfies coherence properties similar to Spearman’s correlation, such as lying between -1 and 1 , with -1 , 0 and 1 corresponding to countermonotonicity, independence and comonotonicity, respectively. Spearman’s correlation is a weighted average of layer dependence across all percentiles. Alternate overall dependence measures are arrived by varying the weights. Layer dependence is an important input to copula modeling by extracting the dependence structure from past data and incorporating expert opinion if necessary.

Keywords: Local dependence; rank dependence; Spearman’s correlation; layers; conditional tail expectation; concordance.

1. Local dependence and layer dependence

Dependence between two random variables generally varies with percentile. For example extreme movements in stock markets are likely to be highly related whereas minor fluctuations may be relatively independent. Catastrophes create significant insurance losses for several classes of business at the same time, while attritional losses are typically weakly dependent.

Local dependence measures aim to capture the dependence structure of a bivariate distribution. This contrasts with measures of overall dependence such as Pearson correlation, Spearman’s ρ and Kendall’s τ (Embrechts et al., 2002). Local dependence measures include the univariate tail concentration (Venter, 2002), correlation curve (Bjerve and Doksum, 1993), and bivariate measures by Bairamov et al. (2003), Jones (1996) and Holland and Wang (1987).

This paper introduces, illustrates and analyzes an alternate local dependence measure called “layer dependence.” Layer dependence is the covariance between a random variable and a single “layer” of another. Layer dependence is also the “gap” between upper and lower conditional tail expectations. Layer dependence is calculated from the copula underlying the joint distribution. Hence of interest is rank dependence rather than dependence between random variables in their original scale: the latter is often distorted by marginal distributions.

Layer dependence satisfies “coherence” properties similar to Spearman’s ρ : it is between -1 and 1 , constant and equal to -1 , 0 and 1 for countermonotonic, independent and comonotonic random variables, sign switching when the ranking order reverses, and taking on higher values when dependence is stronger. Taking a weighted average of layer dependence values across the joint distribution yields Spearman’s ρ and alternate coherent measures of overall dependence.

Layer dependence provides a more appropriate and accurate measure of local dependence compared to existing measures. Higher dispersion between scatter points from the 45° line reduces layer dependence and vice versa. For a Gumbel copula exhibiting upper tail dependence, layer dependence starts from a lower value and increases to 1 while the opposite applies to a Clayton copula with lower tail dependence.

Calculating layer dependence at the first instance extracts essential and interpretable information – the dependence structure. This is often more informative than positing parametric copula, as the implication of its parametric form and parameters on the dependence structure is indirect. Computed layer dependence facilitates the selection and fitting of an appropriate copula.

Remaining sections are as follows. Section 2 discusses the concepts leading to definition of layer dependence. Section 3 demonstrates how layer dependence extracts the dependence structure from common copulas. Section 4 explains the behaviour of layer dependence by decomposing it into a negative function of discordance and dispersion. Section 5 describes coherence properties of layer dependence. Links to existing tail dependence measures are highlighted in §6. Further properties of layer dependence are described in §7. Section 8 forms alternate coherent measures of overall dependence apart from Spearman’s ρ , using weighted averages of layer dependence. Section 9 discusses how layer dependence can be applied to copula modeling. Section 10 concludes.

2. Layer dependence – motivation and definition

A familiar construct in the study of bivariate dependence is Spearman’s correlation (Embrechts et al., 2002) defined as the linear correlation between ranks of two random variables. Rank dependence avoids distortion arising from marginal distributions as with for example Pearson’s correlation (McNeil et al., 2005) measuring the degree of linear relationship between random variables in their original scale. Spearman’s correlation can also be applied to estimate copula parameters using the method of moments (Kojadinovic and Yan (2010), Bouyé et al. (2000)). However Spearman’s correlation suffers from shortcomings and, as an aggregate measure, is inappropriate for assessing local dependence when dependence varies across the joint distribution including the tails.

Another familiar construct, in reinsurance, is a loss layer (Wang, 1995). For example the 95%–96% layer of a random continuous loss x is the portion of x between its 95th and 96th percentile

$$\min \{ (x - x_{0.95})^+, x_{0.96} - x_{0.95} \} ,$$

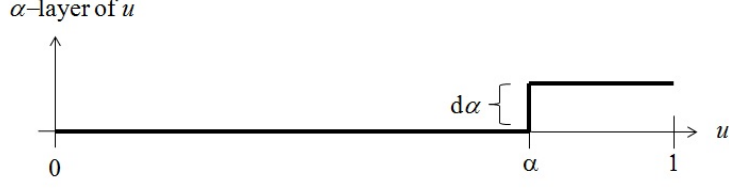


Figure 1: Illustration of α -layer of u , written as $I_\alpha(u)d\alpha$. The α -layer of u is only sensitive to movements in u at α and ignores other movements.

where $()^+$ indicates the positive part of the expression inside the brackets and subscripts indicate the percentile. Layers also arise in the context of derivative payouts and debt tranches in collateralised debt obligations (Mandel et al., 2012). With rank dependence, the 95%–96% layer reduces to

$$\min \{ (u - u_{0.95})^+, u_{0.96} - u_{0.95} \} = \min \{ (u - \alpha)^+, d\alpha \} , \quad (1)$$

where u is the percentile rank of x and in this example $\alpha = 0.95$ and $d\alpha = 0.01$. Note $u_\alpha = \alpha$.

2.1. Percentile layer decomposition

The final expression in (1) can be written and approximated by, using the familiar infinitesimal notation,

$$-d(u - \alpha)^+ \approx I_\alpha(u)d\alpha , \quad I_\alpha(u) \equiv \begin{cases} 0 , & u \leq \alpha , \\ 1 , & u > \alpha . \end{cases} \quad (2)$$

The left hand side expression in (2) is called the α -layer of u . The approximation becomes exact as $d\alpha \rightarrow 0$. The α -layer of u is an infinitesimally small increment $d\alpha$ if u exceeds α and is zero otherwise. Hence the α -layer captures the movement in u at α and ignores movements elsewhere. Figure 1 illustrates the α -layer of u .

As often exploited in reinsurance, any non-negative random loss or variable can be thought of as a sum of layers:

$$u = \int_0^u 1d\alpha = \int_0^1 I_\alpha(u)d\alpha . \quad (3)$$

Hence u is formed from infinitely many α -layers, each capturing the movement of u at a different α .

2.2. Constructing layer dependence

Spearman's correlation is the correlation and also the standardized covariance between two percentile rank random variables v and u :

$$\rho \equiv \text{cor}(u, v) = \frac{\text{cov}(v, u)}{\sqrt{\text{cov}(v, v)\text{cov}(u, u)}} = \frac{\text{cov}(v, u)}{\text{cov}(u, u)} = 12\text{cov}(v, u) , \quad (4)$$

where cor and cov indicate covariance and correlation, respectively. In this paper assume the copula of u and v is exchangeable (Nelson, 1999).

Using the decomposition of u in (3),

$$\text{cov}(v, u) = \text{cov} \left\{ v, \int_0^1 I_\alpha(u) d\alpha \right\} = \int_0^1 \text{cov}\{v, I_\alpha(u)\} d\alpha. \quad (5)$$

Hence the covariance $\text{cov}(v, u)$ can be thought of as the sum of infinitely many covariances $\text{cov}\{v, I_\alpha(u)\}$ for $0 \leq \alpha \leq 1$. Each covariance in the sum measures the dependence between v and the α -layer of u . In a reinsurance setting, this covariance measures dependence between a particular layer of a loss (in percentile rank terms) and another factor. Alternatively the covariance captures dependence between movements in u at α and v .

Similar to ρ , scaling the layer covariances $\text{cov}\{v, I_\alpha(u)\}$ with the same when $v = u$ leads to the definition of layer dependence

$$\ell_\alpha \equiv \frac{\text{cov}\{v, I_\alpha(u)\}}{\text{cov}\{u, I_\alpha(u)\}}, \quad (6)$$

where the denominator

$$\text{cov}\{u, I_\alpha(u)\} = E[\{u - E(u)\}I_\alpha(u)] = \int_\alpha^1 \left(u - \frac{1}{2}\right) d\alpha = \frac{\alpha(1 - \alpha)}{2}. \quad (7)$$

Combining (4), (5), (6) and (7) yields

$$\rho = 12 \int_0^1 \ell_\alpha \text{cov}\{u, I_\alpha(u)\} d\alpha = \int_0^1 \ell_\alpha 6\alpha(1 - \alpha) d\alpha. \quad (8)$$

Hence ρ is a weighted average of ℓ_α for $0 \leq \alpha \leq 1$ with weights $w_\alpha = 6\alpha(1 - \alpha)$ integrating to 1, and ℓ_α decomposes ρ into local dependence values. Note w_α has minimum 0 at $\alpha = 0$ and 1 and increases symmetrically to maximum at $\alpha = 0.5$. Hence Spearman's correlation places little emphasis on the tails which may be undesirable in finance or insurance where tail dependence is critical. Modifying the weights w_α leads to alternate measures of overall dependence further discussed in §8.

Layer dependence summarises the dependence structure of a copula. Layer dependence provides additional information compared to Spearman's correlation: how dependence varies across the joint distribution. As any summary measure, layer dependence can mislead but is less misleading than ρ and other measures of overall dependence. Layer dependence is a more meaningful characterisation of a copula compared to the parameters of copula families such as the Clayton or Gumbel (McNeil et al., 2005). Properties of and arguments for using layer dependence are explored below.

2.3. Key properties of layer dependence

In §3 and §4 it is illustrated and shown that $\ell_\alpha = 1$ occurs if and only if $I_\alpha(u)$ and $I_\alpha(v)$ are both simultaneously 1 or 0: that is u and v are both either greater than α or less than α . Additionally $\ell_\alpha = 1$ for $c \leq \alpha \leq d$ implies $u = v$ over $c \leq u, v \leq d$.

Layer dependence ℓ_α satisfies the following coherence properties for all $0 \leq \alpha \leq 1$. These properties are shared with Spearman's correlation and are formalised in §5. Similar to correlation, $0 \leq \ell_\alpha \leq 1$, and $\ell_\alpha = -1, 0$ and 1 if u and v are countermonotonic, independent and comonotonic, respectively. In addition ℓ_α increases with the correlation order of (u, v) . Replacing u or v with their complement leads to straightforward changes in layer dependence.

The following is an alternative expression for ℓ_α :

$$\ell_\alpha = \frac{E(v|u > \alpha) - E(v|u \leq \alpha)}{E(u|u > \alpha) - E(u|u \leq \alpha)} = 2 \{E(v|u > \alpha) - E(v|u \leq \alpha)\} . \quad (9)$$

Proofs are in an appendix. The middle expression in (9) is the expected change in v relative to the expected change in u when u crosses α . The latter is 0.5 for all α , yielding the final expression in (9). Hence large ℓ_α implies v is sensitive to movements in u across α , indicating strong dependence between v and u at α . When $\ell_\alpha = 0$, v is unchanged on average when u crosses α .

3. Illustration of layer dependence curves for various copulas

The nine panels in Figure 2 display (u, v) scatterplots of well known exchangeable copulas. Less standard copulas are also shown for illustration. All copulas are calibrated to have equal Spearman's correlation $\rho = 0.6$. Layer dependence curves ℓ_α for all $0 \leq \alpha \leq 1$ are plotted against α on (u, v) scatterplots to demonstrate the link between the scatter and ℓ_α .

The scatterplots in Figure 2 emphasize that copulas with the same overall dependence can exhibit a variety of local dependence structures and these structures are captured by layer dependence curves. Given α , ℓ_α is larger if scatter points are more clustered around (α, α) and vice versa. In particular ℓ_α increases to one in the tails of copulas exhibiting strong tail dependence. Hence ℓ_α tracks the clustering of scatter points across the 45° line. This is formalised in §4.

4. Discordance and dispersion

Layer dependence is intimately connected to discordance and dispersion. Again assuming the copula C of (u, v) is exchangeable then

$$\ell_\alpha = 1 - 2(1 + \gamma_\alpha)\delta_\alpha , \quad (10)$$

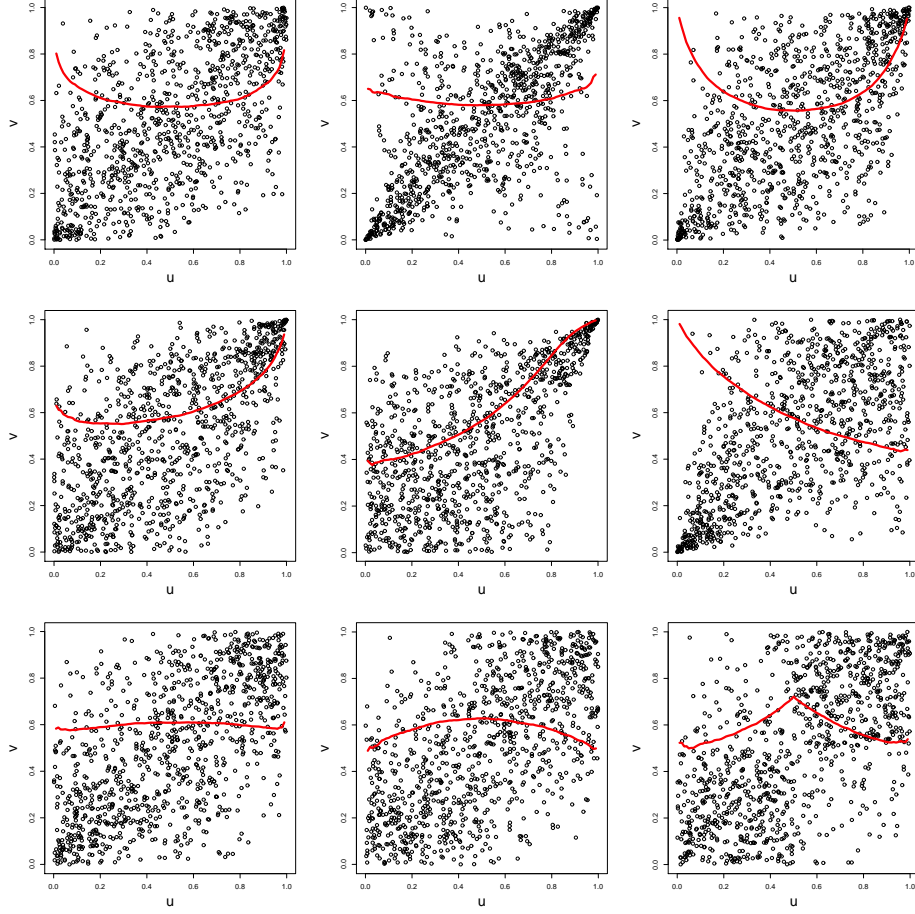


Figure 2: Copulas all with the same $\rho = 0.6$ but different layer dependence curves ℓ_α over $0 \leq \alpha \leq 1$ (in red). The top 3 copulas display relatively strong tail dependence in both lower and upper tail. The first two copulas in the middle three have relatively strong upper tail dependence and while the third has a high degree of lower tail dependence. The bottom three copulas have relatively high local dependence in the middle of the distribution and less correlation in the tails. Overall the panels illustrate how a given overall level of ρ can mask a range of dependence structures and that layer dependence is an appropriate tool for assessing local dependence. The top left panel displays a Gaussian copula, followed by a Student's t copula on the right. The leftmost panel in the middle three is a Gumbel copula, and the rightmost is a Clayton copula. The bottom left panel is a Frank copula. Remaining copulas are constructed from a factor model and are included to illustrate how layer dependence captures local dependence.

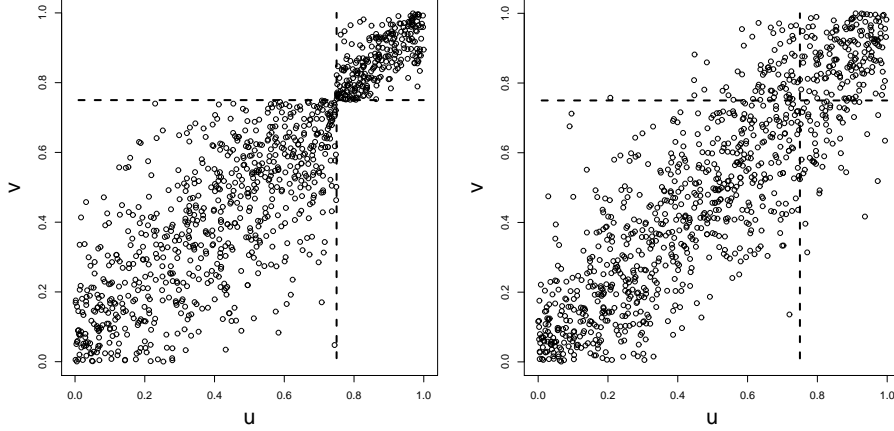


Figure 3: The left and right panel show copulas where $\ell_{0.75} = 1$ and $\ell_{0.75} = 0.86$, respectively. In the left panel, $\gamma_{0.75} = -1$ (no discordant points) and $\delta_{0.75} = 0$ (zero dispersion). In the right panel, $\gamma_{0.75} = -0.65$ (some discordant points) and $\delta_{0.75} = 0.21$ (some dispersion between discordant points).

where

$$\gamma_\alpha \equiv \text{cor}\{1 - I_\alpha(u), I_\alpha(v)\} = \text{cor}\{I_\alpha(u), 1 - I_\alpha(v)\} = \frac{\alpha^2 - C(\alpha, \alpha)}{\alpha(1 - \alpha)},$$

and

$$\delta_\alpha \equiv \text{E}\{|u - v| | (u - \alpha)(v - \alpha) < 0\}.$$

A proof of (10) is in the appendix. Equation (10), as shown below, explains the behaviour of layer dependence curves in Figure 2 where ℓ_α for any α increases as the scatter become more tightly concentrated around the point (α, α) , that is lower discordance and dispersion relative to the same point.

The correlation $-1 \leq \gamma_\alpha \leq 1$ measures the tendency for (u, v) to be discordant at α : either u or v is above α and the other is below α . The expectation $0 \leq \delta_\alpha \leq 1$ measures the average dispersion between discordant u and v at α , noting $(u - \alpha)(v - \alpha) < 0$ is equivalent to $u > \alpha$ and $v \leq \alpha$ or $u \leq \alpha$ and $v > \alpha$.

Figure 3 illustrates (10) using two copulas specifically constructed such that, for $\alpha = 0.75$, $\ell_\alpha = 1$ for one and $\ell_\alpha = 0.86$ for the other. For the copula where $\ell_\alpha = 1$, there is no discordance or dispersion at α . This observation is confirmed by setting $\ell_\alpha = 1$ in (10) yielding $\gamma_\alpha = -1$ or $\delta_\alpha = 0$. In this case u and v are perfectly dependent at α , and are simultaneously below or above α . It is also straightforward to extend this result such that if $\ell_\alpha = 1$ over $c \leq \alpha \leq d$ then $u = v$ over $c \leq u, v \leq d$. As ℓ_α decreases from 1, as for the second copula, the extent of discordance γ_α and dispersion δ_α increases. Again this result can be confirmed from (10) noting ℓ_α is negatively related to γ_α and δ_α .

Applying (10) and generalising the illustration in Figure 3 explains the behaviour of layer dependence curves in Figure 2. Layer dependence ℓ_α is larger if scatter points are more clustered around (α, α) , that is, fewer discordant points at α , and discordant points at α are closer to the 45° degree line. The former indicates smaller γ_α and the latter indicates smaller δ_α . Opposite observations apply for small ℓ_α .

5. Coherence properties of layer dependence

Layer dependence ℓ_α satisfies the following five “coherence” properties. These properties are extensions of properties applying to Spearman’s correlation.

- **Bounds:** Layer dependence lies between -1 and 1 : $-1 \leq \ell_\alpha \leq 1$ for all α . Hence layer dependence is bounded in the same way as ρ .
- **Perfect dependence:** Constant layer dependence of -1 or 1 are equivalent to countermonotonicity and comonotonicity, respectively. Thus $\ell_\alpha = -1$ for all α if and only if $v = 1 - u$ while $\ell_\alpha = 1$ for all α if and only if $v = u$.
- **Independence:** If u and v are independent then $\ell_\alpha = 0$ for all α . The converse is not true – zero layer dependence does not imply independence as shown by the following counterexample. Assume $v = u$ and $v = 1 - u$ with equal probability. Then $E(v|u = \alpha) = 0.5$ for all $0 \leq \alpha \leq 1$ implying $E(v|u > \alpha) = E(v|u \leq \alpha) = 0.5$. Hence $\ell_\alpha = 0$ from (9). However u and v are not independent.
- **Symmetry:** Replacing v with $1 - v$ yields layer dependence curve $-\ell_\alpha$. Doing the same to u (the random variable decomposed into layers) yields layer dependence curve $-\ell_{1-\alpha}$ hence a flip is performed about $\alpha = 0.5$ in addition to a sign change. Replacing both u and v with their complements yields layer dependence curve $\ell_{1-\alpha}$.
- **Ordering:** Higher correlation order (Dhaene et al., 2009) leads to higher layer dependence. Specifically, consider bivariate uniform (u^*, v^*) exceeding (u, v) in correlation order: $C^*(a, b) \geq C(a, b)$ for all $0 \leq a, b \leq 1$, where C^* is the copula of (u^*, v^*) . Then $\ell_\alpha^* \geq \ell_\alpha$, $0 \leq \alpha \leq 1$ where ℓ_α^* denotes the α -layer dependence of (u^*, v^*) . Hence greater dependence leads to a higher layer dependence curve.

Independence and symmetry properties follow from the definition of layer dependence in (6). From (9), constant layer dependence of one implies $E(v|u > \alpha) = (\alpha+1)/2$ and $E(v|u = \alpha) = \alpha$, for all $0 \leq \alpha \leq 1$, hence $v = u$. Similarly constant layer dependence of minus one implies $v = 1 - u$. The ordering property holds since higher correlation order implies larger covariances (Dhaene et al., 2009). Prove the bounds property by combining ordering and perfect dependence properties, and noting countermonotonicity and comonotonicity represent minimum

and maximum correlation order, respectively. Detailed proofs are shown in the appendix.

Most of the layer dependence coherence properties can be expressed using copulas shown in Table 1.

Copula	Description	LD
uv	Independent u and v	0
$\min(u, v)$	Comonotonic u and v	1
$\max(u + v - 1, 0)$	Countermonotonic u and v	-1
$u - C(u, 1 - v)$	Replace v with $1 - v$	$-\ell_\alpha$
$v - C(1 - u, v)$	Replace u with $1 - u$	$-\ell_{1-\alpha}$
$u + v - 1 + C(1 - u, 1 - v)$	Replace v and u with complements	$\ell_{1-\alpha}$

Table 1: Layer dependence (LD) for different copulas and transformations. Note the final copula is the survival copula of C .

6. Connections to measures of tail dependence

Measures have been proposed to capture the degree of tail dependence. Tail dependence is dependence between extreme values of random variables, in this case values of u and v near 0 or 1. Strong tail dependence creates catastrophic events such as multiple bank failures and market crashes. Layer dependence is intimately connected to two existing tail dependence measures – coefficient of tail dependence and tail concentration function. Sweeting and Fotiou (2013) and Durante et al. (2014) further discuss tail dependence measures.

6.1. Coefficients of tail dependence

Joe (1997) defines coefficients of lower and upper tail dependence as

$$\lambda_L \equiv \lim_{\alpha \downarrow 0} P(v \leq \alpha | u \leq \alpha), \quad \lambda_U \equiv \lim_{\alpha \uparrow 1} P(v > \alpha | u > \alpha).$$

Unit coefficients indicate perfect positive tail dependence, and occur if and only if u converging to 0 or 1 implies the same for v . Coefficients of negative tail dependence replace $v \leq \alpha$ and $v > \alpha$ in the above expressions with $v > 1 - \alpha$ and $v \leq 1 - \alpha$, respectively. Sweeting and Fotiou (2013) discusses the drawback of these coefficients and suggests a modification by weakening the limits, yielding links to tail concentration functions discussed below.

From (9), $\ell_0 = \lim_{\alpha \downarrow 0} \ell_\alpha = 1 - 2E(v|u = 0)$ which equals 1 if and only if $E(v|u = 0) = 0$, or v approaching 0 if the same happens to u . Hence $\ell_0 = 1$ is equivalent to $\lambda_L = 1$. Similarly $\ell_1 = 2E(v|u = 1) - 1 = 1$ if and only if $u = 1$ implies $v = 1$, which is equivalent to $\lambda_U = 1$. Similar results apply to perfect negative tail dependence.

Hence layer dependence characterises perfect tail dependence in the same way as the measures of Joe (1997): variables attaining their maximum or minimum values simultaneously.

6.2. Tail concentration

Tail concentration (Venter, 2002) is a local dependence measure formed from lower and upper conditional tail probabilities. Similar to layer dependence, tail concentration functions describe the dependence structure of copulas and are used to identify families of copulas. The following shows the connection between layer dependence and tail concentration and shows layer dependence refines tail concentration by incorporating additional information on dispersion defined in §4.

The tail concentration at α is

$$\tau_\alpha \equiv \begin{cases} P(v \leq \alpha | u \leq \alpha) = \frac{C(\alpha, \alpha)}{\alpha}, & \alpha \leq 0.5, \\ P(v > \alpha | u > \alpha) = \frac{1 - 2\alpha + C(\alpha, \alpha)}{1 - \alpha}, & \alpha > 0.5. \end{cases}$$

A larger τ_α implies u and v are more likely to fall in the same tail – lower tail if $\alpha \leq 0.5$ and upper tail if $\alpha > 0.5$. Letting $\alpha \rightarrow 0$ and $\alpha \rightarrow 1$ yields the coefficients of tail dependence λ_L and λ_U discussed above. Properties of tail concentration and its applications to distinguish families of copulas are discussed in Durante et al. (2014).

To show the connection between layer dependence and tail concentration, first standardise τ_α by subtracting τ_α^0 , its value under independence, and dividing by $\tau_\alpha^+ - \tau_\alpha^0$ where τ_α^+ is the value of τ_α if $u = v$, yielding:

$$\tau_\alpha^* = \frac{\tau_\alpha - \tau_\alpha^0}{\tau_\alpha^+ - \tau_\alpha^0} = \frac{C(\alpha, \alpha) - \alpha^2}{\alpha(1 - \alpha)}, \quad \tau_\alpha^0 = \begin{cases} \alpha, & \alpha \leq 0.5 \\ 1 - \alpha, & \alpha > 0.5 \end{cases}, \quad \tau_\alpha^+ = 1.$$

Then τ_α^* equals 1 or 0, if $u = v$ or if u and v are independent, respectively.

Hence $\tau_\alpha^* = -\gamma_\alpha$ where γ_α measures discordance between u and v at α as discussed in §4. Combining this result with (10) yields

$$\ell_\alpha = 1 - 2\delta_\alpha(1 - \tau_\alpha^*) = 1 - 2\delta_\alpha + 2\delta_\alpha\tau_\alpha^*$$

where δ_α as defined in §4 measures average dispersion between discordant u and v at α . Note ℓ_α is increasing in τ_α^* and τ_α . It is also straightforward to show that $\ell_\alpha = 1$ is equivalent to $\tau_\alpha^* = 1$ which is in turn equivalent to $\tau_\alpha = 1$. Hence layer dependence and tail concentration have equivalent characterisations of perfect local dependence at any point. This extends the above result that perfect tail dependence has identical implications on layer dependence and coefficients of tail dependence by Joe (1997).

Hence tail concentration and layer dependence are closely connected. Tail concentration (in its standardised form) is only one of two factors forming layer

dependence. The other factor is the dispersion between discordant points, measured by δ_α . Using the right panel of Figure 3, τ_α or τ_α^* only reflects the number of discordant points at α , whereas layer dependence combines this information and average dispersion between discordant points. Hence layer dependence refines tail concentration by standardising and including information on dispersion.

7. Further properties of layer dependence

This section lists and explores further properties and results of layer dependence.

7.1. Copula integration

Layer dependence ℓ_α can be written as a standardised integral of the copula:

$$\ell_\alpha = \frac{\int_0^1 \text{cov}\{I_\alpha(u), I_\beta(v)\} d\beta}{\alpha(1-\alpha)/2} = \frac{2 \int_0^1 C(\alpha, \beta) d\beta - \alpha}{\alpha(1-\alpha)}. \quad (11)$$

The result follows from (6) by decomposing $v = \int_0^1 I_\beta(v) dv$ similar to (3) and noting $\text{cov}\{I_\alpha(u), I_\beta(v)\} = C(\alpha, \beta) - \alpha\beta$.

Thus ℓ_α summarises a copula: reducing the dimension from two and one, and scales the result to ensure it lies between ± 1 . This computation extracts the dependence structure from the copula. In comparison, tail concentration performs the dependence extraction by computing the diagonal section $C(\alpha, \alpha)$.

7.2. Layer dependence preserves convex combination

Layer dependence is preserved under convex combinations of copulas. Suppose (u^*, v^*) has α -layer dependence ℓ_α^* and copula C^* . Then a mixture $b(u, v) + (1-b)(u^*, v^*)$ where b is Bernoulli with $E(b) = \pi$ has copula $\pi C + (1-\pi)C^*$ and layer dependence $\pi\ell_\alpha + (1-\pi)\ell_\alpha^*$. Hence layer dependence preserves convex combinations of copulas. The proof follows directly from (11).

It is also straightforward to show that layer dependence preserves multiple and continuous convex combinations of copulas.

7.3. One-sided conditional tail expectations

Since $E(v) = \alpha E(v|u \leq \alpha) + (1-\alpha)E(v|u > \alpha)$ it follows from (9) that

$$\ell_\alpha = \frac{E(v|u > \alpha) - E(v)}{E(u|u > \alpha) - E(u)} = \frac{E(v|u \leq \alpha) - E(v)}{E(u|u \leq \alpha) - E(u)},$$

the gap between upper or lower conditional tail expectations of v and the unconditional expectation. Denominators are again scaling factors ensuring $\ell_\alpha = 1$ if u and v are comonotonic and $\ell_\alpha = -1$ if countermonotonic.

7.4. Layer dependence for a non-exchangeable copula

If the copula of u and v is not exchangeable, $C(u, v) \neq C(v, u)$, then layer dependence changes when v is decomposed into layers rather than u :

$$\frac{\text{cov}\{v, I_\alpha(u)\}}{\text{cov}\{u, I_\alpha(u)\}} \neq \frac{\text{cov}\{u, I_\alpha(v)\}}{\text{cov}\{v, I_\alpha(v)\}},$$

that is, the dependence between v and α -layer of u differs from the dependence between u and α -layer of v . This is similar to regression where the regression of y on x is not equivalent to the regression of x on y unless the joint distribution is exchangeable.

8. Alternate measures of overall dependence

It may still be convenient to use an overall dependence measure such as Spearman's correlation. However as discussed in §2, Spearman's correlation downplays tail dependence and may be inappropriate when tail dependence is critical. The following formulates alternative measures of overall dependence by taking different weighted averages of layer dependence:

$$\rho_W \equiv \int_0^1 w_\alpha \ell_\alpha d\alpha = \int_0^1 w_\alpha \frac{\text{cov}\{v, I_\alpha(u)\}}{\text{cov}\{u, I_\alpha(u)\}} d\alpha = \text{cov}\{v, W(u)\},$$

where

$$W(u) \equiv \int_0^1 \frac{w_\alpha I_\alpha(u)}{\text{cov}\{u, I_\alpha(u)\}} d\alpha = 2 \int_0^u \frac{w_\alpha}{\alpha(1-\alpha)} d\alpha, \quad \int_0^1 w_\alpha d\alpha = 1.$$

The weighting function w_α indicates the weight or importance of the α -layer of u and its dependence with v .

Hence ρ_W is the covariance between v and a transformation of u . For Spearman's correlation ρ , $w_\alpha = 6\alpha(1-\alpha)$ and $W(u) = 12u$. Even though the expression for ρ_W is asymmetric in u and v , the result is identical when u and v are switched if the copula C is exchangeable.

Since ρ_W averages ℓ_α using non-negative weights which integrate to 1, all coherence properties of ℓ_α described in §5 apply to ρ_W . Specifically $-1 \leq \rho_W \leq 1$, with $\rho_W = -1, 0$ and 1 , under countermonotonicity, independence and comonotonicity, respectively. Further ρ_W switches its sign when u or v are replaced by their complement. These properties mimic those of Spearman's correlation.

The following are examples of w_α yielding alternate rank dependence measures ρ_W :

- Suppose dependence at different percentiles are equally important. Then $w_\alpha = 1$ yielding

$$\rho_1 = 2\text{cov} \left\{ v, \log \left(\frac{u}{1-u} \right) \right\} = \frac{\pi}{3} \text{cor} \left\{ v, \log \left(\frac{u}{1-u} \right) \right\} ,$$

a multiple of the correlation between v and the logit of u . The final expression follows by noting v and $\log\{u/(1-u)\}$ have standard deviations $1/\sqrt{12}$ and $\pi/\sqrt{3}$, respectively. If tail dependence is pronounced, $\rho_1 > \rho$ since ρ_1 weights tail dependence more heavily compared to ρ . An illustration ρ_1 and other alternates to ρ is shown below.

- If $w_\alpha = 3\alpha^2$ then dependence at higher percentiles are considered more important. This formulation is applicable when upper tail dependence is critical, for example the simultaneous occurrence of large insurance losses in different lines of business. Then

$$\rho_2 = 6\text{cov} \left\{ v, \log \left(\frac{e^{-u}}{1-u} \right) \right\} = \sqrt{3}\text{cor} \{v, -\log(1-u)\} - \frac{\rho}{2} ,$$

noting $-\log(1-u)$ has standard deviation 1. This dependence measure combines ρ and the correlation between v and logarithmic transform of u . If dependence is higher above the median then $\rho_2 > \rho$.

- If dependence over percentiles below the median is more important then for example $w_\alpha = 3(1-\alpha)^2$ yielding

$$\rho_3 = 6\text{cov} \{v, \log(ue^{-u})\} = \sqrt{3}\text{cor}(v, \log u) - \frac{\rho}{2} ,$$

similar to ρ_2 with an opposite logarithmic transform on u .

- Suppose w_α is derived from V_α , the inverse distribution of x with derivative V'_α :

$$w_\alpha = \frac{\text{cov}\{u, I_\alpha(u)\}V'_\alpha}{\int_0^1 \text{cov}\{u, I_\alpha(u)\}V'_\alpha d\alpha} = \frac{0.5\alpha(1-\alpha)V'_\alpha}{\text{cov}(u, V_u)} = \frac{0.5\alpha(1-\alpha)V'_\alpha}{\text{cov}(u, x)} ,$$

where $\int_0^1 I_\alpha(u)V'_\alpha d\alpha = \int_0^u V'_\alpha d\alpha = V_u$ and $x = V_u$. Using these weights yields the Gini correlation (Schechtman and Yitzhaki, 1999)

$$\rho_4 = \frac{\text{cov}(v, V_u)}{\text{cov}(u, V_u)} = \frac{\text{cov}\{G(y), x\}}{\text{cov}\{F(x), x\}} ,$$

where $F \equiv V^-$ and G are distribution functions of observed random variables x and y with percentile ranks u and v , respectively. In this example the weights w_α depend on the marginal distribution of x . More skewness in x leads to more steeply increasing V'_α and w_α hence greater emphasis on upper tail dependence.

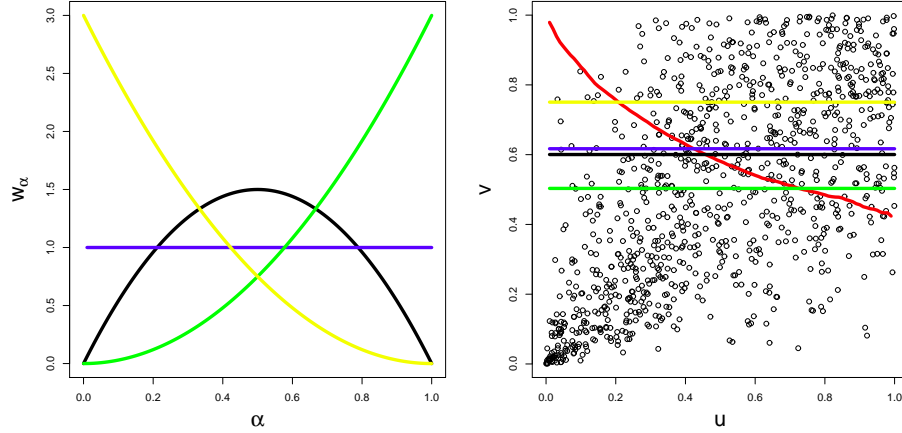


Figure 4: The left panel shows the weight functions $w_\alpha = 6\alpha(1 - \alpha)$ (black), $w_\alpha = 1$ (blue), $w_\alpha = 3\alpha^2$ (green), $w_\alpha = 3(1 - \alpha)^2$ (yellow). The right panel shows a Clayton copula with a red layer dependence curve and overall dependence ρ , ρ_1 , ρ_2 and ρ_3 indicated using the same colors as corresponding weight functions in the left panel. Note the weight function $w_\alpha = 3(1 - \alpha)^2$ delivers the highest overall dependence for the Clayton copula due to the emphasis on lower tail dependence. The weight function $w_\alpha = 3\alpha^2$ delivers the lowest overall dependence due to the emphasis on upper tail dependence.

Figure 4 illustrates the first three alternates to Spearman's correlation listed above using a Clayton copula shown in Figure 2. Note ρ , ρ_1 , ρ_2 and ρ_3 are all weighted averages of ℓ_α . However their values differ depending on the weight function used. As the Clayton copula exhibits strong lower tail dependence and weak upper tail dependence, ℓ_α is decreasing in α . Hence $\rho_1 > \rho$ as ρ_1 places relatively higher weight on large values of ℓ_α in the lower tail. ρ_3 is the largest as it places the most weight on the lower tail, whereas ρ_2 is the smallest as it focusses on upper tail dependence. Hence ρ_3 is a sensitive measure of lower tail dependence and is appropriate when lower tail dependence is of main concern, and vice versa for ρ_2 .

9. Discussion of copula fitting involving layer dependence

This section discusses how layer dependence can be applied to select and fit copulas using past data and expert judgement.

The literature discusses copula fitting extensively. A common approach is to first select parametric copula and marginal distributions, then estimate parameters by maximising joint likelihood (Denuit et al., 2005). A semi-parametric approach replaces marginal distributions in joint likelihood with empirical ranks (Oakes, 1989). The method of moments involves setting up equations to match statistics computed from data and the parametric model (Kojadinovic and Yan,

2010). Kojadinovic and Yan (2010) also compares common copula fitting methods. The choice of parametric copula may be restricted to a specific class of copulas, such as Archimedean copulas (McNeil et al., 2005). Genest and Rivest (1993) suggests an approach to select the generator function of Archimedean copulas. Alternatively a visual assessment of data may suggest an appropriate parametric copula with similar dependence structure, for example the Gumbel or Clayton copulas if tail dependence is strong. At the other extreme is to use the empirical copula, if the volume of past data is sufficiently large. Czado (2010) discusses a semi-parametric approach to model multivariate copulas, based on vine copulas.

Layer dependence can be applied to the copula modeling process in several ways:

- Layer dependence is first computed from data at intervals of say 0.01, that is $\ell_{0.01}, \ell_{0.02}, \dots$. The selected granularity depends on the volume of data. Layer dependence values may be smoothed parametrically or non-parametrically to reveal the dependence structure of the data.
- There may be an intermediate step where the computed layer dependence curve is adjusted to incorporate expert opinion. For example one may wish to increase layer dependence in the upper tail in anticipation of stronger upper tail dependence than observed historically.
- A parametric copula may then be selected by matching the shape of its layer dependence curve to the data. Shapes of layer dependence curves of typical copulas are shown in Figure 2. Parameters can be fitted using either maximum likelihood or method of moments.
- A mixture of parametric copulas, for example where the parameters follow a probability distribution, may be applied such that the layer dependence curve perfectly matches the data. The difficulty is closed form expressions may be unavailable.

Applying layer dependence in copula fitting refines existing approaches in several ways. Firstly layer dependence guides the selection of a copula family so that the dependence structure of past data is mirrored closely. In addition layer dependence is a convenient medium for incorporating expert opinion on the dependence structure. Further layer dependence is robust to data inadequacies as it summarises data into conditional tail mean values and smoothing is applied.

10. Conclusion

Layer dependence captures dependence structures in bivariate copulas, and shares the same practical properties as Spearman’s correlation. Layer dependence is connected to, and refines, current approaches to measure tail dependence. Taking weighted averages of layer dependence curves yields Spearman’s

correlation and alternate measures of overall dependence which emphasize different areas of the dependence structure.

Using layer dependence in copula fitting enhances the process by capturing and reflecting the dependence structures in past data, whilst flexibly accommodating expert opinion.

11. Appendix

11.1. Proof of equation (9)

Since $E(v) = \alpha E(v|u \leq \alpha) + (1 - \alpha)E(v|u > \alpha)$, the numerator of the middle expression is (9) is

$$\begin{aligned} E(v|u > \alpha) - E(v|u \leq \alpha) &= E(v|u > \alpha) - \frac{E(v) - (1 - \alpha)E(v|u > \alpha)}{\alpha} \\ &= \frac{E(v|u > \alpha) - E(v)}{\alpha} = \frac{E\{vI_\alpha(u)\} - E(v)E\{I_\alpha(u)\}}{\alpha(1 - \alpha)} = \frac{\text{cov}\{v, I_\alpha(u)\}}{\alpha(1 - \alpha)}, \end{aligned}$$

and similarly the denominator is

$$E(u|u > \alpha) - E(u|u \leq \alpha) = \frac{\text{cov}\{u, I_\alpha(u)\}}{\alpha(1 - \alpha)} = 0.5,$$

noting $\text{cov}\{u, I_\alpha(u)\} = 0.5\alpha(1 - \alpha)$ in (7). Combining these two results yields the first and third expressions of (9). This completes the proof.

11.2. Proof of equation (10)

Firstly the measure of discordance can be expressed as

$$\gamma_\alpha = -\frac{\text{cov}\{1 - I_\alpha(u), 1 - I_\alpha(v)\}}{\text{cov}\{I_\alpha(u), I_\alpha(u)\}} = \frac{\alpha^2 - C(\alpha, \alpha)}{\alpha(1 - \alpha)},$$

and the measure of dispersion is (assuming the copula of (u, v) is exchangeable),

$$\begin{aligned} \delta_\alpha &= 2E\{(u - v)I_v(u)|(u - \alpha)(v - \alpha) < 0\} \\ &= \frac{2E[(u - v)I_v(u)I_\alpha(u)\{1 - I_\alpha(v)\}]}{2E[I_\alpha(u)\{1 - I_\alpha(v)\}]} = \frac{E[(u - v)I_\alpha(u)\{1 - I_\alpha(v)\}]}{\alpha - C(\alpha, \alpha)} \\ &= \frac{E\{(u - v)I_\alpha(u)\} - E\{(u - v)I_\alpha(u)I_\alpha(v)\}}{\alpha - C(\alpha, \alpha)} = \frac{E\{(u - v)I_\alpha(u)\}}{\alpha - C(\alpha, \alpha)}. \end{aligned}$$

Substituting the above expressions for γ_α and δ_α into the right hand side of (10) yields the expression for ℓ_α in (9), completing the proof.

11.3. Proof of coherence properties in section 5

Independence property

If u and v are independent, then layer dependence is

$$\ell_\alpha = \frac{\text{cov}\{v, I_\alpha(u)\}}{\text{cov}\{u, I_\alpha(u)\}} = 0, \quad 0 \leq \alpha \leq 1$$

since the numerator is zero.

Perfect dependence property

If u and v are comonotonic or $v = u$, then layer dependence

$$\ell_\alpha = \frac{\text{cov}\{v, I_\alpha(u)\}}{\text{cov}\{u, I_\alpha(u)\}} = \frac{\text{cov}\{u, I_\alpha(u)\}}{\text{cov}\{u, I_\alpha(u)\}} = 1, \quad 0 \leq \alpha \leq 1.$$

In addition $\ell_\alpha = 1$ implies from (9) $E(v|u > \alpha) - E(v|u \leq \alpha) = 0.5$ for all α . Substituting $E(v|u \leq \alpha) = \{E(v) - (1 - \alpha)E(v|u > \alpha)\}/\alpha$ yields $E(v|u > \alpha) = 0.5(\alpha + 1)$. Hence the conditional expectation

$$E(v|u = \alpha) = -\frac{d}{d\alpha} \{(1 - \alpha)E(v|u > \alpha)\} = \alpha, \quad 0 \leq \alpha \leq 1,$$

implying $v = u$, completing the proof. A similar proof holds for countermonotonicity where $v = 1 - u$.

Symmetry

Replacing v with $1 - v$ yields layer dependence

$$\frac{\text{cov}\{1 - v, I_\alpha(u)\}}{\text{cov}\{u, I_\alpha(u)\}} = -\frac{\text{cov}\{v, I_\alpha(u)\}}{\text{cov}\{u, I_\alpha(u)\}} = -\ell_\alpha,$$

whilst replacing u with $1 - u$ yields

$$\frac{\text{cov}\{v, I_\alpha(1 - u)\}}{\text{cov}\{1 - u, I_\alpha(1 - u)\}} = -\frac{\text{cov}\{v, 1 - I_{1-\alpha}(u)\}}{\text{cov}\{u, 1 - I_{1-\alpha}(u)\}} = -\frac{\text{cov}\{v, I_{1-\alpha}(u)\}}{\text{cov}\{u, I_{1-\alpha}(u)\}} = -\ell_{1-\alpha}.$$

Using a similar proof, replacing v and u with $1 - v$ and $1 - u$, respectively, yields layer dependence $\ell_{1-\alpha}$.

Correlation order

If (u^*, v^*) exceeds (u, v) in correlation order, then

$$\text{cov}\{f(u^*), g(v^*)\} \geq \text{cov}\{f(u), g(v)\}$$

for any non-decreasing functions f and g (Dhaene et al., 2009). Therefore the numerator of layer dependence in (6) $\text{cov}\{v^*, I_\alpha(u^*)\} \geq \text{cov}\{v, I_\alpha(u)\}$ since $I_\alpha(u)$ is increasing in u , and the denominators are the same. Thus (u^*, v^*) has higher layer dependence than (u, v) .

Bounds

Since layer dependence preserves correlation order, $\ell_\alpha \leq 1$ since $\ell_\alpha = 1$ if and only if u and v are comonotonic and comonotonicity represents maximum correlation order (Dhaene et al., 2009). Similarly $\ell_\alpha \geq -1$ noting countermonotonicity represents minimum correlation order and leads to $\ell_\alpha = -1$.

References

- Bairamov, I., S. Kotz, and T. Kozubowski (2003). A new measure of linear local dependence. *Statistics: A Journal of Theoretical and Applied Statistics* 37(3), 243–258.
- Bjerve, S. and K. Doksum (1993). Correlation curves: Measures of association as functions of covariate values. *The Annals of Statistics*, 890–902.
- Bouyé, E., V. Durrleman, A. Nikeghbali, G. Riboulet, and T. Roncalli (2000). Copulas for finance-a reading guide and some applications. *Available at SSRN 1032533*.
- Czado, C. (2010). Pair-copula constructions of multivariate copulas. In *Copula theory and its applications*, pp. 93–109. Springer.
- Denuit, M., J. Dhaene, M. Goovaerts, and R. Kaas (2005). Actuarial theory for dependent risks: Measures, orders and models. *John Wiley&Sons*.
- Dhaene, J., M. Denuit, and S. Vanduffel (2009). Correlation order, merging and diversification. *Insurance: Mathematics and Economics* 45(3), 325–332.
- Durante, F., J. Fernández-Sánchez, and R. Pappadà (2014). Copulas, diagonals, and tail dependence. *Fuzzy Sets and Systems*.
- Embrechts, P., A. McNeil, and D. Straumann (2002). Correlation and dependence in risk management: properties and pitfalls. *Risk management: value at risk and beyond*, 176–223.
- Genest, C. and L.-P. Rivest (1993). Statistical inference procedures for bivariate archimedean copulas. *Journal of the American statistical Association* 88(423), 1034–1043.
- Holland, P. W. and Y. J. Wang (1987). Dependence function for continuous bivariate densities. *Communications in Statistics-Theory and Methods* 16(3), 863–876.
- Joe, H. (1997). *Multivariate models and dependence concepts*, Volume 73. CRC Press.
- Jones, M. (1996). The local dependence function. *Biometrika* 83(4), 899–904.
- Kojadinovic, I. and J. Yan (2010). Comparison of three semiparametric methods for estimating dependence parameters in copula models. *Insurance: Mathematics and Economics* 47(1), 52–63.
- Mandel, B., D. Morgan, and C. Wei (2012). The role of bank credit enhancements in securitization. *Federal Reserve Bank of New York Economic Policy Review* 18(2), 35–46.

- McNeil, A., R. Frey, and P. Embrechts (2005). *Quantitative risk management*. Princeton University Press.
- Nelson, R. (1999). An Introduction to Copulas. *Lecture Notes in Statistics* 139.
- Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association* 84(406), 487–493.
- Schechtman, E. and S. Yitzhaki (1999). On the proper bounds of the gini correlation. *Economics letters* 63(2), 133–138.
- Sweeting, P. and F. Fotiou (2013). Calculating and communicating tail association and the risk of extreme loss. *British Actuarial Journal* 18(01), 13–72.
- Venter, G. G. (2002). Tails of copulas. In *Proceedings of the Casualty Actuarial Society*, Volume 89, pp. 68–113.
- Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics* 17(1), 43–54.

Chapter 3

Mean and risk densities

The following paper introduces, analyzes and illustrates mean and risk densities which capture mean and risk behaviour across layers of a loss.

Mean and risk densities and their applications to risk management

Abstract

This paper proposes a framework to analyze mean and distortion risk across layers forming a random loss. Layers are standard insurance and financial constructs representing insurance coverage, capital shortfall, derivative payouts and debt tranches. Layers are expressed using percentiles or Values-at-Risk which adjust to the shape and scale of the probability distribution. The proposed framework yields insights and solutions to common risk management problems which are difficult or awkward to solve using standard statistical methods.

Keywords: Layers; density; distortion; Value-at-Risk; reinsurance; tranches.

1. Introduction to layers and this paper

This section first summarises the well established concept of layers and their practical applications. Key contributions of this paper are then highlighted and are developed in subsequent sections.

The layer $[a, b]$ of a random loss $x \geq 0$ is defined as the excess of x over a with the excess capped at $b - a$:

$$\min\{\max(x - a, 0), b - a\} = L_b(x) - L_a(x), \quad L_k(x) \equiv \min(x, k),$$

where $L_k(x)$ is x capped at k . Of interest in this paper are layers over various k for a single random variable x , thus drop x from the notation and write the layer $[a, b]$ of x as $L_b - L_a$. However keep in mind that layers are functions of x and are hence random variables.

Layers are standard insurance and financial constructs. For example insurance and reinsurance with excess a and limit b cover the layer $[a, b]$ of loss x . A capital buffer k divide a loss into two layers: capital consumed $[0, k]$ and capital shortfall $[k, \infty]$. Derivatives and collateralised debt obligations also involve layers, sometimes known as tranches. For example the payout on a call option on x with exercise price k is the layer $[k, \infty]$, whereas the payout on a put option is k minus the $[0, k]$ layer. A debt may be split into tranches $[a, b]$ where a is the attachment point and b is the detachment point. High layers capture rare, extreme outcomes and low layers characterise common, attritional outcomes.

Statistical properties of layers are discussed in Campana and Ferretti (2014), Wang (1998), Wang (1995) and Miccolis (1977). Lee (1988) adopts a graphical approach to explain key concepts and results. The insurance pricing of loss layers is discussed in Evans (2001) and Salzmman (1963). Mandel et al. (2012) and Duffie and Garleanu (2001) discuss tranches in collateralised debt obligations and how they partially enhance the credit quality of the debt.

The following result is central to this paper:

$$x = L_\infty - L_0 = \int_0^\infty dL_k = \int_0^\infty L'_k dk = \int_0^\infty I_k(x) dk, \quad (1)$$

where L'_k is the derivative of L_k with respect to k and is equal to

$$I_k(x) \equiv \begin{cases} 0, & x \leq k, \\ 1, & x > k, \end{cases}$$

an indicator for $x > k$.

In (1), x is composed of infinitesimally small layers dL_k , called the k -layer of x . The k -layer of x is equal to $I_k(x)dk$, an increment dk if x exceeds k . Each k -layer is a random variable and is comonotonic with all other layers since dL_k is non-decreasing in x . The layer decomposition of x in (1) is key to analyzing sources of mean and risk in a loss as shown subsequently in this paper.

Integrating k -layers of x over a to b reproduces the $[a, b]$ layer:

$$\int_a^b dL_k = L_b - L_a = \min(x, b) - \min(x, a) = \min\{\max(x - a, 0), b - a\}.$$

Assuming x has distribution function F , the mean of layer dL_k is

$$E(dL_k) = E\{I_k(x)dk\} = \{1 - F(k)\}dk$$

where E is the expectation. Wang (1996) calls $1 - F$ the layer premium density of x as it indicates the mean value or insurance premium of each layer and it integrates to the overall premium: $\int_0^\infty \{1 - F(k)\}dk = E(x)$. Wang (1996) also distorts $1 - F$ to deliver risk-adjusted premiums.

This paper modifies and extends results in Wang (1996) by defining “VaR layers”: layers expressed on the percentile or Value-at-Risk (VaR) scale. VaRs adjust to the shape and scale of the loss distribution, and are hence comparable across loss distributions. VaRs are standard insurance and financial constructs. For example Solvency II insurance regulation applies 90% and 99.5% VaRs (Eling et al., 2007) to capital requirements. Banking regulations Basel II and Basel III also reference VaRs in risk measurement (Chernobai et al., 2008).

A framework is constructed to analyse mean and distortion risk across VaR layers forming a random loss. The framework yields insights and solutions to risk management problems such as setting optimal capital buffers, analyzing

the credit quality of debt tranches, and applying reinsurance to transform the loss distribution. These problems are typically difficult or awkward to express and solve using standard statistical approaches. In addition optimal quantities derived from the framework, such as capital buffers, are expressed in VaRs which are standard references in finance and insurance as mentioned above.

Remaining sections of this paper are structured as follows. VaR layers are first defined in section 2. Sections 3 to 5 formalise, analyze and illustrate layer means or “mean densities.” Risk densities are covered in sections 6 to 9. Section 10 shows how mean and risk densities are calculated from data and illustrates using historical stock returns. Remaining sections apply mean and risk densities to insurance pricing, collateralised debt obligations, capital setting and reinsurance loss transformation.

2. Decomposing a loss into VaR layers

Consider a continuous random loss $x \geq 0$ with distribution function F and inverse distribution function F^- . For any constant $0 \leq \alpha \leq 1$, the α -VaR or VaR_α of x is $V_\alpha \equiv F^-(\alpha)$. Substitute $k = V_\alpha$ into the first expression for x in (1) to yield a modified decomposition

$$x = L_1^* - L_0^* = \int_0^1 dL_\alpha^* = \int_0^1 (L_\alpha^*)' d\alpha = \int_0^1 I_{V_\alpha}(x) V_\alpha' d\alpha ,$$

where V_α' is the derivative of V_α with respect to α and $L_\alpha^* \equiv L_{V_\alpha} = \min(x, V_\alpha)$ is the capped loss with the cap expressed in VaRs instead of dollars. VaRs are used for the rest of this paper, hence drop the superscript $*$ in the notation. In addition define $u \equiv F(x)$ as the percentile rank of x and hence $I_{V_\alpha}(x) = I_\alpha(u)$.

Therefore rewrite the above result using the simplified notation

$$x = L_1 - L_0 = \int_0^1 dL_\alpha = \int_0^1 L_\alpha' d\alpha = \int_0^1 I_\alpha(u) V_\alpha' d\alpha , \quad (2)$$

where L_α is now defined as $\min(x, V_\alpha)$ for the rest of this paper.

Similar to (1), (2) decomposes x into infinitesimally small VaR layers. The V_α -layer of x is $dL_\alpha = I_\alpha(u) V_\alpha' d\alpha$, an increment $V_\alpha' d\alpha$ if $x > V_\alpha$ or $u > \alpha$. The increment is proportional to V_α' called the V_α -spacing, analogous to spacings between order statistics (Shaked and Shanthikumar, 2007). VaR spacings vary across a loss distribution with larger spacings indicating areas of greater skewness. Increasing the scale of a loss distribution increases VaR spacings by the same scale. The V_α -spacing can be written as

$$V_\alpha' = \frac{d}{d\alpha} F^-(\alpha) = \frac{1}{f(V_\alpha)}, \quad f \equiv F'$$

where f is the loss density. Hence V_α -spacing is increasing in the upper tail of a right skewed distribution where the density gradually tails off to 0.

Similar to §1, integrating V_α -layers forms a larger VaR-layer:

$$\int_\alpha^\beta dL_\pi = L_\beta - L_\alpha = \min(V_\beta, x) - \min(V_\alpha, x) = \min\{\max(x - V_\alpha, 0), V_\beta - V_\alpha\}$$

the excess of x over V_α with the excess capped at $V_\beta - V_\alpha$. The VaR-layer $[V_\alpha, V_\beta]$ is positive when $u > \alpha$, with probability $1 - \alpha$, and reaches its maximum $V_\beta - V_\alpha$ when $u \geq \beta$, with probability $1 - \beta$.

Hence a VaR-layer captures a portion of losses identified from their relative position in the probability distribution. For example the VaR-layer $[V_{0.5}, V_{0.75}]$ captures the top 50% of losses, and the top 25% of losses are capped at the 75th percentile. Endpoints of VaR layers adjust to the shape and scale of the loss distribution. Increasing the scale of the loss distribution increases layer endpoints by the same scale, and similarly when skewness increases. VaR layers are therefore comparable between loss distributions with different scale or shape. In contrast layers determined in dollar terms as in §1 may be attritional or extreme layers depending on the loss distribution.

3. Mean density: expected values of VaR layers

Following Wang (1996), define the mean density of x as the relative mean value of the V_α -layer, yielding

$$m_\alpha \equiv E(L'_\alpha) = E\{I_\alpha(u)V'_\alpha\} = (1 - \alpha)V'_\alpha, \quad (3)$$

for $0 \leq \alpha \leq 1$. The mean density m_α is the product of $1 - \alpha$, the probability of x reaching V_α , and the V_α -spacing V'_α , the width of the V_α -layer of x .

The behaviour of m_α with α depends on relative movements in $1 - \alpha$ and V'_α , and characterises the loss distribution as illustrated in §4. As α increases, $1 - \alpha$ decreases due to a reduced likelihood of x reaching larger VaRs. On the other hand for a right skewed loss distribution where the probability density is decreasing, $V'_\alpha = 1/f(V_\alpha)$ increases with α .

The term “density” is used to refer to m_α due to analogous properties shared with probability densities. Using (2),

$$\int_0^1 m_\alpha d\alpha = E\left(\int_0^1 dL_\alpha\right) = E(x),$$

thus the total area under the mean density is the overall mean loss. The mean density hence spreads the overall mean across layers forming the loss, and identifies layers with high mean contribution. Integrating the mean density over the interval $[\alpha, \beta]$ yields the mean of the $[V_\alpha, V_\beta]$ VaR-layer:

$$\int_\alpha^\beta m_\pi d\pi = E\left(\int_\alpha^\beta dL_\pi\right) = E(L_\beta - L_\alpha). \quad (4)$$

Analogously, integrating a probability density over its entire support yields total probability 1, whilst integrating over other intervals yields the probability over the same interval.

Mean values of VaR-layers represent, for example, the pure premium under an excess-of-loss reinsurance arrangement, expected capital shortfall or consumed, and expected derivative payouts. Section 6 defines risk values of VaR-layers. Combining mean and risk values of VaR-layers yields insights to common risk management problems discussed further in §11.

4. Example mean densities

The following are mean densities for common probability distributions. In all cases $b > 0$ is a scale parameter.

- The uniform distribution over $[0, b]$ yields

$$V_\alpha = b\alpha, \quad V'_\alpha = b, \quad m_\alpha = (1 - \alpha)b,$$

hence the mean density linearly decreases to zero. The mean density is decreasing since uniform VaR-spacings are offset by the lower likelihood of losses reaching higher VaRs. The mean density is also proportional to scale b , and this is the case for all mean densities described below.

- The exponential distribution with mean b yields

$$F(x) = 1 - e^{-x/b}, \quad V_\alpha = -b \ln(1 - \alpha), \quad V'_\alpha = \frac{b}{1 - \alpha}, \quad m_\alpha = b,$$

and the mean density is constant since increasing VaR-spacings exactly offset decreasing probabilities of losses reaching larger VaRs. Hence VaR layers of an exponential distribution have uniform contribution to the overall mean.

- The Pareto distribution with shape parameter $c > 1$ yields

$$F(x) = 1 - \left(\frac{b}{b + x} \right)^c, \quad V_\alpha = b \left\{ (1 - \alpha)^{-1/c} - 1 \right\},$$

$$V'_\alpha = \frac{b}{c(1 - \alpha)^{1/c+1}}, \quad m_\alpha = \frac{b}{c(1 - \alpha)^{1/c}}.$$

The mean density increases to infinity as $\alpha \rightarrow 1$: increasing VaR spacings more than offset decreasing probabilities of reaching larger VaRs. Mean contributions are concentrated in high VaR layers of a Pareto distribution.

- For the Weibull distribution with shape parameter $k > 0$,

$$F(x) = 1 - e^{-(x/b)^k}, \quad V_\alpha = b \{-\ln(1 - \alpha)\}^{1/k},$$

$$V'_\alpha = \frac{b \{-\ln(1 - \alpha)\}^{1/(k-1)}}{k(1 - \alpha)}, \quad m_\alpha = \frac{b \{-\ln(1 - \alpha)\}^{1/(k-1)}}{k}.$$

Setting $k = 1$ yields the exponential distribution and $m_\alpha = b$. Reducing k increases skewness and yields increasing m_α . Vice versa if $k > 1$.

Figure 1 plots mean densities for the above loss distributions with parameters selected to satisfy $E(x) = 1$. Hence mean densities in Figure 1 all integrate to one, and their shape indicates relative contributions from different layers of the loss. As described above, the exponential distribution has a flat mean density and the overall mean value is spread uniformly across VaR layers. For the Pareto distribution, the bulk of the overall mean value is contributed by high VaR layers as reflected by an increasing mean density. The opposite applies to the uniform distribution and the Weibull distribution ($k > 1$ in this case).

The constant mean density for the exponential distribution provides a benchmark for assessing skewness at various parts of a loss distribution. A decreasing mean density (for example uniform distribution and Weibull distribution for $k > 1$) implies VaR-spacings increase at a lower rate compared to the exponential, and hence lower skewness comparatively. Vice versa for an increasing mean density (for example Pareto distribution and Weibull distribution for $k < 1$). The derivative of m_α with respect to α is

$$m'_\alpha = (1 - \alpha)V''_\alpha - V'_\alpha.$$

Hence the mean density is increasing at α if and only if $(1 - \alpha)V''_\alpha/V'_\alpha > 1$. Rewriting this inequality in terms of the loss density f yields the condition

$$\frac{(1 - \alpha)V''_\alpha}{V'_\alpha} = -\frac{(1 - \alpha)f'(V_\alpha)}{\{f(V_\alpha)\}^2} > 1, \quad V''_\alpha = -\frac{f'(V_\alpha)}{\{f(V_\alpha)\}^3}.$$

The probability distribution of x has greater skewness than the exponential over layers where the above inequality holds, and vice versa.

5. Properties of mean densities

5.1. Monotonic loss transformations

Straightforward changes apply to the mean density when a monotonic transformation is applied to the loss. Consider the transformation $y = g(x)$ where $g > 0$ is an increasing function. Then y has $\text{VaR}_\alpha V_\alpha(y) = g\{V_\alpha(x)\}$ and the mean density of y at the VaR_α layer is

$$m_\alpha(y) = (1 - \alpha)V'_\alpha(y) = (1 - \alpha)g'\{V_\alpha(x)\}V'_\alpha(x) = m_\alpha(x)g'\{V_\alpha(x)\},$$

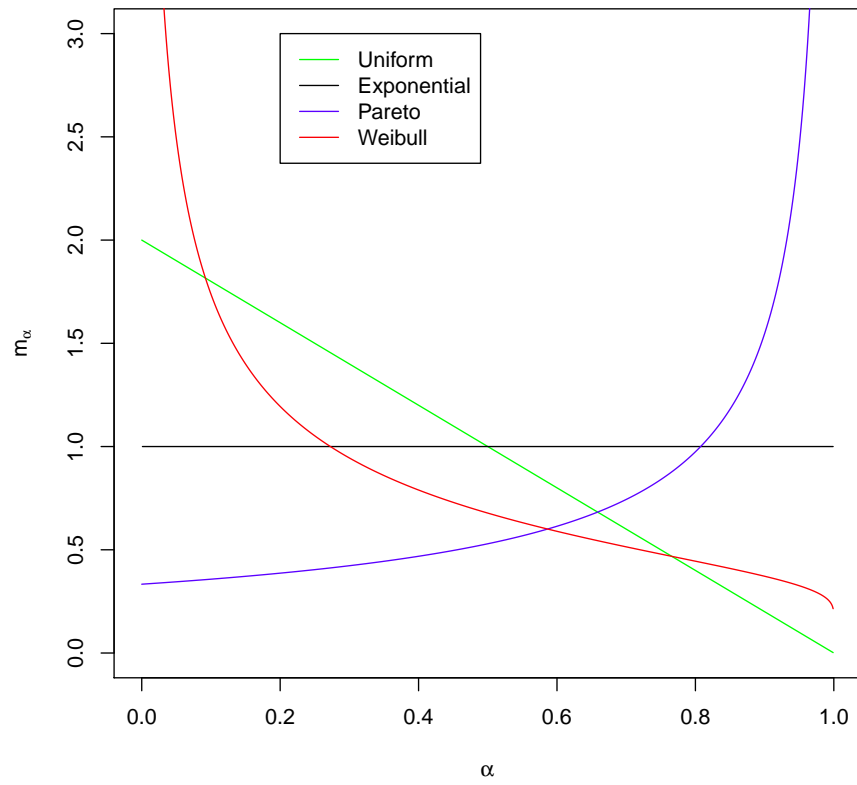


Figure 1: Mean densities for uniform ($b = 2$), exponential ($b = 1$), Pareto ($b = 0.5, c = 1.5$) and Weibull ($b = 1.13, k = 2$) distributions. All distributions have mean 1 hence the area underneath each curve is 1.

where $V_\alpha(x)$ and $m_\alpha(x)$ are respectively the VaR_α and mean density of x . Hence the mean density of y is the mean density of x multiplied by the derivative $g'\{V_\alpha(x)\}$. In particular if $y = cx$ where c is constant then $m_\alpha(y) = cm_\alpha(x)$. This result implies mean densities are proportional to the scale of the loss distribution, as demonstrated by the examples in §4.

It is also straightforward to show that if $y = g(x)$ is a decreasing transformation then $V_\alpha(y) = g\{V_{1-\alpha}(x)\}$ and $m_\alpha(y) = m_{1-\alpha}(x)g'\{V_{1-\alpha}(x)\}$. Hence in this case the mean density of x at the $V_{1-\alpha}$ -layer is referenced.

5.2. Unique characterisation of the loss distribution

Assuming $V_0 = 0$, the mean density uniquely characterises the loss distribution. Re-arranging (3) and integrating yields

$$F^-(\alpha) = V_\alpha = \int_0^\alpha \frac{m_\pi}{1-\pi} d\pi .$$

VaR_α is a weighted cumulative of the mean density up to the VaR_α layer. Larger mean density values lead to greater VaRs.

5.3. Connection to hazard function

Substituting $V'_\alpha = 1/f(V_\alpha)$ to (3) yields

$$m_\alpha = \frac{1-\alpha}{f(V_\alpha)} = \left\{ \frac{f(V_\alpha)}{1-F(V_\alpha)} \right\}^{-1}, \quad F(V_\alpha) = \alpha .$$

Therefore the mean density at the V_α -layer of x is the reciprocal of the hazard function (Hogg and Klugman, 2009) at $x = V_\alpha$. The hazard at V_α indicates the instantaneous probability of $x = V_\alpha$ given it reaches V_α . The above result implies a lower hazard at V_α leads to a higher mean value of the V_α -layer.

6. Risk density: distortion risks of VaR layers

This section defines risk densities which indicate the distortion risk of VaR layers. Similar to mean densities, risk densities identify contributions by various VaR layers to overall risk. Risks are important inputs to decisions involving volatility. For example premium loadings, capital buffers and expected investment returns are driven by risk values. Mean values are insufficient decision making inputs as they do not capture the spread of outcomes.

Wang (1996) defines distortion risk based on the mean value under a distorted distribution function $\Phi \circ F$ where Φ is an increasing, convex distortion operator satisfying $\Phi(0) = 0$ and $\Phi(1) = 1$. The distortion risk of x is hence

$$r \equiv \mathcal{E}(x) - E(x) = \int_0^\infty [1 - \Phi\{F(x)\}] dx - \int_0^\infty \{1 - F(x)\} dx ,$$

where \mathcal{E} calculates expectations under the distorted distribution $\Phi \circ F$. Choo and De Jong (2009) shows the equivalence between distortion risks, loss aversion premiums and spectral risks (Acerbi, 2002). Examples of distortion risks using various Φ are discussed in Wang (1995), Wang (2000) and Choo and De Jong (2009), and include the proportional hazards risk, conditional-tail-expectation and expected-maximal-loss. Distortion risks are coherent (Artzner et al., 1999): positively homogenous, translation invariant, monotonic and subadditive.

The risk density of x indicates the distortion risk of the V_α -layer of x and applying similar calculations as above yields

$$r_\alpha \equiv \mathcal{E}(L'_\alpha) - \mathbb{E}(L'_\alpha) = \{1 - \Phi(\alpha)\}V'_\alpha - (1 - \alpha)V'_\alpha = \{\alpha - \Phi(\alpha)\}V'_\alpha . \quad (5)$$

This result follows since $\mathcal{E}\{I_\alpha(u)\}$ is the distorted probability of $u > \alpha$ which is $1 - \Phi(\alpha)$, whereas the original probability is $1 - \alpha$. Note $r_\alpha \geq 0$ for all α since Φ is convex implying $\Phi(\alpha) \leq \alpha$. In addition once the distribution of x is established, V_α and hence V'_α , defining the α -VaR layer and spacing of x , are unaffected by distortion for all $0 \leq \alpha \leq 1$.

The risk density r_α in (5) is composed of two factors. The difference $\alpha - \Phi(\alpha)$ represents the extent of distortion at V_α , with $\alpha = \Phi(\alpha)$ indicating zero distortion and $r_\alpha = 0$. The other factor is layer width or V_α -spacing V'_α . Increasing either factor increases r_α .

Similar to mean densities, integrating risk densities yields the risk of larger VaR layers. Integrating r_α over all α yields the overall distorted risk of x :

$$\int_0^1 r_\alpha d\alpha = \int_0^1 \{\mathcal{E}(L'_\alpha) - \mathbb{E}(L'_\alpha)\} d\alpha = \mathcal{E}(x) - \mathbb{E}(x) = r .$$

Alternatively, V_α -layers over all $0 \leq \alpha \leq 1$ are comonotonic, and distortion risks of comonotonic random variables are additive (Wang et al., 1997). Integrating r_α over $[\alpha, \beta]$ yields the distorted risk of VaR-layer $[V_\alpha, V_\beta]$ of x :

$$\int_\alpha^\beta r_\pi d\pi = \mathcal{E}(L_\alpha - L_\beta) - \mathbb{E}(L_\alpha - L_\beta) .$$

Again an analogous property applies to probability densities.

7. VaR risk ratios

The risk ratio at the V_α -layer is the ratio between risk and mean densities:

$$r_\alpha^* \equiv \frac{r_\alpha}{m_\alpha} = \frac{\{\alpha - \Phi(\alpha)\}V'_\alpha}{(1 - \alpha)V'_\alpha} = \frac{\alpha - \Phi(\alpha)}{1 - \alpha} , \quad (6)$$

noting V'_α appears in both numerator and denominator and hence cancels after division. Risk ratios indicate the risk of a VaR-layer relative to its mean. Risk

ratios only depend on the distortion operator Φ and are independent of the loss distribution. Risk ratios are also larger at higher layers, since

$$\frac{dr_\alpha^*}{d\alpha} = \frac{1 - \Phi(\alpha) - (1 - \alpha)\Phi'(\alpha)}{(1 - \alpha)^2} \geq 0 .$$

The inequality holds since

$$1 - \Phi(\alpha) = \int_\alpha^1 \Phi'(u) du \geq \int_\alpha^1 \Phi'(\alpha) du = (1 - \alpha)\Phi'(\alpha) ,$$

noting Φ is convex hence Φ' is increasing. Thus higher VaR-layers are always relatively riskier than lower VaR-layers as a proportion of their mean value. Risk ratios at the lowest and highest VaR-layers are $r_0^* = 0$ and $r_1^* = \Phi'(1) - 1$ respectively, where the latter is derived using the L'Hôpital's rule.

The overall risk ratio of x can be written in terms of r_α^* as

$$\frac{r}{E(x)} = \frac{\int_0^1 r_\pi d\pi}{\int_0^1 m_\pi d\pi} = \frac{\int_0^1 r_\pi^* m_\pi d\pi}{\int_0^1 m_\pi d\pi} .$$

Therefore the overall risk ratio is a weighted average of individual risk ratios. Weights are given by mean density values. As r_α^* is increasing in α and independent of the loss distribution as mentioned above, the overall risk ratio is high if the mean density is high at higher layers resulting in higher risk ratios being weighted more heavily. Hence skewed loss distributions are relatively riskier, consistent with intuition. Using the example mean densities in §4, the Pareto has a higher overall risk ratio than the uniform, exponential and Weibull due to an increasing mean density. Examples are further discussed in the next section.

8. Example risk ratios and risk densities

The following are risk ratios based on distortion operators discussed in Wang (1995) and Choo and De Jong (2009). As highlighted in the previous section, risk ratios only depend on the distortion operator. Multiplying risk ratios with the mean density yields the risk density.

- Assume $\Phi(\alpha) = I_c(\alpha)(\alpha - c)/(1 - c)$ where $0 \leq c \leq 1$ is a parameter. The overall distortion risk of x involves the conditional-tail-expectation and is given by $E(x|x > V_c) - E(x)$, the expected value of losses above V_c in excess of the overall mean loss. The risk ratio using (6) is

$$r_\alpha^* = \{1 - I_c(\alpha)\} \frac{\alpha}{1 - \alpha} + I_c(\alpha) \frac{c}{1 - c} .$$

Risk ratios increase from 0 at the lowest layer and reaches $c/(1 - c)$ at the V_c -layer, and remains constant at $c/(1 - c)$ for higher layers. Note larger c yields higher r_α^* for all $0 \leq \alpha \leq 1$, hence c indicates overall risk aversion.

- Assume the power distortion operator $\Phi(\alpha) = \alpha^n$ where $n \geq 1$. If n is an integer, the distortion risk is $E\{\max(x_1, \dots, x_n)\} - E(x)$ or the expected-maximal-loss in excess of the mean loss, where x_1, \dots, x_n are independent copies of x . The risk ratio in this case is

$$r_\alpha^* = \frac{\alpha - \alpha^n}{1 - \alpha} = \sum_{i=1}^{n-1} \alpha^i,$$

where the final expression assumes integer n , and is a polynomial of degree $n - 1$ with unit coefficients. Larger n yields higher risk ratios for all α , hence n indicates risk aversion.

- Assume the proportional hazards transform, $\Phi(\alpha) = 1 - (1 - \alpha)^{1/\gamma}$ where $\gamma \geq 1$. The distortion risk is $\int_0^\infty \{S(x)\}^{1/\gamma} dx - \int_0^\infty S(x) dx$ where S is the survival function, that is $S \equiv 1 - F$. The risk ratio is

$$r_\alpha^* = \frac{\alpha - \{1 - (1 - \alpha)^{1/\gamma}\}}{1 - \alpha} = (1 - \alpha)^{1/\gamma - 1} - 1,$$

and increases to ∞ as α approaches 1. Higher γ increases risk ratios across all layers and indicates overall risk aversion.

The four panels in Figure 2 graph mean densities shown in Figure 1 and risk densities using a power distortion operator $\Phi(\alpha) = \alpha^3$. The risk ratio is $r_\alpha^* = \alpha(1 + \alpha)$. For the uniform loss distribution, risk is highest around median VaR-layers since increasing risk ratios are applied to a decreasing mean density. Risk densities are increasing for exponential, Pareto and Weibull loss distributions. Pareto has the highest overall risk (the area under the risk density), since higher risk ratios are applied to higher mean densities at higher VaR layers.

9. Properties of risk densities and risk ratios

9.1. Unique characterisation of distortion operator

Risk ratios r_α^* over $0 \leq \alpha \leq 1$ uniquely characterise the distortion operator Φ . Given r_α^* , rearranging (6) yields the distortion operator

$$\Phi(\alpha) = \alpha - (1 - \alpha)r_\alpha^*.$$

Thus an alternative formulation of Φ is to first specify risk ratios across all VaR-layers and then set Φ using the above result. The condition $\Phi(0) = 0$ requires $r_0^* = 0$ and $\Phi(1) = 1$ is satisfied generally. Increasing and convex Φ respectively require $1 + r_\alpha^* - (1 - \alpha)(r_\alpha^*)' \geq 0$ and $2(r_\alpha^*)' - (1 - \alpha)(r_\alpha^*)'' \geq 0$.

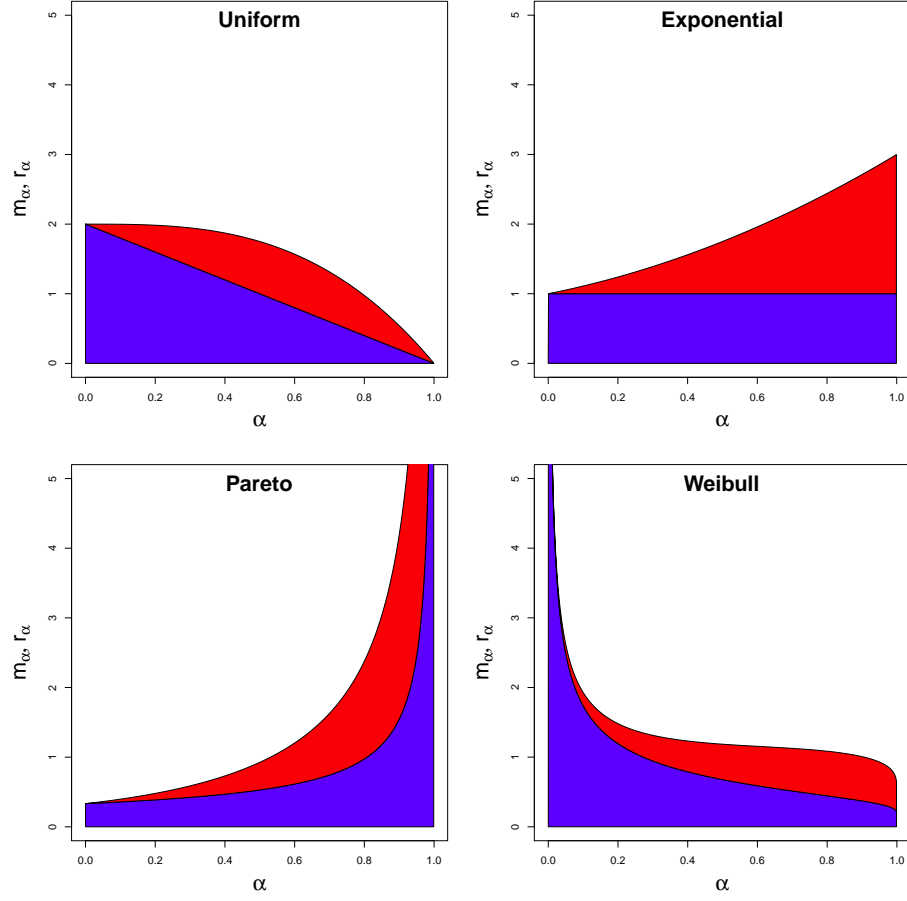


Figure 2: Mean densities (blue area) and risk densities (red area) for the 4 loss distributions in Figure 1. The distortion operator is $\Phi(\alpha) = \alpha^3$ and the risk ratio is $r_\alpha^* = \alpha(1 + \alpha)$ across all loss distributions. Note how risks are concentrated in different layers for different loss distributions. For the Pareto, risk is concentrated in the highest layers whereas risk is concentrated in middle layers for the uniform.

9.2. Monotonic loss transformations

Similar to mean densities, straightforward changes apply to risk densities under monotonic loss transformations. Again assume $y = g(x)$ is an increasing transformation. The risk density of y is

$$r_\alpha(y) = m_\alpha(y)r_\alpha^* = m_\alpha(x)g'\{V_\alpha\}r_\alpha^* = r_\alpha(x)g'\{V_\alpha(x)\}$$

where $r_\alpha(x)$ is the risk density of x and noting risk ratios r_α^* are independent of the loss distribution and hence unchanged. Therefore mean and risk densities change identically under increasing transforms. In addition risk densities are proportional to the scale of the loss distribution, similar to mean densities.

Similarly if g is decreasing then the risk density of y is $r_{1-\alpha}(x)g'\{V_{1-\alpha}(x)\}$.

10. Empirical calculations of mean and risk densities

It is common for insurance and financial companies to collect data or simulate future outcomes to make risk management decisions. Given an ordered sample of losses ℓ_1, \dots, ℓ_n , calculate mean and risk densities as follows:

- Set empirical VaRs equal to ordered observations: $\hat{V}_{i/n} = \ell_i$. The $V_{i/n}$ -layer is $I_{\ell_i}(x)(\ell_{i+1} - \ell_i)$ where x is a random observation.
- The empirical mean density is $\hat{m}_{i/n} = (1 - i/n)\hat{V}'_{i/n}$, for $i = 0, \dots, n-1$ where the empirical derivative $\hat{V}'_{i/n} = n(\hat{V}_{(i+1)/n} - \hat{V}_{i/n})$ and $\hat{V}_0 = 0$.
- Risk ratios are computed using the specified distortion operator: $r_{i/n}^* = \{i/n - \Phi(i/n)\}/(1 - i/n)$. The empirical risk density is $\hat{r}_{i/n} = \hat{m}_{i/n} * r_{i/n}^*$.
- If necessary, apply parametric or non-parametric smoothing to empirical mean and risk densities to reveal underlying patterns.

As $n \rightarrow \infty$, empirical VaRs approach true VaRs and the above algorithm converges: $\hat{m}_{i/n} \rightarrow m_\alpha$ and $\hat{r}_{i/n} \rightarrow r_\alpha$ where $\alpha = i/n$.

Overall mean and risk are the respective sums (rather than integrals)

$$\sum_{i=0}^{n-1} \frac{m_{i/n}}{n}, \quad \sum_{i=0}^{n-1} \frac{r_{i/n}}{n}.$$

Taking sums over other subsets yield mean and risk over corresponding VaR-layers. In addition manipulating the first summation above yields the arithmetic mean $\sum_{i=1}^n \ell_i/n$ of the loss sample.

The following calculates mean and risk densities from NASDAQ, S&P and FTSE daily returns between 1985 and 2015 assuming a stationary distribution. Form a hypothetical portfolio of \$100 in each market index. Focus on daily

investment losses rather than gains, hence switch the signs of historical returns and gains are treated as negative losses. In addition apply the distortion operator $\Phi(\alpha) = I_c(\alpha)(\alpha - c)/(1 - c)$ where $c = 0.75$, implying overall risk is the difference between expected losses above $\text{VaR}_{0.75}$ and the overall mean. Note the condition $x \geq 0$ is relaxed in this illustration, as the interest is in the shape of mean and risk densities rather than their actual values.

Figure 3 illustrates results. Empirical densities and VaRs of daily investment losses are shown in addition to empirical mean and risk densities. The following are key observations:

- NASDAQ has more skewed losses and gains than S&P and FTSE, and in particular larger VaR-spacings. As a result NASDAQ has larger mean and risk densities, and larger mean loss and risk overall. S&P and FTSE behave similarly particularly in the tails.
- Empirical mean densities are large at low VaR-layers, then decrease rapidly at higher VaR-layers. High mean density values at low layers are formed by large VaR-spacings in left tails of loss distributions, and high probabilities of exceeding low VaRs.
- Relatively stable mean densities at higher VaR-layers indicate upper tails of loss distributions have similar shape as the exponential distribution.
- Empirical risk densities paint a different picture from mean densities: risk is mainly contributed by higher layers, unlike for the mean. This observation is consistent with common knowledge that extreme market losses, whilst rare, contribute significantly to portfolio volatility.

11. Applications of mean and risk densities

Remaining sections apply mean and risk densities to common risk management problems involving reinsurance, capital buffers and debt tranching. Current statistical tools provide identical solutions in some cases. However mean and risk densities offer a more elegant solution and yield additional insights.

12. Pricing and insuring loss layers

Suppose the $[V_\alpha, V_\beta]$ VaR-layer of loss x is insured or reinsured. Hence V_a is the excess and V_b is the limit of the coverage. The premium is

$$\begin{aligned} P &= \int_\alpha^\beta (m_\pi + r_\pi) d\pi = \int_\alpha^\beta \{1 - \pi + \pi - \Phi(\pi)\} V'_\pi d\pi = \int_\alpha^\beta \{1 - \Phi(\pi)\} V'_\pi d\pi \\ &= \{1 - \Phi(\beta)\} V_\beta - \{1 - \Phi(\alpha)\} V_\alpha + \int_\alpha^\beta V_\pi \Phi'(\pi) d\pi, \end{aligned}$$

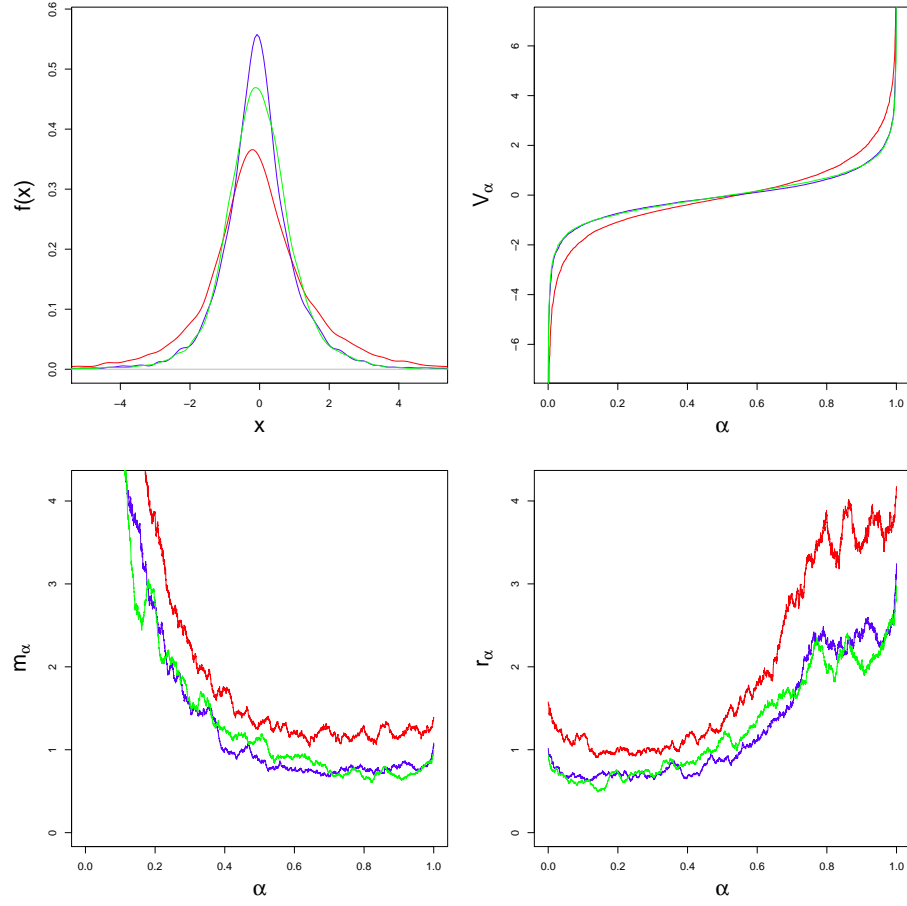


Figure 3: Top left and right panels plot empirical loss densities and distributions, respectively. Bottom left and right panels plot smoothed empirical mean and risk densities, respectively. Red, blue and green represent NASDAQ, S&P and FTSE, respectively.

where $\int_{\alpha}^{\beta} m_{\pi} d\pi$ is the pure premium and $\int_{\alpha}^{\beta} r_{\pi} d\pi$ is the risk loading forming the overall premium P . Both components increase as coverage widens: lower α or higher β . Changes to overall premium due to changes in the excess or limit are the partial derivatives

$$\frac{\partial P}{\partial \alpha} = -(m_{\alpha} + r_{\alpha}), \quad \frac{\partial P}{\partial \beta} = m_{\beta} + r_{\beta}.$$

The previous illustration in Figure 2 shows the pure premium and risk loading for various α and β assuming uniform, exponential, Pareto and Weibull loss distributions and the distortion operator $\Phi(\alpha) = I_c(\alpha)(\alpha - c)/(1 - c)$. The overall premium is the total area over $[\alpha, \beta]$. For example, for an exponential loss distribution, the pure premium increases at the same rate as the excess and limit are changed. However the risk loading increases at an increasing rate when the limit increases, and forms a larger component of the overall premium when high VaR-layers are insured.

Intuitively, the purchaser of insurance or reinsurance would select an excess and limit covering the bulk of the area under mean and risk densities. This implies higher excesses and limits for exponential and Pareto loss distributions. On the other hand, lower excesses and limits would suffice for uniform (and potentially Weibull) loss distributions, as higher layers have lower mean and risk contributions. The optimal excess and limit depends on the interaction between several factors including the premium cost, the purchaser's budget and utility function.

13. Credit quality of debt tranches

Consider a debt with fixed principal p and random default loss $0 \leq x \leq p$. The credit rating of the debt is typically based on the following quantities:

$$\text{PD} = P(x > 0), \quad \text{PEL} = \frac{E(x)}{p}, \quad \text{RR} = \frac{r}{E(x)}$$

where PD is the default probability, PEL is the expected default loss as a proportion of principal and RR is the risk of the default loss relative to its mean value. A high credit rating may be issued for example if a combination of PD, PEL and RR falls below specified limits (Das and Stein, 2011).

Suppose the debt is split into tranches. The default loss on the tranche from V_{α} to V_{β} is $\max\{\min(x - V_{\alpha}, 0), V_{\beta} - V_{\alpha}\}$ and is hence the VaR-layer $[V_{\alpha}, V_{\beta}]$ of x . The principal of this tranche is layer width $V_{\beta} - V_{\alpha}$. In addition V_{α} and V_{β} are called the attachment and detachment points of the tranche, respectively. There is a default loss if $x > V_{\alpha}$, and a complete loss of principal if $x \geq V_{\beta}$.

Splitting a debt into tranches is used in securitisation to enhance the credit quality of portions of the debt (Mandel et al., 2012). Debt tranches are common in collateralised debt obligations (Duffie and Garleanu, 2001). Lower, junior

tranches are more likely to suffer a default loss and are assigned a lower credit rating. Higher, senior tranches only suffer default losses after losses have passed through junior tranches and hence typically receive a higher credit rating.

Corresponding credit rating statistics for the tranche or layer $[V_\alpha, V_\beta]$, using mean and risk densities defined in (3) and (5), are

$$\text{PD} = \text{P}(x > V_\alpha) = 1 - \alpha, \quad \text{PEL} = \frac{\int_\alpha^\beta m_\pi d\pi}{V_\beta - V_\alpha} = \frac{\int_\alpha^\beta (1 - \pi) V'_\pi d\pi}{\int_\alpha^\beta V'_\pi d\pi},$$

$$\text{RR} = \frac{\int_\alpha^\beta r_\pi d\pi}{\int_\alpha^\beta m_\pi d\pi} = \frac{\int_\alpha^\beta r_\pi^* m_\pi d\pi}{\int_\alpha^\beta m_\pi d\pi},$$

where as before V'_π is the derivative of V_π with respect to π and $r_\pi^* = r_\pi/m_\pi$ is the risk ratio of the V_π -layer of x . Hence PEL is a weighted average of default probabilities $1 - \pi$ from $\pi = \alpha$ to β , where weights are VaR derivatives or spacings. Setting $\beta \rightarrow \alpha$ leads to $\text{PEL} \rightarrow 1 - \alpha$ which is the PD. RR is a weighted average of risk ratios from α to β . Note PEL in this case is also the rate-on-line in excess-of-loss reinsurance ignoring the risk loading.

Both PD and PEL on tranche $[V_\alpha, V_\beta]$ decrease as α or β increases. Hence the probability and expected proportion of default are smaller for higher tranches, consistent with common knowledge. However RR increases since r_π^* is increasing in π . Hence the default loss becomes riskier in higher tranches, even though a default loss is less likely and on average smaller. In addition, as default losses on tranches are comonotonic, there is no diversification benefit from holding multiple tranches originating from the same loss. This lack of diversification is a reason for catastrophic financial losses during the 2008 crisis (Kolb, 2010).

14. Setting capital buffers

Suppose capital V_c is held to cover loss x . There is a shortfall if $x > V_c$ and surplus if $x \leq V_c$. The capital shortfall is the VaR-layer $[V_c, V_1]$ of x and capital consumed is the VaR-layer $[0, V_c]$. The following derives c or V_c in two different ways, using mean and risk densities.

14.1. Limiting expected shortfall

Suppose capital V_c is held to restrict expected shortfall to a proportion s of overall expected loss where s is small, say 1%. Hence c is solved from

$$\frac{\text{E}(L_1 - L_c)}{\text{E}(L_1 - L_0)} = \frac{\int_c^1 m_\alpha d\alpha}{\int_0^1 m_\alpha d\alpha} = s. \quad (7)$$

An equivalent standard statistical expression is

$$\frac{\text{E}\{\max(x - V_c, 0)\}}{\text{E}(x)} = s.$$

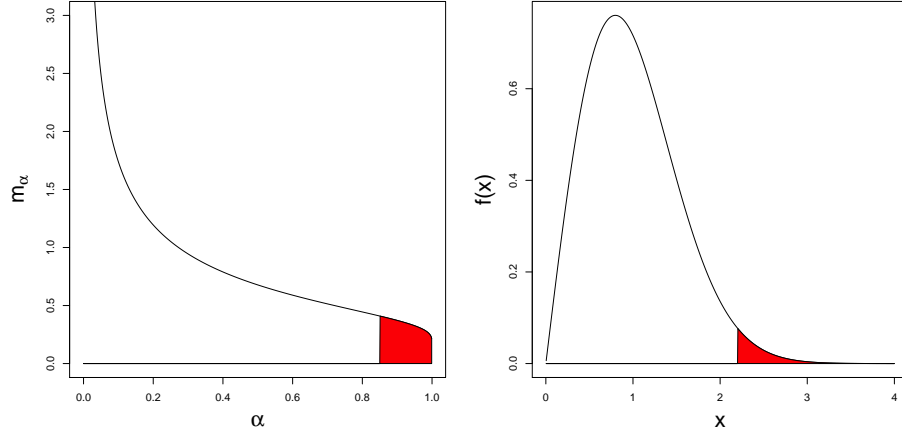


Figure 4: The left panel graphs the mean density and highlights capital satisfying a target expected shortfall as a proportion of the overall expected loss (red area). The right panel illustrates an analogous approach using the probability density, where the resulting capital is the familiar VaR and achieves a fixed shortfall probability (red area).

Equation (7) implies c is selected so that the relative area under the upper tail of mean density m_α is s . Figure 4 provides an illustration using a Weibull loss distribution. Smaller s requires higher c and V_c .

A similar, common approach is to set capital to restrict the *probability* of shortfall, as opposed to *expected* shortfall, to say again s . Under this approach, capital is selected such that the upper tail area under the probability density (rather than the mean density) is s . It is straightforward to show that required capital is V_{1-s} . This approach is also illustrated in Figure 4. Insurance and banking regulations typically set capital at V_{1-s} with pre-determined s (Eling et al. (2007), Chernobai et al. (2008)).

Setting capital based on expected shortfall is more appropriate than using shortfall probability as the former reflects the shape of the loss distribution. This is demonstrated by the following examples. For an exponential loss distribution, the mean density is flat (see §4) and (7) reads $1 - c = s$ therefore capital is V_{1-s} . For example if expected shortfall is 1% of overall mean loss then capital is $V_{0.99}$. Setting shortfall probability at s also yields capital V_{1-s} , hence both approaches are the same for exponential loss distributions. For a Pareto loss distribution with shape parameter γ , solving (7) yields

$$\int_c^1 \frac{1}{(1-\alpha)^{1/\gamma}} d\alpha = s \int_0^1 \frac{1}{(1-\alpha)^{1/\gamma}} d\alpha, \quad c = 1 - s^{\gamma/(\gamma-1)} \geq 1 - s.$$

In this case capital $V_c \geq V_{1-s}$. Therefore referring to expected shortfall rather than shortfall probability increases capital. The reason is the Pareto distribution has a thicker upper tail than the exponential, hence increasing expected

shortfall. Reducing the skewness of the Pareto by increasing γ reduces V_c . For any increasing mean density, $\int_{1-s}^1 m_\alpha d\alpha > s \int_0^1 m_\alpha d\alpha$ thus $c > 1 - s$, and vice versa for a decreasing mean density. Hence when capital is set based on expected shortfall, the capital VaR threshold varies with the shape of the loss distribution. Higher skewness requires a higher VaR threshold and reduces shortfall probability. On the other hand a higher shortfall probability is appropriate for a less skewed loss distribution.

Extend (7) by replacing expectations with risk-adjusted expectations. Then c is solved from

$$\frac{\mathcal{E}(L_1 - L_c)}{\mathcal{E}(L_1 - L_0)} = \frac{\int_c^1 (m_\alpha + r_\alpha) d\alpha}{\int_0^1 (m_\alpha + r_\alpha) d\alpha} = \frac{\int_c^1 m_\alpha (1 + r_\alpha^*) d\alpha}{\int_0^1 m_\alpha (1 + r_\alpha^*) d\alpha} = s. \quad (8)$$

Since r_α^* is increasing in α , $m_\alpha(1 + r_\alpha^*)$ increases at a relatively faster rate than m_α . For example refer to the exponential loss distribution in Figure 1. The mean density is flat but increasing after risk adjustment. Hence c from (8) exceeds c from (7), using the above argument. Risk-adjusted expected shortfall increases relative to overall risk-adjusted expected loss, resulting in higher c .

The scale of the loss distribution does not affect c in (7) and (8), since scale factors are present in numerator and denominator. Scale is reflected when V_c is computed from c . Also note that a consequence of expressing loss layers as VaR layers is that derived quantities are in VaR terms rather than in dollars. This argues for VaR layers given the prevalence of VaR references in practice.

14.2. Balancing expected shortfall and surplus

The following derives capital by considering expected capital surplus in addition to shortfall. Considering capital surplus recognises opportunity costs of holding capital. All else equal, a higher opportunity cost reduces capital held.

Suppose an opportunity cost i is attached to every dollar of capital surplus. For example i may be the expected return on alternative investments. Per dollar cost of capital shortfall is j , for example the borrowing cost when in distress. Given capital V_c , the expected total cost of capital shortfall and surplus expressed in VaR-layers and the mean density is

$$\begin{aligned} \Gamma &= iE(V_c - L_c) + jE(L_1 - L_c) = i \left(V_c - \int_0^c m_\alpha d\alpha \right) + j \int_c^1 m_\alpha d\alpha \\ &= iE\{\max(V_c - x, 0)\} + jE\{\max(x - V_c, 0)\}, \end{aligned}$$

where the final expression uses standard statistical notation and is shown for comparison. Setting the derivative of Γ with respect to c to 0 yields

$$i(V'_c - m_c) - jm_c = 0, \quad ic - j(1 - c) = 0, \quad c = \frac{j}{i + j}.$$

The middle expression is derived from the first by dividing by V_c . Substituting $c = j/(i + j)$ into the second derivative of Γ confirms a minimum.

Therefore the capital VaR threshold in this setup is the unit shortfall cost j relative to the total unit cost $i + j$. If $i = j$ then $c = 0.5$ and capital is held at the median. Higher j relative to i increases c and V_c . In addition any capital V_c implies relative shortfall cost $j/i = c/(1 - c)$. For example, capital $V_{0.99}$ implies $j/i = 99$, thus the unit cost of shortfall is implicitly 99 times the unit cost of surplus. Higher c implies higher j relative to i .

Again introducing risk-adjustment yields a revised expected total cost

$$\Gamma = i \left\{ V_c - \int_0^c (m_\alpha + r_\alpha) d\alpha \right\} + j \int_c^1 (m_\alpha + r_\alpha) d\alpha .$$

Setting the derivative $V'_c[i\Phi(c) + j\{1 - \Phi(c)\}] = 0$ yields c as the solution to

$$\Phi(c) = \frac{j}{i + j} .$$

The solution $c > j/(i + j)$ since Φ is convex or $\Phi(c) \leq c$. Risk adjustment hence increases capital since expected shortfall and its cost are magnified relative to surplus. In addition c increases as distortion increases, and reduces to $j/(i + j)$ when there is no distortion. For example the distortion operator $\Phi(c) = c^n$ where $n \geq 1$ yields $c = \{j/(i + j)\}^{1/n}$ which is increasing in n and approaches 1 as $n \rightarrow \infty$.

The capital VaR threshold c derived by minimising expected cost of shortfall and surplus is independent of the loss distribution. In contrast the first approach where expected shortfall is set as a proportion of expected loss yields a value of c varying with the shape of the loss distribution. In both cases derived capital is expressed directly in VaR terms instead of dollars.

15. Loss transformation using reinsurance

The following illustrates how reinsurance transforms the probability, mean and risk profiles of a random loss. Reinsurance can also be structured to deliver a desired probability, mean or risk profile.

Suppose an insurer faces loss x , and enters into a reinsurance arrangement where a portion $0 \leq t_\alpha \leq 1$ of every V_α -layer is reinsured and the remaining portion $1 - t_\alpha$ is retained. Hence V_α -layer and its mean and risk all reduce by t_α . The retained loss is formed by combining retained VaR-layers

$$\tilde{x} \equiv \int_0^1 (1 - t_\alpha) dL_\alpha = \int_0^1 (1 - t_\alpha) I_\alpha(u) V'_\alpha d\alpha \quad (9)$$

and the corresponding mean and risk densities are, respectively,

$$\tilde{m}_\alpha \equiv (1 - t_\alpha) m_\alpha, \quad \tilde{r}_\alpha \equiv (1 - t_\alpha) r_\alpha . \quad (10)$$

Risk ratios are unchanged after reinsurance: $\tilde{r}_\alpha/\tilde{m}_\alpha = r_\alpha/m_\alpha$. The loss distribution is altered by reinsurance and is further discussed below.

The set of values of t_α over $0 \leq \alpha \leq 1$ defines the “reinsurance structure.” Typical reinsurance structures are “quota share”: $t_\alpha = t$ and “excess-of-loss”: $t_\alpha = I_d(\alpha)$. Quota share covers a constant proportion t of all layers. Excess-of-loss covers all VaR-layers above d . A combination of quota share and excess-of-loss reinsurance can also be applied to deliver a desired loss distribution as shown in the example at the end of this section.

The overall risk ratio after reinsurance is

$$\frac{\int_0^1 \tilde{r}_\alpha d\alpha}{\int_0^1 \tilde{m}_\alpha d\alpha} = \frac{\int_0^1 (1 - t_\alpha) m_\alpha r_\alpha^* d\alpha}{\int_0^1 (1 - t_\alpha) m_\alpha d\alpha}, \quad r_\alpha^* = \frac{r_\alpha}{m_\alpha}.$$

Compare the overall risk ratio after reinsurance with the same prior to reinsurance in (6). To reduce the overall risk ratio, reinsure high VaR layers with high r_α^* , particularly if m_α is high. This observation implies excess-of-loss is more risk effective than quota share. Quota share maintains the overall risk ratio since $t_\alpha = t$ for all α , whereas the overall risk ratio with excess-of-loss is

$$\frac{\int_0^d r_\alpha d\alpha}{\int_0^d m_\alpha d\alpha} = \frac{\int_0^d m_\alpha r_\alpha^* d\alpha}{\int_0^d m_\alpha d\alpha},$$

a weighted average of risk ratios up to the V_d -layer.

The VaR $_\alpha$ -spacing after reinsurance is $\tilde{V}'_\alpha \equiv (1 - t_\alpha)V'_\alpha$. Therefore integrate this expression to construct the VaR $_\alpha$ of the retained loss:

$$\tilde{V}_\alpha = \int_0^\alpha (1 - t_s) V'_s ds = (1 - t_\alpha) V_\alpha + \int_0^\alpha V_s t'_s ds,$$

where the second expression follows from integration by parts. For quota share, $t'_\alpha = 0$ therefore $\tilde{V}_\alpha = (1 - t)V_\alpha$. Losses reduce by a portion t . For excess-of-loss, $t'_\alpha = (\alpha = d)$, the Dirac delta function, and $\tilde{V}_\alpha = \{1 - I_d(\alpha)\}V_\alpha + I_d(\alpha)V_d$. Hence all losses below V_d are retained and losses above V_d are capped at V_d , implying the retained loss is the VaR-layer $[0, V_d]$ of x .

The following forms a reinsurance structure to deliver a desired probability, mean or risk profile of the retained loss. If the desired mean density is \tilde{m}_α then from (10) the reinsurance structure is $t_\alpha = 1 - \tilde{m}_\alpha/m_\alpha$. Similarly for a desired risk density. If the desired loss distribution has VaR $_\alpha$ \tilde{V}_α then $t_\alpha = 1 - \tilde{V}'_\alpha/V'_\alpha$. In all cases the reinsurance structure computes the ratio between desired and original quantities.

Consider a Pareto loss distribution with shape parameter $\gamma > 1$ and scale parameter $b > 0$. A reinsurance structure is required to transform the loss distribution into an exponential with scale $\tilde{b} > 0$. The reinsurance structure is

$$t_\alpha = 1 - \tilde{V}'_\alpha/V'_\alpha = 1 - \frac{\tilde{b}(1 - \alpha)^{-1}}{b\gamma^{-1}(1 - \alpha)^{-(1/\gamma+1)}} = 1 - \tilde{b}b^{-1}\gamma(1 - \alpha)^{1/\gamma}.$$

The reinsured portion is $1 - \tilde{b}b^{-1}\gamma$ at the zero VaR layer, and increases monotonically to 1 at the highest VaR layer. This result is intuitive since the Pareto is thicker tailed than the exponential. Hence in order to transform the loss from Pareto to exponential, minimum reinsurance is required for small losses, and larger losses are increasingly reinsured. Note $\tilde{b} \leq b/\gamma$ since $0 \leq t_\alpha \leq 1$ for all α .

16. Conclusion

This paper develops a framework to analyze mean and risk contributions by VaR-layers of a loss. Expressing layers using VaRs establishes connections with the standard use of VaRs in insurance and finance.

Applying mean and risk densities yields insights to common risk management problems involving capital, reinsurance and derivatives.

References

- Acerbi, C. (2002). Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking & Finance* 26(7), 1505–1518.
- Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). COHERENT MEASURES OF RISK. *Mathematical Finance* 9(3), 203–228.
- Campana, A. and P. Ferretti (2014). Risk measures in solvency regulation: Reinsurance layers and unexpected loss. *Applied Mathematical Sciences* 8(116), 5783–5794.
- Chernobai, A. S., S. T. Rachev, and F. J. Fabozzi (2008). *Operational risk: a guide to Basel II capital requirements, models, and analysis*, Volume 180. John Wiley & Sons.
- Choo, W. and P. De Jong (2009). Loss reserving using loss aversion functions. *Insurance: Mathematics and Economics* 45(2), 271–277.
- Das, A. and R. M. Stein (2011). Differences in tranching methods: Some results and implications. *New York: Moodys Research Labs*.
- Duffie, D. and N. Garleanu (2001). Risk and valuation of collateralized debt obligations. *Financial Analysts Journal* 57(1), 41–59.
- Eling, M., H. Schmeiser, and J. Schmit (2007). The solvency ii process: Overview and critical analysis. *Risk Management and Insurance Review* 10(1), 69.
- Evans, J. (2001). Exposure rating loss layers: Unifying the property perspective of severity with the liability perspective of frequency. *Including the Reserves Call Papers*, 493.
- Hogg, R. V. and S. A. Klugman (2009). *Loss distributions*, Volume 249. John Wiley & Sons.
- Kolb, R. (2010). *Lessons from the financial crisis: Causes, consequences, and our economic future*, Volume 12. John Wiley & Sons.
- Lee, Y.-S. (1988). The mathematics of excess of loss coverages and retrospective rating—a graphical approach. *PCAS LXXV* 49.
- Mandel, B., D. Morgan, and C. Wei (2012). The role of bank credit enhancements in securitization. *Federal Reserve Bank of New York Economic Policy Review* 18(2), 35–46.
- Miccolis, R. S. (1977). On the theory of increased limits and excess of loss pricing. *PCAS LXIV*, 27.
- Salzmann, R. E. (1963). Rating by layer of insurance. *PC, 4S*, 15–26.

- Shaked, M. and J. G. Shanthikumar (2007). *Stochastic orders*. Springer Science & Business Media.
- Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics* 17(1), 43–54.
- Wang, S. (1996). Premium Calculation by Transforming the Premium Layer Density. *ASTIN Bulletin* 26(1), 71–92.
- Wang, S. (1998). An actuarial index of the right-tail risk. *North American Actuarial Journal* 2(2), 88–101.
- Wang, S. (2000). A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance* 67(1), 15–36.
- Wang, S., V. Young, and H. Panjer (1997). Axiomatic characterization of insurance prices. *Insurance Mathematics and Economics* 21(2), 173–183.

Chapter 4

Analyzing systematic risk and diversification

The following paper applies risk densities to analyze systematic risk and diversification when imperfectly dependent losses are aggregated.

Insights to systematic risk and diversification across a joint probability distribution

Abstract

This paper analyses and develops insights to systematic risk and diversification when random, imperfectly dependent, losses are aggregated. Systematic risk and diversification are shown to vary across layers of component losses according to local dependence and volatility structures. Systematic risk is high and diversification is weak overall if high risk layers are heavily dependent on the aggregate loss. This result explains weak diversification observed in financial markets despite weak to moderate correlations overall. A coherent risk setup is assumed in this paper, where risks are measured using distortion and allocated using the Euler principle.

Keywords: Distortion risk; spectral risk; Euler allocation; systematic risk; diversification; layer; Value-at-Risk.

1. Introduction to systematic risk and diversification

Suppose x is one of several continuous, non-negative and random component losses aggregating to x_+ . For example x may be the loss from an insurance class and x_+ is the loss aggregated across all classes. Or x may be the credit loss on a portfolio of loans and x_+ is the aggregate credit loss across all portfolios. Component losses are imperfectly dependent leading to risk diversification as formalised below.

This paper applies the following risk setup. Suppose $\phi \geq 0$ is an increasing risk aversion function integrating to 1. Standalone risk of x , systematic risk of x and aggregate risk of x_+ are respectively

$$r = \text{cov}\{x, \phi(u)\}, \quad \bar{r} = \text{cov}\{x, \phi(u_+)\}, \quad r_+ = \text{cov}\{x_+, \phi(u_+)\}, \quad (1)$$

where cov denotes covariance and u and u_+ are percentile ranks of x and x_+ . If F and F_+ are distribution functions of x and x_+ then $u \equiv F(x)$ and $u_+ \equiv F_+(x_+)$. Standalone and systematic risks of other component losses forming x_+ are calculated using similar covariance expressions.

The risk setup (1) is well established in the literature. Choo and De Jong (2009) refers to r and r_+ as loss aversion risks of x and x_+ and establishes their equivalence with distortion risks (Wang, 1996) and spectral risks (Acerbi, 2002).

Examples of ϕ leading to different risk measures are discussed in Choo and De Jong (2009). In addition \bar{r} is an allocation of r_+ to x by applying the Euler principle (McNeil et al., 2005) or game theory (Tasche, 2007). The allocation is extensively discussed in Choo and De Jong (2010), Buch and Dorfleitner (2008) and Tsanakas and Christofides (2006).

The risk setup (1) is coherent in the sense of Artzner et al. (1999), Denault (2001) and Kalkbrener (2005): positively homogenous, translation invariant, monotonic and subadditive. Further the allocation is complete: adding \bar{r} across component losses yields r_+ , and there is no undercut: $\bar{r} \leq r$ no matter how x is carved out from x_+ .

The term "systematic risk" is consistent with the capital asset pricing model in finance (Luenberger, 1998), where the volatility or risk of an asset return is divided into diversifiable and non-diversifiable or systematic risk. Systematic risk increases with the correlation between asset and market returns, and is a key driver of risk premiums: expected asset returns above market return.

This paper compares r , the risk of x before aggregation, with \bar{r} , the risk of x after aggregation. Since x and u_+ are imperfectly dependent, $\bar{r} \leq r$ and the difference $r - \bar{r}$ is a diversification benefit. Diversification benefits are critical in risk management by enabling risks to be managed viably and efficiently as a group, a classic case being insurance. The systematic risk ratio is

$$\theta \equiv \frac{\bar{r}}{r} = \frac{\text{cov}\{x, \phi(u_+)\}}{\text{cov}\{x, \phi(u)\}} = \frac{\text{cor}\{x, \phi(u_+)\}}{\text{cor}\{x, \phi(u)\}} \leq 1, \quad (2)$$

where cor represents correlation. A lower systematic risk ratio indicates greater diversification. The final expression holds since u and u_+ are both uniform hence $\phi(u)$ and $\phi(u_+)$ have equal standard deviations. Stronger dependence between x and x_+ leads to weaker diversification due to higher numerators in (2) hence larger θ . Conversely weaker dependence leads to stronger diversification.

In this paper, systematic risk and diversification benefit are shown to vary across the distribution of x based on the local dependence structure underlying x and x_+ . This result is important when formulating risk management strategies to maximise diversification, as demonstrated by a case study using historical stock returns. The analysis focusses on x , however equivalent results apply to other component losses forming x_+ .

The remaining paper is structured as follows. Section 2 gives an overview of mean and risk densities and Value-at-Risk (VaR) loss layers established in the second paper of this thesis (chapter 3). Section 3 defines systematic risk densities analogous to standalone risk densities, and establishes links to local dependence structures underlying component and aggregate losses. Section 4 uses systematic risk densities to explore drivers of systematic risk and diversification such as in financial markets. A theoretical example of the proposed analytical framework is shown in §5. Section 6 attributes aggregate mean and risk densities to component losses. The attribution is important when risk management strategies target the aggregate loss. Section 7 examines the case where

component losses are comonotonic, a benchmark for measuring diversification. Section 8 concludes with an illustration of the framework using historical stock returns.

2. Overview of VaR layers, mean and risk densities

This section outlines VaR layers, mean and risk densities developed in chapter 3 of this thesis. Mean and risk densities track the mean and standalone risk of a loss across its VaR layers. Risk densities are critical to the analysis of systematic risk and diversification as shown in subsequent sections.

The α -Value-at-Risk (VaR_α) of a continuous random loss $x \geq 0$ with distribution function F is the percentile $V_\alpha \equiv F^-(\alpha)$, where F^- is the inverse distribution function of x . In addition define $L_\alpha \equiv \min(x, V_\alpha)$, the loss capped at its VaR_α . Consider the following two expressions:

$$L_\beta - L_\alpha = \min\{\max(x - V_\alpha, 0), V_\beta - V_\alpha\}, \quad x = L_1 - L_0 = \int_0^1 dL_\alpha.$$

The first expression above is the VaR layer $[V_\alpha, V_\beta]$ of x : the excess of x over V_α with the excess capped at $V_\beta - V_\alpha$. The second expression decomposes x into infinitesimal VaR layers dL_α over $0 \leq \alpha \leq 1$. The infinitesimal V_α -layer of x is

$$dL_\alpha = L'_\alpha d\alpha = I_\alpha(u) V'_\alpha d\alpha,$$

where L'_α and V'_α are derivatives of L_α and V_α with respect to α , and $I_\alpha(u)$ is an indicator equal to 1 if $u > \alpha$ or equivalently $x > V_\alpha$, and 0 otherwise. Hence the V_α -layer of x is an increment proportional to V'_α if $x > V_\alpha$ and 0 otherwise.

Layers are standard insurance and financial constructs, and are also called tranches in finance. VaR layers self-adjust to the shape and scale of the loss distribution and are hence comparable across loss distributions. Refer to chapter 3 for further discussion of layers and reasons for defining layers using VaRs.

The risk density indicates the standalone risk of infinitesimal VaR layers of x , and based on (1) is the covariance

$$r_\alpha = \text{cov}\{L'_\alpha, \phi(u)\} = \text{cov}\{I_\alpha(u) V'_\alpha, \phi(u)\} = \{\alpha - \Phi(\alpha)\} V'_\alpha,$$

where

$$\Phi(\alpha) \equiv \int_0^\alpha \phi(u) du$$

cumulates ϕ . Integrating r_α yields standalone risks of larger layers:

$$\int_\alpha^\beta r_\pi d\pi = \text{cov}\left\{\int_\alpha^\beta L'_\pi d\pi, \phi(u)\right\} = \text{cov}\{L_\beta - L_\alpha, \phi(u)\},$$

and the entire area under r_α is the overall standalone risk of x : $\int_0^1 r_\alpha d\alpha = r$.

The mean density indicates the mean of infinitesimal VaR layers and is

$$m_\alpha = E(L'_\alpha) = E\{I_\alpha(u)V'_\alpha\} = (1 - \alpha)V'_\alpha.$$

Similar to r_α , integrating m_α yields the mean of larger layers. Hence mean and risk densities are akin to probability densities: integrating yields quantities over larger regions. Refer to chapter 3 for further properties, illustrations and applications of mean and risk densities.

3. Systematic risk densities and links to local dependence

Systematic risk densities are akin to standalone risk densities in §2, and describe systematic risks of infinitesimal VaR layers forming a random loss. Comparing systematic and standalone risk densities reveals the extent of diversification at various layers of the loss distribution. The comparison also leads to the dependence structure involving component and aggregate losses.

The systematic risk density of x according to (1) is the covariance

$$\bar{r}_\alpha \equiv \text{cov}\{L'_\alpha, \phi(u_+)\} = \text{cov}\{I_\alpha(u), \phi(u_+)\}V'_\alpha.$$

Similar to r_α , integrating \bar{r}_α yields systematic risks of larger layers and the overall systematic risk of x :

$$\int_\alpha^\beta \bar{r}_\pi d\pi = \text{cov}\{L_\beta - L_\alpha, \phi(u_+)\}, \quad \int_0^1 \bar{r}_\alpha d\alpha = \text{cov}\{L_1, \phi(u_+)\} = \bar{r}.$$

The ratio between systematic and standalone risk densities is the proportion of remaining risk in each VaR layer after diversification, and is given by

$$\theta_\alpha \equiv \frac{\bar{r}_\alpha}{r_\alpha} = \frac{\text{cov}\{I_\alpha(u)V'_\alpha, \phi(u_+)\}}{\text{cov}\{I_\alpha(u)V'_\alpha, \phi(u)\}} = \frac{\text{cov}\{I_\alpha(u), \phi(u_+)\}}{\text{cov}\{I_\alpha(u), \phi(u)\}}, \quad (3)$$

where the last expression follows since the constant V'_α in \bar{r}_α and r_α cancel with division. If x and x_+ are comonotonic or $u = u_+$ then $\theta_\alpha = 1$ for all $0 \leq \alpha \leq 1$. Otherwise $\theta_\alpha < 1$.

For any α , large θ_α implies weak diversification at V_α -layer of x since \bar{r}_α is close to r_α and the diversification benefit $r_\alpha - \bar{r}_\alpha$ is small relative to r_α . Vice versa diversification at V_α -layer is strong if θ_α is small or even negative. An analysis of θ_α against α hence reveals layers with weak diversification, which can be mitigated using reinsurance or hedging, and layers with strong diversification which should be preserved since they reduce overall risk. In addition θ_α varies over α in line with local dependence between x and x_+ . The notion is explored in the rest of this section.

The systematic risk ratio θ_α is calculated entirely from the joint distribution of (u, u_+) or equivalently the copula C of (x, x_+) . Marginal distributions are

not directly involved in θ_α , although the marginal distribution of x affects C^1 . The denominator of the last term in (3) is $\alpha - \Phi(\alpha)$ and the numerator is

$$\int_0^1 \text{cov}\{I_\alpha(u), I_\beta(u_+)\} d\phi(\beta) = \int_0^1 \{C(\alpha, \beta) - \alpha\beta\} \phi'(\beta) d\beta,$$

a partial integration of the copula C weighted by the derivative ϕ' .

The set of θ_α values over $0 \leq \alpha \leq 1$ reflects and characterises the dependence structure of (x, x_+) . Given α , θ_α is the local dependence between V_α -layer of x and a function of the aggregate loss. Of interest is rank dependence since θ_α is defined in terms of percentile ranks u and u_+ . If (x, x_+) exhibits strong upper tail dependence and weak lower tail dependence, then θ_α starts small at $\alpha = 0$ and increases to one as α approaches one. An illustration of θ_α given a particular ϕ is shown in §5.

Systematic risk ratios have similar properties as linear and Spearman's correlation which are measures of overall dependence. For all $0 \leq \alpha \leq 1$, $\theta_\alpha \leq 1$, $\theta_\alpha = 1$ if $u = u_+$ and $\theta_\alpha = 0$ if u and u_+ are independent. Negative dependence yields $\theta_\alpha \leq 0$, and $u_+ = 1 - u$ leads to $\theta_\alpha = -1$ if $\phi(u)$ is symmetric about $u = 0.5$ such that $\phi(u - 0.5) = \phi(0.5 - u)$. In addition stronger correlation order (Dhaene et al., 2009) between x and x_+ increases θ_α for all α . Proofs are straightforward from (3).

Systematic risk ratios are similar to layer dependence (discussed in chapter 2 of this thesis) which measures local dependence between arbitrary uniform random variables u and v :

$$\ell_\alpha \equiv \frac{\text{cov}\{I_\alpha(u), v\}}{\text{cov}\{I_\alpha(u), u\}}.$$

Layer dependence ℓ_α assumes a linear ϕ in systematic risk ratio θ_α .

4. Explaining and exploring systematic risk and diversification

High overall systematic risk ratio in (2), or low overall diversification, arises when strong local dependence coincides with high local skewness or volatility, such as in financial markets. High systematic risk or low diversification overall may arise even when overall correlation is weak to moderate. These intuitive concepts are formalised and explored below.

Noting $\bar{r}_\alpha = \theta_\alpha r_\alpha$ from (3), overall systematic risk and risk ratio of x defined in (1) and (2) are respectively written as

$$\bar{r} = \int_0^1 r_\alpha \theta_\alpha d\alpha, \quad \theta = \frac{\bar{r}}{r} = \frac{\int_0^1 r_\alpha \theta_\alpha d\alpha}{\int_0^1 r_\alpha d\alpha}. \quad (4)$$

¹For example suppose x dominates other losses forming x_+ . Then x and x_+ are strongly dependent compared to the case where x is dominated by other losses.

Similar expressions apply when considering systematic risk over $[V_\alpha, V_\beta]$ VaR-layer of x : change integration limits in (4) from $[0, 1]$ to $[\alpha, \beta]$.

From (4), the overall systematic risk ratio of x is a “risk weighted” average of individual systematic risk ratios across its VaR layers. As mentioned in the previous section, individual systematic risk ratios characterise the dependence structure of (x, x_+) . Risk weights are formed by the standalone risk density $r_\alpha = \{\alpha - \Phi(\alpha)\}V'_\alpha$, and are large for a loss distribution if V'_α is large. The derivative V'_α measures local volatility since it is the relative gap between successive VaRs. Suppose a sample of infinite size is drawn from the loss distribution and hence ordered observations are close to VaRs. Large V'_α implies observations around V_α are highly spread out, or volatile locally.

The result in (4) explains high systematic risk ratios and low diversification observed in financial markets, for example during the 2008 global financial crisis (Kolb, 2010), despite moderate correlations across time. From (4), large θ arises when strong local dependence θ_α coincides with large risk weights or local volatility reflected in r_α . Financial returns are heavy tailed (Cont (2001), Hsieh (1988)) and exhibit strong tail dependence (Rodriguez (2007), Hartmann et al. (2004)). The former implies large r_α and the latter implies large θ_α , both over large α . These two implications combine to create large θ , even though θ_α may be small across most other smaller values of α . Hence risk weighting of individual systematic risk ratios leads to diversification being heavily influenced by local dependence in volatile or risky layers.

Conversely diversification is strong when risks of x are concentrated in layers which are weakly dependent on the aggregate loss. Large r does not necessarily imply x is undesirable if other component losses are structured to yield low θ and \bar{r} . For example suppose x is heavy-tailed, such as Pareto, yielding a large standalone risk concentrated in high layers. Constructing other component losses so that they have favourable values when x is large leads to high layers of x being weakly dependent on the aggregate loss. As a result most of the risk of x is diversified upon aggregation, and x becomes acceptable or even desirable if it attracts a large risk premium based on its standalone risk.

5. Theoretical example using the conditional–tail–expectation

Assume a stepped risk aversion function $\phi(u) = I_t(u)/(1-t)$ where $0 \leq t \leq 1$ is a parameter. Then according to Choo and De Jong (2009) and Choo and De Jong (2010),

$$r = E(x|u > t) - E(x), \quad \bar{r} = E(x|u_+ > t) - E(x), \quad \theta = \frac{E(x|u_+ > t) - E(x)}{E(x|u > t) - E(x)}.$$

Risk in this case is the gap between the conditional–tail–expectation (Rockafellar and Uryasev, 2002) of x and its unconditional expectation. Standalone risk considers the tail event $u > t$ of x whereas systematic risk considers the

aggregate tail event $u_+ > t$. The overall systematic risk ratio θ of x captures the relative impact of the two tail events.

The systematic risk ratio at V_α -layer using (3) is

$$\begin{aligned}\theta_\alpha &= \frac{\text{cov}\{I_\alpha(u), I_t(u_+)/(1-t)\}}{\text{cov}\{I_\alpha(u), I_t(u)/(1-t)\}} = \frac{\text{cov}\{I_\alpha(u), I_t(u_+)\}}{\text{cov}\{I_\alpha(u), I_t(u)\}} \\ &= \frac{P(u > \alpha, u_+ > t) - P(u > \alpha)P(u_+ > t)}{P\{u > \max(\alpha, t)\} - P(u > \alpha)P(u > t)} = \frac{P(u > \alpha|u_+ > t) - P(u > \alpha)}{P(u > \alpha|u > t) - P(u > \alpha)},\end{aligned}$$

where P calculates probability. Given α , θ_α increases with the probability of $u > \alpha$ jointly with, or conditional on, $u_+ > t$. Other above probabilities are marginal probabilities, and are scaling factors such that $\theta_\alpha = 1$ when $u = u_+$ and $\theta_\alpha = 0$ when u and u_+ are independent. In addition results are unchanged when inequalities are reversed. Therefore individual systematic risk ratios in this example reflect the likelihood of joint tail events in x and x_+ .

For any α , $\theta_\alpha = 1$ implies $P(u > \alpha|u_+ > t) = P(u > \alpha|u > t) = \min\{1, (1 - \alpha)/(1 - t)\}$, whilst $\theta_\alpha = 0$ implies $P(u > \alpha|u_+ > t) = P(u > \alpha) = 1 - \alpha$. The former is the maximum conditional tail probability and the latter is the unconditional tail probability assuming independence locally.

An alternative expression for θ_α in terms of the copula C of (x, x_+) in this example is

$$\theta_\alpha = \frac{C(\alpha, t) - \alpha t}{\min(\alpha, t) - \alpha t}.$$

Suppose θ_α is specified for all $0 \leq \alpha \leq 1$ and for all thresholds $0 \leq t \leq 1$. Then C is completely specified. Hence the set of individual systematic risk ratios over all parameters completely characterises the dependence structure of (x, x_+) .

The panels in Figure 1 plot θ_α against α for $t = 0.5, 0.75$ and 0.9 , assuming Gumbel and Clayton copulas for (u, u_+) . Apart from a kink at $\alpha = t$, θ_α traces local dependence between u and u_+ , indicated by the extent of clustering along the 45° line. For the Gumbel copula, systematic risk ratios are higher at both tails due to tail dependence, and lower in the middle. For the Clayton copula, systematic risk ratios are higher at the left tail and lower at the right tail, characterising lower tail dependence.

6. Sub-aggregate mean and risk densities

Write the mean and risk densities of aggregate loss x_+ as respectively

$$m_{\beta,+} = (1 - \beta)V'_{\beta,+}, \quad r_{\beta,+} = \{\beta - \Phi(\beta)\}V'_{\beta,+}$$

where $V_{\beta,+}$ is the VaR_β of x_+ and $V'_{\beta,+}$ is the derivative of $V_{\beta,+}$ with respect to β . Aggregate risk reflects diversification and is therefore composed entirely of systematic risks of component losses forming x_+ . In addition the VaR_β -layer of x_+ is $I_\beta(u_+)V'_{\beta,+}d\beta$ where u_+ is the percentile rank of x_+ .

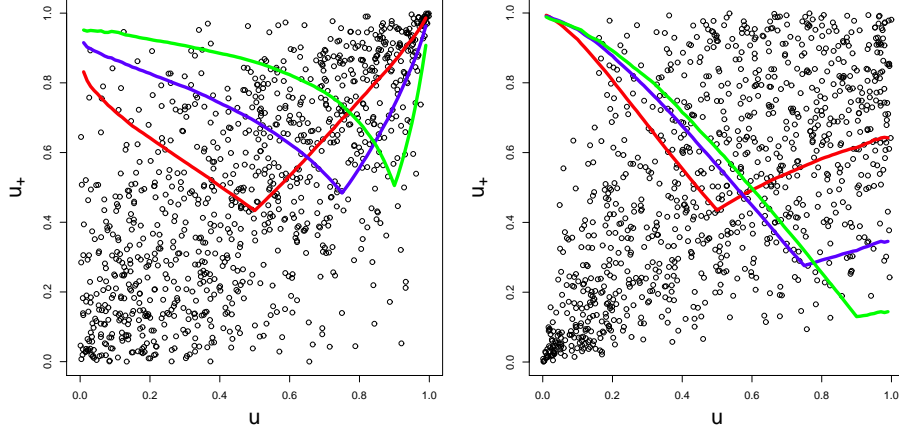


Figure 1: Left and right panels plot θ_α against α for Gumbel and Clayton copulas respectively. θ_α is computed assuming a risk aversion function $\phi(u) = (u > t)/(1-t)$ for $t = 0.5$ (red), 0.75 (blue) and 0.9 (green). The calculation is performed over 100 intervals of α forming $[0, 1]$. Note how θ_α tracks local dependence (clustering along the 45° line) in both copulas, apart from the kinks at $\alpha = t$.

The following allocates aggregate mean and risk densities and aggregate VaR layers to component losses including x . The allocation is critical when risk management strategies are applied to layers of x_+ , and the impact and cost of the strategies are attributed to component losses. Example of such strategies are stop-loss reinsurance and aggregate hedging of a portfolio of losses. The proposed allocation shown below is unbiased and aligns with the overall mean and systematic risk of component losses. The approach relies on conditional mean sharing by Denuit and Dhaene (2012).

The aggregate loss is $x_+ = \sum_i x_i$ where x_i represents a component loss such as x . Using iterated expectations, write x_+ as

$$x_+ = \sum_i g_i(x_+), \quad g_i(x_+) \equiv E(x_i | x_+),$$

where g_i calculates the mean value of x_i conditional on x_+ . Denuit and Dhaene (2012) allocates $g_i(x_+)$ to x_i , called conditional mean sharing. The allocation is unbiased since $E\{g_i(x_+)\} = E\{E(x_i | x_+)\} = E(x_i)$. Substituting $x_+ = V_{\beta,+}$ into the above result and taking derivatives with respect to β yields

$$V'_{\beta,+} = \sum_i \{g_i(V_{\beta,+})\}' = V'_{\beta,+} \sum_i g'_i(V_{\beta,+}), \quad \sum_i g'_i(V_{\beta,+}) = 1. \quad (5)$$

(5) allocates a fraction $g'_i(V_{\beta,+})$ of $V'_{\beta,+}$ to x_i , and fractions sum to one across i . Apply this fractional allocation to aggregate VaR layers and aggregate mean and risk densities, since all are proportional to $V'_{\beta,+}$. Hence the allocated or

“sub-aggregate” mean and risk densities of x are respectively

$$\dot{m}_{\beta,+} \equiv m_{\beta,+}g'(V_{\beta,+}), \quad \dot{r}_{\beta,+} \equiv r_{\beta,+}g'(V_{\beta,+}),$$

where $g(x_+) = E(x|x_+)$. Identical densities are obtained by performing ground-up calculations on $g(x_+)$ assuming g is increasing and VaR_β of $g(x_+)$ is $g(V_{\beta,+})$. Given β , $\dot{m}_{\beta,+}$ and $\dot{r}_{\beta,+}$ are portions of the mean and risk of VaR_β -layer of x_+ attributable to x .

Sub-aggregate mean and risk densities of x are aligned with the overall mean and systematic risk of x . Integrating $\dot{m}_{\beta,+}$ over all β yields

$$\int_0^1 \dot{m}_{\beta,+} d\beta = \int_0^1 (1-\beta)\{g(V_{\beta,+})\}' d\beta = E\{g(x_+)\} = E(x),$$

and similarly

$$\begin{aligned} \int_0^1 \dot{r}_{\beta,+} d\beta &= \int_0^1 \text{cov}\{I_\beta(u_+), \phi(u_+)\}\{g(V_{\beta,+})\}' d\beta \\ &= \text{cov}\{g(x_+), \phi(u_+)\} = \text{cov}\{x, \phi(u_+)\} = \bar{r}. \end{aligned}$$

Integrating $\dot{m}_{\beta,+}$ and $\dot{r}_{\beta,+}$ over a subset of the unit interval yields an allocation of the mean and risk of the corresponding VaR layer of x_+ to x .

Sub-aggregate mean and risk densities are different from mean and systematic risk densities: $\dot{m}_{\beta,+} \neq m_\alpha$ and $\dot{r}_\beta \neq \bar{r}_\alpha$ for $\alpha \neq \beta$, although they integrate to the same result. The former relates to VaR layers of x_+ whereas the latter relates to VaR layers of x . However equality applies when component losses are comonotonic. This special case is discussed in the next section.

7. Comonotonicity as a diversification benchmark

Comonotonic x and x_+ implies $u = u_+$: x and x_+ are always at equal percentiles, and $E(x|x_+) = x$: x_+ pinpoints the value of x . Dhaene et al. (2002) and Wang and Dhaene (1998) further discuss the concept of comonotonicity. Comonotonicity yields maximum systematic risk across all layers and is hence a benchmark for assessing diversification as shown below.

Comonotonicity between x and x_+ implies:

- Systematic and standalone risk densities of x are equal, and there is no diversification at any VaR layer of x . Since $u = u_+$,

$$\bar{r}_\alpha = \text{cov}\{I_\alpha(u), \phi(u_+)\}V'_\alpha = \text{cov}\{I_\alpha(u), \phi(u)\}V'_\alpha = r_\alpha,$$

implying $\theta_\alpha = 1$. In addition $\bar{r} = r$ and $\theta = 1$. This result is discussed in section 3. Hence \bar{r}_α is maximised across all α under comonotonicity.

- Sub-aggregate densities are also equal to standalone densities. Since $x = E(x|x_+) = g(x_+)$, $V_\beta = g(V_{\beta,+})$, hence $V'_\beta = \{g(V_{\beta,+})\}'$ and

$$\dot{m}_{\beta,+} = m_\beta, \quad \dot{r}_{\beta,+} = r_\beta.$$

Since comonotonic x_+ and x reach their VaR_β -layers simultaneously, the attributed mean and risk of the VaR_β -layer of x_+ to x is the mean and risk (systematic or standalone) of the VaR_β -layer of x .

The standalone risk density of x thus indicates maximum systematic risk across all layers, and comparing it with systematic and sub-aggregate risk densities of x reveals the extent of diversification in each layer. Section 3 already compares \bar{r}_α with r_α and the ratio characterises the dependence structure of (x, x_+) . The following shows a similar interpretation when $\dot{r}_{\beta,+}$ is compared with r_β . Define the ratio

$$\begin{aligned} \gamma_\beta &\equiv \frac{\dot{r}_{\beta,+}}{r_\beta} = \frac{\{\beta - \Phi(\beta)\}\{g(V_{\beta,+})\}'}{\{\beta - \Phi(\beta)\}V'_\beta} = \frac{\{g(V_{\beta,+})\}'}{V'_\beta} = \frac{g'(V_{\beta,+})V'_{\beta,+}}{V'_\beta} \\ &= \frac{\frac{d}{d\beta}E(x|u_+ = \beta)}{\frac{d}{d\beta}E(x|u = \beta)} = \frac{\text{cov}\left\{x, \frac{d}{d\beta}\delta_\beta(u_+)\right\}}{\text{cov}\left\{x, \frac{d}{d\beta}\delta_\beta(u)\right\}}. \end{aligned} \quad (6)$$

where $\delta_\beta(u)$ is the Dirac delta function which approaches ∞ when $u = \beta$ and is 0 otherwise. Similar to θ_α , $\gamma_\beta = 1$ if $u = u_+$ and $\gamma_\beta = 0$ if u and u_+ are independent. Unlike θ_α , γ_β is computed entirely from the joint distribution of (x, x_+) and does not involve the risk aversion function ϕ . In addition γ_β may exceed one since the numerator in its definition relates to the VaR_β -layer of x_+ whereas the denominator relates to the VaR_β -layer of x and there is no strict inequality between numerator and denominator.

As per θ_α , γ_β measures local dependence between x and x_+ but in a different manner. The second last expression in (6) measures the sensitivity of the conditional expectation of x to changes in x_+ at VaR_β , relative to the same calculation when $u = u_+$. The last expression in (6) computes the covariance between x and a function of u_+ , again relative to the comonotonic case.

Figure 2 plots γ_β against β using copulas in Figure 1. Assume exponential and normal distributions for x . Note γ_β is independent of location and scale of x . Calculations show $\gamma_\beta > 1$ over some values of β hence γ_β is scaled by its maximum value over β so that resulting values do not exceed 1. Similar to θ_α , γ_β traces the local dependence structure of (x, x_+) : the dispersion between scatter points along the 45° line. However calculated values of γ_β are more volatile than θ_α as the former involve conditional expectations over narrow windows whereas the latter use conditional tail expectations.

8. Case study using historical stock returns

This section applies the proposed systematic risk and diversification analytical framework to daily NASDAQ, S&P and FTSE returns between 1985 and

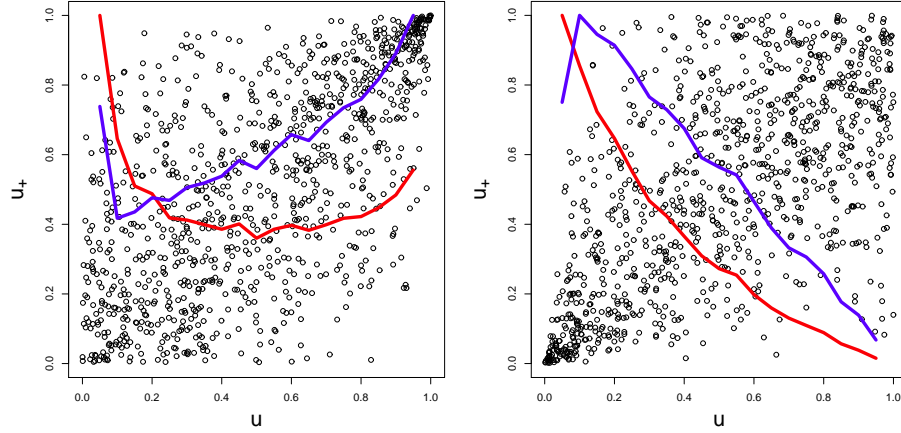


Figure 2: Left and right panels plot $\gamma_\beta / \max_\beta(\gamma_\beta)$ against β for Gumbel and Clayton copulas respectively. γ_β is computed assuming x is exponential (red) and normal (blue). The calculation is performed over 20 intervals of β forming $[0, 1]$ to produce smoother curves. Note how the curves trace the dependence structure of (u, u_+) represented by the scatter of points along the 45° line.

2015. Form a hypothetical portfolio of \$100 in each market index, and assume the empirical joint probability distribution. To be consistent with the proposed framework, focus on investment losses rather than gains. Hence switch the signs of investment returns, with gains being negative losses. Use the risk aversion function $\phi(\alpha) = 4I_{0.75}(\alpha)$, hence risk is the mean loss above $\text{VaR}_{0.75}$ compared to the overall mean.

Top four panels in Figure 3 display empirical marginal probability distributions and copulas of hypothetical index losses. NASDAQ losses are the most skewed, while S&P losses are the most peaked. In addition NASDAQ and S&P losses are highly dependent, and both are less dependent overall on FTSE losses (presumably due to different geographical location). All three market indices exhibit significant upper and lower tail dependence: extreme losses and gains are highly dependent across markets.

Bottom two panels in Figure 3 graph standalone and systematic risk densities for each market index, calculated from the empirical joint probability distribution. Density values are smoothed to reduce volatility across VaR layers. Before aggregation, NASDAQ has the highest risk density due to greater skewness of its probability distribution. S&P and FTSE have similar risk densities, although S&P has a slightly lower density in middle VaR layers due to greater peakedness of its probability density. All three risk densities are reduced after diversification, most notably FTSE. Subsequent figures compare and analyse standalone and systematic risk densities.

Figure 4 compares standalone and systematic risk densities for each market

index. Systematic risk ratios, and copulas between index and aggregate losses, are also shown. Note from §3 that systematic risk ratios are computed entirely from the copulas and summarise the dependence structure. Also note the kink in systematic risk ratios at $\text{VaR}_{0.75}$ -layers due to the selection of the risk aversion function. FTSE has lower systematic risk ratios than NASDAQ and S&P across all VaR layers due to its weaker dependence with the aggregate loss. Hence FTSE experiences stronger diversification. All three market indices have systematic risk ratios close to 1 in extreme VaR layers, due to tail dependence. Therefore there is minimal risk diversification in extreme VaR layers of all three indices.

Table 1 shows overall risks for each market index, before and after diversification. Overall systematic risk ratios are also shown. As expected NASDAQ has the highest standalone and systematic risk. S&P and FTSE have similar standalone risks, although FTSE has higher diversification and therefore lower systematic risk. NASDAQ and S&P have similar overall systematic risk ratios.

Portfolio	r	\bar{r}	θ
NASDAQ	\$2.00	\$1.86	0.93
S&P	\$1.32	\$1.22	0.92
FTSE	\$1.31	\$0.87	0.66
Overall	\$4.63	\$3.95	0.85

Table 1: Overall standalone risk, systematic risk and risk ratio for each market index.

Figure 5 shows mean and risk densities of the aggregate portfolio loss and its breakdown into sub-aggregate densities. NASDAQ generally has a higher sub-aggregate mean and risk density, notably in high VaR layers of the aggregate loss, due to its higher skewness and systematic risk contribution. Note from §6 that sub-aggregate mean and risk densities integrate to the overall mean and systematic risk of each market index.

Figures 3, 4 and 5 guide the formulation of effective risk management strategies. For example standalone and systematic risks are concentrated in high VaR layers and can be significantly reduced with put options. In addition, from Figure 4, dependence is strong at high VaR layers and weaker at lower VaR layers, hence diversification is improved when high VaR layers are eliminated. Consider two scenarios: put options on each market index and an aggregate put option on the portfolio. Assume an exercise price of $\text{VaR}_{0.95}$ of the loss referenced in the put option. Risks after put options are areas under risk densities up to the exercise price. Summary results of the first scenario are shown in table 2 in a similar format as table 1. Under the second scenario, aggregate risk after diversification is 3.55 hence the first scenario is more risk-effective without considering option prices.

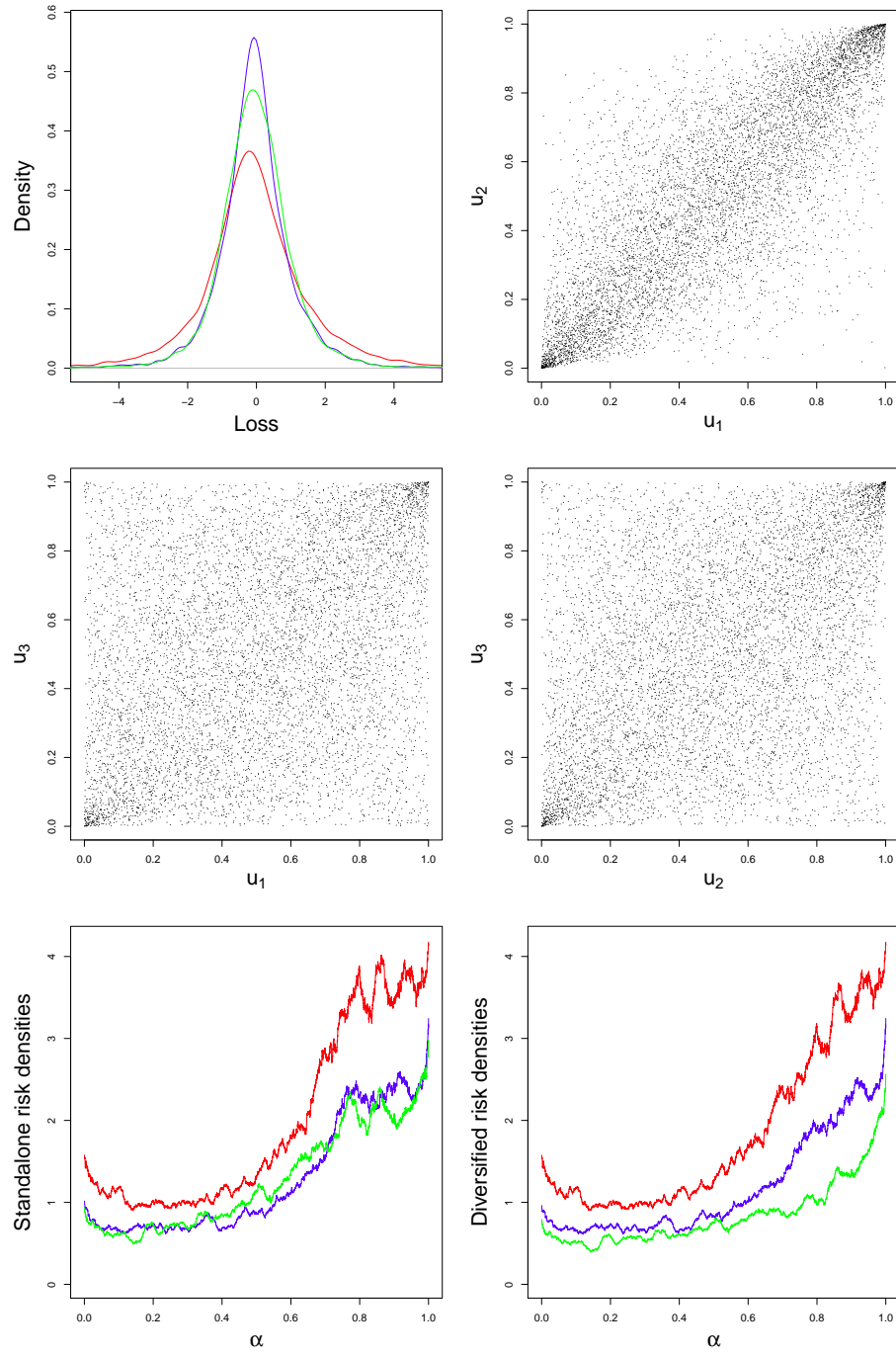


Figure 3: Top left panel plots empirical probability densities. Next 3 top panels plot empirical copulas (u_1 :NASDAQ, u_2 : S&P, u_3 : FTSE). Bottom left and right panels plot calculated standalone and systematic risk densities, respectively. Red, blue and green represent NASDAQ, S&P and FTSE, respectively.

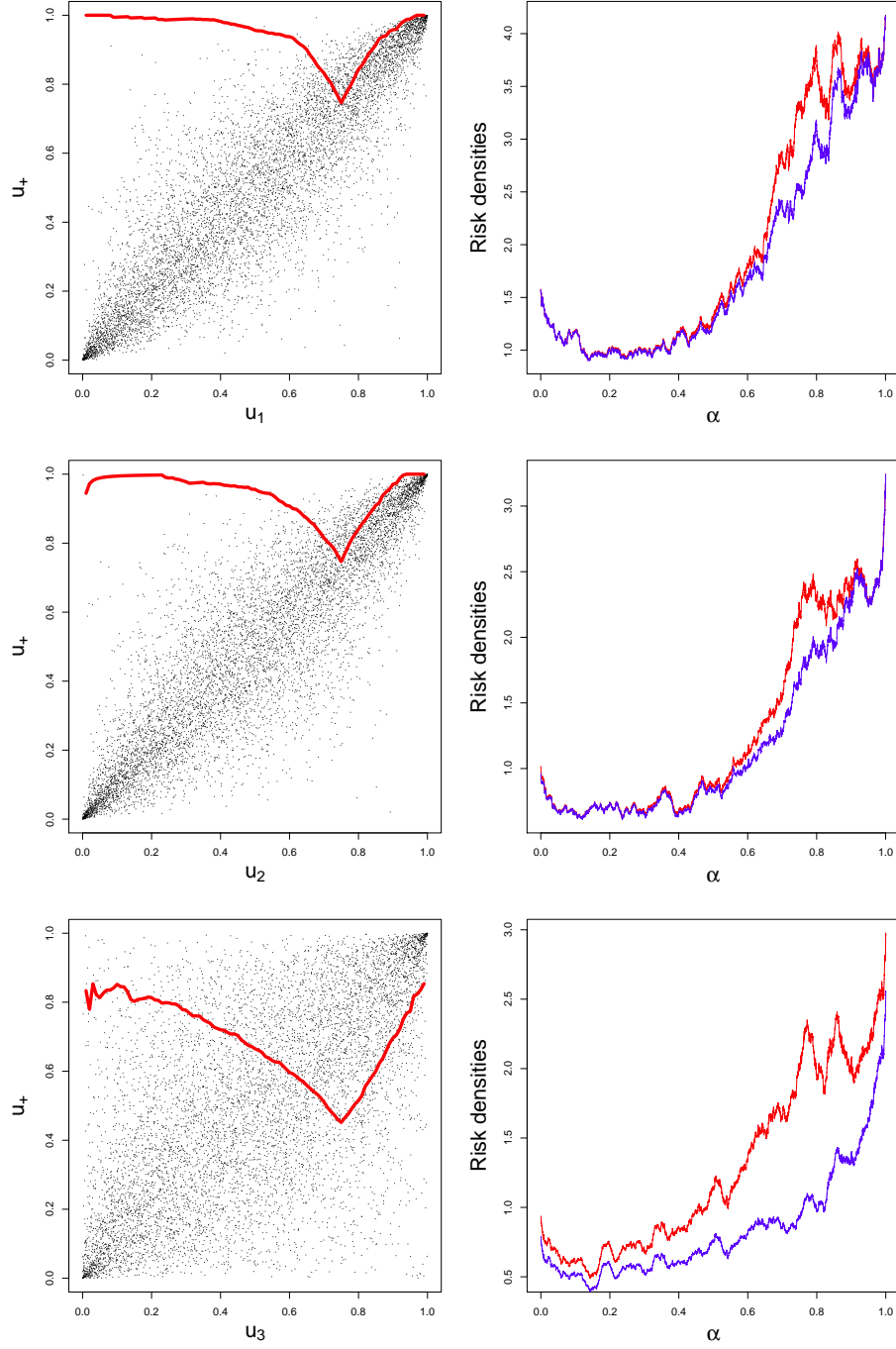


Figure 4: Left panels plot historical (u_i, u_+) and, in red, (α, θ_α) . Right panels plot calculated (α, r_α) in red and (α, \bar{r}_α) in blue. Note $\theta_\alpha = \bar{r}_\alpha / r_\alpha$. Top, middle and bottom panels are for NASDAQ, S&P and FTSE, respectively.

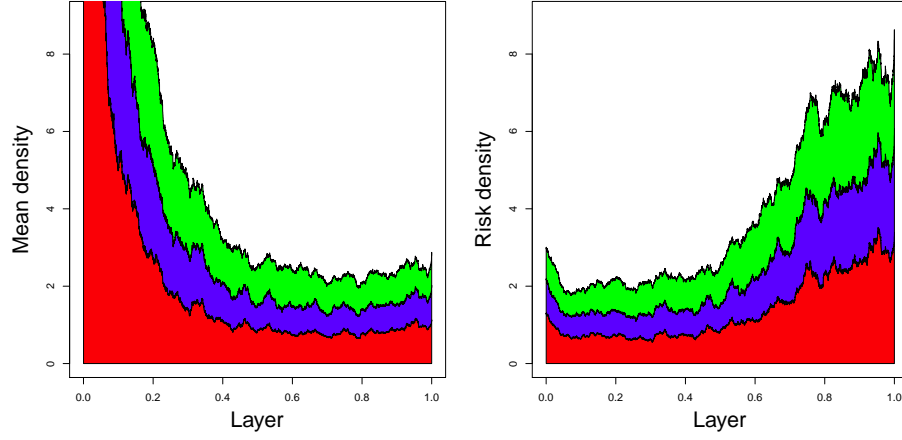


Figure 5: Left panel shows aggregate mean density and its breakdown into sub-aggregate mean densities. Right panel shows the same, for risk densities. Red, blue and green represent NASDAQ, S&P and FTSE, respectively.

Portfolio	r	\bar{r}	θ
NASDAQ	\$1.81	\$1.67	0.92
S&P	\$1.17	\$1.07	0.91
FTSE	\$1.17	\$0.75	0.64
Overall	\$4.15	\$3.49	0.84

Table 2: Overall standalone risk, systematic risk and risk ratio for each market index, after purchasing put options at exercise price $V_{0.95}$ for each index.

References

- Acerbi, C. (2002). Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking & Finance* 26(7), 1505–1518.
- Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). COHERENT MEASURES OF RISK. *Mathematical Finance* 9(3), 203–228.
- Buch, A. and G. Dorfleitner (2008). Coherent risk measures, coherent capital allocations and the gradient allocation principle. *Insurance: Mathematics and Economics* 42(1), 235–242.
- Choo, W. and P. De Jong (2009). Loss reserving using loss aversion functions. *Insurance: Mathematics and Economics* 45(2), 271–277.
- Choo, W. and P. De Jong (2010). Determining and allocating diversification benefits for a portfolio of risks. *Astin Bulletin* 40(1), 257–269.

- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues.
- Denault, M. (2001). Coherent allocation of risk capital. *Journal of risk* 4, 1–34.
- Denuit, M. and J. Dhaene (2012). Convex order and comonotonic conditional mean risk sharing. *Insurance: Mathematics and Economics*.
- Dhaene, J., M. Denuit, M. Goovaerts, R. Kaas, and D. Vyncke (2002). The concept of comonotonicity in actuarial science and finance: theory. *Insurance: Mathematics and Economics* 31(1), 3–33.
- Dhaene, J., M. Denuit, and S. Vanduffel (2009). Correlation order, merging and diversification. *Insurance: Mathematics and Economics* 45(3), 325–332.
- Hartmann, P., S. Straetmans, and C. G. De Vries (2004). Asset market linkages in crisis periods. *Review of Economics and Statistics* 86(1), 313–326.
- Hsieh, D. A. (1988). The statistical properties of daily foreign exchange rates: 1974–1983. *Journal of international economics* 24(1), 129–145.
- Kalkbrener, M. (2005). AN AXIOMATIC APPROACH TO CAPITAL ALLOCATION. *An International Journal of Mathematics, Statistics and Financial Economics* 15(3), 425–437.
- Kolb, R. (2010). *Lessons from the financial crisis: Causes, consequences, and our economic future*, Volume 12. John Wiley & Sons.
- Luenberger, D. (1998). *Investment Science*. Oxford University Press, USA.
- McNeil, A., R. Frey, and P. Embrechts (2005). *Quantitative risk management*. Princeton University Press.
- Rockafellar, R. and S. Uryasev (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance* 26(7), 1443–1471.
- Rodriguez, J. C. (2007). Measuring financial contagion: A copula approach. *Journal of Empirical Finance* 14(3), 401–423.
- Tasche, D. (2007). Capital allocation to business units and sub-portfolios: the euler principle. *Arxiv preprint arXiv:0708.2542*.
- Tsanakas, A. and N. Christofides (2006). Risk exchange with distorted probabilities. *Astin Bulletin* 36(1), 219.
- Wang, S. (1996). Premium Calculation by Transforming the Premium Layer Density. *ASTIN Bulletin* 26(1), 71–92.
- Wang, S. and J. Dhaene (1998). Comonotonicity, correlation order and premium principles. *Insurance: Mathematics and Economics* 22(3), 235–242.

Chapter 5

Tradeoff premiums

The following paper introduces the tradeoff premium which extends distortion risk measures by considering upside risk in addition to downside risk.

The tradeoff insurance premium as a two-sided generalisation of the distortion premium

Abstract

This paper introduces and analyzes the “tradeoff premium,” generalising the loss aversion reserve, distortion premium, spectral risk, and their duals. The tradeoff premium is a weighted average loss where weights increase as loss outcomes deviate from a subjective “loss appetite,” rather than from zero. The U-shaped weights replicate subjective probability adjustment in cumulative prospect theory, and minimise pricing error in a competitive market where overpricing and underpricing are both undesired.

Keywords: Weighted premium; loss aversion reserve; distortion premium; spectral risk; two-sided; loss appetite.

1. Introduction and overview

Premium principles, or risk measures, map a loss distribution to a real number. The mapping is used to calculate insurance premiums or manage risk. An example premium principle is the loss aversion reserve (Choo and De Jong, 2009), a weighted average loss where weights are a non-decreasing function of the loss percentile rank. Loss aversion reserves are equivalent to distortion premiums (Wang, 1996) and spectral risks (Acerbi, 2002). Generalised premium principles based on weighted average losses are discussed in Furman and Zitikis (2008). Gerber (1985) discusses an alternative premium based on the certainty equivalent loss under utility theory, the exponential premium (Deprez and Gerber, 1985) being a specific example. Other common premiums or risks are discussed in McNeil et al. (2005) and Young (2004), including Value-at-Risk, conditional-tail-expectation and the standard deviation premium. Artzner et al. (1999) discusses “coherence” properties of a premium principle or risk measure, namely translation invariance, positive homogeneity, monotonicity, and subadditivity.

Existing premium principles and risk measures are mostly “one-sided”, focussing on large loss outcomes and adding a positive loading to the expected loss. In pricing, a positive loading avoids inevitable ruin in the long run. In risk management, a positive loading captures worse than expected losses, which are the main concern. For example, the loss aversion reserve (or distortion premium and spectral risk) magnifies higher loss percentiles. The standard deviation premium sets the loading as a positive multiple of standard deviation. Assuming

a concave utility function, representing risk aversion, the certainty equivalent premium exceeds the expected loss. The Dutch premium (Van Heerwaarden and Kaas, 1992) also assumes a positive loading above the expected loss, based on a multiple of the expected excess loss.

A “two-sided” premium reflecting the importance of smaller loss outcomes, and potentially having a negative loading, is required for business reasons. In a competitive market, a negative loading may apply in the short term to increase business volume or to avoid loss of business, the latter if other market players are also applying negative loadings. Hence conflicting considerations exist – a positive loading is financially sustainable in the long term, however competitive pressure may force a negative loading to ensure short term survival. A negative loading may sustain for a longer term if the financial loss is offset by profit from other products with positive loading.

The tradeoff premium (ToP) is a novel premium principle addressing “two-sided” concerns highlighted in the previous paragraph. The ToP is a weighted average loss, with U-shaped weights increasing as loss outcomes deviate from a subjective “loss appetite.” U-shaped weights reflect the importance of both small and large loss outcomes, and are consistent with cumulative prospect theory (Tversky and Kahneman, 1992) where extreme outcomes (both positive and negative) are magnified and moderate outcomes are diminished.

The ToP is shown to reduce with loss appetite, with zero loss appetite yielding the loss aversion reserve (equivalently distortion premium or spectral risk). A maximum loss appetite implies monotonic decreasing weights, and yields a negative loading. In addition the resulting ToP is shown to be the dual (Wang, 2000) of the distortion premium with zero loss appetite.

Examples in this paper express the ToP as a two-sided generalisation of one-sided premiums or risks, including Value-at-Risk, conditional-tail-expectation and expected-maximal-loss (Choo and De Jong, 2009). The two-sided generalisation reflects the undesirability of both small and large loss outcomes, with the loss appetite controlling their relative representation in the ToP.

The equilibrium ToP corresponds to the loss appetite. Hence at equilibrium, the appetite for loss is equal to the premium collected. Undesired deviations of loss outcomes from the loss appetite represent premium surplus or shortfall. The equilibrium ToP is shown to be a tail-magnified measure of central tendency, refining the mean and median measures. Hence for a right skewed loss distribution, the equilibrium ToP exceeds the expected loss, and vice versa for a left skewed loss distribution.

The remaining paper is structured as follows. Section 2 defines, illustrates and justifies the ToP. Section 3 discusses properties of the ToP, in particular coherence. The ToP does not satisfy the subadditivity property of coherence, due to its two-sided nature. Section 4 discusses the equilibrium ToP. Section 5 identifies links between the ToP and existing literature. Section 6 decomposes the ToP into the expected loss, and a discount and loading respectively reflecting loss volatility below and above the loss appetite. Section 7 provides a numerical

example of the ToP using a gamma loss distribution and power aversion pattern. Section 8 concludes.

2. The tradeoff premium

The tradeoff premium (ToP) is a weighted premium (Furman and Zitikis, 2008). ToP generalises loss aversion reserves (Choo and De Jong, 2009), distortion premiums (Wang, 1996) and spectral risks (Acerbi, 2002) by assigning higher penalty weights to loss outcomes further from a “loss appetite” rather than to larger loss outcomes. Penalty weights forming the ToP are aligned with cumulative prospect theory (Tversky and Kahneman, 1992), where extreme outcomes, relative to a reference point, are magnified and moderate outcomes are diminished. The ToP also minimises pricing error in a competitive market.

To explain the ToP, first consider the loss aversion reserve of a random loss $x \geq 0$ with distribution function F and percentile rank $u \equiv F(x)$. Then u is uniformly distributed on the unit interval and indicates “loss severity,” with 0 being least severe and 1 being most severe. Given an increasing aversion function $\phi \geq 0$ integrating to one, the loss aversion reserve of x is

$$\mathbb{E}\{x\phi(u)\} = \int_0^\infty [1 - \Phi\{F(x)\}] dx = \int_0^1 V_u \phi(u) du, \quad (1)$$

$$\Phi(u) \equiv \int_0^u \phi(v) dv, \quad V_u \equiv F^-(u),$$

where \mathbb{E} computes expectation and V_u is the Value-at-Risk or VaR (McNeil et al., 2005) with sufficiency probability u . The second and third expressions in (1) are the distortion premium and spectral risk of x , respectively.

The loss aversion reserve (1) is a weighted average loss with higher weight on larger severities. Weights are on average one: $\mathbb{E}\{\phi(u)\} = 1$, since ϕ integrates to one. The distortion premium is the expected loss computed using a distorted distribution $\Phi \circ F$ dominating F , noting Φ is convex. The spectral risk is a weighted average of VaRs where higher VaRs are weighted more heavily. Hence loss aversion reserves, distortion premiums and spectral risks are “one-sided”, concerned only with larger severities. Concern is characterised by ϕ reweighing initially equally weighted VaRs or severities. Equal weighting $\phi = 1$ leads to the original expected loss $\mathbb{E}(x)$. Choo and De Jong (2009) shows the equivalence between loss aversion reserves, distortion premiums and spectral risks.

“Loss appetite” is central to the ToP. To introduce the concept of loss appetite, loss aversion reserves assume zero loss appetite. Then larger losses are always feared more and penalised more heavily. However a positive loss appetite may be relevant or even optimal in a business setting. In a competitive market, premiums may be deliberately reduced to win or prevent loss of market share. Premium reduction is achieved with a positive loss appetite. Penalty weights decrease up to the loss appetite, and increase thereafter. Maximum loss appetite

occurs if penalty weights always decrease with severity and contrasts with zero loss appetite where penalty weights always increase with severity. The aversion function redistributes weights, in regions below and above loss appetite.

To achieve the above redistribution of penalty weights first define “satiation error” based on loss severity u :

$$\psi_\ell(u) \equiv (u \leq \ell) \frac{\ell - u}{\ell} + (u > \ell) \frac{u - \ell}{1 - \ell}, \quad 0 \leq \ell \leq 1,$$

where the above bracketed inequalities are indicators. Here ℓ is a subjectively specified loss appetite on the percentile rank scale, with $\ell = 0$ indicating no appetite for losses and $\ell = 1$ indicating a complete appetite for losses. The deviation $u - \ell$ is standardised so that $\psi_\ell(0) = \psi_\ell(1) = 1$ for all $0 < \ell < 1$ with linear behaviour over the segments $[0, \ell]$ and $[\ell, 1]$.

Similar to loss severity u , satiation error $\psi_\ell(u)$ is uniformly distributed over the unit interval. Satiation error is also uniform over $u \leq \ell$ and over $u > \ell$. Zero loss appetite, $\ell = 0$, reduces satiation error to loss severity: $\psi_0(u) = u$. Maximum loss appetite, $\ell = 1$ yields satiation error $\psi_1(u) = 1 - u$. The role and selection of the loss appetite parameter ℓ is further discussed below.

The ToP is defined analogous to (1) by imposing an aversion function ϕ on satiation error $\psi_\ell(u)$, yielding

$$T_\ell \equiv \mathbb{E}\{x\phi_\ell(u)\} = \int_0^\infty [1 - \Phi_\ell\{F(x)\}] dx = \int_0^1 V_u \phi_\ell(u) du, \quad (2)$$

where

$$\phi_\ell \equiv \phi \circ \psi_\ell, \quad \Phi_\ell(u) \equiv \int_0^u \phi_\ell(v) dv.$$

Note T_0 reduces to the expression in (1) since $\phi_0(u) = \phi(u)$. Conversely maximum loss appetite $\ell = 1$ yields satiation error $\phi_1(u) = \phi(1 - u)$. The resulting ToP, T_1 , is the “dual” of the original distortion premium (Wang, 2000):

$$\int_0^\infty [1 - \Phi^*\{F(x)\}] dx, \quad \Phi^*(u) \equiv 1 - \Phi(1 - u).$$

The ToP is a “two-sided” generalisation of the loss aversion reserve, noting weights $\phi_\ell(u) = \phi\{\psi_\ell(u)\}$ increase as loss severity u deviates from loss appetite ℓ , rather than from zero. Overall weights below and above the loss appetite are

$$\int_0^\ell \phi_\ell(u) du = \ell, \quad \int_\ell^1 \phi_\ell(u) du = 1 - \ell,$$

respectively, noting satiation error $\psi_\ell(u)$ is uniform over both $u \leq \ell$ and $u > \ell$. Hence piecewise uniformity of $\psi_\ell(u)$ preserves the overall weight placed on severities below and above the loss appetite ℓ . The loss appetite ℓ serves to redistribute severity weights, using the aversion function ϕ . An example illustrating the ToP is shown in the following subsection. Subsequent subsections further justify the formulation of the ToP based on cumulative prospect theory and minimisation of pricing error in a competitive market.

2.1. Minmax ToP example

The following example illustrates T_ℓ given power aversion function $\phi(u) = nu^{n-1}$ with $n \geq 1$. In this case

$$T_\ell = \ell E \left\{ \min_{i=1,\dots,n} (x_i) \middle| \max_{i=1,\dots,n} (x_i) \leq V_\ell \right\} + (1-\ell) E \left\{ \max_{i=1,\dots,n} (x_i) \middle| \min_{i=1,\dots,n} (x_i) > V_\ell \right\}$$

where the x_i are n independent copies of x . With minimum and maximum loss appetite, or $\ell = 0, 1$, the ToPs are respectively

$$T_0 = E \left\{ \max_{i=1,\dots,n} (x_i) \right\}, \quad T_1 = E \left\{ \min_{i=1,\dots,n} (x_i) \right\}.$$

Further for $0 \leq \ell \leq 1$, $T_1 \leq T_\ell \leq T_0$. If $n = 1$ then $\phi(u) = 1$ and the ToP is, for all ℓ ,

$$T_\ell = \ell E(x|x \leq V_\ell) + (1-\ell)E(x|x > V_\ell) = E(x).$$

The following key properties of the ToP are inferred from the above example, and are formalised in subsequent sections:

- The ToP is formed by combining “aversion adjusted” conditional expected losses below and above the loss appetite. The expected loss below the loss appetite is reduced by assuming the expected minimum. The expected loss above the loss appetite is raised by assuming the expected maximum. Without aversion adjustment ($n = 1$) the ToP is the expected loss.
- Loss appetite $0 \leq \ell \leq 1$ controls the size of the ToP. T_ℓ is at maximum if $\ell = 0$, that is if there is no appetite for loss. Increasing loss appetite reduces the ToP. A proof is shown in §3.
- Relative left and right tail skewness about ℓ also affect the ToP, given the loss appetite. Increasing the skewness of the right tail while keeping $E(x)$ constant increases the expected maximum of losses above loss appetite, resulting in a higher ToP.
- Increasing loss aversion by increasing n in $\phi(u) = nu^{n-1}$ accentuates the impact of relative tail skewness on the ToP. Consider the case where the right tail above V_ℓ is more skewed than the left tail below V_ℓ . Increasing n has a greater impact on the expected maximum of losses above the loss appetite than the expected minimum of losses below the loss appetite, thus increasing the ToP overall. With neutrality $n = 1$ or $\phi = 1$, relative tail skewness is ignored and the ToP is always the expected loss $E(x)$. In addition changing ℓ alters relative tail skewness and hence the impact of increasing n will differ according to ℓ .
- The same aversion attitude applies to losses below and above the loss appetite. In this example aversion adjustment focusses on the expected maximum or minimum over n copies.

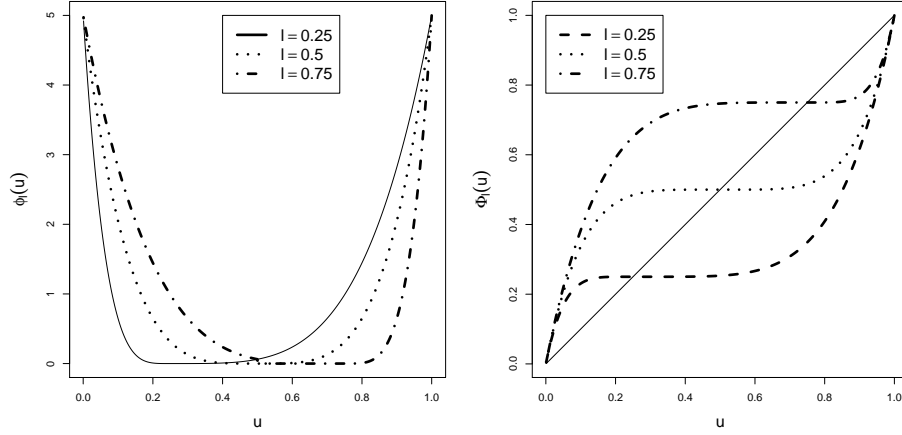


Figure 1: Plots of $\phi_\ell(u)$ (left panel) and $\Phi_\ell(u)$ (right panel) against u assuming $\phi(v) = 5v^4$ and $\ell = 0.25, 0.5$ and 0.75 .

2.2. ToP and cumulative prospect theory

In cumulative prospect theory, extreme outcomes are over-weighted while average outcomes are under-weighted. This psychological phenomenon is represented by a U-shaped weight function on probabilities, and a S-shaped transformation of cumulative probabilities. The ToP effects a similar modification. The weight function ϕ_ℓ in the first expression of the ToP in (2) is U-shaped: decreasing below ℓ and increasing above ℓ . In the second expression, the transformation Φ_ℓ applied to the distribution F is S-shaped: concave below ℓ and convex above ℓ . Left and right panels in Figure 1 illustrate the U-shaped ϕ_ℓ and S-shaped Φ_ℓ , respectively, assuming $\phi(v) = 5v^4$, for various values of ℓ .

2.3. ToP and the minimisation of pricing error

In a competitive market, overpricing leads to loss of business. Hence smaller loss outcomes are feared. Larger loss outcomes are also feared, due to underpricing and financial loss. Loss appetite ℓ and the associated VaR V_ℓ sets the boundary between perceived “small” and “large” losses. A lower loss appetite ℓ yields a higher proportion of “large” losses and vice versa. U-shaped penalty weights $\phi_\ell(u)$ models the undesirability of various severities.

Given premium π , then

$$\mathbb{E} \{ (\pi - V_u)^2 \phi_\ell(u) \} = \int_0^1 (\pi - V_u)^2 d\Phi_\ell(u), \quad (3)$$

is the overall pricing error. Squared pricing errors $(\pi - V_u)^2$ are weighed according to the aversion function ϕ_ℓ , given loss appetite ℓ . Large pricing errors, those

associated with smaller or larger severities u , are featured more prominently in the overall pricing error via larger squared differences and penalty weights. The ToP, T_ℓ , is that premium π minimising overall pricing error in (3).

2.4. Role and selection of loss appetite

The loss appetite parameter ℓ is central to the ToP and distinguishes it from one-sided premium principles. As mentioned just after (2), zero loss appetite $\ell = 0$ yields the usual distorted premium, whereas the dual is obtained with $\ell = 1$, the maximum loss appetite.

Loss appetite ℓ is subjectively selected within the unit interval. The chosen ℓ is a neutrality, “preference” or “comfort” point with no aversion: $\phi_\ell(\ell) = 0$ and no distribution adjustment: $\Phi_\ell(\ell) = \ell$. Aversion increases as $|u - \ell|$ increases.

Loss appetite ℓ balances the fear of “small” severities, suggesting overpricing and possible loss of business, relative to the fear of “large” severities, indicating underpricing and an overall loss. A low loss appetite, $\ell \approx 0$, indicates little fear of overpricing and loss of business, and large fear of underpricing and overall loss and vice versa. In the above example, the loss appetite is the weight attached to aversion adjusted expected losses below the loss appetite, while the complement is the weight attached to the corresponding expectation of losses above the loss appetite.

An equilibrium occurs if $F(T_\ell) = \ell$, that is if $T_\ell = V_\ell$, the VaR at ℓ . In this case the loss appetite ℓ equals the probability of an adequate premium. In addition, since the equilibrium ToP is equal to the loss appetite, satiation error $\phi_\ell(u)$, measuring deviations between loss outcomes and the loss appetite on the percentile rank scale, also indicates the extent of premium surplus or shortfall. Equilibrium ToPs are further discussed in §4.

2.5. Aversion symmetry about the loss appetite

The ToP assumes symmetric aversion to loss severities about the loss appetite. From the above example, the aversion adjustment applies the expected maximum or minimum over n copies for severities below and above the loss appetite. The symmetry is the result of two factors in the setup of the ToP: satiation error $\psi_\ell(u)$ varies uniformly over the unit interval for severities both below and above the loss appetite ℓ , and a single aversion function ϕ is applied to satiation error in the formulation of penalty weights.

Hence a single aversion attitude applies to losses below and above the loss appetite. There is no bias when performing “aversion adjustment” to left and right tails. This unbiasedness is different from, and should not be confused with, the relative fear towards left and right tails, controlled by loss appetite.

3. Features of the ToP

3.1. Relationship with loss appetite

As noted above T_ℓ is monotonic decreasing in ℓ : if $\ell \leq \ell_*$ then $T_\ell \geq T_{\ell_*}$. The proof follows by noting Φ_ℓ is increasing in ℓ , or $\Phi_\ell(u) \leq \Phi_{\ell_*}(u)$ for all u if $\ell \leq \ell_*$. Thus the distorted distribution $\Phi_\ell \circ F$ is also increasing in ℓ , implying the ToP is decreasing in ℓ , based on the second expression in (2).

Thus higher loss appetite reduces T_ℓ . Zero loss appetite $\ell = 0$ and maximum loss appetite $\ell = 1$ yield maximum and minimum ToPs, respectively. Higher loss appetite indicates higher tolerance of larger loss outcomes, hence a greater willingness to charge lower premiums.

3.2. Relation to the standard deviation premium

If μ_x and σ_x are the mean and standard deviation of x respectively and σ_ϕ is standard deviation of $\phi(v)$ where v is uniform, then express the ToP as

$$T_\ell = \mu_x + \text{cov}\{x, \phi_\ell(u)\} = \mu_x + \sigma_x \times [\sigma_\phi \text{cor}\{x, \phi_\ell(u)\}] , \quad (4)$$

where cov and cor denotes the covariance and correlation, respectively. The first equality holds since $E\{\phi_\ell(u)\} = E\{\phi(u)\} = 1$. Thus T_ℓ can be thought of as a standard deviation premium (Young, 2004): the mean μ_x plus a multiple of the standard deviation σ_x of the loss. The multiple $\sigma_\phi \text{cor}\{x, \phi_\ell(u)\}$ can be positive or negative and depends on two factors:

- The aversion standard deviation σ_ϕ , representing the overall aversion to satiation error. High σ_ϕ implies aversion increases dramatically when satiation error increases, while neutrality to satiation error implies $\phi = 1$ and hence $\sigma_\phi = 0$. The value of σ_ϕ does not depend on loss appetite ℓ .
- The loss appetite ℓ , controlling the correlation term. Zero loss appetite $\ell = 0$ maximises the multiple, while increasing loss appetite reduces the multiple. Maximum loss appetite $\ell = 1$ yields a negative standard deviation multiple, since $\phi_1(u) = \phi(1 - u)$ is negatively correlated with x . Note ℓ affects T_ℓ only through the standard deviation multiple.

3.3. Linearity

The standard deviation multiple $\sigma_\phi \text{cor}\{x, \phi_\ell(u)\}$ in (4) is invariant to location and scale changes in x . Hence the ToP is linear in the loss random variable:

$$T_\ell(\alpha + \beta x) = \alpha + \beta T_\ell(x) , \quad 0 \leq \ell \leq 1 ,$$

where α and $\beta \geq 0$ are constants and $T_\ell(x)$ is the ToP of random loss x . The linearity property corresponds to translation invariance and positive homogeneity properties of coherent risk measures (Artzner et al., 1999). Other properties of coherent risk measures are discussed below.

3.4. Stochastic dominance and monotonicity

Suppose loss y stochastically dominates x in first order, or the distribution function of y is less than x at all points. Then $T_\ell(y) > T_\ell(x)$ for any loss appetite ℓ . Hence “larger” loss random variables have higher ToPs, given the same loss appetite. The proof follows from the second expression for the ToP in (2), noting Φ_ℓ is the same for both x and y .

The ToP also preserves statewise dominance, the monotonicity property of coherent risk measures. When loss y exceeds x over all states, y also stochastically dominates x in first order, thus y has a larger ToP compared to x .

3.5. Non-subadditivity

Unlike coherent risk measures, the ToP is not always sub-additive: the ToP of a sum of losses may exceed the sum of ToPs for each individual loss.

Non-subadditivity follows from the “two-sidedness” of the ToP. One-sided loss aversion reserves, distortion premiums and spectral risks are sub-additive since undesired large severities are “diversified” upon aggregation, resulting in a reduced overall reserve, premium or risk:

$$T_0\left(\sum_i x_i\right) \leq \sum_i T_0(x_i) ,$$

where x_i are possibly dependent losses. Proofs are given in Wang (1996) and Choo and De Jong (2009). In contrast the dual is super-additive: the premium increases upon aggregation since undesired smaller severities are diversified:

$$T_1\left(\sum_i x_i\right) \geq \sum_i T_1(x_i) .$$

For the ToP where $0 \leq \ell \leq 1$, both small and large severities are undesired, and are simultaneously diversified upon aggregation. The overall impact on the ToP depends on the loss appetite ℓ : low ℓ implies large severities have dominant concern, resulting in subadditivity. Conversely high ℓ indicates dominant concern on small severities, yielding a super-additive ToP.

3.6. Additivity for comonotonic losses

Suppose x and y are comonotonic losses (Dhaene et al., 2002) such that x , y and $x + y$ have equal percentile rank in any state. Then

$$T_\ell(x + y) = T_\ell(x) + T_\ell(y) , \quad 0 \leq \ell \leq 1 ,$$

since the aversion weights $\phi_\ell(u)$ are equal for x , y and $x + y$ under comonotonicity, and applying the first expression in (2).

3.7. No unfair premium

The ToP lies with the range of the loss for any loss appetite ℓ :

$$\min(x) \leq T_\ell \leq \max(x) ,$$

noting the ToP is a weighted average loss with non-negative weights $\phi_\ell(u)$.

3.8. Relation to integral operators

Multiplying the last expression for the ToP in (2) by $\phi_\ell^-(v)$, defined below, and integrating with respect to ℓ over the unit interval yields

$$\int_0^1 T_\ell \phi_\ell^-(v) d\ell = \int_0^1 V_u \left\{ \int_0^1 \phi_\ell(u) \phi_\ell^-(v) d\ell \right\} du = V_v ,$$

where the final equality holds if the above expression in curly brackets is the Dirac-delta function ($u = v$). Hence $\phi_\ell^-(v)$ is obtained by solving

$$\int_0^1 \phi_\ell(u) \phi_\ell^-(v) d\ell = (u = v) . \quad (5)$$

In the above setup $\phi_\ell(u)$ is an integral operator yielding ToP at any loss appetite ℓ based on all VaRs V_u . In addition $\phi_\ell^-(v)$ is the corresponding inverse integral operator yielding any VaR V_v , given ToPs at all loss appetites ℓ . Both $\phi_\ell(u)$ and $\phi_\ell^-(v)$ only depend on the aversion function ϕ . A closed form expression for $\phi_\ell^-(v)$ does not exist. However integrating (5) with respect to u over the unit interval yields

$$\int_0^1 \phi_\ell^-(v) d\ell = \int_0^1 (u = v) du = 1 ,$$

noting $\phi_\ell(u)$ integrates to one. Hence the inverse weights $\phi_\ell^-(v)$ placed on ToPs, over all loss appetites ℓ , integrate to one. A similar property applies to original weights $\phi_\ell(u)$.

4. Equilibrium tradeoff premium

An equilibrium occurs if $F(T_\ell) = \ell$ or $T_\ell = V_\ell$: that is if the ToP T_ℓ is the ℓ -VaR corresponding to loss appetite ℓ . Hence at equilibrium the probability of premium sufficiency equals the loss appetite. A unique solution ℓ_* to $T_\ell = V_\ell$ exists since T_ℓ is decreasing in ℓ and V_ℓ is increasing in ℓ , and V_ℓ covers a larger range of values over $0 \leq \ell \leq 1$ compared to T_ℓ .

The equilibrium condition suggests the iterative scheme

$$T_* = \lim_{n \rightarrow \infty} T_{\ell_n} , \quad \ell_{n+1} = F(T_{\ell_n}) , \quad n = 0, 1, \dots , \quad (6)$$

where ℓ_0 is an initial loss appetite and $T_* \equiv T_{\ell_*}$ denotes the equilibrium ToP. In addition define $V_* \equiv V_{\ell_*} = T_*$ as the VaR corresponding to the equilibrium loss appetite ℓ_* .

The iteration (6) has a practical interpretation. Suppose $\ell < \ell_*$ implying $T_\ell > T_* = V_* > V_\ell$ where the first inequality applies since T_ℓ is decreasing in ℓ . Then the VaR at loss appetite ℓ falls below the ToP, that is the amount the insurer is willing to lose is less than the premium collected. This “disequilibrium” creates a tendency to increase loss appetite or reduce premium. An analogous argument applies if ℓ is above ℓ_* . Inconsistency between the updated loss appetite and ToP results in further calculation, until convergence.

The equilibrium ToP is further discussed in the following subsections. A numerical example is given in §7.

4.1. Special equilibrium cases

A special case of equilibrium ToP occurs with neutrality to satiation error, $\phi = 1$. In this case $T_\ell = \mu_x$, the mean loss, for all ℓ , and hence $T_* = \mu_x$. The equilibrium loss appetite is $\ell_* = F(\mu_x)$.

The equilibrium ToP is also equal to the mean loss if the loss distribution is symmetric, and $\ell_* = F(\mu_x) = 0.5$. This result is established by noting satiation error $\psi_{0.5}(u)$ and hence aversion weights $\phi_{0.5}(u)$ are always symmetric about $u = 0.5$. In addition loss percentiles V_u are also symmetric about $u = 0.5$ for a symmetric distribution, therefore the ToP at $\ell = 0.5$ achieves equilibrium: $T_{0.5} = E\{V_u \phi_{0.5}(u)\} = V_{0.5} = \mu_x$, completing the proof.

4.2. Central tendency and impact of skewed loss distributions

The equilibrium ToP is a measure of central tendency akin to the mean and median. The equilibrium ToP magnifies both tails of the loss distribution, and thus includes a positive loading over the mean loss for a right skewed distribution and vice versa for a left skewed distribution.

Re-arranging the equilibrium condition $T_* = V_*$ and replacing T_ℓ with the definition in (2) yields

$$E\{|V_* - x|^+ \phi_*(u)\} = E\{|x - V_*|^+ \phi_*(u)\} , \quad \phi_* \equiv \phi_{\ell_*} , \quad (7)$$

and $|x|^+$ indicates the “positive part of x .” Hence the equilibrium ToP equalises expected premium surplus $|T_* - x|^+$ and premium shortfall $|x - T_*|^+$. The aversion adjustment ϕ_* magnifies larger surpluses or shortfalls except in the case of neutrality $\phi \equiv 1$.

The equilibrium condition in (7) is analogous to the well known equilibrium conditions for the mean and median:

$$E(|\mu_x - x|^+) = E(|x - \mu_x|^+) , \quad P(x \leq m_x) = P(x > m_x) .$$

where m_x is the median of x and P calculates probability using F . The mean loss equalises expected surplus and shortfall without aversion adjustment, that is assuming neutrality towards pricing error. The median loss equalises probabilities of surplus and shortfall ignoring their magnitude.

Hence T_* is a measure of central tendency similar to the mean and median. The equilibrium ToP generalises the mean by including an aversion adjustment weighing larger deviations more heavily: both tails of the loss distribution are amplified. The result of amplification is discussed below. For example if the aversion function is $\phi(v) = nv^{n-1}$ then weights increase as a power of the deviation from T_* .

Amplifying both tails of the loss distribution creates a “loading” and “discount” over the mean loss for a right and left skewed distribution, respectively. Hence for a right skewed distribution, $T_* > \mu_x$ and vice versa. The proof follows by noting, for a right skewed distribution, substituting $V_* = \mu_x$ and $\ell_* = F(\mu_x)$ into (7) creates disequilibrium:

$$E \{ |\mu_x - x|^+ \phi_{F(\mu_x)}(u) \} < E \{ |x - \mu_x|^+ \phi_{F(\mu_x)}(u) \} ,$$

noting the impact of tail amplification is larger for the right tail compared to the left tail. Since the left hand side of (7) is increasing in ℓ_* while the right hand side is decreasing in ℓ_* , V_* must increase from its current value μ_x to achieve equilibrium. Hence at equilibrium $T_* = V_* > \mu_x$, completing the proof. A similar proof applies for a left skewed loss distribution.

4.3. Existence and persistence of disequilibrium

As mentioned above, disequilibrium arises if the ToP differs from the loss appetite: $T_\ell \neq V_\ell$. Disequilibrium may exist, and persist, due to conscious decision. Suppose $\ell < \ell_*$ resulting in $T_\ell > V_\ell$, or loss appetite is less than the ToP. Disequilibrium persists if a decision is made to set loss appetite less than the ToP, for example when underpricing and financial loss is of utmost concern. A change in loss appetite only occurs when new factors create pressure on disequilibrium to be reduced or eliminated, such as fear of overpricing and loss of future business. In this case the loss appetite increases, resulting in a lower ToP. The adjustment ceases when equilibrium is achieved: $T_\ell = V_\ell$.

5. Connection to literature

As mentioned in §2, the ToP is a weighted premium (Furman and Zitikis, 2008) and reduces to loss aversion reserves (Choo and De Jong, 2009), distortion premiums (Wang, 1996) and spectral risks (Acerbi, 2002) when loss appetite is zero. The weights forming the ToP are consistent with cumulative prospect theory (Tversky and Kahneman, 1992). Lastly the ToP is a standard deviation premium (Young, 2004), where the standard deviation multiple depends on loss appetite and aversion to pricing error.

The ToP also generalises the zero utility premium (Heilpern, 2003) using rank-dependent utility theory (Quiggin, 1982) and assuming a linear utility function. This connection is established by writing T_ℓ as

$$E_\ell \{U(T_\ell - x)\} = 0, \quad U(w) = a + bw,$$

where E_ℓ calculates expectations with respect to $\Phi_\ell \circ F$, the modified distribution of the loss x , and U is an utility function. Rank-dependent utility theory assumes $\ell = 0$, or zero loss appetite. In contrast the ToP allows for subjective selection of loss appetite, based on the relative concern of over and underpricing.

Kaluszka and Krzeszowiec (2011) applies cumulative prospect theory to the generalised zero utility premium mentioned above, and sets the premium as the loss appetite. The resulting premium corresponds to the equilibrium ToP discussed in §4, again assuming a linear utility function. The ToP considers disequilibrium cases where loss appetite is deliberately set to be different from the calculated premium.

Van der Hoek and Sherris (2001) use differential probability adjustment to outcomes above and below a “reference point,” by using two different distortion operators. The resulting generalized distortion premium is applied to asset allocation. However both distortion operators applied by Van der Hoek and Sherris (2001) are convex, implying monotonic increasing concern on larger loss outcomes, or a zero loss appetite, similar to rank dependent utility. On the other hand the ToP combines convex and concave distortions yielding a S-shaped transformation of cumulative probabilities. In addition Van der Hoek and Sherris (2001) expresses the reference point in absolute terms, whereas the ToP specifies the loss appetite (the reference point) as a percentile rank.

6. Decomposing the ToP

The ToP is composed of two premiums separately addressing over and underpricing given loss appetite. The premium for losses below the loss appetite places higher weight on smaller severities, and vice versa for the premium relating to losses above the loss appetite. Further manipulation of the ToP yields a discount and markup applied to the expected loss, proportional to left and right tail volatilities, respectively.

Write $T_\ell(x|A)$ as the ToP for loss x conditional on event A and given loss appetite ℓ . Note the percentile rank, or severity, of conditional losses below and above the loss appetite are $1 - \psi_\ell(u)$ and $\psi_\ell(u)$, respectively. Applying iterated expectations to the ToP yields

$$T_\ell = E\{x\phi_\ell(u)\} = \ell T_1(x|x \leq V_\ell) + (1 - \ell)T_0(x|x > V_\ell). \quad (8)$$

Recall T_0 is the one-sided ToP focussed totally on larger severities, while T_1 is solely concerned on smaller severities.

The decomposition (8) emphasizes features of the ToP explained in §2. There is aversion to smaller severities below the loss appetite ℓ and larger severities above the loss appetite. Severities above the loss appetite are priced using the loss aversion reserve, distortion premium or spectral risk, whilst severities below the loss appetite are priced using the corresponding dual. The loss appetite ℓ determines what constitutes “small” and “large” severities. A low loss appetite indicates most severities are considered “large” with higher concern on larger severities resulting in a higher ToP, and vice versa for a high loss appetite. Weights attached to the two premiums are the probabilities of losses falling below and above the loss appetite.

The decomposition of the ToP in (8) generalises the example in §2 using the power aversion function $\phi(v) = nv^{n-1}$, where T_1 is the expected minimum loss and T_0 is the expected maximum loss, both over n copies. Other examples using other aversion functions are given and discussed below.

Further rewrite the ToP as follows. Let σ_x^- and σ_x^+ denote the standard deviation of losses below and above the loss appetite, respectively. In addition apply the decomposition of the ToP in (4) into standard deviation and correlation terms. Then the ToP, by rewriting (8), is

$$T_\ell = \mu_x + \sigma_\phi \{ -\ell \tilde{\sigma}_x^- + (1 - \ell) \tilde{\sigma}_x^+ \} , \quad (9)$$

where

$$\tilde{\sigma}_x^- \equiv \sigma_x^- \text{cor} \{ -x, \phi_\ell(u) | x \leq V_\ell \} , \quad \tilde{\sigma}_x^+ \equiv \sigma_x^+ \text{cor} \{ x, \phi_\ell(u) | x > V_\ell \} ,$$

are the “aversion adjusted” loss volatility in the left and right tails, respectively. Aversion adjusted volatilities are formed by multiplying loss standard deviations with correlations or “correction factors” between 0 and 1, the latter measuring similarity between movements in loss outcomes and aversion. Choo and De Jong (2009) further discusses correction factors.

Based on (9) the ToP comprises of the expected loss μ_x , a discount $\ell \sigma_\phi \tilde{\sigma}_x^-$ and a markup $(1 - \ell) \sigma_\phi \tilde{\sigma}_x^+$. The discount reflects concern towards smaller severities below the loss appetite whereas the markup reflects concern towards larger severities above the loss appetite. The discount and markup are proportional to the overall aversion to mispricing σ_ϕ , and loss volatility in the respective tail.

The ToP exceeds the expected loss if and only if $(1 - \ell) \tilde{\sigma}_x^+ > \ell \tilde{\sigma}_x^-$, or the expected loss volatility in the right tail exceeds the expected loss volatility in the left tail. This occurs if the right tail is more skewed, resulting in high σ_x^+ , or the loss appetite is low resulting in high $1 - \ell$. The overall aversion σ_ϕ scales the impact of relative tail size and choice of loss appetite on the ToP. High aversion magnifies the impact, whereas neutrality $\phi = 1$ or $\sigma_\phi = 0$ eliminates the impact with $T_\ell = \mu_x$ regardless of relative tail size or loss appetite.

6.1. Examples of decomposing the ToPs

The following illustrates the decomposition of the ToP in (8) by applying three aversion functions described in Choo and De Jong (2009). These aversion

functions yield well known examples of loss aversion reserves, distortion premiums or spectral risks. The corresponding ToP is a “two-sided” generalisation of existing examples, with emphasis on both smaller and larger losses relative to the loss appetite.

- Suppose $\phi(v) = (v = \alpha)$, a Dirac delta function where $0 \leq \alpha \leq 1$. Aversion exists to a single satiation error α , and other errors are ignored. The ToP in this case is

$$T_\ell = \ell V_{(1-\alpha)\ell} + (1 - \ell) V_{1-(1-\alpha)(1-\ell)} ,$$

a weighted average of lower and upper VaRs. The corresponding loss aversion reserve or distortion premium, and the dual, are V_α and $V_{1-\alpha}$ respectively, the upper and lower VaRs. VaR is an extreme percentile in either tail, whereas the ToP in this case combines extreme percentiles in both tails to reflect concerns towards smaller and larger severities. Hence call this ToP the “two-sided VaR.”

In this example ϕ is not increasing, but it nevertheless highlights the two-sided feature of ToPs. In addition T_0 and T_1 are not subadditive and superadditive, respectively – for example see Denuit et al. (2005).

- Suppose $\phi(v) = (v > \alpha)/(1 - \alpha)$ for $0 \leq \alpha \leq 1$, the step function equal to $1/(1 - \alpha)$ if $v > \alpha$ and 0 otherwise. Satiation errors below α are ignored while larger errors above α are magnified by the factor $1/(1 - \alpha)$. The resulting ToP is a weighted average of small and large losses:

$$T_\ell = \ell E\{x|x \leq V_{(1-\alpha)\ell}\} + (1 - \ell) E\{x|x > V_{1-(1-\alpha)(1-\ell)}\} ,$$

and is called the “two-sided conditional-tail-expectation.” The one-sided loss aversion reserve or distortion premium and the dual are, respectively

$$T_0 = E(x|x > V_\ell) , \quad T_1 = E(x|x \leq V_{1-a}) ,$$

known as the “conditional-tail-expectation” (McNeil et al., 2005). Whilst the conditional-tail-expectation takes the average of extreme losses in one or other tail, the ToP averages extreme losses in both tails.

- The power aversion function $\phi(v) = nv^{n-1}$ is discussed in §2. The ToP as shown in §2 is

$$\begin{aligned} & \ell E \left\{ \min_{i=1,\dots,n} (x_i) \middle| \max_{i=1,\dots,n} (x_i) \leq V_\ell \right\} \\ & + (1 - \ell) E \left\{ \max_{i=1,\dots,n} (x_i) \middle| \min_{i=1,\dots,n} (x_i) > V_\ell \right\} \equiv T_{\ell,n} , \end{aligned}$$

or the “two-sided expected-maximal-loss,” noting the one-sided equivalents are the expected maximal or minimal loss.

Construct another aversion function weighting positive integer values of n in nv^{n-1} , using the Poisson distribution with mean λ and truncating away zero. The resulting aversion function is

$$\phi(v) = \sum_{n=1}^{\infty} \{nv^{n-1}p_n\} = \frac{e^{\lambda v}}{E(e^{\lambda v})}, \quad p_n \equiv \frac{e^{-\lambda}\lambda^n}{n!(1-e^{-\lambda})}, \quad n \geq 1.$$

Therefore the aversion function is exponentially increasing at rate λ . The expected value of the parameter n is

$$\sum_{n=1}^{\infty} np_n = \frac{\lambda}{1-e^{-\lambda}} \approx \lambda$$

where the approximation assumes sufficiently large λ . Hence rather than fixing the parameter n in the power aversion function, a mean value is selected and the actual value varies according to a truncated Poisson distribution. Using the resulting exponential aversion function, one-sided ToPs assuming zero and maximum loss appetite are, respectively,

$$T_0 = \frac{E(xe^{\lambda u})}{E(e^{\lambda u})}, \quad T_1 = \frac{E\{xe^{\lambda(1-u)}\}}{E(e^{\lambda u})},$$

which are percentile rank versions of Esscher premiums (Van Heerwaarden et al., 1989): $E(xe^{\lambda x})/E(e^{\lambda x})$. In addition the ToP for $0 < \ell < 1$ is

$$\begin{aligned} T_\ell &= \frac{\ell E\left\{xe^{\lambda(\frac{\ell-u}{\ell})}\middle|x \leq V_\ell\right\} + (1-\ell)E\left\{xe^{\lambda(\frac{u-\ell}{1-\ell})}\middle|x > V_\ell\right\}}{E(e^{\lambda u})} \\ &= \sum_{n=1}^{\infty} T_{\ell,n} p_n = \ell \sum_{n=1}^{\infty} \left[nv^{n-1} \times E\left\{\min_{i=1,\dots,n}(x_i)\middle|x_i \leq V_\ell\right\} \right] \\ &\quad + (1-\ell) \sum_{n=1}^{\infty} \left[nv^{n-1} \times E\left\{\max_{i=1,\dots,n}(x_i)\middle|x_i > V_\ell\right\} \right]. \end{aligned}$$

where the last two expressions above are weighted averages of two-sided expected-maximal-losses using identical Poisson weights as those applied in generating the aversion function. Aversion weights $\phi_\ell(u)$ are decreasing exponentially up to $u = \ell$, and increasing exponentially thereafter.

Choo and De Jong (2009) constructs other aversion functions based on weighted averages $w\phi_1 + (1-w)\phi_2$ and compositions $(\Phi_1 \circ \Phi_2)'$, where ϕ_1 and ϕ_2 are aversion functions and Φ_1 and Φ_2 are corresponding distortion operators such that $\Phi'_1 = \phi_1$ and $\Phi'_2 = \phi_2$.

7. Numerical examples

This section illustrates key properties of the ToP and its equilibrium value, assuming a gamma loss distribution and power aversion function. Premiums are standardised to eliminate location and scale effects. The standardised ToP and its equilibrium value are, respectively,

$$T_\ell^* \equiv \frac{T_\ell - \mu_x}{\sigma_x} = \sigma_\phi \text{cor}\{x, \phi_\ell(u)\} , \quad T_*^* \equiv \frac{T_* - \mu_x}{\sigma_x} ,$$

equal to the standard deviation multiple of the premium above the mean loss. Calculations in this section show that the ToP and its equilibrium value increase with right skewness of the loss distribution, consistent with results noted in §4 and §6. In addition a higher overall aversion to satiation error increases the ToP when loss appetite is low and the dominant concern is on underpricing and large loss severities. Vice versa when loss appetite is high.

Assume a gamma loss distribution with density

$$f(x) \propto x^{\alpha-1} \exp(-x/\beta) , \quad x \geq 0 , \quad \alpha, \beta > 0$$

and a power aversion function

$$\phi(v) = nv^{n-1} , \quad 0 \leq v \leq 1 , \quad n \geq 1 .$$

For the gamma loss distribution, increasing the shape parameter α reduces right skewness, with $\alpha \rightarrow \infty$ yielding normality. In addition β is a scale parameter. Assume $\alpha = 2$ and $\beta = 1$ as a base case. The mean and standard deviation of x are $\mu_x = \alpha\beta = 2$ and $\sigma_x = \sqrt{\alpha}\beta = \sqrt{2}$, respectively. Similar to the illustration in §2, assume $n = 5$ for the aversion function, hence the overall aversion to pricing error is $\sigma_\phi = (n-1)/\sqrt{2n-1} = 4/3$. An illustration of the aversion or penalty weights is shown in the top left panel of Figure 2. Parameter values for the loss distribution and aversion function are subsequently varied from the base case to assess their impact on the ToP and its equilibrium value.

The solid line in the top right panel in Figure 2 graphs standardised ToP T_ℓ^* against loss appetite ℓ . As noted in §3, T_ℓ^* is decreasing in ℓ . The maximum T_ℓ^* is 1.3 at $\ell = 0$, corresponding to the expected maximum loss over 5 copies. The minimum T_ℓ^* is -0.8 at $\ell = 1$, corresponding to the expected minimum. In addition $T_\ell^* > 0$ or equivalently $T_\ell > \mu_x$, a positive premium loading, for $\ell \leq 0.75$, and $T_\ell < \mu_x$ for $\ell > 0.75$. The ToP exceeds the mean loss over most loss appetites in this case since the gamma loss distribution is right skewed, hence positive deviations from loss appetite have a greater impact on the ToP compared to negative deviations.

The standardised equilibrium ToP T_*^* is also shown in the top right panel as the intersection between T_ℓ^* and standardised VaR $(V_\ell - \mu_x)/\sigma$, the latter indicated by the dotted line. Note $T_*^* > 0$, hence the equilibrium ToP exceeds

the mean loss. As noted in §4, a right skewed loss distribution yields an equilibrium ToP exceeding the mean loss, since aversion adjustment has a greater impact on the right tail compared to the left tail.

The bottom left panel in Figure 2 demonstrates that reducing the right skewness of the loss distribution, from increasing the parameter α , reduces the ToP and its equilibrium value. As noted from (9) in §6, reducing right skewness decreases the loading compared to the discount, relative to the mean loss, yielding a smaller ToP. A similar discussion in §4 applies to the equilibrium ToP.

The bottom right panel in Figure 2 shows the impact of increasing overall aversion to satiation error, by increasing n . Increasing n places higher weight on extreme loss outcomes (both large and small), and yields a higher equilibrium ToP for a right skewed loss distribution (see §4). On the other hand the ToP only increases with n when the loss appetite is small to moderate, where large loss outcomes have dominant importance compared to small loss outcomes. For a high loss appetite, higher n reduces the ToP as the focus shifts to smaller loss outcomes.

8. Conclusion

The ToP generalises loss aversion reserves, distortion premiums and spectral risks by imposing increasing concern on both smaller and larger loss severities. The formulation is consistent with cumulative prospect theory and the minimisation of pricing error in a competitive market. The loss appetite controls the relative emphasis on small and large severities. Zero loss appetite yields the loss aversion reserve (distortion premium or spectral risk), with sole concern on large severities, while maximum loss appetite yields the dual.

Examples of the ToP using various aversion functions yield two-sided extensions of the well known VaR, conditional-tail-expectation and expected-maximal-loss.

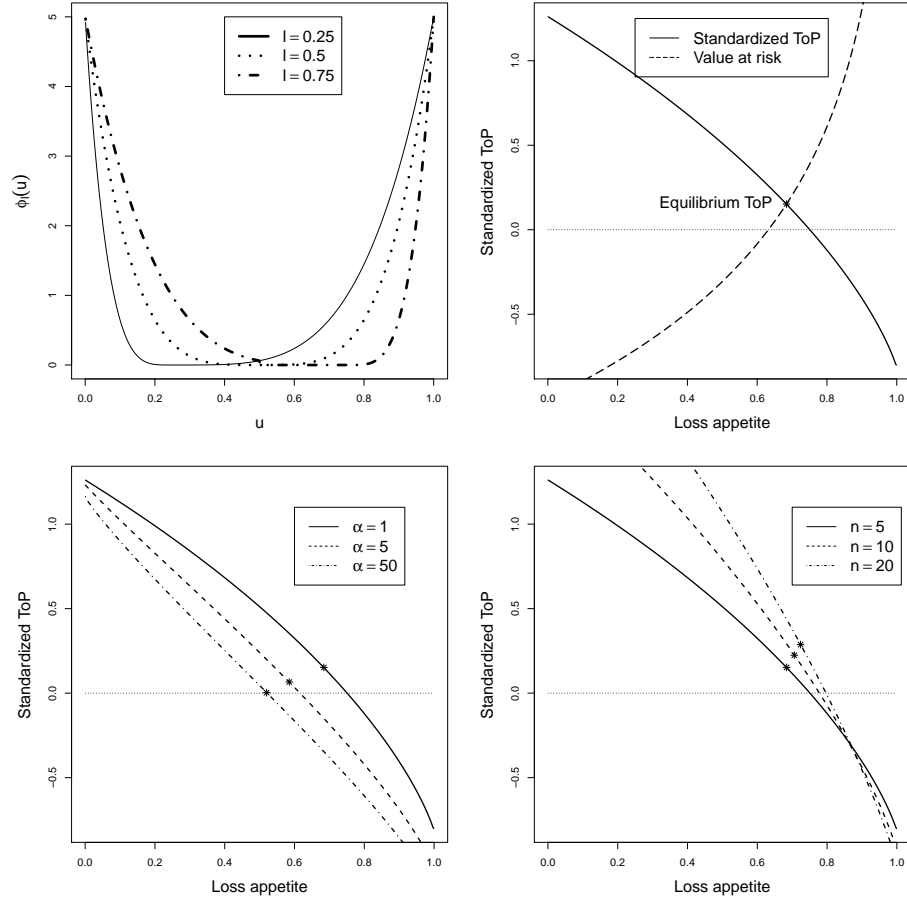


Figure 2: The top left panel plots penalty weights $\phi_\ell(u)$ against u for various loss appetites ℓ . The top right panel plots standardised ToP T_ℓ^* and standardised VaR $(V_\ell - \mu_x)/\sigma_x$ against loss appetite ℓ . The bottom left and right panels plot T_ℓ^* against ℓ for varying α and n , respectively. Equilibrium ToPs in the bottom panels are indicated by asterisks.

References

- Acerbi, C. (2002). Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking & Finance* 26(7), 1505–1518.
- Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). COHERENT MEASURES OF RISK. *Mathematical Finance* 9(3), 203–228.
- Choo, W. and P. De Jong (2009). Loss reserving using loss aversion functions. *Insurance: Mathematics and Economics* 45(2), 271–277.
- Denuit, M., J. Dhaene, M. Goovaerts, and R. Kaas (2005). Actuarial theory for dependent risks: Measures, orders and models. *John Wiley & Sons*.
- Deprez, O. and H. Gerber (1985). On convex principles of premium calculation* 1. *Insurance: Mathematics and Economics* 4(3), 179–189.
- Dhaene, J., M. Denuit, M. Goovaerts, R. Kaas, and D. Vyncke (2002). The concept of comonotonicity in actuarial science and finance: theory. *Insurance: Mathematics and Economics* 31(1), 3–33.
- Furman, E. and R. Zitikis (2008). Weighted premium calculation principles. *Insurance: Mathematics and Economics* 42(1), 459–465.
- Gerber, H. (1985). On additive principles of zero utility. *Insurance: Mathematics and Economics* 4(4), 249–251.
- Heilpern, S. (2003). A rank-dependent generalization of zero utility principle. *Insurance: Mathematics and Economics* 33(1), 67–73.
- Kaluszka, M. and M. Krzeszowiec (2011). Pricing insurance contracts under cumulative prospect theory. *Insurance: Mathematics and Economics*.
- McNeil, A., R. Frey, and P. Embrechts (2005). *Quantitative risk management*. Princeton University Press.
- Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization* 3(4), 323–343.
- Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5(4), 297–323.
- Van der Hoek, J. and M. Sherris (2001). A class of non-expected utility risk measures and implications for asset allocations. *Insurance: Mathematics and Economics* 28(1), 69–82.
- Van Heerwaarden, A. and R. Kaas (1992). The dutch premium principle. *Insurance: Mathematics and Economics* 11(2), 129–133.
- Van Heerwaarden, A., R. Kaas, and M. Goovaerts (1989). Properties of the esscher premium calculation principle. *Insurance: Mathematics and Economics* 8(4), 261–267.

- Wang, S. (1996). Premium Calculation by Transforming the Premium Layer Density. *ASTIN Bulletin* 26(1), 71–92.
- Wang, S. (2000). A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance* 67(1), 15–36.
- Young, V. (2004). Premium principles. *Encyclopedia of Actuarial Science*.

Chapter 6

Integrating proposed tools and future research areas

6.1 Introduction

As mentioned and briefly discussed in chapter 1, proposed tools in this thesis employ consistent concepts such as VaRs, layers and distortion risks and hence combine to form a coherent quantitative risk management framework. This integration is detailed in this chapter. Future research areas to expand proposed tools are also outlined and discussed.

6.2 Integrated quantitative risk management

Consider an organisation facing multiple, dependent random losses. For example a bank may suffer credit default losses from different portfolios which may be correlated particularly during a severe economic downturn. An insurer may be exposed to losses from multiple lines of business such as motor and property which are typically unrelated except when a natural catastrophe strikes. Insurance and financial companies are typically of interest, although the same discussion applies to any scenario involving risk and dependence.

Suppose the organisation aims to analyze risks arising from its losses, and manage the risks by holding capital buffers, purchasing reinsurance, hedging, exiting or expanding specific portfolios, and setting tolerance limits. Proposed tools in this thesis combine to form a consistent, holistic and coherent

analytical and management framework as outlined in the following subsections. Assume either sufficient data is available or probability distributions are known. Consider a single time period such as a month or a year.

6.2.1 Calculating individual and aggregate risks

Risks values are first attached to individual and aggregate losses. Risks are calculated using distortion: the increase in expected value by moving from the original to distorted probability distribution. Distortion is an established coherent risk measurement approach as mentioned in chapter 1.

Risk values provide an initial, overall indication of the riskiness and diversification¹ in the portfolio. Calculated risks and diversification are subsequently analyzed and managed using proposed tools described below.

6.2.2 Analyzing risk behaviour across layers

Once individual risks are calculated, they are decomposed and spread across layers of underlying losses using risk densities proposed in chapter 3. Risk densities indicate risk contributions of each layer and are hence useful when analyzing large individual risks. For example large individual risks may arise from extreme, rare layers or low, attritional layers. Layers are expressed in VaRs, providing a “common language” across individual loss distributions with different shape and scale. Note the analysis still assumes individual losses are “standalone” and ignores diversification effects from aggregation.

Mean densities in chapter 3 complement risk densities by indicating contributions of loss layers to mean values. Risk ratios, the ratio between risk and mean densities, indicate the risk of each loss layer relative to its mean.

6.2.3 Analyzing diversification and dependence

Analyzing diversification is important to understand the risk reduction when individual losses are aggregated. Chapter 4 constructs systematic risk densities akin to standalone risk densities, indicating post-diversification or systematic risks of layers forming individual losses. Comparing systematic with standalone risk densities reveals the extent of risk diversification in each layer and its contribution to overall diversification. Diversification is hence

¹The difference between the sum of individual risks and aggregate risk.

decomposed in two stages: first across individual losses, then across layers of individual losses.

Chapter 4 also shows that local dependence between individual and aggregate losses drives the risk diversification in each layer, and the local dependence relates to layer dependence set up in chapter 2. Hence analyzing risk diversification is akin to analyzing dependence structures. Strong diversification arises when risky layers of individual losses are weakly dependent with the aggregate loss.

6.2.4 Formulating risk management actions

The above analysis of risk and diversification behaviour leads to actions to minimise risk or maximise diversification. For example, loss layers with large systematic risk are reinsured or hedged. These layers have large standalone risks and strong dependence with the aggregate loss, and can be identified from risk densities and layer dependence curves. Loss layers with large standalone risk but weak dependence with the aggregate loss have small contributions to aggregate risk, and hence may be retained or even expanded.

Chapter 4 also discusses how the framework of mean and risk densities is used to analyse the credit rating of debt tranches, transform loss distributions using proportional and excess-of-loss reinsurance, and set capital buffers under various guiding principles. Derived actions are framed in VaRs since mean and risk densities are defined over VaR layers. VaR based actions adjust to the shape and scale of probability distributions, and are aligned to the widespread use of VaRs in finance and insurance.

6.2.5 Reflecting upside risk

Finally, upside risk may be important due to competition, opportunity costs of conservatism, or limited capital available. Tradeoff premiums defined in chapter 5 extend distortion risk measurement by explicitly allowing for both upside and downside risks defined relative to a selected loss appetite. Loss appetites are defined in VaR terms, consistent with other proposed methods in this thesis.

Revising risk densities in chapters 3 and 4 using two-sided distortion risk measurement, i.e. tradeoff premiums, leads to an extended quantitative risk management framework where upside risks are managed in conjunction with downside risks. This development is not considered in this thesis and is a

future research area. Other future research areas are described in the next section.

6.3 Areas of future research

The following outlines potential areas of future research to extend proposed methods in this thesis.

6.3.1 Time series extension

It is common to model risk and dependence over multiple time periods. One such model is the Generalised Autoregressive Conditional Heteroskedasticity model with Dynamic Conditional Correlation, or GARCH–DCC (Engle (2002), Engle (2001)). This multivariate time series model captures cyclical volatility and dependence. A period of high volatility and dependence may characterise for example turmoil in global stock markets during 2008.

The proposed quantitative risk management framework in this thesis can be extended and applied across multiple time periods. Risk densities and layer dependence curves reveal risk, volatility and dependence behaviour at various points in time. Certain loss layers may be systematically risky in periods of stability, and new loss layers may emerge as being systematically most risky during turmoil. Optimal risk management actions will hence vary, depending on the phase of volatility and dependence cycles.

6.3.2 Layer dependence in the multivariate case

Layer dependence defined in this thesis is a univariate function summarising the dependence structure of a bivariate copula, measuring dependence between a random variable and layers of another. A natural extension is hence layer dependence for multivariate copulas.

Extended layer dependence may be multi-dimensional functions with one less dimension than the multivariate copula. Alternatively several univariate layer dependence curves may combine to characterise the dependence structure of the copula. Depending on the setup of extended layer dependence, one or more random variables may be decomposed into layers.

6.3.3 Fitting a copula to given layer dependence curves

One may specify layer dependence using past data and expert opinion. The next step, which is not covered in this thesis, is to fit a copula to the given dependence structure in order to compute the aggregate probability distribution and risk, for example. As mentioned in chapter 2, layer dependence curves do not imply unique copulas since the former represents summarised information whilst the latter contains complete information.

A general fitting approach is to first construct a copula with controllable and flexible layer dependence. Parameters of this copula are then selected to satisfy the specified layer dependence curve. The copula may be a mixture of “basic” copulas as building blocks. Alternatively consider the copula generated by the single factor model (Krupskii and Joe 2013)

$$x = f + \epsilon_1 \quad , \quad y = f + \epsilon_2$$

where f , ϵ_1 and ϵ_2 are independent random variables. The common factor f generates and controls dependence between x and y . For example if f has a highly right skewed probability distribution then large values of x and y are dominated by f instead of ϵ_1 and ϵ_2 , implying strong layer dependence in the upper tail. The probability distribution of f can be iteratively solved to achieve the given layer dependence curve, whilst ϵ_1 and ϵ_2 can be standard Gaussian for example.

6.3.4 Two-sided risk densities

The previous section suggests extending the proposed quantitative risk management framework to consider upside in addition to downside risks, using tradeoff premiums. A specific extension relates to risk densities, which currently only capture downside risk. As tradeoff premiums are additive over comonotonic random variables, one can calculate the tradeoff premium for each VaR layer of a random loss which then add to the overall tradeoff premium. The resulting “tradeoff premium density” is a two-sided risk density.

The extension applies to both standalone and systematic risk densities. Varying the loss appetite shifts the relative focus on upside and downside risks, and a zero loss appetite leads to original one-sided risk densities.

Bibliography

- Acerbi, C. (2002). Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking & Finance* 26(7), 1505–1518.
- Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). COHERENT MEASURES OF RISK. *Mathematical Finance* 9(3), 203–228.
- Bairamov, I., S. Kotz, and T. Kozubowski (2003). A new measure of linear local dependence. *Statistics: A Journal of Theoretical and Applied Statistics* 37(3), 243–258.
- Bjerve, S. and K. Doksum (1993). Correlation curves: Measures of association as functions of covariate values. *The Annals of Statistics*, 890–902.
- Buch, A. and G. Dorfleitner (2008). Coherent risk measures, coherent capital allocations and the gradient allocation principle. *Insurance: Mathematics and Economics* 42(1), 235–242.
- Campana, A. and P. Ferretti (2014). Risk measures in solvency regulation: Reinsurance layers and unexpected loss. *Applied Mathematical Sciences* 8(116), 5783–5794.
- Chernobai, A. S., S. T. Rachev, and F. J. Fabozzi (2008). *Operational risk: a guide to Basel II capital requirements, models, and analysis*, Volume 180. John Wiley & Sons.
- Choo, W. and P. De Jong (2009). Loss reserving using loss aversion functions. *Insurance: Mathematics and Economics* 45(2), 271–277.
- Choo, W. and P. De Jong (2010). Determining and allocating diversification benefits for a portfolio of risks. *Astin Bulletin* 40(1), 257–269.
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues.
- Cummins, J. (2000). Allocation of capital in the insurance industry. *Risk Management and Insurance Review* 3(1), 7–27.

- Denault, M. (2001). Coherent allocation of risk capital. *Journal of risk* 4, 1–34.
- Denuit, M. and J. Dhaene (2012). Convex order and comonotonic conditional mean risk sharing. *Insurance: Mathematics and Economics*.
- Denuit, M., J. Dhaene, M. Goovaerts, and R. Kaas (2005). Actuarial theory for dependent risks: Measures, orders and models. *John Wiley & Sons*.
- Dhaene, J., M. Denuit, M. Goovaerts, R. Kaas, and D. Vyncke (2002). The concept of comonotonicity in actuarial science and finance: theory. *Insurance: Mathematics and Economics* 31(1), 3–33.
- Dhaene, J., M. Goovaerts, and R. Kaas (2003). Economic capital allocation derived from risk measures. *North American Actuarial Journal*, Vol. 7, No. 2, pp. 44–59, 2003.
- Dhaene, J., A. Tsanakas, E. Valdez, and S. Vanduffel (2012). Optimal capital allocation principles. *Journal of Risk and Insurance* 79(1), 1–28.
- Dowd, K. and D. Blake (2006). After var: The theory, estimation, and insurance applications of quantile-based risk measures. *Journal of Risk and Insurance* 73(2), 193–229.
- Duffie, D. and N. Garleanu (2001). Risk and valuation of collateralized debt obligations. *Financial Analysts Journal* 57(1), 41–59.
- Durante, F., J. Fernández-Sánchez, and R. Pappadà (2014). Copulas, diagonals, and tail dependence. *Fuzzy Sets and Systems*.
- Durante, F., R. Mesiar, and C. Sempì (2006). On a family of copulas constructed from the diagonal section. *Soft Computing* 10(6), 490–494.
- Eling, M., H. Schmeiser, and J. Schmit (2007). The solvency ii process: Overview and critical analysis. *Risk Management and Insurance Review* 10(1), 69.
- Embrechts, P. and M. Hofert (2014). Statistics and quantitative risk management for banking and insurance. *Annual Review of Statistics and Its Application* 1, 493–514.
- Engelmann, B. and R. Rauhmeier (2006). *The Basel II risk parameters*. Springer.
- Engle, R. (2001). Garch 101: The use of arch/garch models in applied econometrics. *Journal of economic perspectives*, 157–168.

- Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* 20(3), 339–350.
- Evans, J. (2001). Exposure rating loss layers: Unifying the property perspective of severity with the liability perspective of frequency. *Including the Reserves Call Papers*, 493.
- Feller, W. (2008). *An introduction to probability theory and its applications*, Volume 2. John Wiley & Sons.
- Finger, R. J. (1976). Estimating pure premiums by layer-an approach. *PCAS LXIII* 2.
- Fredricks, G. A. and R. B. Nelsen (1997). Copulas constructed from diagonal sections. In *Distributions with Given Marginals and Moment Problems*, pp. 129–136. Springer.
- Furman, E. and R. Zitikis (2008a). Weighted premium calculation principles. *Insurance: Mathematics and Economics* 42(1), 459–465.
- Furman, E. and R. Zitikis (2008b). Weighted risk capital allocations. *Insurance: Mathematics and Economics* 43(2), 263–269.
- Genest, C. and L.-P. Rivest (1993). Statistical inference procedures for bivariate archimedean copulas. *Journal of the American statistical Association* 88(423), 1034–1043.
- Gerber, H. (1985). On additive principles of zero utility. *Insurance: Mathematics and Economics* 4(4), 249–251.
- Hartmann, P., S. Straetmans, and C. G. De Vries (2004). Asset market linkages in crisis periods. *Review of Economics and Statistics* 86(1), 313–326.
- Heilpern, S. (2003). A rank-dependent generalization of zero utility principle. *Insurance: Mathematics and Economics* 33(1), 67–73.
- Hillson, D. (2002). Extending the risk process to manage opportunities. *International Journal of project management* 20(3), 235–240.
- Hillson, D. (2003). *Effective opportunity management for projects: Exploiting positive risk*. CRC Press.
- Hogg, R. V. and S. A. Klugman (2009). *Loss distributions*, Volume 249. John Wiley & Sons.
- Holland, P. W. and Y. J. Wang (1987). Dependence function for continuous bivariate densities. *Communications in Statistics-Theory and Methods* 16(3), 863–876.

- Hsieh, D. A. (1988). The statistical properties of daily foreign exchange rates: 1974–1983. *Journal of international economics* 24(1), 129–145.
- Hürlimann, W. (1998). On distribution-free safe layer-additive pricing. *Insurance: Mathematics and Economics* 22(3), 277–285.
- Joe, H. (1997). *Multivariate models and dependence concepts*, Volume 73. CRC Press.
- Jones, M. (1996). The local dependence function. *Biometrika* 83(4), 899–904.
- Kalkbrener, M. (2005). AN AXIOMATIC APPROACH TO CAPITAL ALLOCATION. *An International Journal of Mathematics, Statistics and Financial Economics* 15(3), 425–437.
- Kaluszka, M. and M. Krzeszowiec (2011). Pricing insurance contracts under cumulative prospect theory. *Insurance: Mathematics and Economics*.
- Kaufmann, R., A. Gadmer, and R. Klett (2001). Introduction to dynamic financial analysis. *Astin Bulletin* 31(01), 213–249.
- King, P. and H. Tarbert (2011). Basel iii: an overview. *Banking & Financial Services Policy Report* 30(5), 1–18.
- Kolb, R. (2010). *Lessons from the financial crisis: Causes, consequences, and our economic future*, Volume 12. John Wiley & Sons.
- Kotz, S. and S. Nadarajah (2000). *Extreme value distributions*, Volume 31. World Scientific.
- Krupskii, P. and H. Joe (2013). Factor copula models for multivariate data. *Journal of Multivariate Analysis* 120, 85–101.
- Ladoucette, S. A. and J. L. Teugels (2006). Analysis of risk measures for reinsurance layers. *Insurance: Mathematics and Economics* 38(3), 630–639.
- Lamoureux, C. G. and W. D. Lastrapes (1990). Heteroskedasticity in stock return data: volume versus garch effects. *The Journal of Finance* 45(1), 221–229.
- Lee, Y.-S. (1988). The mathematics of excess of loss coverages and retrospective rating—a graphical approach. *PCAS LXXV* 49.
- Luenberger, D. (1998). *Investment Science*. Oxford University Press, USA.
- Mandel, B., D. Morgan, and C. Wei (2012). The role of bank credit enhancements in securitization. *Federal Reserve Bank of New York Economic Policy Review* 18(2), 35–46.

- McNeil, A., R. Frey, and P. Embrechts (2005). *Quantitative risk management*. Princeton University Press.
- Miccolis, R. S. (1977). On the theory of increased limits and excess of loss pricing. *PCAS LXIV*, 27.
- Myers, S. and J. Read Jr (2001). Capital allocation for insurance companies. *Journal of Risk and Insurance*, 545–580.
- Nelson, R. (1999). An Introduction to Copulas. *Lecture Notes in Statistics* 139.
- Nocco, B. W. and R. M. Stulz (2006). Enterprise risk management: theory and practice. *Journal of Applied Corporate Finance* 18(4), 8–20.
- Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association* 84(406), 487–493.
- Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization* 3(4), 323–343.
- Rockafellar, R. and S. Uryasev (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance* 26(7), 1443–1471.
- Rodriguez, J. C. (2007). Measuring financial contagion: A copula approach. *Journal of Empirical Finance* 14(3), 401–423.
- Salzmann, R. E. (1963). Rating by layer of insurance. *PC, 4S*, 15–26.
- Shapley, L. (1974). Values of non-atomic games.
- Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance* 19(3), 425–442.
- Sherris, M. (2006). Solvency, capital allocation, and fair rate of return in insurance. *Journal of Risk and Insurance* 73(1), 71–96.
- Tasche, D. (2007). Capital allocation to business units and sub-portfolios: the euler principle. *Arxiv preprint arXiv:0708.2542*.
- Tsanakas, A. and N. Christofides (2006). Risk exchange with distorted probabilities. *Astin Bulletin* 36(1), 219.
- Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5(4), 297–323.
- Van Gulick, G., A. De Waegenaere, and H. Norde (2012). Excess based allocation of risk capital. *Insurance: Mathematics and Economics* 50(1), 26–42.

- Van Heerwaarden, A., R. Kaas, and M. Goovaerts (1989). Properties of the esscher premium calculation principle. *Insurance: Mathematics and Economics* 8(4), 261–267.
- Venter, G. (2004). Capital Allocation Survey with Commentary. *NORTH AMERICAN ACTUARIAL JOURNAL* 8(2), 96–125.
- Venter, G. G. (2002). Tails of copulas. In *Proceedings of the Casualty Actuarial Society*, Volume 89, pp. 68–113.
- Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics* 17(1), 43–54.
- Wang, S. (1996). Premium Calculation by Transforming the Premium Layer Density. *ASTIN Bulletin* 26(1), 71–92.
- Wang, S. (1998). An actuarial index of the right-tail risk. *North American Actuarial Journal* 2(2), 88–101.
- Wang, S. (2000). A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance* 67(1), 15–36.
- Young, V. (2004). Premium principles. *Encyclopedia of Actuarial Science*.