Zitikis, Ricardas (Ed.); Ren, Jiandong (Ed.); Sendova, Kristina (Ed.)

Book — Published Version

Risk, ruin and survival: Decision making in insurance and finance

Risk, Ruin and Survival

Decision Making in Insurance and Finance

Edited by

Ricardas Zitikis, Jiandong Ren and Kristina Sendova

Printed Edition of the Special Issue Published in *Risks*

www.mdpi.com/journal/risks

MDPI

Risk, Ruin and Survival

Risk, Ruin and Survival

Decision Making in Insurance and Finance

Special Issue Editors

Ricardas Zitikis
Jiandong Ren
Kristina Sendova

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin

Special Issue Editors
Ricardas Zitikis
Western University
Canada

Jiandong Ren
Western University
Canada

Kristina Sendova
Western University
Canada

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Risks* (ISSN 2227-9091) (available at: https://www.mdpi.com/journal/risks/special_issues/Risk_Ruin_Survival).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Article Number*, Page Range.

ISBN 978-3-03928-516-7 (Pbk)
ISBN 978-3-03928-517-4 (PDF)

Contents

About the Special Issue Editors . . . . . **vii**

Jiandong Ren, Kristina Sendova, and Ričardas Zitikis
Special Issue "Risk, Ruin and Survival: Decision Making in Insurance and Finance"
Reprinted from: *Risks* **2019**, *7*, 96, doi:10.3390/risks7030096 . . . . . **1**

Huong Dieu Dang
National Culture and Corporate Rating Migrations
Reprinted from: *Risks* **2018**, *6*, 130, doi:10.3390/risks6040130 . . . . . **9**

Francesca Greselin, Fabio Piacenza and Ričardas Zitikis
Practice Oriented and Monte Carlo Based Estimation of the Value-at-Risk for Operational Risk Measurement
Reprinted from: *Risks* **2019**, *7*, 50, doi:10.3390/risks7020050 . . . . . **37**

Nadezhda Gribkova and Ričardas Zitikis
A User-Friendly Algorithm for Detecting the Influence of Background Risks on a Model
Reprinted from: *Risks* **2018**, *6*, 100, doi:10.3390/risks6030100 . . . . . **57**

Erwan Koch
Spatial Risk Measures and Rate of Spatial Diversification
Reprinted from: *Risks* **2019**, *7*, 52, doi:10.3390/risks7020052 . . . . . **69**

Shuanming Li and Yi Lu
On the Moments and the Distribution of Aggregate Discounted Claims in a Markovian Environment
Reprinted from: *Risks* **2018**, *6*, 59, doi:10.3390/risks6020059 . . . . . **95**

Xin Liu, Jiang Wu, Chen Yang and Wenjun Jiang
A Maximal Tail Dependence-Based Clustering Procedure for Financial Time Series and Its Applications in Portfolio Selection
Reprinted from: *Risks* **2018**, *6*, 115, doi:10.3390/risks6040115 . . . . . **111**

Mohamed Amine Lkabous and Jean-François Renaud
A VaR-Type Risk Measure Derived from Cumulative Parisian Ruin for the Classical Risk Model
Reprinted from: *Risks* **2018**, *6*, 85, doi:10.3390/risks6030085 . . . . . **137**

Sooie-Hoe Loke and Enrique Thomann
Numerical Ruin Probability in the Dual Risk Model with Risk-Free Investments
Reprinted from: *Risks* **2018**, *6*, 110, doi:10.3390/risks6040110 . . . . . **149**

Fouad Marri, Franck Adékambi and Khouzeima Moutanabbir
Moments of Compound Renewal Sums with Dependent Risks Using Mixing Exponential Models
Reprinted from: *Risks* **2018**, *6*, 86, doi:10.3390/risks6030086 . . . . . **163**

Vadim Semenikhine, Edward Furman and Jianxi Su
On a Multiplicative Multivariate Gamma Distribution with Applications in Insurance
Reprinted from: *Risks* **2018**, *6*, 79, doi:10.3390/risks6030079 . . . . . **181**

About the Special Issue Editors

Ricardas Zitikis, Ph.D., is a full-time researcher and teacher affiliated with the School of Mathematical and Statistical Sciences at Western University, Canada. He consults for non-profit and for-profit entities, and is an author and co-author of more than 150 peer-reviewed publications. Over the past several decades, he has successfully completed numerous research projects for the Society of Actuaries, Casualty Actuarial Society, Natural Sciences and Engineering Research Council of Canada, and the Canadian national research organization for Mathematics of Information Technology and Complex Systems. His supervised students have held a variety of academic and industrial positions, including internships, specializing in topics such as InsurTech, FinTech, machine learning, and artificial intelligence.

Jiandong Ren, Ph.D., is a professor in the Department of Statistical and Actuarial Sciences at Western University. He has a M.S. in electrical engineering from Southwest Jiaotong University, China. His Ph.D. is in risk management and insurance from Temple University, USA. Dr. Ren is an associate member of the Casualty Actuarial Society and is an author of more than 30 refereed publications. He has supervised many M.S. and Ph.D. students, who are currently in a variety of academic and industrial positions.

Kristina Sendova, Ph.D., has a B.S. in Applied Mathematics and a dual M.S. in Mathematical Economics and Applied Statistics from Sofia University, Bulgaria. Her Ph.D. is in actuarial science from the University of Waterloo, Canada. After graduating, Prof. Sendova completed a post-doc at the University of Toronto and then started her employment at Western University, where she has been since 2005. Prof. Sendova has been Undergraduate and then Graduate Chair of the Department of Statistical and Actuarial Sciences at Western University, Canada, and is now its Chair. For five years, she received a University Faculty Award from the Natural Sciences and Engineering Council of Canada. She is an active researcher and is also an Associate Editor of the *Journal of Applied Mathematics and Computation*.

risks

MDPI

Editorial

Special Issue "Risk, Ruin and Survival: Decision Making in Insurance and Finance"

Jiandong Ren, Kristina Sendova and Ričardas Zitikis *,†

Department of Statistical and Actuarial Sciences, Western University, London, ON N6A 5B7, Canada
* Correspondence: rzitikis@uwo.ca
† Lead Guest Editor.

Received: 30 July 2019; Accepted: 21 August 2019; Published: 7 September 2019

It has been six years since Editor-in-Chief Steffensen (2013) wrote in his editorial that "to *Risks* inclusiveness, inter-disciplinarity, and open-mindedness is the very starting point." An one-liner of his editorial also stated that "[w]hat is complicated is not necessarily insightful and what is insightful is not necessarily complicated: *Risks* welcomes simple manuscripts that contribute with insight, outlook, understanding and overview."

This philosophy has contributed immensely to *Risks* becoming one of the most interesting, enlightening, and inspiring journals dealing with various facets of risk. It was, therefore, most humbling and exciting to accept the invitation by *Risks* to organize a special issue, and we decided to cover three inevitable stages of our lives: taking risks, sometimes getting ruined, and hopefully surviving. Success along this path requires well thought out strategy and tactics, which involve measuring and modeling risks, selecting and testing decisions. These tasks benefit greatly from exploratory simulation studies and, needless to say, from well-tested ancient wisdom. As Master Sun (Sun-tzu 2007) said during the Warring States Period more than 2000 years ago,

> He who knows the enemy and himself
> Will never in a hundred battles be at risk;
> He who does not know the enemy but knows himself
> Will sometimes win and sometimes lose;
> He who knows neither the enemy nor himself
> Will be at risk in every battle. (Sun-tzu 2007, p. 87)

We are blessed to have witnessed the immense interest that the Special Issue has generated, and we thank all the authors whose contributions have, or have not, been selected for this Issue, and we are of course grateful to all the referees for their effective and tireless work during the selection process.

We are indebted to the entire Editorial and Production Teams of *Risks* for making our work on the Issue as easy as one-two-three, but we would nevertheless be remiss if we did not especially mention Managing Editor Janine Li for her constant and frank advice that has guided our work.

We shall next introduce each of the ten papers, in the usual alphabetical order based on authorship, with the intent of generating enough interest among the *Risks* readership to look at the papers in depth and, in turn, to write up their thoughts and suggestions as follow-up contributions to *Risks*. To ensure the accuracy of each of the ten introductions, we have sought and received generous and much appreciated feedback from the authors.

Dang (2018):

Investors from different cultures exhibit different cognitive biases (Afego 2018). Behavioral pitfalls such as overreaction and herding are related to uncertainty avoidance and collectivism (embeddedness) dimensions, respectively (Chang and Lin 2015; Hofstede et al. 2010). Empirical studies document the effects of culture values on investor sentiments and market mood, and the latter makes an impact on

credit ratings. President of the European Commission José Manuel Barroso remarked at the European Parliament in May 2010 that credit ratings are "too cyclical, too reliant on the general market mood rather than on fundamentals—regardless of whether market mood is too optimistic or too pessimistic." The informal constraints that stem from culture have made an extensive impact on daily behavior, and this impact extends far beyond formal laws (North 1990). Culture directly influences managerial decision making, thereby indirectly influencing corporate risk taking (Li et al. 2013), capital structure (Chui et al. 2002), debt maturity (Zhen et al. 2012), cash management (Chang and Noorbakhsh 2009; Ramirez and Tadesse 2009), corporate investment (Shao et al. 2013), dividend policy (Bae et al. 2012; Fidrmuc and Jacob 2010; Shao et al. 2010), financial disclosures, financial report quality (Gray 1988; Hope 2003), and the degree of earnings management (Desender et al. 2011; Han et al. 2010). Guiso et al. (2016) further suggest that cultural differences between Germany and Greece hindered Greece's negotiations to avoid a default. There is increasing recognition of the role that culture plays in corporate and sovereign credit risk. An interesting question that has not been addressed in the literature is whether and how culture affects changes in credit risk, specifically, in credit ratings. Dang (2018) offers a cross-disciplinary explanation of this phenomenon and establishes a link between culture and corporate rating migration. In particular, Dang (2018) shows that studying culture helps to enrich our understanding of credit rating decisions, which in turn can be helpful in developing predictive models of corporate rating changes across countries.

Greselin et al. (2019):

Quantile, also known as percentile and value-at-risk (VaR), is a fundamental quantity in many areas of research and practice. It is quite often thought to be an easy parameter to estimate, and this is indeed true in the sense that any quantile can easily be calculated from any data set. Deriving its confidence intervals, however, has turned out to be a Herculean task: obtaining limiting distributions require mathematically elegant (e.g., Beirlant et al. 2004; Csörgő 1983; De Haan and Ferreira 2006) but nevertheless practically-formidable assumptions; and resampling techniques (e.g., Shao and Tu 1995) designed to circumvent some of the difficulties quickly run into a host of challenges (e.g., Bickel and Sakov 2008). Yet, regulatory frameworks in insurance and banking require estimating pre-specified quantiles and then use them as risk measures. Attempts to utilize optimal, or nearly optimal, results derived by mathematical statisticians under optimal, or nearly optimal, assumptions quickly become disconcerting, due to the need to explain to the managers and shareholders, among others, why and when the assumptions hold. This is when decision-makers posit the formidable challenge to researchers to come up with practically-appreciable confidence intervals under practically-justifiable conditions. Obviously, "practically appreciable" does not necessarily mean asymptotically accurate—the definition of "practically appreciable" depends on managers, decision-makers, and others involved in making the company profitable, or at least safeguarding it from defaulting. It is this viewpoint that has guided the research of Greselin et al. (2019) on VaR in the context of operational risk (OpRisk) measurement.

Gribkova and Zitikis (2018):

The notion of systemic, also known as background, risk permeates various research areas: insurance, finance, engineering, system security, and so on. It is a multifaceted topic, with various methods and techniques available for detection, analysis, and decision making. Many engineering-related studies employ techniques in the frequency domain, while Gribkova and Zitikis (2018) pursue the task in the time domain. The latter paper is a part of the tetralogy by Gribkova and Zitikis (2018, 2019a, 2019b, 2019c) who develop a comprehensive classification and testing methodology for dealing with potential effects of systemic risk on systems at their input and/or output stages. The importance of such research is due to the fact, among other reasons, that even though the decision maker may be aware of the existence of systemic risk and would thus incorporate it into the statistical model, the decision maker cannot be sure that the resulting model complexity is really justified. Indeed, the effects of systemic risk on the system may be statistically insignificant, and thus a simpler and

more tractable model could be used. The aforementioned tetralogy offers a comprehensive resolution of the problem.

Koch (2019):

An accurate assessment of the risk of extreme environmental events is essential, inter alia for the (re)insurance industry, and this is increasingly so in the context of climate change. Although the theory of risk measures has been under development for a long time, the spatial and infinite-dimensional settings have essentially been tackled only since Koch (2017), where the author introduces a notion of spatial risk measure that explicitly considers the spatial region and the cost field generated by a specific hazard over the region. Koch (2017) also proposes a related set of axioms that describes how the spatial risk is expected to evolve with respect to spatial features of the region, such as its location and size. Results of this kind facilitate the quantification of a number of parameters of interest for (re)insurance companies, including the rate of spatial diversification. In the present research, Koch (2019) makes further advances in the development of the concept of spatial risk and corresponding axioms, and supplements them with in-depth explanations of their uses in actuarial science and practice. In the case of a general cost field, Koch (2019) specifies conditions that give the rate of spatial diversification for spatial risk measures linked with expectation, variance, Value-at-Risk (VaR) and Expected Shortfall (ES). These conditions are then refined when the cost field is a function of max-stable random fields, which are well suited for modeling spatial extremes and are thus beneficial in practice. Finally, in Koch (2017 2019), the dependence between the individual risks are fully characterized by means of the spatial dependence structure of the cost field. Note in this regard that the classical actuarial individual and collective risk models lack the localization of risks, thus making their dependence modeling somewhat arbitrary and possibly less reliable from the practical point of view.

Li and Lu (2018):

The aggregate discounted claims for insurance portfolios are the present values of the total amounts to be paid by the company up to a certain time in the future. They depend on the arrival times (frequency) and sizes (severity) of the claims, as well as on the forces of interest for discounting. The aggregate discounted claims may also be influenced by the background or environmental conditions; for example, weather or climate conditions may impact property and casualty insurance claims. Hence, it is important for the insurer to know various distributional characteristics (e.g., the average, variance, distribution, etc.) of the aggregate discounted claims, as they represent the insurer's future liabilities at present time. In light of this, Li and Lu (2018) study the distribution and moments of the aggregate discounted claims occurring in a cluster of states based on a Markovian arrival process in which the claim arrivals, claim amounts, and forces of interest are influenced by an underlying Markov process. They also examine the correlations of the aggregate discounted claims occurring under two different clusters of environmental conditions, or with two different time lengths. To illuminate the results, Li and Lu (2018) provide an illustrative numerical study based on a two-state Markov-modulated model.

Liu et al. (2018):

As illustrated by Furman et al. (2015), existing methods for measuring tail dependence may sometimes underestimate the true interdependence between extreme co-movements of risks. This might be disconcerting from the practical point of view, and has therefore given rise to the notion of maximum tail dependence (MTD). Liu et al. (2018) have adopted this notion in their real data based explorations of a portfolio selection problem initiated by De Luca and Zuccolotto (2011), who have proposed to cluster financial assets by tail dependence. In Liu et al. (2018), the tail dependence coefficient (TDC) used by De Luca and Zuccolotto (2011) is replaced by the MTD coefficient, and the corresponding techniques of clustering are developed and discussed. The obtained results suggest that

MTD-based portfolios outperform TDC-based portfolios on avoiding extremely low rates of return. The importance of this article is at least two-fold: it lends support to the MTD methodology from a practical perspective, and proposes a way to improve the performance of portfolios from the risk management perspective.

Lkabous and Renaud (2018):

In actuarial mathematics, the occupation time of an insurer's surplus process below a given threshold level is particularly critical when assessing the insurer's solvency risk (e.g., Guérin and Renaud 2017; Landriault et al. 2014). Such problems are closely related to Parisian ruin models, under which insurers are granted a grace period to re-emerge above the threshold level before a ruin is deemed to occur. Naturally, therefore, Guérin and Renaud (2017) introduce the concept of cumulative Parisian ruin, which is based on the "time spent in the red" by the underlying surplus process. The time of the cumulative Parisian ruin is the first time the surplus process stays cumulatively below a critical level longer than the pre-determined grace period. Several dynamic risk measures, which are those based on ruin-theoretic quantities, have been studied by, e.g., Trufin et al. (2011), Mitric and Trufin (2016), and Loisel and Trufin (2014). These results have, in turn, motivated Lkabous and Renaud (2018) to explore a VaR-type risk measure based on the cumulative Parisian ruin for the classical risk model. In particular, they relate their measure to other ones that have appeared in the literature, and they also verify various properties of the measure. Finally, Lkabous and Renaud (2018) perform a sensitivity analysis of their risk measure in the case of a Cramér–Lundberg process with exponential claims.

Loke and Thomann (2018):

The dual risk model exemplifies the surplus process of a company that incurs expenses at a constant rate and earns random positive gains at random times. The model is also known as the "negative claims model" because it can be obtained by negating the signs of premiums and claims in the classical Cramér-Lundberg model. A great variety of applications have been tackled using this model. To illustrate, in life insurance and pensions (e.g., Grandell 2012), continuous payments are made by the company to the policyholder, and upon the death of the latter, the gross reserve becomes available to the company as a profit. Another example concerns companies that engage in research and discovery, such as petroleum or pharmaceutical companies (e.g., Avanzi et al. 2007), with random gains corresponding to, e.g., the discoveries of oil or the development of new patents, respectively. In greater generality, many subsequent risk models have also incorporated investments with constant force of interest, which could be, e.g., investments of the entire surplus in bonds, as earlier argued by Segerdahl (1942). These arguments have motivated Loke and Thomann (2018) to further explore the problems in detail, particularly from the view of practical implementation needs, such as speedy decision making for which numerical algorithms become indispensable. Hence, the authors have put forward and examined the performance of a numerical algorithm for tackling the ruin probability in the dual risk model with risk-free investments under arbitrary gain distributions. Crucially for the algorithm, the ruin probability has been shown to satisfy an integro-differential equation, which the authors subsequently discretized and reduced to a linear matrix equation. This has enabled Loke and Thomann (2018) to efficiently compute the ruin probability for any jump distribution. Furthermore, the authors have employed an analogous numerical method to tackle other Gerber–Shiu type functions, such as the Laplace transform of the time of ruin.

Marri et al. (2018):

The problem of risk aggregation for insurance portfolios has been extensively studied from various perspectives. For example, Léveillé and Garrido (2001a, 2001b) employ renewal theory to derive closed-form expressions for the first two moments of the discounted aggregate claims, and Léveillé and Hamel (2013) study aggregate discounted payments and the expense process for medical

malpractice insurance. As is the case with many pioneering studies, they have assumed that the inter-arrival times and the claim amounts are independent. Empirical observations have revealed, however, that a departure from independence would considerably increase the practical appeal of such studies. To illustrate, in non-life insurance, the same catastrophic event (e.g., a flood or an earthquake) may lead to frequent and high losses. Inspired by such observations, Marri et al. (2018) study discounted renewal aggregate claims under full dependence structure: they allow dependence among the inter-claim times, dependence among the claim sizes, and also dependence between the inter-claim times and the claim sizes.

Semenikhine et al. (2018):

Loss distributions have been a staple of actuarial studies for many years. When modeling losses, actuaries are particularly interested in distributions that are supported on the non-negative real half-line, have positive skewness, and allow for varying degrees of heavy-tailness. One of such distributions, arguably the most prominent one in insurance applications, is the gamma distribution. When insurance losses arise from multiple business lines and need to be considered jointly due to their interdependence, multivariate extensions of the gamma distribution are naturally called upon. Although there are numerous multivariate gamma models in the literature, real data-driven applications impose significant constraints on the model choice. In particular, practitioners often wish to work with multivariate distributions that (i) admit meaningful and relevant interpretations, (ii) allow for adequate fits (marginally or jointly) to a wide range of multivariate data, (iii) possess desirable distributional properties for insurance valuation and risk management, and (iv) can be readily implemented. The multivariate gamma family proposed by Semenikhine et al. (2018) is exactly such. Although the family is exceptionally general, the authors have succeeded in thoroughly exploring its properties. In particular, Semenikhine et al. (2018) have linked the family to the multiplicative background risk model, derived an explicit formula for the distribution of the aggregate risk, specified the corresponding copula function, and determined measures of nonlinear correlation, including the index of maximal tail dependence (Furman et al. 2015).

Funding: The authors gratefully acknowledge research support from the Natural Sciences and Engineering Research Council (NSERC) of Canada, and the National Research Organization "Mathematics of Information Technology and Complex Systems" (MITACS) of Canada.

Conflicts of Interest: The authors declare no conflict of interest.

References

Afego, P. N. 2018. Index shocks, investor action and long-run stock performance in Japan: A case of cultural behaviouralism? *Journal of Behavioral and Experimental Finance* 18: 54–66.

Avanzi, B., H. U. Gerber, and E. S. W. Shiu. 2007. Optimal dividends in the dual model. *Insurance: Mathematics and Economics* 41: 111–23.

Bae, S. C., K. Chang, and E. Kang. 2012. Culture, corporate governance and dividend policy: International evidence. *Journal of Financial Research* 35: 289–316.

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

Bickel, P. J., and A. Sakov. 2008. On the choice of m in the m out of n bootstrap and confidence bounds for extrema. *Statistica Sinica* 18: 967–85.

Chang, C.-H., and S.-J. Lin. 2015. The effects of national culture and behavioral pitfalls on investors' decision-making: Herding behavior in international stock markets. *International Review of Economics and Finance* 37: 380–92.

Chang, K., and A. Noorbakhsh. 2009. Does national culture affect international corporate cash holdings? *Journal of Multinational Financial Management* 19: 323–42.

Chui, A. C. W., A. E. Lloyd, and C. C. Y. Kwok. 2002. The determination of capital structure: Is national culture a missing piece to the puzzle? *Journal of International Business Studies* 33: 99–127.

Csörgő, M. 1983. *Quantile Processes with Statistical Applications*. Philadelphia: SIAM.
Dang, H. D. 2018. National culture and corporate rating migrations. *Risks* 6: 130.
De Haan, L., and A. Ferreira. 2006. *Extreme Value Theory: An Introduction*. New York: Springer.
De Luca, G., and P. Zuccolotto. 2011. A tail dependence-based dissimilarity measure for financial time series clustering. *Advances in Data Analysis and Classification* 5: 323–40.
Desender, K. A., C. E. Castro, and S. A. Escamilla De León. 2011. Earnings management and cultural values. *American Journal of Economics and Sociology* 70: 639–70.
Fidrmuc, J. P., and M. Jacob. 2010. Culture, agency costs and dividends. *Journal of Comparative Economics* 38: 321–39.
Furman, E., J. Su, and R. Zitikis. 2015. Paths and indices of maximal tail dependence. *ASTIN Bulletin: The Journal of the International Actuarial Association* 45: 661–78.
Grandell, J. 2012. *Aspects of Risk Theory*. New York: Springer.
Gray, S. J. 1988. Towards a theory of cultural influence on the development of accounting systems internationally. *ABACUS: A Journal of Accounting, Finance and Business Studies* 24: 1–15.
Greselin, F., F. Piacenza, and R. Zitikis. 2019. Practice oriented and Monte Carlo based estimation of the value-at-risk for operational risk measurement. *Risks* 7: 50.
Gribkova, N., and R. Zitikis. 2018. A user-friendly algorithm for detecting the influence of background risks on a model. *Risks* 6: 100.
Gribkova, N., and R. Zitikis. 2019a. Assessing transfer functions in control systems. *Journal of Statistical Theory and Practice* 13: 35.
Gribkova, N., and R. Zitikis. 2019b. Statistical detection and classification of background risks affecting inputs and outputs. *Metron—International Journal of Statistics* 77: 1–18.
Gribkova, N., and R. Zitikis. 2019c. Detecting intrusions in control systems: A rule of thumb, its justification and illustrations. *Journal of Statistics and Management Systems*, in press.
Guérin, H., and J.-F. Renaud. 2017. On the distribution of cumulative Parisian ruin. *Insurance: Mathematics and Economics* 73: 116–23.
Guiso, L., H. Herrera, and M. Morelli. 2016. Cultural differences and institutional integration. *Journal of International Economics* 99: 97–113.
Han, S., T. Kang, S. Salter, and Y. K. Yoo. 2010. A cross-country study on the effects of national culture on earnings management. *Journal of International Business Studies* 41: 123–41.
Hofstede, G., G. J. Hofstede, and M. Minkov. 2010. *Cultures and Organizations: Software of the Mind*, 3rd ed. New York: McGraw-Hill.
Hope, O.-K. 2003. Firm-level disclosures and the relative roles of culture and legal origin. *Journal of International Financial Management and Accounting* 14: 218–48.
Koch, E. 2017. Spatial risk measures and applications to max-stable processes. *Extremes* 20: 635–70.
Koch, E. 2019. Spatial risk measures and rate of spatial diversification. *Risks* 7: 52.
Landriault, D., J.-F. Renaud, and X. Zhou. 2014. An insurance risk model with Parisian implementation delays. *Methodology and Computing in Applied Probability* 16: 583–607.
Léveillé, G., and J. Garrido. 2001a. Moments of compound renewal sums with discounted claims. *Insurance: Mathematics and Economics* 28: 217–31.
Léveillé, G., and J. Garrido. 2001b. Recursive moments of compound renewal sums with discounted claims. *Scandinavian Actuarial Journal* 2: 98–110.
Léveillé, G., and E. Hamel. 2013. A compound renewal model for medical malpractice insurance. *European Actuarial Journal* 3: 471–90.
Li, K., D. Griffin, H. Yue, and L. Zhao. 2013. How does culture influence corporate risk-taking? *Journal of Corporate Finance* 23: 1–22.
Li, S., and Y. Lu. 2018. On the moments and the distribution of aggregate discounted claims in a Markovian environment. *Risks* 6: 59.
Liu, X., J. Wu, C. Yang, and W. Jiang. 2018. A maximal tail dependence-based clustering procedure for financial time series and its applications in portfolio selection. *Risks* 6: 115.
Lkabous, M. A., and J.-F. Renaud. 2018. A VaR-type risk measure derived from cumulative Parisian ruin for the classical risk model. *Risks* 6: 85.
Loisel, S., and J. Trufin. 2014. Properties of a risk measure derived from the expected area in red. *Insurance: Mathematics and Economics* 55: 191–99.

Loke, S.-H., and E. Thomann. 2018. Numerical ruin probability in the dual risk model with risk-free investments. *Risks* 6: 110.

Marri, F., F. Adékambi, and K. Moutanabbir. 2018. Moments of compound renewal sums with dependent risks using mixing exponential models. *Risks* 6: 86.

Mitric, I.-R., and J. Trufin. 2016. On a risk measure inspired from the ruin probability and the expected deficit at ruin. *Scandinavian Actuarial Journal* 10: 932–51.

North, D. C. 1990. *Institutions, Institutional Change and Economic Performance.* Cambridge: Cambridge University Press.

Ramirez, A., and S. Tadesse. 2009. Corporate cash holdings, uncertainty avoidance, and the multinationality of firms. *International Business Review* 18: 387–403.

Semenikhine, V., E. Furman, and J. Su. 2018. On a multiplicative multivariate gamma distribution with applications in insurance. *Risks* 6: 79.

Segerdahl, C.-O. 1942. Über einige risikotheoretische fragestellungen. *Scandinavian Actuarial Journal* 61: 43–83.

Shao, J., and D. Tu. 1995. *The Jackknife and Bootstrap.* New York: Springer.

Shao, L., C. C. Y. Kwok, and O. Guedhami. 2010. National culture and dividend policy. *Journal of International Business Studies* 41: 1391–414.

Shao, L., C. C. Y. Kwok, and R. Zhang. 2013. National culture and corporate investment. *Journal of International Business Studies* 44: 745–63.

Steffensen, M. 2013. Surrounding *Risks*. *Risks* 1: 43–44.

Sun-tzu. 2007. *The Art of War.* Translated, with an introduction and commentary, by Roger T. Ames. Preface by Rupert Smith. London: Folio Society.

Trufin, J., H. Albrecher, and M. Denuit. 2011. Properties of a risk measure derived from ruin theory. *The Geneva Risk and Insurance Review* 36: 174–88.

Zheng, X., S. El Ghoul, O. Guedhami, and C. C. Y. Kwok. 2012. National culture and corporate debt maturity. *Journal of Banking and Finance* 36: 468–88.

risks

MDPI

Article

National Culture and Corporate Rating Migrations

Huong Dieu Dang

Department of Economics and Finance, University of Canterbury, Christchurch 8140, New Zealand; Huong.dang@canterbury.ac.nz

Received: 14 October 2018; Accepted: 7 November 2018; Published: 14 November 2018

Abstract: The informal constraints that arise from the national culture in which a firm resides have a pervasive impact on managerial decision making and corporate credit risk, which in turn impacts on corporate ratings and rating changes. In some cultures, firms are naturally predisposed to rating changes in a particular direction (downgrade *or* upgrade) while, in other cultures, firms are more likely to migrate from the current rating in either direction. This study employs a survival analysis framework to examine the effect of national culture on the probability of rating transitions of 5360 firms across 50 countries over the period 1985–2010. Firms located in *long-term oriented* cultures are less likely to be downgraded and, in some cases, more likely to be upgraded. Downgrades occur more often in strong *uncertainty-avoiding* countries and less often in large *power distance* (*hierarchy*) and *embeddedness* countries. There is some evidence that *masculinity* predisposes firms to more rating transitions. Studying culture helps enrich our understanding of corporate rating migrations, and helps develop predictive models of corporate rating changes across countries.

Keywords: national culture; survival analysis; hazard model; rating migrations

1. Introduction

Hofstede, Hofstede and Minkov (Hofstede et al. 2010) define culture as "the collective programming of the mind that distinguishes the members of one category of people from those of another." In a global survey conducted in May 2008 by PricewaterhouseCoopers, 78% of the participants considered culture and excessive risk taking the contributing factors that led to the credit crisis (PricewaterhouseCoopers 2008). Contributing to this view is the evidence that culture affects corporate risk taking (Li et al. 2013), bank risk taking (Ashraf et al. 2016), and bank performance (Boubakri et al. 2017).

Pan et al. (2017) observe similarity in risk attitudes inside U.S. firms, which are rooted in the founders' risk attitudes and retained through the appointments of leaders with similar mindsets. Unlike Pan et al., this study does not focus on founders' risk attitudes and corporate risk culture.[1] As Ramirez and Tadesse (2009) suggest, the culture in which an individual resides influences his/her perception of uncertainty and his/her "mechanism" to cope with uncertainty and ambiguity. Corporate managers, regardless of their original cultural backgrounds, have to tailor their business priorities and corporate policies to the cultural contexts in which their firms operate. This study therefore focuses on the national culture in which a firm resides.

In practice, credit rating has been widely used as a measure of credit risk. A strong understanding of credit rating decisions has important implications to market participants. As the President of the European Commission José Manuel Barroso remarked at the European Parliament on 5 May 2010, ratings are "too cyclical, too reliant on the general market mood rather than on fundamentals—regardless of whether market mood is too optimistic or too pessimistic".

[1] Pan et al. acknowledge that they do not account for social influences and shared experience inside a firm, thus not capturing a firm's risk culture entirely.

Investors from different cultures react differently to an information shock and are affected by different cognitive biases (Afego 2018). Behavioral pitfalls, such as overreaction and herding, are related to culture values. For example, people in strong *uncertainty avoiding* countries, such as Greece and Italy, show a high level of anxiety and a natural sense of urgency (Hofstede et al. 2010). People in *collectivism* (*embeddedness*) cultures tend to exhibit herding behavior (Chang and Lin 2015). These culture values affect investor sentiment and market mood, and market mood impacts on credit ratings, as stated by the President of the European Commission. Thus, it is expected that culture indirectly influences rating decisions.

Empirical studies suggest that culture feeds into the rating review process through several channels (Appendix A). Culture directly influences managerial decision making thereby indirectly influencing corporate risk taking (Li et al. 2013), capital structure (Chui et al. 2002), debt maturity (Zheng et al. 2012), cash management (Ramirez and Tadesse 2009; Chang and Noorbakhsh 2009), and corporate investment (Shao et al. 2013). Corporate risk taking, corporate investment, capital structure, debt maturity and cash management are important criteria used in the rating review process (S&P RatingsDirect 2013).

The informal constraints that stem from culture have an extensive influence on daily behaviors, and this influence extends far beyond formal laws (North 1990). Culture directly affects managers' perceptions of questionable business practices (Vitell et al. 1993; Cohen et al. 1996) and indirectly influences managers' choices relating to financial disclosures, financial report quality (Gray 1988; Hope 2003) and the degree of earnings management (Han et al. 2010; Desender et al. 2011). Jorion et al. (2009) consider increased earnings management the contributing factor that has led to a deterioration in the quality of corporate investment ratings.

The conflict of interest between shareholders and creditors, and the extent of protection toward shareholders and creditors vary markedly across countries (Li et al. 2013). Paredes and Wheatley (2017) find that measures of investor protection are "subsumed" by culture. Culture affects managers' decisions to hire a Big Four auditor (Hope et al. 2008). High quality audits reduce information asymmetries and agency conflicts between corporate managers (shareholders) and creditors (Jensen and Meckling 1976). Culture also affects dividend policy (Shao et al. 2010; Fidrmuc and Jacob 2010; Bae et al. 2012), and dividend policy has been widely used as a monitoring mechanism to minimize agency problems. Culture may also complicate the negotiations to avoid a default between corporate managers (shareholders) and creditors.[2]

There is increasing recognition of the role culture plays in corporate risk taking and corporate credit risk. An interesting question which has not been addressed in the literature is if and how culture affects *changes* in credit risk, specifically, *changes* in ratings. Though culture is very stable, I argue that some cultures may predispose to more rating changes while other cultures may lead to directional predisposition for rating changes over time. As discussed in Section 2.3, a strong uncertainty avoidance trait may lead to more downgrades while a long-term orientation trait may result in fewer downgrades as time passes. Analysts may be more prone to change corporate ratings in countries with a high score for masculinity, individualism, and power distance.

Culture is expected to affect rating migrations through two channels (Appendix A). First, culture influences macro-economic activities[3] through its roles as an informal constraint and through its effects on managerial decision making (Zheng et al. 2012). Ratings move pro-cyclically (Lobo et al. 2017) and rating regrades are primarily affected by macro-economic conditions (Blume et al. 1991). Accelerated downgrades and defaults occur more often during economic contractions while upgrades tend to

---

[2] Contributing to this view is the evidence that the cultural differences between Greece and Germany made Greece's negotiations to avoid a default much more difficult (Guiso et al. 2016).

[3] For example, long-term oriented countries have a higher national saving rate and a higher growth rate. Individualistic countries achieve a higher GNI per capita (Hofstede et al. 2010, pp. 38, 263–65).

outweigh downgrades during economic expansions (Bangia et al. 2002; Lobo et al. 2017). Rating volatility is intensified during business cycle troughs and subdued during business cycle peaks (Nickell et al. 2000).

Second, culture feeds into the rating review process (as discussed above) and thus affects a firm's rating and its rating history. Empirical studies suggest that the current rating and various aspects of rating history impact on subsequent rating changes. For example, issuers in the boundary between investment and speculative rating grades (BBB−/BB+) exhibit different migration dynamics compared with other issuers (Carty and Fons 1994; Johnson 2004). Issuers downgraded to a given rating are more likely to be downgraded than those upgraded to that rating grade.[4]

The above discussion motivates this study to examine the impact of national culture on corporate rating migrations. The effects of culture are examined after accounting for variables that have been found to be significant in previous studies on rating migrations. The results of this study provide robust evidence that national culture significantly impacts the probability of rating changes of 17,109 ratings (5360 firms) across 50 countries over the period 1985–2010. The evidence in favor of including culture variables is stronger for rating downgrades and when numeric scores are used to represent culture values.

This study finds that *long-term orientation (LTO)* value is associated with a lower downgrade hazard, and in some cases, a higher upgrade probability. The effect of *LTO* on downgrades is robust to alternative samples (with and without U.S. firms), alternative measures of culture (Hofstede's and Schwartz's culture scores) and alternative study periods (crisis and non-crisis times). *Uncertainty avoidance* trait makes a downgrade more likely whereas *power distance (hierarchy)* and *embeddedness* trait reduces the risk of a downgrade. There is some evidence that *masculinity* dimension predisposes firms to more rating transitions.

This study extends the literature in two ways. First, this is the first study to explore the effects of culture values established by Hofstede et al. (2010) and Schwartz (1994) on the rating migration hazards of 5360 firms in 50 countries over a 26-year period. The study suggests a cross-disciplinary explanation and establishes a link between culture literature and rating migration literature. Second, the study applies a survival analysis framework (Allison 1995) and develops Cox's dynamic hazard model (Cox 1972) which offers three attractive features: It accounts for the sequence of rating migrations of the same firm, includes both time-fixed and time-varying variables in the estimation process, and allows for non-Markovian behaviors, such as within-rating heterogeneity and time-heterogeneity, in corporate rating migration dynamics.

The remainder of this paper is organized as follows. Section 2 describes the data and discusses culture values established by Hofstede et al. (2010) and Schwartz (1994). Section 3 presents the method, variables and samples. Section 4 discusses the results, and Section 5 summarizes the key findings.

2. National Culture and Corporate Rating Migrations

2.1. Rating Data

This study employs Standard and Poor's foreign currency issuer ratings retrieved from Ratings Xpress database on 28 September 2010. Firm-specific data other than credit ratings are not available. Standard and Poor's (S&P) ratings are used as S&P's decisions are generally timelier than Moody's (Güttler 2011). Foreign currency ratings are used as the persistence of national culture dimensions is relevant for firms which issue debts in international markets.

Financial institutions and utilities are excluded from this study for two reasons. First, previous studies suggest that financial institutions exhibit different rating migration dynamics compared with industrial firms (Nickell et al. 2000; Lando and Skodeberg 2002). Second, financial institutions and

[4] For the evidence of downward momentum in rating migration dynamics, see Altman and Kao (1992); Carty and Fons (1994); Altman (1998); Bangia et al. (2002); Lando and Skodeberg (2002); Güttler and Wahrenburg (2007); Figlewski et al. (2012); Dang and Partington (2014)

utilities operate in highly regulated environments and are closely monitored by their respective regulatory agencies. This suggests less latitude for corporate managers in these sectors to exhibit flexibility in decision making. Thus, financial institutions and utilities deserve a separate study and this article focuses entirely on other sectors. The study spans the period from 1 January 1985[5] to 28 September 2010.

2.2. Culture Data

Hofstede's culture values have been widely employed in various business disciplines including finance. This study employs Hofstede et al. (2010)'s five culture values, namely *power distance index*, *uncertainty avoidance index*, *individualism versus collectivism*, *masculinity versus femininity*, and *long-term versus short-term orientation*. A country's score for a culture value does not represent an absolute position but reflects its ranking relative to other countries. The validity of culture values persists for a long period of time. Hofstede (2001, p. 255) stresses that in the long term "cultures shift, but they shift in formation so that the differences between them remain intact."

This study employs the numerical scores of Hofstede's five culture values and five dummy variables, each based on a country's score relative to the mean score. Scores greater than or equal to the mean take a value of one or zero otherwise. As the differences between cultures remain stable over time, the use of dummy variables helps maintain the relative rankings between cultures over the study period.

Most culture values introduced in recent studies are conceptually related to and empirically correlated with Hofstede's culture values (Leung et al. 2005).[6] Two widely known frameworks that categorize culture along similar values are the GLOBE study conducted by Robert House and his team in 1991 and Schwartz (1994). GLOBE values are not used in this study as, despite using the same terms, their meanings are different compared with Hofstede's values. The approach the GLOBE team used to formulate survey questions suffers from major disadvantages (Hofstede et al. 2010, pp. 42–43).

There are significant correlations between Schwartz's scores and Hofstede's scores (Hofstede et al. 2010, p. 41). Given their comparability, this study employs the numeric scores of two of Schwartz's culture measures, *embeddedness*[7] and *hierarchy*, as alternative measures of Hofstede's *collectivism* (the opposite pole of *individualism*) and *power distance index*, respectively. As in Zheng et al. (2012), this study does not employ Schwartz's *mastery* trait because it captures such values as success, independence, ambition, and capability, which overlap with those of Hofstede's *individualism* and *masculinity* traits.

2.3. Culture Dimensions

The following discussion elaborates on the possible links between culture values and corporate rating migrations, as outlined in Appendix A. In some cultures, firms are naturally predisposed to rating changes in a particular direction (downgrade *or* upgrade) while in other cultures firms tend to migrate from the current rating in either direction.

2.3.1. Power Distance Index (PDI) or Hierarchy

Hofstede's *power distance index* (PDI) and Schwartz's *hierarchy* reflect the extent to which people accept an unequal, hierarchical distribution of power, authority and wealth (Hope 2003; Licht et al. 2005). Small power distance (PD) cultures strive to equalize the distribution of power, and justification

---

[5] S&P rating scales were changed in 1983. To calculate the annual changes of employed macro-economic variables in 1985, the values in 1984 and 1985 were needed. Thus, 1985 was chosen as the starting year of the study.

[6] Tang and Koveos (2008) suggest that institutional factors, such as language, religion, climate and legal origin, are subsumed by Hofstede's *uncertainty avoidance* and *masculinity* traits. The correlations between Hofstede's culture scores and other measures "do not tend to become weaker over time" (Hofstede et al. 2010, p. 39).

[7] *Embeddedness* is referred to as *conservatism* in some studies, such as Johnson and Lenartowicz (1998); Chui et al. (2002); Shao et al. (2010).

is required for any inequalities. Large PD cultures accept a hierarchical order in which everybody has a place and no further justification is required (Hofstede et al. 2010).

Agency conflicts are less severe in large PD countries as there is greater acceptance of wealth and power inequalities (Fidrmuc and Jacob 2010). As a result, firms adopt a lower dividend payout policy and are less likely to hire a Big Four auditor (Hope et al. 2008). It is difficult for firms to get access to long-term credit markets given the greater risk of violence in domestic politics. Thus, firms employ more short-term debt than long-term debt and exhibit a lower degree of financial risk (Zheng et al. 2012). Contributing to this view is the evidence that banks in large PD countries tend to take fewer risks (Ashraf et al. 2016).

A high PD score is related to a high degree of earnings management (Paredes and Wheatley 2017). Corporate managers exhibit strong influences on financial reporting choices and prefer not to disclose information to preserve power inequalities (Zarzeski 1996). People often view a questionable business practice as ethical (Cohen et al. 1996, p. 58) and are more tolerant of corruption and tax evasion (Husted 1999; Tsakumis et al. 2007). Policy debates and political discussions are not often seen in larger PD societies, which indicates a lower degree of transparency and accountability.

The above discussion suggests that a high PD score may predispose firms to rating changes in either direction (downgrade *and* upgrade).

2.3.2. Individualism (IDV) versus Collectivism (Embeddedness)

People in individualistic countries (high IDV score) show autonomy and are encouraged to pursue personal goals and stand up for their rights (Licht et al. 2005). Hofstede's *individualism* trait may predispose firms to rating changes in either direction (downgrade *and* upgrade), as discussed below.

Individualistic countries have strong economies (Hofstede 2001, p. 519) and tend to adopt market-based financial systems, which encourage corporate risk taking (Li et al. 2013). Their preferred Anglo-Saxon corporate governance systems focus on shareholders' interests (Griffin et al. 2015). Firms employ a high degree of leverage (Chui et al. 2002), invest more in long-term assets, R&D projects (instead of physical assets), and employ excess cash to increase R&D (instead of increasing dividends) (Shao et al. 2013). Aggressive risk-taking results in more volatile operating income (Li et al. 2013). Consequently, managers are more likely to manage earnings (Han et al. 2010) and engage in income smoothing (Fonseca and Gonzalez 2008). Agency conflicts are inherently more severe (Chui et al. 2002) as corporate insiders exhibit a strong tendency to pursue their own personal interests rather than adhere to different stakeholders' preferences.

On the positive side, individualistic countries have effective regulatory systems. Laws and rights are equal for all groups, and formal institutions are established to protect the rights of competing parties, such as shareholders and creditors (Licht et al. 2005). Firms tend to adopt a high dividend payout policy to minimize agency problem concerns (Shao et al. 2010; Fidrmuc and Jacob 2010). The business environment in individualistic cultures is more competitive and less secretive than in collectivistic cultures (Gray 1988). Firms are open to extensive accounting disclosures (Gray and Vint 1995) and are more likely to hire a Big Four auditor (Hope et al. 2008). People are less tolerant of corruption and tax evasion (Tsakumis et al. 2007).

In collectivistic (embeddedness) cultures (low IDV score), people have less need for autonomy and an active determination of their own lives. Schwartz's *embeddedness* trait is expected to lower downgrade hazard and raise upgrade probability, as discussed below.

Collectivistic (embeddedness) cultures focus on maintaining a harmonious relationship within a society and preserving public images. People value conformity and adherence to societal norms and regulations (Hofstede 1980). Firms tend to adopt autocratic and paternalistic management systems. Corporate managers are very concerned about the liquidation costs to its stakeholders (Titman 1984). Capital structure is decided by corporate leaders who have strong views that there should be no detrimental effect on employees. As a high degree of financial leverage restricts firms' flexibility and raises the probability of bankruptcy, firms in embeddedness cultures tend to employ less debt

(Chui et al. 2002) and use more short-term debt (Zheng et al. 2012). Avoiding bankruptcy is also consistent with preserving a firm's public images.

Collectivistic (embeddedness) cultures are more likely to have bank-based financial systems and relation-based corporate governance systems, which place a greater emphasis on the interests of a firm's different stakeholders (Griffin et al. 2015). Agency conflicts are less severe and excessive risk taking is not encouraged as managers tend to act in line with the interests of all stakeholders (Jensen and Meckling 1976; Fidrmuc and Jacob 2010; Shao et al. 2013). Corporate managers share a strong sense of responsibility and are less likely to differentiate between their own and others' welfares.

Overall, *individualism* may predispose firms to more rating regrades whereas *embeddedness* is expected to raise upgrade hazard and reduce downgrade risk.

2.3.3. Masculinity (MAS) versus Femininity

Masculine cultures (high MAS score) value assertiveness, ambition, competition, challenge, recognition, material accomplishment, and success. *Feminine* cultures (low MAS score) value cooperation, modesty, tenderness, security, caring for the weak, and quality of life (Hofstede et al. 2010).

Masculine cultures encourage taking risky decisions and achieving performance goals. Men often exhibit overconfidence or self-attribution biases (Barber and Odean 2001). Compared with women, men are less likely to assess the accurate levels of risk associated with an assigned task (Byrnes et al. 1999). Men are also more likely to experience failure or unfavorable outcomes, and can be more aggressive than women about default. Contributing to this view is the evidence that banks in masculine countries are more likely to incur a large loss during a crisis (Kanagaretnam et al. 2011).

Masculine cultures are open to head-on confrontation (Licht et al. 2005) and encourage aggressive behaviors (Kanagaretnam et al. 2011). Conflicts are often resolved by fighting in masculine countries and by negotiation in feminine countries.[8] The political environment in masculine countries tends to be adversarial, as opposed to cooperative coalitions in feminine countries (Hofstede et al. 2010, pp. 173–80).

Countries with high MAS scores are less likely to perceive ethical issues in business practices (Vitell et al. 1993, p. 758). However, masculine countries incorporate economic interests in legal form (Licht et al. 2005), and are less permissive in dealing with lawbreakers. Masculine countries focus more on punishment while feminine countries are more lenient and place greater emphasis on correction and rehabilitation (Hofstede 2001, p. 319).

Firms in masculine cultures are open to extensive information sharing (Hope 2003; Doupnik and Tsakumis 2004) as societies are more business oriented and value visible achievement. Corporate managers are concerned that debt covenants may restrict their aggressive business plans and fears of financial distress may interfere with their bold investment strategies. Firms thus prefer less debt financing (Chui et al. 2002) and use more short-term debt (Zheng et al. 2012). By contrast, feminine cultures may accept higher indebtedness to support their welfare systems "even at the expense of accomplishment and performance" (Weaver 2001, p. 9).

The above discussion suggests that a high MAS score may predispose firms to rating migrations in either direction (downgrade *and* upgrade).

---

[8] An example is the handling of the Åland Islands crisis and the Falkland Islands crisis. The Åland Islands crisis was resolved by negotiations in 1921 between feminine countries Finland and Sweden. The Åland Islands remained Finnish but the pro-Swedish islands gained substantial regional autonomy. The Falklands Islands crisis in 1982 involved Argentinean military and British expeditionary forces. The crisis between two masculine countries cost "725 Argentinean and 225 British lives and enormous financial expense." The Falklands Islands have remained a disputed territory and required "constant British subsidies and military presence" (Hofstede et al. 2010, p. 173).

2.3.4. Uncertainty Avoidance Index (UAI)

Uncertainty avoidance (UA) is the extent to which members of a society feel uncomfortable with uncertainty and find ambiguity stressful (Chui et al. 2002). People from strong UA countries are subject to a natural sense of urgency and are more likely to suffer from anxiety (Hofstede et al. 2010). A strong UA trait tends to exacerbate investor panic and create more uncertainty during market turmoil. Deteriorating market mood and heightened concerns about uncertainty often lead to drastic and accelerated downgrades.[9]

Uncertainty avoidance differs from risk avoidance (Tsakumis et al. 2007). High UA cultures perceive ambiguity as a continuous threat. People tend to "engage in risky behavior in order to reduce ambiguities, such as starting a fight with a potential opponent rather than sitting back and waiting". They are open to familiar risks such as driving faster and having more fatal accidents but they show fears of ambiguity and unfamiliar risks (Hofstede et al. 2010, pp. 197–98).

Strong UA countries tend to adopt bank-based financial systems (Kwok and Tadesse 2006) and maintain a low degree of economic freedom. Firms in these countries have a negative view of competition. Debt financing is minimized as the use of debt imposes constraints and may lead to financial instability (Chui et al. 2002). Corporate managers are concerned about potential financial distress and the cash shortage that may require capital raising under unfavorable market conditions or fire sales of a firm's assets (Chang and Noorbakhsh 2009). Access to long-term credit markets is difficult given the greater risk of political instability.[10] So, firms prefer to use more short-term debt (Zheng et al. 2012), adopt a lower dividend payout policy (Fidrmuc and Jacob 2010) and hold more cash (Ramirez and Tadesse 2009; Chang and Noorbakhsh 2009). Firms are not open to extensive accounting disclosure (Gray and Vint 1995) and are less likely to hire a Big Four auditor (Hope et al. 2008). Information asymmetries and agency conflicts between managers (shareholders) and creditors tend to be greater.

In strong UA countries, "ineffective rules can satisfy people's emotional need for formal structure" (Hofstede et al. 2010, p. 209). Corporate managers are less likely to perceive ethical problems and are not opposed to contravening an unjust law. In situations where outcomes cannot be determined with certainty, corruption is viewed as an option to secure a predictable result. Tax evasion is considered a means of reducing ambiguity (Vitell et al. 1993, p. 757; Husted 1999; Tsakumis et al. 2007; Hofstede et al. 2010, p. 223). By contrast, in weak UA countries rules are often more likely to be followed (Hofstede et al. 2010), investors' legal rights are stronger (Licht et al. 2005), and people are less likely to view tax evasion as a viable option.

In strong UA societies, people are also slower in paying their bills (De Mooij 2004, p. 154). These countries have more precise formal laws, informal rules, rigid safety and security measures in place. However, it takes more time for their citizens to comply with two simple civil procedures: collecting a bounced check which had been refused by a bank and evicting a tenant for non-payment of rent (Hofstede et al. 2010, pp. 216–17). Debt burdened firms in strong UA countries may not waste time negotiating with creditor(s) to reduce their delinquent debts. Driven by a high level of anxiety and a natural sense of urgency, these firms may simply walk away to avoid any uncertainty associated with the debt restructure process.

In the light of this discussion, a high UA score is expected to make downgrades more likely and upgrades less likely.

---

[9] Contributing to this view is the remark of the President of the European Commission José Manuel Barroso at the European Parliament in May 2010 that ratings are "too cyclical, too reliant on the general market mood rather than on fundamentals...".

[10] Strong UA countries are intolerant of political ideologies, are "more likely to harbor extremist minorities within their political landscape" and have more "native terrorists" (Hofstede et al. 2010, p. 221).

2.3.5. Long-Term (LT) versus Short-Term (ST) Orientation

Long-term oriented societies (high LTO score) value pragmatic virtues related to the future, such as savings and adapting to changing circumstances (Hofstede et al. 2010, p. 239). Firms invest in building up strong market positions even at the expense of poor short-term performance. Managers are allowed time and resources to make sustained efforts with an aim to serve stakeholders and future generations (Hofstede et al. 2010, p. 244). Young people are taught the importance of savings and learn to put money aside for future uses.

Short-term oriented societies (low LTO score)—for example, Arab countries—value virtues related to the past and present such as national pride, respect for tradition, preservation of public images, and fulfilling social obligations (Hofstede et al. 2010, p. 239). Young people, affected by a culture of consumption, often live beyond their means. Corporate managers focus on the immediate results and are judged by their short-term performance. The cost of short-term decisions in terms of "pecuniary considerations, myopic decisions, work process control, hasty adoption, and quick abandonment of novel ideas" is evident (Hofstede et al. 2010, pp. 244–45).

The economic success of Taiwan, South Korean, Singapore, Hong Kong and Japan in the early 1990s highlights the value of their LTO cultures. These countries encourage thrift, perseverance (Anderson et al. 2011), and support entrepreneurial activities. People are persistent in the pursuit of their goals and are less tolerant of questionable business activities (Cohen et al. 1996). Their thrift translates into higher savings and growth rates (Hofstede et al. 2010, pp. 38, 263–65) and the availability of capital for reinvestment. Firms are less likely to rely heavily on debt financing. The sense of shame prevalent in LTO societies encourages interrelatedness through social contacts and stresses the importance of keeping commitments (Hofstede et al. 2010, pp. 243–44). Firms have strong ethical motivations to pay debts on time, and the severity of agency conflicts is inherently lower given their focus on stakeholders and future generations.

In the light of this discussion, a high LTO score is expected to make upgrades more likely and downgrades less likely.

3. Models and Variables

3.1. Estimation Model

This study applies a survival analysis framework (Allison 1995) and develops Cox's dynamic hazard model (Cox 1972) to examine the effect of culture on the probability of corporate rating migrations. Previous studies suggest that upgrade and downgrade follow different dynamics (Figlewski et al. 2012; Dang and Partington 2014). Thus, upgrade and downgrade are treated as competing risks, and separate models are estimated for these two migration outcomes.

Rating observations are arranged in event time (gap time) risk sets; each risk set includes all the firms that are at risk of a migration of interest at event time t. The clock is reset when a firm is assigned a rating grade. An upgrade (downgrade) is treated as a migration of interest (a censored observation) when estimating the upgrade model, and vice versa. Ratings which started before the beginning of the study or ended after the end of the study are also treated as censored. The survival time of a rating is the time a firm maintains a rating grade measured from the time it enters the rating grade subsequent to the start of the study until the time it either migrates to another rating grade or becomes censored.

Cox's hazard model is the premier technique in survival analysis. The main attraction of the estimated model is that it is convenient to handle repeated migrations. A typical firm experienced several migrations during the study period, which may lead to dependence among ratings. Accounting for repeated migrations is important as 42.6% (18%) of the firms in the main sample experienced between 2 and 15 downgrades (upgrades); the most volatile firms experienced 26 downgrades (25 upgrades). This highlights the need to consider unobserved heterogeneity in the underlying hazard of a rating change and to ensure that a firm is not considered at risk of a rating change before all previous rating changes have already occurred (Hosmer et al. 2008).

The problem arising from repeated migrations is minimized in two ways. First, the Wei et al. (1989)'s method for multiple failure time data is applied to account for the dependence among ratings of the same firm. The advantage of this method is that it does not require any assumptions about the nature or the structure of the dependence (Allison 1995, p. 242). Second, the estimation of the stratified hazard model, as in as Dang and Partington (2014), takes into account the sequence of rating changes. Each stratum includes ratings that have the same number of prior changes. The underlying (baseline) hazard of a rating change differs according to the number of prior changes whereas the effect of a variable is assumed to be the same across strata.

Another attractive feature of the estimated hazard model is that it accommodates time-varying outlook and rating age. A firm may be assigned a negative, positive, stable or developing outlook while its rating remains unchanged. A negative (positive) outlook signals the deterioration (improvement) in a firm's credit quality, and indicates S&P's opinion regarding the potential direction of a long-term credit rating over the intermediate term (six months to two years) (S&P RatingsDirect 2009). Estimating stratified dynamic hazard models with time-varying variables on a sample of 17,109 ratings (5360 firms) across 50 countries over a 26-year period creates substantial computational challenges in this study.

The hazard of a rating change m in stratum s is given by the product of the underlying (baseline) hazard $h_{(0,s)}(t)$ and the effect of the risk factors (covariates), and can be expressed as follows:

$$h_{m,s}(t, Z, Z(t)) = h_{(0,s)}(t)\ \exp[Z_j^m \beta_j + Z_p^m(t)\beta_p\ \Big], \tag{1}$$

where: $h_{m,s}(t, Z, Z(t))$ is the migration hazard of rating m in stratum s at time t given its time-fixed covariate vector Z_j^m and its time-varying covariate vector $Z_p^m(t)$; $h_{(0,s)}(t)$ is the baseline hazard of a migration in stratum s at time t; β_p is the vector of estimated coefficients for time-varying covariates $Z_p^m(t)$; and β_j is the vector of estimated coefficients for time-fixed covariates Z_j^m.

The full rating migration model is estimated by multiplying together the individual likelihood functions for all the strata in the sample.

3.2. Variables

The definitions of variables employed, the data sources, and the references to examples of the relevant literature are given in Table 1.

The first group of variables includes Hofstede's five culture values and two of Schwartz's culture values, as discussed in the previous section.

The second group of variables accounts for the current rating (*current rating grade),* its proximity to investment/speculative rating threshold *(dummy investment boundary, dummy junk boundary*), rating history (*logarithm of age since first rated, dummy lag one downgrade, lag one duration, dummy prior fallen angel, dummy large downgrade, dummy large upgrade, rating volatility*) and time-varying rating outlook (*dummy negative outlook, dummy positive outlook*). Previous studies on corporate rating migrations have widely documented that the direction of lagged rating change, the duration of lagged rating, and rating outlook are the key determinants of rating changes.[11] The use of lagged dependent variable (*lag one duration*) and important control variables (*dummy lag one downgrade*, *dummy negative outlook, dummy positive outlook*) minimizes endogeneity concerns, Li (2016).

[11] See, for example, Altman and Kao (1992); Carty and Fons (1994); Altman (1998); Bangia et al. (2002); Lando and Skodeberg (2002); Vazza et al. (2005b); Güttler and Wahrenburg (2007); Figlewski et al. (2012); Dang and Partington (2014).

Table 1. Definitions of variables used in the analysis.

Variable	Definition	References
Hofstede's culture dimensions Hofstede (1980) conducted surveys with IBM employees in over 50 countries and used the survey responses to identify four national culture dimensions that were virtually uncorrelated. In Hofstede et al. (2010), the scores on the four dimensions were listed for 76 countries. The fifth dimension, *long-term* versus *short-term orientation*, was introduced by Michael Bond in his Chinese Value Survey conducted in 23 countries in 1987, and extended to 93 countries in 2010 by Michael Minkov (Hofstede et al. 2010)		
Power distance index	Power distance index expresses the degree to which members of a society accept and expect that power and authority is distributed unequally.	Licht et al. (2005); Hope et al. (2008); Hofstede et al. (2010); Fidrmuc and Jacob (2010); Kanagaretnam et al. (2011); Zheng et al. (2012); Paredes and Wheatley (2017)
Individualism vs. collectivism	Individualism encourages the pursuit of personal interests, autonomy and an active determination of one's destiny. Collectivism stresses conformity and adherence to societal norms and regulations	Chui et al. (2002); Licht et al. (2005); Tsakumis et al. (2007); Fidrmuc and Jacob (2010); Han et al. (2010); Hofstede et al. (2010); Shao et al. (2010); Kanagaretnam et al. (2011); Zheng et al. (2012); Li et al. (2013); Shao et al. (2013)
Masculinity vs. femininity	Masculine countries strive for a performance society and value assertiveness, material accomplishment, ambition, competition and success. Feminine countries strive for a welfare society and value cooperation, modesty, caring for the weak and quality of life	Vitell et al. (1993); Chui et al. (2002); Licht et al. (2005); Hofstede et al. (2010); Anderson et al. (2011); Kanagaretnam et al. (2011); Zheng et al. (2012)
Uncertainty avoidance index	The uncertainty avoidance index expresses the degree to which the members of a society feel uncomfortable with uncertainty and ambiguity.	Licht et al. (2005); Kwok and Tadesse (2006); Ramirez and Tadesse (2009); Tsakumis et al. (2007); Hope et al. (2008); Fidrmuc and Jacob (2010); Han et al. (2010); Hofstede et al. (2010); Anderson et al. (2011); Kanagaretnam et al. (2011); Zheng et al. (2012); Li et al. (2013)
Long-term vs. short-term orientation	Long-term oriented cultures are oriented toward the future and value perseverance and thrift. Short-term oriented societies foster virtues related to the past and present.	Cohen et al. (1996); Hofstede et al. (2010); Anderson et al. (2011)
Schwartz's culture dimensions Schwartz (1994) collected survey data from school teachers and university students in more than 60 countries. He classified national cultures into six dimensions. Two dimensions *embeddedness* and *hierarchy* are employed in this study		
Embeddedness (conservatism)	Embedded cultures value social relationships, emphasize maintaining the status quo and restraining actions that may disrupt in-group solidarity and traditional order	Johnson and Lenartowicz (1998); Chui et al. (2002); Licht et al. (2005); Shao et al. (2010); Zheng et al. (2012)
Hierarchy	Hierarchical cultures view the unequal distribution of power, roles, and wealth as legitimate and even desirable.	Chui et al. (2002); Licht et al. (2005); Zheng et al. (2012)

Table 1. *Cont.*

Variable	Definition	References
S&P's rating data: Source: Standard & Poor's Ratings Xpress		
Current rating grade	The current rating (start rating) for the rating transition being observed.	Carty and Fons (1994); Figlewski et al. (2012); Dang and Partington (2014)
Investment rating boundary	The dummy takes the value of one if the current rating is in the investment grade boundary (BBB−, BBB, BBB+) or zero otherwise	Carty and Fons (1994); Johnson (2004); Dang and Partington (2014)
Junk rating boundary	The dummy takes the value of one if the current rating is in the speculative (junk) grade boundary (BB−, BB, BB+) or zero otherwise	Carty and Fons (1994); Johnson (2004); Dang and Partington (2014)
Logarithm of age since first rated	*Age since first rated* is a time-varying variable measuring the duration since a firm was first rated. This variable is updated whenever a migration of interest occurs in the sample	Altman (1998); Figlewski et al. (2012); Dang and Partington (2014)
Dummy lag one downgrade	The dummy takes the value of one if the lag one rating ends with a downgrade, and zero otherwise	Carty and Fons (1994); Lando and Skodeberg (2002); Bangia et al. (2002); Figlewski et al. (2012); Dang and Partington (2014)
Lag one duration (years)	The duration of the rating immediately preceding the current rating	Carty and Fons (1994); Lando and Skodeberg (2002); Dang and Partington (2014)
Dummy prior fallen angel	This variable takes the value of one if a firm had experienced a downgrade from an investment-grade rating to a speculative-grade rating as of the start of the current rating, and zero otherwise	Mann et al. (2003); Vazza et al. (2005a); Güttler and Wahrenburg (2007); Dang and Partington (2014)
Dummy large downgrade	This variable takes the value of one if a firm had experienced a big downgrade of at least three rating notches as of the start of the current rating, and zero otherwise	Carty and Fons (1994); Dang and Partington (2014)
Dummy large upgrade	This variable takes the value of one if a firm had experienced a big upgrade of at least three rating notches as of the beginning of the current rating, and zero otherwise	Dang and Partington (2014)
Rating volatility	This is the average number of migrations per year over a firm's rating history. It is calculated as the number of migrations a firm had experienced as of the beginning of the current rating divided by *age since first rated*.	Dang and Partington (2014)
S&P's outlook Source: Standard & Poor's Ratings Xpress S&P issues an outlook to indicate its opinion regarding the potential direction of a long-term credit rating over the intermediate term (six months–two years) (S&P RatingsDirect 2009). Outlooks can be positive (rating may be raised), negative (rating may be lowered), stable (rating unlikely to change), or developing (rating may be raised/ lowered)		
Dummy negative outlook	This time-varying variable takes the value of one if a firm was assigned a negative outlook by S&P, and zero otherwise.	Vazza et al. (2005a); Hill et al. (2010)
Dummy positive outlook	This time-varying variable takes the value of one if a firm was assigned a positive outlook by S&P, and zero otherwise.	Vazza et al. (2005a); Hill et al. (2010)

Table 1. *Cont.*

Variable	Definition	References
Macro-economic and financial conditions Source: World Bank databases unless otherwise stated		
Dummy prior default	This dummy takes the value of one if a country where a firm resides had a foreign currency-denominated debt default prior to the start of the rating under study, and zero otherwise. Source: S&P Global Ratings' S&P Global Ratings' Credit Research (2013)	Mora (2006); Hill et al. (2010)
Dummy debt crisis	This variable takes a value of one if a rating commences during a period of sovereign debt/ banking crisis as listed in Manasse et al. (2003), Laeven and Valencia (2008), or De Paoli et al. (2009), and zero otherwise.	Ferri et al. (1999); Mora (2006)
Dummy OECD member	This variable takes a value of one if the country where a firm resides is a member of the OECD at the start of the current rating, and zero otherwise.	Ferri et al. (1999); Mora (2006)
Logarithm of GDP per capita	The logarithm of real GDP per capita	Ramirez and Tadesse (2009); Zheng et al. (2012); Figlewski et al. (2012); Shao et al. (2013); Li et al. (2013); Dang and Partington (2014)
Change in real GDP growth rate	The change in the real GDP growth rate over the year prior to the start of the rating.	Ferri et al. (1999); Mora (2006); Hill et al. (2010); Shao et al. (2013)
Change in inflation	The change in the inflation rate over the year prior to the start of the rating.	Ramirez and Tadesse (2009); Zheng et al. (2012)
Change in current account surplus/GDP	The change in the current account surplus or deficit divided by GDP	Ferri et al. (1999); Mora (2006); Hill et al. (2010)
Change in term trade	The change in terms of trade. The terms of trade effect equals capacity to imports less exports of goods and services in constant prices. Data are in constant local currency.	
Logarithm of ratio stock market capitalization/GDP	The logarithm of the ratio of stock market capitalization to GDP	Zheng et al. (2012); Li et al. (2013); Shao et al. (2013)
Return of world stock market index	The average return of the World-Datastream stock market index, which is calculated using daily data over a 63-trading day rolling window prior to the start of the rating under study. Source: Datastream	Hill et al. (2010)
Political rights and civil liberties Source: International Country Risk Guide database. The political risk rating comprises the scores of 12 metrics including government stability, bureaucracy quality, corruption, democratic accountability, external conflict, ethnic tensions, internal conflict, investment profile, law and order, military in politics, religion in politics, and socioeconomics conditions.		
Dummy high political risk	This dummy takes a value of one if a country's political rating score is less than or equal to 40, and zero otherwise 40	
Dummy low political risk	This dummy takes a value of one if a country's political rating score is higher than or equal to 80, and zero otherwise.	

Table 1 defines the variables used in Cox's hazard models. The table provides the data sources, variable names, the definitions of the variables, and references to examples of the literature that have used the variables.

The third group of variables takes into account country characteristics. Potential variables are identified from previous studies on rating migrations. The choices of variables eliminate multi-collinearity concerns and aim at creating a sample from a large number of cultures. Selected variables are proxies for changes in macro-economic conditions (*change in real GDP growth rate, change in inflation, change in current account surplus/GDP, change in term trade*), the level of economic development (*logarithm of GDP per capita, dummy OECD member*), the degree of stock market capitalization (*logarithm of ratio stock market capitalization/GDP*), the existence of a sovereign debt/banking crisis (*dummy debt crisis*), and a history of sovereign foreign currency debt defaults (*dummy prior default)*. Selected variables also account for the performance of global stock market (*return of world stock market index*), and the political risk in each country (*dummy high political risk, dummy low political risk*). By definition, *dummy high political risk* and *dummy low political risk* capture institutional factors such as religion and law (Table 1). The use of these variables further minimizes endogeneity concerns.

The variables are of two types: static (time-fixed) and dynamic (time-varying). The time-varying variables (*dummy negative outlook, dummy positive outlook, logarithm of age since first rated)* are updated over the duration of a rating as its outlook changes or as a migration of interest occurs. The time-fixed variables take the values observed at the start or closest to the start of a rating observation and are not changed over the duration of a rating.[12] The values of static variables only change when the rating changes. The durations of ratings, particularly for downgrades, were small, as depicted in Figure 1. Thus, substantial changes in the values of static variables during the duration of a rating were unlikely.

3.3. Samples

Since the data on outlook and the last rating change are required, only firms having at least one rating outlook and experiencing at least one prior migration during the study period are included in the final dataset. The main sample, sample A(H), includes 17,109 ratings of 5360 firms from 50 countries. This sample consists of ratings which have no missing data on Hofstede (H)'s five culture values and control variables.

Several robustness tests are conducted using alternative samples, alternative study periods, and alternative culture measures. In the first test, Schwartz's numeric scores for *embeddedness* and *hierarchy* are used as alternative measures of Hofstede's numeric scores for *collectivism* (the opposite pole of *individualism*) and *power distance index*, respectively. The requirement of data availability for Schwartz (S)'s culture scores causes a small reduction in the size of sample A(H), resulting in sample A(H-S) of 16,966 ratings. Each rating in sample A(H-S) has no missing data for three of Hofstede (H)'s culture values (*masculinity*, *uncertainty avoidance index*, and *long-term orientation*), two of Schwartz (S)'s culture values (*embeddedness* and *hierarchy*), and all control variables.

In the second robustness test, U.S. firms are excluded from samples A(H) and A(H-S), which results in samples B(H) and B(H-S), respectively. Sample B(H) (with Hofstede's culture values) includes 4745 ratings of 1717 firms from 49 countries. Sample B(H-S) (with Hofstede and Schwartz's culture values) includes 4602 ratings.

In the third test, ratings in sample A(H) are pooled across countries at the time they were (were not) experiencing a debt/a banking crisis, as defined in Manasse et al. (2003), Laeven and Valencia (2008), and De Paoli et al. (2009). Restricting the study period to crisis and non-crisis times results in samples C(H-1) and C(H-2), respectively. Sample C(H-1) (crisis times) includes 3927 ratings of 2088 firms from 32 countries. Sample C(H-2) (non-crisis times) includes 13,182 ratings of 4614 firms from 50 countries. Excluding U.S firms from sample C(H-2) results in sample C(H-3) of 3714 ratings (1535 firms).

---

[12] Most static variables are updated annually. *Return of world stock market index* is calculated using daily data over a 63-trading day rolling window prior to the beginning of each rating. *Dummy OECD member, dummy debt crisis and dummy prior default* are updated at the beginning of each rating.

3.4. Statistics

A numerical rating scale is employed to represent S&P's alphabetical ratings, ranging from AAA as 21 to C as 1. The descriptive statistics of the current rating grades (*start rating)* for observations in the three samples with Hofstede's culture scores are presented in Panel A of Table 2. On average, firms in all three samples had a speculative grade median rating (BB for the whole sample and BB+ for the sample excluding U.S firms). During crisis times the median rating dropped by two notches from BB to B+, suggesting a substantial decline in credit quality.

The descriptive statistics of the survival times for downgrades and upgrades in the three samples with Hofstede's culture scores are respectively given in Panel B and Panel C of Table 2. The median survival time was 1 year (1.74 years) for downgrades (upgrades) in the entire sample A(H), and about 0.38 year (1.01 year) for downgrades (upgrades) in the crisis sample C(H-1).

Downgrades (upgrades) account for 60.85% (23.48%), 56.08% (23.1%) and 46.47% (13.3%) of the entire sample A(H), non-U.S. sample B(H), and crisis sample C(H-1), respectively. The surprisingly low frequency of downgrades in the crisis sample is due to the lower than average frequency of downgrades of U.S. firms during crisis times. This is not too surprising as the study does not consider financial sector and thus does not examine accelerated downgrades which occurred to a large number of U.S. financial institutions during the financial crisis 2007–2009. Additional analysis shows that for non-U.S. firms, the frequency of downgrades is comparable (56%) across business cycles. For U.S. firms, the frequency of downgrades is markedly higher in non-crisis times (68.63%) than in crisis times (43.23%). For both U.S. and non-U.S. firms, upgrades occur much less often in crisis times.

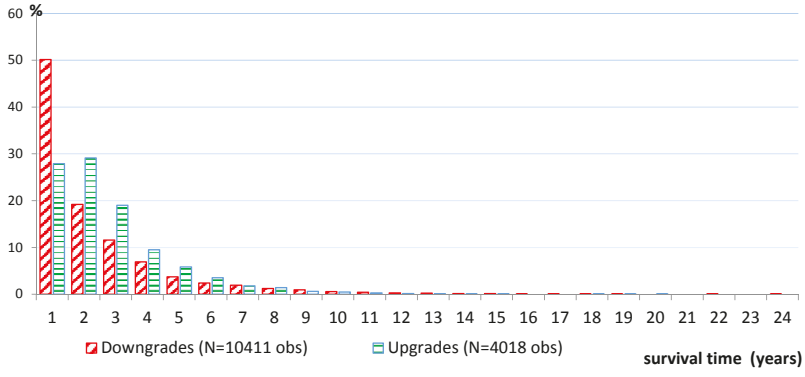
Table 2. Statistics of rating grades and survival time.

Panel A: Statistics of S&P's Numerical Rating Grades			
	Sample A(H): All Firms	**Sample B(H): Non-US Firms**	**Sample C(H-1): Crisis Sample**
Sample size	17,109	4745	3927
Mean	10.29	11.08	8.79
Median	10 (BB)	11 (BB+)	8 (B+)
Std dev	4.18	4.22	3.9
Min	1 (C)	1 (C)	2 (CC)
Max	21 (AAA)	21 (AAA)	20 (AA+)
Panel B: Statistics of Survival Time for Downgrades			
	Sample A(H): All Firms	**Sample B(H): Non-US Firms**	**Sample C(H-1): Crisis Sample**
Number of downgrades	10,411	2661	1825
Frequency of downgrades	60.85%	56.08%	46.5%
Mean (years)	1.82	1.55	0.66
Median (years)	1	0.92	0.38
Std dev	2.25	1.78	0.76
Min (years)	~0	0.01	0.01
Max (years)	23.43	14.28	7.41
Panel C: Statistics of Survival Time for Upgrades			
	Sample A(H): All Firms	**Sample B(H): Non-US Firms**	**Sample C(H-1): Crisis Sample**
Number of upgrades	4018	1096	521
Frequency of upgrades	23.48%	23.1%	13.3%
Mean (years)	2.25	1.99	1.34
Median (years)	1.74	1.55	1.01
Std dev	1.96	1.63	1.32
Min (years)	0.02	0.02	0.04
Max (years)	19.72	10.3	10.3

Panel A of Table 2 shows the statistics of rating grades, Panel B and Panel C present the statistics of survival time for downgrades and upgrades, respectively, in the whole sample A(H), non-U.S. sample B(H), and crisis sample C(H-1). For brevity reasons, this table presents the statistics for ratings in the three samples with Hofstede's culture values.

The histograms of survival times for downgrades and upgrades in sample A(H) (the entire sample) and sample C(H-1) (crisis times) are depicted in Figure 1. Overall, downgrades had a shorter survival time than upgrades and were heavily massed in durations shorter than a year. About 50% of downgrades during the study period and about 80% of downgrades during crisis times retained their current rating for less than a year.

Panel A: Distribution of rating observations by duration, 1985-2010



Panel B: Distribution of ratings by duration in crisis time

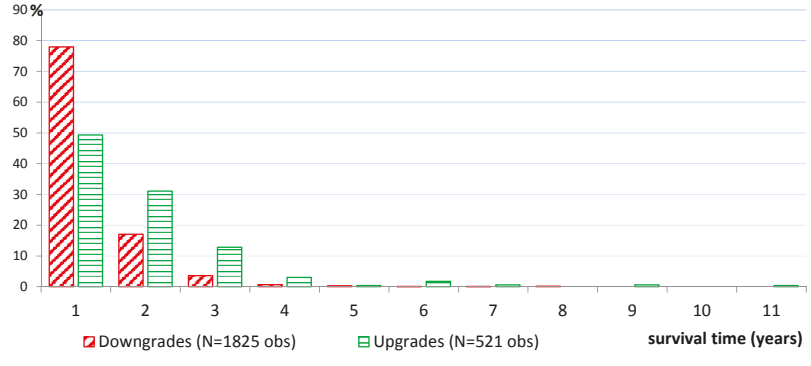


Figure 1. Distribution of rating observations by duration (survival time). Panel A and Panel B of Figure 1 shows the distribution of migrated ratings by duration (survival time) for the entire sample A(H) and crisis sample C(H-1), respectively. Duration is the length of time a firm stays in a rating grade measured from the time it enters the rating grade subsequent to the beginning of the study until the time it either migrates to another rating grade or becomes censored. A typical firm contributes multiple ratings to the dataset. Sample A(H) and sample C(H-1) includes 17,109 ratings and 3927 ratings, respectively.

Panels A and B of Table 3 give the descriptive statistics of the numeric scores on Hofstede's five culture values and Schwartz's two culture values for firms in samples A(H) and B(H), respectively. There is a wide range of variation in the scores for each of Hofstede's culture values. Compared with firms in sample A(H), non-U.S. firms in sample B(H) show characteristics of larger power distance (*PDI*), more feminine (*MAS*), more collectivistic (*IDV*), stronger risk-avoiding (*UAI*), and more long-term oriented (*LTO*) cultures.

Table 3. Descriptive statistics of national culture scores.

Panel A: Descriptive Statistics of Culture Values for Samples A(H) and A(H-S) of All Firms					
	Mean	**Median**	**Std Dev**	**Min**	**Max**
Hofstede (H) culture values (N = 17,109)					
Power distance index (PDI)	42.7	40	10.43	11	104
Individualism vs. collectivism (IDV)	83.26	91	17.64	12	91
Masculinity vs. femininity (MAS)	60.15	62	10.4	5	110
Uncertainty avoidance index (UAI)	50.28	46	13.46	8	112
Long-term vs. short-term orientation (LTO)	32.91	26	15.99	13	100
Schwartz (S) culture values (N = 16,966)					
Embeddedness	3.62	3.67	0.16	3.03	4.35
Hierarchy	2.34	2.37	0.17	1.49	3.23
Panel B: Descriptive Statistics of Culture Values for Sample B(H) and B(H-S) of Non-U.S. Firms					
	Mean	**Median**	**Std Dev**	**Min**	**Max**
Hofstede (H) culture values (N = 4745)					
Power distance index (PDI)	49.75	39	17.98	11	104
Individualism vs. collectivism (IDV)	63.08	71	23.63	12	90
Masculinity vs. femininity (MAS)	55.32	56	18.91	5	110
Uncertainty avoidance index (UAI)	61.44	53	21.94	8	112
Long-term vs. short-term orientation (LTO)	50.93	51	21.74	13	100
Schwartz (S) culture values (N = 4602)					
Embeddedness	3.49	3.46	0.27	3.03	4.35
Hierarchy	2.25	2.22	0.32	1.49	3.23

Table 3 shows the descriptive statistics for scores of Hofstede (H)'s and Schwartz (S)'s culture values for the whole samples and non-U.S. samples. The definitions of the culture variables are as in Table 1.

4. Results

Separate models are estimated for downgrades and upgrades. Overall, culture values have significant effects on downgrades and upgrades after accounting for time-varying rating outlook, various aspects of rating history, macro-economic conditions and political risks.

4.1. Models for the Whole Sample (Samples A(H) and A(H-S))

The results of the models for the whole sample are given in Table 4. For each migration outcome, three models are estimated and three sets of results are given: first, using dummy variables where each dummy is set according to a country's score on Hofstede's culture value relative to the mean (model 1); second, using numeric scores for Hofstede's culture values (model 2); and, third using numeric scores for Hofstede's and Schwarz's culture values (model 3).

As presented in Table 4, most of the retained variables are significant at the 10% level or better based on a Wald chi-square test. Culture values have stronger effects on downgrades than upgrades. The downgrade model with Hofstede and Schwartz's scores (model 3) features the significance of all five culture measures. The sign of a significant culture variable in model (3) is consistent with its sign (if significant) in model (1) and model (2).

The three models for downgrades feature one common culture value, which is *long-term orientation (LTO)/dummy LTO*, and its sign is as expected. Being in a *LTO* country reduces the downgrade probability by 25.2% (model 1) whereas a one-unit increase in *LTO* score makes a downgrade 0.9% (model 2) or 1.1% (model 3) less likely.[13]

Power distance index (PDI) is significant in model (2) for downgrades while *hierarchy* is present in model (3) for both downgrades and upgrades. Schwartz's score of *hierarchy* has a stronger effect than Hofstede's score of *PDI*. For example, a one-unit increase in *hierarchy* makes a downgrade 29.8% less

[13] Subtracting one from the hazard ratio (HR) gives the change in risk for a one-unit change in the independent variable. *Dummy LTO*'s HR of 0.748 represents a 25.2% reduction in downgrade risk for firms in a *LTO* country (model 1). *Dummy LTO* (model 1) has a stronger impact than *LTO* (models 2 and 3). A larger effect of *dummy LTO* is not unusual in hazard modelling, often because a switch from short-term to long-term orientation represents a substantial change.

likely and an upgrade 73.7% more likely (model 3). A one-unit increase in *PDI* reduces the downgrade hazard by only 1% (model 2). This is not surprising as the score of Hofstede's *PDI* value for a country is much higher than the respective score of Schwartz's *hierarchy* value (Table 3). The risk being changed is very small for a one-unit increase in *PDI*.

Apart from *LTO* and *PDI/hierarchy*, models (2) and (3) for downgrades share another common culture value which is *uncertainty avoidance index (UAI)*. As expected, firms residing in strong UA countries have unfavorable experiences toward downgrades.

Individualism (IDV)/ Dummy IDV is not significant in models (1) and (2). However, *embeddedness* is significant in model (3) for downgrades. *Embeddedness* is comparable to *collectivism* which is the opposite pole of *individualism*. As expected, firms in *embeddedness* cultures are less likely to experience a downgrade (model 3).

For upgrades, *hierarchy* is the only significant variable in model (3) (as discussed above) while *masculinity (MAS)/ dummy MAS* is the only significant variable in models (1) and (2). *MAS*'s positive sign in models (1) and (2) (for upgrades) and model 3 (for downgrades) suggests that firms in *MAS* cultures are more likely to experience rating changes in either direction.

With regard to control variables, the following discussion focuses on those significant in all three models for downgrades (upgrades). Firms rated around the speculative rating boundary (BB+/BB/BB) are more likely to become rising stars. Firms rated close to the investment rating threshold (BBB+/BBB/BBB−) or those with a positive outlook have lower downgrade risk and higher upgrade probability. Consistent with prior literature, issuers with a lagged downgrade or a negative outlook exhibit a strong tendency to travel downward on the rating spectrum. Older firms, firms with a high current rating, a volatile rating history or those which have experienced a fallen angel event tend to retain the current rating grade. Firms with a longer lagged rating or those with a prior substantial downgrade are more likely to experience a subsequent migration.

With regard to macro-economic environment, a strong global equity market is associated with a lower rating volatility, raising the probability that a firm will retain its current rating grade. Firms in OECD countries or countries with strong current account balances tend to go up the rating scales. Firms in countries high on inflation or countries with equity market-based economies experience fewer upgrades. A high GDP growth rate, a strong current account balance, and an emergence from a sovereign default lower the probability of corporate downgrades.

In terms of political risk, the effect of *dummy low political risk* is in contrast to initial expectations. Firms in countries with low political risk are more likely to be downgraded and less likely to be upgraded. Perhaps this surprising effect is due to the sample being dominated by U.S. firms.

A natural question is whether the impact of culture remains robust when the sample is restricted to non-U.S. firms and the study period is restricted to crisis/ non-crisis times.

4.2. Robustness Tests

4.2.1. Models for Non-U.S. Firms (Samples B(H) and B(H-S))

The results of the models for non-U.S. firms are given in Table 5. The evidence in favor of including culture variables is weaker in the upgrade models and stronger in the downgrade models (compared with the respective models in Table 4).

For downgrades, *dummy LTO/LTO* remains significant in three models. Consistent with the results presented in Table 4, firms in *LTO* cultures are less likely to be downgraded. Model (3) for downgrades features only four significant culture values compared with five significant culture values in the respective model for the whole sample (Table 4). *UAI* is no longer present in model (3) whereas *dummy UAI* and *dummy large PDI* become significant in model (1). The negative sign of *dummy large PDI* (model 1) is consistent with the negative sign of *PDI* (model 2), *hierarchy* (model 3) and the results presented in Table 4. Similarly, the positive sign of *dummy UAI* (model 1) is consistent with the positive sign of *UAI* (model 2).

For upgrades, model (1) does not feature any significant culture dummies. In model (3), *hierarchy* is no longer significant and is replaced by *MAS*. The positive effect of *MAS* on upgrades (models 2 and 3) and on downgrades (model 3) is as expected. Firms in *MAS* cultures are more volatile and exhibit a tendency to migrate from the current rating in either direction.

4.2.2. Models for Crisis and Non-Crisis Periods (Samples C(H-1), C(H-2), C(H-3))

Table 6 presents the results of the models for crisis sample C(H-1), non-crisis sample C(H-2), and non-crisis non-U.S. sample C(H-3). For reasons of brevity, only the results of the models with Hofstede's numeric scores are reported in Table 6.

Overall, the statistical significance of culture variables for crisis and non-crisis samples is weaker. Consistent with the results presented in Tables 4 and 5, *LTO* is the only culture value which is significant in the downgrade models for three samples. Firms in *LTO* cultures are less likely to be downgraded and during crisis times, are more likely to be upgraded.

UAI is only present in the downgrade model for the crisis sample. A strong UA culture (*UAI*) predisposes to more downgrades during crisis times. The effect is consistent with its effect on the whole sample (Table 4) and non-U.S. sample (Table 5). *PDI* is significant in the downgrade model for the crisis sample and the upgrade model for the non-crisis non-U.S. sample. A high PDI score reduces the downgrade hazard during crisis times and raises the upgrade probability of non-U.S. firms during non-crisis periods. The impact of *PDI* on downgrades during crisis times is consistent with its impact on the whole sample (Table 4) and non-U.S. sample (Table 5). *MAS* is not present in any downgrade models while *IDV* is not significant in any models.

The upgrade models for three samples do not share any significant common culture variable. However, *MAS* is significant in the upgrade models for the two non-crisis samples. Its positive sign suggests that firms in *MAS* cultures are more likely to be upgraded during non-crisis periods.

4.2.3. Other Robustness Tests

For each migration outcome, model (1) is re-estimated for three samples using dummy variables where each dummy is set according to a country's score on Hofstede's culture value relative to the *median* score. Three samples are used in this test: the whole sample (sample A(H)), non-U.S. sample (sample B(H)) and crisis sample (sample C(H-1)). Untabulated results across the three samples consistently show that firms in *LTO* countries (*dummy LTO*) are less likely to go down the rating scales. *Dummy LTO* is significant in both downgrade and upgrade models for the whole sample, and its opposite effects are as expected.

Dummy masculinity (MAS) is also significant in the models for the two large samples A(H) and B(H). The positive effect of *dummy MAS* on upgrades in sample A(H) is consistent with the result of the respective model when *dummy MAS* is set based on a country score relative to the mean score (Table 4, model 1 for upgrades).

While *dummy individualism/ IDV* is not significant in any models presented in Tables 4–6, it is significant when each dummy is set based on a country score relative to the median score. Firms in individualistic countries are more likely to be downgraded (for the whole sample and non-U.S. sample) and less likely to be upgraded (for the crisis sample).

Table 4. Cox's regression models for rating changes of all firms.

	Culture Dummy Mean (Hofstede, N = 17,109)				Numeric Culture Score (Hofstede, N = 17,109)				Numeric Score (Hofstede & Schwartz, N = 16,966)			
	Downgrade (1)		Upgrade (1)		Downgrade (2)		Upgrade (2)		Downgrade (3)		Upgrade (3)	
Variables	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**
Hofstede's national culture dimensions												
Power distance index (PDI)	NA		NA		−0.01 ***	0.99			NA		NA	
Individualism vs. collectivism (IDV)	NA		NA						NA		NA	
Masculinity vs. femininity (MAS)	NA		NA				0.004 **	1.004	0.006 ***	1.006		
Uncertainty avoidance index (UAI)	NA		NA		0.009 ***	1.009			0.004 ***	1.004		
Long-term vs. short-term orientation (LTO)	NA		NA		−0.009 ***	0.991			-0.011 ***	0.989		
Dummy large power distance index					NA		NA		NA		NA	
Dummy individualism					NA		NA		NA		NA	
Dummy masculine			0.17907 ***	1.196	NA		NA		NA		NA	
Dummy strong uncertainty avoidance					NA		NA		NA		NA	
Dummy long-term orientation	−0.29086 ***	0.748			NA		NA		NA		NA	
Schwartz's national culture dimensions												
Embeddedness	NA		NA		NA		NA		−0.274 **	0.76		
Hierarchy	NA		NA		NA		NA		−0.353 ***	0.702	0.552 ***	1.737
S&P's rating data												
Current rating grade	−0.01021 ***	0.99	−0.10447 ***	0.901	−0.009 **	0.991	−0.106 ***	0.9	−0.01 **	0.99	−0.108 ***	0.898
Investment rating boundary (BBB−/BBB/BBB+)	−0.23186 ***	0.793	0.16734 ***	1.182	−0.241 ***	0.786	0.167 ***	1.182	−0.24 ***	0.786	0.164 ***	1.179
Junk rating boundary (BB−/BB/BB+)			0.14226 ***	1.153			0.143 ***	1.154			0.141 ***	1.152
Dummy negative outlook (time-varying)	0.29138 ***	1.338	−1.7834 ***	0.168	0.285 ***	1.33	−1.783 ***	0.168	0.282 ***	1.326	−1.776 ***	0.169
Dummy positive outlook (time-varying)	−1.73029 ***	0.177	1.206 ***	3.34	−1.721 ***	0.179	1.209 ***	3.35	−1.713 ***	0.18	1.207 ***	3.345
Logarithm of age since first rated (time-varying)	−1.57095 ***	0.208	−1.45531 ***	0.233	−1.556 ***	0.211	−1.439 ***	0.237	−1.552 ***	0.212	−1.434 ***	0.238
Dummy lag one downgrade	0.66822 ***	1.951			0.659 ***	1.933			0.665 ***	1.945		
Lag one rating duration	0.06974 ***	1.072	0.03321 ***	1.034	0.07 ***	1.072	0.033 ***	1.034	0.07 ***	1.072	0.033 ***	1.033
Dummy prior fallen angel event(s)	−0.13936 ***	0.87			−0.137 ***	0.872			−0.119 ***	0.888		
Dummy large downgrade	0.15244 ***	1.165	0.28276 ***	1.327	0.171 ***	1.187	0.285 ***	1.33	0.163 ***	1.177	0.281 ***	1.325
Dummy large upgrade									0.088 *	1.092		
Rating volatility	−0.05591	0.946	−0.13391 ***	0.875	−0.053	0.948	−0.134 ***	0.874	−0.051	0.951	−0.135 ***	0.874
Macro-economic and financial conditions												
Dummy prior default	−0.47471 ***	0.622			−0.536 ***	0.585			−0.442 ***	0.643	−0.24 *	0.787
Dummy debt crisis			−0.1048 *	0.901			−0.112 **	0.894				
Dummy OECD member			0.23915 ***	1.27	−0.249 ***	0.78	0.306 ***	1.358	−0.269 ***	0.764	0.212 *	1.236
Logarithm of GDP per capita											0.164 **	1.178
Change in real GDP growth rate	−0.02713 ***	0.973			−0.028 ***	0.972			−0.028 ***	0.973		
Change in inflation			−0.02144 ***	0.979			−0.022 ***	0.979			−0.022 ***	0.978
Change in current account surplus/GDP	−0.04317 ***	0.958	0.07859 ***	1.082	−0.04 ***	0.961	0.079 ***	1.082	−0.044 ***	0.957	0.069 ***	1.071
Change in term trade												
Logarithm of ratio stock market cap/GDP			−0.09241 ***	0.912	0.06 ***	1.062	−0.078 ***	0.925	0.08 ***	1.083	−0.161 ***	0.852
Return of world stock market index	−0.58725 ***	0.556	−0.24199 **	0.785	−0.525 ***	0.592	−0.228 *	0.796	−0.532 ***	0.588	−0.251 **	0.778

Table 4. *Cont.*

	Culture Dummy Mean (Hofstede, N = 17,109)				Numeric Culture Score (Hofstede, N = 17,109)				Numeric Score (Hofstede & Schwartz, N = 16,966)			
	Downgrade (1)		Upgrade (1)		Downgrade (2)		Upgrade (2)		Downgrade (3)		Upgrade (3)	
Variables	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**
Political risks												
Dummy low political risk	0.09715 ***	1.102	−0.08271 **	0.921	0.067 **	1.069	−0.094**	0.911	0.072 ***	1.074	−0.066 *	0.936
Dummy high Political risk												
Events/ sample size	60.85%		23.48%		60.85%		23.48%		60.97%		23.44%	
Likelihood ratio χ^2	8319.85 ***		5022.41 ***		8379.83 ***		5017.3 ***		8350.8 ***		4977.2 ***	

Table 4 presents the results of Cox's hazard models for upgrades and downgrades in the entire samples. Downgrades (upgrades) were treated as events (censored) in the downgrade model and vice versa. The scores for Hofstede's culture values are coded as dummy variables (model 1) and numeric variables (model 2 and model 3). The scores for Schwartz's culture values are coded as numeric variables (model 3). In model (1), scores greater than or equal to the mean take a value of one or zero otherwise. Significant variables were retained in the model using the stepwise selection procedure. A variable has to be significant at the 0.25 level before it can be entered into the model, and a variable has to be significant at the 0.1 level for it to remain in the model. Parameter estimates are given first followed by the corresponding *p*-values based on Wald chi-square test, with ***, **, * representing significance at the 0.01, 0.05 and 0.10 levels, respectively, using the Wald test. Subtracting 1 from the hazard ratio (HR) gives the percentage change in the hazard for a one-unit change in the independent variable. The likelihood ratio is based on a comparison of the model with variables added versus a model without variables, and tests the null hypothesis that all variables included in the model have coefficients of 0.

Table 5. Cox's regression models for rating changes of non-U.S. firms.

	Culture Dummy Mean (Hofstede, N = 4745)				Numeric Culture Score (Hofstede, N = 4745)				Numeric Score (Hofstede & Schwartz, N = 4602)			
	Downgrade (1)		Upgrade (1)		Downgrade (2)		Upgrade (2)		Downgrade (3)		Upgrade (3)	
Variables	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**
Hofstede's national culture dimensions												
Power distance index (PDI)	NA		NA		−0.008 ***	0.993			NA		NA	
Individualism vs. collectivism (IDV)	NA		NA						NA		NA	
Masculinity vs. femininity (MAS)	NA		NA				0.003 *	1.003	0.004 ***	1.004	0.003 *	1.003
Uncertainty avoidance index (UAI)	NA		NA		0.004 ***	1.004						
Long-term vs. short-term orientation (LTO)	NA		NA		−0.007 ***	0.993			−0.008 ***	0.992		
Dummy large power distance index	−0.37814 ***	0.685			NA		NA		NA		NA	
Dummy individualism					NA		NA		NA		NA	
Dummy masculine					NA		NA		NA		NA	
Dummy strong uncertainty avoidance	0.17973 ***	1.197			NA		NA		NA		NA	
Dummy long-term orientation	−0.1631 ***	0.85			NA		NA		NA		NA	
Schwartz's national culture dimensions												
Embeddedness	NA		NA		NA		NA		−0.448 ***	0.639		
Hierarchy	NA		NA		NA		NA		−0.337 ***	0.714		
S&P's rating data												
Current rating grade	−0.10145 ***	0.904	−0.10145 ***	0.904			−0.103 ***	0.902			−0.102 ***	0.904
Investment rating boundary (BBB−/BBB/BBB+)	0.25624 ***	1.292	0.25624 ***	1.292	−0.292 ***	0.747	0.25 ***	1.284	−0.263 ***	0.769	0.219 ***	1.244
Junk rating boundary (BB−/BB/BB+)	0.17256 **	1.188	0.17256 **	1.188			0.17 **	1.185			0.16 **	1.173

Table 5. *Cont.*

	Culture Dummy Mean (Hofstede, N = 4745)				Numeric Culture Score (Hofstede, N = 4745)				Numeric Score (Hofstede & Schwartz, N = 4602)			
	Downgrade (1)		Upgrade (1)		Downgrade (2)		Upgrade (2)		Downgrade (3)		Upgrade (3)	
Variables	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**
Dummy negative outlook (time-varying)	−1.61499 ***	0.199	−1.61499 ***	0.199	0.354 ***	1.425	−1.618 ***	0.198	0.341 ***	1.406	−1.576 ***	0.207
Dummy positive outlook (time-varying)	1.27989 ***	3.596	1.27989 ***	3.596	−1.787 ***	0.167	1.282 ***	3.605	−1.767 ***	0.171	1.306 ***	3.691
Logarithm of age since first rated (time-varying)	−2.77147 ***	0.063	−2.77147 ***	0.063	−2.299 ***	0.1	−2.765 ***	0.063	−2.335 ***	0.097	−2.742 ***	0.064
Dummy lag one downgrade	−0.13508 **	0.874	−0.13508 **	0.874	0.598 ***	1.818	−0.133 **	0.876	0.575 ***	1.777	−0.13 **	0.878
Lag one rating duration	0.15485 ***	1.167			0.167 ***	1.181	0.156 ***	1.169	0.184 ***	1.202	0.16 ***	1.173
Dummy prior fallen angel event(s)	0.22181 **	1.248					0.227 **	1.255	0.174 **	1.19	0.225 **	1.253
Dummy large downgrade												
Dummy large upgrade	−0.24784 **	0.78			−0.232 *	0.793			−0.189	0.828		
Rating volatility	−0.07356	0.929	−0.21696 **	0.805	−0.075	0.928	−0.215 **	0.807	−0.067	0.936	−0.204 **	0.816
Macro-economic and financial conditions												
Dummy prior default	−0.45461 ***	0.635			−0.4 ***	0.671			−0.391 ***	0.676		
Dummy debt crisis	0.16082 **	1.174			0.172 ***	1.188			0.171 **	1.186		
Dummy OECD member									−0.191 **	0.826		
Logarithm of GDP per capita												
Change in real GDP growth rate	−0.02736 ***	0.973			−0.03 ***	0.971			−0.031 ***	0.97		
Change in inflation	−0.016 ***	0.984	−0.01629 ***	0.984	−0.017 ***	0.983	−0.016 ***	0.984	−0.017 ***	0.984	−0.017 ***	0.983
Change in current account surplus/GDP			0.03483 ***	1.035			0.035 ***	1.036			0.037 ***	1.038
Change in term trade												
Logarithm of ratio stock market cap/GDP												
Return of world stock market index	−0.60958 ***	0.544			−0.609 ***	0.544			−0.683 ***	0.505		
Political risks												
Dummy low political risk												
Dummy high Political risk												
Events/ sample size	56.08%		23.1%		56.08%		23.1%		56.37%		22.92%	
Likelihood ratio χ^2	2858.04 ***		1681.21 ***		2874 ***		1684.7 ***		2860 ***		1630.5 ***	

Table 5 presents the results of Cox's hazard models for rating upgrades and downgrades in the samples of non-U.S. firms. Downgrades (upgrades) were treated as events (censored) in the downgrade model and vice versa. The scores for Hofstede's culture values are coded as dummy variables (model 1) and numeric variables (model 2 and model 3). The scores for Schwartz's culture values are coded as numeric variables (model 3). In model (1), scores greater than or equal to the mean take a value of one and zero otherwise. Significant variables were retained in the model using the stepwise selection procedure. A variable has to be significant at the 0.25 level before it can be entered into the model, and a variable in the model has to be significant at the 0.1 level for it to remain in the model. Parameter estimates are given first followed by the corresponding *p*-values based on Wald chi-square test, with ***, **, * representing significance at the 0.01, 0.05 and 0.10 levels, respectively, using the Wald test. Subtracting 1 from the hazard ratio (HR) gives the percentage change in the hazard for a one-unit change in the independent variable. The likelihood ratio is based on a comparison of the model with variables added versus a model without variables, and tests the null hypothesis that all variables included in the model have coefficients of 0.

Table 6. Cox's regression model for corporate rating changes: crisis vs. non-crisis periods.

	Crisis Sample (Hofstede, N = 3927)				Non-Crisis Sample (Hofstede, N = 13,182)				Non-Crisis Non-U.S. Sample (Hofstede, N = 3714)			
	Downgrade		Upgrade		Downgrade		Upgrade		Downgrade		Upgrade	
Variables	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**	**Coefficient**	**HR**
Hofstede's national culture dimensions												
Power distance index (PDI)	−0.019 ***	0.981									0.00405 **	1.004
Individualism vs. collectivism (IDV)												
Masculinity vs. femininity (MAS)							0.00442 ***	1.004			0.00326 *	1.003
Uncertainty avoidance index (UAI)	0.009 ***	1.009										
Long-term vs. short-term orientation (LTO)	−0.005 ***	0.995	0.015 ***	1.015	−0.0068 ***	0.993			−0.00473 ***	0.995		
S&P's rating data												
Current rating grade	−0.055 ***	0.947	−0.281 ***	0.755			−0.09783 ***	0.907			−0.105 ***	0.9
Investment rating boundary (BBB−/BBB/BBB+)	−0.417 ***	0.659	0.445 *	1.56	−0.21511 ***	0.806	0.18773 ***	1.207	−0.22824 ***	0.796	0.29756 ***	1.347
Junk rating boundary (BB−/BB/BB+)			0.472 ***	1.603			0.15556 ***	1.168			0.17947 **	1.197
Dummy negative outlook (time-varying)	0.406 ***	1.501	−1.512 ***	0.22	0.2374 ***	1.268	−1.93868 ***	0.144	0.26543 ***	1.304	−1.97581 ***	0.139
Dummy positive outlook (time-varying)	−1.668 ***	0.189	1.327 ***	3.769	−1.72525 ***	0.178	1.17788 ***	3.247	−1.91491 ***	0.147	1.20023 ***	3.321
Logarithm of age since first rated (time-varying)	−1.048 ***	0.351	−0.69 ***	0.502	−1.73288 ***	0.177	−1.47874 ***	0.228	−3.19282 ***	0.041	−2.99776 ***	0.05
Dummy lag one downgrade	0.877 ***	2.403			0.61724 ***	1.854			0.58564 ***	1.796		
Lag one rating duration	0.066 ***	1.068			0.07298 ***	1.076	0.03799 ***	1.039	0.18513 ***	1.203	0.17943 ***	1.197
Dummy prior fallen angel event(s)					−0.19549 ***	0.822					0.30926 ***	1.362
Dummy large downgrade	0.194 ***	1.214			0.17731 ***	1.194	0.32772 ***	1.388			0.17525	1.192
Dummy large upgrade					0.14199 ***	1.153					−0.26509 **	0.767
Rating volatility					−0.107	0.899	−0.10982 **	0.896	−0.46349 ***	0.629	−0.19194 **	0.825
Macro-economic and financial conditions												
Dummy prior default	−0.654 ***	0.52										
Dummy OECD member	−1.008 ***	0.365			0.28907 ***	1.335	0.2497 ***	1.284	0.42189 ***	1.525		
Logarithm of GDP per capita			0.564 ***	1.758								
Change in real GDP growth rate	−0.055 ***	0.946										
Change in inflation	−0.019 ***	0.981	−0.027 ***	0.973								
Change in current account surplus/GDP	−0.063 ***	0.939	0.142 ***	1.152			0.04871 ***	1.05				
Change in term trade			0.000002 ***	1			0.000037 ***	1			0.00003 ***	1
Logarithm of ratio stock market cap/GDP					0.11571 ***	1.123	−0.10045 ***	0.904				
Return of world stock market index	−0.367 ***	0.693	−0.61 ***	0.544	−0.45234 ***	0.636			−0.47447 ***	0.622		
Political risks												
Dummy low political risk	−0.198 ***	0.82			0.11886 ***	1.126	-0.10071 **	0.904				
Dummy high Political risk												
Events/ sample size	46.47%		13.3%		65.13%		26.53%		56.22%		26.17%	
Likelihood ratio χ^2	1825.3 ***		697 ***		6808.65 ***		4414.7 ***		2367.02 ***		1568.68 ***	

Table 6 presents the results of Cox's hazard models for corporate ratings pooled across countries during the time they were (were not) experiencing a debt crisis or a banking crisis as defined in Manasse et al. (2003), Laeven and Valencia (2008), and De Paoli et al. (2009). For brevity reasons, only the results of the models with Hofstede's numeric scores are reported. Downgrades (upgrades) were treated as events (censored) in the downgrade model and vice versa. Significant variables were retained in the model using the stepwise selection procedure. A variable has to be significant at the 0.25 level before it can be entered into the model, and a variable has to be significant at the 0.1 level for it to remain in the model. Parameter estimates are given first followed by the corresponding *p*-values based on Wald chi-square test, with ***, **, * representing significance at the 0.01, 0.05 and 0.10 levels, respectively, using the Wald test. Subtracting 1 from the hazard ratio (HR) gives the percentage change in the hazard for a one-unit change in the independent variable. The likelihood ratio is based on a comparison of the model with variables added versus a model without variables, and tests the null hypothesis that all variables included in the model have coefficients of 0.

5. Conclusions

The informal constraints (social influences) that arise from the national culture in which a firm resides have a pervasive impact on managerial decision making and corporate credit risk, which in turn impacts on corporate ratings and rating changes. Some cultures may lead to more rating changes whereas other cultures may result in a directional predisposition for rating changes (downgrade *or* upgrade). This study is the first attempt to explore the effects of national culture values established by Hofstede et al. (2010) and Schwartz (1994) on the rating migration dynamics of 5360 firms across 50 countries over the period 1985–2010. The study enriches the literature by presenting empirical evidence that national culture provides a better explanation of corporate rating migrations. The effects of culture are significant after accounting for variables which have been found significant in previous studies on rating migrations. The study overcomes computational challenges in estimating stratified dynamic hazard models (with time-varying variables) for a large sample of 17,109 ratings.

Overall, the evidence in favor of including culture variables is generally stronger in the downgrade models or when numeric scores are employed to represent culture. The evidence of statistical significance is strongest for the effect of *long-term orientation*. Firms located in *long-term oriented* cultures (*LTO*) are less likely to be downgraded and in some cases, more likely to climb up the rating scales. The effect of *LTO* on downgrades is robust to alternative samples, alternative measures of culture and alternative study periods. There is some evidence that downgrades occur more often in strong *uncertainty-avoiding* countries and less often in large *power distance* (*hierarchy*) and *embeddedness* countries. *Masculinity* predisposes to a higher rating volatility, raising the probability of downgrades and upgrades.

This study is somewhat limited as apart from credit rating, other firm-specific data is not available. In the absence of variables such as firm size, it is *not* possible to examine the effects of culture on rating *levels*. This study therefore focuses on the effects of national culture on the probability of rating *migrations*. The estimated models of rating changes include the current rating, which captures firm-specific characteristics such as firm size and leverage, and time-varying rating outlooks, which signal the potential direction of a long-term credit rating over the intermediate term. Thus, the lack of firm-specific data such as firm size is not a substantial concern for this study.

This study emphasizes the need to understand the role of culture in managerial decision making and corporate policies, which feed into the rating *review* process. Studying culture helps enrich our understanding of corporate rating migration dynamics. This knowledge in turn can be helpful in developing predictive models of corporate rating changes across countries. The results of this study have practical implications to investors who use credit ratings to make investment decisions.

Acknowledgments: I would like to thank two anonymous reviewers, the editors, and Jean Helwege, Gady Jacoby, Andreas Knetsch, Bing Li, Graham Partington, as well as participants at the 2017 Cross Country Perspectives in Finance Symposium for helpful suggestions. Thanks are also due to George Griffiths for proofreading this article, and to the Knowledge Unlatched and the MDPI for supporting this manuscript. The rating data used in this study was purchased thanks to the generous funding from the University of Munich (LMU). I am grateful for Andreas Richter, Susanne Weber, and the Center for Advanced Studies at LMU for the funding support.

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A

Appendix A outlines different channels over which national culture impacts on corporate rating migration dynamics.

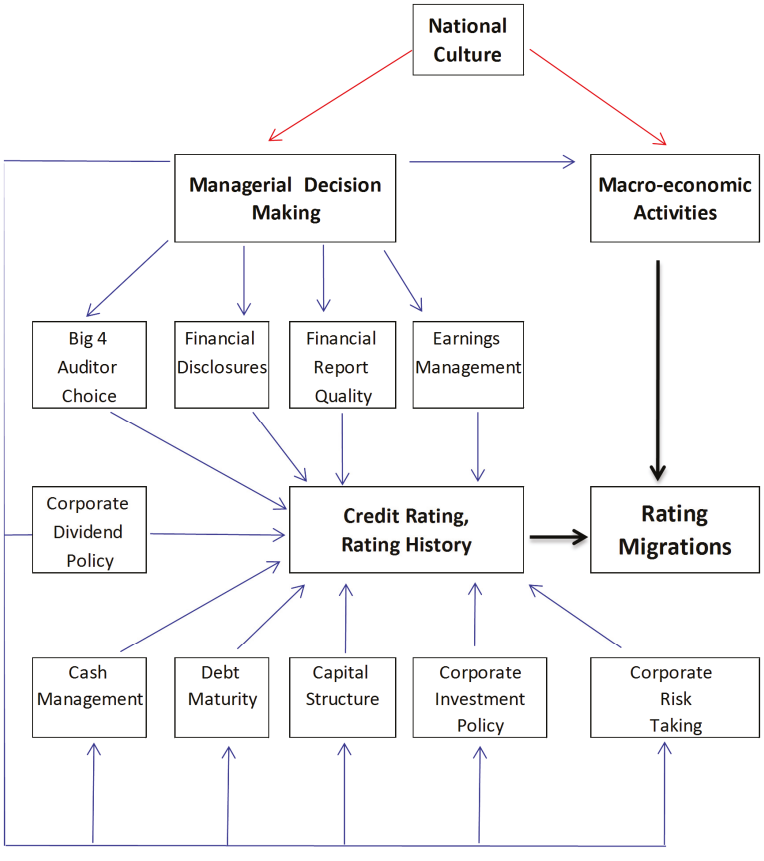


Figure A1. Possible links between national culture and corporate rating migrations.

References

Afego, Pyemo N. 2018. Index shocks, investor action and long-run stock performance in Japan: A case of cultural behaviouralism? *Journal of Behavioral and Experimental Finance* 18: 54–66. [CrossRef]

Allison, Paul D. 1995. *Survival Analysis Using SAS: A Practical Guide*. Cary: SAS Institute Inc. ISBN 978-1-59994-884-3.

Altman, Edward. 1998. The importance and subtlety of credit rating migration. *Journal of Banking and Finance* 22: 1231–47. [CrossRef]

Altman, Edward I., and Duen L. Kao. 1992. The implications of corporate bond rating drifts. *Financial Analysts Journal* 48: 64–75. [CrossRef]

Anderson, Christopher W., Mark Fedenia, Mark Hirschey, and Hilla Skiba. 2011. Cultural influences on home bias and international diversification by institutional investors. *Journal of Banking and Finance* 35: 916–34. [CrossRef]

Ashraf, Badar N., Changjun Zheng, and Sidra Arshad. 2016. Effects of national culture on bank risk-taking behavior. *Research in International Business and Finance* 27: 309–26. [CrossRef]

Bae, Sung C., Kiyoung Chang, and Eun Kang. 2012. Culture, corporate governance and dividend policy: International evidence. *The Journal of Financial Research* 35: 289–316. [CrossRef]

Bangia, Anil, Francis X. Diebold, André Kronimus, Christian Schagen, and Til Schuermann. 2002. Rating Migrations and the Business Cycle, with Applications to Credit Portfolio Stress Testing. *Journal of Banking and Finance* 26: 445–74. [CrossRef]

Barber, Brad M., and Terrance Odean. 2001. Boys will be boys: Gender, overconfidence and common stock investment. *Quarterly Journal of Economics* 116: 261–92. [CrossRef]

Blume, Marshall E., Donald B. Keim, and Sandeep A. Patel. 1991. Returns and volatility of low-grade bonds, 1977–1989. *Journal of Finance* 46: 49–74. [CrossRef]

Boubakri, Narjess, Ali Mirzaei, and Anis Samet. 2017. National culture and bank performance: Evidence from the recent financial crisis. *Journal of Financial Stability* 29: 36–56. [CrossRef]

Byrnes, James P., David C. Miller, and William Schafer. 1999. Gender Differences in Risk Taking: A Meta-Analysis. *Psychology Bulletin* 75: 367–83. [CrossRef]

Carty, Lee V., and Jerome S. Fons. 1994. Measuring Changes in Corporate Credit Quality. *Journal of Fixed Income* 4: 27–41. [CrossRef]

Chang, Chih-Hsiang, and Shih-Jia Lin. 2015. The effects of national culture and behavioral pitfalls on investors' decision-making: Herding behavior in international stock markets. *International Review of Economics and Finance* 37: 380–92. [CrossRef]

Chang, Kiyoung, and Abbas Noorbakhsh. 2009. Does national culture affect international corporate cash holdings? *Journal of Multinational Financial Management* 19: 323–42. [CrossRef]

Chui, Andy, Alison E. Lloyd, and Chuck C. Y. Kwok. 2002. The determination of capital structure: Is national culture a missing piece to the puzzle? *Journal of International Business Studies* 33: 99–127. [CrossRef]

Cohen, Jeffrey R., Laurie W. Pant, and David J. Sharp. 1996. A methodological note on cross-cultural accounting ethics research. *International Journal of Accounting* 31: 55–66. [CrossRef]

Cox, David R. 1972. Regression models and life tables. *Journal of Royal Statistical Society Series B (Methodological)* 34: 187–220. [CrossRef]

Dang, Huong, and Graham Partington. 2014. Rating migrations: The effects of rating history and time. *ABACUS* 50: 174–202. [CrossRef]

De Mooij, Marieke. 2004. *Consumer Behaviour and Culture: Consequence for Global Marketing and Advertising*. Thousand Oaks: Sage. ISBN 9781412979900.

De Paoli, Bianca, Glenn Hoggarth, and Victoria Saporta. 2009. Output Costs of Sovereign Crises: Some Empirical Estimates. Bank of England Working Paper No. 362. Available online: https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2009/output-costs-of-sovereign-crises-some-emperical-estimates.pdf (accessed on 2 November 2018).

Desender, Kurt A., Christian E. Castro, and Sergio A. Escamilla De Leon. 2011. Earnings management and cultural values. *American Journal of Economics and Sociology* 70: 639–70. [CrossRef]

Doupnik, Timothy S., and George T. Tsakumis. 2004. A critical review of tests of Gray's theory of cultural relevance and suggestions for future research. *Journal of Accounting Literature* 23: 1–48.

Ferri, Giovanni, Li-gang Liu, and Joseph E. Stiglitz. 1999. The procyclical role of rating agencies: Evidence from the East Asian crisis. *Economic Notes* 28: 335–55. [CrossRef]

Fidrmuc, Jana P., and Marcus Jacob. 2010. Culture, agency costs and dividends. *Journal of Comparative Economics* 38: 321–39. [CrossRef]

Figlewski, Stephen, Halina Frydman, and Weijian Liang. 2012. Modeling the Effects of Macroeconomic Factors on Corporate Default and Credit Rating Transitions. *International Review of Economics and Finance* 21: 87–105. [CrossRef]

Fonseca, Ana R., and Francisco Gonzalez. 2008. Cross country determinants of bank income smoothing by managing loan-loss provisions. *Journal of Banking and Finance* 32: 217–28. [CrossRef]

Gray, Sidney. 1988. Towards a theory of cultural influence on the development of accounting systems internationally. *ABACUS* 24: 1–15. [CrossRef]

Gray, Sidney J., and Hazel M. Vint. 1995. The impact of culture on accounting disclosures: Some international evidence. *Asia Pacific Journal of Accounting* 21: 33–43. [CrossRef]

Griffin, Dale, Omrane Guedhami, Chuck C. Y. Kwok, Kai Li, and Liang Shao. 2015. National Culture, Corporate Governance Practices and Firm Performance. Working Paper. Available online: https://memento.epfl.ch/public/upload/files/Paper_Li.pdf (accessed on 2 November 2018).

Guiso, Luigi, Helios Herrera, and Massimo Morelli. 2016. Cultural differences and institutional integration. *Journal of International Economics* 99: 97–113. [CrossRef]

Güttler, André. 2011. Lead-lag Relationships and Rating Convergence among Credit Rating Agencies. *Journal of Credit Risk* 7: 95–119. [CrossRef]

Güttler, André, and Mark Wahrenburg. 2007. The adjustment of credit ratings in advance of defaults. *Journal of Banking and Finance* 31: 751–67. [CrossRef]

Han, Sam, Tony Kang, Stephen Salter, and Yong Keun Yoo. 2010. A cross-country study on the effects of national culture on earnings management. *Journal of International Business Studies* 41: 123–41. [CrossRef]

Hill, Paula, Robert Brooks, and Robert Faff. 2010. Variations in sovereign credit quality assessments across rating agencies. *Journal of Banking and Finance* 34: 1324–43. [CrossRef]

Hofstede, Geert. 1980. *Culture's Consequences: International Differences in Work-Related Values*. Beverly Hills: Sage. ISBN 9780803913066.

Hofstede, Geert. 2001. *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. Thousand Oaks: Sage. ISBN 9780803973244.

Hofstede, Geert, Geert Hofstede, and Michael Minkov. 2010. *Cultures and Organizations: Software of the Mind*, 3rd ed. New York: McGraw-Hill. ISBN 978-0071664189.

Hope, Ole-Kristian. 2003. Firm-level disclosures and the relative roles of culture and legal origin. *Journal of International Financial Management and Accounting* 14: 218–48. [CrossRef]

Hope, Ole-Kristian, Tony Kang, Wayne Thomas, and Yong Keun Yoo. 2008. Culture and auditor choice: A test of the secrecy hypothesis. *Journal of Accounting Public Policy* 27: 357–73. [CrossRef]

Hosmer, David W., Stanley Lemeshow, and Susanne May. 2008. *Applied Survival Analysis: Regression Modeling of Time-to-Event Data*, 2nd ed. New York: Wiley. ISBN 978-0-471-75499-2.

Husted, Bryan. 1999. Wealth culture and corruption. *Journal of International Business Studies* 3: 339–59. [CrossRef]

Jensen, Michael, and William Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3: 305–60. [CrossRef]

Johnson, Richard. 2004. Rating agency actions around the investment grade boundary. *Journal of Fixed Income* 13: 25–37. [CrossRef]

Johnson, James, and Tomasz Lenartowicz. 1998. Culture freedom and economic growth: Do cultural values explain economic growth? *Journal of World Business* 33: 332–56. [CrossRef]

Jorion, Philippe, Charles Shi, and Sanjian Zhang. 2009. Tightening Credit Standards: The Role of Accounting Quality. *Review of Accounting Study* 14: 123–60. [CrossRef]

Kanagaretnam, Kiridaran, Chee Yeow Lim, and Gerald Lobo. 2011. Effects of National Culture on Earnings Quality of Banks. *Journal of International Business Studies* 42: 853–74. [CrossRef]

Kwok, Chuck C. Y., and Solomon Tadesse. 2006. National culture and financial systems. *Journal of International Business Studies* 37: 227–47. [CrossRef]

Laeven, Luc, and Fabian Valencia. 2008. Systemic Banking Crises: A New Database. IMF Working Paper WP/08/224. Available online: https://www.imf.org/external/pubs/ft/wp/2008/wp08224.pdf (accessed on 2 November 2018).

Lando, David, and Torben M. Skodeberg. 2002. Analyzing ratings transitions and rating drift with continuous observations. *Journal of Banking and Finance* 26: 423–44. [CrossRef]

Leung, Kwok, Rabi S. Bhagat, Nancy R. Buchan, Miriam Erez, and Cristina B. Gibson. 2005. Culture and international business: Recent advances and their implications for future research. *Journal of International Business Studies* 36: 357–78. [CrossRef]

Li, Zhichuan. 2016. Endogeneity in CEO power: A survey and experiment. *Investment Analysts Journal* 45: 149–62. [CrossRef]

Li, Kai, Dale Griffin, Heng Yue, and Longkai Zhao. 2013. How does culture influence corporate risk-taking? *Journal of Corporate Finance* 23: 1–22. [CrossRef]

Licht, Amir, Chanan Goldschmidt, and Shalom Schwartz. 2005. Culture, law and corporate governance. *International Review of Law and Economics* 25: 229–55. [CrossRef]

Lobo, Gerald, Luc Paugam, Herve Stolowy, and Pierre Astolfi. 2017. The Effect of Business and Financial Market Cycles on Credit Ratings: Evidence from the Last Two Decades. *ABACUS* 53: 59–93. [CrossRef]

Manasse, Paolo, Nouriel Roubini, and Axel Schimmelpfennig. 2003. Predicting Sovereign Debt Crisis. IMF Working Paper 03/221. Available online: https://www.imf.org/en/Publications/WP/Issues/2016/12/30/Predicting-Sovereign-Debt-Crises-16951 (accessed on 2 November 2018).

Mann, Christopher, David Hamilton, Praveen Varma, and Richard Cantor. 2003. *What Happens to Fallen Angels? A Statistical Review 1982–2003*. Moody's Investor Service Special Comment. Available online: https://www.moodys.com/sites/products/DefaultResearch/2002000000425343.pdf (accessed on 2 November 2018).

Mora, Nada. 2006. Sovereign credit ratings: Guilty beyond reasonable doubt? *Journal of Banking and Finance* 30: 2041–62. [CrossRef]

Nickell, Pamela, William Perraudin, and Simone Varotto. 2000. Stability of ratings transitions. *Journal of Banking and Finance* 24: 203–22. [CrossRef]

North, Douglass. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press. ISBN 9780521397346.

Pan, Yihui, Stephen Siegel, and Tracy Yue Wang. 2017. Corporate Risk Culture. *Journal of Financial and Quantitative Analysis* 52: 2327–67. [CrossRef]

Paredes, Angel A. P., and Clark Wheatley. 2017. The influence of culture on real earnings management. *International Journal of Emerging Markets* 12: 38–57. [CrossRef]

PricewaterhouseCoopers. 2008. Reward: A New Paradigm? Available online: https://www.pwc.com/gx/en/banking-capital-markets/pdf/reward.pdf (accessed on 2 November 2018).

Ramirez, Andrés, and Solomon Tadesse. 2009. Corporate Cash Holdings, Uncertainty Avoidance and the Multinationality of Firms. *International Business Review* 18: 387–403. [CrossRef]

S&P Global Ratings' Credit Research. 2013. *Default Study: Sovereign Defaults and Rating Transition Data, 2012 Update*. New York: Alacra Store.

S&P RatingsDirect. 2009. Use of CreditWatch and Outlooks. September 14. Available online: http://www.maalot.co.il/publications/MT20131212112758c.pdf (accessed on 2 November 2018).

S&P RatingsDirect. 2013. Corporate Methodology. November 19. Available online: https://www.spratings.com/scenario-builder-portlet/pdfs/CorporateMethodology.pdf (accessed on 2 November 2018).

Schwartz, Shalom H. 1994. Beyond Individualism-Collectivism: New Dimensions of Values. In *Individualism and Collectivism: Theory, Method and Application*. Edited by Uichol Kim, Harry C. Triandis, Çiğdem Kâğitçibaşi, Sang-Chin Choi and Gene Yoon. Thousand Oaks: Sage. ISBN 978-0803957633.

Shao, Liang, Chuck C. Y. Kwok, and Omrane Guedhami. 2010. National culture and dividend policy. *Journal of International Business Studies* 41: 1391–414. [CrossRef]

Shao, Liang, Chuck C. Y. Kwo, and Ran Zhang. 2013. National culture and corporate investment. *Journal of International Business Studies* 44: 745–63. [CrossRef]

Tang, Linghui, and Peter Koveos. 2008. A framework to update Hofstede's cultural value indices: Economic dynamics and institutional stability. *Journal of International Business Studies* 39: 1045–63. [CrossRef]

Titman, Sheridan. 1984. The effect of capital structure on a firm's liquidation decision. *Journal of Financial Economics* 13: 137–52. [CrossRef]

Tsakumis, George, Anthony Curatola, and Thomas Porcano. 2007. The relation between national cultural dimensions and tax evasion. *Journal of International Accounting, Auditing and Taxation* 16: 131–47. [CrossRef]

Vazza, Diane, Devi Aurora, and Ryan Schneck. 2005a. *Crossover Credit: A 24-Year Study of Fallen Angel Rating Behavior*. New York: Standard & Poor's Global Fixed Income Research.

Vazza, Diane, Edward Leung, Marya Alsati, and Mike Katz. 2005b. *Credit Watch and Ratings Outlooks: Valuable Predictors of Rating Behaviour*. New York: Standard & Poor's Global Fixed Income Research.

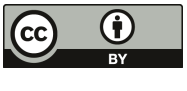
Vitell, Scott, Saviour Nwachukwu, and James Barnes. 1993. The effects of culture on ethical decisions making: An application of Hofstede's typology. *Journal of Business Ethics* 12: 753–60. [CrossRef]

Weaver, Gary R. 2001. Ethics Programs in Global Businesses: Culture's Role in Managing Ethics. *Journal of Business Ethics* 30: 3–15. [CrossRef]

Wei, Lee-Jen, Danyu Lin, and Lisa Weissfeld. 1989. Regression analysis of multivariate incomplete failure time data by modelling marginal distributions. *Journal of the American Statistical Association* 84: 1065–73. [CrossRef]

Zarzeski, Marilyn T. 1996. Spontaneous harmonization effects of culture and market forces on accounting disclosure practices. *Accounting Horizons* 10: 18–37.

Zheng, Xiaolan, Sadok Ghoul, Omrane Guedhami, and Chuck C. Y. Kwok. 2012. National culture and corporate debt maturity. *Journal of Banking and Finance* 36: 468–88. [CrossRef]



risks

MDPI

Article

Practice Oriented and Monte Carlo Based Estimation of the Value-at-Risk for Operational Risk Measurement

Francesca Greselin [1,†], Fabio Piacenza [2,*,†] and Ričardas Zitikis [3,†,‡]

[1] Department of Statistics and Quantitative Methods, Milano-Bicocca University, 20126 Milano, Italy; francesca.greselin@unimib.it
[2] Group Operational and Reputational Risks, UniCredit S.p.A., 20154 Milano, Italy
[3] School of Mathematical and Statistical Sciences, Western University, London, ON N6A 5B7, Canada; rzitikis@uwo.ca
* Correspondence: fabio.piacenza@unicredit.eu; Tel.: +39-02-886-24035
† These authors contributed equally to this work.
‡ Research of the third author has been supported by the Natural Sciences and Engineering Research Council of Canada under the title "From Data to Integrated Risk Management and Smart Living: Mathematical Modelling, Statistical Inference, and Decision Making," as well as by a Mitacs Accelerate Award from the national research organization Mathematics of Information Technology and Complex Systems, Canada, in partnership with Sun Life Financial, under the title "Risk Aggregation Beyond the Normal Limits."

Received: 22 March 2019; Accepted: 25 April 2019; Published: 1 May 2019

Abstract: We explore the Monte Carlo steps required to reduce the sampling error of the estimated 99.9% quantile within an acceptable threshold. Our research is of primary interest to practitioners working in the area of operational risk measurement, where the annual loss distribution cannot be analytically determined in advance. Usually, the frequency and the severity distributions should be adequately combined and elaborated with Monte Carlo methods, in order to estimate the loss distributions and risk measures. Naturally, financial analysts and regulators are interested in mitigating sampling errors, as prescribed in EU Regulation 2018/959. In particular, the sampling error of the 99.9% quantile is of paramount importance, along the lines of EU Regulation 575/2013. The Monte Carlo error for the operational risk measure is here assessed on the basis of the binomial distribution. Our approach is then applied to realistic simulated data, yielding a comparable precision of the estimate with a much lower computational effort, when compared to bootstrap, Monte Carlo repetition, and two other methods based on numerical optimization.

Keywords: advanced measurement approach; confidence interval; Monte Carlo; operational risk; value-at-risk

1. Introduction

International financial institutions typically calculate capital requirements for operational risk via the advanced measurement approach (AMA). The AMA is based on statistical models that are internally defined by institutions and comply with regulatory requirements (see European Parliament and Council of the European Union 2013). In particular, the regulations define which data sources must be used to measure operational risk:

- internal loss data;
- external loss data;
- scenario analysis; and
- business environmental and internal control factors.

Another significant requirement specifies that the capital charge has to be calculated at the 99.9% confidence level, with the holding period of one year. This means that the financial institution may experience an annual loss higher than the capital requirement once every 1000 years, on average.

The most-adopted implementation of AMA models is the loss distribution approach (see Frachot et al. 2001, 2007), where the objective is to estimate the probability distribution of the annual loss amount. In more detail, the loss data is categorized by Operational Risk Categories (ORCs), such that the independent and identical distribution (iid) hypothesis is applicable within each ORC. For each ORC, the severity distribution (i.e., the probability distribution of a single loss amount) and the frequency distribution (i.e., the probability distribution of the number of losses in one year) are estimated. The annual loss distribution for each ORC is obtained through a convolution process, and practically implemented with approximation or numerical techniques, since it is not usually feasible to represent the convolution function in closed form. For this aim, the main available approaches are:

- Monte Carlo method,
- Fourier transform-related methods,
- Panjer algorithm, and
- single loss approximation.

Naturally, these methodologies have their pros and cons, which have to be carefully considered when they are adopted. In particular:

- Monte Carlo method (see Owen 2013) is the most flexible and widely used technique, but it requires intensive computations and converges slowly to the correct result. In particular, the sampling error decreases proportionally to the square root of the number of steps.
- Fourier transform-related methods and the Panjer algorithm (see Klugman et al. 2012; Embrechts and Frei 2009) allow for a higher accuracy based on less intensive computations, but they require to carefully define the discretization of the severity distribution, to avoid underflow or overflow cases. The complexity of a valid discretization for the severity distribution has, usually, limited the usage of these techniques in the AMA models. In fact, while the slow convergence of the Monte Carlo method can affect the accuracy of the estimates, the underflow or overflow cases can potentially lead to totally unrealistic results.
- Single loss approximation (SLA) (see Böcker and Klüppelberg 2005; Böcker and Sprittulla 2006) allows approximating a high level quantile (corresponding to a probability close to 1) of the annual loss distribution on the basis of a particular quantile of the severity distribution and of the average of the frequency distribution. In particular, the SLA says that the 99.9% quantile of the annual loss distribution G can be approximated by the quantile of the severity distribution F at the level $1-(1-0.999)/\lambda$, where λ is the average of the frequency distribution, i.e., $G^{-1}(0.999) \approx F^{-1}(1-(1-0.999)/\lambda)$. Therefore, the shape of the frequency distribution does not significantly affect the 99.9% quantile of the loss distribution. This technique is very useful to obtain a quick proxy for the 99.9% quantile of the annual loss distribution but, being an asymptotic result, it is valid only for high level quantiles and cannot be used to define the entire annual loss distribution.

As suggested by the SLA, the purpose of AMA models is to estimate the severity distribution. This calculation phase of the AMA models is heavily affected by the "extrapolation issue" (see Mignola and Ugoccioni 2006; Cope et al. 2009). In fact, since the 99.9% quantile of the annual loss distribution approximately corresponds to the $1-(1-0.999)/\lambda$ quantile of the severity distribution, this means that $\lambda/(1-0.999) = \lambda/0.001 = 1000\lambda$ losses would be needed to estimate the 99.9% quantile of the annual loss distribution without resorting to extrapolation. This is because, in order to calculate the quantile of probability p (where $0 < p < 1$) without extrapolation, we need at least $1/(1-p)$ data. In recent years, banks have been collecting data, besides reporting current losses, and an effort has been made to build data bases going back in time. In any case, no more than ten to

fifteen years of loss data are available, and so the capital requirement needs a delicate extrapolation. To extrapolate beyond the data, therefore, parametric models for the severity distribution can be applied. In fact, non-parametric estimation does not provide any guidance outside of the range of observed data. Usually, parametric models are applied only above a sufficiently high loss threshold to fit the tail of the severity distribution; while the losses below the threshold (i.e., the body of the severity distribution) can be easily modeled using non-parametric methods, e.g., the empirical distribution. The lognormal, loglogistic, Weibull, Pareto and generalized Pareto are among the most common model choices (see Bolancel et al. 2012; Klugman et al. 2012). The parameters (e.g., (μ, σ) for the lognormal distribution) can be estimated through the maximum likelihood method, adopting truncated or shifted distributions (see Cavallo et al. 2012; Luo et al. 2007) to fit the models above a specified threshold.

Many other methods have been proposed to reduce the issues due to the extrapolation, among them are the following ones:

- Internal loss data are complemented with external loss data and a scenario analysis in the estimation of the severity distributions. Such an integration can be coherently performed using Bayesian methods (see Lambrigger et al. 2007; Shevchenko 2011; Dalla Valle and Giudici 2008; Figini et al. 2014; Shevchenko and Wüthrich 2006).
- Robust methods are employed to yield model parameters, like the weighted likelihood or the Optimal B-Robust estimation method (see Colombo et al. 2015; Danesi and UniCredit Business Integrated Solutions 2015; Danesi et al. 2016, respectively).

The AMA methods usually incorporate a fitting procedure to select the best model and threshold (i.e., tail parametric model) for each ORC (see Chernobai et al. 2005; Lavaud and Leherisse 2014; Panjer 2006). It has to be noted that a sound procedure to select the best tail model should include measures of both the goodness-of-fit vs loss data and the uncertainty of the estimated capital requirement (see Larsen 2016, 2019).

Whenever the shape of the frequency distribution does not significantly affect the capital requirement, as it has been discussed in the SLA approach, then the frequency distribution is estimated through the simplest possible model, which is the Poisson distribution. This discrete probability distribution has only one parameter λ, representing both the average and the variance of the number of loss data in one year. Once, for each ORC, the severity and the frequency distributions have been estimated, the annual loss distribution can be approximated through one of the above mentioned techniques, e.g., the Monte Carlo method. Starting with the annual loss distributions of the ORCs as margins, the overall annual loss distribution is usually obtained by applying a copula function. In fact, on the basis of Sklar's theorem, if F is an n-dimensional cdf with continuous margins $F_1, \ldots, F_n$, then it has the unique copula representation $F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n))$, where C is the copula function, i.e., a multivariate cdf with margins uniformly distributed on [0,1], with the following properties:

- $C : [0,1]^n \to [0,1]$;
- C is grounded, i.e., $C_i(u) = C(u_1, \ldots, u_{i-1}, 0, u_{i+1}, \ldots, u_n) = 0$ for all $i = 1, \ldots, n$, and n-increasing;
- C has margins $C_i (i = 1, \ldots, n)$ satisfying $C_i(u) = C(1, \ldots, 1, u, 1, \ldots, 1) = u$ for all $u \in [0,1]$.

In practice, we see that for continuous multivariate distribution functions, the univariate margins and the multivariate dependence structure can be separated. The dependence structure can be represented by a proper copula function, e.g., the Student t-copula (see Di Clemente and Romano 2004). The capital requirement is then, usually, defined through the Value-at-Risk (VaR) measure, as the 99.9% quantile of the overall annual loss distribution. Deductions for insurance (see Bazzarello et al. 2006) and expected losses are often considered in AMA models. For expected loss deduction, the European Parliament and Council of the European Union (2013) states that a financial institution shall calculate its capital requirement for operational risk as comprising both the expected loss and the unexpected loss, unless the expected loss is adequately captured in its internal business practices. We can see the

99.9% quantile of the overall annual loss distribution as the sum between the expected loss and the unexpected loss, from which the expected loss could be eventually deducted. To obtain this deduction, it is necessary to comply with the requirements reported in European Parliament and Council of the European Union (2018):

- the expected loss estimation process is done by ORC and is consistent over time;
- the expected loss is calculated using statistics that are less influenced by extreme losses, including median and trimmed mean, especially in the case of medium- or heavy-tailed data;
- the maximum offset for expected loss applied by the institution is bounded by the total expected loss;
- the maximum offset for expected loss in each ORC is bound by the relevant expected losses calculated according to the institution's operational risk measurement system applied to that category;
- the offsets the institution allows for expected loss in each ORC are capital substitutes or available to cover expected loss with a high degree of certainty (e.g., provisions) over the one-year period;
- specific reserves for exceptional operational risk loss events, that have already occurred, cannot be used as expected loss offsets.

In case the annual loss distributions for the ORCs and the related overall annual loss distribution are obtained through the Monte Carlo method, such distributions have to be represented by a sufficiently high number of simulated data. Following the terminology adopted by the EU regulators, the number of Monte Carlo steps refers to the sample size. For example, with 5 million data points, the 99.9% quantile can be estimated via the empirical estimator as the value at position 4,995,000 in the increasing ordered sequence of the simulated losses. It provides the capital requirement for operational risk.

2. The Monte Carlo Method: An Introduction

The Monte Carlo method has a long history in statistics, where it was also called 'model sampling'. It was used to verify the properties of estimates by reproducing the settings for which they were designed. W. S. Gosset, writing as Student (1908), before deriving analytic result on Student's t distribution, did some simulations, using height and left middle finger measurements from 3000 criminals as written on pieces of cardboard. Additional history on the Monte Carlo method can be found in Kalos and Whitlock (2008) including computations made by Fermi in the 1930s. This method became more important in the 1940s and early 1950s, when it was also used to solve problems in physics, related to atomic weapons. The name itself was assigned in this period, taken from the famous casino located in Monte Carlo.

There are several papers in which the Monte Carlo method is widely used for new problems. For example, Metropolis et al. (1953) presented the Metropolis algorithm, which was the first Markov chain Monte Carlo method. Boyle (1977) showed how to apply Monte Carlo methods to financial options pricing. Efron (1979) defined the bootstrap, using Monte Carlo method to derive statistical properties without specifying distributional assumptions.

Approximately during the same period, the Quasi-Monte Carlo techniques were defined. The term itself was assigned by Richtmyer (1952). He thought to improve Monte Carlo by using sequences with better uniformity than truly random sequences would have.

The uniform distribution on $[0,1]$ is fundamental in the theory of Monte Carlo method. In fact the simulation methods derive their randomness from sequences of independent, uniformly distributed, random variables on $[0,1]$. A continuous random variable U is uniformly distributed on $[0,1]$ if it has the following probability density function

$$f_U(u) = \begin{cases} 1, & \text{if } u \in [0,1]; \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

In order to generate a realization from a non-uniform distribution, random numbers uniformly distributed on $[0,1]$ can be transformed in observations extracted from the desired distribution. Given a

random variable X, its cumulative distribution function $F_X(x)$ is an increasing function on $[0,1]$. If $F_X(x)$ is continuous and strictly increasing (i.e., $F_X(x_1) < F_X(x_2)$ if $x_1 < x_2$), then it assumes all the values between 0 and 1. In this case, the inverse function $F_X^{-1}(y)$ (i.e., the quantile function) can be defined, such that $y = F_X(x) \iff x = F_X^{-1}(y)$. More generally, the quantile function can be defined as $F_X^{-1}(y) = \inf\{x : F_X(x) \geq y\}$. Hence, to obtain a random number c from a random variable X, one can generate a number r from the uniform distribution on $[0,1]$ and calculate $c = F_X^{-1}(r)$.

Actually, no sequence generated through an algorithm, and hence through a computer, can be really considered 'random'. In fact, we can only generate sequences of pseudo-random numbers, i.e., numbers which are apparently random. A pseudo-random sequence is built by a sequence of values x_j, for which the first k values $x_0, x_1, \ldots, x_{k-1}$ are chosen, and each of the next values x_j, $j \geq k$ is obtained by applying a function on the k previous values $x_{j-1}, x_{j-2}, \ldots, x_{j-k}$. The sequence of the first k values $x_0, x_1, \ldots, x_{k-1}$ is named 'seed'. Therefore, the entire sequence of generated numbers depends on the value of the seed. Generally, the generation of random sequences is based on a linear congruential generator defined as follows

$$x_{j+1} = (a \cdot x_j + b) \mod m \tag{2}$$

where a is the multiplier, b is the shift, and m is the module. The best of these quickly produce very long streams of numbers, and have fast and portable implementations in many programming languages. Among these high quality generators, the Mersenne twister, MT19937, of Matsumoto and Nishimura (1998) has become the most prominent, though it is not the only high quality random number generator. We also cite the 'combined multiple-recursive generator' from L'Ecuyer (1999).

In a simple Monte Carlo problem (see Owen 2013) we define the quantity to be estimated as the expected value of a random variable Y , i.e., $\mu = \mathbb{E}(Y)$. Please note that the estimation of average is the most basic application of Monte Carlo method, which is generally used to introduce the approach. The estimation of quantile, which is of main interest in operational risk measurement, will be discussed from the following sections. To begin with, we generate values $Y_1, \ldots, Y_n$ independently and randomly from the distribution of Y and take their average

$$\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^{n} Y_i \tag{3}$$

as the estimate of μ. Usually, we have $Y = f(X)$ where the random variable $X \in D \subset \mathbb{R}^d$ has a probability density function $p(x)$, and f is a real-valued function defined over D. Then $\mu = \int_D f(x)p(x)dx$.

The simple Monte Carlo is justified through the laws of large numbers, assuring that under simple conditions (μ has to exist), the sample-average estimator converges to the actual value. This tells us that Monte Carlo will eventually produce an error as small as we like, but it does not tell us how large n has to be for this to happen. Supposing that Y has a finite variance ($\sigma^2 < \infty$), $\hat{\mu}_n$ is a random variable having mean

$$\mathbb{E}(\hat{\mu}_n) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(Y_i) = \mu, \tag{4}$$

and variance

$$\mathbb{E}((\hat{\mu}_n - \mu)^2) = \frac{\sigma^2}{n}. \tag{5}$$

Therefore, the standard deviation of $\hat{\mu}_n$ is equal to $\sqrt{\mathbb{E}((\hat{\mu}_n - \mu)^2)} = \frac{\sigma}{\sqrt{n}}$. This means that to get one more decimal digit of accuracy is equivalent to asking for a standard deviation one tenth as large, and that requires a 100-fold increase in computation. The value of σ^2 can be estimated by

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \hat{\mu}_n)^2. \tag{6}$$

The variance estimate s^2 tells us that the standard error is of the order of $s/\sqrt{n}$. From the central limit theorem (CLT), we also know that $\hat{\mu}_n - \mu$ has approximately a normal distribution with mean 0 and variance s^2/n. This yields the familiar 95% confidence interval

$$\hat{\mu}_n - 1.96\frac{s}{\sqrt{n}} \leq \mu \leq \hat{\mu}_n - 1.96\frac{s}{\sqrt{n}}. \tag{7}$$

Since the Monte Carlo method typically has an error variance of the form $\frac{\sigma^2}{n}$, we get a more accurate estimate by sampling with a larger value of n. However, the computing time grows with n. Therefore several techniques have been defined to reduce σ (see Owen 2013; Kleijnen et al. 2010). To do this, new Monte Carlo methods can be constructed, having the same answer as the simple one but with a lower σ. These methods are known as variance reduction techniques:

- Antithetic sampling: let $\mu = \mathbb{E}(X)$ for $X \sim p$, where p is a symmetric density on the symmetric set D. Here, symmetry is with respect to reflection through the center point of D. If we reflect $x \in D$ through c we get the point $\tilde{x}$ with $\tilde{x} - c = -(x - c)$, that is $\tilde{x} = 2c - x$. Symmetry means that $p(\tilde{x}) = p(x)$ including the constraint that $x \in D$ if and only if $\tilde{x} \in D$. The antithetic sampling estimate of μ is

  $$\hat{\mu}_{anti} = \frac{1}{n}\sum_{i=1}^{n/2}(f(X_i) + \tilde{f}(X_i)), \tag{8}$$

  where $X_i \sim p$, and n is an even number. It can be proved that the variance in antithetic sampling is

  $$Var(\hat{\mu}_{anti}) = \frac{\sigma^2}{n}(1 + \rho), \tag{9}$$

  where $\rho = Corr(f(X); \tilde{f}(X))$. From $-1 \leq \rho \leq 1$ we obtain $0 \leq \sigma^2(1+\rho) \leq 2\sigma^2$. In the best case, antithetic sampling gives the exact answer from just one pair of function evaluations. In the worst case it doubles the variance. Both cases do arise. It is clear that a negative correlation is favorable.
- Stratification: the idea in stratified sampling is to split up the domain D of X into separate regions, take a sample of points from each such region, and combine the results to estimate $\mathbb{E}(f(X))$. Intuitively, if each region gets its fair share of points then we should get a better answer. One might be able to do better still by oversampling within the important strata and undersampling those in which f is nearly constant.
- Common Random Numbers (CRN): this technique applies when we are comparing two or more alternative configurations (of a system) instead of investigating a single configuration. CRN requires synchronization of the random number streams, which ensures that in addition to using the same random numbers to simulate all configurations, a specific random number used for a specific purpose in one configuration is used for exactly the same purpose in all other configurations.
- Control variate: control variates provide ways to exploit closed form results. With control variates we use some other problems, quite similar to our given one, but for which an exact answer is known. The precise meaning of 'similar' depends on how we use this other problem, and more than one method is given below. As for 'exact', we mean it literally, but in practice it may just mean known with an error negligible when compared to Monte Carlo errors.
- Importance sampling: the idea behind importance sampling is that certain values of the input random variables in a simulation have more impact on the parameters being estimated than

others. If these 'important' values are sampled more frequently, then the variance can be reduced. The basic methodology in importance sampling is to choose a distribution which 'encourages' the important values. This use of 'biased' distributions will result in a biased estimator, if it is applied directly in the simulation. However, if the simulation outputs are weighted to correct for the use of the biased distribution, this will ensure that the new importance sampling estimator is unbiased. The weight is given by the likelihood ratio, that is, the Radon–Nikodym derivative of the true underlying distribution with respect to the biased simulation distribution. The fundamental issue in implementing importance sampling is the choice of a biased distribution which encourages the important regions of the input variables. Choosing a good biased distribution is the core part of importance sampling. The rewards for a good distribution can be huge run-time savings; the penalty for a bad distribution can be longer run times than for a simple Monte Carlo simulation without importance sampling.

We also mention the Quasi-Monte Carlo methods, which use low-discrepancy sequences (also called quasi-random sequences or sub-random sequences). This is in contrast to the simple Monte Carlo method, which is based on sequences of pseudorandom numbers. The advantage of using low-discrepancy sequences is a faster rate of convergence. Quasi-Monte Carlo has a rate of convergence of order $1/n$, whereas—as already mentioned above—the rate for the simple Monte Carlo method is $1/\sqrt{n}$.

In this paper we treat and apply the simple Monte Carlo method, which is also known as crude Monte Carlo to distinguish it from more sophisticated methods. In fact, all of the more sophisticated methods offer several promising technical venues for implementing the capital requirement estimation methods that we discuss. However, we have chosen the simple method because we have found it to be the easiest to be applied by practitioners working in the industry, where simplicity is often preferred to more complex methods. Moreover, calculations with simple Monte Carlo method can be easily distributed to more CPUs, allowing for a parallel computation, in order to make them less time consuming. In fact, most of adopted softwares in statistics offer built-in routines for parallel computations (e.g., parallel package in R by R Core Team (2016)), allowing reducing the calculation time of simple Monte Carlo proportionally to the number of used CPUs.

3. The Monte Carlo Method: Basic Definitions for Operational Risk Measurement

As mentioned in Section 1, Monte Carlo simulations are widely employed in the operational risk measurement. A detailed description of the Monte Carlo method applied to the calculation of annual loss distribution and capital requirement for operational risk can be found in Guharay et al. (2016). Here we will recall briefly the main idea behind it.

Let S be the random variable defined as

$$S = \sum_{i=1}^{N} X_i \tag{10}$$

describing the annual loss distribution, where X_i is the random variable representing the severity of the i-th operational risk loss event, and N is the random variable representing the number of single operational risk loss events, for each year. On the basis of the above notation, the Monte Carlo method can be implemented by executing the following sequence of instructions at the generic step j:

- from the frequency distribution, draw a value n_j representing the number of loss events for year j;
- from the severity distribution, draw the amounts $x_{1j}, \ldots, x_{n_j j}$ of loss events;
- calculate the annual loss amount for year j, that is $s_j = \sum_{i=1}^{n_j} x_{ij}$;

and iterate the above three steps for a sufficiently large number J of steps. The capital requirement can be evaluated as the empirical 99.9% quantile of the simulated sample of the annual losses $s_1, s_2, \ldots, s_J$.

In practice, after ordering the sequence $\{s_j\}_{j=1,...,J}$ to obtain $s_{1:J} \leq s_{2:J} \leq \cdots \leq s_{J:J}$, the capital requirement corresponds to the annual loss $s_{j:J}$ with $j = 99.9\%J$.

It is well known (as already mentioned in Section 2) that the sampling accuracy of the capital requirement, calculated through the Monte Carlo method, increases proportionally to the square root of the number of steps (see Owen 2013). In other words, increasing the number of steps by 100 times, the random sampling error is expected to decrease by 10 times. Therefore, when the annual loss distribution and the related 99.9% quantile are approximated by the Monte Carlo method, the good practice would require to consider and, possibly, measure the consequent sampling error. Moreover, the recently published EU Regulatory Technical Standards (European Parliament and Council of the European Union 2018) require that irrespective of the techniques adopted to obtain the annual loss distribution and the related 99.9% quantile, the institution adopts criteria that mitigate sampling and numerical errors and provides a measure of their magnitude. With reference to the Monte Carlo method, the Regulatory Technical Standards 2018/959 require that:

- the number of steps to be performed is consistent with the shape of the distributions and with the confidence level to be achieved;
- where the distribution of losses is heavy-tailed and measured at a high confidence level, the number of steps is sufficiently large to reduce sampling variability below an acceptable threshold.

To comply with these requirements, the sampling variability, i.e., the random sampling error, has to be assessed. Having estimated such an error, given the acceptable threshold (reasonably defined in terms of a low percentage of the 99.9% quantile), it is possible to verify whether the number of steps is sufficiently large and, consequently, consistent with the shape of the estimated distribution. It is worth mentioning that the topic of the sampling error related to the Monte Carlo method is also considered by Mignola and Ugoccioni (2006), based on the previous work of Chen and Kelton (2001).

There is a discussion in the recent literature on extreme quantile estimation and the adoption of some types of Monte Carlo techniques, as in Boos (1984); Danielsson and De Vries (1996); Peng and Qi (2006); Gomes and Pestana (2007). The proposed methodologies are mainly based on the Extreme Value Theory (EVT). However, in the operational risk measurement context, the EVT is no longer intensively used by practitioners, despite its interesting theoretical properties. In fact, when the EVT is applied to actual operational risk data, several practical problems arise, due to the failure of achieving the asymptotic regime described by the theory. Additionally, the capital requirement measures based on EVT are affected by the high volatility due to the use of Pareto-type tails (see Mignola and Ugoccioni 2005; Makarov 2006). In any case, EVT methods can be conjugated with the present approach, but such development is beyond the scope and purpose of this paper—it can be an interesting topic for further studies.

4. The Monte Carlo Error in Operational Risk Measurement

Suppose that using Monte Carlo techniques applied to historical data, we have obtained a set of n losses, which are positive real numbers, assuming that we do not take into account null values for losses. Hence, we are dealing with outcomes of n positive i.i.d. random variables (rv's)

$$X_1, X_2, ..., X_n \tag{11}$$

with cdf $F(x)$ and quantile function $F^{-1}(p)$. The focus of actuaries is on large losses, and thus on large order statistics of the Xs. In this framework, percentiles are rich informative measures that can be represented using the order statistics $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ of rv's (11). From the mathematical point of view, by ordering rv's we obtain a monotone sequence and thus extract additional information from the original rv's (11). In a general setting, we want to obtain a confidence interval for $VaR(p) = F^{-1}(p)$, for some p in (0,1), but the focus here is to measure the error when estimating a very high (an extreme) quantile $F^{-1}(p)$, say for $p = 0.999$ for the operational risk measurement. Once we have a CI, we can define:

- the Monte Carlo error (E) as the width of the CI, and
- the relative Monte Carlo error (RE) as the ratio between the width of the CI and the best available estimate of the 99.9% quantile itself.

To construct a CI, we consider (and compare) the following methods:

- the method based on the binomial distribution applied on the quantile, as proposed by Owen (2013);
- the bootstrap (or re-sampling) method (i.e., the classical n-bootstrap and the m-out-of-n-bootstrap), where we get the probability distribution of the 99.9% quantile by drawing random samples with replacement from the annual loss distribution generated by the Monte Carlo method;
- the Monte Carlo repetition method, where we simulate the probability distribution of the 99.9% quantile, repeating several times the Monte Carlo method;
- two methods based on constrained optimization.

Since we suppose that the severity and the frequency distributions are already known or, in any case, already estimated, some typical operational risk distribution functions and parameterizations can be considered, in order to run the above mentioned methods and to analyze results. In particular, the *lognormal* pdf

$$f_X(x) = \frac{1}{x\,\sigma\,\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\,\sigma^2}\right),$$

for $x > 0$, with $-\infty < \mu < +\infty$ and $\sigma > 0$, can be used to model the severity (with , e.g., parameters $\mu = 9$ and $\sigma = 2$), and the *Poisson* distribution can be used to model the frequency (with, e.g., parameter $\lambda = 100$, i.e., assuming 100 losses per year, on average). Such choices are motivated by previous works in the literature; see, for example, Frachot et al. (2001, 2007); Mignola and Ugoccioni (2005); Böcker and Sprittulla (2006).

Naturally, other models can be adopted for the severity distribution. We may for example consider, as a valuable alternative, the three-and four-parameter lognormal distributions, or gamma-type, or beta-type, or Pareto-type size distributions (see Kleiber and Kotz (2003) for a detailed review). Recent papers on thick-tailed distributions have adopted a semi-nonparametric approach (the so-called log-SNP, see, e.g., Cortés et al. (2017)) that generalizes the lognormal incorporating not only thick tails but also non-monotonic decay, multimodality and many other salient empirical features of the distribution of losses. This approach presents a valuable avenue for future research on operational risk.

For simplicity, we suppose that the overall annual loss distribution is generated on the basis of a single ORC, without affecting the generality of our results. Therefore, applying the Monte Carlo method, we simulate a sample of $n = 5$ million annual losses. We obtain the estimated 99.9% quantile as $VaR(0.999) = 47.8037 \times 10^6$, and we are now ready to evaluate the different CIs for the latter.

4.1. The Binomial Distribution Method

To begin with, we employ the approach based on the binomial distribution (see Owen 2013). The idea behind this method is that the number of simulated annual losses below the 99.9% quantile follows the binomial distribution $B_{n,\theta}$, where n is the number of simulated annual losses and $\theta = 0.999$.

Let S have a continuous distribution. Then $\mathbb{P}(S \le Q^\theta) = \theta$ holds and so $Q^\theta = \eta$ if and only if $\mathbb{E}(\mathbb{1}_{S\le\eta}) = \theta$. If θ is not included in the confidence interval for $\mathbb{E}(\mathbb{1}_{S\le\eta})$, then we can reject $Q^\theta = \eta$. As a result, we can obtain confidence intervals for Q^θ using confidence intervals for a binomial proportion. As a candidate value η increases, the number of S_i below it changes only when η crosses one of the order statistics $s_{i:n}$ in the sample. As a result, the confidence interval takes the form $[s_{L:n}, s_{R:n}]$ for integers L and R. Then

$$\mathbb{P}(s_{L:n} \le Q^\theta \le s_{R:n}) = \mathbb{P}(L \le W \le R)$$

where $W = \sum_{i=1}^{n} \mathbb{1}_{S_i \le Q^\theta} \sim B_{n,\theta}$.

This method allows defining a confidence interval for the 99.9% quantile given, e.g., the following information:

- ordered sample of simulated annual losses $s_{1:n} \leq s_{2:n} \leq \cdots \leq s_{n:n}$,
- confidence level $1-\alpha$,
- quantile level $\theta = 99.9\%$,
- cdf $B_{n,\theta}$ of the $Binomial(n;\theta)$ distribution,

by adopting the following steps:

- $L = B_{n,\theta}^{-1}(\alpha/2)$,
- if $B_{n,\theta}(L) \leq \alpha/2$, then $L = L+1$,
- $R = B_{n,\theta}^{-1}(1-\alpha/2)$,
- the confidence interval is defined as $[s_{L:n}, s_{R:n}]$.

Please note that the cases $L = 0$ and $R = n+1$ can occur when n is small. In these cases, the confidence interval can be defined by considering $s_{0:n} = 0$ (since the losses are non-negative by definition) and $s_{n+1:n} = +\infty$, respectively. For numerical reasons, the test on L is not $B_{n,\theta}(L) = \alpha/2$, but it is $B_{n,\theta}(L) \leq \alpha/2$. In fact, we could have $B_{n,\theta}(L) < \alpha/2$, even if it is actually $B_{n,\theta}(L) = \alpha/2$, just because of some rounding in the implementation of the method.

Applying the above described algorithm to 5 million simulated annual losses from the *lognormal* $(\mu = 9, \sigma = 2)$ severity and *Poisson* $(\lambda = 100)$ frequency, we obtain the following CI at the confidence level $1-\alpha = 0.99$:

$$[L^{Bin}; R^{Bin}] = [47.3667 \times 10^6; 48.2897 \times 10^6]. \tag{12}$$

We get the error

$$E^{Bin} = R^{Bin} - L^{Bin} = 0.9231 \times 10^6 \tag{13}$$

and the relative error

$$RE^{Bin} = (R^{Bin} - L^{Bin})/VaR(0.999) = 0.019310 = 1.9310\%. \tag{14}$$

4.2. The n-Bootstrap Method

In this section we compare our previous results with those that can be obtained using the bootstrap method. The bootstrap is a resampling technique that allows estimating the standard error of an estimator. It can be used to estimate the variance of the estimator, its bias, construct confidence intervals, and test hypotheses by calculating p-values. The method was introduced by Efron (1979) and has been widely used, since it is easy to understand and implement. A very useful version of the method is the non-parametric bootstrap, which does not require any prior knowledge of the underlying distribution.

Consider a statistic T_n calculated based on a sample of size n from a distribution F. Let F_n be an estimator of F. The bootstrap method is based on generating a large number B of samples from F_n, i.e., drawing n elements with replacement from the original sample of size n, obtaining $X_1, \ldots, X_n$. In the case of non-parametric bootstrap, the empirical estimator $F_n(x) = \frac{1}{n}\sum_i^n \mathbb{1}_{[0,x]}X_i$ is the natural choice for estimating $F(x)$. The B samples are then used to estimate the sampling distribution of T_n, i.e., from each sample, a realization of T_n is calculated until we obtain realizations $t_1, \ldots, t_B$, which are used to simulate the sampling distribution of T_n. If the estimator F_n is "close" enough to F, then the bootstrap (hereafter n-bootstrap) method works. Other resampling techniques are the jackknife and cross-validation (see Efron 1979, 1982).

Therefore, given the empirical estimator of the distribution of annual losses, we simulate the distribution of the quantile by drawing $B = 100$ random samples with replacement from the 5 million of simulated annual losses. We get the 99.9% quantile estimate from the sample average of the estimated 99.9% quantiles calculated on the 100 random samples:

$$VaR^{n-boot}(0.999) = 47.7977 \times 10^6. \tag{15}$$

We employ the empirical 0.05 and 0.95 quantiles of the $B = 100$ realizations of the 99.9% quantiles obtained from the bootstrapped samples, in order to obtain the corresponding CI

$$[L^{n\text{-}boot}; R^{n\text{-}boot}] = [47.2914 \times 10^6; 48.2306 \times 10^6]. \tag{16}$$

As before, we evaluate the error

$$E^{n\text{-}boot} = (R^{n\text{-}boot} - L^{n\text{-}boot}) = 0.9392 \times 10^6 \tag{17}$$

and the relative error

$$RE^{n\text{-}boot} = (R^{n\text{-}boot} - L^{n\text{-}boot})/VaR^{n\text{-}boot}(0.999) = 0.019649 = 1.9649\%. \tag{18}$$

Please note that this method provides very similar results to those obtained earlier using the binomial method.

4.3. The m-out-of-n-Bootstrap Method

There are cases where the bootstrap method, in its most classical version, fails. Guidelines to identify cases where we should have problems are reported by Bickel et al. (1997). However, it is not trivial to a priori identify cases where the bootstrap method could not work. When the classical bootstrap fails, several authors suggest to consider, instead, samples of size $m = o(n)$ (Bickel et al. 1997; Sakov 1998; Bickel and Sakov 2008; Gribkova and Helmers 2011). However, some issues arise:

- one does not know, a priori, whether or not the bootstrap works in each particular case;
- the choice of m can be crucial, in case the classical bootstrap fails.

Assume that $X_1, \ldots, X_n$ are i.i.d. random variables with distribution F. Let $T_n = T_n(X_1, \ldots, X_n; F)$ be an estimator with cdf $L_n(x) = \mathbb{P}(T_n \leq x)$. We assume a known rate of convergence of T_n to a non-degenerate limiting distribution L, i.e., $L_n \Rightarrow L$. We aim at estimating a confidence interval for $\theta = \gamma(L)$, where γ is a functional.

The bootstrap method can be used to estimate θ, which in turn is used to estimate L, applying the functional γ. For any positive integer m, let $X_1^*, \ldots, X_m^*$ be a bootstrap sample drawn from the empirical cdf F_n, and denote the m-bootstrap version of T_n by $T_m^* = T_m^*(X_1^*, \ldots, X_m^*; F_n)$, with the bootstrap distribution $L_{m,n}^*(x) = \mathbb{P}^*(T_m^* \leq x) = \mathbb{P}(T_m^* \leq x | F_n)$.

We say that the bootstrap "works" if $L_{m,n}^*$ converges weakly to L in probability for all $m, n \to \infty$ and, in particular, for $m = n$. When the n-bootstrap does not "work," using a smaller bootstrap sample size rectifies the problem. It has been proved that under minimal conditions, although $L_{n,n}^*$ does not have the correct limiting distribution, $L_{m,n}^*$ with "small" but not too small m does work (Bickel et al. 1997, Bickel and Sakov 2008). Please note that for any $m < n$, bootstrap samples may be drawn with or without replacement, and we adopt the second choice as it allows for the case of $m = n$.

Using a smaller bootstrap sample requires a choice of m, and the method is known as m-out-of-n-bootstrap. When m is in the "right range" of values, the bootstrap distributions for different possible samples are "close" to each other (Bickel and Sakov 2008). When m is "too large" or fixed, the bootstrap distributions can be different. This suggests looking at a sequence of values of m and their corresponding bootstrap distributions. A measure of discrepancy between these distributions should have large values when m is "too large" or fixed. The discrepancies should be small when m is of the "right order".

The failure of the n-bootstrap is, usually, of the following type: $L_{n,n}^*$ when viewed as a probability distribution on the space of all probability distributions does not converge to a point mass at the correct limit L, but rather converges to a non-degenerate distribution, denote it by L_1 on that space. If $m \to \infty$, $m/n \to \lambda$, $0 < \lambda < 1$, one gets convergence to a non-degenerate distribution, say L_λ, which is typically

different from L_1. We would expect that $L_0 = L$. This behavior suggests the following adaptive rule for choosing m, which is described by Bickel and Sakov (2008).

1. Consider a sequence of values for m, defined as follows:

$$m_j = [q^j n], \text{ for } j = 0, 1, 2, \ldots, \quad 0 < q < 1, \tag{19}$$

   where $[\alpha]$ denotes the smallest integer $\geq \alpha$.
2. For each m_j , find $L^*_{m_j,n}$ using the m-out-of-n-bootstrap.
3. Let ρ be some metric consistent with convergence in law, and set

$$\hat{m} = \underset{m_j}{\operatorname{argmin}}\, \rho(L^*_{m_j,n}, L^*_{m_{j+1},n}). \tag{20}$$

   If the difference is minimized for a few values of m_j , then pick the largest among them. Denote the j corresponding to $\hat{m}$ by $\hat{j}$.
4. The estimator of L is $\hat{L} = L^*_{m_{\hat{j}},n}$. Estimate θ by $\hat{\theta}_n = \gamma(\hat{L})$ and use the quantiles of $\hat{L}$ to construct a confidence interval for θ.

Bickel and Sakov (2008) propose to use the Kolmogorov-Smirnov distance , i.e., $\rho(F, G) = \sup_x |F(x) - G(x)|$, for measuring the distance. The adaptive rule described above has been applied on our example to choose m (where n is equal to 5 millions), with parameter $q = 0.8$ to define m_j. Figure 1 shows the Kolmogorov-Smirnov distances, calculated for each m_j over the grid of values in 1–5 millions as in (19) with $q = 0.8$, to select the value of $\hat{m}$ which minimizes the discrepancies among bootstrap distributions. The Kolmogorov-Smirnov distance has been minimized for $\hat{m}$ = 1,310,721.

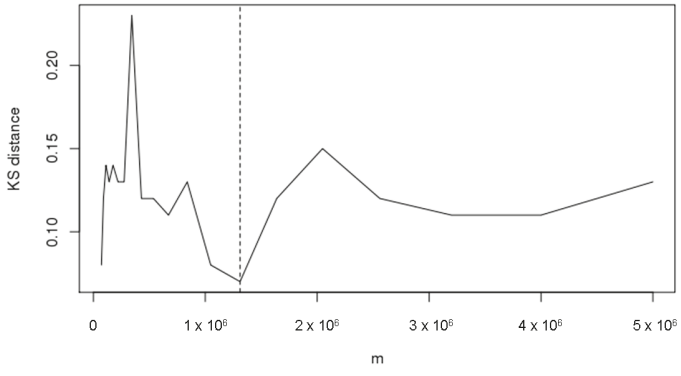


Figure 1. Choice of the optimal m in terms of KS distance from bootstrap distributions (n = 5,000,000).

We get the 99.9% quantile estimate

$$VaR^{m\text{-}boot}(0.999) = 47.8173 \times 10^6 \tag{21}$$

from the sample average of the $B = 100$ estimated 99.9% quantiles calculated using the Monte Carlo samples of size $\hat{m}$ = 1,310,721. We employ the empirical 0.05 and 0.95 quantiles of the $B = 100$ realizations of the 99.9% quantiles calculated using the Monte Carlo samples to obtain the CI

$$[L^{m\text{-}boot}; R^{m\text{-}boot}] = [46.9388 \times 10^6; 48.9690 \times 10^6]. \tag{22}$$

As before, we evaluate the error

$$E^{m\text{-}boot} = (R^{m\text{-}boot} - L^{m\text{-}boot}) = 2.0302 \times 10^6 \tag{23}$$

and the relative error

$$RE^{m\text{-}boot} = (R^{m\text{-}boot} - L^{m\text{-}boot})/VaR^{m\text{-}boot}(0.999) = 0.042457 = 4.2457\%. \quad (24)$$

We note that by applying the m-out-of-n-bootstrap with $\hat{m}$ = 1,310,721, the error and the relative error are slightly higher than the double of the same values calculated using the n-bootstrap. We can observe that this result is consistent with the fact that the Monte Carlo accuracy increases (decreases) proportionally to the square root of the number of steps: since the ratio between n = 5,000,000 and $\hat{m}$ = 1,310,721 is equal to around 3.8, we can expect that the relative error increases by approximately $\sqrt{3.8} = 1.95$, which is actually close to the observed result. To see the impact of the selection of $\hat{m}$, Figures 2 and 3 depict the errors and the relative errors of the CIs calculated on the grid of values of m as in (19).

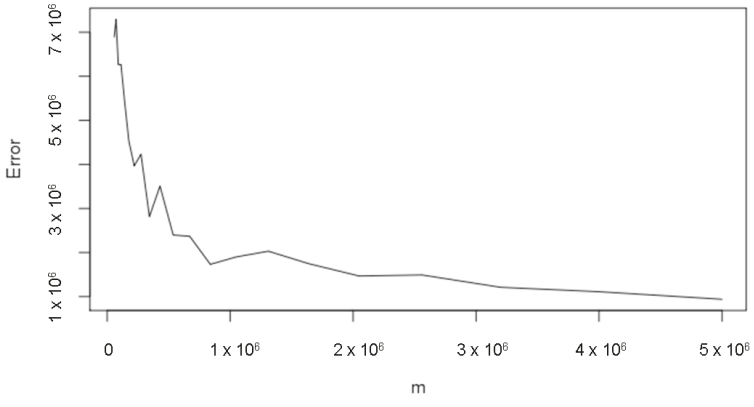


Figure 2. Error for different m values.

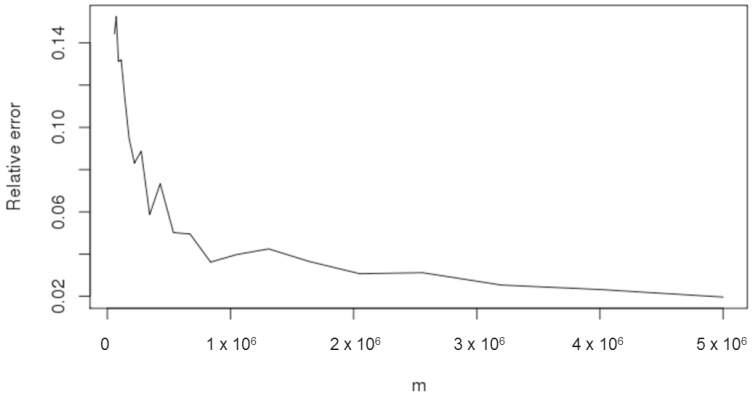


Figure 3. Relative error for different m values.

We observe that the errors decrease as the value of m increases, which is consistent with the aforementioned square root rule for the Monte Carlo accuracy.

4.4. The Monte Carlo Repetition Method

In this section, we compare the previous results with those we shall next obtain using the Monte Carlo repetition method. We first note that repeating the Monte Carlo method several times,

without fixing the initial seed of each repetition, provides a realistic probability distribution of the possible results that we can obtain using this numerical procedure. Please note that it is fundamental to initialize the seed only for the first repetition since, as reminded in Section 2, the seed defines the entire sequence of the generated random numbers. In fact, if we initialize the same seed for each Monte Carlo repetition, we will alway obtain the same results leading to a degenerate probability distribution. Although the Monte Carlo repetition method can be intuitively considered as the most accurate to estimate the Monte Carlo error (if the number of repetitions is sufficient), it cannot be extensively used because it is the most computationally intensive. Therefore, to compare the Monte Carlo repetition method with previous results, we simulate the distribution of the quantile by repeating 100 times the Monte Carlo method. We get the 99.9% quantile estimate

$$VaR^{MC100}(0.999) = 47.4246 \times 10^6 \tag{25}$$

from the sample average of the 100 estimated 99.9% quantiles, calculated on the 100 Monte Carlo samples. We employ the empirical 0.05 and 0.95 quantiles of the 99.9% quantiles calculated from the 100 Monte Carlo samples to obtain the CI

$$[L^{MC100}; R^{MC100}] = [47.0261 \times 10^6; 47.8522 \times 10^6]. \tag{26}$$

As before, we evaluate the error

$$E^{MC100} = (R^{MC100} - L^{MC100}) = 0.8261 \times 10^6 \tag{27}$$

and the relative error

$$RE^{MC100} = (R^{MC100} - L^{MC100})/VaR^{MC100}(0.999) = 0.017420 = 1.7420\%. \tag{28}$$

We note that the current method provides slightly lower errors than the previous two methods.

To see how the number of Monte Carlo repetitions affects the results, we have reapplied the current method by repeating 1000 times the Monte Carlo method. We obtain the following results:

$$VaR^{MC1000}(0.999) = 47.4345 \times 10^6; \tag{29}$$

$$[L^{MC1000}; R^{MC1000}] = [47.0207 \times 10^6; 47.8503 \times 10^6]; \tag{30}$$

$$E^{MC1000} = (R^{MC1000} - L^{MC1000}) = 0.8297 \times 10^6; \tag{31}$$

$$RE^{MC1000} = (R^{MC1000} - L^{MC1000})/VaR^{MC1000}(0.999) = 0.017491 = 1.7491\%. \tag{32}$$

Please note that even though we have increased the number of repetitions from 100 to 1000, the error results basically remain the same.

To obtain a further confirmation of the sufficiency of 1000 repetitions, we check the normality of the distribution of the 99.9% quantile. First of all, we estimate the parameter of the normal distribution from the sample average $\mu = 47.4345 \times 10^6$ and the sample standard deviation $\sigma = 0.2540 \times 10^6$. Figure 4 depicts the histogram of the simulated 99.9% quantiles derived from the 1000 samples. It is apparent that it follows the normal distribution.

As an additional check for the normality assumption, we apply the Shapiro-Wilk test using the R function shapiro.test, see R Core Team (2016). The p-value of the test is 0.4642, implying that the normality distribution cannot be rejected due to being considerably higher than 0.05.

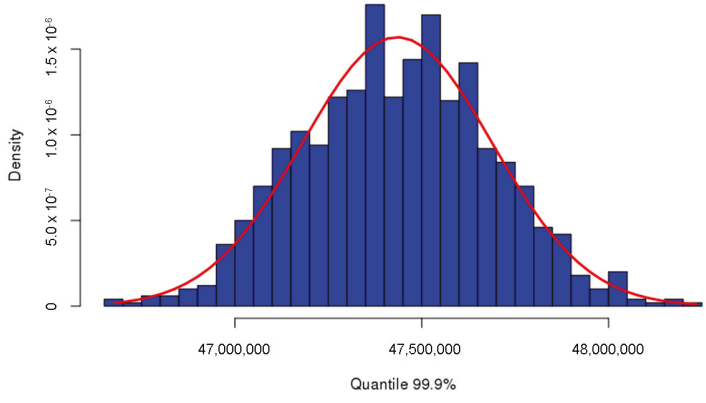


Figure 4. Histogram of 99.9% quantiles derived from 1000 Monte Carlo repetitions, compared with the estimated normal distribution.

4.5. Constrained Optimization Methods

In this section, we introduce two alternative methods for estimating the Monte Carlo error. In fact, we have not seen these methods introduced earlier in the literature.

Adopting the earlier notation introduced and setting $k = [np]$, it is natural to base the interval estimation for the 99.9% quantile on the order statistic $X_{k:n} = F_n^{-1}(p)$. We look for the shortest interval $[a, b] \subset \mathbb{R}$ such that

$$\mathbb{P}\{a < X_{k:n} \leq b\} = 0.9. \tag{33}$$

Let $\alpha, \beta \in [0, 1]$ be such that and $F^{-1}(\alpha) = a$ and $F^{-1}(\beta) = b$, and let

$$\mathbb{P}\{a < X_{k:n} \leq b\} = \mathbb{P}\{F^{-1}(\alpha) < X_{k:n} \leq F^{-1}(\beta)\} = 0.9. \tag{34}$$

We know that $\mathbb{P}\{X_{k:n} \leq x\} = \sum_{j=k}^{n} \binom{n}{j} F^j(x)(1 - F(x))^{n-j}$. Therefore

$$\mathbb{P}\{F^{-1}(\alpha) < X_{k:n} \leq F^{-1}(\beta)\} = \sum_{j=k}^{n} \binom{n}{j} \beta^j (1-\beta)^{n-j} - \sum_{j=k}^{n} \binom{n}{j} \alpha^j (1-\alpha)^{n-j}$$

and the CI for $F^{-1}(p)$ is obtained by searching for α and β such that

- $\sum_{j=k}^{n} \binom{n}{j} \beta^j (1-\beta)^{n-j} - \sum_{j=k}^{n} \binom{n}{j} \alpha^j (1-\alpha)^{n-j} = 0.9$, and
- $(\beta - \alpha)$ is minimal.

This constrained optimization problem can be solved numerically by considering a grid of α and β values close to p.

Equivalently, we can restate our research question into the search of α and β such that

- $\mathbb{P}\{\alpha < U_{k:n} \leq \beta\} = 0.9$, and
- $(\beta - \alpha)$ is minimal,

where $U_{k:n} = F(X_{k:n})$ is a uniform r.v. on [0,1]. Here we have used the fact that if X is a continuous random variable with cdf F, then the random variable $U = F(X)$ has a uniform distribution on [0,1]. Therefore α and β can be obtained by repeated draws from uniform random variables on [0,1].

Coming back to the CI for VaR, we have

$$\mathbb{P}\{F^{-1}(\alpha) < F_n^{-1}(p) \leq F^{-1}(\beta)\} = 0.9,$$

that is

$$\mathbb{P}\{F_n^{-1}(p) \leq F^{-1}(\beta)\} - \mathbb{P}\{F_n^{-1}(p) \leq F^{-1}(\alpha)\} = 0.9. \tag{35}$$

Note the equations

$$\mathbb{P}\{F_n^{-1}(p) \leq F^{-1}(\beta)\} = \mathbb{P}\left\{\frac{F^{-1}(p)}{F^{-1}(\beta)} F_n^{-1}(p) \leq F^{-1}(p)\right\}$$

and

$$\mathbb{P}\{F_n^{-1}(p) \leq F^{-1}(\alpha)\} = \mathbb{P}\left\{F_n^{-1}(p)\frac{F^{-1}(p)}{F^{-1}(\alpha)} \leq F^{-1}(p)\right\}.$$

Consequently, Equation (35) becomes

$$\mathbb{P}\left\{\frac{F^{-1}(p)}{F^{-1}(\beta)} F_n^{-1}(p) \leq F^{-1}(p) \leq \frac{F^{-1}(p)}{F^{-1}(\alpha)} F_n^{-1}(p)\right\} = 0.9$$

and implies the following CI

$$[L\ ;\ R] = \left[\frac{\left(F_n^{-1}(p)\right)^2}{F_n^{-1}(\beta)}\ ;\ \frac{\left(F_n^{-1}(p)\right)^2}{F_n^{-1}(\alpha)}\right]$$

where β and α are functions of n and p obtained using any of the two methods presented earlier.

Next are the results obtained using the first constrained optimization method:

$$VaR^{O1}(0.999) = 47.8037 \times 10^6, \tag{36}$$

$$[L^{O1}; R^{O1}] = [47.3399 \times 10^6; 48.2386 \times 10^6], \tag{37}$$

$$E^{O1} = (R^{O1} - L^{O1}) = 0.8987 \times 10^6, \tag{38}$$

$$RE^{O1} = (R^{O1} - L^{O1})/VaR^{O1}(0.999) = 0.018799 = 1.8799\%. \tag{39}$$

The second constrained optimization method yields the following results:

$$VaR^{O2}(0.999) = 47.8037 \times 10^6, \tag{40}$$

$$[L^{O2}; R^{O2}] = [47.4584 \times 10^6; 49.0432 \times 10^6], \tag{41}$$

$$E^{O2} = (R^{O2} - L^{O2}) = 1.5848 \times 10^6, \tag{42}$$

$$RE^{O2} = (R^{O2} - L^{O2})/VaR^{O2}(0.999) = 0.033153 = 3.3153\%. \tag{43}$$

We observe that the second method provides a higher relative error, probably due to the less accurate evaluation of α, obtained by redrawing from the uniform distribution on [0,1]. In fact, while in the first constrained optimization method we used a grid of 10,000 points close to p, in the second one we drew 10,000 values from the uniform distribution on [0,1]. These configurations were used to have a comparable computational effort between the two methods. The obtained results suggest that the second constrained optimization method would require a higher computational effort to reach a relative error comparable to the first one.

4.6. Comparison among the Described Methods

All the results that we have obtained using the methods discussed in this paper are summarized in Table 1. We applied different methodologies for constructing CI's for the VaR to realistic data samples, mimicking annual loss distributions and obtained via the Monte Carlo method with 5 million steps based on the *lognormal (9,2)* severity distribution and the *Poisson (100)* frequency distribution.

First of all, we see from Table 1 that all the approaches (apart from the *m*-out-of-*n*-bootstrap and the constrained optimization II, for the above mentioned reasons) are consistent with each other, and the method based on Monte Carlo repetitions is slightly less conservative than the other methods. On the other hand, the binomial method allows obtaining consistent results with the great advantage of requiring much less computational effort. In fact, the binomial method allows constructing the CI for the VaR using closed forms and, on the contrary to other methods, does not require to perform further simulations. Applying the Monte Carlo repetition method, in both cases considering 100 and 1000 samples, and the binomial, *n*-bootstrap and constrained optimization I methods, we obtain a relative error not higher than 2%. Therefore, whenever we fix the tolerance threshold for the relative error equal to 2%, we can conclude that 5 million steps are sufficient to mitigate sampling errors, as prescribed in EU Regulation 2018/959. For tolerance thresholds lower than 2%, we presumably have to increase the number of Monte Carlo steps or, alternatively, we have to adopt more sophisticated approaches, such as the variance reduction techniques or the Quasi-Monte Carlo methods.

Table 1. Estimated Monte Carlo errors (values in millions of Euro).

Method	VaR(0.999)	*L*	*R*	*E*	*RE*
Binomial	47.8037	47.3667	48.2897	0.9231	1.9310%
n-Bootstrap (100 samples)	47.7977	47.2914	48.2306	0.9392	1.9649%
m-out-of-*n*-Bootstrap (100 samples)	47.8173	46.9388	48.9690	2.0302	4.2457%
MC repetitions (100 samples)	47.4246	47.0261	47.8522	0.8261	1.7420%
MC repetitions (1000 samples)	47.4345	47.0207	47.8503	0.8297	1.7491%
Constrained optimization I	47.8037	47.3399	48.2386	0.8987	1.8799%
Constrained optimization II	47.8037	47.4584	49.0432	1.5848	3.3153%

5. Conclusions and Further Work

Motivated by the recently published EU Regulatory Technical Standards on AMA models, we have analyzed and compared six methods to estimate the Monte Carlo error when estimating VaR in operational risk measurement. The purpose is to assess the capital charge for the financial institution, usually defined through the VaR measure as the 99.9% quantile of the annual loss distribution with the holding period of one year. Presently, financial analysts and regulators are more and more focused on how to mitigate sampling errors, and this concerns the assessment of CI's for the Value-at-Risk along the lines of EU Regulation 959/2018. In this paper we have aimed at offering to practitioners an easy-to-follow rule of thumb for choosing approaches to evaluate such CI's and, if needed, for tuning them in terms of the number of required Monte Carlo replications. In particular, we started with the more straightforward method based on the binomial distribution, and we compared it with other methods such as the *n*-bootstrap, the *m*-out-of-*n*-bootstrap, the Monte Carlo repetition method, and two other techniques based on constrained optimization. All the methods have been applied to realistic simulated data. Our results show that the method based on Monte Carlo repetitions is slightly less conservative than the other methods. Furthermore, we have seen that the *m*-out-of-*n*-bootstrap and the second method based on constrained optimization give less precise results, in terms of relative error. The binomial method, which requires a much lower computational effort, seems to be particularly useful for the purpose, due to its precision. In fact, the binomial method is just slightly more conservative than the Monte Carlo repetition method, which can be intuitively considered as the most accurate for estimating the Monte Carlo error. Moreover, it is worth mentioning that the more sophisticated Monte Carlo methods, such as the variance reduction techniques and Quasi-Monte Carlo methods, offer several promising technical avenues for implementing the operational risk capital requirement estimation methods that we discussed. They can be subjects for future research.

Author Contributions: Conceptualization, F.G., F.P. and R.Z.; methodology, F.G., F.P. and R.Z.; software, F.P.; validation, F.G. and R.Z.; formal analysis, F.G., F.P. and R.Z.; investigation, F.G., F.P. and R.Z.; resources, F.G., F.P.

and R.Z.; data curation, F.G., F.P. and R.Z.; writing—original draft preparation, F.G., F.P. and R.Z.; writing—review and editing, F.G., F.P. and R.Z.; visualization, F.G., F.P. and R.Z.; supervision, F.G. and R.Z.; project administration, F.G., F.P. and R.Z.

Funding: This research received no external funding.

Acknowledgments: We are indebted to four anonymous reviewers for careful reading, constructive criticism, and insightful suggestions that guided our work on the revision.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AMA	Advanced Measurement Approach
cdf	cumulative distribution function
CI	Confidence Interval
E	Error
EU	European Union
EVT	Extreme Value Theory
ORC	Operational Risk Category
pdf	probability density function
LDA	Loss Distribution Approach
RE	Relative Error
RTS	Regulatory Technical Standards
SLA	Single Loss Approximation
VaR	Value-at-Risk

References

Bazzarello, Davide, Bert Crielaard, Fabio Piacenza, and Aldo Soprano. 2006. Modeling insurance mitigation on operational risk capital. *Journal of Operational Risk* 1: 57–65. [CrossRef]

Bickel, Peter J., and Anat Sakov. 2008. On the choice of m in the m out of n bootstrap and confidence bounds for extrema. *Statistica Sinica* 18: 967–85.

Bickel, Peter J., Friedrich Götze, and Willem R. van Zwet. 1997. Resampling fewer than n observations: Gains, losses, and remedies for losses. *Statistica Sinica* 7: 1–31.

Böcker, Klaus, and Claudia Klüppelberg. 2005. Operational VaR: A closed-form approximation. *Risk Magazine* 18: 90–93.

Böcker, Klaus, and Jacob Sprittulla. 2006. Operational VaR: Meaningful means. *Risk Magazine* 19: 96–98.

Bolancé, Catalina, Montserrat Guillén, Jim Gustafsson, and Jens Perch Nielsen. 2012. *Quantitative Operational Risk Models*. Abingdon: CRC Press, Taylor & Francis Group.

Boos, Dennis D. 1984. Using extreme value theory to estimate large percentiles. *Technometrics* 26: 33–39. [CrossRef]

Boyle, Phelim P. 1977. Options: A monte carlo approach. *Journal of Financial Economics* 4: 323–38. [CrossRef]

Cavallo, Alexander, Benjamin Rosenthal, Xiao Wang, and Jun Yan. 2012. Treatment of the data collection threshold in operational risk: A case study with the lognormal distribution. *Journal of Operational Risk* 7: 3–38. [CrossRef]

Chen, E. Jack, and W. David Kelton. 2001. Quantile and histogram estimation. Paper Presented at the 2001 Winter Simulation Conference, Arlington, VA, USA, December 9–12.

Chernobai, Anna, Svetlozar T. Rachev, and Frank J. Fabozzi. 2005. Composite goodness-of-fit tests for left-truncated loss samples. *Handbook of Financial Econometrics and Statistics* 575–96. [CrossRef]

Colombo, Andrea, Alessandro Lazzarini, and Silvia Mongelluzzo. 2015. A weighted likelihood estimator for operational risk data: Improving the accuracy of capital estimates by robustifying maximum likelihood estimates. *Journal of Operational Risk* 10: 47–108. [CrossRef]

Cope, Eric W., Giulio Mignola, G. Antonini, and R. Ugoccioni. 2009. *Challenges in Measuring Operational Risk From Loss Data*. Technical Report. Torino: IBM Zurich Research Lab, Intesa Sanpaolo SpA.

Cortés, Lina M., Andrés Mora-Valencia, and Javier Perote. 2017. Measuring firm size distribution with semi-nonparametric densities. *Physica A* 485: 35–47. [CrossRef]

Dalla Valle, Luciana, and Paolo Giudici. 2008. A bayesian approach to estimate the marginal loss distributions in operational risk management. *Computational Statistics & Data Analysis* 52: 3107–27.

Danesi, Ivan Luciano, and UniCredit Business Integrated Solutions. 2015. *Robust Estimators for Operational Risk: A Computationally Efficient Approach*. Technical Report. Working Paper Series No. 63; Milano: UniCredit & Universities.

Danesi, Ivan Luciano, Fabio Piacenza, Erlis Ruli, and Laura Ventura. 2016. Optimal B-robust posterior distributions for operational risk. *Journal of Operational Risk* 11: 1–20. [CrossRef]

Danielsson, Jon, and Casper G. De Vries. 1996. Tail index and quantile estimation with very high frequency data. *Journal of Empirical Finance* 4: 241–57. [CrossRef]

Di Clemente, Annalisa, and Claudio Romano. 2004. A copula-extreme value theory approach for modelling operational risk. In *Operational Risk Modelling and Analysis: Theory and Practice*. Edited by M. Cruz. London: Risk Books, vol. 9, pp. 189–208.

Efron, Bradley. 1979. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics* 7: 1–26. [CrossRef]

Efron, Bradley. 1982. *The Jackknife, the Bootstrap and Other Resampling Plans*. Number 38 in Regional Conference Series in Applied Mathematics. Philadelphia: Society for Industrial and Applied Mathematics.

Embrechts, Paul, and Marco Frei. 2009. Panjer recursion versus fft for compound distributions. *Mathematical Methods of Operations Research* 69: 497–508. [CrossRef]

European Parliament and Council of the European Union. 2013. Capital Requirement Regulations. Technical Report, European Union. Available online: https://eur-lex.europa.eu/eli/reg/2013/575/oj (accessed on 26 June 2013).

European Parliament and Council of the European Union. 2018. Regulatory Technical Standards of the Specification of the Assessment Methodology Under Which Competent Authorities Permit Institutions to Use AMA for Operational Risk. Technical Report, European Union. Available online: https://eur-lex.europa.eu/legal-content/GA/TXT/?uri=CELEX:32018R0959 (accessed on 14 March 2008).

Figini, Silvia, Lijun Gao, and Paolo Giudici. 2014. Bayesian operational risk models. *Journal of Operational Risk* 10: 45–60. [CrossRef]

Frachot, Antoine, Pierre Georges, and Thierry Roncalli. 2001. Loss Distribution Approach for Operational Risk. Technical Report. Lyonnais. Available online: http://www.thierry-roncalli.com/download/lda-practice.pdf (accessed on 7 May 2003).

Frachot, Antoine, Olivier Moudoulaud, and Thierry Roncalli. 2007. Loss distribution approach in practice. In *The Basel Handbook: A Guide for Financial Practitioners*. Edited by Michael Ong and Jorg Hashagen. London: Risk Books, pp. 527–55.

Gomes, M. Ivette, and Dinis Pestana. 2007. A sturdy reduced-bias extreme quantile (var) estimator. *Journal of the American Statistical Association* 102: 280–92. [CrossRef]

Gribkova, Nadezhda V., and Roelof Helmers. 2011. On the consistency of the m << n bootstrap approximation for a trimmed mean. *Theory of Probability and Its Applications* 55: 42–53.

Guharay, Sabyasachi, K. C. Chang, and Jie Xu. 2016. Robust estimation of value-at-risk through correlated frequency and severity model. Paper presented at 19th International Conference on Information Fusion, Heidelberg, Germany, July 5–8.

Kalos, Malvin H., and Paula A. Whitlock. 2008. *Monte Carlo Methods*. New York: Wiley.

Kleiber, Christian, and Samuel Kotz. 2003. *Statistical Size Distributions in Economics and Actuarial Sciences*. New York: Wiley.

Kleijnen, Jack P. C., Ad A. N. Ridder, and Reuven Y. Rubinstein. 2010. Variance reduction techniques in monte carlo methods. *Encyclopedia of Operations Research and Management Science*, 1598–610. [CrossRef]

Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot. 2012. *Loss Models: From Data to Decisions*, 4th ed. New York: Wiley.

Lambrigger, Dominik D., Pavel V. Shevchenko, and Mario V. Wüthrich. 2007. The quantification of operational risk using internal data, relevant external data and expert opinion. *Journal of Operational Risk* 2: 3–27. [CrossRef]

Larsen, Paul. 2016. Operational risk models and asymptotic normality of maximum likelihood estimation. *Journal of Operational Risk* 11: 55–78. [CrossRef]

Larsen, Paul. 2019. Maximum likelihood estimation error and operational value-at-risk stability. *Journal of Operational Risk* 14: 1–23. [CrossRef]

Lavaud, Sophie, and Vincent Lehérissé. 2014. Goodness-of-fit tests and selection methods for operational risk. *Journal of Operational Risk* 9: 21–50. [CrossRef]

L'ecuyer, Pierre. 1999. Good parameters and implementations for combined multiple recursive random number generators. *Operations Research* 47: 159–64. [CrossRef]

Luo, Xiaolin, Pavel V. Shevchenko, and John B. Donnelly. 2007. Addressing the impact of data truncation and parameter uncertainty on operational risk estimates. *Journal of Operational Risk* 2: 3–26. [CrossRef]

Makarov, Mikhail. 2006. Extreme value theory and high quantile convergence. *Journal of Operational Risk* 1: 51–57. [CrossRef]

Matsumoto, Makoto, and Takuji Nishimura. 1998. Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation* 8: 3–30. [CrossRef]

Metropolis, Nicholas, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. 1953. Equations of state calculations by fast computing machines. *Journal of Chemical Physics* 21: 1087–91. [CrossRef]

Mignola, Gulio, and Roberto Ugoccioni. 2005. Tests of extreme value theory. *Operational Risk* 6: 32–35.

Mignola, Gulio, and Roberto Ugoccioni. 2006. Sources of uncertainty in modeling operational risk losses. *Journal of Operational Risk* 1: 33–50. [CrossRef]

Owen, Art B. 2013. *Monte Carlo Theory, Methods and Examples*. Stanford: Department of Statistics, Stanford University.

Panjer, Harry H. 2006. *Operational Risk: Modeling Analytics*. New York: Wiley.

Peng, Liang, and Yongcheng Qi. 2006. Confidence regions for high quantiles of a heavy tailed distribution. *Ann. Statist.* 34: 1964–86. [CrossRef]

R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Richtmyer, Robert D. 1952. *The Evaluation of Definite Integrals, and a Quasi-Monte Carlo Method Based on the Properties of Algebraic Numbers*. Technical Report, LA-1342. Los Angeles: University of California.

Sakov, Anat. 1998. *Using the m out n Bootstrap in the Hypothesis Testing*. Berkeley: University of California.

Shevchenko, Pavel V. 2011. *Modelling Operational Risk Using Bayesian Inference*. New York: Springer.

Shevchenko, Pavel V., and Mario V. Wuthrich. 2006. The structural modelling of operational risk via bayesian inference: Combining loss data with expert opinions. *Journal of Operational Risk* 1: 3–26. [CrossRef]

Student. 1908. The probable error of a mean. *Biometrika* 6: 1–25. [CrossRef]

risks

MDPI

Article

A User-Friendly Algorithm for Detecting the Influence of Background Risks on a Model

Nadezhda Gribkova [1] and Ričardas Zitikis [2,*]

[1] Faculty of Mathematics and Mechanics, St. Petersburg State University, St. Petersburg 199034, Russia; nv.gribkova@gmail.com
[2] School of Mathematical and Statistical Sciences, Western University, London, ON N6A 5B7, Canada
* Correspondence: rzitikis@uwo.ca; Tel.: +1-519-432-7370

Received: 28 August 2018; Accepted: 10 September 2018; Published: 14 September 2018

Abstract: Background, or systematic, risks are integral parts of many systems and models in insurance and finance. These risks can, for example, be economic in nature, or they can carry more technical connotations, such as errors or intrusions, which could be intentional or unintentional. A most natural question arises from the practical point of view: is the given system really affected by these risks? In this paper we offer an algorithm for answering this question, given input-output data and appropriately constructed statistics, which rely on the order statistics of inputs and the concomitants of outputs. Even though the idea is rooted in complex statistical and probabilistic considerations, the algorithm is easy to implement and use in practice, as illustrated using simulated data.

Keywords: background risk; systematic risk; transfer function; information processing; order statistic; concomitant

1. Introduction

Actuarial, financial, and economic literature is abundant with models and analyses of background, or systematic, risks that affect decision making (cf., e.g., Finkelshtain et al. 1999; Franke et al. 2006, 2011; Nachman 1982; Pratt 1998; Guo et al. 2018; Furman et al. 2018; and references therein). Various models have been proposed, including additive, multiplicative, and more intricate ones that couple underlying losses (or, generally speaking, inputs) with background risks. For recent far-reaching contributions to this area, we refer to Perote et al. (2015), Su (2016), Su and Furman (2017a, 2017b) Semenikhine et al. (2018), Guo et al. (2018), as well as to the extensive lists of references therein.

Systems and thus their models are prone to a myriad of intentional or unintentional disruptions, which could affect inputs and/or outputs. The literature on the topic is vast, and some of the recent contributions include those tackling deliberate intrusions (e.g., Cárdenas et al. 2011; Premathilaka et al. 2013), and false data injections (e.g., Liang et al. 2017). A number of sophisticated methods have been developed for tackling the problems (e.g., Huang et al. 2016; Onoda 2016; He et al. 2017; Potluri et al. 2017), to name a few.

Whether or not these risks affect the underlying input variables and thus decision making is a problem of immense interest. From the conceptual point of view, broadly speaking, two scenarios arise. First, if it is suspected that the outputs are affected, then testing whether or not this is indeed the case falls, in a sense, within the context of regression analysis, though additional statistical challenges arise (e.g., Perote and Perote-Peña 2004; Perote et al. 2015.; Chen et al. 2018; Gribkova and Zitikis 2018). The second scenario, which is the main topic of the present paper, deals with the case when it is the inputs that are possibly affected by risks.

Statistically speaking, given the input and output random variables X and Y, respectively, which in the risk-free scenario are connected by a "transfer" function h via the equation

$$Y = h(X), \tag{1}$$

we wish to have an algorithm that would tell us whether risk-free model (1) is true or the risk-contaminated one

$$Y = h(X + \delta), \tag{2}$$

where δ is an exogenous risk, sometimes called input-reading error, that directly affects the input X and thus, indirectly, the output variable as well. We note that Chen et al. (2018) consider model (1) with deterministic inputs, like those to be defined in Equation (3) below. Gribkova and Zitikis (2018) explore risk-free model (1), which can be viewed as the "null hypothesis" in the context of the present paper. Hence, model (2) can be viewed as the "alternative hypothesis," and the algorithm to be constructed and illustrated in this paper will distinguish between the two hypotheses.

The rest of the paper is organized as follows. In Section 2, we lay out the foundations for assessing the presence, or absence, of input-affecting risks. In Section 3, we describe the algorithm itself. It relies on two statistics whose roles, interrelationship, and asymptotic properties are presented in Sections 4 and 5. Section 6 concludes the paper with a brief overview of main findings.

2. The Model

Systems are usually associated with finite-length transfer windows, say $[a, b] \subset \mathbf{R}$, and also with transfer functions $h : [a, b] \to \mathbf{R}$. Let $X_1, \ldots, X_n$ be input random variables, which we assume to be pre-whitened (e.g., Box et al. 2015), that is, independent and identically distributed (iid). Denote their marginal cumulative distribution functions (cdf) by $F(x)$, whose support is the transfer window $[a, b]$. Hence, the input values are always in $[a, b]$. We assume that the cdf $F(x)$ is strictly increasing on the interval $[a, b]$, with $F(a) = 0$ and $F(b) = 1$. In fact, to simplify mathematics and still cover a wide variety of applications, we assume that the cdf is continuously differentiable and its probability density function (pdf) is bounded away from 0 on the transfer window $[a, b]$.

Denote the input-affecting risks by $\delta_1, \ldots, \delta_n \sim F_\delta$, which act upon the inputs $X_1, \ldots, X_n$ as visualized in Figure 1.

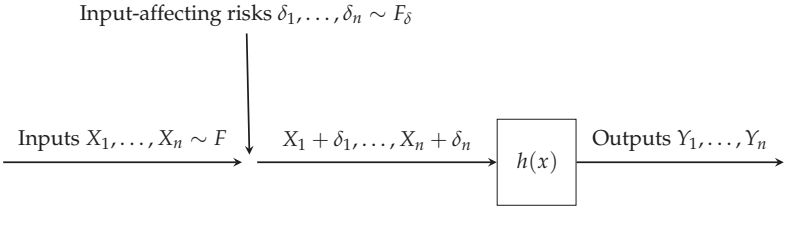


Figure 1. Are the input-affecting risks absent (i.e., degenerate at 0) or not?

We assume that the input-affecting risks are pre-whitened, that is, iid random variables, and we also assume that they are independent of the input variables $X_1, \ldots, X_n$ and are affecting their values in the additive way. The inputs X_i take values in the interval $[a, b]$, but the risks δ_i, being exogenous variables, are not restricted to any domain and can therefore take any real values. Our goal in this paper is to offer a practical way for detecting whether or not the risks are absent, or present. Two following notes relate our research to the topics in statistical literature.

First, the problem that we tackle is different from that dealing with errors-in-variables, where observations already contain errors, whereas in our case, the inputs X_i are uncontaminated but possibly

become such while being transferred into the filter, also known as the transmission channel in the engineering literature. That is, in the errors-in-variables scenario, we would observe $X_i + \delta_i$, whereas in the present context we observe the original inputs X_i and want to know whether or not they are affected by δ_i.

Second, there is a connection between our research and classical regression, and we have already noted contributions by Perote and Perote-Peña (2004), Perote et al. (2015), where we also find extensive lists of related references. Namely, given the outputs $Y_i = h(X_i + \delta_i)$ and assuming for the sake of argument that the risks δ_i are small, the Taylor formula gives the approximation $Y_i \approx h(X_i) + h'(X_i)\delta_i$, which places the input-based scenario into the output-based scenario $Y_i = h(X_i) + \varepsilon_i$, but the risks $\varepsilon_i \approx h'(X_i)\delta_i$ depend on the inputs X_i via the term $h'(X_i)$. This dependence feature presents a major hurdle, which we circumvent in our following considerations and produce a user-friendly algorithm for detecting δ_i's when they are present.

Throughout the paper we assume that the transfer function $h(x)$ has a bounded and continuous first derivative, and we also assume that the derivative is not identically equal to 0, thus ruling out the trivial case of constant transfer functions. Actually, throughout the paper we also exclude the case $h(a) = h(b)$, which causes some technical complications but is hardly of practical relevance, as we shall explain in the next section. If, however, due to some considerations we would need to depart from these conditions, then there is room for relaxing the conditions, though naturally at the expense of more complex considerations.

3. The Algorithm

We first elaborate on the definition of outputs. Indeed, even though X_i's are in the transfer window $[a, b]$, the affected inputs $X_i + \delta_i$ may or may not be in $[a, b]$, which is the domain of definition of the transfer function $h(x)$. Hence, the actual outputs are

$$\begin{aligned} Y_i &= h\big(\max\{a, \min\{X_i + \delta_i, b\}\}\big) \\ &= g(X_i + \delta_i), \quad i = 1, \ldots, n, \end{aligned}$$

where

$$g(x) = h\big(\max\{a, \min\{x, b\}\}\big).$$

Since the cdf of X is continuous, we can uniquely order the random variables $X_1, \ldots, X_n$. The resulting order statistics $X_{1:n} < \cdots < X_{n:n}$ give rise to the concomitants $Y_{1,n}, \ldots, Y_{n,n}$ (e.g., David and Nagaraja 2003). Based on them and using the notation $x_+ = \max\{x, 0\}$, we define the statistics

$$A_n := \frac{1}{\sqrt{n}} \sum_{i=2}^{n} \big(Y_{i,n} - Y_{i-1,n}\big)_+$$

and

$$B_n := \frac{1}{\sqrt{n}} \sum_{i=2}^{n} \big|Y_{i,n} - Y_{i-1,n}\big|,$$

and then, in turn, their ratio

$$I_n := \frac{A_n}{B_n}.$$

The algorithm, to be introduced in a moment, for detecting input-affecting risks is based on asymptotics, when n gets large, of I_n and B_n, which we call the *pivot* and its *supporter*, thus hinting at their main and supporting roles, respectively. Before formulating the algorithm, we make the natural assumption that the risks, when they exist, should not be so large that the performance of the system would be derailed to such an extent that it becomes unnecessary to run any algorithm. For the purpose of rigour, in the following definition we summarize the circumstances under which there is ambiguity

as to the absence, or presence, of input reading risks, and thus employing the algorithm becomes warranted.

Definition 1. *The presence of input-affecting risk is suspected, and thus becomes a subject for testing, when it is believed that there is a set $T \subset [a, b]$ such that the event $X \in T$ has a (strictly) positive probability and, for all $x \in T$, the random variable $g(x + \delta)$ is non-degenerate, due to the random δ.*

We note at the outset that Definition 1 is a user-friendly reformulation of technically-looking condition (10) to be presented in Section 5 below, where it plays a pivotal role in setting rigorous mathematical foundations for our algorithm. In this regard, we note that the condition is tightly tied to the indefinite growth of B_n when the sample size n grows, as we shall see in Theorem 3 below. Hence, if the subject-matter knowledge is not sufficiently convincing for the decision maker to see whether or not the circumstances delineated by Definition 1 hold, then data-based checking of the asymptotic behaviour of B_n for large n should clarify the situation.

Definition 1 implies that the system's output $Y = g(X + \delta)$ varies not just because of X but also because of δ, assuming of course that the latter is present, that is, is not degenerate at 0. This, for example, excludes situations (as unquestionably obvious) when $g(x + \delta) = g(a)$ for every x (i.e., when $-\delta > 0$ is very large), or when $g(x + \delta) = g(b)$ for every x (i.e., when $\delta > 0$ is very large). In either of these extreme cases, the decision maker would immediately see the system's malfunction because of the outputs constantly lingering on, or near, the boundaries $g(a)$ and $g(b)$, and thus no special testing would be warranted.

We are now ready to formulate the algorithm for detecting the input-affecting risk when its presence is suspected.

Case 1: The pivot I_n is *not* approaching $1/2$.

(i) If I_n decisively tends to a limit other than $1/2$, then we advise the decision maker about the *absence* of the risk.

(ii) If I_n seems to tend to a limit other than $1/2$ but there is some doubt as to whether this is true, then we check if the supporter B_n is asymptotically bounded, and if yes, then we advise the decision maker about the *absence* of the risk.

Case 2: The pivot I_n is approaching $1/2$.

(i) If the supporter B_n tends to infinity, then we advise the decision maker about the *presence* of the risk.

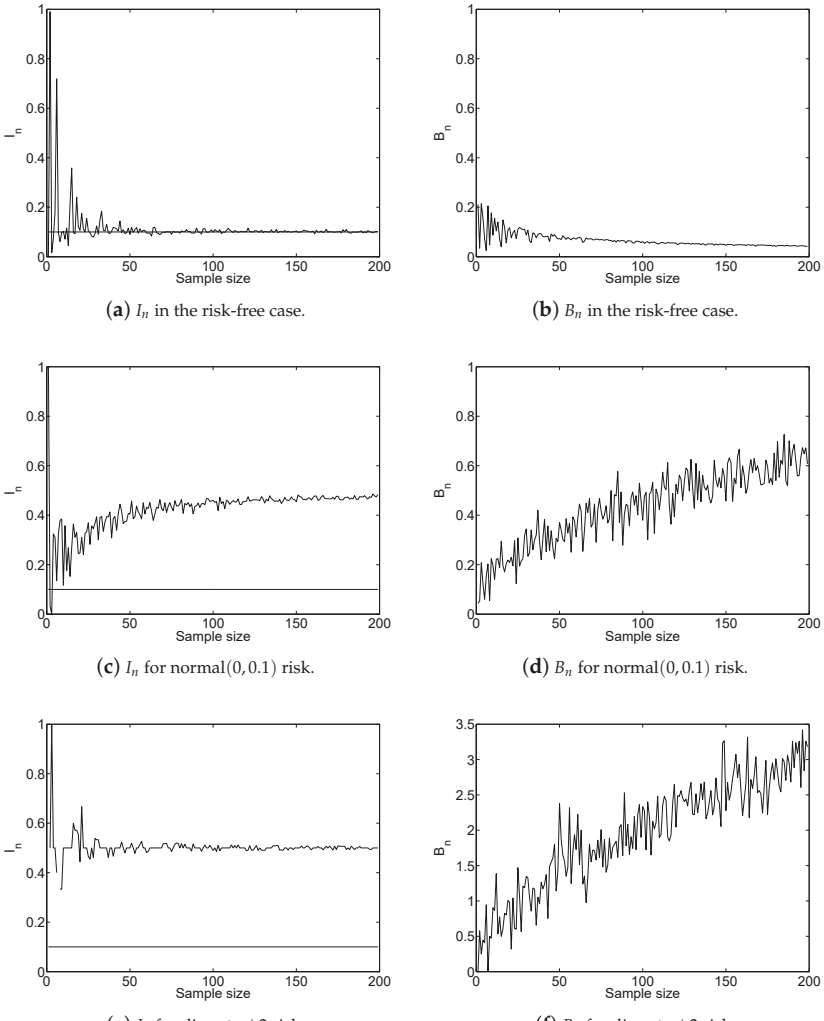
(ii) If the supporter B_n is asymptotically bounded, then $h(a)$ and $h(b)$ are likely to be insufficiently different to have already triggered Case 1 above, and we thus advise the decision maker about the *absence* of the risk.

In the next two sections, we present rigorous results upon which the above algorithm relies. We note in passing that irrespective of whether the algorithm detects risks or not, in either case we can still wish to double-check the findings. It can also be necessary to check the system's vulnerability (e.g., Hug and Giampapa 2012; and references therein). In such cases, we can use artificially constructed inputs, such as

$$x_{i,n} = a + (b - a)\frac{i - 1}{n - 1}, \quad i = 1, \ldots, n. \tag{3}$$

We conclude this section with an example that shows how the algorithm works in practice. For this, let the transfer function be $h(x) = 1 - (x - 0.25)^2$ for $x \in [0, 1]$. Furthermore, upon recalling that the (unconditional) Lomax cdf is $1 - (1 + x/\beta))^{-\alpha}$ for $x \geq 0$, with shape and scale parameters $\alpha > 0$ and $\beta > 0$, we assume that the input X follows the Lomax(α, β) distribution *conditioned* on the transfer interval $[0, 1]$. Throughout the illustration, we set $\alpha = 1.5$ and $\beta = 1$.

Let δ follow the normal distributions with the mean 0 and standard deviation σ. In the risk-free case (i.e., $\sigma = 0$), the asymptotics of I_n and B_n is depicted in panels (a) and (b) of Figure 2, and when $\sigma = 0.1$, their asymptotics is depicted in panels (c) and (d). We also check the performance of the algorithm when the risk δ is discrete, specifically, when it is equal to -2 with probability 0.7 and to 2 with probability 0.3. The asymptotics of I_n and B_n is depicted in panels (e) and (f) of Figure 2.



(**a**) I_n in the risk-free case.

(**b**) B_n in the risk-free case.

(**c**) I_n for normal$(0, 0.1)$ risk.

(**d**) B_n for normal$(0, 0.1)$ risk.

(**e**) I_n for discrete ± 2 risk.

(**f**) B_n for discrete ± 2 risk.

Figure 2. The risk-detection algorithm via the asymptotics of the pivot I_n and its supporter B_n, with the horizontal line in the three left-hand panels at the height of I_h to be defined by Equation (4) below.

We see from the left-hand panels that the pivot I_n converges to the limit other than $1/2$ (i.e., to the value of I_h to be defined by Equation (4) in the next section) only in the risk-free case. The increasing pattern of B_n in panels (d) and (f) confirms the presence of input risk in both scenarios, which have initially been detected by the pivot I_n (due to its convergence to $1/2$) in panels (c) and (e). Note that

the convergence to $1/2$ in panel (e) is decisive, whereas the convergence in panel (c) may not be so well pronounced, and thus the increasing pattern of B_n in panel (d) provides reassurance.

4. Asymptotics of the Pivot I_n

We begin with the case when the input-affecting risk is absent, and thus the system is functioning properly. This is the starting point of many works (e.g., Cárdenas et al. 2011, p. 360) dealing with intrusion detection (e.g., Debar et al. 1999; Premathilaka et al. 2013), false data injections (e.g., Liang et al. 2017), and other disruptions. Recall the notation $x_+ = \max\{x, 0\}$ for any $x \in \mathbf{R}$.

Theorem 1 (Gribkova and Zitikis, 2018)**.** *If δ is absent, then, when $n \to \infty$, the pivot I_n converges to*

$$I_h := \frac{\int_a^b (h'(u))_+ \mathrm{d}u}{\int_a^b |h'(u)| \mathrm{d}u}. \tag{4}$$

For another perspective on the meaning of I_h, we refer to Davydov and Zitikis (2017) where I_h arises as the solution to an optimization problem. The importance of Theorem 1 in the present paper follows from the fact that when the cdf F_δ is non-degenerate, then (details in Section 5 below) the pivot I_n converges to $1/2$ when $n \to \infty$. Of course, the limit $1/2$ can also manifest when δ is absent, that is, in the context of Theorem 1, but this can happen only when $h(a) = h(b)$. Indeed, as it is easy to check using the equations $|x| = x_+ + x_-$ and $x = x_+ - x_-$ with $x_- = \max\{-x, 0\}$, we have $I_h = 1/2$ if and only if $h(a) = h(b)$. The latter property is, however, an exception rather than the rule: it manifests in such cases when, for example, the system is down and thus $h(x)$ takes the same value irrespective of $x \in [a, b]$. Hence, unless explicitly noted otherwise, throughout the paper we assume

$$h(a) \neq h(b), \tag{5}$$

as we have already mentioned earlier.

We next discuss how to check whether or not the risk δ is degenerate. Naturally, in order to detect anomalies, the original state of the system has to be in reasonable working order (cf., e.g., Cárdenas et al. 2011, p. 360). Gribkova and Zitikis (2018) have put forward an argument in favour of the following definition.

Definition 2. *A system is in reasonable working order whenever in the absence of input-affecting risk (i.e., when $\delta = 0$ almost surely), the sequence B_n is asymptotically bounded in probability. In mathematical terms, we write this as $B_n = O_{\mathbf{P}}(1)$ when $n \to \infty$.*

Given that in the absence of input-affecting risk we are exploring the asymptotic behaviour of the pivot I_n, which is the ratio of A_n and B_n, both of which are asymptotically bounded in probability, the requirement $B_n = O_{\mathbf{P}}(1)$ is natural. This can be seen from the following argument involving the mean-value theorem:

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=2}^{n} \big|h(X_{i:n}) - h(X_{i-1:n})\big| &= \frac{1}{\sqrt{n}} \sum_{i=2}^{n} \big|h'(\xi_i)\big(X_{i:n} - X_{i-1:n}\big)\big| \\ &\leq \frac{\|h'\|(b-a)}{\sqrt{n}} \end{aligned} \tag{6}$$

for some ξ_i between $X_{i-1:n}$ and $X_{i:n}$, where

$$\|h'\| := \sup_{a \leq x \leq b} |h'(x)| < \infty.$$

As a side-note, the right-hand side of bound (6) implies that, if needed, the boundedness of the first derivative of the transfer function can be relaxed and the system can still remain in reasonable working order, as per Definition 2.

We next present an example that shows what happens with the system when the input-affecting risk is present, that is, when the cdf F_δ is non-degenerate. Before starting the example, we recall (David and Nagaraja 2003) that the concomitants $Y_{1,n}, \ldots, Y_{n,n}$ can be written as follows

$$Y_{i,n} = g\big(X_{i:n} + \delta_{[i]}\big),$$

where $\delta_{[i]}$ is the random variable among $\delta_1, \ldots, \delta_n$ that corresponds to $X_{i:n}$. As noted by David and Nagaraja (2003, p. 145), the random variables $\delta_{[1]}, \ldots, \delta_{[n]}$ are iid and follow the cdf F_δ of the original risk δ.

Example 1. *Let δ take value $c > 0$ with probability $p \in (0,1)$ and $-c$ with probability $1-p$, and let $c \geq b - a$. The latter assumption implies that irrespective of the value of $X_{i:n}$, the value of $X_{i:n} + \delta_{[i]}$ is above b with probability p and below a with probability $1-p$. Hence, the concomitant $Y_{i,n}$ is equal to $h(b)$ with probability p and to $h(a)$ with probability $1-p$. Since each concomitant can take only two values, $|Y_{i,n} - Y_{i-1,n}|$ is equal to $|h(b) - h(a)|$ when $\delta_{[i]} \neq \delta_{[i-1]}$ and 0 otherwise. Consequently,*

$$|Y_{i,n} - Y_{i-1,n}| = \frac{|h(b) - h(a)|}{2c} \big|\delta_{[i]} - \delta_{[i-1]}\big|,$$

which implies

$$\begin{aligned} B_n &= \frac{|h(b) - h(a)|}{2c} \left\{ \frac{1}{\sqrt{n}} \sum_{i=2}^{n} \Big(\big|\delta_{[i]} - \delta_{[i-1]}\big| - \mathbf{E}\big[\big|\delta_{[i]} - \delta_{[i-1]}\big|\big] \Big) + \frac{n-1}{\sqrt{n}} \mathbf{E}\big[\big|\delta_{[i]} - \delta_{[i-1]}\big|\big] \right\} \\ &= \frac{|h(b) - h(a)|}{2c} \sqrt{n}\, \mathbf{E}\big[\big|\delta_{[i]} - \delta_{[i-1]}\big|\big] + O_{\mathbf{P}}(1). \end{aligned} \tag{7}$$

Since the variables $\delta_{[1]}, \ldots, \delta_{[n]}$ are iid and follow the same cdf F_δ as the original δ, the mean $\mathbf{E}[|\delta_{[i]} - \delta_{[i-1]}|]$ is equal to $4cp(1-p)$ and thus Equation (7) implies

$$B_n = 2\sqrt{n}\, p(1-p)|h(b) - h(a)| + O_{\mathbf{P}}(1). \tag{8}$$

From this we conclude that if p is neither 0 nor 1, which we assume, and if $h(b) \neq h(a)$, which we also assume, then $B_n \xrightarrow{\mathbf{P}} \infty$ when $n \to \infty$. Analogous arguments lead to

$$A_n = \sqrt{n}\, p(1-p)|h(b) - h(a)| + O_{\mathbf{P}}(1). \tag{9}$$

Combining statements (8) and (9), we have $I_n = A_n / B_n \xrightarrow{\mathbf{P}} 1/2$ when $n \to \infty$, which in turn implies that the system is affected by the risk. This concludes Example 1.

The above example has been constructed to show—in a somewhat dramatic way—what happens when the input-affecting risk pushes the input outside the transfer window, but the same conclusion can be reached under much weaker assumptions on δ, as we shall show in the next section.

5. Growth of the Supporter B_n

The following general result plays a major role in the justification of the earlier presented algorithm.

Theorem 2. *Gribkova and Zitikis (2018) Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be independent copies of a generic random pair (X, Y), with X having continuous cdf and Y having finite second moment. If $B_n \xrightarrow{\mathbf{P}} \infty$, then $I_n \xrightarrow{\mathbf{P}} 1/2$ when $n \to \infty$.*

We know from statement (6) and the arguments around it that if δ is degenerate, then $B_n \xrightarrow{\mathbf{P}} \infty$ cannot be true. In the next theorem, we give a necessary and sufficient condition for $B_n \xrightarrow{\mathbf{P}} \infty$ to hold, which, according to Theorem 2, implies $I_n \xrightarrow{\mathbf{P}} 1/2$.

Theorem 3. *The statement $B_n \xrightarrow{\mathbf{P}} \infty$ holds if and only if*

$$\int_0^1 (2t-1)\left(\int_0^1 F^{-1}_{g(x_s+\delta)}(t)\mathrm{d}s\right)\mathrm{d}t > 0, \tag{10}$$

where $x_s = F^{-1}(s)$ is the sth percentile of X, and $F^{-1}_{g(x_s+\delta)}(t)$ denotes the quantile function of the random variable $g(x_s+\delta)$.

Condition (10) arises naturally, but its formulation is not user friendly. Remarkably, its meaning is very simple and has already been conveyed in Definition 1. Before proving Theorem 3, we next illuminate the meaning of condition (10) by revisiting Example 1 through the lens of the condition.

Example 2. *Let δ take value $c > 0$ with probability $p \in (0,1)$ and $-c$ with probability $1-p$, and let $c \geq b-a$. Since for every $s \in (0,1)$ we have $x_s = F^{-1}(s) \in [a,b]$, the random variable $g(x_s+\delta)$ has the probability distribution*

$$g(x_s+\delta) = \begin{cases} h(a) & \text{with probability} \quad 1-p, \\ h(b) & \text{with probability} \quad p. \end{cases}$$

To obtain its quantile function, we start with the case $h(a) \leq h(b)$ and have the formula

$$F^{-1}_{g(x_s+\delta)}(t) = \begin{cases} h(a) & \text{when} \quad 0 < t \leq 1-p, \\ h(b) & \text{when} \quad 1-p < t \leq 1. \end{cases}$$

Consequently,

$$\begin{aligned}\int_0^1 (2t-1)\left(\int_0^1 F^{-1}_{g(x_s+\delta)}(t)\mathrm{d}s\right)\mathrm{d}t &= h(a)\int_0^{1-p}(2t-1)\mathrm{d}t + h(b)\int_{1-p}^1 (2t-1)\mathrm{d}t \\ &= p(1-p)\big(h(b)-h(a)\big).\end{aligned}$$

Analogous calculations when $h(a) \geq h(b)$ give the answer $p(1-p)(h(a)-h(b))$, thus establishing the equation

$$\int_0^1 (2t-1)\left(\int_0^1 F^{-1}_{g(x_s+\delta)}(t)\mathrm{d}s\right)\mathrm{d}t = p(1-p)\big|h(b)-h(a)\big|$$

irrespective of the values of $h(a)$ and $h(b)$. We can therefore conclude that as long as $h(b) \neq h(a)$ and p is neither 0 nor 1, condition (10) is satisfied. Thus, we have $B_n \xrightarrow{\mathbf{P}} \infty$ according to Theorem 3.

Proof of Theorem 3. We first show that if condition (10) is satisfied, then $B_n \xrightarrow{\mathbf{P}} \infty$. We start with the bound

$$\begin{aligned} B_n &= \frac{1}{\sqrt{n}}\sum_{i=2}^n \big|g(X_{i:n}+\delta_{[i]}) - g(X_{i-1:n}+\delta_{[i-1]})\big| \\ &\geq \frac{1}{\sqrt{n}}\sum_{i=2}^n \big|g(X_{i-1:n}+\delta_{[i]}) - g(X_{i-1:n}+\delta_{[i-1]})\big| - \frac{1}{\sqrt{n}}\sum_{i=2}^n \big|g(X_{i:n}+\delta_{[i]}) - g(X_{i-1:n}+\delta_{[i]})\big|. \end{aligned} \tag{11}$$

Since the transfer function $h(x)$ has a bounded derivative on the interval $[a, b]$, the function $g(x)$ is Lipschitz continuous on the entire real line, that is, $|g(x) - g(y)| \leq \|h'\||x - y|$ for all $x, y \in \mathbf{R}$. Continuing with bound (11), we have

$$\begin{aligned} B_n &\geq \frac{1}{\sqrt{n}} \sum_{i=2}^{n} \left| g(X_{i-1:n} + \delta_{[i]}) - g(X_{i-1:n} + \delta_{[i-1]}) \right| - \frac{\|h'\|(b-a)}{\sqrt{n}} \\ &= \frac{1}{\sqrt{n}} \sum_{i=2}^{n} \left| g(X_{i-1} + \delta_i) - g(X_{i-1} + \delta_{i-1}) \right| - \frac{\|h'\|(b-a)}{\sqrt{n}} \\ &= \frac{n-1}{\sqrt{n}} \mathbf{E}\left[\left| g(X + \delta_2) - g(X + \delta_1) \right| \right] - \frac{\|h'\|(b-a)}{\sqrt{n}} + O_{\mathbf{P}}(1), \end{aligned} \tag{12}$$

because (i) the inputs X_i and the risks δ_i are independent, (ii) the inputs X_i have the same cdf F, and (iii) the risks δ_i have the same cdf F_δ. Hence, if the expectation on the right-hand side of bound (12) does not vanish, then we must have $B_n \xrightarrow{\mathbf{P}} \infty$ when $n \to \infty$. The proof of the converse (i.e., if $B_n \xrightarrow{\mathbf{P}} \infty$, then condition (13) is satisfied) follows from the same arguments but now with "+" instead of "−" and the reversed inequalities in bounds (11) and (12).

We are left to show that the statement

$$\mathbf{E}\left[\left| g(X + \delta_2) - g(X + \delta_1) \right| \right] > 0 \tag{13}$$

holds if and only if condition (10) is satisfied. We do so with the help of the equation

$$\mathbf{E}\left[\left| g(X + \delta_2) - g(X + \delta_1) \right| \right] = 2 \int_0^1 (2t - 1) \left(\int_{\mathbf{R}} F^{-1}_{g(x+\delta)}(t) \mathrm{d}F(x) \right) \mathrm{d}t, \tag{14}$$

which trivially follows from

$$\begin{aligned} \mathbf{E}\left[\left| g(X + \delta_2) - g(X + \delta_1) \right| \right] &= \mathbf{E}\left[\mathrm{GMD}(X) \right] \\ &= \int_0^1 \mathrm{GMD}(x_s) \mathrm{d}s, \end{aligned}$$

where $\mathrm{GMD}(x)$ is the Gini mean difference of the variable $g(x + \delta)$, defined by

$$\begin{aligned} \mathrm{GMD}(x) :=& \mathbf{E}\left[\left| g(x + \delta_2) - g(x + \delta_1) \right| \right] \\ =& 2 \int_0^1 (2t - 1) F^{-1}_{g(x+\delta)}(t) \mathrm{d}t. \end{aligned}$$

The right-most equation holds due to the well-known representation of the Gini mean difference as a Choquet integral (e.g., Giorgi 1993; Yitzhaki and Schechtman 2013; Furman et al. 2017; and references therein). We conclude with the note that the Gini mean difference is known to be (strictly) positive whenever the underlying random variable is non-degenerate, which in our case is $g(x + \delta)$. Hence, by assuming non-degeneracy of $g(x + \delta)$ for every $x \in T \subseteq [a, b]$ such that $\mathbf{P}[X \in T] > 0$, we arrive at condition (13) and thus, in turn, at (10). The proof of Theorem 3 is finished. □

6. Concluding Notes

The need for an algorithm that distinguishes between the "null hypothesis" $Y = g(X)$ and the "alternative" $Y = g(X + \delta)$ for exogenous background risk δ arises in many problems of economics, insurance, and finance. In the present paper, we have developed a user-friendly algorithm for distinguishing between the aforementioned two hypotheses. The algorithm is based on the asymptotic behaviour of two statistics: the pivot I_n and its supporter B_n, which are constructed using the order statistics of inputs and the corresponding concomitants of outputs. We have supplemented our

theoretical considerations with illustrative examples, graphs, and discussions, thus facilitating the use of the algorithm in practice.

As we have noted in the Introduction, practical considerations give rise to alternatives which couple X and δ not just in the additive way but possibly in a more intricate way, which we generally formulate as $Y = g(X, \delta)$. In this regard we also note that X and δ might be dependent random variables, and even multivariate ones, thus giving rise to a highly non-trivial follow-up problem.

Author Contributions: Both authors have contributed equally to the paper.

Funding: Research of the second author has been supported by the Natural Sciences and Engineering Research Council of Canada.

Acknowledgments: We are indebted to three anonymous reviewers for careful reading, constructive criticism, and numerous suggestions that resulted in a major revision of the original submission.

Conflicts of Interest: The authors declare no conflict of interest.

References

Box, George Edward Pelham, Gwilym Meirion Jenkins, Gregory C. Reinsel, and Greta M. Ljung. 2015. *Time Series Analysis: Forecasting and Control*, 5th ed. New York: Wiley.

Cárdenas, Alvaro A., Saurabh Amin, Zong-Syun Lin, Yu-Lun Huang, Chi-Yen Huang, and Shankar Sastry. 2011. Attacks against process control systems: Risk assessment, detection, and response. Paper presented at the 6th ACM Symposium on Information, Computer and Communications Security, Hong Kong, March 22–24; pp. 355–66.

Chen, Lingzhi, Youri Davydov, Nadezhda Gribkova, and Ričardas Zitikis. 2018. Estimating the index of increase via balancing deterministic and random data. *Mathematical Methods of Statistics* 27: 83–102. [CrossRef]

David, Herbert A., and Haikady N. Nagaraja. 2003. *Order Statistics*, 3rd ed. New York: Wiley.

Davydov, Youri, and Ričardas Zitikis. 2017. Quantifying non-monotonicity of functions and the lack of positivity in signed measures. *Modern Stochastics: Theory and Applications* 4: 219–31. [CrossRef]

Debar, Hervé, Marc Dacier, and Andreas Wespi. 1999. Towards a taxonomy of intrusion-detection systems. *Computer Networks* 31: 805–22. [CrossRef]

Finkelshtain, Israel, Offer Kella, and Marco Scarsini. 1999. On risk aversion with two risks. *Journal of Mathematical Economics* 31: 239–50. [CrossRef]

Franke, Günter, Harris Schlesinger, and Richard C. Stapleton. 2006. Multiplicative background risk. *Management Science* 52: 146–53. [CrossRef]

Franke, Günter, Harris Schlesinger, and Richard C. Stapleton. 2011. Risk taking with additive and multiplicative background risks. *Journal of Economic Theory* 146: 1547–68. [CrossRef]

Furman, Edward, Ruodu Wang, and Ričardas Zitikis. 2017. Gini-type measures of risk and variability: Gini shortfall, capital allocations, and heavy-tailed risks. *Journal of Banking and Finance* 83: 70–84. [CrossRef]

Furman, Edward, Alexey Kuznetsov, and Ričardas Zitikis. 2018. Weighted risk capital allocations in the presence of systematic risk. *Insurance: Mathematics and Economics* 79: 75–81.

Giorgi, Giovanni. 1993. A fresh look at the topical interest of the Gini concentration ratio. *Metron* 51: 83–98.

Gribkova, Nadezhda, and Ričardas Zitikis. 2018. Assessing transfer functions in control systems. *arXiv* arXiv:1805.10633.

Guo, Xu, Andreas Wagener, Wing-Keung Wong, and Lixing Zhu. 2018. The two-moment decision model with additive risks. *Risk Management* 20: 77–94. [CrossRef]

Guo, Xu, Raymond Honfu Chan, Wing-Keung Wong, and Lixing Zhu. 2018. Mean-variance, mean-VaR, and mean-CVaR models for portfolio selection with background risk. *Risk Management* doi:10.1057/s41283-018-0043-2. [CrossRef]

He, Youbiao, Gihan J. Mendis, and Jin Wei. 2017. Real-rime detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism. *IEEE Transactions on Smart Grid* 8: 2505–16. [CrossRef]

Huang, Yi, Jin Tang, Yu Cheng, Husheng Li, Kristy A. Campbell, and Zhu Han. 2016. Real-time detection of false data injection in smart grid networks: An adaptive CUSUM method and analysis. *IEEE Systems Journal* 10: 532–43. [CrossRef]

Hug, Gabriela, and Joseph Andrew Giampapa. 2012. Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks. *IEEE Transactions on Smart Grid* 3: 1362–70. [CrossRef]

Liang, Gaoqi, Junhua Zhao, Fengji Luo, Steven R. Weller, and Zhao Yang Dong. 2017. A review of false data injection attacks against modern power systems. *IEEE Transactions on Smart Grid* 8: 1630–38. [CrossRef]

Nachman, David. 1982. Preservation of "more risk averse" under expectations. *Journal of Economic Theory* 28: 361–68. [CrossRef]

Onoda, Takashi. 2016. Probabilistic models-based intrusion detection using sequence characteristics in control system communication. *Neural Computing and Applications* 27: 1119–27. [CrossRef]

Perote, Javier, and Juan Perote-Peña. 2004. Strategy-proof estimators for simple regression. *Mathematical Social Sciences* 47: 153–76. [CrossRef]

Perote, Javier, Juan Perote-Peña, and Marc Vorsatz. 2015. Strategic behavior in regressions: An experimental study. *Theory and Decision* 79: 517–46. [CrossRef]

Potluri, Sasanka, Christian Diedrich, and Girish Kumar Reddy Sangala. 2017. Identifying false data injection attacks in industrial control systems using artificial neural networks. Paper presented at the 22nd IEEE International Conference on Emerging Technologies and Factory Automation, Limassol, Cyprus, December 21; pp. 1–8.

Pratt, John W. 1998. Aversion to one risk in the presence of others. *Journal of Risk and Uncertainty* 1: 395–413. [CrossRef]

Premathilaka, Nalaka Arjuna, Achala Chathuranga Aponso, and Naomi Krishnarajah. 2013. Review on state of art intrusion detection systems designed for the cloud computing paradigm. Paper presented at 47th International Carnahan Conference on Security Technology, Medellin, Colombia, October 8–11; pp. 1–6.

Semenikhine, Vadim, Edward Furman, and Jianxi Su. 2018. On a multiplicative multivariate gamma distribution with applications in insurance. *Risks* 6: 79. [CrossRef]

Su, Jianxi. 2016. Multiple Risk Factors Dependence Structures with Applications to Actuarial Risk Management. Ph.D. dissertation, York University, Toronto, ON, Canada.

Su, Jianxi, and Edward Furman. 2017a. A form of multivariate Pareto distribution with applications to financial risk measurement. *ASTIN Bulletin* 47: 331–57. [CrossRef]

Su, Jianxi, and Edward Furman. 2017b. Multiple risk factor dependence structures: Distributional properties. *Insurance: Mathematics and Economics* 76: 56–68. [CrossRef]

Yitzhaki, Shlomo, and Edna Schechtman. 2013. *The Gini Methodology: A Primer on a Statistical Methodology*. New York: Springer.

risks

MDPI

Article

Spatial Risk Measures and Rate of Spatial Diversification

Erwan Koch

EPFL, Chair of Statistics STAT, EPFL-SB-MATH-STAT, MA B1 433 (Bâtiment MA), Station 8, 1015 Lausanne, Switzerland; erwan.koch@epfl.ch

Received: 21 December 2018; Accepted: 4 April 2019; Published: 2 May 2019

Abstract: An accurate assessment of the risk of extreme environmental events is of great importance for populations, authorities and the banking/insurance/reinsurance industry. Koch (2017) introduced a notion of spatial risk measure and a corresponding set of axioms which are well suited to analyze the risk due to events having a spatial extent, precisely such as environmental phenomena. The axiom of asymptotic spatial homogeneity is of particular interest since it allows one to quantify the rate of spatial diversification when the region under consideration becomes large. In this paper, we first investigate the general concepts of spatial risk measures and corresponding axioms further and thoroughly explain the usefulness of this theory for both actuarial science and practice. Second, in the case of a general cost field, we give sufficient conditions such that spatial risk measures associated with expectation, variance, value-at-risk as well as expected shortfall and induced by this cost field satisfy the axioms of asymptotic spatial homogeneity of order 0, -2, -1 and -1, respectively. Last but not least, in the case where the cost field is a function of a max-stable random field, we provide conditions on both the function and the max-stable field ensuring the latter properties. Max-stable random fields are relevant when assessing the risk of extreme events since they appear as a natural extension of multivariate extreme-value theory to the level of random fields. Overall, this paper improves our understanding of spatial risk measures as well as of their properties with respect to the space variable and generalizes many results obtained in Koch (2017).

Keywords: central limit theorem; insurance; max-stable random fields; rate of spatial diversification; reinsurance; risk management; risk theory; spatial dependence; spatial risk measures and corresponding axiomatic approach

1. Introduction

Hurricane Irma, which affected many Caribbean islands and parts of Florida in September 2017 caused at least 134 deaths and catastrophic damage exceeding 64.8 billion USD in value. Such an example shows the prime importance for civil authorities and for the insurance[1] industry of the accurate assessment of the risk of natural disasters, particularly as, in a climate change context, certain types of extreme events become more and more frequent (e.g., Bevere and Mueller 2014).

Motivated by the spatial feature of natural disasters, Koch (2017) introduced a new notion of spatial risk measure, which makes explicit the contribution of the space and enables one to account for at least part of the spatial dependence in the risk measurement. He also introduced a set of axioms describing how the risk is expected to evolve with respect to the space variable, at least under some conditions. These notions constitute relevant tools for risk assessment. For instance, the knowledge of the order of asymptotic spatial homogeneity allows the quantification of the rate of

[1] Throughout the paper, insurance also refers to reinsurance.

spatial diversification. Hence, they may be appealing for the banking/insurance industry. It should be highlighted that the literature about risk measures in a spatial context is very limited. To the best of our knowledge, the paper by Koch (2017) constitutes the first attempt to establish a theory about risk measures in a spatial context where the risks spread over a continuous geographical region.

In the following, the spatial risk measure associated with a classical risk measure Π and induced by a cost random field C (e.g., modelling the cost due to damage caused by a natural disaster) consists in the function of space arising from the application of Π to the normalized integral of C on various geographical areas. The contribution of this paper is threefold. First, we further explore the notions of spatial risk measure and corresponding axioms introduced in Koch (2017). Among others, we show that, for a given region, the distribution of the normalized spatially aggregated loss is entirely determined by the finite-dimensional distributions of the cost field, and propose alternative definitions of the concepts developed in Koch (2017). Additionally, we deeply explain why this whole theory about spatial risk measures is fruitful for both actuarial science and practice; e.g., we show that considering the risk related to the normalized loss does not prevent our theory from being successful for the study of the risk related to the non-normalized loss. We also point out how it can be used by insurance companies to tackle concrete issues. Second, in the case of a general cost field, we give sufficient conditions such that spatial risk measures associated with expectation, variance, value-at-risk (VaR) as well as expected shortfall (ES) and induced by this cost field satisfy the axiom of asymptotic spatial homogeneity of order 0, -2, -1 and -1, respectively. Last but not least, we focus on the case where the cost field is a function of a max-stable random field. We provide sufficient conditions on both the function and the max-stable field such that spatial risk measures associated with expectation, variance, VaR as well as ES and induced by the resulting cost field satisfy the axiom of asymptotic spatial homogeneity of order 0, -2, -1 and -1, respectively. Max-stable random fields naturally appear when one is interested in extreme events having a spatial extent since they constitute an extension of multivariate extreme-value theory to the level of random fields (in the case of stochastic processes, see, e.g., de Haan 1984; de Haan and Ferreira 2006). They are particularly well suited to model the temporal maxima of a given variable (for instance a meteorological variable) at all points in space since they arise as the pointwise maxima taken over an infinite number of appropriately rescaled independent and identically distributed random fields. On the whole, this study improves our comprehension of spatial risk measures as well as of their properties with respect to the space variable and generalizes many results by Koch (2017).

The remainder of the paper is organized as follows. In Section 2, we recall and further study the notion of spatial risk measure and the corresponding set of axioms introduced in Koch (2017). Furthermore, we thoroughly demonstrate their usefulness for both actuarial science and practice. Then, we introduce some concepts about mixing and central limit theorems for random fields. Finally, we provide some insights about max-stable random fields. Then, Section 3 presents our results relating to the properties of some spatial risk measures. Finally, Section 4 contains a short summary as well as some perspectives.

Throughout the paper, $(\Omega, \mathcal{F}, \mathbb{P})$ is an adequate probability space and $\overset{d}{=}$ and $\overset{d}{\to}$ designate equality and convergence in distribution, respectively. In the case of random fields, distribution has to be understood as the set of all finite-dimensional distributions. Finally, we denote by ν the Lebesgue measure.

2. Spatial Risk Measures and Other Concepts

2.1. Spatial Risk Measures and Corresponding Axioms

First, we describe the setting required for a proper definition of spatial risk measures. Let $\mathcal{A}$ be the set of all compact subsets of $\mathbb{R}^2$ with a positive Lebesgue measure and $\mathcal{A}_c$ be the set of all convex

elements of $\mathcal{A}$. Denote by $\mathcal{C}$ the set of all real-valued and measurable[2] random fields on $\mathbb{R}^2$ having almost surely (a.s.)[3] locally integrable sample paths. Let $\mathcal{P}$ be the family of all possible distributions of random fields belonging to $\mathcal{C}$. Each random field represents the economic or insured cost caused by the events belonging to specified classes and occurring during a given time period, say $[0, T_L]$. In the following, T_L is considered as fixed and does not appear anymore for the sake of notational parsimony. Each class of events (e.g., European windstorms or hurricanes) will be referred to as a hazard in the following. Let $\mathcal{L}$ be the set of all real-valued random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$. A risk measure typically will be some function $\Pi : \mathcal{L} \to \mathbb{R}$. This kind of risk measure will be called a classical risk measure in the following. A classical risk measure Π is termed law-invariant if, for all $\tilde{X} \in \mathcal{L}$, $\Pi(\tilde{X})$ only depends on the distribution of $\tilde{X}$.

We first remind the reader of the definition of the normalized spatially aggregated loss, which enables one to disentangle the contribution of the space and the contribution of the hazards and underpins our definition of spatial risk measure.

Definition 1 (Normalized spatially aggregated loss as a function of the distribution of the cost field)**.** *For $A \in \mathcal{A}$ and $P \in \mathcal{P}$, the normalized spatially aggregated loss is defined by*

$$L_N(A, P) = \frac{1}{\nu(A)} \int_A C_P(\mathbf{x})\, \nu(\mathrm{d}\mathbf{x}), \tag{1}$$

where the random field $\{C_P(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2}$ belongs to $\mathcal{C}$ and has distribution P.

The quantity

$$L(A, P) = \int_A C_P(\mathbf{x})\, \nu(\mathrm{d}\mathbf{x}) \tag{2}$$

corresponds to the total economic or insured loss over region A due to specified hazards. For technical reasons and to favour a more intuitive understanding, we base our definition of spatial risk measures on $L_N(A, P)$, which is the loss per surface unit and can be interpreted, in a discrete setting[4] and in an insurance context, as the mean loss per insurance policy. Among other advantages, this normalization enables a fair comparison of the risks related to regions having different sizes.

Since the field C_P is measurable, $L(A, P)$ and $L_N(A, P)$ are well-defined random variables. Moreover, they are a.s. finite as A is compact and C_P has a.s. locally integrable sample paths. The following proposition gives a sufficient condition for a random field to have a.s. locally integrable sample paths.

Proposition 1. *Let $d \geq 1$ and $\{Q(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ be a measurable random field. If the function*

$$\begin{array}{rccc} E & : & \mathbb{R}^d & \to & \mathbb{R} \\ & & \mathbf{x} & \mapsto & \mathbb{E}[|Q(\mathbf{x})|] \end{array}$$

is locally integrable, then Q has a.s. locally integrable sample paths.

Proof. Let A be a compact subset of $\mathbb{R}^d$. First, since Q is measurable, $\int_A |Q(\mathbf{x})|\, \nu(\mathrm{d}\mathbf{x})$ is a well-defined random variable. By Fubini's theorem, we have

$$\mathbb{E}\left[\int_A |Q(\mathbf{x})|\, \nu(\mathrm{d}\mathbf{x})\right] = \int_A \mathbb{E}[|Q(\mathbf{x})|]\, \nu(\mathrm{d}\mathbf{x}) < \infty,$$

---

2 Throughout the paper, when applied to random fields, the adjective "measurable" means "jointly measurable".
3 Unless otherwise stated, by a.s., we mean P-a.s.
4 See Section 2.2.

which necessarily implies that

$$\int_A |Q(\mathbf{x})|\, \nu(\mathrm{d}\mathbf{x}) < \infty \quad \text{a.s.}$$

Since this is true for all A being a compact subset of $\mathbb{R}^d$, we obtain the result. □

We now recall the notion of spatial risk measure introduced by Koch (2017), which makes explicit the contribution of the space in the risk measurement.

Definition 2 (Spatial risk measure as a function of the distribution of the cost field)**.** *A spatial risk measure is a function $\mathcal{R}_\Pi$ that assigns a real number to any region $A \in \mathcal{A}$ and distribution $P \in \mathcal{P}$:*

$$\begin{array}{rccc} \mathcal{R}_\Pi & : & \mathcal{A} \times \mathcal{P} & \to & \mathbb{R} \\ & & (A, P) & \mapsto & \Pi(L_N(A, P)), \end{array}$$

where Π is a classical and law-invariant risk measure and $L_N(A, P)$ is defined in (1)*.*

This extends the notion of classical risk measure to the spatial and infinite-dimensional setting as we now have a function of both the space and the distribution of a random field (or directly a random field, see below) instead of a function of a unique real-valued random variable. Note that law-invariance of Π is necessary for spatial risk measures to be defined in this way; see below for more details. For a given Π and a fixed $P \in \mathcal{P}$, the quantity $\mathcal{R}_\Pi(\cdot, P)$ is referred to as the spatial risk measure associated with Π and induced by P. A nice feature is that, for many useful classical risk measures Π such as, e.g., variance, VaR and ES, this notion of spatial risk measure allows one to take (at least) part of the spatial dependence structure of the field C_P into account. We could define spatial risk measures in the same way but using the non-normalized spatially aggregated loss; this is not what we do for reasons explained above and in Remark 2 below.

Now, we remind the reader of the set of axioms for spatial risk measures developed in Koch (2017). It concerns the spatial risk measures properties with respect to the space and not to the cost distribution, the latter being considered as given by the problem at hand. For any $A \in \mathcal{A}$, let $\mathbf{b}_A$ denote its barycenter.

Definition 3 (Set of axioms for spatial risk measures induced by a distribution)**.** *Let Π be a classical and law-invariant risk measure. For a fixed $P \in \mathcal{P}$, we define the following axioms for the spatial risk measure associated with Π and induced by P, $\mathcal{R}_\Pi(\cdot, P)$:*

1. *Spatial invariance under translation:*
   for all $\mathbf{v} \in \mathbb{R}^2$ and $A \in \mathcal{A}$, $\mathcal{R}_\Pi(A + \mathbf{v}, P) = \mathcal{R}_\Pi(A, P)$, where $A + \mathbf{v}$ denotes the region A translated by the vector $\mathbf{v}$.
2. *Spatial sub-additivity:*
   for all $A_1, A_2 \in \mathcal{A}$, $\mathcal{R}_\Pi(A_1 \cup A_2, P) \leq \min\{\mathcal{R}_\Pi(A_1, P), \mathcal{R}_\Pi(A_2, P)\}$.
3. *Asymptotic spatial homogeneity of order $-\gamma, \gamma \geq 0$:*
   for all $A \in \mathcal{A}_c$,

$$\mathcal{R}_\Pi(\lambda A, P) \underset{\lambda \to \infty}{=} K_1(A, P) + \frac{K_2(A, P)}{\lambda^\gamma} + o\left(\frac{1}{\lambda^\gamma}\right),$$

   where λA is the area obtained by applying to A a homothety with center $\mathbf{b}_A$ and ratio $\lambda > 0$, and $K_1(\cdot, P) : \mathcal{A}_c \to \mathbb{R}$, $K_2(\cdot, P) : \mathcal{A}_c \to \mathbb{R}\backslash\{0\}$ are functions depending on P.

It is also reasonable to introduce the axiom of spatial anti-monotonicity: for all $A_1, A_2 \in \mathcal{A}$, $A_1 \subset A_2 \Rightarrow \mathcal{R}_\Pi(A_2, P) \leq \mathcal{R}_\Pi(A_1, P)$. The latter is equivalent to the axiom of spatial sub-additivity. These axioms appear natural and make sense at least under some conditions on the cost field C_P

(e.g., stationarity[5] in the case of spatial invariance under translation and spatial sub-additivity) and for some classical risk measures Π. The axiom of spatial sub-additivity indicates spatial diversification. If it is satisfied with strict inequality, an insurance company would be well advised to underwrite policies in both regions A_1 and A_2 instead of only one of them. This axiom involves the minimum operator because the concept of spatial risk measure is based on the normalized spatially aggregated loss; using the summation operator instead would not provide information about spatial diversification. On the other hand, if spatial risk measures were defined using the non-normalized loss, then summation would make sense; see Remark 2 below for more details. Originally, Koch (2017) used the term "sub-additivity", among other reasons, by analogy with the axiom of sub-additivity by Artzner et al. (1999), which also conveys a diversification idea. The axiom of asymptotic spatial homogeneity of order $-\gamma$ quantifies the rate of spatial diversification when the region becomes large. Consequently, determining the value of γ is of interest for the insurance industry; see Section 2.2 for further details.

The axioms of spatial invariance under translation and spatial sub-additivity a priori make sense only if the cost field satisfies at least some kind of stationarity. If an insurance company covers a region A_1 which is much less risky than a region A_2, it is very unlikely that the company reduces its risk by covering $A_1 \cup A_2$. For a given hazard (e.g., hurricanes), the cost resulting from a single specific event (e.g., a particular hurricane) generally varies across space, making any particular realization of the cost field spatially inhomogeneous. Nevertheless, the cost field (and not one realization of it) related to this hazard can be stationary or, at least, piecewise stationary; see immediately below.

In concrete actuarial applications, the cost field (for a given hazard) is often non-stationary over the entire region covered by the insurance company, unless it is a very small area. In many cases, however, it can reasonably be considered as locally stationary; see, e.g., Dahlhaus (2012) for an excellent review about locally stationary processes, and Eckley et al. (2010) as well as Anderes and Stein (2011) for papers dealing with local non-stationarity in the case of random fields. Locally stationary processes can be well approximated by piecewise stationary processes (e.g., Ombao et al. 2001, Section 2.2) and, assuming this to be also true for random fields, we can reasonably consider the cost field to be stationary over sub-regions, at least in most cases. In the latter, the axioms of spatial invariance under translation and spatial sub-additivity make sense separately on each sub-region over which the field is stationary. Let S_{ub} be such a sub-region (a subset of $\mathbb{R}^2$) and $\mathcal{S}_{\text{ub}}$ be the set of all compact subsets of S_{ub} with a positive Lebesgue measure. The axiom of spatial invariance under translation becomes: for all $\mathbf{v} \in \mathbb{R}^2$ and $A \in \mathcal{S}_{\text{ub}}$ such that $A + \mathbf{v} \in \mathcal{S}_{\text{ub}}$, $\mathcal{R}_\Pi(A + \mathbf{v}, P) = \mathcal{R}_\Pi(A, P)$; spatial sub-additivity is now written: for all $A_1, A_2 \in \mathcal{S}_{\text{ub}}$, $\mathcal{R}_\Pi(A_1 \cup A_2, P) \leq \min\{\mathcal{R}_\Pi(A_1, P), \mathcal{R}_\Pi(A_2, P)\}$.

Of course, the fact that the axioms of Definition 3 are satisfied depends on both the classical risk measure Π and the cost field C_P. It may be interesting to determine for which classical risk measures the axioms are satisfied for the broadest class of cost fields. These classical risk measures could be considered as "adapted" to the spatial context.

Remark 1. *Although the concept of spatial risk measure and related axioms naturally apply in an insurance context (see Section 2.2 for further details), they can also be used in the banking industry and on financial markets. A potential application is the assessment of the risk related to event-linked securities such as catastrophe bonds. Furthermore, they can be used for a wider class of risks than those linked with damage due to environmental events. These concepts are actually insightful as soon as the risks spread over a geographical region. One might think, e.g., about the loss in value of real estate due to adverse economic conditions.*

We close this section by deeply commenting on the previous concepts and giving slightly modified and more natural versions of previous definitions. First, we need the following useful result.

[5] Throughout the paper, stationarity refers to strict stationarity.

Theorem 1. *Let $d \geq 1$ and $\{H(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ be a measurable random field having a.s. locally integrable sample paths. Moreover, let A be a compact subset of $\mathbb{R}^d$ with positive Lebesgue measure. Then the distribution of*

$$L_N(A,H) = \frac{1}{\nu(A)} \int_A H(\mathbf{x})\, \nu(\mathrm{d}\mathbf{x})$$

only depends on A and the finite-dimensional distributions of H.

Proof. The proof is partly inspired from the proof of Theorem 11.4.1 in Samorodnitsky and Taqqu (1994). We assume that the random field H is defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For a fixed $\omega \in \Omega$, we denote by H_ω the corresponding realization of H on $\mathbb{R}^d$ and by $H_\omega(\mathbf{x})$ the realization of H at location $\mathbf{x}$. By definition, we have, for almost every $\omega \in \Omega$, that

$$L_N(A,H_\omega) = \frac{1}{\nu(A)} \int_A H_\omega(\mathbf{x})\, \nu(\mathrm{d}\mathbf{x}). \tag{3}$$

Now, let $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ be a probability space different from the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathbf{U}$ be a random vector defined on $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ and following the uniform distribution on A, with density $f_{\mathbf{U}}(\mathbf{x}) = \mathbb{I}_{\{\mathbf{x}\in A\}}/\nu(A), \mathbf{x} \in \mathbb{R}^d$. From (3), it directly follows that, for almost every $\omega \in \Omega$,

$$L_N(A,H_\omega) = \int_{\mathbb{R}^d} H_\omega(\mathbf{x}) f_{\mathbf{U}}(\mathbf{x})\, \nu(\mathrm{d}\mathbf{x}). \tag{4}$$

Let us denote by $\mathbb{E}_1$ the expectation with respect to the probability measure $\mathbb{P}_1$. We have

$$\mathbb{E}_1[H_\omega(\mathbf{U})] = \int_{\mathbb{R}^d} H_\omega(\mathbf{x}) f_{\mathbf{U}}(\mathbf{x})\, \nu(\mathrm{d}\mathbf{x}),$$

giving, using (4),

$$L_N(A,H_\omega) = \mathbb{E}_1[H_\omega(\mathbf{U})].$$

Now, let $\mathbf{U}_1, \ldots, \mathbf{U}_n$ be independent replications of $\mathbf{U}$ (which are independent of the random field H). The strong law of large numbers gives that, for almost every $\omega \in \Omega$,

$$L_N(A,H_\omega) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} H_\omega(\mathbf{U}_i)\ \mathbb{P}_1\text{-a.s.} \tag{5}$$

Therefore, using Fubini's theorem, we deduce that, for $\mathbb{P}_1$-almost every $\omega_1 \in \Omega_1$,

$$L_N(A,H_\omega) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} H_\omega(\mathbf{U}_i(\omega_1))\ \mathbb{P}\text{-a.s.} \tag{6}$$

Now, we choose $\omega_0 \in \Omega_1$ such that the (non-random) sequence $(\mathbf{U}_1(\omega_0), \mathbf{U}_2(\omega_0), \ldots)$ satisfies (6). We obtain

$$L_N(A,H_\omega) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} H_\omega(\mathbf{U}_i(\omega_0))\ \mathbb{P}\text{-a.s.} \tag{7}$$

Equation (7) says that the distribution of $L_N(A,H)$ is determined by the finite-dimensional distributions at the points belonging to the set $\{\mathbf{U}_i(\omega_0) : i \in \mathbb{N}\}$. This yields the result. □

It is more natural, especially in terms of interpretation, to introduce the normalized spatially aggregated loss as a function of the cost field instead of its distribution, as shown immediately below.

Definition 4 (Normalized spatially aggregated loss as a function of the cost field)**.** *The normalized spatially aggregated loss function is defined by*

$$\begin{array}{rccc} L_N & : & \mathcal{A}\times\mathcal{C} & \to & \mathbb{R} \\ & & (A,C) & \mapsto & \dfrac{1}{\nu(A)}\displaystyle\int_A C(\mathbf{x})\,\nu(\mathrm{d}\mathbf{x}). \end{array} \tag{8}$$

Let $C_P \in \mathcal{C}$ be a random field with distribution P. Although a particular realization of $L_N(A, C_P)$ obviously depends on C_P (through its corresponding realization), we know from Theorem 1 that its distribution is entirely characterized by A and P. This explains our notation $L_N(A, P)$ instead of $L_N(A, C_P)$ in Definition 1. More precisely, let $C_P^{(1)}, C_P^{(2)} \in \mathcal{C}$ be random fields having the same distribution P. Then, $C_P^{(1)}$ and $C_P^{(2)}$ have the same finite-dimensional distributions, which implies that $L_N(A, C_P^{(1)}) \overset{d}{=} L_N(A, C_P^{(2)})$.

Similarly, it can appear more natural to define spatial risk measures as functions of the cost field instead of its distribution. Moreover, this allows spatial risk measures to be defined even when the classical risk measure Π is not law-invariant.

Definition 5 (Spatial risk measure as a function of the cost field)**.** *A spatial risk measure is a function $\mathcal{R}_\Pi$ that assigns a real number to any region $A \in \mathcal{A}$ and random field $C \in \mathcal{C}$:*

$$\begin{array}{rccc} \mathcal{R}_\Pi & : & \mathcal{A}\times\mathcal{C} & \to & \mathbb{R} \\ & & (A,C) & \mapsto & \Pi(L_N(A,C)), \end{array}$$

where Π is a classical risk measure.

For a given classical and law-invariant risk measure Π and a given region $A \in \mathcal{A}$, the value of the spatial risk measure of Definition 5 is completely determined by the distribution of $L_N(A, C)$ by law-invariance of Π. Consequently, using Theorem 1, it is completely determined by A and the distribution of the cost field C. This explains why Koch (2017) has introduced the notion of spatial risk measure as a function of the distribution of C (see the reminder in Definition 2); if Π is law-invariant, the spatial risk measures described in Definitions 2 and 5 refer to the same notion. For a given Π and a fixed $C \in \mathcal{C}$, $\mathcal{R}_\Pi(\cdot, C)$ is referred to as the spatial risk measure associated with Π and induced by C.

Of course, we can also express the axioms recalled in Definition 3 for the spatial risk measures induced by a cost field $C \in \mathcal{C}$ introduced immediately above. On top of being more natural, it enables one to leave out the assumption of law-invariance for the classical risk measure Π.

Definition 6 (Set of axioms for spatial risk measures induced by a cost field)**.** *Let Π be a classical risk measure. For a fixed $C \in \mathcal{C}$, we define the following axioms for the spatial risk measure associated with Π and induced by C, $\mathcal{R}_\Pi(\cdot, C)$:*

1. *Spatial invariance under translation:*
   for all $\mathbf{v} \in \mathbb{R}^2$ and $A \in \mathcal{A}$, $\mathcal{R}_\Pi(A + \mathbf{v}, C) = \mathcal{R}_\Pi(A, C)$, where $A + \mathbf{v}$ denotes the region A translated by the vector $\mathbf{v}$.
2. *Spatial sub-additivity:*
   for all $A_1, A_2 \in \mathcal{A}$, $\mathcal{R}_\Pi(A_1 \cup A_2, C) \leq \min\{\mathcal{R}_\Pi(A_1, C), \mathcal{R}_\Pi(A_2, C)\}$.
3. *Asymptotic spatial homogeneity of order $-\gamma, \gamma \geq 0$:*
   for all $A \in \mathcal{A}_c$,

$$\mathcal{R}_\Pi(\lambda A, C) \underset{\lambda\to\infty}{=} K_1(A, C) + \frac{K_2(A, C)}{\lambda^\gamma} + o\left(\frac{1}{\lambda^\gamma}\right),$$

   where λA is the area obtained by applying to A a homothety with center $\mathbf{b}_A$ and ratio $\lambda > 0$, and $K_1(\cdot, C) : \mathcal{A}_c \to \mathbb{R}$, $K_2(\cdot, C) : \mathcal{A}_c \to \mathbb{R}\backslash\{0\}$ are functions depending on C.

All the explanations and interpretations given for Definitions 1–3 remain valid in the case of Definitions 4–6. For the reasons mentioned above, our opinion is that Definitions 4–6 rather than previous ones should be used. This is what is done in the following.

2.2. Concrete Applications to Insurance

This section is dedicated to the connections between the concepts described above and actuarial risk theory as well as real insurance practice. We especially show how they can be used for concrete purposes. In an insurance context, the quantity

$$L(A, C) = \int_A C(\mathbf{x})\, \nu(\mathrm{d}\mathbf{x}) \tag{9}$$

appearing in Definition 4 (or equivalently in (2)) can be seen as a continuous and more complex version of the classical actuarial individual risk model. The latter can be formulated as

$$L_{\text{ind}} = \sum_{i=1}^{N} X_i, \tag{10}$$

where L_{ind} is the total loss, N denotes the number of insurance policies and, for $i = 1, \ldots, N$, X_i is the claim related to the i-th policy. The X_i's are generally assumed to be independent but not necessarily identically distributed. In $L(A, C)$, each location $\mathbf{x}$ corresponds to a specific insurance policy and thus each $C(\mathbf{x})$ is equivalent to a X_i in (10). By the way, by choosing ν to be a counting measure instead of the Lebesgue measure, the integral in (9) can be reduced to a sum, e.g., $\sum_{\mathbf{x}\in A'} C(\mathbf{x})$, where A' is a finite set of locations in $\mathbb{R}^2$ (e.g., part of a lattice in $\mathbb{Z}^2$). It is worth mentioning that the ideas of this paper can easily be applied to such a framework.

Even if dependence between the X_i, $i = 1, \ldots, N$, in (10) was allowed, considering $L(A, C)$ (see (9)) would appear more promising. Indeed, the geographical information of each risk (i.e., insurance policy) is explicitly taken into account and, consequently, the dependence between all risks can be modelled in a more realistic way than in (10). The dependence between the risks directly inherits from their respective associated geographical positions and, thus, ignoring their localizations as in (10) makes the modelling of their dependence more arbitrary and likely less reliable. In our approach, this dependence is fully characterized by the spatial dependence structure of the cost field C. Potential central limit theorems (see below) would have stronger implications because the dependence is more realistic. For these reasons, Models (8) and (9) allow a more accurate assessment of spatial diversification. The same remarks hold if we compare our loss models with the classical actuarial collective risk model.

Our risk models (8) and (9) and more generally our theory about spatial risk measures may be particularly relevant for an insurance company willing to adapt its policies portfolio. For example, the axioms of spatial sub-additivity and asymptotic spatial homogeneity can help it to assess the potential relevance of extending its activity to a new geographical region. Such an analysis requires the company to have an accurate view of the dependence between its risks (inter alia between the possible new risks and those already present in the portfolio), as allowed by Models (8) and (9) through the cost field C. Model (10) would not enable the insurer to precisely account for the dependence between the new risks and those already in the portfolio and hence to properly quantify the impact of a geographical expansion, i.e., of an increase of the number of contracts N.

At present, we show that, consistently with our intuition, considering the risk related to the normalized spatially aggregated loss is also insightful when the insurer is interested in the risk related to its non-normalized counterpart, which is often the case. Let Π be a positive homogeneous and translation invariant classical risk measure and p_r denote either the claims reserves, revenues or

any relevant related quantity[6] per surface unit (possibly the mean premium per surface unit) of an insurance company Ins.

We first consider the axiom of spatial sub-additivity, which is assumed to be satisfied. Ins covers region A_1 for a given hazard and potentially aims at covering also a region A_2 disjoint of A_1. We assume that Ins properly hedges its risk on A_1, i.e.,

$$\nu(A_1)p_r \geq \Pi(L(A_1, C)), \quad \text{i.e.,} \quad p_r \geq \Pi(L_N(A_1, C)), \tag{11}$$

by positive homogeneity. Using again the same property,

$$\Pi(L(A_1 \cup A_2, C)) = \nu(A_1 \cup A_2)\Pi(L_N(A_1 \cup A_2, C)).$$

Combined with

$$\Pi(L_N(A_1 \cup A_2, C)) \leq \Pi(L_N(A_1, C)),$$

this yields

$$\Pi(L(A_1 \cup A_2, C)) \leq \frac{\nu(A_1 \cup A_2)}{\nu(A_1)}\Pi(L(A_1, C)).$$

Hence, by translation invariance,

$$\begin{aligned}\Pi(L(A_1 \cup A_2, C) - \nu(A_1 \cup A_2)p_r) &= \Pi(L(A_1 \cup A_2, C)) - \nu(A_1 \cup A_2)p_r \\ &\leq \frac{\nu(A_1 \cup A_2)}{\nu(A_1)}\Pi(L(A_1, C)) - \nu(A_1 \cup A_2)p_r. \end{aligned} \tag{12}$$

It follows from (11) that

$$p_r[\nu(A_1 \cup A_2) - \nu(A_1)] \geq \frac{\Pi(L(A_1, C))}{\nu(A_1)}[\nu(A_1 \cup A_2) - \nu(A_1)],$$

which gives

$$\frac{\nu(A_1 \cup A_2)}{\nu(A_1)}\Pi(L(A_1, C)) - \nu(A_1 \cup A_2)p_r \leq \Pi(L(A_1, C)) - \nu(A_1)p_r. \tag{13}$$

The combination of (12) and (13) yields that

$$\Pi(L(A_1 \cup A_2, C) - \nu(A_1 \cup A_2)p_r) \leq \Pi(L(A_1, C) - \nu(A_1)p_r).$$

The last inequality is strict if that in the axiom of spatial sub-additivity or in (11) is so. Thus, if Ins suitably hedges its risk on A_1, the risk is even better hedged on $A_1 \cup A_2$. Exactly the same reasoning holds for A_2.

Remark 2. *For spatial risk measures defined using the non-normalized spatially aggregated loss, we could propose the following axiom of spatial sub-additivity: for all disjoint $A_1, A_2 \in \mathcal{A}$, $\Pi(L(A_1 \cup A_2, C)) \leq \Pi(L(A_1, C)) + \Pi(L(A_2, C))$. Nevertheless, this property is trivially satisfied as soon as the classical risk measure Π is sub-additive and therefore its validity does not depend on the properties of the cost field C. Basing the axiom of spatial sub-additivity on the normalized spatially aggregated loss as we did is more appealing since it allows a diversification effect coming from C (and not only from Π). This argument is in favour of defining spatial risk measures using the normalized spatially aggregated loss.*

[6] It is out of the scope of this paper to enter into accounting details.

We now consider the axiom of asymptotic spatial homogeneity of order $-\gamma$, $\gamma \geq 0$. Assume that it is satisfied with $\gamma > 0$ (e.g., we will see that for Π being VaR or ES, γ typically equals 1). It follows from Definition 6, Point 3, that

$$\begin{aligned}\Pi(L(\lambda A, C) - \nu(\lambda A)p_r) &\underset{\lambda\to\infty}{=} \lambda^2\nu(A)K_1(A,C) + \nu(A)K_2(A,C)\lambda^{2-\gamma} + o\left(\lambda^{2-\gamma}\right) - \lambda^2\nu(A)p_r \\ &\underset{\lambda\to\infty}{=} \lambda^2\nu(A)(K_1(A,C) - p_r) + \nu(A)K_2(A,C)\lambda^{2-\gamma} + o\left(\lambda^{2-\gamma}\right).\end{aligned}$$

Since $\gamma > 0$, the dominant term as $\lambda \to \infty$ is $\lambda^2\nu(A)(K_1(A,C) - p_r)$. Assume that $K_1(A,C) > 0$ and $K_2(A,C) > 0$. This is true under the conditions of Section 3 for VaR and ES: we have $K_1(A,C) = \mathbb{E}[C(\mathbf{0})]$, which is positive as the cost field can be assumed to be non-negative and not a.s. equal to 0; regarding $K_2(A,C)$, this is always true for ES and, provided that the confidence level α is greater than $1/2$, also for VaR. Consequently, for λ large enough, the total risk of the company, $\Pi(L(\lambda A, C) - \nu(\lambda A)p_r)$, is a decreasing function of λ as soon as the revenue per surface unit (or claims reserves, ...) satisfies $p_r > K_1(A,C)$. Under the conditions of Section 3, for VaR and ES, $K_1(A,C) = \mathbb{E}[C(\mathbf{x})]$ for all $\mathbf{x} \in \mathbb{R}^2$, and therefore the latter inequality entails that the revenue per surface unit (e.g., the mean premium) exceeds the expected cost at each location, which appears natural. The term $2-\gamma$ corresponds to the second highest power with respect to λ. Provided that $K_2(A,C) > 0$ and $0 < \gamma < 2$ (which is true for VaR and ES under the conditions of Section 3), the corresponding term, $\nu(A)K_2(A,C)\lambda^{2-\gamma}$, increases the total risk of the company as λ increases. However, the highest the value of γ, the fastest the decrease of the total risk as λ increases owing to the term in λ^2. For λ large, the values of γ, $K_1(A,C)$, $K_2(A,C)$ and p_r allow one to determine the value of λ necessary to reach a targeted sufficiently low level of the total risk. Note that in the case of the variance, at least under the conditions of Section 3, $K_1(A,C) = 0$ and $\gamma = 2$.

Remark 3. *The axioms of spatial invariance under translation and asymptotic spatial homogeneity could also be defined for spatial risk measures based on the non-normalized spatially aggregated loss. Spatial invariance under translation would be unchanged and asymptotic spatial homogeneity of order $-\gamma$, $\gamma \geq 0$, would become: for all $A \in \mathcal{A}_c$,*

$$\Pi(L(\lambda A, C)) \underset{\lambda\to\infty}{=} \lambda^2\nu(A)K_1(A,C) + \nu(A)K_2(A,C)\lambda^{2-\gamma} + o\left(\lambda^{2-\gamma}\right).$$

In this case, we would obtain the risk related to the non-normalized loss without assuming that Π is positive homogeneous.

Finally, we discuss a possible way for a company to develop an adequate model for the cost field C in regions where it is still inactive. The general model for the cost field introduced in Koch (2017, Section 2.3), is written

$$\{C(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2} = \{E(\mathbf{x})\ D\left(Z(\mathbf{x})\right)\}_{\mathbf{x}\in\mathbb{R}^2}, \tag{14}$$

where $\{E(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2}$ is the exposure field, D a damage function and $\{Z(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2}$ the random field of the environmental variable generating risk. The cost is assumed to be only due to a unique class of events, i.e., to a unique natural hazard. The latter (e.g., heat waves or hurricanes) is described by the random field of an environmental variable (e.g., the temperature or the wind speed, respectively), Z. We assume that Z is representative of the risk during the whole period $[0, T_L]$. The application of the damage function (also referred to as vulnerability curve in the literature) D to the natural hazard random field gives the destruction percentage at each location. Finally, multiplying the destruction percentage by the exposure gives the cost at each location. For more details, we refer the reader to Koch (2017, Section 2.3). In order to obtain an adequate model C in regions where it has no policies yet, the company can for instance consider crude estimates of the exposure field in the new region, develop

a detailed statistical model[7] for the environmental field Z responsible of the risk insured (e.g., wind speed in the case of hurricanes) using appropriate data and apply the same damage functions as in the region it already covers. The company can then simulate from this cost model, hence obtaining an empirical distribution of the loss appearing in (8) and (9). This makes it possible to check whether the axiom of spatial sub-additivity is satisfied or not. Furthermore, if the spatial domain is large (which is generally the case for reinsurance companies), considering potential central limit theorems and determining the order of asymptotic spatial homogeneity (by checking if the conditions of Section 3 are satisfied) is useful as it allows the company to quantify the rate of spatial diversification.

Remark 4. *Strictly speaking, the terms of the insurance policies should be accounted for in Model* (14)*. By the way, the latter model can be interpreted differently from what is done here. For instance, we can imagine that Z represents the random field of the real cost and D accounts for the terms of the policies.*

2.3. Mixing and Central Limit Theorems for Random Fields

We first remind the reader of the definition of the α- and β-mixing coefficients which will be used in Section 3. Let $\{X(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ be a real-valued random field. For $S \subset \mathbb{R}^d$ a closed subset, we denote by $\mathcal{F}_S^X$ the σ-field generated by the random variables $\{X(\mathbf{x}) : \mathbf{x} \in S\}$. Let $S_1, S_2 \subset \mathbb{R}^d$ be disjoint closed subsets. The α-mixing coefficient (introduced by Rosenblatt 1956) between the σ-fields $\mathcal{F}_{S_1}^X$ and $\mathcal{F}_{S_2}^X$ is defined by

$$\alpha^X(S_1, S_2) = \sup\left\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{F}_{S_1}^X, B \in \mathcal{F}_{S_2}^X\right\}. \tag{15}$$

The β-mixing coefficient or absolute regularity coefficient (attributed to Kolmogorov in Volkonskii and Rozanov 1959) between the σ-fields $\mathcal{F}_{S_1}^X$ and $\mathcal{F}_{S_2}^X$ is given by

$$\beta^X(S_1, S_2) = \frac{1}{2}\sup\left\{\sum_{i=1}^{I}\sum_{j=1}^{J}|\mathbb{P}(A_i \cap B_j) - \mathbb{P}(A_i)\mathbb{P}(B_j)|\right\},$$

where the supremum is taken over all partitions $\{A_1, \ldots, A_I\}$ and $\{B_1, \ldots, B_J\}$ of Ω with the A_i's in $\mathcal{F}_{S_1}^X$ and the B_j's in $\mathcal{F}_{S_2}^X$. These coefficients satisfy the useful inequality

$$\alpha^X(S_1, S_2) \leq \frac{1}{2}\beta^X(S_1, S_2), \quad \text{for all } S_1, S_2 \subset \mathbb{R}^d. \tag{16}$$

Now, we recall the concepts of Van Hove sequence and central limit theorem (CLT) in the case of random fields. This will be useful, since, for instance, asymptotic spatial homogeneity of order -1 of spatial risk measures associated with VaR (at a confidence level $\alpha \in (0,1)\backslash\{1/2\}$) and induced by a cost field $C \in \mathcal{C}$ is satisfied as soon as C fulfills the CLT and has a constant expectation (see below). For $V \subset \mathbb{R}^d$ and $r > 0$, we introduce $V^{+r} = \{\mathbf{x} \in \mathbb{R}^d : \text{dist}(\mathbf{x}, V) \leq r\}$, where dist stands for the Euclidean distance. Additionally, we denote by ∂V the boundary of V. A Van Hove sequence in $\mathbb{R}^d$ is a sequence $(V_n)_{n\in\mathbb{N}}$ of bounded measurable subsets of $\mathbb{R}^d$ satisfying $V_n \uparrow \mathbb{R}^d$, $\lim_{n\to\infty}\nu(V_n) = \infty$, and $\lim_{n\to\infty}\nu((\partial V_n)^{+r})/\nu(V_n) = 0$ for all $r > 0$. The assumption "bounded" does not always appear in the definition of a Van Hove sequence. Let Cov denote the covariance. In the following, we say that a random field $\{X(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ such that, for all $\mathbf{x} \in \mathbb{R}^d$, $\mathbb{E}\left[[X(\mathbf{x})]^2\right] < \infty$, satisfies the CLT if

$$\int_{\mathbb{R}^d}|\text{Cov}(X(\mathbf{0}), X(\mathbf{x}))|\,\nu(\mathrm{d}\mathbf{x}) < \infty,$$

$$\sigma_X = \left(\int_{\mathbb{R}^d}\text{Cov}(X(\mathbf{0}), X(\mathbf{x}))\,\nu(\mathrm{d}\mathbf{x})\right)^{\frac{1}{2}} > 0,$$

[7] Potentially different from those developed in the natural catastrophes industry: e.g., a max-stable model.

and, for any Van Hove sequence $(V_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^d$,

$$\frac{1}{[\nu(V_n)]^{1/2}}\int_{V_n}(X(\mathbf{x})-\mathbb{E}[X(\mathbf{x})])\,\nu(\mathrm{d}\mathbf{x})\xrightarrow{d}\mathcal{N}(0,\sigma_X^2),\quad \text{as } n\to\infty,$$

where $\mathcal{N}(\mu,\sigma^2)$ denotes the normal distribution with expectation $\mu\in\mathbb{R}$ and variance $\sigma^2>0$. In the case of a random field satisfying the CLT, we have the following result.

Theorem 2. *Let $\{C(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2}\in\mathcal{C}$. Assume moreover that C has a constant expectation (i.e., for all $\mathbf{x}\in\mathbb{R}^2$, $\mathbb{E}[C(\mathbf{x})]=\mathbb{E}[C(\mathbf{0})]$) and satisfies the CLT. Then, we have, for all $A\in\mathcal{A}_c$, that*

$$\lambda\left(L_N(\lambda A,C)-\mathbb{E}[C(\mathbf{0})]\right)\xrightarrow{d}\mathcal{N}\left(0,\frac{\sigma_C^2}{\nu(A)}\right),\quad \textit{for } \lambda\to\infty.$$

Proof. The result is essentially based on part of the proof of Theorem 4 in Koch (2017). We refer the reader to this proof for details and only provide the main ideas here. First, we show (see the third paragraph of the proof of Theorem 4 in Koch 2017) that, for any $A\in\mathcal{A}_c$ and any positive non-decreasing sequence $(\lambda_n)_{n\in\mathbb{N}}\in\mathbb{R}$ such that $\lim_{n\to\infty}\lambda_n=\infty$, the sequence $(\lambda_n A)_{n\in\mathbb{N}}$ is a Van Hove sequence. Therefore, since C satisfies the CLT and has a constant expectation, we obtain

$$\lambda_n\left(L_N(\lambda_n A,C)-\mathbb{E}[C(\mathbf{0})]\right)\xrightarrow{d}\mathcal{N}\left(0,\frac{\sigma_C^2}{\nu(A)}\right),\quad \text{for } n\to\infty.$$

Second, we deduce (see the proof of Theorem 4, after (44), in Koch 2017) that, for all $A\in\mathcal{A}_c$,

$$\lambda\left(L_N(\lambda A,C)-\mathbb{E}[C(\mathbf{0})]\right)\xrightarrow{d}\mathcal{N}\left(0,\frac{\sigma_C^2}{\nu(A)}\right),\quad \text{for } \lambda\to\infty.$$

This concludes the proof. □

This theorem will be useful in the following since it will allow us to prove asymptotic spatial homogeneity of order respectively -2, -1 and -1 for spatial risk measures associated with variance, VaR as well as ES and induced by a cost field satisfying the CLT and additional conditions. Moreover, if λ is large enough, it gives an approximation of the distribution of the normalized spatially aggregated loss:

$$L_N(\lambda A,C)\approx\mathcal{N}\left(\mathbb{E}[C(\mathbf{0})],\frac{\sigma_C^2}{\lambda^2\nu(A)}\right),$$

where $\approx$ means "approximately follows". Such an approximation can be fruitful in practice, e.g., for an insurance company.

2.4. Max-Stable Random Fields

This concise introduction to max-stable fields is partly based on Koch et al. (2018, Section 2.2). Below, "$\bigvee$" denotes the supremum when the latter is taken over a countable set. In any dimension $d\geq 1$, max-stable random fields are defined as follows.

Definition 7 (Max-stable random field)**.** *A real-valued random field $\{Z(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ is said to be max-stable if there exist sequences of functions $(a_T(\mathbf{x}),\mathbf{x}\in\mathbb{R}^d)_{T\geq 1}>0$ and $(b_T(\mathbf{x}),\mathbf{x}\in\mathbb{R}^d)_{T\geq 1}\in\mathbb{R}$ such that, for all $T\geq 1$,*

$$\left\{\frac{\bigvee_{t=1}^T\{Z_t(\mathbf{x})\}-b_T(\mathbf{x})}{a_T(\mathbf{x})}\right\}_{\mathbf{x}\in\mathbb{R}^d}\overset{d}{=}\{Z(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d},$$

where the $\{Z_t(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}, t=1,\ldots,T$, are independent replications of Z.

A max-stable random field is termed simple if it has standard Fréchet margins, i.e., for all $\mathbf{x} \in \mathbb{R}^d$, $\mathbb{P}(Z(\mathbf{x}) < z) = \exp(-1/z), z > 0$.

Now, let $\{\tilde{T}_i(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}, i = 1, \ldots, n$, be independent replications of a random field $\{\tilde{T}(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$. Let $(c_n(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d)_{n\geq 1} > 0$ and $(d_n(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d)_{n\geq 1} \in \mathbb{R}$ be sequences of functions. If there exists a non-degenerate random field $\{G(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ such that

$$\left\{ \frac{\bigvee_{i=1}^{n} \{\tilde{T}_i(\mathbf{x})\} - d_n(\mathbf{x})}{c_n(\mathbf{x})} \right\}_{\mathbf{x}\in\mathbb{R}^d} \xrightarrow{d} \{G(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}, \quad \text{for } n \to \infty,$$

then G is necessarily max-stable; see, e.g., de Haan (1984). This explains the relevance and significance of max-stable random fields in the modelling of spatial extremes.

Any simple max-stable random field Z can be written (see, e.g., de Haan 1984) as

$$\{Z(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d} \stackrel{d}{=} \left\{ \bigvee_{i=1}^{\infty} \{U_i Y_i(\mathbf{x})\} \right\}_{\mathbf{x}\in\mathbb{R}^d}, \tag{17}$$

where the $(U_i)_{i\geq 1}$ are the points of a Poisson point process on $(0, \infty)$ with intensity $u^{-2}\nu(\mathrm{d}u)$ and the $Y_i, i \geq 1$, are independent replications of a random field $\{Y(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ such that, for all $\mathbf{x} \in \mathbb{R}^d$, $\mathbb{E}[Y(\mathbf{x})] = 1$. The field Y is not unique and is called a spectral random field of Z. Conversely, any random field of the form (17) is a simple max-stable random field. Hence, (17) enables the building up of models for max-stable fields. We now present one of the most famous among such models, the Brown–Resnick random field, which is defined in Kabluchko et al. (2009) as a generalization of the stochastic process introduced in Brown and Resnick (1977). We recall that a random field $\{W(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ is said to have stationary increments if the distribution of the random field $\{W(\mathbf{x}+\mathbf{x}_0) - W(\mathbf{x}_0)\}_{\mathbf{x}\in\mathbb{R}^d}$ does not depend on $\mathbf{x}_0 \in \mathbb{R}^d$. Provided the increments of W have a finite second moment, the variogram of W, γ_W, is defined by

$$\gamma_W(\mathbf{x}) = \mathrm{Var}(W(\mathbf{x}) - W(\mathbf{0})), \quad \mathbf{x} \in \mathbb{R}^d,$$

where Var denotes the variance. The Brown–Resnick random field is specified as follows.

Definition 8 (Brown–Resnick random field). *Let $\{W(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ be a centred Gaussian random field with stationary increments and with variogram γ_W. Let us consider the random field Y defined by*

$$\{Y(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d} = \left\{ \exp\left(W(\mathbf{x}) - \frac{\mathrm{Var}(W(\mathbf{x}))}{2} \right) \right\}_{\mathbf{x}\in\mathbb{R}^d}.$$

Then the simple max-stable random field defined by (17) with Y is referred to as the Brown–Resnick random field associated with the variogram γ_W. In the following, we will also call this field the Brown–Resnick random field built with W.[8]

The Brown–Resnick field is stationary (Kabluchko et al. 2009, Theorem 2) and its distribution only depends on the variogram (Kabluchko et al. 2009, Proposition 11).

Now, let $(U_i, \mathbf{C}_i)_{i\geq 1}$ be the points of a Poisson point process on $(0, \infty) \times \mathbb{R}^d$ with intensity function $u^{-2}\nu(\mathrm{d}u) \times \nu(\mathrm{d}\mathbf{c})$. Independently, let $f_i, i \geq 1$, be independent replicates of some non-negative random

[8] In the following, when W is sample-continuous, what we refer to as the Brown–Resnick random field built with W is obtained by taking replications of W (see (17)) which are also sample-continuous.

function f on $\mathbb{R}^d$ satisfying $\mathbb{E}\left[\int_{\mathbb{R}^d} f(\mathbf{x})\,\nu(\mathrm{d}\mathbf{x})\right] = 1$. Then, it is known that the mixed moving maxima (M3) random field

$$\{Z(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d} = \left\{\bigvee_{i=1}^{\infty}\{U_i f_i(\mathbf{x}-\mathbf{C}_i)\}\right\}_{\mathbf{x}\in\mathbb{R}^d} \tag{18}$$

is a stationary and simple max-stable field. The so-called Smith random field introduced by Smith (1990) is a specific case of M3 random field and is defined immediately below.

Definition 9 (Smith random field)**.** *Let Z be written as in* (18) *with f being the density of a d-variate Gaussian random vector with mean $\mathbf{0}$ and covariance matrix Σ. Then, the field Z is referred to as the Smith random field with covariance matrix Σ.*

As the Brown–Resnick and Smith fields are defined using the random fields-based and M3 representations (17) and (18), respectively, it is usual in the spatial extremes literature to distinguish both models, although the Smith field with covariance matrix Σ corresponds to the Brown–Resnick field associated with the variogram $\gamma_W(\mathbf{x}) = \mathbf{x}'\Sigma^{-1}\mathbf{x}, \mathbf{x}\in\mathbb{R}^d$, where $'$ designates transposition; see, e.g., Huser and Davison (2013).

Finally, we briefly present the extremal coefficient (see, e.g., Schlather and Tawn 2003) which is a well-known measure of spatial dependence for max-stable random fields. Let $\{Z(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ be a simple max-stable random field. In the case of two locations, the extremal coefficient function θ is defined by

$$\mathbb{P}\left(Z(\mathbf{x}_1)\le u, Z(\mathbf{x}_2)\le u\right) = \exp\left(-\frac{\theta(\mathbf{x}_1,\mathbf{x}_2)}{u}\right), \quad \mathbf{x}_1,\mathbf{x}_2\in\mathbb{R}^d,$$

where $u>0$.

3. Properties of Some Induced Spatial Risk Measures

In this section, we provide sufficient conditions on the cost field such that some induced spatial risk measures satisfy the axioms presented in Definition 6. First, we consider the case of a general cost field before investigating the relevant case of a cost field being a function of a max-stable random field. In the following, for $\alpha\in(0,1)$, q_α and ϕ denote the quantile at level α and the standard Gaussian density, respectively. We recall that for a random variable $\tilde{X}$ with distribution function F, its value-at-risk at confidence level $\alpha\in(0,1)$ is written $\mathrm{VaR}_\alpha(\tilde{X}) = \inf\{x\in\mathbb{R} : F(x)\ge\alpha\}$. Moreover, provided $\mathbb{E}[|\tilde{X}|]<\infty$, its expected shortfall at confidence level $\alpha\in(0,1)$ is defined as

$$\mathrm{ES}_\alpha\left(\tilde{X}\right) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u\left(\tilde{X}\right)\,\nu(\mathrm{d}u).$$

Typical values for α are 0.95 and 0.99. It should be noted that in the actuarial literature, ES is sometimes called tail value-at-risk (see, e.g., Definition 2.4.1 in Denuit et al. 2005).

In the following, we mainly consider the spatial risk measures

$$\begin{aligned}
&\mathcal{R}_1(A,C) = \mathbb{E}[L_N(A,C)], \quad A\in\mathcal{A}, C\in\mathcal{C},\\
&\mathcal{R}_2(A,C) = \mathrm{Var}(L_N(A,C)), \quad A\in\mathcal{A}, C\in\mathcal{C},\\
&\mathcal{R}_{3,\alpha}(A,C) = \mathrm{VaR}_\alpha(L_N(A,C)), \quad A\in\mathcal{A}, C\in\mathcal{C},\\
&\mathcal{R}_{4,\alpha}(A,C) = \mathrm{ES}_\alpha(L_N(A,C)), \quad A\in\mathcal{A}, C\in\mathcal{C}.
\end{aligned}$$

As a classical risk measure, the expectation is not very satisfying since it does not provide any information about variability. Moreover, as will be seen, the associated spatial risk measures do not take into account the spatial dependence of the cost field. An advantage of variance, VaR and ES lies in the fact that their associated spatial risk measures all take into account (at least) part of this spatial dependence. Historically, the variance has been the dominating risk measure in finance, primarily

due to the huge influence of the portfolio theory of Markowitz which uses variance as a measure of risk. However, using the variance is only possible when the normalized spatially aggregated loss has a finite second moment. Moreover, since it allocates the same weight to positive and negative deviations from the expectation, variance is a good risk measure only for distributions which are approximately symmetric around the expectation. Currently, VaR is probably the most widely used risk measure in the finance/insurance industry. Nevertheless, it does not provide any information about the severity of losses which occur with a probability lower than $1-\alpha$, which is obviously a serious shortcoming. Moreover, VaR is in general not sub-additive and hence not coherent in the sense of Artzner et al. (1999). ES overcomes these two drawbacks of VaR. Pertaining to the first one, it can be seen from the fact that, if a random variable $\tilde{X}$ has a continuous distribution function, then

$$\mathrm{ES}_\alpha\left(\tilde{X}\right) = \mathbb{E}\left[\tilde{X}\middle|\tilde{X} > \mathrm{VaR}_\alpha(\tilde{X})\right].$$

Hence, the Basel Committee on Banking Supervision proposed the use of ES instead of VaR for the internal models-based approach (Basel Committee on Banking Supervision 2012, Section 3.2.1). However, contrary to VaR, ES is not elicitable (Gneiting 2011), implying that backtesting for ES is more difficult than for VaR.

3.1. General Cost Field

The next result provides sufficient conditions on the cost field C such that the induced spatial risk measure $\mathcal{R}_1(\cdot, C)$ satisfies the axioms presented in Definition 6.

Theorem 3. *Let $\{C(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2}$ be a measurable random field having a constant expectation and such that, for all $\mathbf{x} \in \mathbb{R}^2$, $\mathbb{E}[|C(\mathbf{x})|] = \mathbb{E}[|C(\mathbf{0})|] < \infty$. Then, we have, for all $A \in \mathcal{A}$, that $\mathcal{R}_1(A, C) = \mathbb{E}[C(\mathbf{0})]$. Hence, the spatial risk measure induced by C $\mathcal{R}_1(\cdot, C)$ satisfies the axioms of spatial invariance under translation and spatial sub-additivity. If, moreover, $\mathbb{E}[C(\mathbf{0})] \neq 0$, then $\mathcal{R}_1(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order 0 with $K_1(A, C) = 0$ and $K_2(A, C) = \mathbb{E}[C(\mathbf{0})]$, $A \in \mathcal{A}_c$.*

Proof. By assumption, the function $\mathbf{x} \mapsto \mathbb{E}[|C(\mathbf{x})|]$ is constant and hence obviously locally integrable. Consequently, as C is measurable, Proposition 1 gives that C has a.s. locally integrable sample paths. Using Fubini's theorem and the fact that C has a constant expectation, we have, for all $A \in \mathcal{A}$, that

$$\mathcal{R}_1(A, C) = \frac{1}{\nu(A)} \int_A \mathbb{E}[C(\mathbf{0})]\, \nu(\mathrm{d}\mathbf{x}) = \mathbb{E}[C(\mathbf{0})].$$

Thus, for all $\mathbf{v} \in \mathbb{R}^2$ and $A \in \mathcal{A}$, $\mathcal{R}_1(A + \mathbf{v}, C) = \mathcal{R}_1(A, C)$. Moreover, for all $A_1, A_2 \in \mathcal{A}$,

$$\mathcal{R}_1(A_1 \cup A_2, C) = \mathcal{R}_1(A_1, C) = \mathcal{R}_1(A_2, C) = \min\{\mathcal{R}_1(A_1, C), \mathcal{R}_1(A_2, C)\}.$$

Finally, for all $A \in \mathcal{A}_c$ and $\lambda > 0$, we have $\mathcal{R}_1(\lambda A, C) = \mathbb{E}[C(\mathbf{0})]$. As $|\mathbb{E}[C(\mathbf{0})]| \leq \mathbb{E}[|C(\mathbf{0})|] < \infty$, we have $|\mathbb{E}[C(\mathbf{0})]| \in (0, \infty)$, which concludes the proof. □

The next result is a generalization of Theorem 2 in Koch (2017) and will be useful in the following.

Theorem 4. *Let $\{C(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2} \in \mathcal{C}$ and such that, for all $\mathbf{x} \in \mathbb{R}^2$, $\mathbb{E}\left[[C(\mathbf{x})]^2\right] < \infty$. Moreover, assume that, for all $A \in \mathcal{A}$,*

$$\int_A \int_A |\mathbb{E}\left[C(\mathbf{x})C(\mathbf{y})\right]|\, \nu(\mathrm{d}\mathbf{x})\, \nu(\mathrm{d}\mathbf{y}) < \infty. \tag{19}$$

Then, for all $A \in \mathcal{A}$ and $\lambda > 0$, we have

$$\mathcal{R}_2(\lambda A, C) = \frac{1}{\lambda^4 [\nu(A)]^2} \int_{\lambda A} \int_{\lambda A} \mathrm{Cov}(C(\mathbf{x}), C(\mathbf{y}))\, \nu(\mathrm{d}\mathbf{x})\, \nu(\mathrm{d}\mathbf{y}).$$

Condition (19) *is satisfied for instance in the following cases:*

1. *For any* $A \in \mathcal{A}$,
$$\sup_{\mathbf{x}\in A}\left\{\mathbb{E}\left[[C(\mathbf{x})]^2\right]\right\} < \infty. \tag{20}$$
2. *For all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$,
$$\mathrm{Cov}(C(\mathbf{x}), C(\mathbf{y})) = \mathrm{Cov}(C(\mathbf{0}), C(\mathbf{x}-\mathbf{y})), \tag{21}$$
and
$$\int_{\mathbb{R}^2} |\mathrm{Cov}(C(\mathbf{0}), C(\mathbf{x}))|\, \nu(\mathrm{d}\mathbf{x}) < \infty. \tag{22}$$

Proof. For all $A \in \mathcal{A}$, we consider $L(A, C) = \nu(A)\, L_N(A, C)$. Thus, using Fubini's theorem and (19), we obtain

$$\begin{aligned}\mathbb{E}\left[[L(A,C)]^2\right] = \mathbb{E}\left[\left(\int_A C(\mathbf{x})\,\nu(\mathrm{d}\mathbf{x})\right)^2\right] &= \mathbb{E}\left[\int_A C(\mathbf{x})\,\nu(\mathrm{d}\mathbf{x})\ \int_A C(\mathbf{y})\,\nu(\mathrm{d}\mathbf{y})\right] \\ &= \int_A\int_A \mathbb{E}\left[C(\mathbf{x})C(\mathbf{y})\right]\,\nu(\mathrm{d}\mathbf{x})\,\nu(\mathrm{d}\mathbf{y}).\end{aligned} \tag{23}$$

Moreover, it is clear that

$$\left(\mathbb{E}\left[L(A,C)\right]\right)^2 = \int_A\int_A \mathbb{E}\left[C(\mathbf{x})\right]\mathbb{E}\left[C(\mathbf{y})\right]\,\nu(\mathrm{d}\mathbf{x})\,\nu(\mathrm{d}\mathbf{y}). \tag{24}$$

The combination of (23) and (24) gives that

$$\mathbb{E}\left[[L(A,C)]^2\right] - \left(\mathbb{E}\left[L(A,C)\right]\right)^2 = \int_A\int_A \mathrm{Cov}(C(\mathbf{x}), C(\mathbf{y}))\,\nu(\mathrm{d}\mathbf{x})\,\nu(\mathrm{d}\mathbf{y}),$$

which implies that

$$\mathcal{R}_2(A,C) = \frac{1}{[\nu(A)]^2}\int_A\int_A \mathrm{Cov}(C(\mathbf{x}), C(\mathbf{y}))\,\nu(\mathrm{d}\mathbf{x})\,\nu(\mathrm{d}\mathbf{y}).$$

The result is obtained by replacing A with λA.

We now prove the second part of the theorem, concerning (19). Let $A \in \mathcal{A}$. In the first case, we obtain, using Cauchy–Schwarz inequality and (20),

$$\begin{aligned}\int_A\int_A |\mathbb{E}\left[C(\mathbf{x})C(\mathbf{y})\right]|\,\nu(\mathrm{d}\mathbf{x})\,\nu(\mathrm{d}\mathbf{y}) &\le \int_A\int_A \left(\mathbb{E}\left[[C(\mathbf{x})]^2\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[[C(\mathbf{y})]^2\right]\right)^{\frac{1}{2}}\,\nu(\mathrm{d}\mathbf{x})\,\nu(\mathrm{d}\mathbf{y}) \\ &\le \int_A\int_A \left(\sup_{\mathbf{x}\in A}\left\{\mathbb{E}\left[[C(\mathbf{x})]^2\right]\right\}\right)^{\frac{1}{2}}\left(\sup_{\mathbf{x}\in A}\left\{\mathbb{E}\left[[C(\mathbf{y})]^2\right]\right\}\right)^{\frac{1}{2}}\,\nu(\mathrm{d}\mathbf{x})\,\nu(\mathrm{d}\mathbf{y}) \\ &< \infty.\end{aligned}$$

In the second case, it follows from (21) and (22) that

$$\begin{aligned}\int_A\int_A |\mathrm{Cov}(C(\mathbf{x}), C(\mathbf{y}))|\,\nu(\mathrm{d}\mathbf{x})\,\nu(\mathrm{d}\mathbf{y}) &= \int_A\int_A |\mathrm{Cov}(C(\mathbf{0}), C(\mathbf{x}-\mathbf{y}))|\,\nu(\mathrm{d}\mathbf{x})\,\nu(\mathrm{d}\mathbf{y}) \\ &= \int_A\left[\int_{A-\mathbf{y}} |\mathrm{Cov}(C(\mathbf{0}), C(\mathbf{z}))|\,\nu(\mathrm{d}\mathbf{z})\right]\,\nu(\mathrm{d}\mathbf{y}) \\ &\le \int_A\left[\int_{\mathbb{R}^2} |\mathrm{Cov}(C(\mathbf{0}), C(\mathbf{z}))|\,\nu(\mathrm{d}\mathbf{z})\right]\,\nu(\mathrm{d}\mathbf{y}) \\ &= \nu(A)\tilde{\sigma}_C < \infty,\end{aligned}$$

where

$$\tilde{\sigma}_C = \int_{\mathbb{R}^2} |\mathrm{Cov}(C(\mathbf{0}), C(\mathbf{x}))| \, \nu(\mathrm{d}\mathbf{x}).$$

Thus, (19) is obviously satisfied. □

We recall that for a random field $\{C(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2}$ such that, for all $\mathbf{x} \in \mathbb{R}^2$, $\mathbb{E}\left[[C(\mathbf{x})]^2\right] < \infty$, we note

$$\sigma_C = \left(\int_{\mathbb{R}^2} \mathrm{Cov}(C(\mathbf{0}), C(\mathbf{x})) \, \nu(\mathrm{d}\mathbf{x}) \right)^{\frac{1}{2}}.$$

The next theorem provides the main results of this subsection. In particular, it gives sufficient conditions on the cost field C such that the induced spatial risk measures $\mathcal{R}_2(\cdot, C)$, $\mathcal{R}_{3,\alpha}(\cdot, C)$ and $\mathcal{R}_{4,\alpha}(\cdot, C)$ satisfy the axioms of asymptotic spatial homogeneity of order -2, -1 and -1, respectively.

Theorem 5. *Let* $\{C(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2} \in \mathcal{C}$.

1. *Assume that C is stationary. Then, provided it exists, any spatial risk measure associated with a law-invariant classical risk measure Π and induced by C satisfies the axiom of spatial invariance under translation.*
2. *Assume that C is such that, for all $\mathbf{x} \in \mathbb{R}^2$,*

$$\mathbb{E}\left[[C(\mathbf{x})]^2\right] < \infty, \tag{25}$$

and satisfies (21) and (22). Then, we have, for all $A \in \mathcal{A}_c$, that

$$\mathcal{R}_2(\lambda A, C) \underset{\lambda\to\infty}{=} \frac{\sigma_C^2}{\lambda^2 \nu(A)} + o\left(\frac{1}{\lambda^2}\right). \tag{26}$$

Hence, if $\sigma_C > 0$, $\mathcal{R}_2(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -2 with $K_1(A, C) = 0$ and $K_2(A, C) = \sigma_C^2/\nu(A)$, $A \in \mathcal{A}_c$.
3. *Assume that C has a constant expectation and satisfies the CLT. Then, we have, for all $A \in \mathcal{A}_c$, that*

$$\mathcal{R}_{3,\alpha}(\lambda A, C) \underset{\lambda\to\infty}{=} \mathbb{E}[C(\mathbf{0})] + \frac{\sigma_C q_\alpha}{\lambda [\nu(A)]^{\frac{1}{2}}} + o\left(\frac{1}{\lambda}\right). \tag{27}$$

Hence, if $\alpha \in (0,1)\backslash\{1/2\}$, $\mathcal{R}_{3,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A, C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A, C) = \sigma_C q_\alpha / [\nu(A)]^{\frac{1}{2}}$, $A \in \mathcal{A}_c$.
4. *Assume that C has a constant expectation, satisfies the CLT and is such that the random variables $\lambda\left(L_N(\lambda A, C) - \mathbb{E}[C(\mathbf{0})]\right)$, $\lambda > 0$, are uniformly integrable. Then, we have, for all $A \in \mathcal{A}_c$, that*

$$\mathcal{R}_{4,\alpha}(\lambda A, C) \underset{\lambda\to\infty}{=} \mathbb{E}[C(\mathbf{0})] + \frac{\sigma_C \phi(q_\alpha)}{\lambda [\nu(A)]^{\frac{1}{2}} (1-\alpha)} + o\left(\frac{1}{\lambda}\right). \tag{28}$$

Hence, $\mathcal{R}_{4,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A, C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A, C) = \sigma_C \phi(q_\alpha) / \{[\nu(A)]^{\frac{1}{2}}(1-\alpha)\}$, $A \in \mathcal{A}_c$.

Proof. 1. Let $A \in \mathcal{A}$, $\mathbf{v} \in \mathbb{R}^2$ and Π be a law-invariant classical risk measure. Using the fact that $\nu(A + \mathbf{v}) = \nu(A)$ and a change of variable, we obtain

$$\mathcal{R}_\Pi(A + \mathbf{v}, C) = \Pi\left(\frac{1}{\nu(A+\mathbf{v})} \int_{A+\mathbf{v}} C(\mathbf{x}) \, \nu(\mathrm{d}\mathbf{x}) \right) = \Pi\left(\frac{1}{\nu(A)} \int_A C(\mathbf{y}+\mathbf{v}) \, \nu(\mathrm{d}\mathbf{y}) \right). \tag{29}$$

Due to the stationarity of C, we have, for all $\mathbf{v} \in \mathbb{R}^2$, that $\{C(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2} \stackrel{d}{=} \{C(\mathbf{x}+\mathbf{v})\}_{\mathbf{x}\in\mathbb{R}^2}$, yielding, since Π is law-invariant,

$$\Pi\left(\frac{1}{\nu(A)}\int_A C(\mathbf{y}+\mathbf{v})\,\nu(\mathrm{d}\mathbf{y})\right) = \Pi\left(\frac{1}{\nu(A)}\int_A C(\mathbf{x})\,\nu(\mathrm{d}\mathbf{x})\right) = \mathcal{R}_\Pi(A,C). \tag{30}$$

The combination of (29) and (30) provides the result.

2. As (21) and (22) are satisfied, we know from Theorem 4 that, for all $A \in \mathcal{A}_c$ and $\lambda > 0$, $\mathcal{R}_2(\lambda A, C)$ is well-defined. The result follows from an adapted version of the proof of Theorem 3, Point 3, in Koch (2017). We refer the reader to that proof for the technical parts. We only highlight some of the main steps as well as the main differences here.

The first part consists in showing that

$$\lim_{\lambda\to\infty} \lambda^2 \nu(A)\mathcal{R}_2(\lambda A, C) = \sigma_C^2. \tag{31}$$

Using Theorem 4 and (21), it follows that

$$\mathcal{R}_2(\lambda A, C) = \frac{1}{\lambda^4[\nu(A)]^2}\int_{\lambda A}\int_{\lambda A} \mathrm{Cov}(C(\mathbf{0}), C(\mathbf{x}-\mathbf{y}))\,\nu(\mathrm{d}\mathbf{x})\,\nu(\mathrm{d}\mathbf{y}).$$

Let $A_\lambda = \lambda A, \lambda > 0$. Then, we consider the quantity

$$T_\lambda = \frac{1}{\lambda^2\nu(A)}\int_{A_\lambda}\int_{A_\lambda} k(\mathbf{x}-\mathbf{y})\;\nu(\mathrm{d}\mathbf{x})\,\nu(\mathrm{d}\mathbf{y}), \quad \lambda > 0,$$

where

$$k(\mathbf{x}) = \mathrm{Cov}(C(\mathbf{0}), C(\mathbf{x})), \quad \mathbf{x} \in \mathbb{R}^2.$$

The next step consists in showing that

$$\lim_{\lambda\to\infty} T_\lambda = \sigma_C^2. \tag{32}$$

For this purpose, we proceed similarly as in Koch (2017), proof of Theorem 3, Point 3. The only difference consists in the fact that here k is not necessarily non-negative. Hence, in order to bound $|T_{1,\lambda}|$ and $|T_{3,\lambda}|$ (quantities defined in Koch 2017) from above, k must be replaced with its absolute value in the corresponding integrals. This is where Condition (22) plays a role. Finally, since, for all $\lambda > 0$,

$$T_\lambda = \lambda^2\nu(A)\mathcal{R}_2(\lambda A, C),$$

(31) follows from (32).

In a second part, we easily derive (26) from (31). Now, as a compact subset of $\mathbb{R}^2$, A is bounded, giving that $\nu(A) \in (0,\infty)$. Since, moreover, $\sigma_C^2 \in (0,\infty)$, $\sigma_C^2/\nu(A) \in (0,\infty)$. Hence, the second part of the result follows from (26).

3. Theorem 2 gives that, for all $A \in \mathcal{A}_c$,

$$\lambda\left(L_N(\lambda A, C) - \mathbb{E}[C(\mathbf{0})]\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma_C^2}{\nu(A)}\right), \quad \text{for } \lambda\to\infty.$$

Hence, the fact that the quantile function of a normal random variable is continuous on $(0,1)$, Proposition 0.1 in Resnick (1987) and easy computations (see Koch 2017, proof of Theorem 5) yield (27). Since C satisfies the CLT, we have $\mathbb{E}\left[[C(\mathbf{0})]^2\right] < \infty$ and thus $\mathbb{E}\left[C(\mathbf{0})\right] < \infty$. Additionally, as $\alpha \neq 1/2$, we have $q_\alpha \neq 0$. Moreover, as $\sigma_C > 0$ (because C satisfies the CLT) and $\nu(A) > 0$, we obtain that $\sigma_C q_\alpha/[\nu(A)]^{\frac{1}{2}} \neq 0$. Finally, since $\alpha \notin \{0,1\}$, $|q_\alpha| < \infty$. Furthermore, $\sigma_C < \infty$ and $\nu(A) < \infty$, giving that $|\sigma_C q_\alpha/[\nu(A)]^{\frac{1}{2}}| < \infty$. The result follows by definition.

4. Since C satisfies the CLT, we have, for all $\mathbf{x} \in \mathbb{R}^2$, $\mathbb{E}\left[[C(\mathbf{x})]^2\right] < \infty$, which implies that, for all $\mathbf{x} \in \mathbb{R}^2$, $\mathbb{E}\left[|C(\mathbf{x})|\right] < \infty$. We easily deduce, using Fubini's theorem, that $\mathbb{E}[|L_N(\lambda A, C)|]$ is finite, and, therefore, that $\mathcal{R}_{4,\alpha}(\lambda A, C)$ is well-defined for all $A \in \mathcal{A}_c$ and $\lambda > 0$. Theorem 2 gives that, for all $A \in \mathcal{A}_c$,

$$\lambda \left(L_N(\lambda A, C) - \mathbb{E}[C(\mathbf{0})]\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma_C^2}{\nu(A)}\right), \quad \text{for } \lambda \to \infty.$$

Now, ES is known to be continuous with respect to convergence in distribution in the case of uniformly integrable random variables. For details, see, e.g., Wang et al. (2018), Theorem 3.2 and Example 2.2, Point (ii); the authors' results concern bounded random variables but the mentioned result can be extended to the case of integrable random variables. Hence, it follows from the fact that the random variables $\lambda \left(L_N(\lambda, C) - \mathbb{E}[C(\mathbf{0})]\right)$, $\lambda > 0$, are uniformly integrable, and the expression of ES_α for the Gaussian distribution, that

$$\lim_{\lambda \to \infty} \frac{1}{1-\alpha} \int_\alpha^1 \mathrm{VaR}_u(\lambda[L_N(\lambda A, C) - \mathbb{E}[C(\mathbf{0})]])\, \nu(\mathrm{d}u) = \frac{\sigma_C \phi(q_\alpha)}{[\nu(A)]^{\frac{1}{2}}(1-\alpha)}. \tag{33}$$

Moreover, we have

$$\begin{aligned}\frac{1}{1-\alpha} \int_\alpha^1 \mathrm{VaR}_u(\lambda[L_N(\lambda A, C) - \mathbb{E}[C(\mathbf{0})]])\, \nu(\mathrm{d}u) &= \frac{1}{1-\alpha} \int_\alpha^1 \lambda \left(\mathrm{VaR}_u(L_N(\lambda A, C)) - \mathbb{E}[C(\mathbf{0})]\right)\, \nu(\mathrm{d}u) \\ &= \lambda \left(\mathcal{R}_{4,\alpha}(\lambda A, C) - \mathbb{E}[C(\mathbf{0})]\right).\end{aligned}$$

Thus, (33) gives, for all $A \in \mathcal{A}_c$,

$$\lambda \left(\mathcal{R}_{4,\alpha}(\lambda A, C) - \mathbb{E}[C(\mathbf{0})]\right) \underset{\lambda \to \infty}{=} \frac{\sigma_C \phi(q_\alpha)}{[\nu(A)]^{\frac{1}{2}}(1-\alpha)} + o(1),$$

which yields (28). Now, we have $\mathbb{E}\left[C(\mathbf{0})\right] < \infty$. Moreover, using the fact that, for all $\alpha \in (0,1)$, $\phi(q_\alpha) \in (0, \infty)$, and arguments stated at the end of the proof of Point 3, we obtain $|\sigma_C \phi(q_\alpha) / \{[\nu(A)]^{\frac{1}{2}}(1-\alpha)\}| \in (0, \infty)$. Consequently, the result follows by definition. □

Remark 5. *In order to establish Points 3 and 4, we take advantage of the fact that both VaR and ES are continuous with respect to convergence in distribution under appropriate assumptions. Hence, similar results may hold for other classical risk measures satisfying continuity with respect to convergence in distribution.*

Theorem 5 entails the following important result.

Corollary 1. *Let $\{C(\mathbf{x})\}_{\mathbf{x} \in \mathbb{R}^2} \in \mathcal{C}$. Moreover, assume that C satisfies (21) and the CLT. Then, we have, for all $A \in \mathcal{A}_c$, that*

$$\mathcal{R}_2(\lambda A, C) \underset{\lambda \to \infty}{=} \frac{\sigma_C^2}{\lambda^2 \nu(A)} + o\left(\frac{1}{\lambda^2}\right). \tag{34}$$

Hence, $\mathcal{R}_2(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -2 with $K_1(A, C) = 0$ and $K_2(A, C) = \sigma_C^2 / \nu(A)$, $A \in \mathcal{A}_c$.

Proof. Since C satisfies the CLT, it satisfies (22), (25) and $\sigma_C > 0$. Thus, the result follows from Theorem 5, Point 2. □

The next result provides a convenient condition ensuring the uniform integrability required in Theorem 5, Point 4.

Proposition 2. *Let $\{C(\mathbf{x})\}_{\mathbf{x} \in \mathbb{R}^2} \in \mathcal{C}$. Assume moreover that C has a constant expectation and satisfies the CLT. If C satisfies (21), then the random variables $\lambda \left(L_N(\lambda A, C) - \mathbb{E}[C(\mathbf{0})]\right)$, $\lambda > 0$, are uniformly integrable.*

Proof. Let, for $\lambda > 0$, $M_\lambda = \lambda \left(L_N(\lambda A, C) - \mathbb{E}[C(\mathbf{0})] \right)$. Theorem 2 gives that, for all $A \in \mathcal{A}_c$, $M_\lambda \xrightarrow{d} M$, for $\lambda \to \infty$, where $M \sim \mathcal{N}\left(0, \sigma_C^2 / \nu(A)\right)$. Therefore, by the continuous mapping theorem, we obtain

$$M_\lambda^2 \xrightarrow{d} M^2, \quad \text{for } \lambda \to \infty. \tag{35}$$

Now, it is clear that, for all $\lambda > 0$, $\mathrm{Var}(M_\lambda) = \lambda^2 \mathcal{R}_2(\lambda A, C)$. Hence, it follows from (34) that $\mathrm{Var}(M_\lambda) \underset{\lambda \to \infty}{\to} \sigma_C^2 / \nu(A)$, which gives, since for all $\lambda > 0$, $\mathbb{E}[M_\lambda] = 0$, that $\mathbb{E}\left[M_\lambda^2\right] \underset{\lambda \to \infty}{\to} \mathbb{E}\left[M^2\right]$. Additionally, M^2 is non-negative and integrable. Furthermore, the M_λ^2 are non-negative and, for all $\lambda > 0$, $\mathbb{E}\left[M_\lambda^2\right] = \lambda^2 \mathcal{R}_2(\lambda A, C)$, which is finite according to Theorem 4 as (19) is satisfied. Therefore, the M_λ^2 are integrable. Consequently, using (35) and Theorem 3.6 in Billingsley (1999), we know that the random variables M_λ^2, $\lambda > 0$, are uniformly integrable. This directly yields that the random variables M_λ, $\lambda > 0$, are uniformly integrable. □

3.2. Cost Field Being a Function of a Max-Stable Random Field

We now consider a cost field model written as in (14), i.e.,

$$\{C(\mathbf{x})\}_{\mathbf{x} \in \mathbb{R}^2} = \{E(\mathbf{x})\, D\left(Z(\mathbf{x})\right)\}_{\mathbf{x} \in \mathbb{R}^2}, \tag{36}$$

where Z is max-stable and the exposure is uniformly equal to unity. The relevance of using max-stable random fields has been previously highlighted.

In the following, all theorems and corollaries assume Z to be simple, although max-stable fields fitted to real data have generalized extreme-value (GEV) univariate marginal distributions with location, scale and shape parameters $\eta \in \mathbb{R}$, $\tau > 0$ and $\xi \in \mathbb{R}$. However, this does not cause any loss of generality. If $\{Z(\mathbf{x})\}_{\mathbf{x} \in \mathbb{R}^2}$ is a max-stable field with such GEV parameters, we can write

$$Z(\mathbf{x}) = \begin{cases} \eta + \tau(\tilde{Z}(\mathbf{x})^\xi - 1)/\xi & \text{if} \quad \xi \neq 0, \\ \eta + \tau \log(\tilde{Z}(\mathbf{x})) & \text{if} \quad \xi = 0, \end{cases} \quad \mathbf{x} \in \mathbb{R}^2, \tag{37}$$

where $\{\tilde{Z}(\mathbf{x})\}_{\mathbf{x} \in \mathbb{R}^2}$ is simple max-stable. Thus, there exists a function D_1 such that $Z(\mathbf{x}) = D_1(\tilde{Z}(\mathbf{x}))$ and Model (36) can be written $C(\mathbf{x}) = \tilde{D}(\tilde{Z}(\mathbf{x}))$, where $\tilde{Z}$ is simple max-stable and $\tilde{D} = D \circ D_1$, with "$\circ$" denoting function composition. On $(0, \infty)$, for any $\xi \neq 0$ the transformation $\tilde{z} \mapsto \eta + \tau(\tilde{z}^\xi - 1)/\xi$ is increasing and the same holds for the transformation $\tilde{z} \mapsto \eta + \tau \log(\tilde{z})$, implying that D_1 is increasing. Most often, the damage function D is also increasing (e.g., the higher the wind speed, temperature or rainfall amount, the higher the cost) and thus the same is true for $\tilde{D} = D \circ D_1$. Consequently, the requirement in Corollaries 3–5 (see below) on the function applied to the simple max-stable field to be non-decreasing and non-constant is generally satisfied in the applications motivating the present work.

For the sake of notational simplicity, in the following, we denote by Z (instead of $\tilde{Z}$) the simple max-stable field and by D (instead of $\tilde{D}$) the quantity $D \circ D_1$. Accordingly, the reader should pay attention to the fact that Z models the standardized environmental field (and not the real one) and D consists in the composition of the marginal transformation of Z and the damage function.

We first give sufficient conditions on the function D and the field Z such that the spatial risk measure $\mathcal{R}_1(\cdot, D(Z))$ induced by the cost field $D(Z)$ satisfies the axioms presented in Definition 6.

Corollary 2. *Let $\{Z(\mathbf{x})\}_{\mathbf{x} \in \mathbb{R}^2}$ be a simple max-stable random field and D a measurable function such that $\{C(\mathbf{x})\}_{\mathbf{x} \in \mathbb{R}^2} = \{D(Z(\mathbf{x}))\}_{\mathbf{x} \in \mathbb{R}^2} \in \mathcal{C}$ and $\mathbb{E}[|C(\mathbf{0})|] < \infty$. Then, for all $A \in \mathcal{A}$, $\mathcal{R}_1(A, C) = \mathbb{E}[C(\mathbf{0})]$. Hence, $\mathcal{R}_1(\cdot, C)$ satisfies the axioms of spatial invariance under translation and spatial sub-additivity. If, moreover, $\mathbb{E}[C(\mathbf{0})] \neq 0$, then $\mathcal{R}_1(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order 0 with $K_1(A, C) = 0$ and $K_2(A, C) = \mathbb{E}[C(\mathbf{0})]$, $A \in \mathcal{A}_c$.*

Proof. Since Z has identical margins, for all $\mathbf{x} \in \mathbb{R}^2$, $\mathbb{E}\left[|C(\mathbf{x})|\right] = \mathbb{E}\left[|C(\mathbf{0})|\right]$. Therefore, the result directly follows from Theorem 3. □

The result below gives sufficient conditions on D and Z such that the spatial risk measure $\mathcal{R}_2(\cdot, D(Z))$ induced by the cost field $D(Z)$ satisfies the axiom of asymptotic spatial homogeneity of order -2.

Theorem 6. *Let $\{Z(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2}$ be a simple and sample-continuous max-stable random field and D a measurable function such that $\{C(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2} = \{D(Z(\mathbf{x}))\}_{\mathbf{x}\in\mathbb{R}^2} \in \mathcal{C}$ and such that there exist $p, q > 0$ satisfying $2/p + 1/q = 1$ such that*

$$\mathbb{E}\left[|C(\mathbf{0})|^p\right] < \infty \tag{38}$$

and

$$\int_{\mathbb{R}^2} [2 - \theta(\mathbf{0}, \mathbf{x})]^{\frac{1}{q}} \, \nu(\mathrm{d}\mathbf{x}) < \infty, \tag{39}$$

where θ is the extremal coefficient function of Z. Then, we have

$$\int_{\mathbb{R}^2} |\mathrm{Cov}(C(\mathbf{0}), C(\mathbf{x}))| \, \nu(\mathrm{d}\mathbf{x}) < \infty.$$

Additionally, assume that C satisfies (21) and $\sigma_C > 0$. Then $\mathcal{R}_2(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -2 with $K_1(A, C) = 0$ and $K_2(A, C) = \sigma_C^2/\nu(A)$, $A \in \mathcal{A}_c$.

Proof. Since Z has identical margins, (38) yields that, for all $\mathbf{x} \in \mathbb{R}^2$, $\mathbb{E}\left[|C(\mathbf{x})|^p\right] < \infty$. Thus, using the fact that $2/p + 1/q = 1$, Davydov's inequality (Davydov 1968, (2.2)) gives that

$$|\mathrm{Cov}(C(\mathbf{0}), C(\mathbf{x}))| \leq 12 \left[\alpha^C(\{\mathbf{0}\}, \{\mathbf{x}\})\right]^{\frac{1}{q}} \left(\mathbb{E}\left[|C(\mathbf{0})|^p\right]\right)^{\frac{1}{p}} \left(\mathbb{E}\left[|C(\mathbf{x})|^p\right]\right)^{\frac{1}{p}}. \tag{40}$$

For all $\mathbf{x} \in \mathbb{R}^2$, since D is measurable, $C(\mathbf{x}) = D(Z(\mathbf{x}))$ is $\mathcal{F}^Z_{\{\mathbf{x}\}}$-measurable. Hence, $\mathcal{F}^C_{\{\mathbf{x}\}} \subset \mathcal{F}^Z_{\{\mathbf{x}\}}$, which gives by (15) that, for all $\mathbf{x} \in \mathbb{R}^2$,

$$\alpha^C\left(\{\mathbf{0}\}, \{\mathbf{x}\}\right) \leq \alpha^Z\left(\{\mathbf{0}\}, \{\mathbf{x}\}\right). \tag{41}$$

Now, using (16) and Corollary 2.2 in Dombry and Eyi-Minko (2012), we obtain that, for all $\mathbf{x} \in \mathbb{R}^2$,

$$\alpha^Z\left(\{\mathbf{0}\}, \{\mathbf{x}\}\right) \leq 2[2 - \theta(\mathbf{0}, \mathbf{x})]. \tag{42}$$

Thus, the combination of (41) and (42) gives that

$$\alpha^C\left(\{\mathbf{0}\}, \{\mathbf{x}\}\right) \leq 2[2 - \theta(\mathbf{0}, \mathbf{x})].$$

Consequently, (40) gives that

$$|\mathrm{Cov}(C(\mathbf{0}), C(\mathbf{x}))| \leq 12\, 2^{\frac{1}{q}} \left(\mathbb{E}\left[|C(\mathbf{0})|^p\right] \mathbb{E}\left[|C(\mathbf{x})|^p\right]\right)^{\frac{1}{p}} [2 - \theta(\mathbf{0}, \mathbf{x})]^{\frac{1}{q}}.$$

Therefore, using (38) and (39), we obtain

$$\int_{\mathbb{R}^2} |\mathrm{Cov}(C(\mathbf{0}), C(\mathbf{x}))| \, \nu(\mathrm{d}\mathbf{x}) < \infty.$$

Since $p, q > 0$ and $2/p + 1/q = 1$, we have $p > 2$. Consequently, for all $\mathbf{x} \in \mathbb{R}^2$, $\mathbb{E}\left[[C(\mathbf{x})]^2\right] < \infty$. Thus, Theorem 5, Point 2, gives the result. □

Until the end, the following results provide sufficient conditions on D and Z such that the induced spatial risk measures $\mathcal{R}_2(\cdot, D(Z))$, $\mathcal{R}_{3,\alpha}(\cdot, D(Z))$ and $\mathcal{R}_{4,\alpha}(\cdot, D(Z))$ satisfy the axioms of asymptotic spatial homogeneity of order -2, -1 and -1, respectively. In order to establish them, we take advantage of the results in Koch et al. (2018) about the existence of a CLT for functions of stationary

max-stable random fields. Let $\mathcal{B}(\mathbb{R})$ and $\mathcal{B}((0,\infty))$ be the Borel σ-fields on $\mathbb{R}$ and $(0,\infty)$, respectively. For $\mathbf{h} = (h_1, h_2)' \in \mathbb{Z}^2$, we adopt the notation $[\mathbf{h}, \mathbf{h}+1] = [h_1, h_1+1] \times [h_2, h_2+1]$. Next theorem considers a general simple, stationary and sample-continuous max-stable random field.

Theorem 7. *Let $\{Z(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2}$ be a simple, stationary and sample-continuous max-stable random field and D be a measurable function from $((0,\infty), \mathcal{B}((0,\infty)))$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ satisfying*

$$\mathbb{E}\left[|D(Z(\mathbf{0}))|^{2+\delta}\right] < \infty, \tag{43}$$

for some $\delta > 0$. Furthermore, assume that, for all $\mathbf{h} \in \mathbb{Z}^2$,

$$\mathbb{E}\left[\min\left\{\sup_{\mathbf{x}\in[0,1]^2}\{Y(\mathbf{x})\}, \sup_{\mathbf{x}\in[\mathbf{h},\mathbf{h}+1]}\{Y(\mathbf{x})\}\right\}\right] \leq K\|\mathbf{h}\|^{-b},$$

for some $K > 0$, $b > 2\max\{2, (2+\delta)/\delta\}$ and where $\{Y(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2}$ is a spectral random field of Z (see (17)). Let $\{C(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2} = \{D(Z(\mathbf{x}))\}_{\mathbf{x}\in\mathbb{R}^2}$. Then, if $\sigma_C > 0$:

1. *$\mathcal{R}_2(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -2 with $K_1(A, C) = 0$ and $K_2(A, C) = \sigma_C^2/\nu(A)$, $A \in \mathcal{A}_c$.*
2. *For all $\alpha \in (0,1)\backslash\{1/2\}$, $\mathcal{R}_{3,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A, C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A, C) = \sigma_C q_\alpha/[\nu(A)]^{\frac{1}{2}}$, $A \in \mathcal{A}_c$.*
3. *For all $\alpha \in (0,1)$, $\mathcal{R}_{4,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A, C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A, C) = \sigma_C \phi(q_\alpha)/\{[\nu(A)]^{\frac{1}{2}}(1-\alpha)\}$, $A \in \mathcal{A}_c$.*

Proof. Since Z is sample-continuous, it is measurable. Thus, the function D being measurable from $((0,\infty), \mathcal{B}((0,\infty)))$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, we obtain that C is measurable. Moreover, it follows from the stationarity of C (due to the stationarity of Z) and Condition (43) that, for all $\mathbf{x} \in \mathbb{R}^2$, $\mathbb{E}[|C(\mathbf{x})|] = \mathbb{E}[|C(\mathbf{0})|] < \infty$. Therefore, the function $\mathbf{x} \mapsto \mathbb{E}[|C(\mathbf{x})|]$ is constant and hence obviously locally integrable. Consequently, Proposition 1 gives that C has a.s. locally integrable sample paths. Therefore, $C \in \mathcal{C}$.

Furthermore, the assumptions enable us to apply Theorem 2 in Koch et al. (2018). The latter yields that the random field C satisfies the CLT. Finally, since C is stationary, it satisfies (21) and has a constant expectation. Hence, Corollary 1 gives the first result. The second result follows from Theorem 5, Point 3. The combination of Proposition 2 and Point 4 in Theorem 5 yields the third result. □

Theorem 7 directly entails the following result.

Corollary 3. *Let Z, D and C be as in Theorem 7 (but without assuming that $\sigma_C > 0$). Moreover, assume that D is non-decreasing and non-constant. Then:*

1. *$\mathcal{R}_2(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -2 with $K_1(A, C) = 0$ and $K_2(A, C) = \sigma_C^2/\nu(A)$, $A \in \mathcal{A}_c$.*
2. *For all $\alpha \in (0,1)\backslash\{1/2\}$, $\mathcal{R}_{3,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A, C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A, C) = \sigma_C q_\alpha/[\nu(A)]^{\frac{1}{2}}$, $A \in \mathcal{A}_c$.*
3. *For all $\alpha \in (0,1)$, $\mathcal{R}_{4,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A, C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A, C) = \sigma_C \phi(q_\alpha)/\{[\nu(A)]^{\frac{1}{2}}(1-\alpha)\}$, $A \in \mathcal{A}_c$.*

Proof. Proposition 1 in Koch et al. (2018) gives that $\sigma_C > 0$. Therefore, Theorem 7 yields the result. □

The next results concern the Brown–Resnick and Smith max-stable random fields. The Brown–Resnick model is of high practical interest since, owing to its flexibility, it appears as one of the best (if not the best) models among currently available max-stable models, at least for environmental data; see, e.g., Davison et al. (2012, Section 7.4), in the case of rainfall.

Theorem 8. *Let $\{Z(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2}$ be the Brown–Resnick random field associated with the variogram $\gamma_W(\mathbf{x}) = m\|\mathbf{x}\|^{\psi}$, where $m > 0$ and $\psi \in (0,2]$, or the Smith random field with covariance matrix Σ, and D be as in Theorem 7. Let $\{C(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2} = \{D(Z(\mathbf{x}))\}_{\mathbf{x}\in\mathbb{R}^2}$. Then, if $\sigma_C > 0$:*

1. *$\mathcal{R}_2(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -2 with $K_1(A,C) = 0$ and $K_2(A,C) = \sigma_C^2/\nu(A)$, $A \in \mathcal{A}_c$.*
2. *For all $\alpha \in (0,1)\backslash\{1/2\}$, $\mathcal{R}_{3,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A,C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A,C) = \sigma_C q_\alpha/[\nu(A)]^{\frac{1}{2}}$, $A \in \mathcal{A}_c$.*
3. *For all $\alpha \in (0,1)$, $\mathcal{R}_{4,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A,C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A,C) = \sigma_C\phi(q_\alpha)/\{[\nu(A)]^{\frac{1}{2}}(1-\alpha)\}$, $A \in \mathcal{A}_c$.*

Proof. We start with the proof in the case of the Brown–Resnick field. As previously mentioned, the Brown–Resnick random field is stationary. Thus, C is stationary and hence satisfies (21) and has a constant expectation. Moreover, we can see from the proof of Theorem 3 in Koch et al. (2018) that Z is sample-continuous. Consequently, the same arguments as in the proof of Theorem 7 yield that $C \in \mathcal{C}$. Furthermore, Theorem 3 in Koch et al. (2018) gives that C satisfies the CLT. Therefore, Corollary 1 yields the first result. The second result follows from Theorem 5, Point 3. The combination of Proposition 2 and Point 4 in Theorem 5 gives the third result.

The Smith random field is stationary as an instance of M3 random field. Thus, C is stationary and consequently satisfies (21) and has a constant expectation. Moreover, as the Smith field is sample-continuous, the same arguments as in the proof of Theorem 7 yield that $C \in \mathcal{C}$. Additionally, Theorem 4 in Koch et al. (2018) gives that C satisfies the CLT. Therefore, Corollary 1 yields the first result. The second result follows from Theorem 5, Point 3. The combination of Proposition 2 and Point 4 in Theorem 5 gives the third result. □

Next corollary easily follows from Theorem 8.

Corollary 4. *Let Z, D and C be as in Theorem 8 (but without assuming that $\sigma_C > 0$). Moreover, assume that D is non-decreasing and non-constant. Then:*

1. *$\mathcal{R}_2(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -2 with $K_1(A,C) = 0$ and $K_2(A,C) = \sigma_C^2/\nu(A)$, $A \in \mathcal{A}_c$.*
2. *For all $\alpha \in (0,1)\backslash\{1/2\}$, $\mathcal{R}_{3,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A,C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A,C) = \sigma_C q_\alpha/[\nu(A)]^{\frac{1}{2}}$, $A \in \mathcal{A}_c$.*
3. *For all $\alpha \in (0,1)$, $\mathcal{R}_{4,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A,C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A,C) = \sigma_C\phi(q_\alpha)/\{[\nu(A)]^{\frac{1}{2}}(1-\alpha)\}$, $A \in \mathcal{A}_c$.*

Proof. As explained in the proof of Theorem 8, both such Brown–Resnick fields and the Smith field are stationary and sample-continuous. Furthermore, they are simple max-stable. Thus, Proposition 1 in Koch et al. (2018) gives that $\sigma_C > 0$. Hence, Theorem 8 yields the result. □

Let $\|.\|$ denote the Euclidean norm in $\mathbb{R}^2$. We introduce $\mathcal{B}_1 = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| = 1\}$, the unit ball of $\mathbb{R}^2$. For two functions g_1 and g_2 from $\mathbb{R}^2$ to $\mathbb{R}$, the notation $g_1(\mathbf{h}) \underset{\|\mathbf{h}\|\to\infty}{=} o(g_2(\mathbf{h}))$ means that $\lim_{h\to\infty} \sup_{\mathbf{u}\in\mathcal{B}_1} \{|g_1(h\mathbf{u})/g_2(h\mathbf{u})|\} = 0$. Moreover, $\lim_{\|\mathbf{h}\|\to\infty} g_1(\mathbf{h}) = \infty$ must be understood as $\lim_{h\to\infty} \inf_{\mathbf{u}\in\mathcal{B}_1} \{g_1(h\mathbf{u})\} = \infty$.

Theorem 9. *Let $\{Z(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2}$ be the Brown–Resnick random field built with a random field $\{W(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2}$ which is sample-continuous and whose variogram satisfies*

$$\sup_{\mathbf{x}\in[0,1]^2} \{\gamma_W(\mathbf{h}) - \gamma_W(\mathbf{x}+\mathbf{h})\} \underset{\|\mathbf{h}\|\to\infty}{=} o(\gamma_W(\mathbf{h})),$$

and

$$\lim_{\|\mathbf{h}\|\to\infty} \frac{\gamma_W(\mathbf{h})}{\ln(\|\mathbf{h}\|)} = \infty.$$

Moreover, let D be as in Theorem 7. Let $\{C(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2} = \{D(Z(\mathbf{x}))\}_{\mathbf{x}\in\mathbb{R}^2}$. Then, if $\sigma_C > 0$:

1. *$\mathcal{R}_2(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -2 with $K_1(A, C) = 0$ and $K_2(A, C) = \sigma_C^2/\nu(A)$, $A \in \mathcal{A}_c$.*
2. *For all $\alpha \in (0,1)\backslash\{1/2\}$, $\mathcal{R}_{3,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A, C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A, C) = \sigma_C q_\alpha/[\nu(A)]^{\frac{1}{2}}$, $A \in \mathcal{A}_c$.*
3. *For all $\alpha \in (0,1)$, $\mathcal{R}_{4,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A, C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A, C) = \sigma_C \phi(q_\alpha)/\{[\nu(A)]^{\frac{1}{2}}(1-\alpha)\}$, $A \in \mathcal{A}_c$.*

Proof. The same arguments as in the proof of Theorem 8 show that C satisfies (21) and has a constant expectation. As W is sample-continuous, Proposition 13 in Kabluchko et al. (2009) gives that Z is sample-continuous. Thus, the same arguments as in the proof of Theorem 7 show that $C \in \mathcal{C}$. Moreover, Remark 3 in Koch et al. (2018) gives that C satisfies the CLT. Hence, Corollary 1 gives the first result. The second result follows from Theorem 5, Point 3. The combination of Proposition 2 and Point 4 in Theorem 5 yields the third result. □

The following result is a direct consequence of Theorem 9.

Corollary 5. *Let Z, D and C be as in Theorem 9 (but without assuming that $\sigma_C > 0$). Moreover, assume that D is non-decreasing and non-constant. Then:*

1. *$\mathcal{R}_2(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -2 with $K_1(A, C) = 0$ and $K_2(A, C) = \sigma_C^2/\nu(A)$, $A \in \mathcal{A}_c$.*
2. *For all $\alpha \in (0,1)\backslash\{1/2\}$, $\mathcal{R}_{3,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A, C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A, C) = \sigma_C q_\alpha/[\nu(A)]^{\frac{1}{2}}$, $A \in \mathcal{A}_c$.*
3. *For all $\alpha \in (0,1)$, $\mathcal{R}_{4,\alpha}(\cdot, C)$ satisfies the axiom of asymptotic spatial homogeneity of order -1 with $K_1(A, C) = \mathbb{E}[C(\mathbf{0})]$ and $K_2(A, C) = \sigma_C \phi(q_\alpha)/\{[\nu(A)]^{\frac{1}{2}}(1-\alpha)\}$, $A \in \mathcal{A}_c$.*

Proof. The random field Z is simple, stationary, sample-continuous (see the proof of Theorem 9) and max-stable. Thus, Proposition 1 in Koch et al. (2018) gives that $\sigma_C > 0$. Consequently, Theorem 9 yields the result. □

We conclude this section by commenting on the damage function $D(z) = \mathbb{I}_{\{z>u\}}, z > 0$, for $u > 0$, which is considered in Koch (2017). This function is measurable from $((0,\infty), \mathcal{B}((0,\infty)))$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Moreover, it is bounded and hence obviously satisfies (43) for every random field Z. Additionally, this function is non-decreasing and non-constant. Consequently, the results of Theorem 3, Point 3 and Theorem 5, Point 2 in Koch (2017) concerning the Brown–Resnick random field associated with the variogram $\gamma_W(\mathbf{x}) = m\|\mathbf{x}\|^\psi$, where $m > 0$ and $\psi \in (0, 2]$, and the Smith random field, are particular cases of Corollary 4.

4. Conclusions

In this paper, we first explore the notions of spatial risk measure and corresponding axioms introduced in Koch (2017) further as well as describe their utility for both actuarial science and practice. Second, in the case of a general cost field, we provide sufficient conditions such that spatial risk measures associated with expectation, variance, VaR as well as ES and induced by this cost field satisfy the axiom of asymptotic spatial homogeneity of order 0, -2, -1 and -1, respectively. Finally, in the case where the cost field is a function of a max-stable random field, we give sufficient conditions on both the function and the max-stable field such that spatial risk measures associated with expectation,

variance, VaR as well as ES and induced by the resulting cost field satisfy the axiom of asymptotic spatial homogeneity of order 0, -2, -1 and -1, respectively. Hence, these conditions allow one to know the rate of spatial diversification when the region under study becomes large, which is valuable for the banking/insurance industry. Overall, this paper improves our comprehension of the concept of spatial risk measure as well as of their properties with respect to the space variable and, among others, generalizes several results to be found in Koch (2017).

Ongoing work consists of the study of concrete examples of spatial risk measures involving max-stable fields and relevant damage functions. Inter alia, we apply our theory to winter storm risk over a specific European region. As previously mentioned, max-stable fields have GEV univariate marginal distributions with three parameters. The first step involves jointly fitting the latter and the dependence parameters of different max-stable models (Smith, Brown–Resnick, ...) to wind speed maxima using, e.g., composite likelihood methods (see, e.g., Padoan et al. 2010). Model selection has then to be performed employing, for instance, the composite likelihood information criterion. The second step consists in choosing an appropriate damage function and exposure field and leads, in combination with the first one, to the cost field model. If the sufficient conditions mentioned in Section 3.2 are met, then we can draw conclusions about the asymptotic rate of spatial diversification, and less importantly spatial invariance under translation. Simulating from the cost field, we obtain realizations of the normalized spatially aggregated loss on chosen sub-regions, which allow the estimation of the spatial risk measures of interest. This enables one to check, e.g., whether the axiom of spatial sub-additivity is satisfied.

Future work will include the study of spatial risk measures associated with other classical risk measures (e.g., more general distorsion risk measures than VaR or ES and expectile risk measures) and/or induced by cost fields involving other kinds of random fields than max-stable fields. For instance, it would be worthwhile to investigate whether spatial risk measures associated with VaR and ES can still satisfy the axiom of asymptotic spatial homogeneity of order -1 in the case where the cost field does not satisfy the CLT.

Funding: This research was partly funded by the Swiss National Science Foundation grant number 200021_178824.

Acknowledgments: The author would like to thank Anthony C. Davison and Christian Y. Robert for some interesting comments. He also acknowledges Paul Embrechts and Ruodu Wang for providing some references about robustness of risk measures as well as Gennady Samorodnitsky for a fruitful exchange about integrals of random fields. Finally, he is also grateful to the associate editor and two anonymous referees for insightful suggestions.

Conflicts of Interest: The author declares no conflict of interest. The funders had no role in the writing of the manuscript and in the decision to publish the results.

References

Anderes, Ethan B., and Michael L. Stein. 2011. Local likelihood estimation for nonstationary random fields. *Journal of Multivariate Analysis* 102: 506–20. [CrossRef]

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. Coherent measures of risk. *Mathematical Finance* 9: 203–28. [CrossRef]

Basel Committee on Banking Supervision. 2012. Fundamental Review of the Trading Book. Available online: https://www.bis.org/publ/bcbs219.htm (accessed on 25 April 2019).

Bevere, Lucia, and Lea Mueller. 2014. Natural catastrophes and man-made disasters in 2013: Large losses from floods and hail; Haiyan hits the Philippines. *Sigma Swiss Re* 1.

Billingsley, Patrick. 1999. *Convergence of Probability Measures*. New York: John Wiley & Sons.

Brown, Bruce M., and Sidney I. Resnick. 1977. Extreme values of independent stochastic processes. *Journal of Applied Probability* 14: 732–39. [CrossRef]

Dahlhaus, Rainer. 2012. Locally stationary processes. In *Handbook of Statistics*. Kidlington: Elsevier, pp. 351–413. [CrossRef]

Davison, Anthony C., Simone A. Padoan, and Mathieu Ribatet. 2012. Statistical modeling of spatial extremes. *Statistical Science* 27: 161–86. [CrossRef]

Davydov, Yu A. 1968. Convergence of distributions generated by stationary stochastic processes. *Theory of Probability and Its Applications* 13: 691–96. [CrossRef]

De Haan, Laurens. 1984. A spectral representation for max-stable processes. *The Annals of Probability* 12: 1194–204. [CrossRef]

De Haan, Laurens, and Ana Ferreira. 2006. *Extreme Value Theory: An Introduction*. New York: Springer. [CrossRef]

Denuit, Michel, Jan Dhaene, Marc Goovaerts, and Rob Kaas. 2005. *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. Chichester: John Wiley & Sons.

Dombry, Clément, and Frédéric Eyi-Minko. 2012. Strong mixing properties of max-infinitely divisible random fields. *Stochastic Processes and their Applications* 122: 3790–811. [CrossRef]

Eckley, Idris A., Guy P. Nason, and Robert L. Treloar. 2010. Locally stationary wavelet fields with application to the modelling and analysis of image texture. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 59: 595–616. [CrossRef]

Gneiting, Tilmann. 2011. Making and evaluating point forecasts. *Journal of the American Statistical Association* 106: 746–62. [CrossRef]

Huser, Raphaël, and Anthony C. Davison. 2013. Composite likelihood estimation for the Brown–Resnick process. *Biometrika* 100: 511–18. [CrossRef]

Kabluchko, Zakhar, Martin Schlather, and Laurens de Haan. 2009. Stationary max-stable fields associated to negative definite functions. *The Annals of Probability* 37: 2042–65. [CrossRef]

Koch, Erwan. 2017. Spatial risk measures and applications to max-stable processes. *Extremes* 20: 635–70. [CrossRef]

Koch, Erwan, Clément Dombry, and Christian Y. Robert. 2018. A central limit theorem for functions of stationary mixing max-stable random fields on $\mathbb{R}^d$. *Stochastic Processes and Their Applications*. [CrossRef]

Ombao, Hernando C., Jonathan A. Raz, Rainer von Sachs, and Beth A. Malow. 2001. Automatic statistical analysis of bivariate nonstationary time series. *Journal of the American Statistical Association* 96: 543–60. [CrossRef]

Padoan, Simone A., Mathieu Ribatet, and Scott A. Sisson. 2010. Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association* 105: 263–77. [CrossRef]

Resnick, Sidney I. 1987. *Extreme Values, Regular Variation, and Point Processes*. New York: Springer. [CrossRef]

Rosenblatt, Murray. 1956. A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences of the United States of America* 42: 43–47. [CrossRef] [PubMed]

Samorodnitsky, Gennady, and Murad S. Taqqu. 1994. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. London: Chapman and Hall/CRC.

Schlather, Martin, and Jonathan A. Tawn. 2003. A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika* 90: 139–56. [CrossRef]

Smith, Richard L. 1990. Max-Stable Processes and Spatial Extremes. Unpublished manuscript. Available online: https://pdfs.semanticscholar.org/033b/fb040b67dbe584a62ae900ad537f5f0eef0b.pdf (accessed on 25 April 2019).

Volkonskii, V A., and Yu A. Rozanov. 1959. Some limit theorems for random functions. I. *Theory of Probability and Its Applications* 4: 178–97. [CrossRef]

Wang, Ruodu, Yunran Wei, and Gordon Willmot. 2018. Characterization, Robustness and Aggregation of Signed Choquet Integrals. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2956962 (accessed on 25 April 2019).

risks

MDPI

Article

On the Moments and the Distribution of Aggregate Discounted Claims in a Markovian Environment

Shuanming Li [1,*] and Yi Lu [2]

[1] Centre for Actuarial Studies, Department of Economics, The University of Melbourne, Melbourne 3010, Australia

[2] Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada; yilu@sfu.ca

* Correspondence: shli@unimelb.edu.au

Received: 7 May 2018; Accepted: 21 May 2018; Published: 23 May 2018

Abstract: This paper studies the moments and the distribution of the aggregate discounted claims (ADCs) in a Markovian environment, where the claim arrivals, claim amounts, and forces of interest (for discounting) are influenced by an underlying Markov process. Specifically, we assume that claims occur according to a Markovian arrival process (MAP). The paper shows that the vector of joint Laplace transforms of the ADC occurring in each state of the environment process by any specific time satisfies a matrix-form first-order partial differential equation, through which a recursive formula is derived for the moments of the ADC occurring in certain states (a subset). We also study two types of covariances of the ADC occurring in any two subsets of the state space and with two different time lengths. The distribution of the ADC occurring in certain states by any specific time is also investigated. Numerical results are also presented for a two-state Markov-modulated model case.

Keywords: aggregate discounted claims; Markovian arrival process; partial integro-differential equation; covariance

1. Introduction

Consider a line of business or an insurance portfolio to be insured by a property and casualty insurance company. Suppose that random claims arrive in the future according to a counting process, denoted by $\{N(t)\}_{t\geq 0}$, i.e., $N(t)$ is the random number of claims up to time t. Assume that $\{T_n\}_{n\geq 1}$ is a sequence of random claim occurrence times and $\{X_n\}_{n\geq 1}$ is a sequence of corresponding random positive claim amounts (also called claim severities), and $\delta(t)$ is the force of interest at time t, which is modeled by a stochastic process. Then $S(t)$ defined by

$$S(t) = \sum_{n=1}^{N(t)} X_n e^{-\int_0^{T_n} \delta(s)ds}, \quad t \geq 0 \tag{1}$$

is the aggregate discounted claims (ADCs) up to certain time t, or the present value of the total amounts paid out by the company up to time t, which describes the random change over time of the insurer's future liabilities at present time. Accordingly, $\{S(t)\}_{t\geq 0}$ is the ADC process (compound discounted claims) for this business. At a fixed time t, the randomness of $S(t)$ comes from the number of claims up to time t, claim occurrence times, and corresponding sizes as well as the values of $\delta(s)$, $0 \leq s \leq t$. It is an important quantity in the sense that, at the time of issue ($t = 0$), this quantity would help insurers set a premium for this particular line of business, and predict and manage their future liabilities.

A simple case of model (1) is one in which the counting process $\{N(t)\}_{t\geq 0}$ is a homogeneous Poisson process, independent of claim amounts, and the force of interest is deterministic. In this paper, we assume that the counting process $\{N(t)\}_{t\geq 0}$ is a Markovian arrival process (MAP) with

representation $(\boldsymbol{\gamma}, \mathbf{D_0}, \mathbf{D_1})$, introduced by Neuts (1979). That is, claim arrivals are influenced by an underlying continuous-time Markov process $\{J(t)\}_{t\geq 0}$ on state space $E = \{1, 2, \ldots, m\}$ with an $m \times m$ intensity matrix $\mathbf{D}$ and initial distribution $\boldsymbol{\gamma}$, where $\mathbf{D} = \mathbf{D_0} + \mathbf{D}_1 = (d_{0,ij}) + (d_{1,ij})$, and is assumed to be irreducible. Precisely, $d_{0,ij}$ represents the intensity of transitions from state i to state j without claim arrivals, while $d_{1,ij}(\geq 0)$ represents the intensity of transitions from state i to state j with an accompanying claim, having a cumulative distribution function F_i, density function f_i, k-th moment $\mu_i^{(k)}$, and Laplace transform $\hat{f}_i(s) = \int_0^\infty e^{-sx} f_i(x)dx$. Here, the process $\{J(t)\}_{t\geq 0}$ models the random environment, which affects the frequency and the severity of claims and thus the insurance business; for example, it is well known that the weather or climate conditions have impacts on automobile, property and casualty insurance claims.

Moreover, we assume that the force of interest process $\{\delta(t)\}_{t\geq 0}$ in (1) is also governed by the same Markov process $\{J(t)\}_{t\geq 0}$ and is assumed constant while staying at certain state, that is, when $J(t) = i$, $\delta(t) = \delta_i(> 0)$, for all $i \in E$. As the force of interest used for evaluation is mainly driven by the local or global economics conditions, we would reasonably model its random fluctuations by a stochastic process that is different from $\{J(t)\}_{t\geq 0}$. Technically, we can assume a two-dimensional Markov process as the environment or background process and other mathematical treatments would be the same as we do below. Hence, we make the above assumption in this paper to simplify notations and presentations. We note that studies of the influence of economic conditions such as interest and inflation on the classical risk theory can be found in papers by Taylor (1979), Delbaen and Haezendonck (1987), Willmot (1989), and Garrido and Léveill (2004).

The MAP has received considerable attention in recent decades due to its versatility and feasibility in modeling stochastic insurance claims dynamics. MAPs include Poisson processes, renewal processes with the inter-arrival times following phase-type distributions, and Markov-modulated Poisson processes as special cases, which are intensively studied in actuarial science literature. Detailed characteristics and properties of MAPs can be found in papers by Neuts (1979) and Asmussen (2003). Below, we present a brief literature review on the related work based on models given by Equation (1) (including its special cases).

Most of the studies on model (1) are under the assumption that $\{\delta(t)\}_{t\geq 0}$ is deterministic. For the ADC, Léveillé and Garrido (2001a) give explicit expressions for its first two moments in the compound renewal risk process by using renewal theory arguments, while Léveillé and Garrido (2001b) further derive a recursive formula for the moments calculation. Léveillé et al. (2010) study the moment generating function (mgf) of the ADC by finite and infinite time under a renewal risk model or a delayed renewal risk model. Recently, Wang et al. (2018) studied the distribution of discounted compound phase-type renewal sums using the analytical results of their mgf obtained by Léveillé et al. (2010). Jang (2004) obtains the Laplace transform of the distribution of the ADC using a shot noise process. Woo and Cheung (2013) derive recursive formulas for the moments of the ADC using techniques used by Léveillé and Garrido (2001b), for a renewal risk process with certain dependence between the claim arrival and the amount caused. The impact of the dependency on the ADC are illustrated numerically. Kim and Kim (2007) derive simple expressions for the first two moments of the ADC when the rates of claim arrivals and the claim sizes depend on the states of an underlying Markov process. Ren (2008) studies the Laplace transform and the first two moments of the ADC following a MAP process, and Li (2008) further derives a recursive formula for the moments of the discounted claims for the same model. Barges et al. (2011) study the moments of the ADC in a compound Poisson model with dependence introduced by a Farlie–Gumbel–Morgenstern (FGM) copula; Mohd Ramli and Jang (2014) further derive Neumann series expression of the recursive moments by using the method of successive approximation.

There are few papers that study models described by Equation (1) with a stochastic process $\{\delta(t)\}_{t\geq 0}$ in the literature of actuarial science. Leveille and Adekambi (2011, 2012) study the covariance and the joint moments of the discounted compound renewal sum at two different times with a stochastic interest rate where the Ho–Lee–Merton and the Vasicek interest rate models are

considered. Their idea of studying the covariance and the joint moments is adopted and extended in this paper. Here, we assume that the components of the ADC process $\{S(t)\}_{t\geq 0}$ described by Equation (1)—the number of claims, the size of the claims, and the force of interest for discounting—are all influenced by the same Markovian environment process, which enhances the flexibility of the model parameter settings. It follows that $S(t)$ depends on the trajectory of this underlying process whose states may represent different external conditions or circumstances that affect insurance claims. The main objective of this paper is to study the moments and the distribution of $S(t)$ given in Equation (1), occurring in certain states (e.g., certain conditions) by time t.

In general, while the expectation of $S(t)$ at any given time t can be used as a reference for the insurer's liability, the higher moments of $S(t)$, describing further characteristics of the random variable such as the variability around the mean and how extreme outcomes could go, may be used to determine the marginals on reserves. Furthermore, the distributional results regarding $S(t)$ would be useful for obtaining the risk measures such as the value at risk and the conditional tail expectation, which may help insurers prevent or minimize their losses from extreme cases.

Our work is basically a generalization of some aforementioned studies. We first obtain formulas for calculating mean, variance, and distribution of the ADC occurring in a subset of states at a certain time. The subset may represent a collection of similar conditions that the insurer would consider them as a whole. We then derive explicit matrix-analytic expressions for covariances of the ADC occurring in two subsets of the state space at a certain time and those occurring in a certain subset of states with two different time lengths. The motivation of studying these two types of covariance is that we believe they can reveal the correlation between the random discounted sums either between different underlying conditions or with different time lengths, and the information would be helpful for insurers to set their capital requirements for preventing future losses, and make strategic and contingency plans. Moreover, we obtain a matrix-form partial integro-differential equation satisfied by the distribution function of the ADC occurring in certain subset of states. The equation can be solved numerically to obtain the probability distribution function of the ADC, which again could be useful for measuring insurers' risks of insolvency.

The rest of the paper is organized as follows. In Section 2, we study the joint Laplace transforms of the ADC occurring in each state by time t and pay attention to some special cases. Recursive formulas for calculating the moments of the ADC occurring in certain states are obtained. A formula for computing the covariance of the ADC occurring in two subsets of the state space is derived in Section 3, while the covariance of the ADC occurring in certain states with two different time lengths is studied in Section 4. The distribution of the ADC occurring in certain states is investigated in Section 5. Finally, some numerical illustrations are presented in Section 6.

2. The Laplace Transforms and Moments

We first decompose $S(t)$ into m components as

$$S(t) = \sum_{j=1}^{m} S_j(t)$$

where

$$S_j(t) = \sum_{n=1}^{N(t)} X_n I(J(T_n) = j) e^{-\int_0^{T_n} \delta(s)ds}$$

is the ADC occurring in state $j \in E$, with $I(\cdot)$ being the indicator function. For a given $k (1 \leq k \leq m)$, $1 \leq l_1 < l_2 < \ldots < l_k \leq m$ denote $E_k = \{l_1, l_2, \ldots, l_k\} \subseteq E$, a sub-state space of E. We then define

$$S_{E_k}(t) = \sum_{j \in E_k} S_j(t)$$

to be the ADC occurring in the subset of state space E_k. In particular, $S_E(t) = S(t)$ and $S_{\{j\}}(t) = S_j(t)$. If $\delta(t) = 0$ and $X_i \equiv 1$ for all $i \in \mathbb{N}^+$, then $S_{E_k}(t) = N_{E_k}(t)$, where $N_{E_k}(t)$ is the number of claims occurring in the sub-state space E_k by time t.

Let $\mathbb{P}_i$ and $\mathbb{E}_i$ denote conditional probability and conditional expectation given $J(0) = i$, respectively. Define

$$ {}_iL(\xi_1, \xi_2, \ldots, \xi_m; t) = \mathbb{E}_i \left[e^{-\sum_{j=1}^m \xi_j S_j(t)} \right], \quad \xi_j \geq 0, t \geq 0, i \in E \tag{2} $$

to be the joint Laplace transform of $S_1(t), S_2(t), \ldots, S_m(t)$, given that the initial state is i. In particular, we have

$$ \begin{aligned} {}_iL(\xi; t) &= \mathbb{E}_i \left[e^{-\xi S(t)} \right] = {}_iL(\xi, \xi, \ldots, \xi; t) \\ {}_iL_{E_k}(\xi; t) &= \mathbb{E}_i \left[e^{-\xi S_{E_k}(t)} \right] = {}_iL(\xi_1, \xi_2, \ldots, \xi_m; t)\big|_{\xi_j = \xi I(j=l_n), n=1,2,\ldots,k} \\ {}_iL_j(\xi_j; t) &= \mathbb{E}_i \left[e^{-\xi_j S_j(t)} \right] = {}_iL(\xi_1, \xi_2, \ldots, \xi_m; t)\big|_{\xi_k = 0, k \neq j}. \end{aligned} $$

We define, for $n \in \mathbb{N}^+$, the n-th moment of $S(t)$, $S_j(t)$, and $S_{E_k}(t)$, respectively, as

$$ \begin{aligned} {}_iV^{(n)}(t) &= \mathbb{E}_i \left[S^n(t) \right], \quad i \in E \\ {}_iV_j^{(n)}(t) &= \mathbb{E}_i \left[S_j^n(t) \right], \quad i, j \in E \\ {}_iV_{E_k}^{(n)}(t) &= \mathbb{E}_i \left[S_{E_k}^n(t) \right], \quad 1 \leq k \leq m \end{aligned} $$

given that the initial state is i.

We write the following column vectors for the Laplace transforms

$$ \begin{aligned} \mathbf{L}(\xi_1, \xi_2, \ldots, \xi_m; t) &= \left({}_1L(\xi_1, \xi_2, \ldots, \xi_m; t), \ldots, {}_mL(\xi_1, \xi_2, \ldots, \xi_m; t) \right)^\top \\ \mathbf{L}(\xi; t) &= \left({}_1L(\xi; t), {}_2L(\xi; t), \ldots, {}_mL(\xi; t) \right)^\top \\ \mathbf{L}_{E_k}(\xi; t) &= \left({}_1L_{E_k}(\xi; t), {}_2L_{E_k}(\xi; t), \ldots, {}_mL_{E_k}(\xi; t) \right)^\top \\ \mathbf{L}_j(\xi_j; t) &= \left({}_1L_j(\xi_j; t), {}_2L_j(\xi_j; t), \ldots, {}_mL_j(\xi_j; t) \right)^\top, \end{aligned} $$

with $\mathbf{L}(0; t) = \mathbf{L}_{E_k}(0; t) = \mathbf{L}_j(0; t) = \mathbf{1} = (1, 1, \ldots, 1)^\top$.

In this section, we first show that $\mathbf{L}(\xi_1, \xi_2, \ldots, \xi_m; t)$ satisfies a matrix-form first-order partial differential equation, and derive recursive formulas for calculating the moments of various ADC depending on the initial state of the underlying Markovian process. We also consider some special cases.

Theorem 1. *$\mathbf{L}(\xi_1, \xi_2, \ldots, \xi_m; t)$ satisfies*

$$ \begin{aligned} &\frac{\partial \mathbf{L}(\xi_1, \xi_2, \ldots, \xi_m; t)}{\partial t} + \boldsymbol{\delta} \sum_{j=1}^m \xi_j \frac{\partial \mathbf{L}(\xi_1, \xi_2, \ldots, \xi_m; t)}{\partial \xi_j} \\ &\quad = \mathbf{D}_0 \mathbf{L}(\xi_1, \xi_2, \ldots, \xi_m; t) + \hat{\mathbf{f}}(\xi_1, \xi_2, \ldots, \xi_m) \mathbf{D}_1 \mathbf{L}(\xi_1, \xi_2, \ldots, \xi_m; t) \end{aligned} \tag{3} $$

where $\boldsymbol{\delta} = diag(\delta_1, \delta_2, \ldots, \delta_m)$ and $\hat{\mathbf{f}}(\xi_1, \xi_2, \ldots, \xi_m) = diag\big(\hat{f}_1(\xi_1), \hat{f}_2(\xi_2), \ldots, \hat{f}_m(\xi_m)\big)$.

Proof. For an infinitesimal $h > 0$, conditioning on three possible events which can occur in $[0, h]$—no change in the MAP phase (state), a change in the MAP phase accompanied by no claims, and a change in the MAP phase accompanied by a claim—we have

$$\begin{aligned} {}_iL(\xi_1, \xi_2, \ldots, \xi_m; t) = {} & [1 + d_{0,ii}h]\, {}_iL\big(\xi_1 e^{-\delta_i h}, \xi_2 e^{-\delta_i h}, \ldots, \xi_m e^{-\delta_i h}; t - h\big) \\ & + \sum_{k=1, k\neq i}^{m} d_{0,ik} h\, {}_kL\big(\xi_1 e^{-\delta_i h}, \xi_2 e^{-\delta_i h}, \ldots, \xi_m e^{-\delta_i h}; t - h\big) \\ & + \sum_{k=1}^{m} d_{1,ik} h \hat{f}_i\big(\xi_i e^{-\delta_i h}\big)\, {}_kL\big(\xi_1 e^{-\delta_i h}, \xi_2 e^{-\delta_i h}, \ldots, \xi_m e^{-\delta_i h}; t - h\big). \end{aligned} \tag{4}$$

As ${}_iL(\xi_1, \xi_2, \ldots, \xi_m; t)$ is differentiable with respect to ξ_i $(i \in E)$ and t (the differentiability of ${}_iL$ with respect to t is justified in Appendix), we have

$$\begin{aligned} & {}_iL\big(\xi_1 e^{-\delta_i h}, \xi_2 e^{-\delta_i h}, \ldots, \xi_m e^{-\delta_i h}; t - h\big) \\ & = {}_iL(\xi_1, \xi_2, \ldots, \xi_m; t) - h \frac{\partial\, {}_iL(\xi_1, \xi_2, \ldots, \xi_m; t)}{\partial t} - \delta_i h \sum_{l=1}^{m} \xi_l \frac{\partial\, {}_iL(\xi_1, \xi_2, \ldots, \xi_m; t)}{\partial \xi_l} + o(h) \end{aligned} \tag{5}$$

where $\lim_{h\to 0}(o(h)/h) = 0$. Substituting the expression above into Equation (4), dividing both sides by h, and letting $h \to 0$, we have

$$\begin{aligned} & \delta_i \sum_{l=1}^{m} \xi_l \frac{\partial\, {}_iL(\xi_1, \xi_2, \ldots, \xi_m; t)}{\partial \xi_l} + \frac{\partial\, {}_iL(\xi_1, \xi_2, \ldots, \xi_m; t)}{\partial t} \\ & = \sum_{k=1}^{m} d_{0,ik}\, {}_kL(\xi_1, \xi_2, \ldots, \xi_m; t) + \sum_{k=1}^{m} d_{1,ik} \hat{f}_i(\xi_i)\, {}_kL(\xi_1, \xi_2, \ldots, \xi_m; t). \end{aligned} \tag{6}$$

Rewriting Equation (6) in matrix form gives Equation (3). □

Remark 1. *Using the same argument, we have the follow results.*

(1) $\mathbf{L}_{E_k}(\xi; t)$ *satisfies the following matrix-form first-order partial differential equation:*

$$\frac{\partial \mathbf{L}_{E_k}(\xi; t)}{\partial t} + \delta\xi \frac{\partial \mathbf{L}_{E_k}(\xi; t)}{\partial \xi} = \mathbf{D_0} \mathbf{L}_{E_k}(\xi; t) + \hat{\mathbf{f}}_{E_k}(\xi) \mathbf{D_1} \mathbf{L}_{E_k}(\xi; t) \tag{7}$$

where $\hat{\mathbf{f}}_{E_k}(\xi)$ *is an* $m \times m$ *diagonal matrix with the* l_i*-th entry being* $\hat{f}_{l_i}(\xi)$*, for* $i = 1, 2, \ldots, k$ *and all other entries being* 1.

(2) $\mathbf{L}(\xi; t)$ *satisfies*

$$\frac{\partial \mathbf{L}(\xi; t)}{\partial t} + \delta\xi \frac{\partial \mathbf{L}(\xi; t)}{\partial \xi} = \mathbf{D_0} \mathbf{L}(\xi; t) + \hat{\mathbf{f}}(\xi) \mathbf{D_1} \mathbf{L}(\xi; t)$$

where $\hat{\mathbf{f}}(\xi) = diag\big(\hat{f}_1(\xi), \hat{f}_2(\xi), \ldots, \hat{f}_m(\xi)\big)$.

(3) $\mathbf{L}_j(\xi_j; t)$ *satisfies*

$$\frac{\partial \mathbf{L}_j(\xi_j; t)}{\partial t} + \delta\xi_j \frac{\partial \mathbf{L}_j(\xi_j; t)}{\partial \xi_j} = \mathbf{D_0} \mathbf{L}_j(\xi_j; t) + \hat{\mathbf{f}}_j(\xi_j) \mathbf{D_1} \mathbf{L}_j(\xi_j; t)$$

where $\hat{\mathbf{f}}_j(\xi_j) = diag\big(1, 1, \ldots, \hat{f}_j(\xi_j), 1, \ldots, 1\big)$.

We now study the moments of the ADC considered in Theorem 1. Denote the vectors of the n-th moment of the corresponding ADC as

$$\begin{aligned}\mathbf{V}_n(t) &= \left({}_1V^{(n)}(t),\ {}_2V^{(n)}(t),\ldots,\ {}_mV^{(n)}(t)\right)^\top\\ \mathbf{V}_{n,E_k}(t) &= \left({}_1V^{(n)}_{E_k}(t),\ {}_2V^{(n)}_{E_k}(t),\ldots,\ {}_mV^{(n)}_{E_k}(t)\right)^\top\\ \mathbf{V}_{n,j}(t) &= \left({}_1V^{(n)}_{j}(t),\ {}_2V^{(n)}_{j}(t),\ldots,\ {}_mV^{(n)}_{j}(t)\right)^\top.\end{aligned}$$

From Equation (7), we obtain in Theorem 2 a matrix-form first-order differential equation satisfied by the moments of $S_{E_k}(t)$, $\mathbf{V}_{n,E_k}(t)$ and then, in Theorem 3, obtain recursive formulas for calculating them.

Theorem 2. *The moments of $S_{E_k}(t)$ satisfy*

$$\mathbf{V}'_{n,E_k}(t)+\left(n\boldsymbol{\delta}-\mathbf{D}_0-\mathbf{D}_1\right)\mathbf{V}_{n,E_k}(t)=\sum_{r=1}^{n}\binom{n}{r}\mathbf{I}_{E_k}\boldsymbol{\mu}_r\,\mathbf{D}_1\,\mathbf{V}_{n-r,E_k}(t),\quad n\in\mathbb{N}^+,\tag{8}$$

with initial conditions $\mathbf{V}_{n,E_k}(0)=\mathbf{0}$ and $\mathbf{V}_{0,E_k}(t)=\mathbf{1}$. In particular,

$$\mathbf{V}'_{1,E_k}(t)+\left(\boldsymbol{\delta}-\mathbf{D}_0-\mathbf{D}_1\right)\mathbf{V}_{1,E_k}(t)=\mathbf{I}_{E_k}\boldsymbol{\mu}_1\mathbf{D}_1\mathbf{1},\quad t\geq 0$$

where $\boldsymbol{\mu}_r=diag\left(\mu_1^{(r)},\mu_2^{(r)},\ldots,\mu_m^{(r)}\right)$, $\mathbf{I}_{E_k}$ is an $m\times m$ diagonal matrix with the l_i-th entry being 1, for $i=1,2,\ldots,k$, and all other diagonal entries being 0.

Proof. By Taylor's expansion (its existence is easily justified as we assume that f_i has moment $\mu_i^{(n)}$ for any $n\in\mathbb{N}^+$), we have

$$\hat{f}_i(\xi)=1+\sum_{n=1}^{\infty}\frac{(-1)^n\xi^n}{n!}\mu_i^{(n)}.$$

In matrix notation,

$$\hat{\mathbf{f}}_{E_k}(\xi)=\mathbf{1}+\sum_{n=1}^{\infty}\frac{(-1)^n\xi^n}{n!}\mathbf{I}_{E_k}\boldsymbol{\mu}_n.\tag{9}$$

Substituting Equation (9) together with

$$\mathbf{L}_{E_k}(\xi;t)=\sum_{n=0}^{\infty}\frac{(-1)^n\xi^n}{n!}\mathbf{V}_{n,E_k}(t)$$

into Equation (7) and equating the coefficients of ξ^n give Equation (8). □

Corollary 1. *We have the following results for the moments of $S(t)$ and $S_j(t)$.*

(i) $\mathbf{V}_n(t)$ *satisfies the matrix-form first-order differential equation:*

$$\mathbf{V}'_n(t)+\left(n\boldsymbol{\delta}-\mathbf{D}_0-\mathbf{D}_1\right)\mathbf{V}_n(t)=\sum_{r=1}^{n}\binom{n}{r}\boldsymbol{\mu}_r\mathbf{D}_1\,\mathbf{V}_{n-r}(t),\quad n\in\mathbb{N}^+$$

where $\mathbf{V}_n(0)=\mathbf{0}$ and $\mathbf{V}_0(t)=\mathbf{1}$. In particular, $\mathbf{V}_1(t)$ satisfies

$$\mathbf{V}'_1(t)+\left(\boldsymbol{\delta}-\mathbf{D}_0-\mathbf{D}_1\right)\mathbf{V}_1(t)=\boldsymbol{\mu}_1\mathbf{D}_1\mathbf{1},\quad t\geq 0.$$

(ii) $\mathbf{V}_{n,j}(t)$ *satisfies*

$$\mathbf{V}'_{n,j}(t)+\left(n\boldsymbol{\delta}-\mathbf{D}_0-\mathbf{D}_1\right)\mathbf{V}_{n,j}(t)=\sum_{r=1}^{n}\binom{n}{r}\mathbf{I}_j\boldsymbol{\mu}_r\mathbf{D}_1\,\mathbf{V}_{n-r,j}(t)$$

where $\mathbf{I}_j = \mathbf{I}_{\{j\}}$ *is a diagonal matrix with the j-th entry being* 1, *and* 0 *otherwise,* $\mathbf{V}_{n,j}(0) = \mathbf{0}$ *and* $\mathbf{V}_{0,j}(t) = \mathbf{1}$. *In particular,* $\mathbf{V}_{1,j}(t)$ *satisfies*

$$\mathbf{V}'_{1,j}(t) + (\boldsymbol{\delta} - \mathbf{D}_0 - \mathbf{D}_1)\mathbf{V}_{1,j}(t) = \mathbf{I}_j\boldsymbol{\mu}_1\mathbf{D}_1\mathbf{1}, \quad t \geq 0.$$

Solving differential Equation (8) with $\mathbf{V}_{n,E_k}(0) = \mathbf{0}$, we obtain the following recursive formulas for $\mathbf{V}_{n,E_k}(t)$.

Theorem 3. *For* $t > 0$ *and* $n \in \mathbb{N}^+$, *we have*

$$\mathbf{V}_{n,E_k}(t) = \sum_{r=1}^{n} \binom{n}{r} \int_0^t e^{-\left(n\,\boldsymbol{\delta} - (\mathbf{D}_0+\mathbf{D}_1)\right)x} \mathbf{I}_{E_k}\boldsymbol{\mu}_r\mathbf{D}_1\mathbf{V}_{n-r,E_k}(t-x)dx\,.$$

In particular,

$$\begin{aligned}\mathbf{V}_{1,E_k}(t) &= \int_0^t e^{-\left(\boldsymbol{\delta} - (\mathbf{D}_0+\mathbf{D}_1)\right)x} dx\, \mathbf{I}_{E_k}\boldsymbol{\mu}_1\,\mathbf{D}_1\mathbf{1} \\ &= [\boldsymbol{\delta} - (\mathbf{D}_0 + \mathbf{D}_1)]^{-1}\left[\mathbf{I} - e^{-[\boldsymbol{\delta} - (\mathbf{D}_0+\mathbf{D}_1)]t}\right] \mathbf{I}_{E_k}\boldsymbol{\mu}_1\,\mathbf{D}_1\mathbf{1}\,.\end{aligned} \tag{10}$$

Clearly, we have $\mathbf{V}_{1,E_k}(t) + \mathbf{V}_{1,E_k^c}(t) = \mathbf{V}_1(t)$, where $E_k^c = E \setminus E_k$.

Corollary 2. *If we set* $E_k = E$ *and* $E_k = \{j\}$ *in Theorem 3, we have the following recursive formulas for the moments of* $S(t)$ *and* $S_j(t)$*:*

$$\begin{aligned}\mathbf{V}_n(t) &= \sum_{k=1}^{n} \binom{n}{k} \int_0^t e^{-\left(n\,\boldsymbol{\delta} - (\mathbf{D}_0+\mathbf{D}_1)\right)x} \boldsymbol{\mu}_k\mathbf{D}_1\mathbf{V}_{n-k}(t-x)dx \\ \mathbf{V}_{n,j}(t) &= \sum_{k=1}^{n} \binom{n}{k} \int_0^t e^{-\left(n\,\boldsymbol{\delta} - (\mathbf{D}_0+\mathbf{D}_1)\right)x} \mathbf{I}_j\boldsymbol{\mu}_k\mathbf{D}_1\mathbf{V}_{n-k,j}(t-x)dx\,.\end{aligned}$$

In particular,

$$\begin{aligned}\mathbf{V}_1(t) &= [\boldsymbol{\delta} - (\mathbf{D}_0 + \mathbf{D}_1)]^{-1}\left[\mathbf{I} - e^{-[\boldsymbol{\delta} - (\mathbf{D}_0+\mathbf{D}_1)]t}\right] \boldsymbol{\mu}_1\mathbf{D}_1\mathbf{1} \\ \mathbf{V}_{1,j}(t) &= [\boldsymbol{\delta} - (\mathbf{D}_0 + \mathbf{D}_1)]^{-1}\left[\mathbf{I} - e^{-[\boldsymbol{\delta} - (\mathbf{D}_0+\mathbf{D}_1)]t}\right] \mathbf{I}_j\boldsymbol{\mu}_1\mathbf{D}_1\mathbf{1}\,.\end{aligned}$$

Remark 2. *When* $t \to \infty$, *we have the following asymptotic results for the moments of the ADC for* $n \in \mathbb{N}^+$*:*

$$\begin{aligned}\mathbf{V}_{n,E_k}(\infty) &= [n\boldsymbol{\delta} - (\mathbf{D}_0 + \mathbf{D}_1)]^{-1} \sum_{r=1}^{n} \binom{n}{r} \mathbf{I}_{E_k}\boldsymbol{\mu}_r\mathbf{D}_1\mathbf{V}_{n-r,E_k}(\infty) \\ \mathbf{V}_n(\infty) &= [n\boldsymbol{\delta} - (\mathbf{D}_0 + \mathbf{D}_1)]^{-1} \sum_{r=1}^{n} \binom{n}{r} \boldsymbol{\mu}_r\,\mathbf{D}_1\mathbf{V}_{n-r}(\infty) \\ \mathbf{V}_{n,j}(\infty) &= [n\boldsymbol{\delta} - (\mathbf{D}_0 + \mathbf{D}_1)]^{-1} \sum_{r=1}^{n} \binom{n}{r} \mathbf{I}_j\,\boldsymbol{\mu}_r\,\mathbf{D}_1\mathbf{V}_{n-r,j}(\infty)\end{aligned}$$

where $\mathbf{V}_{0,E_k}(\infty) = \mathbf{V}_0(\infty) = \mathbf{V}_{0,j}(\infty) = \mathbf{1}$.

3. The Covariance of ADC Occurring in Two Sub-State Spaces

In this section, we first calculate the joint moment of the ADC occurring in two subsets of the state space and then the covariance between them could be calculated.

For $1 \leq l_1 < l_2 < \ldots < l_k \leq m$ and $1 \leq n_1 < n_2 < \ldots < n_j \leq m$, where $2 \leq k+j \leq m$, denote $E_k = \{l_1, l_2, \ldots, l_k\}$ and $E_j = \{n_1, n_2, \ldots, n_j\}$ to be two disjoint subsets of E, i.e., $E_k \cap E_j = \emptyset$. The aggregate discounted claim amounts occurring in E_k and E_j are

$$S_{E_k}(t) = \sum_{i \in E_k} S_i(t), \quad S_{E_j}(t) = \sum_{i \in E_j} S_i(t).$$

Define

$${}_iL_{E_k,E_j}(\xi_k, \xi_j; t) = \mathbb{E}_i \left[e^{-\xi_k S_{E_k}(t) - \xi_j S_{E_j}(t)} \right]$$

to be the joint Laplace transform of $S_{E_k}(t)$ and $S_{E_j}(t)$. Let $\mathbf{L}_{E_k,E_j}(\xi_k, \xi_j; t)$ be a column vector with the i-th entry being ${}_iL_{E_k,E_j}(\xi_k, \xi_j; t)$. Moreover, let

$${}_iV_{E_k,E_j}(t) = \mathbb{E}_i \left[S_{E_k}(t) S_{E_j}(t) \right]$$

be the joint moment of $S_{E_k}(t)$ and $S_{E_j}(t)$. Denote $\mathbf{V}_{E_k,E_j}(t)$ as an $m \times 1$ column vector with the i-th entry being ${}_iV_{E_k,E_j}(t)$. A matrix-form integral expression of $\mathbf{V}_{E_k,E_j}(t)$ and its asymptotic formula when $t \to \infty$ are presented in the theorem below.

Theorem 4. *For two disjoint subsets of E, E_k and E_j, the joint moment of $S_{E_k}(t)$ and $S_{E_j}(t)$ satisfies*

$$\begin{aligned} \mathbf{V}_{E_k,E_j}(t) = & \int_0^t e^{-\left(2\delta - (\mathbf{D}_0 + \mathbf{D}_1)\right)x} \mathbf{I}_{E_k} \boldsymbol{\mu}_1 \mathbf{D_1} \mathbf{V}_{1,E_j}(t-x) dx \\ & + \int_0^t e^{-\left(2\delta - (\mathbf{D}_0 + \mathbf{D}_1)\right)x} \mathbf{I}_{E_j} \boldsymbol{\mu}_1 \mathbf{D_1} \mathbf{V}_{1,E_k}(t-x) dx \end{aligned} \tag{11}$$

where $\mathbf{V}_{1,E_k}(t)$ is given by Equation (10) in Theorem 3. When $t \to \infty$, we have

$$\mathbf{V}_{E_k,E_j}(\infty) = \left[2\delta - (\mathbf{D}_0 + \mathbf{D}_1)\right]^{-1} \left[\mathbf{I}_{E_k} \boldsymbol{\mu}_1 \mathbf{D_1} \mathbf{V}_{1,E_j}(\infty) + \mathbf{I}_{E_j} \boldsymbol{\mu}_1 \mathbf{D_1} \mathbf{V}_{1,E_k}(\infty) \right]. \tag{12}$$

Proof. Following from Equation (3), we have

$$\begin{aligned} & \frac{\partial \mathbf{L}_{E_k,E_j}(\xi_k, \xi_j; t)}{\partial t} + \delta \xi_k \frac{\partial \mathbf{L}_{E_k,E_j}(\xi_k, \xi_j; t)}{\partial \xi_k} + \delta \xi_j \frac{\partial \mathbf{L}_{E_k,E_j}(\xi_k, \xi_j; t)}{\partial \xi_j} \\ & = \mathbf{D}_0 \mathbf{L}_{E_k,E_j}(\xi_k, \xi_j; t) + \hat{\mathbf{f}}_{E_k,E_j}(\xi_k, \xi_j) \mathbf{D}_1 \mathbf{L}_{E_k,E_j}(\xi_k, \xi_j; t) \end{aligned} \tag{13}$$

where $\hat{\mathbf{f}}_{E_k,E_j}(\xi_k, \xi_j)$ is a diagonal matrix with the l_i-th entry being $\hat{f}_{l_i}(\xi_k)$, for $i = 1, 2, \ldots, k$, with the n_i-th entry being $\hat{f}_{n_i}(\xi_j)$, for $i = 1, 2, \ldots, j$, and all other elements being 1.

Taking partial derivatives with respect to ξ_k and ξ_j on both sides of Equation (13), setting $\xi_k = 0$ and $\xi_j = 0$, and noting that

$${}_iV_{E_k,E_j}(t) = \frac{\partial^2 \, {}_iL_{E_k,E_j}(\xi_k, \xi_j; t)}{\partial \xi_k \partial \xi_j} \Big|_{\xi_k = \xi_j = 0},$$

we obtain the following matrix-form first-order differential equation for $\mathbf{V}_{E_k,E_j}(t)$:

$$\mathbf{V}'_{E_k,E_j}(t) + \left[2\delta - \mathbf{D}_0 - \mathbf{D}_1\right] \mathbf{V}_{E_k,E_j}(t) = \mathbf{I}_{E_k} \boldsymbol{\mu}_1 \mathbf{D_1} \mathbf{V}_{1,E_j}(t) + \mathbf{I}_{E_j} \boldsymbol{\mu}_1 \mathbf{D_1} \mathbf{V}_{1,E_k}(t).$$

Solving it gives Equation (11).

Letting $t \to \infty$ in Equation (11), we obtain expression (12) for the joint moment of $S_{E_k}(\infty)$ and $S_{E_j}(\infty)$. □

Remark 3. *If $E_k = \{k\}$ and $E_j = \{j\}$, and $k \neq j$, we have*

$$\begin{aligned}\mathbf{V}_{\{k\},\{j\}}(t) = &\int_0^t e^{-\left(2\delta-(\mathbf{D}_0+\mathbf{D}_1)\right)x}\mathbf{I}_j\boldsymbol{\mu}_1\mathbf{D_1}\mathbf{V}_{1,k}(t-x)dx\\ &+\int_0^t e^{-\left(2\delta-(\mathbf{D}_0+\mathbf{D}_1)\right)x}\mathbf{I}_k\boldsymbol{\mu}_1\mathbf{D_1}\mathbf{V}_{1,j}(t-x)dx\,.\end{aligned}$$

When $t \to \infty$, the joint moment of $S_k(\infty)$ and $S_j(\infty)$ can be expressed as

$$\mathbf{V}_{\{k\},\{j\}}(\infty) = \left[2\delta-(\mathbf{D}_0+\mathbf{D}_1)\right]^{-1}\left[\mathbf{I}_j\boldsymbol{\mu}_1\mathbf{D_1}\mathbf{V}_{1,k}(\infty) + \mathbf{I}_k\boldsymbol{\mu}_1\mathbf{D_1}\mathbf{V}_{1,j}(\infty)\right].$$

Remark 4. *If two subsets E_k and E_j are not disjoint, i.e., $E_k \cap E_j = E_{kj} \neq \emptyset$, then*

$$\begin{aligned}Cov_i\left(S_{E_k}(t), S_{E_j}(t)\right) &= Cov_i\left(S_{E_k\setminus E_{kj}}(t) + S_{E_{kj}}(t),\, S_{E_j\setminus E_{kj}}(t) + S_{E_{kj}}(t)\right)\\ &= Cov_i\left(S_{E_k\setminus E_{kj}}(t), S_{E_{kj}}(t)\right) + Cov_i\left(S_{E_j\setminus E_{kj}}(t), S_{E_{kj}}(t)\right)\\ &\quad + Cov_i\left(S_{E_k\setminus E_{kj}}(t),\, S_{E_j\setminus E_{kj}}(t)\right) + Var_i\left(S_{E_{kj}}(t)\right).\end{aligned}$$

All the covariance terms in the expression above are for ADCs occurring in two disjoint sets.

4. The Covariance of the ADC with Two Different Time Lengths

In this section, we investigate the covariance of the ADCs occurring in two (overlapped) time periods, i.e., we want to evaluate

$$\begin{aligned}\mathrm{Cov}_i(S_{E_k}(t), S_{E_k}(t+h)) &\triangleq \mathrm{Cov}(S_{E_k}(t), S_{E_k}(t+h)\big|J(0)=i)\\ &= \mathbb{E}_i\left[S_{E_k}(t)S_{E_k}(t+h)\right] - \mathbb{E}_i[S_{E_k}(t)]\mathbb{E}_i[S_{E_k}(t+h)]\end{aligned}$$

for $t, h > 0$ and $E_k = \{l_1, l_2, \ldots, l_k\}$ with $k \leq m$. Denote $\mathbf{R}_{E_k}(t, t+h)$ as a column vector with the i-th entry being $\mathbb{E}_i\left[S_{E_k}(t)S_{E_k}(t+h)\right]$. In the following, we first show in a lemma a result that is needed for deriving the expression for $\mathbf{R}_{E_k}(t, t+h)$. We then present an explicit formula of $\mathbf{R}_{E_k}(t, t+h)$ in a theorem below.

As $S_{E_k}(t+h) = S_{E_k}(t+h) - S_{E_k}(t) + S_{E_k}(t)$, we have

$$\mathbb{E}_i\left[S_{E_k}(t)S_{E_k}(t+h)\right] = \mathbb{E}_i\left[S^2_{E_k}(t)\right] + \mathbb{E}_i\left[S_{E_k}(t)\left(S_{E_k}(t+h) - S_{E_k}(t)\right)\right]. \tag{14}$$

Define $\mathcal{F}_t = \sigma(S(v); 0 \leq v \leq t)$ to be σ-algebra generated by the ADC process by time t. Using the law of iterated expectation, we have

$$\begin{aligned}&\mathbb{E}_i\left[S_{E_k}(t)\left(S_{E_k}(t+h) - S_{E_k}(t)\right)\right]\\ &= \mathbb{E}_i\left\{\mathbb{E}\left[S_{E_k}(t)\left(S_{E_k}(t+h) - S_{E_k}(t)\right)\big|\mathcal{F}_t\right]\right\}\\ &= \mathbb{E}_i\left\{S_{E_k}(t)\mathbb{E}\left[\left(S_{E_k}(t+h) - S_{E_k}(t)\right)\big|\mathcal{F}_t\right]\right\}\\ &= \mathbb{E}_i\left\{S_{E_k}(t)e^{-\int_0^t\delta(s)ds}\mathbb{E}\left[\left(S_{E_k}(t,t+h)\right)\big|\mathcal{F}_t\right]\right\}\\ &= \mathbb{E}_i\left\{S_{E_k}(t)e^{-\int_0^t\delta(s)ds}\mathbb{E}\left[\left(S_{E_k}(t,t+h)\right)\big|J(t)\right]\right\}\\ &= \sum_{j=1}^m\mathbb{E}_i\left\{S_{E_k}(t)e^{-\int_0^t\delta(s)ds}\mathbb{E}\left[\left(S_{E_k}(t,t+h)\right)\big|J(t)=j\right]\right\}\mathbb{P}(J(t)=j|J(0)=i)\\ &= \sum_{j=1}^m\mathbb{E}_i\left[S_{E_k}(t)e^{-\int_0^t\delta(s)ds}I(J(t)=j)\right]\mathbb{E}_j[S_{E_k}(h)]\mathbb{P}(J(t)=j|J(0)=i)\end{aligned} \tag{15}$$

where $S_{E_k}(t, t+h)$ is the present value, at time t, of the claims occurring in states within E_k over $(t, t+h]$.

Denote $\mathbf{M}_{E_k}(t) = \left(M_{i,j,E_k}(t)\right)_{m\times m}$, where

$$M_{i,j,E_k}(t) = \mathbb{E}_i\left[S_{E_k}(t)e^{-\int_0^t \delta(s)ds} I(J(t)=j)\right].$$

The following lemma gives a matrix-form integral expression for $\mathbf{M}_{E_k}(t)$.

Lemma 1. *$\mathbf{M}_{E_k}(t)$ is of the form*

$$\mathbf{M}_{E_k}(t) = \int_0^t e^{-\left(2\delta-(\mathbf{D}_0+\mathbf{D}_1)\right)x}\mathbf{I}_{E_k}\boldsymbol{\mu}_1\mathbf{D}_1\mathbf{v}(t-x)dx \tag{16}$$

where $\mathbf{v}(t)$ is a matrix with (i,j)-th element being

$$v_{i,j}(t) = \mathbb{E}_i\left[e^{-\int_0^t \delta(s)ds} I(J(t)=j)\right], \quad i,j\in E.$$

Proof. Conditioning on the events that may occur over an infinitesimal interval $(0,\Delta t)$, we have

$$\begin{aligned} M_{i,j,E_k}(t) &= (1+d_{0,ii}\Delta t)e^{-2\delta_i\Delta t}M_{i,j,E_k}(t-\Delta t) + \sum_{l\neq i} d_{0,il}\Delta t\, e^{-2\delta_i\Delta t}M_{l,j,E_k}(t-\Delta t) \\ &\quad + \sum_{l=1}^{m} d_{1,il}\,\Delta t\, e^{-2\delta_i\Delta t}\left[I(i\in E_k)\mu_i^{(1)}\mathbb{E}_l\left(e^{-\int_{\Delta t}^t \delta(s)ds}I(J(t)=j)\right) + M_{l,j,E_k}(t-\Delta t)\right]. \end{aligned} \tag{17}$$

We can then obtain a matrix-form differential equation for $\mathbf{M}_{E_k}(t)$ from Equation (17) as follows:

$$\mathbf{M}'_{E_k}(t) = (\mathbf{D}_0+\mathbf{D}_1-2\delta)\mathbf{M}_{E_k}(t) + \mathbf{I}_{E_k}\boldsymbol{\mu}_1\mathbf{D}_1\mathbf{v}(t), \tag{18}$$

with $\mathbf{M}_{E_k}(0) = \mathbf{0}$. In fact, it is easy to show that $\mathbf{v}(t) = e^{(\mathbf{D}_0+\mathbf{D}_1-\delta)t}$, with $\mathbf{v}(0) = \mathbf{I}$ and $\mathbf{v}(\infty) = \mathbf{0}$. Solving Equation (18) gives Equation (16). □

Let $q_{i,j}(t) = \mathbb{P}_i(J(t)=j)$. Then $\mathbf{Q}(t) = \left(q_{i,j}(t)\right)_{m\times m}$ is the transition matrix of the underlying Markov process $\{J(t)\}_{t\geq 0}$ at time t. It follows from Ren (2008) that $\mathbf{Q}(t) = e^{(\mathbf{D}_0+\mathbf{D}_1)t}$.

Theorem 5. *$\mathbf{R}_{E_k}(t,t+h)$ can be expressed as*

$$\mathbf{R}_{E_k}(t,t+h) = \mathbf{V}_{2,E_k}(t) + (\mathbf{M}_{E_k}\circ\mathbf{Q})(t)\mathbf{V}_{1,E_k}(h) \tag{19}$$

where $(\mathbf{M}_{E_k}\circ\mathbf{Q})(t)$ is the Hadamard product of $\mathbf{M}_{E_k}(t)$ and $\mathbf{Q}(t)$, i.e., the (i,j)-th element of $(\mathbf{M}_{E_k}\circ\mathbf{Q})(t)$ is $M_{i,j,E_k}(t)\times q_{i,j}(t)$, and $\mathbf{M}_{E_k}(t)$ is given by Equation (16) in Lemma 1.

Proof. Equation (19) follows immediately from Equations (14) and (15). □

Remark 5. *If $E_k = E$ or $E_k = \{k\}$, Equation (19) simplifies to the joint moment of $S(t)$ and $S(t+h)$, or the joint moment of $S_k(t)$ and $S_k(t+h)$.*

5. The Distributions of the ADC

In this section, we investigate the distributions of $S_{E_k}(t)$ and its two special cases, $S(t)$ and $S_k(t)$, for $E_k = \{l_1, l_2, \ldots, l_k\} \subseteq E$. To precede, we define for $x\geq 0$ and $i\in E$,

$$\begin{aligned} G_i(x,t) &= \mathbb{P}_i(S(t)\leq x) \\ G_{i,k}(x,t) &= \mathbb{P}_i(S_k(t)\leq x) \\ G_{i,E_k}(x,t) &= \mathbb{P}_i\left(S_{E_k}(t)\leq x\right), \end{aligned}$$

with the following conditions:

$$\begin{aligned} G_i(x,0) &= G_{i,k}(x,0) = G_{i,E_k}(x,0) = 1, \quad x \geq 0 \\ G_i(0,t) &= \mathbb{P}_i(N(t) = 0) \\ G_{i,k}(0,t) &= \mathbb{P}_i(N_k(t) = 0) \\ G_{i,E_k}(0,t) &= \mathbb{P}_i\big(N_{E_k}(t) = 0\big) \end{aligned}$$

where $N_k(t) = \sum_{l=1}^{N(t)} I(J(T_l) = k)$ is the number of claims occurring in state k and $N_{E_k}(t) = \sum_{j \in E_k} N_j(t)$ is the number of claims occurring in the subset E_k. Denote

$$\begin{aligned} \mathbf{G}(x,t) &= \big(G_1(x,t), G_2(x,t), \ldots, G_m(x,t)\big)^\top \\ \mathbf{G}_k(x,t) &= \big(G_{1,k}(x,t), G_{2,k}(x,t), \ldots, G_{m,k}(x,t)\big)^\top \\ \mathbf{G}_{E_k}(x,t) &= \big(G_{1,E_k}(x,t), G_{2,E_k}(x,t), \ldots, G_{m,E_k}(x,t)\big)^\top. \end{aligned}$$

We present in the theorem below that $\mathbf{G}_{E_k}(x,t)$ satisfies a first-order partial integro-differential equation in matrix form.

Theorem 6. *$\mathbf{G}_{E_k}(x,t)$ satisfies*

$$\begin{aligned} &\frac{\partial \mathbf{G}_{E_k}(x,t)}{\partial t} - x\boldsymbol{\delta}\frac{\partial \mathbf{G}_{E_k}(x,t)}{\partial x} \\ &\quad = (\mathbf{D}_0 + \mathbf{D}_1 - \mathbf{I}_{E_k}\mathbf{D}_1)\mathbf{G}_{E_k}(x,t) + \int_0^x \mathbf{I}_{E_k}\mathbf{f}(y)\mathbf{D}_1\mathbf{G}_{E_k}(x-y,t)dy\,, \end{aligned} \tag{20}$$

with initial conditions

$$\mathbf{G}_{E_k}(x,0) = \mathbf{1}, \quad \mathbf{G}_{E_k}(0,t) = e^{(\mathbf{D}_0+\mathbf{D}_1-\mathbf{I}_{E_k}\mathbf{D}_1)t}\mathbf{1} \tag{21}$$

where $\mathbf{G}_{E_k}(0,t)$ is the solution of the differential equation obtained from Equation (20) by setting $x = 0$.

Proof. Using the same arguments as in Section 2, we have, by conditioning on events that may occur over $(0,h]$,

$$\begin{aligned} G_{i,E_k}(x,t) = [1 + d_{0,ii}h]G_{i,E_k}\left(xe^{\delta_i h}, t-h\right) + \sum_{j=1, j\neq i}^{m} d_{0,ij}hG_{j,E_k}\left(xe^{\delta_i h}, t-h\right) \\ + \sum_{j=1}^{m} d_{1,ij}h\,G_{j,E_k}\left(xe^{\delta_i h}, t-h\right), \quad i \notin E_k\,. \end{aligned} \tag{22}$$

As $G_{i,E_k}(x,t)$ is differentiable with respect to x and t, we have

$$G_{i,E_k}\left(xe^{\delta_i h}, t-h\right) = G_{i,E_k}(x,t) + \delta_i xh\frac{\partial G_{i,E_k}(x,t)}{\partial x} - h\frac{\partial G_{i,E_k}(x,t)}{\partial t} + o(h)\,. \tag{23}$$

The justification of Equation (23) can be done similarly as that for Equation (5) (see Appendix A). Substituting Equation (23) into Equation (22), rearranging terms, dividing both sides by h, and taking limit as $h \to 0$, give

$$\frac{\partial G_{i,E_k}(x,t)}{\partial t} - \delta_i x\frac{\partial G_{i,E_k}(x,t)}{\partial x} = \sum_{j=1}^{m} d_{0,ij}G_{j,E_k}(x,t) + \sum_{j=1}^{m} d_{1,ij}G_{j,E_k}(x,t), \quad i \notin E_k\,.$$

For $i \in E_k = \{l_1, l_2, \ldots, l_k\}$, we have

$$G_{i,E_k}(x,t) = [1 + d_{0,ii}h]G_{i,E_k}\left(xe^{\delta_i h}, t-h\right) + \sum_{j=1, j\neq i}^{m} d_{0,ik}hG_{j,E_k}\left(xe^{\delta_i h}, t-h\right) + \sum_{j=1}^{m} d_{1,ij}h \int_0^{xe^{\delta_i h}} f_i(y)G_{j,E_k}\left(xe^{\delta_i h} - y, t-h\right) dy\,.$$

Taylor's expansion gives

$$\frac{\partial G_{i,E_k}(x,t)}{\partial t} - \delta_i x \frac{\partial G_{i,E_k}(x,t)}{\partial x} = \sum_{j=1}^{m} d_{0,ij}G_{j,E_k}(x,t) + \sum_{j=1}^{m} d_{1,ij}\int_0^x f_i(y)G_{j,E_k}(x-y,t)dy\,, \quad i \in E_k\,.$$

Equations for $i \in E_k$ and $i \notin E_k$ can then be expressed in matrix form (20). □

Remark 6. *If we set $E_k = E$ and $E_k = \{k\}$, respectively, we have the following results:*

$$\begin{aligned} \frac{\partial \mathbf{G}(x,t)}{\partial t} - x\boldsymbol{\delta}\frac{\partial \mathbf{G}(x,t)}{\partial x} &= \mathbf{D}_0\mathbf{G}(x,t) + \int_0^x \mathbf{f}(y)\mathbf{D_1}\mathbf{G}(x-y,t)dy \\ \frac{\partial \mathbf{G}_k(x,t)}{\partial t} - x\boldsymbol{\delta}\frac{\partial \mathbf{G}_k(x,t)}{\partial x} &= (\mathbf{D}_0 + \mathbf{D}_1 - \mathbf{I}_k\,\mathbf{D}_1)\mathbf{G}_k(x,t) + \int_0^x \mathbf{I}_k\mathbf{f}(y)\mathbf{D}_1\mathbf{G}_k(x-y,t)dy\,, \end{aligned} \tag{24}$$

with initial conditions

$$\begin{aligned} \mathbf{G}(x,0) &= \mathbf{1}, \quad \mathbf{G}(0,t) = e^{\mathbf{D_0}t}\mathbf{1} \\ \mathbf{G}_k(x,0) &= \mathbf{1}, \quad \mathbf{G}_k(0,t) = e^{(\mathbf{D_0}+\mathbf{D}_1-\mathbf{I}_k\mathbf{D}_1)t}\mathbf{1}\,. \end{aligned}$$

Here, $\mathbf{G}_k(0,t)$ is the solution of the differential equation obtained from Equation (24) by setting $x = 0$.

Remark 7. *The matrix-form partial integro-differential Equation (20) with the corresponding initial conditions given by Equation (21) may be solved numerically as follows.*

(a) For two infinitesimal h_1 and h_2, we set $\mathbf{G}_{E_k}(lh_1, 0) = \mathbf{1}$, for $l = 1, 2, \ldots$, and we calculate $\mathbf{G}_{E_k}(0, nh_2)$ using Equation (21) for $n = 1, 2, \ldots$.

(b) With Equation (20), $\mathbf{G}_{E_k}(lh_1, nh_2)$ can be calculated recursively, for $n, l = 1, 2, \ldots$, by

$$\begin{aligned} \mathbf{G}_{E_k}(lh_1, nh_2) = \left[\mathbf{I} - lh_2\,\boldsymbol{\delta} - h_2(\mathbf{D}_0 + \mathbf{D}_1 - \mathbf{I}_{E_k}\mathbf{D}_1)\right]^{-1} \\ \times \Big[\mathbf{G}_{E_k}(lh_1, (n-1)h_2) - lh_2\,\boldsymbol{\delta}\mathbf{G}_{E_k}\big((l-1)h_1, nh_2\big) \\ + h_2\,h_1\sum_{j=0}^{l-1}\mathbf{I}_{E_k}\mathbf{f}(jh_1)\mathbf{D}_1\mathbf{G}_{E_k}\big((l-1-j)h_1, nh_2\big)\Big]. \end{aligned}$$

Remark 8. *If $f_i(x) = \beta_i e^{-\beta_i x}, \beta_i > 0$, then $\mathbf{f}(x) = \boldsymbol{\beta}e^{-\boldsymbol{\beta}x}$, with $\mathbf{f}'(x) = -\boldsymbol{\beta}\mathbf{f}(x)$, where $\boldsymbol{\beta} = diag(\beta_1, \beta_2, \ldots, \beta_m)$. Taking partial derivative with respect to x on both sides of Equation (20) and performing some manipulations, we obtain the following matrix-form second-order partial differential equation for $\mathbf{G}_{E_k}(x,t)$:*

$$\begin{aligned} &\frac{\partial^2 \mathbf{G}_{E_k}(x,t)}{\partial t \partial x} - x\boldsymbol{\delta}\frac{\partial^2 \mathbf{G}_{E_k}(x,t)}{\partial x^2} + \mathbf{I}_{E_k}\boldsymbol{\beta}\frac{\partial \mathbf{G}_{E_k}(x,t)}{\partial t} \\ &+ \left(\boldsymbol{\delta} - \mathbf{D}_0 - \mathbf{D}_1 + \mathbf{I}_{E_k}(x\boldsymbol{\delta}\boldsymbol{\beta} + \mathbf{D}_1)\right)\frac{\partial \mathbf{G}_{E_k}(x,t)}{\partial x} - \mathbf{I}_{E_k}(\boldsymbol{\beta}\mathbf{D}_0 + \mathbf{D_1})\mathbf{G}_{E_k}(x,t) = \mathbf{0}\,. \end{aligned} \tag{25}$$

This partial differential equation can also be solved numerically by using forward finite difference methods.

Remark 9. *Li et al. (2015) show that, when $\delta(s) = 0$, $G_i(x,t)$ can be used to find an expression for the density of the time of ruin in a MAP risk model.*

6. Numerical Illustrations

In this section, we consider a two-state Markov-modulated with intensity matrix

$$\mathbf{A} = \begin{pmatrix} -1/4 & 1/4 \\ 3/4 & -3/4 \end{pmatrix}.$$

We also assume that $f_1(x) = e^{-x}$, $f_2(x) = 0.5e^{-0.5x}$, $x > 0$, $\lambda_1 = 1$, $\lambda_2 = 2/3$, $\delta_1 = 0.03$, and $\delta_2 = 0.05$. Table 1 gives the first moments of $S_1(t)$ and $S_2(t)$ and their covariance for $t = 1, 2, 5, 10, 20, 30$, and ∞, given $J(0) = 1$ and $J(0) = 2$, respectively, in which the covariances, for $i = 1, 2$, are calculated by

$$\mathrm{Cov}_i(t) \triangleq \mathrm{Cov}\big(S_1(t), S_2(t) \big| J(0) = i\big) = \mathbb{E}_i\big[S_1(t)S_2(t)\big] - \mathbb{E}_i[S_1(t)]\mathbb{E}_i[S_2(t)].$$

It shows that, as expected, the expected values of $S_1(t)$ and $S_2(t)$ (and hence $S(t)$) are increasing in t given $J(0) = i$ for $i = 1, 2$. It is not surprised to see that $S_1(t)$ and $S_2(t)$ are negatively correlated for any t, as claims occurring in the two states compete with each other. Moreover, the larger the time t, the more the negative correlation between $S_1(t)$ and $S_2(t)$.

Figure 1 plots the variances of $S(t)$, $S_1(t)$, and $S_2(t)$, given $J(0) = 1$, for $0 \le t \le 150$. The variances all increase with time t. The variance of $S(t)$ is bigger than those of $S_1(t)$ and $S_2(t)$ for a fixed t. When time t goes to ∞, the three variances converge.

Tables 2 and 3 display the covariances of the ADC at time t and $t + h$, given $J(0) = 1$, for some selected t values and for $h = 1$ and $h = 5$. It is shown that $S(t)$ and $S(t + h)$, $S_1(t)$ and $S_1(t + h)$, and $S_2(t)$ and $S_2(t + h)$ are all positively correlated. Moreover, when t increases, the covariances increase; moreover, when h increases, the covariances decrease. When $t \to \infty$, the covariances of the pairs $S(t)$ and $S(t + h)$, $S_i(t)$, and $S_i(t + h)$ converge to the variances of $S(\infty)$ and $S_i(\infty)$, respectively. Similar patterns should be expected for $J(0) = 2$.

Finally, we display in Figure 2 the numerical values of the distribution function of $S(t)$ with initial state i, $G_i(x,t) = \mathbb{P}_i(S(t) \le x)$, for $t = 1$ and 4, $0 \le x \le 25$, and $i = 1, 2$. Note that $\mathbf{G}(x,t) = (G_1(x,t), G_2(x,t))^\top$ satisfies the partial differential Equation (25); its solution can be obtained numerically. From the graph, it shows clearly that the probability of $S(t)$ being bigger than a fixed x is smaller for small values of t as expected. For most x values, $G_1(x,t)$ is bigger than $G_2(x,t)$ due to the fact that the underlying Markov process in our example tends to stay in state 1 more often than staying at state 2.

Table 1. Expected values and covariances of $S_1(t)$ and $S_2(t)$.

t	$J(0) = 1$			$J(0) = 2$		
	$\mathbb{E}_1[S_1(t)]$	$\mathbb{E}_1[S_2(t)]$	$\mathrm{Cov}_1(t)$	$\mathbb{E}_2[S_1(t)]$	$\mathbb{E}_2[S_2(t)]$	$\mathrm{Cov}_2(t)$
1	0.8948	0.1196	−0.0599	0.2690	0.9444	−0.1412
2	1.6665	0.3607	−0.2832	0.8117	1.4717	−0.5475
5	3.7056	1.1998	−1.3303	2.6996	2.4452	−1.8361
10	6.6248	2.4695	−2.9252	5.5563	3.6966	−3.4208
20	11.1330	4.4336	−5.0170	9.9757	5.6221	−5.4630
30	14.3123	5.8188	−6.1938	13.0922	6.9800	−6.6142
∞	21.9178	9.1324	−7.9012	20.5479	10.2283	−8.2962

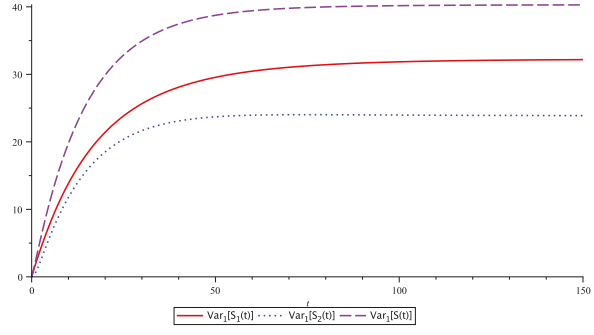


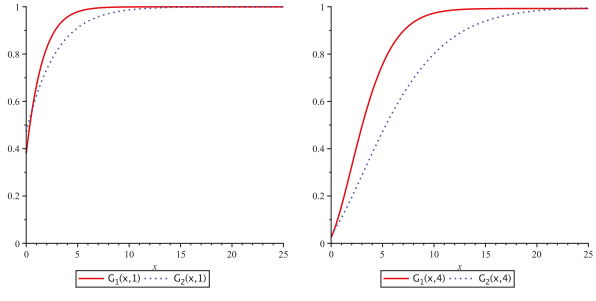
Figure 1. Variances of $S(t)$, $S_1(t)$ and $S_2(t)$ with initial state $J(0) = 1$.

Table 2. Covariances of discounted claims at t and $t+1$.

t	$J(0) = 1$		
	$\mathbf{Cov_1}(S(t), S(t+1))$	$\mathbf{Cov_1}(S_1(t), S_1(t+1))$	$\mathbf{Cov_1}(S_2(t), S_2(t+1))$
1	1.9327	1.7143	0.5328
2	3.9024	3.2169	1.7021
5	9.5545	7.3372	5.8448
10	17.0771	12.8965	11.4471
20	26.6637	20.2686	18.0782
30	31.9796	24.5961	21.2571
∞	40.3073	32.2449	23.8648

Table 3. Covariances of discounted claims at t and $t+5$.

t	$J(0) = 1$		
	$\mathbf{Cov_1}(S(t), S(t+5))$	$\mathbf{Cov_1}(S_1(t), S_1(t+5))$	$\mathbf{Cov_1}(S_2(t), S_2(t+5))$
1	0.8651	1.2213	0.4437
2	1.3181	2.0228	1.4775
5	3.4481	4.5219	5.2676
10	7.5945	8.5121	10.5187
20	15.2651	14.9882	16.9435
30	21.6039	19.7868	20.2186
∞	40.3073	32.2449	23.8648



(**a**) (**b**)

Figure 2. Distribution functions of $S(1)$ in (**a**) and $S(4)$ in (**b**) for $J(0) = 1, 2$.

Author Contributions: The two authors contribute equally to this article.

Acknowledgments: The authors would like to thank two anonymous reviewers for providing helpful comments and suggestions that improved the presentation of this paper. This research for Dr. Yi Lu was supported by the Natural Science and Engineering Research Council (NSERC) of Canada (grant number 611467).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Justification of Equation (5)

It is easy to see that the partial derivation of ${}_iL(\xi_1,\xi_2,\ldots,\xi_m;t)$ in Equation (2) with respect to ξ_i exists for $i \in E$. For its partial derivation with respect to t, we have

$$\begin{aligned}\frac{\partial\, {}_iL(\xi_1,\xi_2,\ldots,\xi_m;t)}{\partial t} &= -\lim_{h\to 0}\frac{{}_iL(\xi_1,\xi_2,\ldots,\xi_m;t) - {}_iL(\xi_1,\xi_2,\ldots,\xi_m;t+h)}{h}\\ &= -\lim_{h\to 0}\frac{1}{h}\mathbb{E}_i\left[e^{-\sum_{j=1}^m \xi_j S_j(t)}\left(1-e^{-\sum_{j=1}^m \xi_j (S_j(t+h)-S_j(t))}\right)\right].\end{aligned}$$

Now, under certain regularity conditions, we obtain

$$\begin{aligned}\lim_{h\to 0}\mathbb{E}_i\left[e^{-\sum_{j=1}^m \xi_j S_j(t)}\frac{1-e^{-\sum_{j=1}^m \xi_j (S_j(t+h)-S_j(t))}}{h}\right] &\le \lim_{h\to 0}\mathbb{E}_i\left[\frac{1-e^{-\sum_{j=1}^m \xi_j (S_j(t+h)-S_j(t))}}{h}\right]\\ &= \lim_{h\to 0}\frac{1-\mathbb{E}_i\left[e^{-\sum_{j=1}^m \xi_j (S_j(t+h)-S_j(t))}\right]}{h}.\end{aligned}$$

Let $\xi = \max(\xi_1,\xi_2,\ldots,\xi_m)$. Then

$$\begin{aligned}\lim_{h\to 0}\frac{1-\mathbb{E}_i\left[e^{-\sum_{j=1}^m \xi_j (S_j(t+h)-S_j(t))}\right]}{h} &\le \lim_{h\to 0}\frac{1-\mathbb{E}_i\left[e^{-\xi\sum_{j=1}^m (S_j(t+h)-S_j(t))}\right]}{h}\\ &= \lim_{h\to 0}\frac{1-\mathbb{E}_i\left[e^{-\xi(S(t+h)-S(t))}\right]}{h}\\ &= \lim_{h\to 0}\frac{1-\mathbb{E}_i\left[e^{-\xi S(h)}\right]}{h}.\end{aligned} \tag{A1}$$

Since

$$\begin{aligned}\mathbb{E}_i\left[e^{-\xi S(h)}\right] &= [1+d_{0,ii}h] + \sum_{k=1,k\neq i}^m d_{0,ik}h + \sum_{k=1}^m d_{1,ik}h\hat{f}_i\left(\xi e^{-\delta_i h}\right) + o(h)\\ &= 1 + \sum_{k=1}^m d_{0,ik}h + \sum_{k=1}^m d_{1,ik}h\hat{f}_i\left(\xi e^{-\delta_i h}\right) + o(h)\\ &= 1 + \sum_{k=1}^m (d_{0,ik}+d_{1,ik})h + \sum_{k=1}^m d_{1,ik}h\left[\hat{f}_i\left(\xi e^{-\delta_i h}\right) - 1\right] + o(h)\\ &= 1 + \sum_{k=1}^m d_{1,ik}h\left[\hat{f}_i\left(\xi e^{-\delta_i h}\right) - 1\right] + o(h),\end{aligned}$$

it follows from Equation (A1) that

$$\lim_{h\to 0}\frac{1-\mathbb{E}_i\left[e^{-\sum_{j=1}^m \xi_j (S_j(t+h)-S_j(t))}\right]}{h} \le \lim_{h\to 0}\frac{1-\mathbb{E}_i\left[e^{-\xi S(h)}\right]}{h} = \left(\sum_{k=1}^m d_{1,ik}\right)\left[1-\hat{f}_i(\xi)\right],$$

which justifies the existence of the partial derivation of ${}_iL(\xi_1,\xi_2,\ldots,\xi_m;t)$ with respect to t.

References

Asmussen, Søren. 2003. *Applied Probability and Queues*. New York: Springer.

Barges, Mathieu, Hélene Cossette, Stéphane Loisel, and Etienne Marceau. 2011. On the moments of the aggregate discounted claims with dependence introduced by a FGM copula. *ASTIN Bulletin* 41: 215–38.

Delbaen, Freddy, and Jean Haezendonck. 1987. Classical risk theory in an economic environment. *Insurance: Mathematics and Economics* 6: 85–116. [CrossRef]

Garrido, José, and Ghislain Léveillé. 2004. Inflation impact on aggregate claims. *Encyclopedia of Actuarial Science* 2: 875–78.

Kim, Bara, and Hwa-Sung Kim. 2007. Moments of claims in a Markov environment. *Insurance: Mathematics and Economics* 40: 485–97.

Jang, Ji-Wook. 2004. Martingale approach for moments of discounted aggregate claims. *Journal of Risk and Insurance* 71: 201–11. [CrossRef]

Leveille, Ghislain, and Franck Adekambi. 2011. Covariance of discounted compound renewal sums with stochastic interest. *Scandinavian Actuarial Journal* 2: 138–53. [CrossRef]

Léveillé, Ghislain, and Franck Adékambi. 2012. Joint moments of discounted compound renewal sums. *Scandinavian Actuarial Journal* 1: 40–55. [CrossRef]

Leveille, Ghislain, and Jose Garrido. 2001a. Moments of compound renewal sums with discounted claims. *Insurance: Mathematics and Economics* 28: 201–11.

Léveillé, Ghislain, and José Garrido. 2001b. Recursive moments of compound renewal sums with discounted claims. *Scandinavian Actuarial Journal* 2: 98–110. [CrossRef]

Léveillé, Ghislain, José Garrido, and Ya Fang Wang. 2010. Moment generating functions of compound renewal sums with discounted claims. *Scandinavian Actuarial Journal* 3: 165–84. [CrossRef]

Li, Shuanming. 2008. Discussion of "On the Laplace transform of the aggregate discounted claims with Markovian arrivals". *North American Actuarial Journal* 12: 208–10. [CrossRef]

Li, Jingchao, David CM Dickson, and Shuanming Li. 2015. Some ruin problems for the MAP risk model. *Insurance: Mathematics and Economics* 65: 1–8. [CrossRef]

Mohd Ramli, Siti Norafidah, and Jiwook Jang. 2014. Neumann series on the recursive moments of copula-dependent aggregate discounted claims. *Risks* 2: 195–210. [CrossRef]

Neuts, Marcel F. 1979. A versatile Markovian point process. *Journal of Applied Probability* 16: 764–79. [CrossRef]

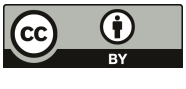
Ren, Jiandong. 2008. On the Laplace transform of the aggregate discounted claims with Markovian arrivals. *North American Actuarial Journal* 12: 198–206. [CrossRef]

Taylor, Gregory Clive. 1979. Probability of ruin under inflationary conditions or under experience rating. *ASTIN Bulletin* 10: 149–62. [CrossRef]

Wang, Ya Fang, José Garrido, and Ghislain Léveillé. 2018. The distribution of discounted compound PH-renewal processes. *Methodology and Computing in Applied Probability* 20: 69–96. [CrossRef]

Willmot, Gordon E. 1989. The total claims distribution under inflationary conditions. *Scandinavian Actuarial Journal* 1: 1–12. [CrossRef]

Woo, Jae-Kyung, and Eric CK Cheung. 2013. A note on discounted compound renewal sums under dependency. *Insurance: Mathematics and Economics* 52: 170–79. [CrossRef]



risks

MDPI

Article

A Maximal Tail Dependence-Based Clustering Procedure for Financial Time Series and Its Applications in Portfolio Selection

Xin Liu [1], Jiang Wu [2,*], Chen Yang [3] and Wenjun Jiang [4]

[1] School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai 200433, China; liu.xin@mail.shufe.edu.cn
[2] School of Economics, Central University of Finance and Economics, Beijing 100081, China
[3] Department of Insurance and Actuary, Wuhan University, Wuhan 430072, Hubei, China; cyang244@whu.edu.cn
[4] Department of Statistical and Actuarial Sciences, University of Western Ontario, London, ON N6A 5B7, Canada; wjiang66@uwo.ca
* Correspondence: jiang_wu@cufe.edu.cn; Tel.: +86-151-9902-2966

Received: 12 August 2018; Accepted: 27 September 2018; Published: 9 October 2018

Abstract: In this paper, we propose a clustering procedure of financial time series according to the coefficient of weak lower-tail maximal dependence (WLTMD). Due to the potential asymmetry of the matrix of WLTMD coefficients, the clustering procedure is based on a generalized weighted cuts method instead of the dissimilarity-based methods. The performance of the new clustering procedure is evaluated by simulation studies. Finally, we illustrate that the optimal mean-variance portfolio constructed based on the resulting clusters manages to reduce the risk of simultaneous large losses effectively.

Keywords: maximal tail dependence; clustering; financial time series; weighted cuts; copula

1. Introduction

It is of great interest in identifying the risk of simultaneous large losses in portfolio selection and financial risk management. If this type of risk is identified properly, candidate assets could be grouped such that asset prices or returns from different groups are unlikely to drop simultaneously. An investment strategy is called portfolio diversification if the portfolio is constructed by selecting one asset from each group. As we could see, the performance of portfolio diversification depends on how the assets are grouped.

In general, the observed prices or returns of assets are essentially time series. To group assets properly, time series clustering techniques are usually involved. Early works on time series clustering include interdependence measure between asset returns such as the (Pearson or Spearman type) cross-correlation coefficients (cf. Kaufman and Rousseeuw 1990). In particular, Mantegna (1999) and Bonanno et al. (2004) quantified the degree of interdependence between the synchronous time evolution of a pair of stock prices and used it in financial time series clustering. Moreover, as another extension of dependence-based method, Baragona (2001) and Brockwell and Davis (2002) developed a new measure of interdependence from the residuals obtained by fitting the data to acceptable time series. In addition, inspired by the dynamic conditional correlation (DCC) model developed by Engle and Sheppard (2001) and Engle (2002), Billio et al. (2006) and Billio and Caporin (2009) proposed the Flexible Dynamic Conditional Correlation (FDCC) multivariate GARCH model and provided an estimate of the dynamics of correlation coefficients within groups of financial assets for asset allocations.

However, cross-correlation coefficients do not always guarantee a sufficient degree of portfolio diversification because these coefficients cannot always capture the possible extreme co-movements of

asset returns in lower tails. Extreme co-movement of asset returns in lower tails plays an important role in studying contagion of financial crisis. Bae et al. (2003) provided evidence of the existence of extreme co-movements in terms of coexceedances when studying the phenomenon of contagion. More formally, the contagion of financial crisis could be defined directly as a significant increase of extreme co-movements if financial crisis occurs in one of the markets (cf. Pericoli and Sbracia 2003, Definition 4). Hence, if a portfolio diversification arrangement fails to diversify the risk of extreme co-movement in the lower tail, it might be vulnerable to the contagion of financial crisis occurring in other markets.

Even when there is no contagion, extreme co-movements of asset returns may also exist due to the similarity of fundamentals from the traditional point of view, investor trading patterns(Barberis et al. 2005), or incomplete information (Veldkamp 2006). To diversify the risk of extreme co-movement in lower tail, De Luca and Zuccolotto (2011) proposed a dissimilarity measure based on tail dependence coefficients (TDC) instead of cross-correlation coefficients to obtain homogeneous groups of time series with an association between extreme low values. Inspired by this work, Durante et al. (2014) developed a time series clustering procedure with a conditional version of Spearman's correlation coefficient for extremely low values introduced by Durante et al. (2014), and a non-parametric estimator of tail dependence provided in Durante et al. (2015). De Luca and Zuccolotto (2015) further proposed a dynamic clustering procedure so that the coefficient employed to measure the lower tail dependence can be time-varying on the basis of historical market volatility.

In this paper, we propose to cluster time series via the coefficients of maximum tail dependence introduced by Furman et al. (2015). The coefficients of maximal tail dependence are direct extensions of TDCs including the tail dependence coefficient λ, the weak tail dependence coefficient χ and the tail order κ. The major difference is that the coefficients of maximal tail dependence are calculated with convergence paths that are possibly other than the diagonal path. As a result, the matrix of coefficients of maximal tail dependence may not be symmetric and thus cannot be used as a similarity (or dissimilarity) matrix in clustering procedures. Instead, such a matrix may be seen as a type of affinity matrix representing directed relations between assets.

The paper is organized as follows. Section 2 is a brief introduction of the coefficients of maximal tail dependence. The proposed clustering procedure of time series is formally described in Section 3. The performance of the proposed procedure is evaluated in Section 4. An application to real exchange rates of G20 countries is presented and analyzed in Section 5. Section 6 concludes.

2. The Coefficients of Maximal Tail Dependence

Several coefficients have been introduced by researchers to measure the extreme co-movements in recent years. For example, one of the most important measures is the lower (upper) tail dependence, which is formally defined by

$$\begin{cases} \lambda_L := \lim\limits_{u\to 0+} \mathbf{P}\left(X \le F_X^{-1}(u) | Y \le F_Y^{-1}(u)\right) \\ \lambda_U := \lim\limits_{u\to 1-} \mathbf{P}\left(X > F_X^{-1}(u) | Y > F_Y^{-1}(u)\right) \end{cases},$$

where random variables X and Y represent the potentially dependent risks. Since by Sklar's Theorem (cf. Nelsen 2006) there is a uniquely determined copula function $C : [0,1]^2 \to [0,1]$ such that

$$F_{X,Y}(x,y) = C\left(F_X(x), F_Y(y)\right),$$

the lower (upper) tail dependence could be defined as the limiting point of a functional of the copula function, namely,

$$\lambda_L := \lambda_L(C) = \lim_{u\to 0+} \frac{C(u,u)}{u}, \quad \lambda_U := \lambda_U(C) = \lim_{u\to 0+} \frac{\hat{C}(u,u)}{u},$$

where $\hat{C}$ is the survival copula with respect to C. Apart from the lower (upper) tail dependence, similar measures include the weak lower (upper) tail dependence

$$\chi_L := \chi_L(C) = \lim_{u\to 0+} \frac{2\log u}{\log C(u,u)} - 1, \quad \chi_U := \chi_U(C) = \lim_{u\to 0+} \frac{2\log u}{\log \hat{C}(u,u)} - 1,$$

(cf. Coles et al. 1999) and lower (upper) tail order κ_L (κ_U) defined via:

$$C(u,u) = \ell_L(u)u^{\kappa_L}, \quad \hat{C}(1-u,1-u) = \ell_U(u)u^{\kappa_U}, \quad u \in (0,1)$$

where ℓ_L and ℓ_U are slowly varying functions of u at 0 (cf. Hua and Joe 2011).

The aforementioned measures of tail dependence are all limiting values of functionals of C as the arguments (u,v) shrink to $(0,0)$ along the diagonal line of the square $[0,1]^2$. However, as pointed out by Furman et al. (2015), these measures may sometimes underestimate the extent of extreme co-movements for dependent risks, and, for this reason, the authors proposed improved versions of these coefficients of tail dependence, named as the *coefficients of the maximal tail dependence*, which are more sensitive to extreme co-movements. Accordingly, a clustering procedure based on such coefficients may provide better clustering results than those based on other coefficients of tail dependence such as those proposed by De Luca and Zuccolotto (2011) and Durante et al. (2014), and the portfolios constructed based on such clustering results may also outperform.

The *coefficients of the maximum tail dependence* are limiting values of the usual versions of the corresponding functionals of C (or $\hat{C}$) converging to the lower-left (or upper-right) vertex along paths of maximal tail dependence. To formally define the paths of maximal tail dependence, consider a function $\varphi : [0,1] \mapsto [0,1]$ satisfying the following admissible conditions (see Furman et al. 2015, Definition 2.1):

1. $\varphi(u) \in [u^2, 1]$ for every $u \in [0,1]$; and
2. both $\varphi(u)$ and $u^2/\varphi(u)$ converge to 0 when $u \downarrow 0$.

The collection of such kind of functions is called the admissible set, denoted as $\mathcal{A}$. Then, a path $\left(\varphi(u), u^2/\varphi(u)\right)_{0\leqslant u\leqslant 1}$ shrinking to the lower-left (or upper-right) vertex is called admissible whenever φ belongs to $\mathcal{A}$. Specifically, the diagonal path used to define the usual coefficients of tail dependence $(u,u)_{0\leqslant u\leqslant 1}$ is admissible as the function $\varphi_0(u) = u$, $u \in [0,1]$ is admissible. According to (Furman et al. 2015, Definition 2.2), the paths of maximal tail dependence is denoted as $\left(\varphi^*(u), u^2/\varphi^*(u)\right)_{0\leqslant u\leqslant 1}$ where

$$\varphi^*(u) = \arg\max_{\varphi\in\mathcal{A}} C\left(\varphi(u), u^2/\varphi(u)\right).$$

To simplify the notations, we denote $\Pi^*(u) = C\left(\varphi^*(u), u^2/\varphi^*(u)\right)$ if the optimal value exists. Then, the lower tail maximal dependence (LTMD) is defined via:

$$\lambda_L^* := \lambda_L^*(C) = \lim_{u\to 0+} \frac{\Pi^*(u)}{u},$$

the weak lower tail maximal dependence (WLTMD) is defined via:

$$\chi_L^* := \chi_L^*(C) = \lim_{u\to 0+} \frac{2\log u}{\log \Pi^*(u)} - 1,$$

and the order of lower tail maximal dependence is defined via:

$$\Pi^*(u) = \ell_L^*(u)u^{\kappa_L^*}, \quad u \in (0,1)$$

where ℓ_L^* is a slowly varying function of u at 0. In particular, for κ_L^*, we have the following result similar to that of the usual tail order κ_L (cf. Hua and Joe 2011):

Proposition 1. *For any bivariate copula function C, if $\varphi^*(u) \in \mathcal{A}$ exists, then the corresponding index of maximal tail dependence $\kappa_L^* \in [1, 2]$.*

The proof is given in Appendix A. As a result, $\chi_L^* \in [0, 1]$ may also not be a desirable affinity measure for common clustering procedures such as k means clustering or hierarchical clustering. However, χ_L^* is a desirable weight for graphs. The larger χ_L^* of two assets is, the stronger the extreme co-movement between these assets is, and hence a bigger weight is posed by χ_L^*.

We refer to Furman et al. (2015) for examples of expressions for λ_L^* and κ_L^* with closed forms in the case of parametric families of distributions. Notably, Furman et al. (2016) proved that, in the Gaussian case, the classical and maximal tail dependence coefficients coincide. In the present paper, however, to speed up practical calculations, we resort to non-parametric approach in the following sections. In particular, we find a clustering procedure based on χ_L^* to be very attractive.

3. Clustering Procedure

Typically, clustering based on dissimilarity matrices such as given by De Luca and Zuccolotto (2011, sct. 3) could be achieved through the hierarchical clustering method directly. However, to cluster using the affinity matrix constructed with χ_L^*, we could not use the hierarchical clustering method because the affinity matrix may not be symmetric. Hence, we have to consider graph based clustering procedures.

Suppose we have n assets in total available for a portfolio construction. Then, the affinity matrix $\mathbf{\Delta} = \left(\mathbf{\Delta}_{ij}\right)_{n\times n}$ is given by

$$\mathbf{\Delta}_{ij} = \chi_{L,ij}^* = \frac{2}{\kappa_{L,ij}^*} - 1 \tag{1}$$

(cf. Furman et al. 2015, sct. 5). Theoretically, $\mathbf{\Delta}$ should be a symmetric matrix which could be seen as an affinity matrix consisting of edge weights for an undirected graph and thus could be used for clustering with the hierarchical clustering method. This is because $\Pi^*(u)$ is unique as long as it exists and thus $\kappa_{L,ij}^* = \kappa_{L,ji}^*$. However, for those asymmetric copulas such as the unexchangeable Marshall–Olkin copulas

$$C_{a,b}^{\mathrm{MO}}(u, v) = \min\{u^{1-a}v, uv^{1-b}\}, \quad a, b \in (0, 1), \quad a \neq b$$

the estimated parameters may differ due to that for two series of observations there are actually two different copulas to be chosen for the parameter estimation procedure: $C_{a,b}^{\mathrm{MO}}(u, v)$ and $C_{b,a}^{\mathrm{MO}}(u, v) = C_{a,b}^{\mathrm{MO}}(v, u)$. In other words, for two arbitrary series of observations, it is impossible to determine which group should be regarded as "u" and the other group as "v" in practice, even though we are sure that these observations are generated from an unexchangeable Marshall–Olkin copula.

For simplicity, when constructing the affinity matrix $\mathbf{\Delta}$, we keep only one of the two possible copulas whenever we have to estimate the parameters of the copula from pairwise observations. The advantage of this idea is that, taking $C_{a,b}^{\mathrm{MO}}(u, v)$ for instance, the lower triangle part of $\mathbf{\Delta}$ is calculated by assuming the pairwise observations are generated from $C_{a,b}^{\mathrm{MO}}(u, v)$ and then estimating the parameters a and b while the upper triangle part of $\mathbf{\Delta}$ is actually calculated by assuming the same pairwise observations are generated from $C_{b,a}^{\mathrm{MO}}(u, v) = C_{a,b}^{\mathrm{MO}}(v, u)$ and then estimating the parameters a and b. In other words, the resulting affinity matrix $\mathbf{\Delta}$ contains information of estimated parameters from both $C_{a,b}^{\mathrm{MO}}(u, v)$ and $C_{b,a}^{\mathrm{MO}}(u, v) = C_{a,b}^{\mathrm{MO}}(v, u)$ in fact.

Since such an affinity matrix may not necessarily be symmetric, the hierarchical clustering method fails to work. In this case, the resulting affinity matrix could be considered as a matrix of weighted edges in a directed graph instead of an undirected graph. The clustering task could be achieved with

the *Weighted Normalized Cuts* (*WNACut* for short) introduced by Meilă and Pentney (2007), which is initially developed to analyze directed graphs related to the link data.

To understand the *WNACut* method, let $\mathcal{C} = \{C_1, \ldots, C_K\}_{1 \le K \le n}$ be an arbitrary partition of the set of all assets. Then, the *cut* of C_k to $C_{k'}$ represents the total influence of Cluster C_k on Cluster $C_{k'}$, namely,

$$Cut(C_k, C_{k'}) = \sum_{i \in C_k} \sum_{j \in C_{k'}} \mathbf{\Delta}_{ij}. \tag{2}$$

Hence, the *total weighted cut* of all clusters is defined via:

$$WCut(\mathcal{C}) = \sum_{k=1}^{K} \sum_{k' \neq k} \frac{Cut(C_k, C_{k'})}{\sum_{i \in C_k} V_i},$$

where $V_i = \sum_{j=1}^{n} \mathbf{\Delta}_{ij}$ for $i = 1, \ldots, n$. Then, our target cluster $\mathcal{C}^*$ is such that

$$WCut(\mathcal{C}^*) = \min_{\mathcal{C}} WCut(\mathcal{C}).$$

This optimization problem could be solved through a spectral clustering algorithm named "Best*WCut*". Similar to the clustering procedure proposed by De Luca and Zuccolotto (2011), this algorithm is also a two-stage clustering procedure, in which the *non-metric multidimensional scaling (MDS)* in the first stage is substituted by a process that transforms the asymmetric affinity matrix consisting of the WLTMDs into k orthonormal columns, where k is the predetermined number of clusters. The details are given in Algorithm A1. In Meilă and Pentney (2007), the *WNACut* method is shown to consistently outperform all other clustering methods chosen to be tested with synthetic data in their experiments. For this reason, we adopt this method to finish the clustering task in our proposed procedure.

4. A Simulation Study of Synthetic Data

As mentioned, our proposed clustering procedure is a two-stage procedure, as shown in Algorithm A1. However, there is no information related to the choice of clustering method for the second stage in this particular situation revealed. Hence, a simulation study is designed in this section to compare the performance of the second stage clustering method in *WNACut* with that of a list of commonly used clustering methods, including the classical k-means method and hierarchical clustering procedure with Ward's minimum variance method (considering both Ward's criterion and Ward's criterion squared, the results are denoted as Ward.D2 and Ward.D, respectively), single linkage method, complete linkage method, average linkage method, McQuitty's linkage method, median linkage method, and centroid linkage method. The performances are measured in two metrics: the *misclassification error* (ME) described in Verma and Meilă (2003) and *variation in information* (VI) introduced in Meilă (2003). Both criteria tend to be smaller if the resulting clusters are more similar to the K known clusters.

To begin with, we assume there are K different known clusters whose numbers of elements are $N_1, \ldots, N_K$, respectively. The dependence structure employed to generate the realizations in each cluster is a particular case of the asymmetric multivariate copula given by Liebscher (2008, eq. 3); namely, for cluster k, we have

$$C_k(u_{k,1}, \ldots, u_{k,n_1}) = u_{k,1}^{\gamma_1/(\gamma_0+\gamma_1)} \left(u_{k,1}^{-1/(\gamma_0+\gamma_1)} + \sum_{i=2}^{n_1} \left(u_{k,i}^{-1/\gamma_0} - 1 \right) \right)^{-\gamma_0},$$

where $\gamma_0 > 0$ and $\gamma_1 \geq 0$. Then, the joint CDF of the distribution used to generate the realizations is given by

$$C(u_{1,1},\ldots,u_{1,n_1},\ldots,u_{K,1},\ldots,u_{K,n_K};\gamma_0,\gamma_1)=\prod_{k=1}^{K}C_k(u_{k,1},\ldots,u_{k,n_1}).$$

Notice that, with the above model settings, the pairwise marginal copulas could be written as either

$$C(u,v)=u^{\gamma_1/(\gamma_0+\gamma_1)}\left(u^{-1/(\gamma_0+\gamma_1)}+v^{-1/\gamma_0}-1\right)^{-\gamma_0},\quad \gamma_1\in[0,\infty], \tag{3}$$

or

$$C(u,v)=v^{\gamma_1/(\gamma_0+\gamma_1)}\left(v^{-1/(\gamma_0+\gamma_1)}+u^{-1/\gamma_0}-1\right)^{-\gamma_0},\quad \gamma_1\in[0,\infty], \tag{4}$$

which includes both the pairwise independent copula $C(u,v)=uv$ as the particular case $\gamma_1=\infty$ and the pairwise standard Clayton copula $C(u,v)=\left(u^{-1/\gamma_0}+v^{-1/\gamma_0}-1\right)^{-\gamma_0}$ as the particular case $\gamma_1=0$. Then, as discussed in Section 3, we only keep (3) for further analysis. When the realizations are generated, the affinity matrix consisting of pairwise WLTMDs could be calculated by first estimating γ_0 and γ_1 with the maximal likelihood method based on (3) and then calculating the WLTMDs using

$$\hat{\kappa}^*_{L,ij}=1+\hat{\gamma}_{1,ij}/(\hat{\gamma}_{1,ij}+2\hat{\gamma}_{0,ij})$$

(cf. Furman et al. 2015, eq. 6.2). Therefore, with all pairwise $\hat{\kappa}^*_{L,ij}$ obtained, the affinity matrix $\mathbf{\Delta}$ could be calculated through Equation (1). Then, given the predetermined number of clusters k, the first stage of the Best*WCut* will transform the affinity matrix $\mathbf{\Delta}$ into k orthonormal columns.

The total number of iterations for our simulation study is 500, which is the same as De Luca and Zuccolotto (2011). Other values of the parameters for the simulation study are given below:

- The number of known clusters $K=4$.
- The number of objects in each cluster $n_1,\ldots,n_K$ are independently sampled from $\{3,4,\ldots,8\}$ at random for each iteration.
- $\gamma_0=4$.
- $\gamma_1=1,8,64$, which result in theoretical $\kappa^*_L=10/9,1.5,17/9$, respectively, if the two series of realizations are neither independent nor dependent with a classical Clayton copula.
- The distances used in hierarchical clustering methods are all Euclidean distance in $\mathbb{R}^K$.

The results of MEs and VIs are given in Tables 1 and 2, respectively, in which we can see that, with Ward's criterion squared, the Ward's minimum variance method is consistently competitive or even outperforms other methods. Moreover, the distributional properties of the simulated ME's are shown in Figures 1–3 for $\gamma_1=1$, $\gamma_1=8$, and $\gamma_1=64$, respectively, and the distributional properties of the simulated VIs are shown in Figures 4–6 for $\gamma_1=1$, $\gamma_1=8$, and $\gamma_1=64$, respectively. All these figures indicate that the Ward's minimum variance method might be the best choice for the second stage of the proposed clustering procedure.

The results shown in Tables 1 and 2 also indicate the sensitivity of the ME/VI to κ^*_L. Taking the Ward's minimum variance method as an example, when κ^*_L increases or decreases by $(3.5/9/1.5=)$ 25.93% (i.e., γ_1 increases from 8 to 64 or decreases from 8 to 1, correspondingly) the average ME increases by 179.25% or decreases by 19.57%, correspondingly, while the average VI increases by 79.8% or decreases by 38.39%, correspondingly. Thus, our proposed clustering procedure seems to perform better as the true WLTMD κ^*_L gets closer to 1, and as the true WLTMD κ^*_L moves towards 2, the performance of our proposed clustering procedure worsens rapidly.

Table 1. Means and variances of simulated MEs using various clustering methods in the second stage of *WNACut* for $\gamma_1 = 1$, $\gamma_1 = 8$ and $\gamma_1 = 64$, respectively.

	$\gamma_1 = 1$		$\gamma_1 = 8$		$\gamma_1 = 64$	
	Mean	**Var**	**Mean**	**Var**	**Mean**	**Var**
k means	0.1269	0.0143	0.1795	0.0132	0.3875	0.0109
ward.D	0.1089	0.0095	0.1354	0.0012	0.3781	0.0113
ward.D2	0.1206	0.0117	0.1326	0.0014	0.3913	0.0111
single	0.3341	0.0271	0.1477	0.0037	0.5311	0.0089
complete	0.1715	0.0205	0.1349	0.0034	0.4244	0.0121
average	0.1988	0.0242	0.1329	0.0021	0.4582	0.0127
McQuitty	0.1995	0.0217	0.1345	0.0023	0.4536	0.0124
median	0.3430	0.0324	0.1422	0.0033	0.5201	0.0088
centroid	0.3223	0.0377	0.1349	0.0019	0.5284	0.0091

Table 2. Means and variances of simulated VIs using various clustering methods in the second stage of *WNACut* for $\gamma_1 = 1$, $\gamma_1 = 8$ and $\gamma_1 = 64$, respectively.

	$\gamma_1 = 1$		$\gamma_1 = 8$		$\gamma_1 = 64$	
	Mean	**Var**	**Mean**	**Var**	**Mean**	**Var**
k means	0.6323	0.2236	1.0234	0.1347	1.8849	0.2088
ward.D	0.6317	0.1959	1.0254	0.0346	1.8437	0.2202
ward.D2	0.6486	0.2043	1.0085	0.0462	1.8738	0.2095
single	1.1335	0.2906	1.0446	0.0383	1.9781	0.1007
complete	0.7914	0.2827	0.9977	0.0659	1.9327	0.1998
average	0.8225	0.2750	1.0000	0.0551	1.8980	0.1622
McQuitty	0.8285	0.2620	1.0085	0.0509	1.9295	0.1804
median	1.2115	0.3518	1.0340	0.0405	1.9935	0.1115
centroid	1.1504	0.3994	1.0158	0.0484	1.9783	0.1013

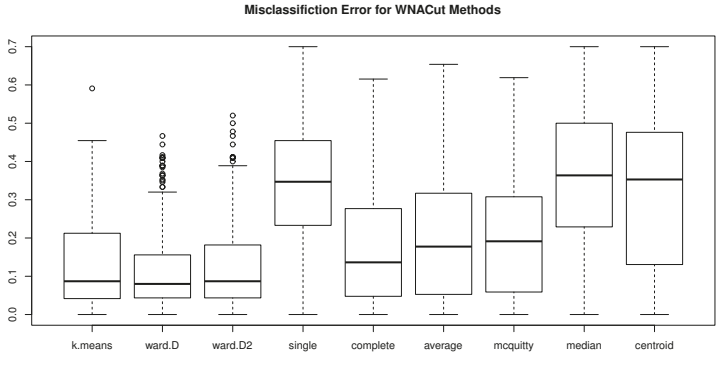


Figure 1. Box plots of simulated MEs using various clustering methods in the second stage of *WNACut* for $\gamma_1 = 1$.

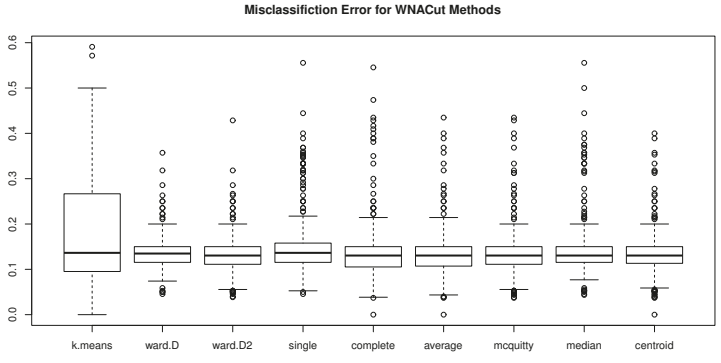


Figure 2. Box plots of simulated MEs using various clustering methods in the second stage of *WNACut* for $\gamma_1 = 8$.

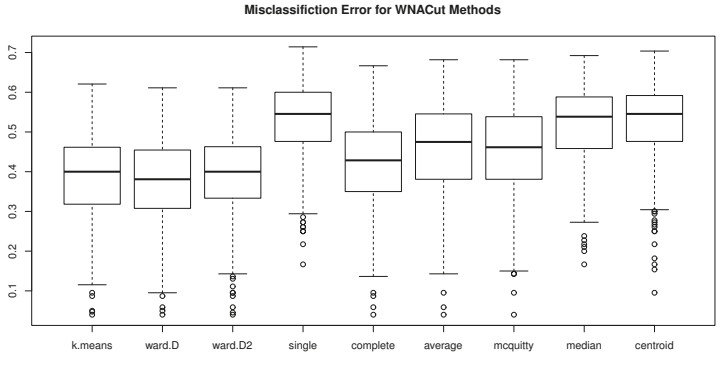


Figure 3. Box plots of simulated MEs using various clustering methods in the second stage of *WNACut* for $\gamma_1 = 64$.

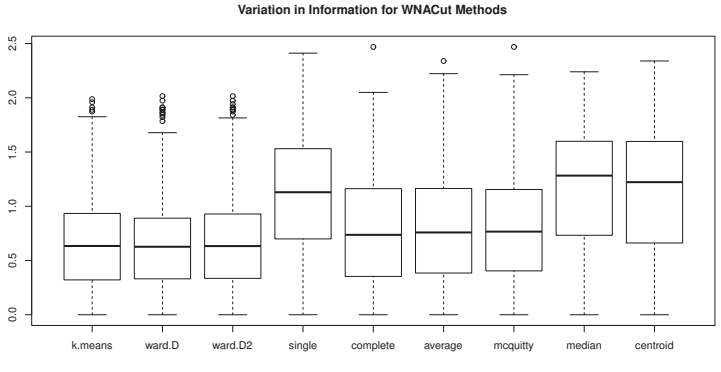


Figure 4. Box plots of simulated VIs using various clustering methods in the second stage of *WNACut* for $\gamma_1 = 1$.

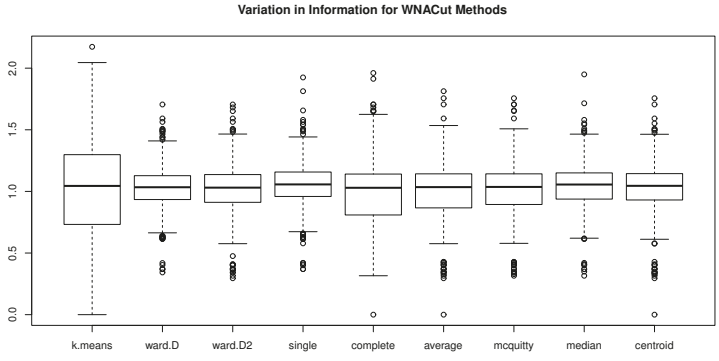


Figure 5. Box plots of simulated MEs for all hierarchical clustering methods when $\gamma = 8$.

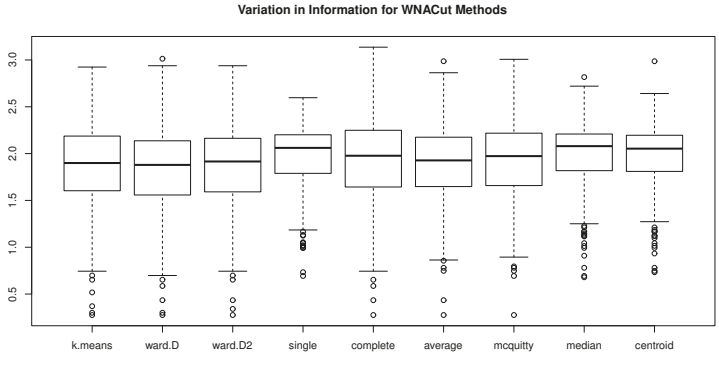


Figure 6. Box plots of simulated VIs using various clustering methods in the second stage of *WNACut* for $\gamma_1 = 64$.

5. Application to Real Data

In this section, we apply the proposed procedure on the log-returns of foreign exchange rates with respect to US dollars from Group of 20 (known as G20). Since US dollar is used as the underlying currency, there are 19 time series of exchange rates in total. The exchange rates of France, Germany and Italy are excluded from our analysis due to their perfect linear correlation with euros, which results in 16 time series of exchange rates available for our clustering analysis[1]. The data were collected weekly from 5 September 2012 to 17 August 2016, covering 207×16 active observations during these four years, which can be downloaded from "PACIFIC Exchange Rate Service", 2016, by Werner Antweiler, University of British Columbia. The descriptive statistics and histograms of observations for the 16 currencies are given in Table 3 and Figure 7.

[1] The currencies include: Argentine Pesos, Australian Dollars, Brazilian Reals, British Pounds, Canadian Dollars, Chinese Renminbi, European Euros, Indian Rupees, Indonesian Rupiah, Japanese Yen, Mexican Pesos, Russian Rubles, Saudi Arabian Riyal, South African Rand, South Korean Won, and Turkish New Lira.

Table 3. Descriptive statistics of (log)returns of 16 selected currencies.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std
ARS	−0.1466	−0.0048	−0.0026	−0.0056	−0.0010	0.0415	0.0112
AUD	−0.0294	−0.0085	−0.0005	−0.0014	0.0060	0.0295	0.0175
BRL	−0.0584	−0.0128	−0.0012	-0.0022	0.0086	0.0474	0.0086
CAD	−0.0239	−0.0071	−0.0008	−0.0013	0.0038	0.0228	0.0028
CNY	−0.0211	−0.0011	0.0000	-0.0002	0.0010	0.0079	0.0106
EUR	−0.0379	−0.0067	−0.0006	−0.0006	0.0058	0.0268	0.0087
GBP	−0.0840	−0.0063	−0.0005	−0.0010	0.0052	0.0304	0.0093
IDR	−0.0366	−0.0059	−0.0015	−0.0015	0.0016	0.0502	0.0118
INR	−0.0408	−0.0050	−0.0010	−0.0009	0.0037	0.0409	0.0089
JPY	−0.0482	−0.0078	−0.0008	−0.0012	0.0051	0.0409	0.0110
KRW	−0.0273	−0.0055	0.0004	0.0001	0.0059	0.0266	0.0245
MXN	−0.0318	−0.0077	−0.0012	−0.0016	0.0052	0.0280	0.0158
RUB	−0.1215	−0.0124	−0.0022	−0.0034	0.0080	0.1253	0.0126
SAR	−0.0052	−0.0008	0.0000	−0.0000	0.0008	0.0045	0.0097
TRY	−0.0484	−0.0099	−0.0006	−0.0023	0.0057	0.0346	0.0112
ZAR	−0.0487	−0.0130	−0.0017	−0.0023	0.0079	0.0437	0.0175

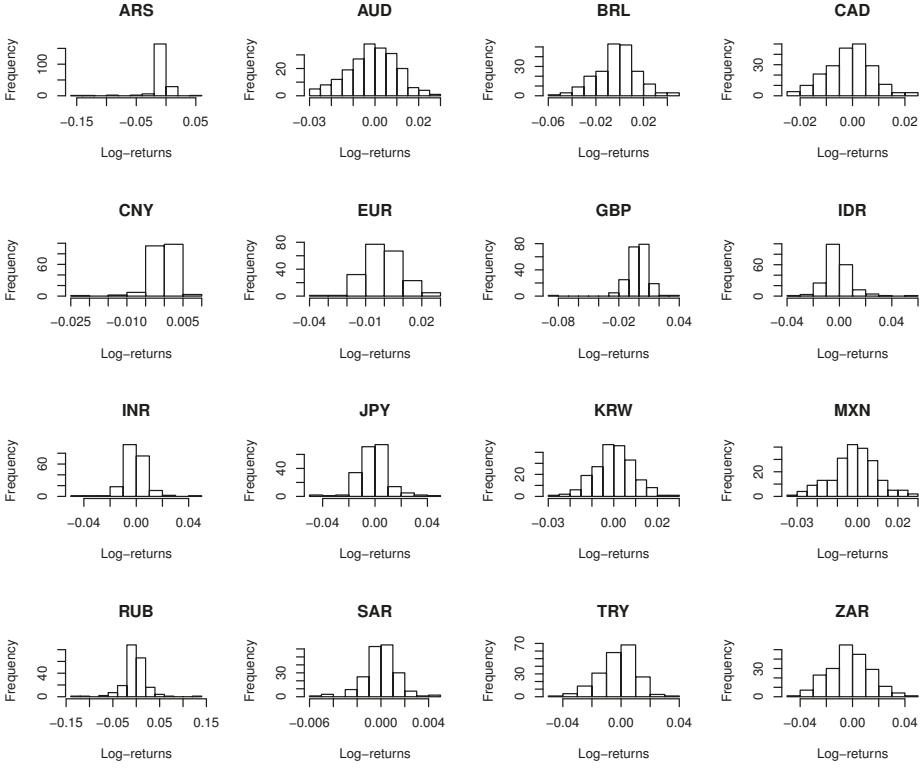


Figure 7. Histograms for the 16 selected currencies.

5.1. Preliminary Analysis

For each series of the exchange rates, the log-returns are obtained by taking logarithm of the fraction between two consecutive weekly exchange rates, part of which are shown in Figure 8. To eliminate the potential autocorrelation and heteroskedasticity of the log-returns, a univariate

generalized error distribution (GED) ARMA-GARCH $(1,1)$ model is applied to each time series of log-returns and the fitted standardized residuals are extracted for the purpose of clustering.

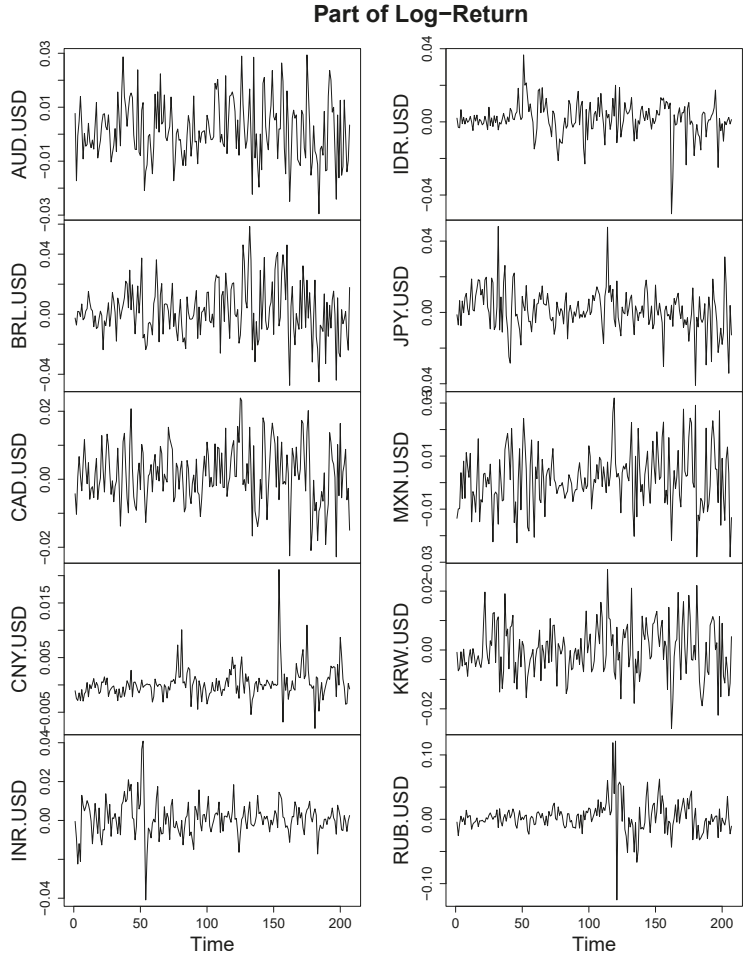


Figure 8. Log return of exchange rates for some of the G20 members.

A preliminary test of correlation indicates that the standardized residuals of Argentine Pesos seem to be uncorrelated to those of all the other currencies. The details of the correlation tests are shown in Table 4, from which we could see that all p-values are greater than 0.05. Thus, we suspect that all WLTMDs between the residuals of ARS and any other currency are approximately 0. To this end, let $T_{X,i} = \min\{R_{ARS,i}, R_{X,i}\}$ where X represents some currency and $R_{X,i}$ is the random variable having the same distribution as the standard residuals obtained by fitting the log-returns of currency X with GED ARMA-GARCH$(1,1)$ model. Then, by Coles et al. (1999, eq. 4.2), we have

$$\mathrm{P}\left(T_{X,i} > t\right) \sim \ell_X(t) t^{-\eta_X}$$

as $t \to \infty$ where $\eta_X = 2/(1 + \chi_{L,ARS}(X))$ and $\chi_{L,ARS}(X)$ is the weak lower-tail dependence between the standard residuals of the log-returns of ARS and currency X. Hence, our concern about whether $\chi_{L,ARS}(X) = 0$ is equivalent to test

$$H_0 : \eta_X = 2 \quad \text{v.s.} \quad H_1 : \eta_X \neq 2.$$

The test statistic, which is asymptotically standard normal, could be obtained using the improved OLS method given in Gabaix and Ibragimov (2011). When 50 pairs of observations are used, the resulting p-values are all greater than 0.1, as shown in Table 4. Thus, H_0 could not be rejected, which leads to $\hat{\eta}_{GI} \approx 2$ and hence $\hat{\chi}_{L,ARS}(X) \approx 0$ for all currencies other than ARS. Notice that by definition WLTMD should always be smaller than or equal to the corresponding weak lower-tail dependence, therefore the WLTMDs between the residuals of ARS and those of any other currency are approximately 0, which allows us to evaluate ARS as an isolated point that will not be taken into account in further analysis.

Table 4. The p-values of correlation tests between the standardized residuals of ARS and those of the other 15 currencies based on Pearson's product moment correlation coefficient, Kendall's τ and Spearman's ρ, respectively. The last column contains the p-values of the test of H_0 using the OLS test statistic.

Currency	Pearson	Kendall	Spearman	$\hat{\eta}_{GI}$
AUD	0.3437	0.2111	0.2248	0.9268
BRL	0.0530	0.6239	0.6545	0.5408
CAD	0.0922	0.1932	0.2157	0.6553
CNY	0.1825	0.9444	0.9718	0.1550
GBP	0.3954	0.8485	0.8263	0.8356
INR	0.6418	0.2816	0.3278	0.8877
IDR	0.3542	0.3807	0.3855	0.2402
JPY	0.6390	0.2889	0.2957	0.3472
KRW	0.7797	0.9702	0.9660	0.6733
MXN	0.8750	0.8962	0.9712	0.4198
RUB	0.4855	0.9944	0.9739	0.3999
SAR	0.7583	0.5040	0.4771	0.9567
ZAR	0.8412	0.1508	0.1746	0.7425
TRY	0.2061	0.0696	0.0741	0.6147
EUR	0.4067	0.2103	0.2173	0.6303

In fact, ARS is not the only currency that should be excluded from further clustering procedure. When testing the correlation between standardized residuals of the fitted GED ARMA-GARCH $(1, 1)$ model for log-returns, we discover that the residuals of SAR have negative correlation with majority of the residuals of the other currencies (see the columns entitled "Sign of estimated coefficients" in Table 5). Furthermore, we can see in Table 5 that only AUD, CNY, GBP and JPY are currencies whose standard residuals may not be negatively correlated with SAR. Unfortunately, none of the corresponding p-values show significant correlation between the standardized residuals of these currencies and SAR. Therefore, it is reasonable to suspect the WLTMDs between the residuals of SAR and any other currency are approximately 0. As expected, the last column in Table 5 verifies this result using the OLS test statistics with only 30 pairs of observations. Therefore, we do not include SAR in our next step of clustering procedure either.

Table 5. The sign of Estimated Correlation Coefficients and p-values of correlation tests between the standardized residuals of SAR and those of the other 14 currencies based on Pearson's product moment correlation coefficient, Kendall's τ and Spearman's ρ, respectively. The last column is the p-values of the test of H_0 using the OLS test statistic.

Currency	Sign of Estimated Coefficients			p-Values			
	Pearson	Kendall	Spearman	Pearson	Kendall	Spearman	$\hat{\eta}_{GI}$
AUD	+	−	−	0.8734	0.7981	0.8015	0.6997
BRL	−	−	−	0.0021	0.0007	0.0008	0.9334
CAD	−	−	−	0.2140	0.0641	0.0514	0.5117
CNY	−	+	+	0.8798	0.7593	0.7691	0.9141
GBP	+	−	−	0.2583	0.8755	0.8342	0.8694
INR	−	−	−	0.0592	0.0240	0.0213	0.6743
IDR	−	−	−	0.8543	0.7563	0.7676	0.9509
JPY	+	+	+	0.4373	0.5078	0.5216	0.1236
KRW	−	−	−	0.9862	0.9492	0.9391	0.9861
MXN	−	−	−	0.0043	0.0044	0.0041	0.4444
RUB	−	−	−	0.0290	0.0022	0.0021	0.2088
ZAR	−	−	−	0.0899	0.0283	0.0233	0.7497
TRY	−	−	−	0.1121	0.0549	0.0471	0.4727
EUR	−	−	−	0.5123	0.3963	0.4209	0.7975

Since there are no more currencies that could be evaluated as isolated points based on correlation tests or OLS test statistics, we retain all of the rest 14 exchange rates in our further clustering analysis.

5.2. Clustering the Exchange Rates Using WLTMD

For each pair of the residuals fitted from the log-return series, a bivariate Clayton copula function (see Liebscher 2008, for instance) is adopted with the form of Equation (3) on the estimated empirical distribution functions (pseudo observations), where $\gamma_0 \geq 0$ and $\gamma_1 \geq 0$ are unknown parameters to be specified. Notice that the Clayton-type copula defined in Equation (3) relaxes the restriction $\gamma_1 = 0$ in the classical Clayton copula and hence is not necessarily symmetric. With this Clayton-type copula, the tail order of maximal dependence is proved to have the following analytic form:

$$\kappa_L^* = 1 + \gamma_1/(\gamma_1 + 2\gamma_0)$$

(cf. Furman et al. 2015, eq. 6.2). The parameters can be estimated by the maximum likelihood method, and hence the estimate of the lower tail of the maximal dependence is given by $\hat{\kappa}_L^* = 1 + \hat{\gamma}_1/(\hat{\gamma}_1 + 2\hat{\gamma}_0)$. To measure the goodness-of-fit of the assumed copula, we employ a test based on the Rosenblatt transformation (see Breymann et al. 2003) between two dependent random variables Y_1 and Y_2, given by

$$S(Y_1, Y_2) = \left[\Phi^{-1}\left(F_1(Y_1)\right)\right]^2 + \left[\Phi^{-1}\left(C(F_2(Y_2)|F_1(Y_1))\right)\right]^2$$

where Φ denotes the standard normal c.d.f. and $C(F_2(Y_2)|F_1(Y_1))$ the conditional copula. Then, the random variable $S(Y_1, Y_2)$ should have a χ_2^2 distribution if C is the true copula, which can be tested by a Kolmogorov–Smirnov test between S as a function of pseudo observations and χ_2^2 for each pair of the standardized residuals for different currencies. The test statistics are given in Table A1 in Appendix C. Notice that $n = 206$ in our situation, thus the 5% critical value for the Kolmogorov–Smirnov test is equal to $1.358/\sqrt{206} = 0.095$, indicating none of the fitted copulas should be rejected as the true copulas.

Using the fitted parameters $\hat{\gamma}_0$ and $\hat{\gamma}_1$, we obtain the fitted tail order of maximal dependence $\hat{\kappa}_L$ as well as the fitted WLTMD $\hat{\chi}_L$. Then, we apply the *WNACut* algorithm for the first stage of the clustering procedure and Ward1 method for the second stage on the affinity matrix constructed by

the $(14 \times 14 = 196)$ $\hat{\chi}_L$s for $k = 2, \ldots, 13$, respectively. The total weights for *WNACut* is k, and the percentages of weights cut of the total weights $WCut(\mathcal{C}^*)/k$ against k are shown in Figure 9, which indicates that a relatively stable clustering result is obtained when the number of clusters $k \geq 4$.

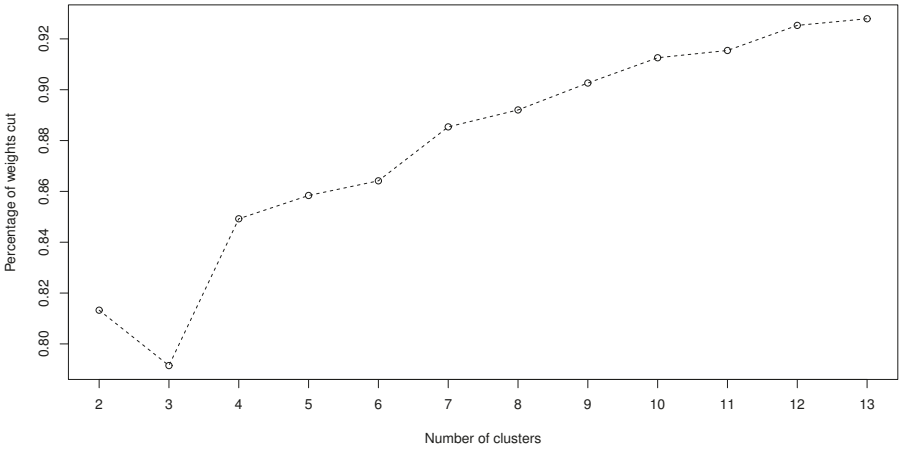


Figure 9. Percentage of weights cut against the number of clusters.

To get a closer look at the clustering result, first we consider the case $k = 3$ whose cluster dendrogram is given by Figure 10.

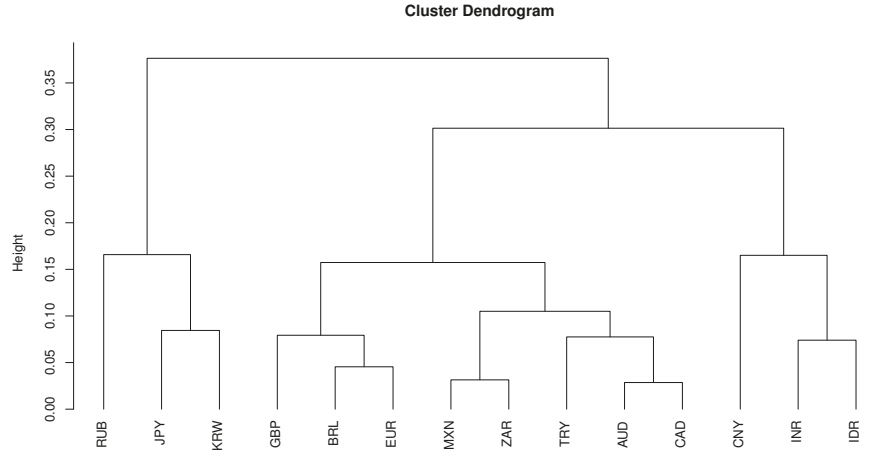


Figure 10. The dendrogram of the clustering procedure based on *WNACut* methods with affinity matrix as WLTMDs for $k = 3$.

The resulting clusters are listed in Table 6 which provides a regional segmentation of the world: Northeast Asia, the East/Southeast Asia and the rest of the world. Obviously, such a cluster result has the lowest percentage of weighted cuts and seems to perform well for currencies of countries in the Northeast Asia, East Asia and Southeast Asia. However, it fails to distinguish the currencies of the rest of the world. For the purpose of comparison, in Table 6, we also provide the clustering results given by the method of De Luca and Zuccolotto (2011) for $k = 3$.

Table 6. Members for each of the clusters when $k = 3$.

	Cluster 1		Cluster 2	Cluster 3
Our proposed procedure	Australia	Mexico	China	Russia
	Brazil	South Africa	India	Japan
	Canada	Turkey	Indonesia	South Korea
	UK	EU		
De Luca and Zuccolotto (2011)	Australia	Mexico	Brazil	Indonesia
	Canada	South Africa	Japan	Turkey
	China	Russia		
	India	South Korea		
	UK	EU		

Next, we compare the clustering result for $k = 4$ to $k = 3$. When $k = 4$, the clustering dendrogram is given by Figure 11.

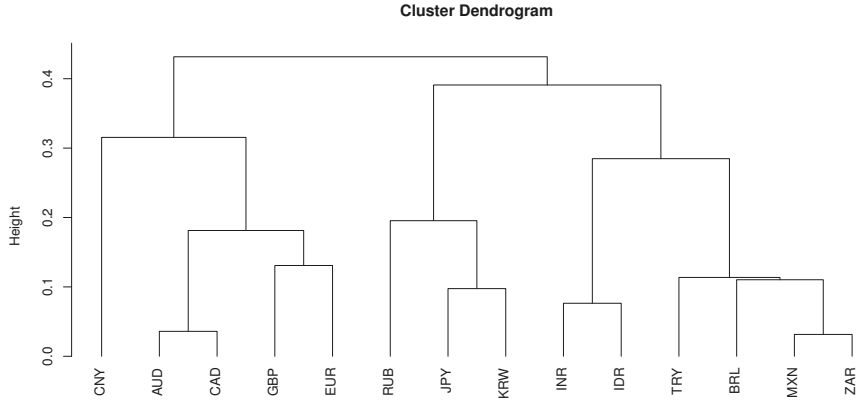


Figure 11. The dendrogram of the clustering procedure based on *WNACut* methods with affinity matrix as WLTMDs for $k = 4$.

The resulting clusters are listed in Table 7 which preserves the regional characteristics for currencies of countries in East/Northeast Asia while the partition of the rest of the world in Table A2 is provided by the IMF. For the purpose of comparison, in Table 7, we also provide a clustering result given by the method of De Luca and Zuccolotto (2011) for $k = 4$. Here are some remarks on the clustering outcomes obtained by our clustering procedure.

- All economies in the first group have strong economic connections with the US (under certain type of free trade agreements). Besides, the nominal per capita GDPs of these economies are all above $\$30,000$ (see Table A2 in Appendix D). Thus, it is reasonable to include currencies of these economies in one group.
- The second group of economies have neither strong economic connection to the US nor high nominal per capita GDPs (less than $\$10,000$). Moreover, all of them are identified as emerging economies by IMF (see Table A2 in Appendix D).
- China is the only member of the third group. Although there is no free trade agreements between US and China, it is well-acknowledged that China has a strong economic connection with the US. In addition, China has a very high nominal GDP (the third highest among all 20 economies) but very low nominal GDP per capita (less than $\$8000$). As a result, it is reasonable to not include China in any other group.
- It is reasonable to include the rest three economies in one group from the geographical perspective.

Table 7. Members for each of the clusters when $k = 4$.

	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**
Our proposed procedure	Australia Canada UK EU	Brazil India Indonesia Mexico South Africa Turkey	China	Russia Japan South Korea
De Luca and Zuccolotto (2011)	Australia Canada Brazil EU Russia South Korea South Africa Turkey	China India Indonesia Mexico	UK	Japan

In conclusion, the clustering result for $k = 4$ seems to more reasonable compared to that of $k = 3$. Besides, the clustering result for $k = 4$ could not be obtained through further partition based on the clustering result for $k = 3$ because it requires a combination of {Brazil, Mexico, South Africa, Turkey} and {India, Indonesia} as a new cluster. Therefore, the clustering result for $k = 3$ seems not to be sufficiently stable from our perspective.

5.3. An Example of Portfolio Management with the Clustering

As stressed in Section 1, one important application of time series clustering is risk management. Since the WLTMD represents the extreme co-movement downwards, the resulting clusters obtained through our proposed procedure represent groups of assets whose returns move in the same direction when both returns are extremely low. Hence, we should avoid including two assets from the same cluster in our portfolio. For instance, if we would like to construct a portfolio with four of the 14 aforementioned currencies, we may consider the following steps:

1. Perform our proposed clustering procedure for $k = 4$. The clustering result is given in Table 7.
2. From the clustering result in Table 7, one has $(4 \times 6 \times 1 \times 3 =)$ 72 choices if he/she tries to avoid selecting two assets from the same cluster in his/her portfolio. All 72 resulting portfolios are listed in Table A3 in Appendix E.
3. Construct portfolio using Markowitz's procedure (minimum variance criteria), namely, computing the sample covariance matrix $\mathbf{S}$ of the log-returns for any combination of currencies in Table A3 and solving
$$\min_{\boldsymbol{w}} \boldsymbol{w}'\mathbf{S}\boldsymbol{w} \tag{5}$$
subject to $w_1 + w_2 + w_3 + w_4 = 1$ and $w_i \geq 0$ for $i = 1, 2, 3, 4$, where $\boldsymbol{w} = \begin{bmatrix} w_1 & w_2 & w_3 & w_4 \end{bmatrix}'$. The resulting weights corresponding to the 72 choices of currencies are also given in Table A3 in Appendix E.

The paths of the returns for all 72 combinations are shown in Figure 12. Notice that there are $(\binom{14}{4} =)$1001 different combinations of four currencies without considering the clustering result (the resulting weights as well as the mean and accumulative returns corresponding to these 1001 portfolios are listed in Table S1 in the Supplementary Materials). We also construct portfolios using Equation (5) for all of the 1001 combinations of currencies and plot the corresponding paths of returns in Figure 12. In contrast, in Figure 12, we also provide the paths of aggregated returns of all portfolios constructed by the method of De Luca and Zuccolotto (2011).

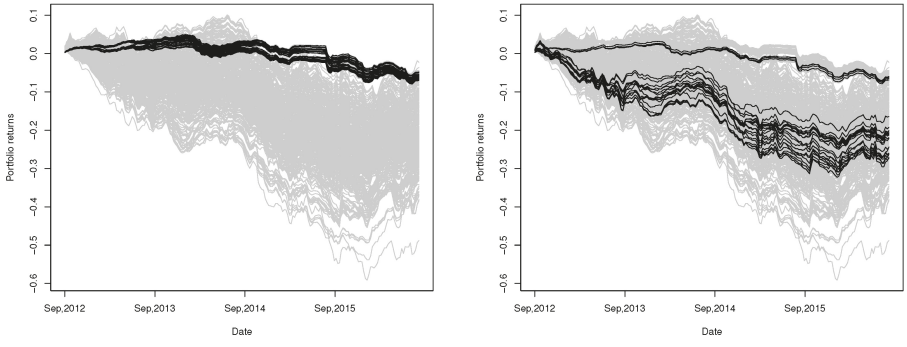


Figure 12. Returns (*black lines*) of the selected minimum variance portfolios based on our clustering procedure (left, 72 portfolios in total) and the procedure proposed by De Luca and Zuccolotto (2011) (right, 32 portfolios in total). Each portfolio consists of four currencies, selected by choosing one currency from each of the four resulting clusters, compared to returns (*gray lines*) of minimum variance portfolios constructed by all combinations (1001 in total) of four currencies out of total 14 currencies.

In Figure 12, the portfolios constructed through our proposed procedure are shown to have uniformly exceptional performance among all possible ways to construct portfolios with four currencies for a long period (September 2012–August 2016). Compared with the portfolios constructed using the method of De Luca and Zuccolotto (2011), all of our portfolios outperform most of their portfolios. Besides, the paths of return provided by our portfolios do not vary significantly across the 72 combinations of currencies and seem to have better performance in the long run (the average and accumulative returns of these 72 portfolios on 17 August 2016 (the last date of the observation period) are also provided in Table A3 in Appendix E, from which we can see that the accumulative returns are from -7% to -4%, corresponding to the range shown in the left of Figure 12).

We also show the path of risk adjusted returns provided by our portfolios in Figure 13. The results also show that our portfolios have outstanding performance most of the time.

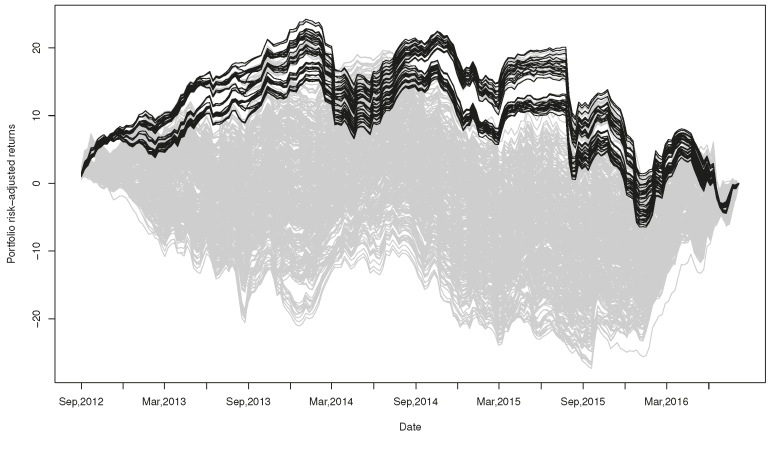


Figure 13. Risk-adjusted returns (*black lines*) of the 72 minimum variance portfolios selected based on our proposed clustering procedure, compared to risk-adjusted returns (*gray lines*) of minimum variance portfolios constructed by all combinations (1001 in total) of four currencies out of total 14 currencies.

For a retrospective study, the performance of the proposed portfolios as well as all possible portfolios from 17 August 2016 to 10 September 2018 are plotted in Figure 14. Although these portfolios fail to be the best choices again, the resulting returns are still shown to have very strong invulnerability against fluctuations and risks.

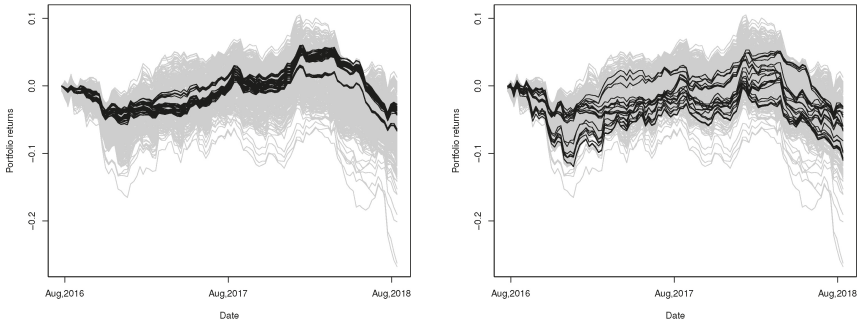


Figure 14. Returns (*black lines*) of the selected minimum variance portfolios based on our clustering procedure (left, 72 portfolios in total) and the procedure proposed by De Luca and Zuccolotto (2011) (right, 32 portfolios in total). Each portfolio consists of four currencies, selected by choosing one currency from each of the four clusters, compared to returns (*gray lines*) of minimum variance portfolios constructed by all combinations (1001 in total) of four currencies out of total 14 currencies.

In contrast, in Figure 14, we also provide the paths of aggregate returns of all portfolios constructed by the method of De Luca and Zuccolotto (2011). Obviously, the performances are quite inconsistent and the returns seem to vary when volatility increases.

For the retrospective study, we also show the path of risk adjusted returns provided by our portfolios in Figure 15. The performance of our portfolios seems to be very good during early 2018 but relatively poor during late 2016 to 2017.

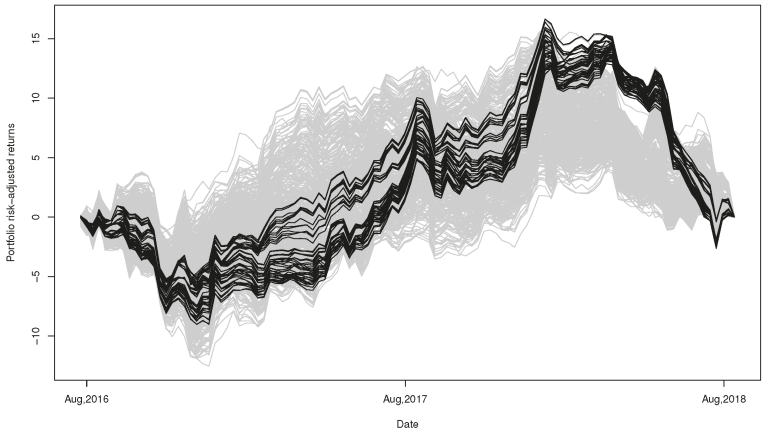


Figure 15. Risk-adjusted returns (*black lines*) of the 72 minimum variance portfolios selected based on our proposed clustering procedure, compared to returns (*gray lines*) of minimum variance portfolios constructed by all combinations (1001 in total) of four currencies out of total 14 currencies.

6. Conclusions

In this paper, we have proposed a clustering procedure based on a new affinity measure indicating the extreme co-movements for financial time series. Unlike the common distance-based affinity matrix, our proposed affinity matrix is not necessarily symmetric and hence cannot be used as the input in hierarchical clustering algorithms directly. As a result, clustering procedures based on weighted cuts are employed and examined, and the *WNACut* method is finally selected as a recommended clustering procedure, based on the performances of compared procedures applicable. The resulting clusters seem to be reasonable when applied to the real exchange rate time series, and the portfolios constructed based on the resulting clusters are shown to outperform those from other clustering procedures, particularly in the long run.

Future research should focus on seeking consistent nonparametric estimators for the coefficients of maximal tail dependence such as WLTMD, since parametric copulas having explicit forms of these coefficients are not always available for the standard residuals extracted from the data. As pointed out in (Furman et al. 2015, sct. 7), the achievement in the area of M-estimators may be a possible way to obtain such estimators as well as the relevant statistical inference. Moreover, the idea of the OLS estimator (see Gabaix and Ibragimov 2011, for instance) may also be a potential way to address this challenging problem. Furthermore, as indexes of maximal tail dependence such as WLTMD are measures of extreme co-movements, they might not reflect the potential causality between quantities and hence could not be directly applied to some important issues in the financial markets such as contagion effects. Nevertheless, our proposed clustering procedure provides an insight to analyze financial time series using asymmetric information matrices such as matrices of contagion measures.

Supplementary Materials: The following are available at http://www.mdpi.com/2227-9091/6/4/115/s1, Table S1: 1001 portfolios constructed for all combinations of 4 currencies out of 14 currencies. The portfolio weights obtained by Markowitz minimal variance criteria is given in the bracket. Also, mean returns and accumulative returns are listed.

Author Contributions: Conceptualization, X.L.; Methodology, J.W.; Formal Analysis, C.Y.; Data Curation, W.J.

Funding: This research was funded by National Natural Science Foundation of China (71603190) and Ministry of Education in China (10YJC790280).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Proposition 2.1

Proof. Since there exists a function $\varphi^*(u) \in \mathcal{A}$ such that

$$\Pi^*(u) = C(\varphi^*(u), u^2/\varphi^*(u)),$$

combined with the upper Fréchet–Hoeffding bound of copulas we must have for any $x \in (0, 2)$,

$$C(u^x, u^{2-x}) \le \Pi^*(u) \le \min\{\varphi^*(u), u^2/\varphi^*(u)\} \tag{A1}$$

for all $u \in (0, 1]$ because $u^x, u^{2-x} \in \mathcal{A}$. From Equation (A1), it is easy to obtain

$$\Pi^*(u) \le u. \tag{A2}$$

Given the Lipschitz condition of copula, we have

$$\left|C(u^x, u^{2-x}) - C(1, u^2)\right| \le |u^x - 1| + u^2 \left|u^{-x} - 1\right| \to 0$$

as $x \downarrow 0$ by noticing $C(1, u^2) = u^2$, then

$$\lim_{x \downarrow 0} C(u^x, u^{2-x}) = u^2 \tag{A3}$$

for all $u \in (0,1]$. Combining Equations (A1) and (A3) yields

$$\Pi^*(u) \geq u^2 \tag{A4}$$

for all $u \in (0,1]$, and since $\Pi^*(u) = \ell^*(u)u^{\kappa_L^*}$, it further follows by Equations (A2) and (A4) that

$$1 - \frac{\log \ell^*(u)}{\log(u)} \leq \kappa_L^* \leq 2 - \frac{\log \ell^*(u)}{\log(u)} \tag{A5}$$

for all $u \in (0,1]$. Provided the Karamata's Representation Theorem, the slowly varying function $\ell^*(u)$ has the following representation:

$$\ell^*(u) = c^*(u) \exp\left\{ \int_1^{1/u} \frac{\epsilon^*(t)}{t} \, \mathrm{d}t \right\}, \tag{A6}$$

where measurable functions $c^*(u)$ and $\epsilon^*(t)$ satisfy

$$\lim_{u \downarrow 0} c^*(u) = c \in (0, \infty)$$

and

$$\lim_{t \to \infty} \epsilon^*(t) = 0.$$

Hence, by Equation (A6), we have for all $u \in (0,1]$

$$\frac{\log \ell^*(u)}{\log(u)} = \frac{1}{\log(u)} \left(\log c^*(u) + \int_1^{1/u} \frac{\epsilon^*(t)}{t} \, \mathrm{d}t \right). \tag{A7}$$

Notice that, in Equation (A7), it is obvious that

$$\lim_{u \downarrow 0} \frac{\log c^*(u)}{\log(u)} = 0, \tag{A8}$$

the major discussion should focus on the limit of the rest part of Equation (A7) as $u \downarrow 0$. Since $\epsilon^*(t) \to 0$ as $t \to \infty$, by defining $T_\epsilon := \sup\{t \in (1,\infty) : |\epsilon^*(t)| \geq \epsilon\}$ for arbitrary $\epsilon > 0$ we know that $T_\epsilon < \infty$. Hence, for small enough $u \in (0,1]$, we could rewrite

$$\frac{1}{\log(u)} \int_1^{1/u} \frac{\epsilon^*(t)}{t} \, \mathrm{d}t = \frac{1}{\log(u)} \left(\int_1^{T_\epsilon} \frac{\epsilon^*(t)}{t} \, \mathrm{d}t + \int_{T_\epsilon}^{1/u} \frac{\epsilon^*(t)}{t} \, \mathrm{d}t \right). \tag{A9}$$

Then, for Equation (A9), it is obvious, that for $u < U_\epsilon := \exp\left\{ -\frac{1}{\epsilon} \int_1^{T_\epsilon} \frac{\epsilon^*(t)}{t} \, \mathrm{d}t \right\}$, we have

$$\left| \frac{1}{\log(u)} \int_1^{T_\epsilon} \frac{\epsilon^*(t)}{t} \, \mathrm{d}t \right| < \epsilon,$$

while for $u < T_\epsilon^{-0.5}$,

$$\left| \frac{1}{\log(u)} \int_{T_\epsilon}^{1/u} \frac{\epsilon^*(t)}{t} \, \mathrm{d}t \right| < \left| \frac{1}{\log(u)} \right| \left| -\log(u) - \log T_\epsilon \right| \epsilon = \left| 1 + \frac{\log T_\epsilon}{\log(u)} \right| \epsilon \leq \epsilon.$$

Therefore, for all $u < \min\{T_\epsilon^{-0.5}, U_\epsilon\}$ we have

$$\left| \frac{1}{\log(u)} \int_1^{1/u} \frac{\epsilon^*(t)}{t} \, \mathrm{d}t \right| < 2\epsilon,$$

which is equivalent to

$$\lim_{u \downarrow 0} \frac{1}{\log(u)} \int_1^{1/u} \frac{\epsilon^*(t)}{t} \, \mathrm{d}t = 0.$$

Finally, combined with (A8) we deduce

$$\lim_{u \downarrow 0} \frac{\log \ell^*(u)}{\log(u)} = 0.$$

Therefore, we may conclude that $\kappa_L^* \in [1, 2]$ by letting $u \downarrow 0$ for both inequalities of Equation (A9). □

Appendix B. Details of the Clustering Algorithm in Simulation

Algorithm A1 Best *WCut* (Meilă and Pentney 2007, Algorithm 4.1).

Require: Affinity matrix, $\mathbf{\Delta}$; diagonal matrix of volume weights, $\mathbf{V}$; diagonal matrix of row weights, $\mathbf{R}$; number of clusters, k;

Ensure: The clustering, $\mathcal{C}^*$;

1: Update $\mathbf{\Delta}$ through the linear transform $\mathbf{R\Delta}$;
2: Define $D_i = \sum_{j=1}^{d} \delta_{ij}$ and $\mathbf{D} = \mathrm{diag}\{D_i\}_{i=1,\ldots,n}$;
3: Define $H(\mathbf{B}) = \frac{1}{2}(\mathbf{B} + \mathbf{B}^\intercal)$, where $\mathbf{B} = \mathbf{V}^{-1/2}(\mathbf{D} - \mathbf{\Delta})\mathbf{V}^{-1/2}$;
4: Compute $\mathbf{Y}$ the $n \times k$ matrix with orthonormal columns containing the eigenvectors of $H(\mathbf{B})$ corresponding to the k smallest eigenvalues;
5: Cluster the rows of $\mathbf{X} = \mathbf{V}^{-1/2}\mathbf{Y}$ as points in $\mathbb{R}^k$.

Appendix C. Goodness-of-Fit Test

Table A1. Estimated Kolmogorov statistics for the goodness-of-fit test.

	AUD	**BRL**	**CAD**	**CNY**	**GBP**	**INR**	**IDR**	**JPY**	**KRW**	**MXN**	**RUB**	**ZAR**	**TRY**	**EUR**
AUD	—	0.0634	0.0778	0.0636	0.0733	0.0614	0.0460	0.0747	0.0685	0.0922	0.0563	0.0815	0.0644	0.0487
BRL	0.0794	—	0.0685	0.0539	0.0425	0.0493	0.0488	0.0501	0.0527	0.0668	0.0766	0.0665	0.0841	0.0454
CAD	0.0899	0.0655	—	0.0502	0.0545	0.0512	0.0561	0.0673	0.0601	0.0568	0.0668	0.0563	0.0555	0.0629
CNY	0.0495	0.0445	0.0469	—	0.0794	0.0375	0.0633	0.0489	0.0666	0.0496	0.0494	0.0516	0.0516	0.0470
GBP	0.0603	0.0579	0.0517	0.0705	—	0.0481	0.0584	0.0580	0.0565	0.0493	0.0494	0.0486	0.0754	0.0414
INR	0.0633	0.0433	0.0470	0.0383	0.0484	—	0.0367	0.0476	0.0472	0.0290	0.0604	0.0462	0.0577	0.0637
IDR	0.0480	0.0531	0.0502	0.0612	0.0584	0.0371	—	0.0675	0.0686	0.0395	0.0577	0.0750	0.0665	0.0505
JPY	0.0687	0.0522	0.0677	0.0534	0.0611	0.0516	0.0705	—	0.0645	0.0637	0.0497	0.0513	0.0587	0.0542
KRW	0.0763	0.0619	0.0581	0.0797	0.0821	0.0569	0.0696	0.0643	—	0.0628	0.0578	0.0797	0.0896	0.0448
MXN	0.0861	0.0648	0.0748	0.0511	0.0524	0.0353	0.0353	0.0637	0.0543	—	0.0718	0.0585	0.0589	0.0446
RUB	0.0558	0.0748	0.0520	0.0550	0.0481	0.0493	0.0559	0.0485	0.0382	0.0512	—	0.0561	0.0565	0.0592
ZAR	0.0715	0.0620	0.0500	0.0566	0.0518	0.0504	0.0565	0.0510	0.0726	0.0786	0.0647	—	0.0687	0.0465
TRY	0.0675	0.0614	0.0542	0.0555	0.0890	0.0512	0.0599	0.0604	0.0719	0.0698	0.0576	0.0753	—	0.0483
EUR	0.0415	0.0574	0.0663	0.0483	0.0418	0.0607	0.0399	0.0548	0.0306	0.0339	0.0540	0.0419	0.0506	—

Appendix D. Economic Summary of G20 Nations

Table A2. Part of the economic summary of G20 nations 2015 by IMF (2014).

Member	Nom. GDP mil. USD	PPP GDP mil. USD	Nom. GDP per capitaUSD	PPP GDP per capltaUSD	HDI	Population	Area	Economic Classification (IMF)
Argentina	585,623	964,300	13,589	22,554	0.836	42,961,000	2,780,400	Emerging
Australia	1,223,887	1,489,000	50,962	47,389	0.935	23,599,000	7,692,024	Advanced
Brazil	1,772,589	3,166,000	8670	16,155	0.755	202,768,000	8,515,767	Emerging
Canada	1,552,386	1,632,000	43,332	44,967	0.913	35,467,000	9,984,670	Advanced
China	10,982,829	19,510,000	7990	14,107	0.727	1,367,520,000	9,572,900	Emerging
France	2,421,560	2,647,000	37,675	41,181	0.888	63,951,000	640,679	Advanced
Germany	3,357,614	3,842,000	40,997	46,893	0.916	80,940,000	357,114	Advanced
India	2,090,706	7,965,000	1617	6162	0.609	1,259,695,000	3,287,263	Emerging
Indonesia	858,953	2,839,000	3362	11,126	0.684	251,490,000	1,904,569	Emerging
Italy	1,815,757	2,174,000	29,867	35,708	0.873	60,665,551	301,336	Advanced
Japan	4,123,258	4,658,000	32,486	38,054	0.891	127,061,000	377,930	Advanced
South Korea	1,376,868	1,849,000	27,195	36,511	0.898	50,437,000	100,210	Advanced
Mexico	1,144,334	2,220,000	9009	17,534	0.756	119,581,789	1,964,375	Emerging
Russia	1,324,734	3,471,000	9055	25,411	0.798	146,300,000	17,098,242	Emerging
Saudi Arabia	653,219	1,683,000	20,813	53,624	0.837	30,624,000	2,149,690	Emerging
South Africa	312,957	723,518	5695	13,165	0.666	53,699,000	1,221,037	Emerging
Turkey	733,642	1,589,000	9437	20,438	0.761	77,324,000	783,562	Emerging
United Kingdom	2,849,345	2,660,000	43,771	41,159	0.907	64,511,000	242,495	Advanced
United States	17,947,000	17,947,000	55,805	55,805	0.915	318,523,000	9,526,468	Advanced
European Union	16,220,370	19,180,000	31,918	37,852	0.876	505,570,700	4,422,773	N/A

Appendix E. Seventy-Two Portfolios Constructed Based on Our Clustering Result in Table 7

Table A3. Sevent-two portfolios constructed based on our clustering result in Table 7. Notice that for each cluster only one currency is selected. The portfolio weights obtained by Markowitz minimal variance criteria is given in the bracket. In addition, mean returns and accumulative returns are listed.

Portfolio	Currency from Cluster 1	Currency from Cluster 2	Currency from Cluster 3	Currency from Cluster 4	Mean Return	Accumulative Return
1	AUD (0.0000)	BRL (0.0220)	CNY (0.9177)	JPY (0.0603)	−0.0321%	−6.6162%
2	CAD (0.0122)	BRL (0.0195)	CNY (0.9089)	JPY (0.0595)	−0.0329%	−6.7676%
3	GBP (0.0221)	BRL (0.0196)	CNY (0.8984)	JPY (0.0599)	−0.0333%	−6.8559%
4	EUR (0.0392)	BRL (0.0164)	CNY (0.8944)	JPY (0.0499)	−0.0313%	−6.4531%
5	AUD (0.0000)	INR (0.0539)	CNY (0.8841)	JPY (0.0620)	−0.0316%	−6.5051%
6	CAD (0.0188)	INR (0.0500)	CNY (0.8709)	JPY (0.0603)	−0.0332%	−6.8367%
7	GBP (0.0236)	INR (0.0508)	CNY (0.8643)	JPY (0.0613)	−0.0331%	−6.8180%
8	EUR (0.0419)	INR (0.0483)	CNY (0.8596)	JPY (0.0502)	−0.0315%	−6.4815%
9	AUD (0.0000)	IDR (0.0322)	CNY (0.9072)	JPY (0.0606)	−0.0320%	−6.5924%
10	CAD (0.0227)	IDR (0.0271)	CNY (0.8913)	JPY (0.0589)	−0.0336%	−6.9266%
11	GBP (0.0278)	IDR (0.0317)	CNY (0.8807)	JPY (0.0597)	−0.0340%	−6.9977%
12	EUR (0.0441)	IDR (0.0263)	CNY (0.8810)	JPY (0.0486)	−0.0315%	−6.4964%
13	AUD (0.0023)	MXN (0.0109)	CNY (0.9221)	JPY (0.0647)	−0.0299%	−6.1674%
14	CAD (0.0300)	MXN (0.0000)	CNY (0.9093)	JPY (0.0608)	−0.0310%	−6.3935%
15	GBP (0.0269)	MXN (0.0066)	CNY (0.9029)	JPY (0.0636)	−0.0310%	−6.3900%
16	EUR (0.0468)	MXN (0.0031)	CNY (0.8995)	JPY (0.0507)	−0.0288%	−5.9339%
17	AUD (0.0085)	ZAR (0.0000)	CNY (0.9290)	JPY (0.0626)	−0.0290%	−5.9725%
18	CAD (0.0300)	ZAR (0.0000)	CNY (0.9093)	JPY (0.0608)	−0.0310%	−6.3935%
19	GBP (0.0282)	ZAR (0.0000)	CNY (0.9091)	JPY (0.0627)	−0.0301%	−6.2074%
20	EUR (0.0476)	ZAR (0.0000)	CNY (0.9024)	JPY (0.0501)	−0.0284%	−5.8410%
21	AUD (0.0085)	TRY (0.0000)	CNY (0.9290)	JPY (0.0626)	−0.0290%	−5.9725%
22	CAD (0.0300)	TRY (0.0000)	CNY (0.9093)	JPY (0.0608)	−0.0310%	−6.3935%
23	GBP (0.0282)	TRY (0.0000)	CNY (0.9091)	JPY (0.0627)	−0.0301%	−6.2074%
24	EUR (0.0476)	TRY (0.0000)	CNY (0.9024)	JPY (0.0501)	−-0.0284%	−5.8410%
25	AUD (0.0000)	BRL (0.0264)	CNY (0.9736)	KRW (0.0000)	−0.0271%	−5.5828%
26	CAD (0.0217)	BRL (0.0218)	CNY (0.9566)	KRW (0.0000)	−0.0285%	−5.8756%
27	GBP (0.0234)	BRL (0.0238)	CNY (0.9528)	KRW (0.0000)	−0.0284%	−5.8427%
28	EUR (0.0606)	BRL (0.0166)	CNY (0.9228)	KRW (0.0000)	−0.0272%	−5.6042%
29	AUD (0.0045)	INR (0.0558)	CNY (0.9397)	KRW (0.0000)	−0.0262%	−5.3969%
30	CAD (0.0304)	INR (0.0511)	CNY (0.9185)	KRW (0.0000)	−0.0286%	−5.8954%
31	GBP (0.0260)	INR (0.0541)	CNY (0.9200)	KRW (0.0000)	−0.0275%	−5.6677%
32	EUR (0.0635)	INR (0.0480)	CNY (0.8885)	KRW (0.0000)	−0.0273%	−5.6167%
33	AUD (0.0058)	IDR (0.0391)	CNY (0.9516)	KRW (0.0036)	−0.0276%	−5.6778%
34	CAD (0.0321)	IDR (0.0346)	CNY (0.9332)	KRW (0.0000)	−0.0299%	−6.1540%
35	GBP (0.0303)	IDR (0.0416)	CNY (0.9282)	KRW (0.0000)	−0.0296%	−6.1004%
36	EUR (0.0643)	IDR (0.0307)	CNY (0.9050)	KRW (0.0000)	−0.0281%	−5.7784%
37	AUD (0.0143)	MXN (0.0000)	CNY (0.9762)	KRW (0.0094)	−0.0233%	−4.7940%
38	CAD (0.0418)	MXN (0.0000)	CNY (0.9581)	KRW (0.0001)	−0.0264%	−5.4331%
39	GBP (0.0289)	MXN (0.0000)	CNY (0.9621)	KRW (0.0089)	−0.0238%	−4.8943%
40	EUR (0.0690)	MXN (0.0000)	CNY (0.9310)	KRW (0.0000)	−0.0242%	−4.9823%
41	AUD (0.0143)	ZAR (0.0000)	CNY (0.9762)	KRW (0.0094)	−0.0233%	−4.7940%
42	CAD (0.0418)	ZAR (0.0000)	CNY (0.9581)	KRW (0.0001)	−0.0264%	−5.4331%
43	GBP (0.0289)	ZAR (0.0000)	CNY (0.9621)	KRW (0.0089)	−0.0238%	−4.8943%
44	EUR (0.0690)	ZAR (0.0000)	CNY (0.9310)	KRW (0.0000)	−0.0242%	−4.9823%
45	AUD (0.0143)	TRY (0.0000)	CNY (0.9762)	KRW (0.0094)	−0.0233%	−4.7940%
46	CAD (0.0418)	TRY (0.0000)	CNY (0.9581)	KRW (0.0001)	−0.0264%	−5.4331%
47	GBP (0.0289)	TRY (0.0000)	CNY (0.9621)	KRW (0.0089)	−0.0238%	−4.8943%
48	EUR (0.0690)	TRY (0.0000)	CNY (0.9310)	KRW (0.0000)	−0.0242%	−4.9823%
49	AUD (0.0000)	BRL (0.0264)	CNY (0.9736)	RUB (0.0000)	−0.0271%	−5.5828%
50	CAD (0.0217)	BRL (0.0218)	CNY (0.9566)	RUB (0.0000)	−0.0285%	−5.8756%
51	GBP (0.0234)	BRL (0.0238)	CNY (0.9528)	RUB (0.0000)	−0.0284%	−5.8427%
52	EUR (0.0606)	BRL (0.0166)	CNY (0.9228)	RUB (0.0000)	−0.0272%	−5.6042%
53	AUD (0.0045)	INR (0.0558)	CNY (0.9397)	RUB (0.0000)	−0.0262%	−5.3969%
54	CAD (0.0304)	INR (0.0511)	CNY (0.9185)	RUB (0.0000)	−0.0286%	−5.8954%
55	GBP (0.0260)	INR (0.0541)	CNY (0.9200)	RUB (0.0000)	−0.0275%	−5.6677%
56	EUR (0.0635)	INR (0.0480)	CNY (0.8885)	RUB (0.0000)	−0.0273%	−5.6167%
57	AUD (0.0071)	IDR (0.0394)	CNY (0.9534)	RUB (0.0000)	−0.0279%	−5.7458%
58	CAD (0.0321)	IDR (0.0346)	CNY (0.9332)	RUB (0.0000)	−0.0299%	−6.1540%
59	GBP (0.0303)	IDR (0.0416)	CNY (0.9282)	RUB (0.0000)	−0.0296%	−6.1004%
60	EUR (0.0643)	IDR (0.0307)	CNY (0.9050)	RUB (0.0000)	−0.0281%	−5.7784%

Table A3. *Cont.*

Portfolio	Currency from Cluster 1	Currency from Cluster 2	Currency from Cluster 3	Currency from Cluster 4	Mean Return	Accumulative Return
61	AUD (0.0182)	MXN (0.0000)	CNY (0.9818)	RUB (0.0000)	−0.0240%	−4.9542%
62	CAD (0.0418)	MXN (0.0000)	CNY (0.9582)	RUB (0.0000)	−0.0264%	−5.4349%
63	GBP (0.0310)	MXN (0.0000)	CNY (0.9690)	RUB (0.0000)	−0.0242%	−4.9888%
64	EUR (0.0690)	MXN (0.0000)	CNY (0.9310)	RUB (0.0000)	−0.0242%	−4.9823%
65	AUD (0.0182)	ZAR (0.0000)	CNY (0.9818)	RUB (0.0000)	−0.0240%	−4.9542%
66	CAD (0.0418)	ZAR (0.0000)	CNY (0.9582)	RUB (0.0000)	−0.0264%	−5.4349%
67	GBP (0.0310)	ZAR (0.0000)	CNY (0.9690)	RUB (0.0000)	−0.0242%	−4.9888%
68	EUR (0.0690)	ZAR (0.0000)	CNY (0.9310)	RUB (0.0000)	−0.0242%	−4.9823%
69	AUD (0.0182)	TRY (0.0000)	CNY (0.9818)	RUB (0.0000)	−0.0240%	−4.9542%
70	CAD (0.0418)	TRY (0.0000)	CNY (0.9582)	RUB (0.0000)	−0.0264%	−5.4349%
71	GBP (0.0310)	TRY (0.0000)	CNY (0.9690)	RUB (0.0000)	−0.0242%	−4.9888%
72	EUR (0.0690)	TRY (0.0000)	CNY (0.9310)	RUB (0.0000)	−0.0242%	−4.9823%

References

Bae, Kee-Hong, G. Andrew Karolyi, and Reneé M. Stulz. 2003. A New Approach to Measuring Financial Contagion. *Review of Financial Studies* 16: 717–63. [CrossRef]

Baragona, Roberto. 2001. A simulation study on clustering time series with metaheuristic methods. *Quaderni di Statistica* 3: 1–26.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler. 2005. Comovement. *Journal of Financial Economics* 75: 283–317 [CrossRef]

Billio, Monica, and Massimiliano Caporin. 2009. A generalized Dynamic Conditional Correlation model for portfolio risk evaluation. *Mathematics and Computers in Simulation* 79: 2566–78. [CrossRef]

Billio, Monica, Massimiliano Caporin, and Michele Gobbo. 2006. Flexible Dynamic Conditional Correlation multivariate GARCH models for asset allocation. *Applied Financial Economics Letters* 2: 123–30. [CrossRef]

Bonanno, Giovanni, Guido Caldarelli, Fabrizio Lillo, Salvatore Miccichè, Nicolas Vandewalle, and Rosario N. Mantegna. 2004. Networks of equities in financial markets. *The European Physical Journal B - Condensed Matter* 38: 363–71. [CrossRef]

Breymann, Wolfgang, Alexandra Dias, and Paul Embrechts. 2003. Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance* 3: 1–14. [CrossRef]

Brockwell, Peter J., and Richard A. Davis. 2002. *Introduction to Time Series and Forecasting*. Springer Texts in Statistics. New York: Springer. [CrossRef]

Coles, Stuart, Janet Heffernan, and Jonathan Tawn. 1999. Dependence measures for extreme value analyses. *Extremes* 2: 339–65. [CrossRef]

De Luca, Giovanni, and Paola Zuccolotto. 2011. A tail dependence-based dissimilarity measure for financial time series clustering. *Advances in Data Analysis and Classification* 5: 323–40. [CrossRef]

De Luca, Giovanni, and Paola Zuccolotto. 2015. Dynamic tail dependence clustering of financial time series. *Statistical Papers* 58: 641–57. [CrossRef]

Durante, Fabrizio, Enrico Foscolo, Piotr Jaworski, and Hao Wang. 2014. A spatial contagion measure for financial time series. *Expert Systems with Applications* 41: 4023–34. [CrossRef]

Durante, Fabrizio, Roberta Pappadà, and Nicola Torelli. 2014. Clustering of financial time series in risky scenarios. *Advances in Data Analysis and Classification* 8: 359–76. [CrossRef]

Durante, Fabrizio, Roberta Pappadà, and Nicola Torelli. 2015. Clustering of time series via non-parametric tail dependence estimation. *Statistical Papers* 56: 701–21. [CrossRef]

Engle, Robert. 2002. Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics* 20: 339–50.

Engle, Robert, and Kevin Sheppard. 2001. *Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH*. Cambridge: National Bureau of Economic Research. [CrossRef]

International Monetary Fund. 2014. *World Economic Outlook Database*. Washington: International Monetary Fund.

Furman, Edward, Alexey Kuznetsov, Jianxi Su, and Ričardas Zitikis. 2016. Tail dependence of the Gaussian copula revisited. *Insurance: Mathematics and Economics* 69: 97–103.

Furman, Edward, Jianxi Su, and Ričardas Zitikis. 2015. Paths and indices of maximal tail dependence. *ASTIN Bulletin* 45: 661–78. [CrossRef]

Gabaix, Xavier, and Rustam Ibragimov. 2011. Rank −1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents. *Journal of Business & Economic Statistics* 29: 24–39.

Hua, Lei, and Harry Joe. 2011. Tail order and intermediate tail dependence of multivariate copulas. *Journal of Multivariate Analysis* 102: 1454–71. [CrossRef]

Kaufman, Leonard, and Peter J. Rousseeuw. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley Series in Probability and Statistics. Hoboken: John Wiley & Sons, Inc. doi:10.1002/9780470316801. [CrossRef]

Liebscher, Eckhard. 2008. Construction of asymmetric multivariate copulas. *Journal of Multivariate Analysis* 99: 2234–50. [CrossRef]

Mantegna, Rosario N. 1999. Hierarchical structure in financial markets. *The European Physical Journal B* 11: 193–97. [CrossRef]

Meilă, Marina. 2003. Comparing Clusterings by the Variation of Information. In *Learning Theory and Kernel Machines*. Berlin and Heidelberg: Springer Nature Switzerland AG., pp. 173–87. [CrossRef]

Meilă, Marina, and William Pentney. 2007. Clustering by weighted cuts in directed graphs. Paper presented at 2007 SIAM International Conference on Data Mining, Philadelphia, PA, USA, 26–28 April.

Nelsen, Roger B. 2006. *An Introduction to Copulas*. Springer Series in Statistics. New York: Springer. [CrossRef]

Pericoli, Marcello, and Massimo Sbracia. 2003. A Primer on Financial Contagion. *Journal of Economic Surveys* 17: 571–608. [CrossRef]

Veldkamp, Laura L. 2006. Information markets and the comovement of asset prices. *The Review of Economic Studies* 73: 823–45. [CrossRef]

Verma, Deepak, and Marina Meilă. 2003. *A Comparison of Spectral Clustering Algorithms*. Technical Report UW-CSE-03-05-01. Seattle: University of Washington.

risks

MDPI

Article

A VaR-Type Risk Measure Derived from Cumulative Parisian Ruin for the Classical Risk Model

Mohamed Amine Lkabous *,† and Jean-François Renaud †

Département de mathématiques, Université du Québec à Montréal (UQAM), Montréal, QC H2X 3Y7, Canada; renaud.jf@uqam.ca

* Correspondence: lkabous.mohamed_amine@courrier.uqam.ca; Tel.: +1-514-987-4186

† These authors contributed equally to this work.

Received: 30 July 2018; Accepted: 23 August 2018; Published: 24 August 2018

Abstract: In this short paper, we study a VaR-type risk measure introduced by Guérin and Renaud and which is based on cumulative Parisian ruin. We derive some properties of this risk measure and we compare it to the risk measures of Trufin et al. and Loisel and Trufin.

Keywords: risk measure; cumulative Parisian ruin; stochastic orders; surplus process

1. Introduction

Over the last few years, several *dynamic* risk measures, i.e., risk measures based on ruin-theoretic quantities, have been studied. For example, in the classical compound Poisson risk model, Trufin et al. (2011) considered a VaR-type risk measure defined as the smallest initial capital needed to ensure a certain probability of solvency throughout the lifetime of the surplus process. This risk measure has been extended by Mitric and Trufin (2016) who defined a risk measure taking into account both the probability of ruin and the expected deficit at ruin. In addition, Loisel and Trufin (2014) used the expected area below the solvency threshold as a risk indicator to introduce a new risk measure with some interesting properties.

Very recently, implementation delays in the recognition of ruin and occupation times of the surplus process have been used as alternative risk management tools to assess the quality of an insurance portfolio. In this direction, Guérin and Renaud (2017) introduced the concept of cumulative Parisian ruin, which is based on the *time spent in the red* by the underlying surplus process. The time of cumulative Parisian ruin is the first time the surplus process stays cumulatively below a critical level longer than a pre-determined grace period. Inspired by the risk measure of Trufin et al. (2011), they defined a VaR-type risk measure based on cumulative Parisian ruin. It is also defined as the smallest amount of capital for which the associated cumulative Parisian ruin probability is less than or equal to a tolerable level.

In this paper, we study this VaR-type risk measure based on cumulative Parisian ruin. In Guérin and Renaud (2017), this risk measure is proposed as a motivational reason to study the concept of cumulative Parisian ruin; the risk measure itself is neither analyzed nor used for any particular application. We derive some of its properties and compare it to the risk measures of Trufin et al. (2011) and Loisel and Trufin (2014).

The rest of the paper is organized as follows. In Section 2, we recall some background on the Cramér–Lundberg model, also known as the classical risk model, and define the concept of cumulative Parisian ruin. In Section 3, we introduce our risk measure and we give some of its properties. Finally, in Section 4, we study our risk measure in the special case of a Cramér–Lundberg process with exponential claims.

2. Insurance Risk Model

The Cramér–Lundberg model was proposed by Lundberg (1903) and further developed by Cramér (1930). In this model, the surplus process of an insurance company is modelled by

$$X_t = x + ct - S_t, \tag{1}$$

where $x \geq 0$ and $c > 0$, and where $S_t = \sum_{i=1}^{N_t} C_i$ is a compound Poisson process with $N = \{N_t, t \geq 0\}$ a Poisson process of intensity $\lambda > 0$ and with $\{C_1, C_2, \dots\}$ positive random variables following a common cumulative distribution function F_C. Recall that in this setup the claim sizes $\{C_1, C_2, \dots\}$ are mutually independent and are also independent of the number-of-claim process N. The process $S = \{S_t, t \geq 0\}$ is known as the aggregate claim amount process. We call x the initial capital and c the premium rate.

We use the following equivalent notations $\mathbb{P}_x(\cdot) \equiv \mathbb{P}(\cdot | X_0 = x)$ to emphasize that the process X starts at level x. The notation $\mathbb{E}_x$ corresponds to $\mathbb{P}_x$. When $X_0 = 0$, we drop the index. In this model, the premium rate c is chosen usually to satisfy the net profit condition $\mathbb{E}[X_1] = c - \lambda \mathbb{E}[C_1] > 0$, which means that we can define the safety loading factor $\eta > 0$ by $\eta := (c - \lambda \mathbb{E}[C_1]) / \lambda \mathbb{E}[C_1]$.

The time of classical ruin associated to X is defined as

$$\tau_0^- = \inf\{t > 0 \colon X_t < 0\}.$$

We denote the corresponding finite-time probability of ruin, for $x \geq 0$ and $t > 0$, by

$$\psi(t, x) = \mathbb{P}_x(\tau_0^- \leq t), \tag{2}$$

and the infinite-time probability of ruin by

$$\psi(x) = \mathbb{P}_x(\tau_0^- < \infty). \tag{3}$$

Of course, we have $\psi(x) = \lim_{t \to \infty} \psi(t, x)$.

In Trufin et al. (2011), assuming that the safety loading η is fixed, the following ruin-consistent VaR-type risk measure is defined and analyzed: for a tolerance level $\epsilon > 0$, set

$$\zeta_\epsilon[C] = \inf\{x \geq 0 \colon \psi(x) \leq \epsilon\}.$$

It is well known that we can compute $\psi(x)$ using the Pollaczeck–Khinchine formula (also known in the actuarial literature as the Beekman's convolution formula, see Beekman (1985)) which states that the probability of classical ruin is equal to the tail distribution function of a compound geometric random variable. First, let us define the aggregate loss at time t by $L_t = S_t - ct$ and the maximal aggregate loss of the process by $L = \max_{t \geq 0} L_t$. The random variable L can be expressed as

$$L = \sum_{i=1}^{M} D_i, \tag{4}$$

where M is the number of record highs, which has a geometric distribution with success probability $\eta/(\eta + 1)$, and where $\{D_1, D_2, \dots\}$ are the ladder heights with common distribution $F_D(u) = \int_0^u (1 - F_C(y)) \mathrm{d}y / \mathbb{E}[C_1]$. The Pollaczeck–Khinchine formula for the probability of ruin is then given by

$$\psi(x) = \mathbb{P}(L > x) = 1 - \frac{\eta}{\eta + 1} \sum_{k=1}^{\infty} \left(\frac{1}{\eta + 1}\right)^k F_D^{*(k)}(x), \tag{5}$$

where $F_D^{*(k)}$ denotes the k-th convolution of the distribution F_D. Therefore, this risk measure can also be written as follows:

$$\zeta_\epsilon[C] = \inf\{x \geq 0 \colon \mathbb{P}(L > x) \leq \epsilon\} = F_L^{-1}(1-\epsilon). \tag{6}$$

In some sense, the focus of this risk measure is shifted from the surplus process X to the distribution of the maximal aggregate loss L. This important relationship is at the core of the analysis done in Trufin et al. (2011). However, this relationship with the maximal aggregate loss L does not exist for the finite-time ruin probability. Note that this is also the case for the VaR-type risk measure defined and analyzed in Mitric and Trufin (2016).

Cumulative Parisian Ruin

Recently, Guérin and Renaud (2017) introduced a new definition of actuarial ruin based on the occupation-time process (below 0) associated with the surplus process X. The occupation-time process $\mathcal{O}^L = \{\mathcal{O}_t^L, t \geq 0\}$ is defined as

$$\mathcal{O}_t^L = \int_0^t \mathbf{1}_{\{X_u<0\}} \mathrm{d}u = \int_0^t \mathbf{1}_{\{L_u>X_0\}} \mathrm{d}u.$$

Then, the time of cumulative Parisian ruin, with delay $r > 0$, is given by

$$\sigma_r = \inf\left\{t > 0 \colon \mathcal{O}_t^L > r\right\}.$$

In the definition of cumulative Parisian ruin, we aggregate the duration of all periods of financial distress until we accumulate r units of time spent in that red zone. Consequently, ruin is not declared as soon as X goes below zero: for $x \geq 0$, $t > 0$ and $r > 0$, we have

$$\mathbb{P}_x(\sigma_r \leq t) \leq \mathbb{P}_x\left(\tau_0^- \leq t\right). \tag{7}$$

Cumulative Parisian ruin is somehow a generalization of classical ruin and, when $r \to 0$, we recover the classical definition (see Guérin and Renaud (2017) for the details and see Figure 1 for a graphical comparison).

We denote the finite-time probability of cumulative Parisian ruin by

$$\psi_r(t,x) = \mathbb{P}_x(\sigma_r \leq t) = \mathbb{P}_x\left(\mathcal{O}_t^L > r\right) \tag{8}$$

and the infinite-time version by

$$\psi_r(x) = \mathbb{P}_x(\sigma_r < \infty).$$

Of course, we have $\psi_r(x) = \lim_{t\to\infty} \psi_r(t,x)$. With this new notation in hand, we can re-write the inequality in Equation (7) as follows: for $x \geq 0$, $t > 0$ and $r > 0$, we have

$$\psi_r(t,x) \leq \psi(t,x). \tag{9}$$

We also have

$$\psi(t,x) = \lim_{r\to 0} \psi_r(t,x) \quad \text{and} \quad \psi(x) = \lim_{r\to 0} \psi_r(x).$$

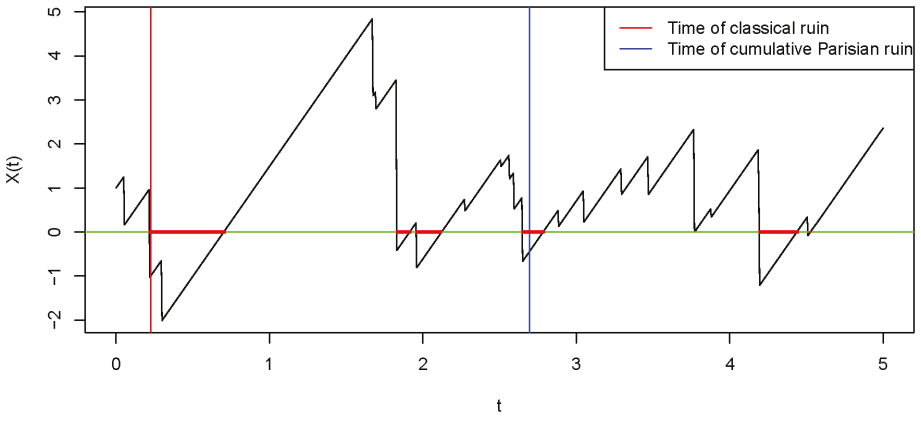


Figure 1. A sample path of a Cramér–Lundberg process X_t. The time of ruin τ_0^- is in red and the time of cumulative Parisian ruin κ_r is in blue.

3. A VaR-type Risk Measure Derived from Cumulative Parisian Ruin

Using the definition of cumulative Parisian ruin, Guérin and Renaud (2017) defined the following VaR-type risk measure: for a time horizon of length t and delay r, and for a given tolerance level $\epsilon > 0$, set

$$\rho_\epsilon^{(r,t)}[X] = \inf\{x \geq 0 \colon \psi_r(t,x) \leq \epsilon\}.$$

It gives the amount of initial capital needed in order to bound the finite-time probability of cumulative Parisian ruin with delay r by ϵ. Since $\psi_r(t,x) = \mathbb{P}_x\left(\mathcal{O}_t^L > r\right)$, we can also write

$$\rho_\epsilon^{(r,t)}[L] = \inf\left\{x \geq 0 \colon \mathbb{P}_x\left(\mathcal{O}_t^L > r\right) \leq \epsilon\right\}.$$

Consequently, this risk measure is based on the distribution of $\mathcal{O}_t^L$. This is the analog of the random variable L for the risk measure in Equation (6). A major improvement is that we can now vary both the time horizon and the implementation delay by changing the values of t and r, respectively. The trade-off is that we need the distribution of a strongly path-dependent random variable, namely $\mathcal{O}_t^L$.

For the rest of this paper, we focus on the properties of this VaR-type cumulative Parisian risk measure. In addition, we compare the infinite-time version to the infinite-time risk measure defined in Trufin et al. (2011). Then, we also study the finite-time version as this is possible as soon as the distribution of $\mathcal{O}_t^L$ is available.

Before going any further, let us give some background material on stochastic dominance.

3.1. Stochastic Dominance

Consider two random variables X and Y, and let $\bar{F}_X$ and $\bar{F}_Y$ be their survival functions. We say that X is smaller than Y in the stochastic dominance order, which is denoted by $X \preceq_{st} Y$, if

$$\bar{F}_X(u) \leq \bar{F}_Y(u), \text{ for all } u. \tag{10}$$

Equivalently, for all non-decreasing functions ϕ, we have

$$\mathbb{E}[\phi(X)] \leq \mathbb{E}[\phi(Y)]. \tag{11}$$

Here is a theorem taken from Shaked and Shanthikumar (2007).

Theorem 1.

(i) *Let $\{X_1, X_2, \ldots, X_m\}$ and $\{Y_1, Y_2, \ldots, Y_m\}$ be two finite sets of independent random variables such that $X_i \preceq_{st} Y_i$, for each $i = 1, \ldots, m$. Then, for any increasing function $g\colon \mathbb{R}^m \to \mathbb{R}$, we have*

$$g(X_1, X_2, \ldots, X_m) \preceq_{st} g(Y_1, Y_2, \ldots, Y_m). \tag{12}$$

(ii) *Consider two sequences of random variables $\{X_1, X_2, \ldots\}$ and $\{Y_1, Y_2, \ldots\}$ and two random variables X and Y such that*

$$X_n \xrightarrow{d} X \quad \text{and} \quad Y_n \xrightarrow{d} Y,$$

where $\xrightarrow{d}$ denotes convergence in distribution. If $X_n \preceq_{st} Y_n$ for each n, then $X \preceq_{st} Y$.

(iii) *Let the positive integer-valued random variable N be independent of the family of random variables $\{C_1, C_2, \ldots\}$ and define $S = \sum_{i=1}^{N} C_i$. Define similarly $\tilde{S} = \sum_{i=1}^{\tilde{N}} \tilde{C}_i$.*

If $N \preceq_{st} \tilde{N}$ and $C_i \preceq_{st} \tilde{C}_i$ for each i, then

$$S \preceq_{st} \tilde{S}. \tag{13}$$

Finally, if $X = \{X_t, t \geq 0\}$ and $Y = \{Y_t, t \geq 0\}$ are stochastic processes, then we write $X \preceq_{st} Y$ if, for each $t \geq 0$, we have

$$X_t \preceq_{st} Y_t.$$

The reader is referred to Shaked and Shanthikumar (2007), Kaas et al. (2008) and Denuit et al. (2005) for more details on stochastic ordering and applications in actuarial science.

3.2. Properties of the Risk Measure $\rho_\epsilon^{(r,t)}$

In the following, let L and $\tilde{L}$ be two aggregate loss amount processes associated with two aggregate claim amount processes S and $\tilde{S}$, themselves from two Cramér–Lundberg risk processes X and $\tilde{X}$ as defined in v(1).

Theorem 2. *For $r > 0$, $\varepsilon > 0$ and $t > 0$, we have:*

(i) ***Invariance by translation:*** *For $a > 0$,*

$$\rho_\epsilon^{(r,t)}[L + a] = \rho_\epsilon^{(r,t)}[L] - a. \tag{14}$$

(ii) ***Positive homogeneity:*** *For $b > 0$,*

$$\rho_\epsilon^{(r,t)}[bL] = b\rho_\epsilon^{(r,t)}[L]. \tag{15}$$

(iii) ***Monotonicity:*** *If $L \preceq_{st} \tilde{L}$, then*

$$\rho_\epsilon^{(r,t)}[L] \leq \rho_\epsilon^{(r,t)}[\tilde{L}]. \tag{16}$$

Proof. First, note that

$$\mathbb{P}_x\left(\mathcal{O}_t^{L+a} > r\right) = \mathbb{P}_x\left(\int_0^t \mathbf{1}_{\{L_u > x-a\}} \mathrm{d}u > r\right) = \mathbb{P}_{x-a}\left(\mathcal{O}_t^{L} > r\right).$$

Consequently,

$$\begin{aligned}\rho_{\epsilon}^{(r,t)}[L+a] &= \inf\left\{x \geq 0 \colon \mathbb{P}_x\left(\mathcal{O}_t^{L+a} > r\right) \geq \epsilon\right\} \\ &= \inf\left\{x \geq 0 \colon \mathbb{P}_{x+a}\left(\mathcal{O}_t^{L} > r\right) \geq \epsilon\right\} \\ &= \rho_{\epsilon}^{(r,t)}[L] + a.\end{aligned}$$

This proves Equation (14).

Similarly, if we note that

$$\mathbb{P}_x\left(\mathcal{O}_t^{bL} > r\right) = \mathbb{P}_x\left(\int_0^t \mathbf{1}_{\{L_u > x/b\}} \mathrm{d}u > r\right) = \mathbb{P}_{x/b}\left(\mathcal{O}_t^{L} > r\right),$$

then Equation (15) follows.

To prove the third property, we fix $t > 0$ and we show that, if $L_u \preceq_{st} \tilde{L}_u$ for all $u \leq t$, then

$$\mathcal{O}_t^L \preceq_{st} \mathcal{O}_t^{\tilde{L}}.$$

First, let us define a sequence of discretized versions of the occupation-time process $\mathcal{O}_t^L$. For each $n \geq 1$, choose $0 = t_0 < t_1 < \ldots < t_n = t$ such that $\max\limits_{0 \leq i \leq n}(t_i - t_{i-1}) \to 0$, as $n \to \infty$, and define

$$\mathcal{O}_t^{(n)} = \sum_{i=1}^{n} (t_i - t_{i-1})\, \mathbf{1}_{\{L_{t_i} > x\}}.$$

We define $\widetilde{\mathcal{O}}_t^{(n)}$ in the obvious way, i.e., when S is replaced by $\tilde{S}$. We can re-write $\mathcal{O}_t^{(n)}$ as follows:

$$\mathcal{O}_t^{(n)} = \phi_n\left(L_{t_1} - L_{t_0}, L_{t_2} - L_{t_1}, \ldots, L_{t_n} - L_{t_{n-1}}\right),$$

where $\phi_n(u_1, \ldots, u_n) = \sum_{i=1}^{n}(t_i - t_{i-1})\mathbf{1}_{\left\{\sum_{j=1}^{i} u_j > x + ct_i\right\}}$.

Since $L_u \preceq_{st} \tilde{L}_u$ for all $u \leq t$, then we have $L_{t_i - t_{i-1}} \preceq_{st} \tilde{L}_{t_i - t_{i-1}}$ for each i. Then, since

$$L_{t_i - t_{i-1}} \overset{d}{=} L_{t_i} - L_{t_{i-1}} \quad \text{and} \quad \tilde{L}_{t_i - t_{i-1}} \overset{d}{=} \tilde{L}_{t_i} - \tilde{L}_{t_{i-1}},$$

we have that $L_{t_i} - L_{t_{i-1}} \preceq_{st} \tilde{L}_{t_i} - \tilde{L}_{t_{i-1}}$ for each i. From Equation (12), we obtain

$$\phi_n\left(L_{t_1} - L_{t_0}, L_{t_2} - L_{t_1}, \ldots, L_{t_n} - L_{t_{n-1}}\right) \preceq_{st} \phi_n\left(\tilde{L}_{t_1} - \tilde{L}_{t_0}, \tilde{L}_{t_2} - \tilde{L}_{t_1}, \ldots, \tilde{L}_{t_n} - \tilde{L}_{t_{n-1}}\right),$$

or equivalently

$$\mathcal{O}_t^{(n)} \preceq_{st} \widetilde{\mathcal{O}}_t^{(n)}.$$

Since $\mathcal{O}_t^{(n)} \xrightarrow{d} \mathcal{O}_t^L$ and $\widetilde{\mathcal{O}}_t^{(n)} \xrightarrow{d} \mathcal{O}_t^{\tilde{L}}$, by the second part of Theorem 1, we get

$$\mathcal{O}_t^L \preceq_{st} \mathcal{O}_t^{\tilde{L}}.$$

This means that

$$\mathbb{P}_x\left(\mathcal{O}_t^L > r\right) \leq \mathbb{P}_x\left(\mathcal{O}_t^{\tilde{L}} > r\right), \text{ for all } r.$$

The property in Equation (16) follows. □

The monotonicity property in Equation (16) says that the risk measure $\rho_{\epsilon}^{(r,t)}[L]$ is increasing with respect to the stochastic dominance order. Note that, if $\mathbb{P}\left(L_t \leq \tilde{L}_t\right) = 1$ for all $t \geq 0$, then we can also prove that

$$\rho_{\epsilon}^{(r,t)}[L] \leq \rho_{\epsilon}^{(r,t)}\left[\tilde{L}\right].$$

If we put together the monotonicity property in Equations (16) and (13), then we can deduce the following intuitive relationship: a smaller frequency and a smaller severity yield less occupation time in the red zone and thus a smaller probability of cumulative Parisian ruin. For example, by the third part of Theorem 1, if C and $\tilde{C}$ are exponentially distributed random variables with parameters α and $\tilde{\alpha}$, respectively, and if $\alpha \geq \tilde{\alpha}$, $\lambda \leq \tilde{\lambda}$ and $c = \tilde{c}$, then, for a given common premium rate c, the initial capital needed at a given tolerance level ϵ is larger for X than for $\tilde{X}$.

It is worth mentioning that, as an immediate consequence of Proposition 1, Theorem 2 is also satisfied for the infinite-time horizon risk measure ζ_ϵ. Thus, we have recovered some of the results in Properties 3.1 and 3.2 of Trufin et al. (2011). In addition, an important consequence of Proposition 1 is the stochastic ordering for the finite-time ruin probability $\psi(t,x)$.

When there is no initial reserve, i.e., when $x = 0$, and for $c = \tilde{c}$, the last theorem generalizes Theorem 4 in Goovaerts and De Vylder (1984) and also Proposition 1 of Lefèvre et al. (2017).

3.3. Relationship with Other Risk Measures

Recall that our main object of study is the following VaR-type risk measure: for $r > 0$, $\epsilon > 0$ and $t > 0$,

$$\rho_\epsilon^{(r,t)}[L] = \inf\{x \geq 0 \colon \psi_r(t,x) \leq \epsilon\} = \inf\left\{x \geq 0 \colon \mathbb{P}_x\left(\mathcal{O}_t^L > r\right) \leq \epsilon\right\}. \tag{17}$$

When $t = \infty$, we write $\rho_\epsilon^{(r)}$.

We are also interested in the risk measure based on the finite-time probability of classical ruin:

$$\zeta_\epsilon^{(t)}[L] = \inf\{x \geq 0 \colon \psi(t,x) \leq \epsilon\}. \tag{18}$$

Using the inequality in Equation (9) and the discussions in the previous section, we deduce the following first proposition:

Proposition 1. *For a given time horizon $0 < t \leq \infty$ and an acceptance level $\epsilon > 0$, the risk measure $\rho_\epsilon^{(r,t)}$ is less conservative than the risk measure $\zeta_\epsilon^{(t)}$, i.e.,*

$$\rho_\epsilon^{(r,t)}[L] \leq \zeta_\epsilon^{(t)}[L], \tag{19}$$

and, when $r \to 0$, it converges to $\zeta_\epsilon^{(t)}$, i.e.,

$$\rho_\epsilon^{(r,t)}[L] \uparrow \zeta_\epsilon^{(t)}[L], \text{ as } r \to 0. \tag{20}$$

When the implementation delay r is replaced by copies of an exponentially distributed random variable e_q with rate $q > 0$, then, for $x \in \mathbb{R}$, we have

$$\mathbb{P}_x\left(\sigma_{\mathrm{e}_q} \leq t\right) = 1 - \mathbb{E}_x\left[\mathrm{e}^{-q\mathcal{O}_t^L}\right]. \tag{21}$$

In addition, in this case, cumulative Parisian ruin corresponds to Parisian ruin with exponential delays, that is

$$\kappa^q = \inf\left\{t > 0 \colon t - g_t > \mathrm{e}_q^{g_t}\right\},$$

where $g_t = \sup\{0 \leq s \leq t \colon X_s \geq 0\}$ is is the last time before t when the process was non-negative.

Hence,

$$\psi_{\mathrm{e}_q}(t,x) = \mathbb{P}_x\left(\sigma_{\mathrm{e}_q} \leq t\right) = \mathbb{P}_x\left(\kappa^q \leq t\right).$$

We can then define the following VaR-type risk measure: for $q, r > 0$, $\epsilon > 0$ and $t > 0$,

$$\rho_\epsilon^{(q,r,t)}[L] = \inf\left\{x \geq 0 \colon \psi_{\mathrm{e}_q}(t,x) \leq \epsilon\right\}, \tag{22}$$

In addition, we have

$$\rho_{\epsilon}^{(q,r,t)}[L] \uparrow \rho_{\epsilon}^{(r,t)}[L], \text{ as } q \to \infty. \tag{23}$$

The risk measure $\rho_{\epsilon}^{(q,r,t)}$ satisfies the properties in Theorem 2. For example, we proved that, if $L \preceq_{st} \tilde{L}$, we have $\mathcal{O}_t^L \preceq_{st} \mathcal{O}_t^{\tilde{L}}$. Then, using Equation (12), we obtain

$$h(\mathcal{O}_t^L) \preceq_{st} h(\mathcal{O}^{\tilde{L}}),$$

where $h(x) = 1 - \mathrm{e}^{-qx}$. Hence, using Equation (11), we get

$$\mathbb{E}_x[h(L)] \leq \mathbb{E}_x\left[h(\tilde{L})\right],$$

and then

$$\rho_{\epsilon}^{(q,r,t)}[L] \leq \rho_{\epsilon}^{(q,r,t)}\left[\tilde{L}\right].$$

In addition, as an improvement of the finite-time version of the (infinite-horizon) risk measure defined by Loisel and Trufin (2014), we can define

$$\omega_a^{(t)}[L] := \inf\left\{x \geq 0 \colon \mathbb{E}_x\left[\mathcal{A}_t^L\right] \leq a\right\},$$

where $a > 0$ is a tolerance level for the expected *area in the red* defined as

$$\mathcal{A}_t^L = \int_0^t (L_u - x)_+ \, \mathrm{d}u,$$

where $(x)_+ = \max(x, 0)$. Furthermore, we can use Theorem 1 of Loisel (2005) and then write

$$\mathbb{E}_x\left[\mathcal{A}_t^L\right] = \int_x^{\infty} \mathbb{E}_v\left[\mathcal{O}_t^L\right] \mathrm{d}v = \int_x^{\infty}\int_0^{\infty} \mathbb{P}_v\left(\mathcal{O}_t^L \geq u\right) \mathrm{d}u\mathrm{d}v. \tag{24}$$

Consequently, if we suppose that $L \preceq_{st} \tilde{L}$, then $\mathcal{O}_t^L \preceq_{st} \mathcal{O}_t^{\tilde{L}}$ and then, from Equations (10) and (24), we have

$$\mathbb{E}_x\left[\mathcal{A}_t^L\right] \leq \mathbb{E}_x\left[\mathcal{A}_t^{\tilde{L}}\right].$$

Thus,

$$\omega_a^{(t)}[L] \leq \omega_a^{(t)}\left[\tilde{L}\right], \tag{25}$$

which corresponds to Property 3.1 in Loisel and Trufin (2014).

Remark 1. *Note that, with the distribution in Theorem 3, it is possible to compute the finite-time version of this risk measure based on the area in the red in the case of a Cramér–Lundberg process with exponential claims.*

4. Example: Cramér–Lundberg Model with Exponential Claims

In this section, we want to see how $\rho_{\epsilon}^{(r,t)}$ reacts to changes in the value of its parameters. In other words, we want to perform a sensitivity analysis.

In general, we could use Monte Carlo simulations to compute values for $\rho_{\epsilon}^{(r,t)}$. However, if we consider a Cramér–Lundberg process with exponentially distributed claims $\{C_1, C_2, \dots\}$ with rate parameter $\alpha > 0$, then there exists an explicit expression for the distribution of the occupation time for a finite-time horizon. Unfortunately, such formulas are not available for most claim distributions.

Theorem 3 (Guérin and Renaud (2017)). *For $t > 0$, we have*

$$\mathbb{P}_x\left(\mathcal{O}_t^L \in \mathrm{d}s\right) = a_t^x \delta_0(\mathrm{d}s) + \left(a_{t-s}^x + k_{t-s}^x\right)\left(\lambda - c\alpha\left(1 - a_s^0\right)\right)\mathbf{1}_{(0,t)}(s)\,\mathrm{d}s,$$

with

$$a_t^x = 1 - \lambda e^{-\alpha x} \int_0^t e^{-(\lambda + c\alpha)s} \left[I_0 \left(2\sqrt{\lambda c \alpha s \left(s + x/c \right)} \right) - \frac{s}{s + x/c} I_2 \left(2\sqrt{\lambda c \alpha s \left(s + x/c \right)} \right) \right] ds$$

and

$$k_t^x = e^{-\alpha x} - 1 + \lambda x \alpha e^{-\alpha x} \int_0^t e^{-(\lambda + c\alpha)s} \left[I_0 \left(2\sqrt{\lambda c \alpha s \left(s + x/c \right)} \right) - I_2 \left(2\sqrt{\lambda c \alpha s \left(s + x/c \right)} \right) \right] ds,$$

where I_v represents the modified Bessel function of the first kind of order v.

In Theorem 3, a_t^x is the survival ruin probability over $[0, t]$, that is

$$\begin{aligned} a_t^x &= 1 - \psi(t, x) \\ &= 1 - \lambda e^{-\alpha x} \int_0^t e^{-(\lambda + c\alpha)s} \left[I_0 \left(2\sqrt{\lambda c \alpha s \left(s + x/c \right)} \right) - \frac{s}{s + x/c} I_2 \left(2\sqrt{\lambda c \alpha s \left(s + x/c \right)} \right) \right] ds. \end{aligned}$$

For an infinite-time horizon, we have the well-known expression:

$$a^x = \lim_{t \to \infty} a_t^x = 1 - \psi(x) = \frac{\lambda}{c\alpha} e^{x(\lambda/c - \alpha)} = \frac{1}{1 + \eta} e^{-x\alpha\eta/(1+\eta)}.$$

From Corollary 2 in Renaud (2014), we can deduce the following expression for the distribution of $\mathcal{O}_\infty^L$, when the claims are exponentially distributed.

Corollary 1. *For any $x \in \mathbb{R}$, we have*

$$\begin{aligned} &\mathbb{P}_x \left(\mathcal{O}_\infty^L \in ds \right) = a^x \delta_0 \left(ds \right) \\ &+ \frac{\lambda}{c} \left(1 - \frac{\lambda}{c\alpha} \right) e^{-cs\alpha} e^{-x(\alpha - \lambda/c)} \left(c + \sum_{i=0}^{\infty} \frac{(\lambda s)^{i+1}}{i!\,(1+i)!} \left(c\Gamma \left(i + 1, s\lambda \right) - \frac{c}{s\lambda} \Gamma \left(i + 2, s\lambda \right) \right) \right), \end{aligned}$$

where $\Gamma(a, x) = \int_0^x e^{-t} t^{a-1} dt$ is the incomplete gamma function.

The explicit formula in Theorem 3 allows for a sensitivity analysis of the value of the probability of cumulative Parisian ruin, when claims are exponentially distributed, with respect to the delay parameter r and the time horizon t. In Figure 2, we observe that for a fixed delay parameter r, the probability of cumulative Parisian ruin increases when the time horizon t increases. This is because we accumulate more occupation time. On the other hand, it decreases when the delay r increases. For a fixed value of the time horizon t, increasing the initial capital x decreases the probability of cumulative Parisian ruin, as expected.

For the corresponding risk measures, Figure 3 illustrates the relationships in Equations (19) and in (20) between $\rho_\epsilon^{(r,t)}$ and $\zeta_\epsilon^{(t)}$. As $r \to 0$, i.e., as the grace period gets smaller, the initial capital needed with $\rho_\epsilon^{(r,t)}$ increases toward that needed with $\zeta_\epsilon^{(t)}$, both at a tolerance level of $\epsilon = 0.3$. When the time horizon t increases, both risk measures increase the initial capital needed for that tolerance level.

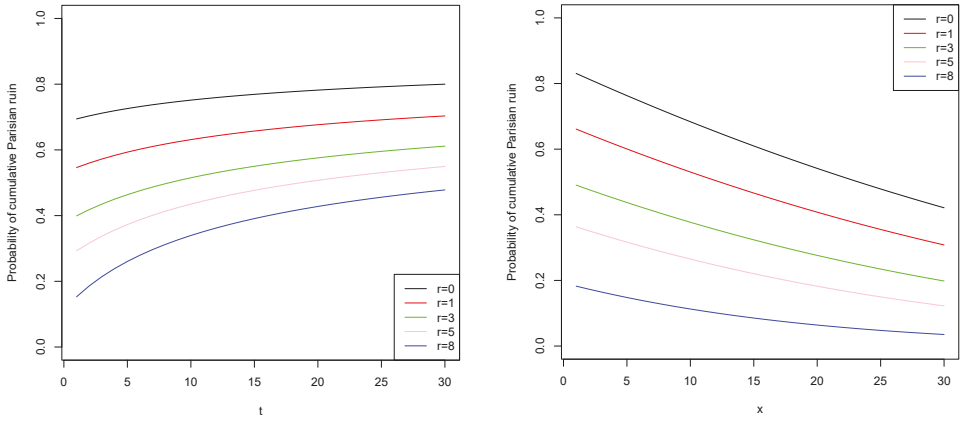


Figure 2. The probability of cumulative Parisian ruin for the Cramér–Lundberg process with $\alpha = 1/8$, $\lambda = 2$, $c = 17$, $r = 1$, $x = 10$ and $t = 10$.

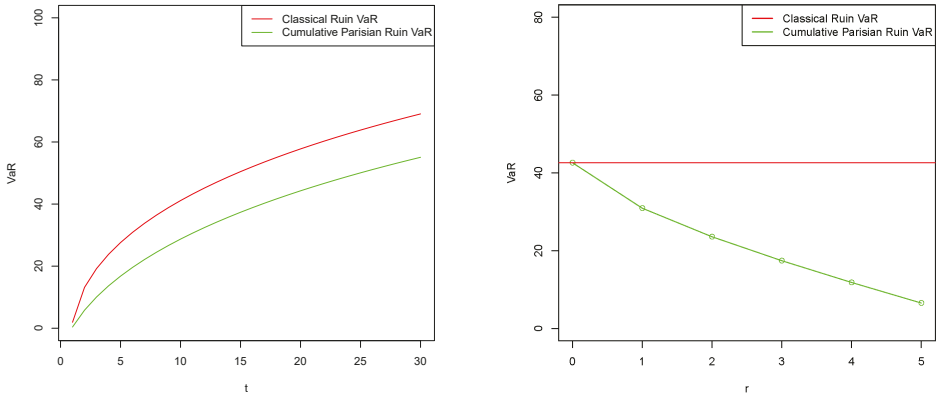


Figure 3. Risk measures $\rho_{\epsilon}^{(r,t)}$ and $\zeta_{\epsilon}^{(t)}$ for the Cramér–Lundberg process with $\alpha = 1/8$, $\lambda = 2$, $c = 17$, $t = 10$, $r = 2$, and $\epsilon = 0.3$.

5. Conclusions

In this paper, we study a VaR-type risk measure derived from cumulative Parisian ruin for the Cramér–Lundberg risk process. Precisely, this measure is defined as the smallest amount of capital for which the associated cumulative Parisian ruin probability is less than or equal to a tolerable level. We derive some properties of this risk measure and we provide some relationships with other risk measures. Finally, for exponentially distributed claims sizes, we performed sensitivity analysis of the values of the probability of cumulative Parisian ruin and the risk measure. Our risk measure could still be used for other risk processes such as the Brownian motion risk model.

Author Contributions: All the authors collaborated equally at every step of the research and in writing the paper.

Acknowledgments: Funding in support of this work was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC). M.A.L. thanks the Institut des sciences mathématiques (ISM) and the Faculté des sciences at UQAM for their doctoral scholarships.

Conflicts of Interest: The authors declare no conflict of interest.

References

Beekman, John A. 1985. A series for infinite time ruin probabilities. *Insurance: Mathematics and Economics* 4: 129–34. [CrossRef]

Cramér, Harald. 1930. *On the Mathematical Theory of Risk*. Alingsås: Centraltryckeriet.

Guérin, Hélène, and Jean-François Renaud. 2017. On the distribution of cumulative Parisian ruin. *Insurance: Mathematics and Economics* 73: 116–23. [CrossRef]

Goovaerts, Marc, and F. Etienne De Vylder. 1984. Dangerous Distributions and Ruin Probabilities in the Classical Risk Model. Paper presented at the 22nd International Congress of Actuaries, Sydney, Australia, October 21–27; pp. 111–20.

Denuit, Michel, Jan Dhaene, Marc Goovaerts, and Rob Kaas. 2005. *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. New York: Wiley.

Shaked, Moshe, and George Shanthikumar. 2007. *Stochastic Orders*. New York: Springer.

Kaas, Rob, Marc Goovaerts, Jan Dhaene, and Michel Denuit. 2008. *Modern Actuarial Risk Theory Using R*. Berlin/Heidelberg: Springer.

Loisel, Stéphane. 2005. Differentiation of some functionals of risk processes, and optimal reserve allocation. *Journal of Applied Probability* 42: 379–92. [CrossRef]

Loisel, Stéphane, and Julien Trufin. 2014. Properties of a risk measure derived from the expected area in red. *Insurance: Mathematics and Economics* 55: 191–9. [CrossRef]

Lefèvre, Claude, Julien Trufin, and Pierre Zuyderhoff. 2017. Some comparison results for finite-time ruin probabilities in the classical risk model. *Insurance: Mathematics and Economics* 77: 143–49. [CrossRef]

Lundberg, Filip. 1903. *I. Approximerad framställning af sannolikhetsfunktionen. II. Återförsäkring af kollektivrisker. Akademisk afhandling, etc*. Uppsala: Almqvist & Wiksell.

Renaud, Jean-François. 2014. On the time spent in the red by a refracted Lévy risk process. *Journal of Applied Probability* 51: 1171–88. [CrossRef]

Mitric, Ilie-Radu, and Julien Trufin. 2016. On a risk measure inspired from the ruin probability and the expected deficit at ruin. *Scandinavian Actuarial Journal* 2016: 932–51. [CrossRef]

Trufin, Julien, Hansjoerg Albrecher, and Michel Denuit. 2011. Properties of a risk measure derived from ruin theory. *The Geneva Risk and Insurance Review* 36: 174–88. [CrossRef]

risks

MDPI

Article

Numerical Ruin Probability in the Dual Risk Model with Risk-Free Investments

Sooie-Hoe Loke [1,*] and Enrique Thomann [2]

[1] Department of Mathematics, Central Washington University, 400 East University Way, Ellensburg, WA 98926, USA
[2] Department of Mathematics, Oregon State University, Corvallis, OR 97331-4605, USA; thomann@math.oregonstate.edu
* Correspondence: SooieHoe.Loke@cwu.edu; Tel.: +1-509-963-2602

Received: 17 August 2018; Accepted: 27 September 2018; Published: 1 October 2018

Abstract: In this paper, a dual risk model under constant force of interest is considered. The ruin probability in this model is shown to satisfy an integro-differential equation, which can then be written as an integral equation. Using the collocation method, the ruin probability can be well approximated for any gain distributions. Examples involving exponential, uniform, Pareto and discrete gains are considered. Finally, the same numerical method is applied to the Laplace transform of the time of ruin.

Keywords: ruin probability; dual risk model; constant interest rate; integral equation; Laplace transform; numerical approximation

1. Introduction

The simplest surplus model in non-life insurance is known as the Cramér–Lundberg model or the classical risk model. It assumes that the company collects premiums at a fixed rate and pays out claims of a random amount at random times. Mathematically, the surplus process can be written as:

$$U_t = u + ct - \sum_{k=1}^{N_t} X_k, \tag{1}$$

where u is the initial capital, c is the constant premium rate, $X_1, X_2, \ldots$ are the claim sizes and N_t is the claim arrival process, which counts the number of claims in the time interval $[0, t]$. It is vital for any company to operate above a certain income level. For convenience, this level is set to be zero. Define the time of ruin by:

$$\tau = \inf\{s > 0 \mid U_s < 0\},$$

where if $U_s \geq 0$ for all s, then $\tau = \inf \emptyset = \infty$. Define the ruin probability with initial surplus $U_0 = u$ by:

$$\psi(u) = P(\tau < \infty \mid U_0 = u).$$

The classical model (1) assumes that the only source of income is from collecting premiums. In the past, many models incorporated investments with constant force of interest, for example, investing all (or part) of the surplus in bonds or time accounts. The study of these risk models dated back to Segerdahl (1942), who considered the constant interest risk model and provided an explicit expression for the ruin probability when the claims are exponentially distributed. Sundt and Teugels (1995) gave an extensive treatment of the ruin probability with constant interest force and obtained approximations, as well as upper and lower bounds. Kasozi and Paulsen (2005) used numerical

methods such as the block-by-block method and the Simpson rule to approximate the ultimate ruin probabilities under a constant rate of interest. Cai et al. (2009) considered the well-known Gerber–Shiu function under the risk model of liquid reserves and constant interest on the surplus, and more recently, Schmidli (2015) studied a variant of the discounted penalty function where a penalty applies when the surplus process leaves a finite interval. Yang and Wang (2010) investigated the asymptotic behavior of the ruin probability of some negatively dependent risk models with a constant interest rate and dominatedly-varying-tailed claims. Renewal risk models with constant interest were well studied by Konstantinides et al. (2010).

A dual model to (1) is obtained by regarding premiums as expenses and claims as gains. In life annuity insurance or pension insurance, a basic model for the surplus process is known as the classical dual risk model, which is given by:

$$U_t = u - ct + \sum_{k=1}^{N_t} X_k. \tag{2}$$

Here, c denotes the expense rate, $\{X_k\}$ represent random gains and N_t is called the gain arrival process. The process N_t is assumed to be a homogenous Poisson process with intensity $\lambda > 0$. Moreover, it is assumed that the claim sizes $X_1, X_2, \ldots$ are independent and identically distributed (i.i.d) with cumulative distribution function (c.d.f.) F and tail distribution function $\overline{F} = 1 - F$ and that the processes $\{X_i\}$ and N_t are independent.

Therefore, (2) models the surplus process of a company with a constant rate of consumption, earning random income at random times. Other examples of such companies are non-profit organizations and petroleum companies where the jumps correspond to random donations (see Chen (2010)) and discoveries of oil (see Avanzi et al. (2007)), respectively. Throughout this paper, the terms "gains", "innovations" and "donations" will be used interchangeably.

Contrary to the vast literature on the insurance models, there were very few results published in the dual model with constant force of interest. Zeng and Xu (2013) considered the perturbed dual risk model with constant interest and a threshold dividend strategy. They used the sinc method to approximate the expected present value of total dividends. Dong and Wang (2008) studied the renewal risk model with constant force of interest and obtained an explicit expression for the ruin probability in terms of infinite series of iterated integrals. Although their renewal model is more general than the Poisson process considered here, the objective and approach of this paper differ from Dong and Wang (2008).

Here, the main objective is to examine the ruin probability numerically in the dual risk model with risk-free investments under an arbitrary gain distribution. Section 2 shows the derivation of the integro-differential equation (IDE) that is satisfied by ruin probability. From the IDE, Section 3 carefully demonstrates that the derivative of the ruin probability satisfies an integral equation (IE). The ruin probability for certain gain sizes (such as exponential and a mixture of exponentials) has a very explicit representation. In Section 4, using the collocation method, the IE is reduced to a linear matrix equation. The ruin probability can then be obtained numerically for any jump distributions. For numerical illustrations, exponential, uniform, Pareto and discrete gains are considered. In Section 5, the same numerical method is applied to other functionals of the time of ruin, such as the Laplace transform of the ruin time. Exponential gains are again considered to illustrate the numerical scheme. Finally, Section 6 concludes the findings of this paper.

2. The Model

All assumptions in the classical dual model (2) are retained, and it is further assumed that the company invests all of its surplus in a risk-free asset with constant force of interest $a > 0$. The surplus process can now be written as:

$$U_t = u - ct + a\int_0^t U_s\, ds + \sum_{i=1}^{N_t} X_i. \tag{3}$$

Define the consumption (or investment) process $\{Z_t\}$ by:

$$Z_t = u - ct + a\int_0^t Z_s\, ds,$$

or equivalently, in differential form,

$$dZ_t = (aZ_t - c)dt, \qquad Z_0 = u.$$

Observe that $\{Z_t\}$ is a deterministic process and that the above is an ordinary differential equation (ODE) with solution:

$$Z_t = \left(u - \frac{c}{a}\right)e^{at} + \frac{c}{a}, \qquad t \geq 0.$$

Since $Z_t \geq 0$ for all t whenever $u \geq \frac{c}{a}$, one has:

$$\psi(u) = 0 \qquad \text{for all } u \geq \frac{c}{a}. \tag{4}$$

In other words, if the initial capital is greater than $\frac{c}{a}$, the contribution from the risk-free investments always offsets the expenses, and so, the company can never be ruined. Hence, for the ruin probability $\psi(u)$, only values of u between zero and $\frac{c}{a}$ are of interest.

Theorem 1. *Assume that the ruin probability $\psi(u)$ is differentiable. For $0 < u < \frac{c}{a}$, the ruin probability satisfies the IDE:*

$$(au - c)\psi'(u) - \lambda\psi(u) + \lambda\int_0^{c/a-u} \psi(u+x)\, dF(x) = 0, \tag{5}$$

with boundary conditions:

$$\psi(0) = 1 \qquad \text{and} \qquad \psi\left(\frac{c}{a}\right) = 0. \tag{6}$$

Proof. Consider the risk process U_t in an infinitesimal time interval $(0, h)$ where $h < \frac{1}{a}\left|\ln\frac{c}{c-au}\right|$. Therefore,

$$\begin{aligned}\psi(u) &= \mathbb{E}(\mathbb{1}_{[\tau<\infty]}\mathbb{1}_{[N(h)=0]} \mid U_0 = u) + \mathbb{E}(\mathbb{1}_{[\tau<\infty]}\mathbb{1}_{[N(h)=1]} \mid U_0 = u) + \mathbb{E}(\mathbb{1}_{[\tau<\infty]}\mathbb{1}_{[N(h)>1]} \mid U_0 = u)\\ &= e^{-\lambda h}\psi(Z_h) + \int_0^h \lambda e^{-\lambda s}\int_0^{c/a-Z_s} \psi(Z_s + x)\, dF(x)\, ds + o(h).\end{aligned} \tag{7}$$

By the chain rule,

$$\lim_{h\to 0}\frac{\psi(Z_h) - \psi(u)}{h} = \psi'(u)\cdot a(u - c/a).$$

Dividing (7) by h and taking the limit as $h \to 0$, one obtains:

$$0 = (au - c)\psi'(u)e^0 + \psi(u)\cdot(-\lambda) + \lambda e^0\int_0^{c/a-u} \psi(u+x)dF(x).$$

□

Remark 1. *Theorem 1 coincides with Equation (20) in Dong and Wang (2008) when $n = 1$.*

Remark 2. *In the classical risk model with risk-free investments, it has been proven that the ruin probability is continuously differentiable on the positive real line, except at points of discontinuity of the jump distribution (for example, see Mishura and Ragulina (2016), Chapter 2, and the references therein). One can verify that the same is true for the dual model considered in this paper.*

3. Integral Equation Approach

Here, the derivative of ψ is shown to satisfy a simple integral equation. This provides a framework for the numerical approximation of the ruin probability in the next section.

For $u \geq 0$, define:

$$\chi(u) \equiv \begin{cases} \psi'(u), & 0 \leq u \leq \frac{c}{a}, \\ 0, & u > \frac{c}{a}. \end{cases}$$

Using (4), for $0 \leq u \leq \frac{c}{a}$, we have:

$$\begin{aligned} \psi(u) - \int_0^{c/a-u} \psi(u+x)\, dF(x) &= \int_0^{c/a-u} (\psi(u) - \psi(u+x))\, dF(x) + \int_{c/a-u}^{\infty} (\psi(u) - 0)\, dF(x) \\ &= \int_0^{\infty} \int_{u+x}^{u} \chi(t)\, dt\, dF(x) \\ &= -\int_u^{\infty} \overline{F}(t-u)\, \chi(t)\, dt. \end{aligned}$$

Therefore, the IDE (5) becomes an IE given by:

$$(au - c)\chi(u) = -\lambda \int_u^{c/a} \overline{F}(t-u)\, \chi(t)\, dt, \qquad 0 < u < \frac{c}{a}.$$

Denoting $\frac{c}{a} = b$ and $\frac{\lambda}{a} = d$, the above IE can be rewritten as:

$$\chi(u) = \frac{d}{b-u} \int_u^b \overline{F}(t-u)\, \chi(t)\, dt, \qquad 0 < u < b, \tag{8}$$

with the integrability condition:

$$\int_0^b \chi(u)\, du = \psi(b) - \psi(0) = -1. \tag{9}$$

Performing the change of variable $\chi(u) = (b-u)^{d-1}\tilde{\chi}(u)$, (8) becomes:

$$(b-u)^{d-1}\tilde{\chi}(u) = \frac{d}{b-u} \int_u^b \overline{F}(t-u)\ (b-t)^{d-1}\tilde{\chi}(t)\, dt, \qquad 0 < u < b.$$

Therefore,

$$\begin{aligned} (b-u)^d \tilde{\chi}(u) &= -\int_u^b \overline{F}(t-u)\, \tilde{\chi}(t)\, \frac{d}{dt}(b-t)^d\, dt \\ &= \overline{F}(t-u)\, \tilde{\chi}(t)(b-t)^d\ |_b^u + \int_u^b (b-t)^d \frac{d}{dt} \left[\overline{F}(t-u)\, \tilde{\chi}(t)\right]\, dt, \end{aligned}$$

and so, one obtains:

$$0 = \int_u^b (b-t)^d \frac{d}{dt} \left[\overline{F}(t-u)\, \tilde{\chi}(t)\right]\, dt, \qquad 0 < u < b. \tag{10}$$

The IE (10) can now be used to obtain numerical approximations for any arbitrary jump distributions. In certain cases, the explicit expression of ψ is simple enough, and this can be used to validate the numerical approximation in the next section.

Example 1. *Assume that the gains are exponentially distributed with mean μ, that is, $\overline{F}(x) = e^{-x/\mu}$. Then, (10) reduces to:*

$$0 = e^{u/\mu} \int_u^b (b-t)^d \frac{d}{dt} \left[e^{-t/\mu}\, \tilde{\chi}(t)\right]\, dt.$$

Since the above holds for all u, the integral term must be equal to zero. Differentiating this equation leads to the ODE $\tilde{\chi}'(u) - \frac{1}{\mu}\tilde{\chi}(u) = 0$, with the solution given by $\tilde{\chi}(u) = e^{u/\mu}$. Therefore,

$$\chi(u) = (b-u)^{\lambda/a-1}e^{u/\mu}$$

and the boundary conditions (6) are then used to obtain:

$$\psi(u) = \frac{\Gamma(\frac{\lambda}{a},0) - \Gamma(\frac{\lambda}{a},\frac{c-au}{a\mu})}{\Gamma(\frac{\lambda}{a},0) - \Gamma(\frac{\lambda}{a},\frac{c}{a\mu})}, \qquad 0 \le u \le c/a, \tag{11}$$

where $\Gamma(b,x) := \int_x^\infty t^{b-1}e^{-t}\,dt$ is the incomplete Gamma function.

Example 2. *Assume that the gains are distributed as a mixture of two exponentials, i.e., $\overline{F}(x) = pe^{-\mu_1 x} + (1-p)e^{-\mu_2 x}$ where w.l.o.g. $\mu_1 > \mu_2$ and $0 \le p < 1$. Then, (10) can be written as:*

$$0 = pe^{\mu_1 u}\int_u^b (b-t)^d \frac{d}{dt}\left[e^{-\mu_1 t}\,\tilde{\chi}(t)\right]\,dt + (1-p)e^{\mu_2 u}\int_u^b (b-t)^d \frac{d}{dt}\left[e^{-\mu_2 t}\,\tilde{\chi}(t)\right]\,dt. \tag{12}$$

Differentiating (12) twice and simplifying yield a second order linear ODE:

$$(b-u)\tilde{\chi}''(u) - ((b-u)(\mu_1+\mu_2)+d)\tilde{\chi}'(u) + (\mu_1\mu_2(b-u) + d(\mu_1 p + \mu_2(1-p)))\tilde{\chi}(u) = 0.$$

From Polyanin and Zaitsev (2002), the general solution to the above ODE is given by:

$$\tilde{\chi}(u) = c_1 e^{\mu_1 u}\,{}_1F_1(d(1-p), d, (\mu_1-\mu_2)(b-u)) + c_2 e^{\mu_1 u}U(d(1-p), d, (\mu_1-\mu_2)(b-u)),$$

where ${}_1F_1$ and U are the confluent hypergeometric functions (see Slater (1960)) and c_1, c_2 are arbitrary constants. Therefore, the ruin probability is given by:

$$\psi(u) = 1 - c_1\int_0^u \Psi_1(s)\,ds - c_2\int_0^u \Psi_2(s)\,ds,$$

where:

$$\Psi_1(u) = e^{\mu_1 u}\left(\frac{c}{a}-u\right)^{\lambda/a-1}{}_1F_1\left(\frac{(1-p)\lambda}{a}, \frac{\lambda}{a}, (\mu_1-\mu_2)\left(\frac{c}{a}-u\right)\right),$$
$$\Psi_2(u) = e^{\mu_1 u}\left(\frac{c}{a}-u\right)^{\lambda/a-1}U\left(\frac{(1-p)\lambda}{a}, \frac{\lambda}{a}, (\mu_1-\mu_2)\left(\frac{c}{a}-u\right)\right).$$

The boundary conditions are then used to determine the constants c_1 and c_2, as shown in Dong and Wang (2008).

Remark 3. *In general, an explicit expression for ψ can be obtained if the c.d.f. of the gains satisfies an ODE with constant coefficients. The techniques used in Examples 1 and 2 can be applied to the case of the mixture of n exponentials, i.e., $\overline{F}(x) = \sum_{i=1}^n p_i e^{-\mu_i x}$ where $0 \le p_i < 1$ with $\sum_{i=1}^n p_i = 1$, which is an important class of distributions since any positive distributions can be approximated by the mixture of exponentials. If the c.d.f. is not of this form, e.g., uniform or Pareto gains, explicit formulas are difficult to obtain. This leads to the next section, which provides numerical approximations for ψ under arbitrary gains.*

4. Numerical Scheme

In this section, a numerical framework for the ruin probability under any gain distributions is presented. The numerical scheme requires solving a simple linear system $A\tilde{\chi} = Y$ where A and

Y are to be determined and the vector $\tilde{\chi}$ consists of values of the function $\tilde{\chi}(u)$ evaluated at some discrete points.

Consider the partition $0 = u_0 < u_1 < \cdots < u_{N-1} < u_N = b$. For each $j = 1, \ldots, N-1$, (10) can be written as:

$$0 = \int_{u_j}^{b} (b-t)^d \frac{d}{dt}[\overline{F}(t-u_j)\tilde{\chi}(t)]\, dt.$$

For each fixed u_j, discretize in the t variable where the length of Δt coincides with the length of Δu. The derivative term is approximated using the forward difference method. Defining $\tilde{\chi}_i \approx \tilde{\chi}(u_i)$, one arrives at the following system of equations:

$$0 = \sum_{k=j}^{N-1} (b-u_k)^d \left[\overline{F}(u_{k+1}-u_j)\tilde{\chi}_{k+1} - \overline{F}(u_k-u_j)\tilde{\chi}_k\right], \qquad j = 1, \ldots, N-1,$$

which can be rewritten as:

$$0 = -(b-u_j)^d \tilde{\chi}_j + \sum_{k=j+1}^{N} \overline{F}(u_k-u_j)\tilde{\chi}_k[(b-u_{k-1})^d - (b-u_k)^d], \;\; j = 1, \ldots, N-1.$$

The integrability condition (9) is approximated by:

$$\sum_{i=1}^{N-1} (b-u_i)^{d-1}\tilde{\chi}_i = -1.$$

Therefore, one obtains the matrix equation:

$$A\tilde{\chi} = Y, \tag{13}$$

where $\tilde{\chi} = [\tilde{\chi}_1, \cdots, \tilde{\chi}_N]^T$, $Y = [0, \cdots, 0, 1]^T$, and $A = (a_{jk})$ where:

$$a_{jk} = \begin{cases} \overline{F}(u_k-u_j)[(b-u_{k-1})^d - (b-u_k)^d], & j = 1, \ldots, N-1,\ k = j+1, \ldots, N \\ -(b-u_j)^d, & j = 1, \ldots, N-1,\ k = j, \\ -(b-u_k)^{d-1}, & j = N,\ k = 1, \ldots, N-1, \\ 0 & \text{else.} \end{cases} \tag{14}$$

To solve (13), the matrix A needs to be invertible. The following lemma asserts this claim.

Lemma 1. *If B is an $N \times N$ matrix of the form:*

$$B = \begin{bmatrix} -b_{1,1} & b_{1,2} & \ldots & b_{1,N-1} & b_{1,N} \\ 0 & -b_{2,2} & \ldots & b_{2,N-1} & b_{2,N} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & -b_{N-1,N-1} & b_{N-1,N} \\ -b_{N,1} & -b_{N,2} & \ldots & -b_{N,N-1} & 0 \end{bmatrix},$$

where each $b_{j,k} > 0$, then $\det(B) \neq 0$.

Proof. It suffices to show that B can be reduced, using elementary row operations, to an upper triangular matrix with non-zero diagonal entries. First, perform $R_1 \leftarrow R_1 - \frac{b_{1,1}}{b_{N,1}} R_N$, which results in a matrix with the first row of the form $[0\, b'_{1,2} \cdots b'_{1,N-1}\, b_{1,N}]$ where each $b'_{1,k}$ is positive. Inductively, perform the row operations $R_1 \leftarrow R_1 + \frac{b'_{1,j}}{b_{j,j}} R_j$ for each $j = 2, \ldots, N-1$. These operations result in a

matrix with the first row of the form $[0\,0\,\cdots\,0\,b^*]$ where $b^* > 0$. Finally, swap the first row with the last row to obtain an upper triangular matrix with non-zero diagonal entries. □

Thus, $\tilde{\chi} = A^{-1}Y$. To recover χ, for each $j = 1, \ldots, N-1$, multiply the j-th component of $\tilde{\chi}$ by $(b-u_j)^{d-1}$. Since $\chi = \psi'$, performing the numerical integration:

$$\psi_j = \sum_{k=j}^{N} \chi_i$$

yields the vector $\psi \equiv [\psi_1, \cdots, \psi_N]^T$ where $\psi_i \approx \psi(u_i)$. Finally, the vector ψ is normalized by imposing the boundary condition $\psi(0) = 1$.

The following four examples show the approximate ruin probabilities under different gain distributions. Naturally, the first example concerns the exponentially-sized gains.

Example 3. *Suppose that* $\overline{F}(x) = e^{-x}$. *Let* $\frac{c}{a} = b = 4$ *and* $\frac{\lambda}{a} = d = 3.5$. *From Example 1,*

$$\psi(u) = \frac{\Gamma(1.5,0) - \Gamma(1.5,4-u)}{\Gamma(1.5,0) - \Gamma(1.5,4)}.$$

Figure 1 shows that the numerical approximation of ψ *approaches the exact solution as the number of subintervals increases. The maximum error is used for error analysis, which is the largest absolute difference between the approximation and the true value of* ψ. *The resulting error suggests that this numerical scheme has a first order of accuracy, which is unsurprising since first order discretization is used for the integral term, as well as for the derivative term.*

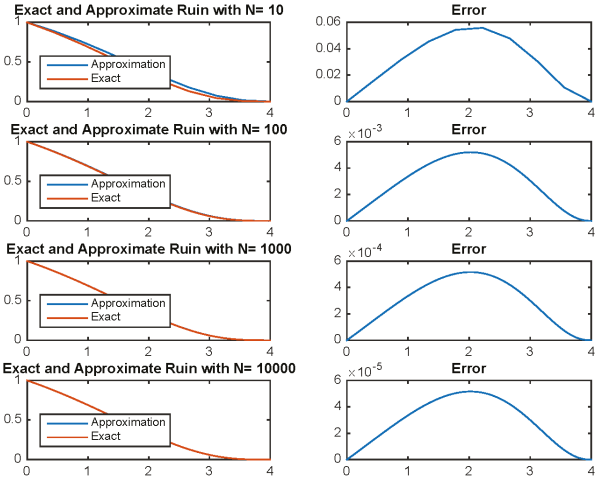


Figure 1. Exact and approximate ruin probabilities for exponential mean one gains for 10, 100, 1000 and 10,000 subintervals.

The next three examples concern gains with c.d.f. that does not satisfy any ODE with constant coefficients, and so, the exact expression for the ruin probability is complicated. In all three examples, varying one of the parameters in the model leads to some interesting yet intuitive results.

Example 4. *Consider uniformly distributed gains on the interval* $[0, \theta]$*. Then,*

$$\overline{F}(x) = \begin{cases} 1, & x \leq 0 \\ 1 - \frac{x}{\theta}, & 0 < x \leq \theta \\ 0, & x > \theta. \end{cases}$$

For the numerical experiment, set $\frac{c}{a} = b = 4$, $\frac{\lambda}{a} = d = 1.5$ *and* $N = 100$*. Figure 2 shows that the numerical approximation of* ψ *approaches a limiting curve as the parameter* θ *increases.*

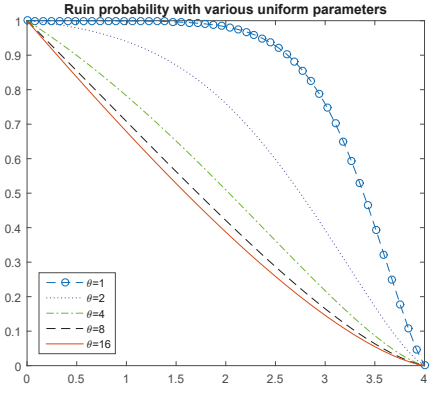


Figure 2. Ruin probabilities for Uniform$[0, \theta]$ gains for $\theta = 1, 2, 4, 8, 16$.

This limiting curve can be found by investigating the IDE (5). For each parameter θ*, denote the associated ruin probability by* ψ_θ*. Then, the IDE is given by:*

$$(au - c)\psi'_\theta(u) - \lambda\psi_\theta(u) + \lambda \int_0^{c/a-u} \frac{1}{\theta} \mathbb{1}_{[0<x<\theta]} \psi_\theta(u + x)\, dx = 0.$$

Now,

$$\frac{\lambda}{\theta} \int_0^{c/a-u} \mathbb{1}_{[0<x<\theta]} \psi_\theta(u + x)\, dx = \frac{\lambda}{\theta} \int_u^{\min\{c/a,\theta\}} \psi_\theta(y)\, dy \to 0$$

as $\theta \to \infty$ *by the dominated convergence theorem since* $0 \leq \psi_\theta \leq 1$*. Hence, as* $\theta \to \infty$*, the IDE becomes:*

$$(au - c)\psi'_\infty(u) - \lambda\psi_\infty(u) = 0,$$

and so, the solution of the above that satisfies the boundary conditions (6) is given by:

$$\psi_\infty(u) = \left(1 - \frac{u}{b}\right)^{\lambda/a}, \qquad 0 < u < b.$$

Observe that there is a natural monotonicity in this case. If $\theta_1 < \theta_2$*, then* $\psi_{\theta_1}(u) > \psi_{\theta_2}(u)$ *for all* $0 < u < b$*. This property agrees with the model, since larger gains will reduce the probability of ruin.*

The next example features a heavy tail gain distribution.

Example 5. *Suppose that the gains follow a Pareto distribution, whose tail distribution is given by:*

$$\overline{F}(x) = \begin{cases} 1, & x \leq x_m \\ \left(\frac{x_m}{x}\right)^\alpha, & x > x_m. \end{cases}$$

For the numerical illustration, set $\frac{c}{a} = b = 4$, $\frac{\lambda}{a} = d = 3.5$ *and* $N = 100$. *Figure 3 shows that the numerical approximation of* ψ *approaches a limiting curve as the parameter* α *increases.*

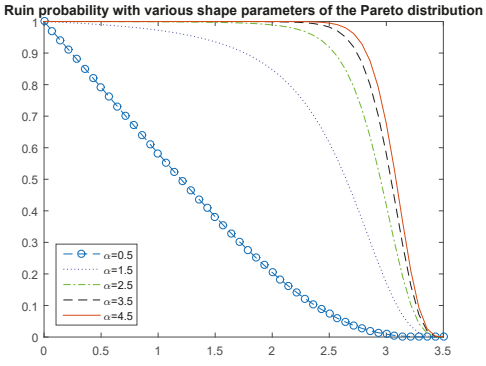


Figure 3. Ruin probabilities for Pareto$(0.1, \alpha)$ gains for $\alpha = 0.5, 1.5, 2.5, 3.5, 4.5$.

The limiting function can be thought of as follows. As $\alpha \to \infty$, X *converges to a constant random variable* x_m, *that is the p.d.f.* $f(x) \to \delta_{x_m}$. *Hence, from the IDE (5), one obtains the following delay differential equation (DDE):*

$$(au - c)\psi'(u) - \lambda\psi(u) + \lambda\psi(u + x_m) = 0, \qquad 0 < x_m < b - u, \tag{15}$$

$$(au - c)\psi'(u) - \lambda\psi(u) = 0, \qquad x_m > b - u. \tag{16}$$

For $u > b - x_m$, *the solution is given by* $\psi(u) = \left(1 - \frac{u}{b}\right)^{\lambda/a}$. *Once this is known, (15) becomes a linear nonhomogenous first order ODE, which is then solvable for* $b - 2x_m < u < b - x_m$. *One can then proceed inductively to obtain the solution to (15). This method of solving DDEs is also called the method of steps.*

Finally, an example in which the gains are discrete random variables is considered. Specifically, the company is assumed to receive two types of incomes: frequent small donations and sparse large gifts.

Example 6. *Suppose that the gains have the probability mass function given by:*

$$P(X = x) = \begin{cases} p, & \text{if } x = \gamma, \\ 1 - p, & \text{if } x = \beta, \end{cases}$$

where $\gamma \ll \beta$ *and* $p \gg 0$. *For the following two sub-examples, fix* $d = 2, b = 8$ *and* $N = 50$. *Figure 4 shows two different scenarios. As the size of the large donations* β *increases, the ruin probability decreases. As the frequency of small donations* γ *increases, the ruin probability increases.*

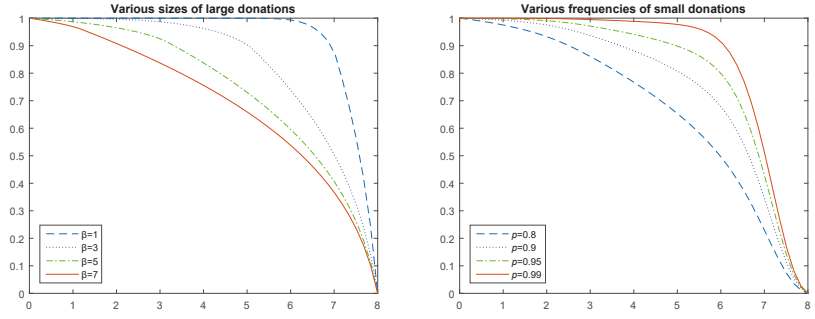


Figure 4. (**a**) (Left) Ruin probabilities for various sizes of large donations, for fixed $\gamma = 0.1$ and $p = 0.85$. (**b**) (Right) Ruin probabilities for various frequencies of small donations, for fixed $\gamma = 0.5$ and $\beta = 5$.

5. The Laplace Transform of the Time of Ruin

The same procedure can be applied to other functionals of the time of ruin. Let $\delta > 0$, and denote the Laplace transform of the time of ruin by:

$$\Phi(u) = \mathbb{E}[e^{-\delta\tau}\mathbb{1}_{[\tau<\infty]} \mid U_0 = u].$$

One can also interpret $\Phi(u)$ as the probability of ruin before an independent exponential clock with mean $1/\delta$. First, it is shown that Φ satisfies an IDE. The proof is different than the proof in Theorem 1. From the IDE of Φ, one recovers the IDE for the ruin probability simply by setting $\delta = 0$.

Theorem 2. *The Laplace transform of the time of ruin satisfies the following IDE:*

$$0 = (au - c)\Phi'(u) - (\lambda + \delta)\Phi(u) + \lambda \int_0^{\frac{c}{a}-u} \Phi(u + y)\, f(y)\, dy. \tag{17}$$

Proof. Let $b = c/a$. Recall that the investment process is given by $Z_t = (u - b)e^{at} + b$ and that the solution to the equation $Z_t = 0$ is given by $t^* = \frac{1}{a}\ln\left(\frac{b}{b-u}\right)$. Therefore, if the time of the first gain is bigger than t^*, ruin is certain. Thus, conditioning on the time and amount of the first gain, we get:

$$\begin{aligned}\Phi(u) &= \int_0^{t^*} \lambda e^{-(\lambda+\delta)t} \int_0^{b-Z_t} \Phi(Z_t + y) f(y)\, dy\, dt + e^{-(\lambda+\delta)t^*} \\ &= \frac{\lambda}{a}\int_0^u \left(\frac{b-u}{b-v}\right)^{\frac{\lambda+\delta}{a}} \int_0^{b-v} \Phi(v + y) f(y)\, dy\, \frac{1}{b-v}\, dv + \left(\frac{b-u}{b}\right)^{\frac{\lambda+\delta}{a}},\end{aligned}$$

where we have used the substitution $t = \frac{1}{a}\ln\left(\frac{b-v}{b-u}\right)$. Applying the operator $\left((b-u)\frac{d}{du} + \frac{\lambda+\delta}{a}\right)$ to the above yields:

$$(b-u)\Phi'(u) + \frac{\lambda+\delta}{a}\Phi(u) = \frac{\lambda}{a}\int_0^{b-u} \Phi(u + y)\, f(y)\, dy.$$

□

Let:

$$\Xi(u) \equiv \begin{cases} \Phi'(u), & 0 \le u \le \frac{c}{a}, \\ 0, & u > \frac{c}{a}. \end{cases}$$

From (17) and following the steps in Section 3, one has:

$$\begin{aligned}(au-c)\Xi(u) &= \delta\Phi(u) + \lambda\int_0^\infty \int_{u+y}^u \Xi(t)\,dt\,f(y)\,dy \\ &= -\delta\int_u^\infty \Xi(t)\,dt - \lambda\int_u^{c/a} \overline{F}(t-u)\,\Xi(t)\,dt,\end{aligned}$$

for $0 < u < \frac{c}{a}$, where the last equality follows from the fact that $\lim_{u\to\infty}\Phi(u) = 0$. Recall that $b = \frac{c}{a}$, and so:

$$\Xi(u) = \frac{\delta}{c-au}\int_u^b \Xi(t)\,dt + \frac{\lambda}{c-au}\int_u^b \overline{F}(t-u)\Xi(t)\,dt, \qquad 0 < u < b.$$

With $\Xi(u) = (b-u)^{d-1}\tilde{\Xi}(u)$, one can perform integration by parts (see the steps leading to (10)) to obtain:

$$0 = \int_u^b (b-t)^d \frac{d}{dt}\left[\overline{F}(t-u)\,\tilde{\Xi}(t)\right]\,dt + \int_u^b \frac{\delta}{a}(b-t)^{d-1}\tilde{\Xi}(t)\,dt, \qquad 0 < u < b, \tag{18}$$

with the integrability condition:

$$\int_0^b (b-u)^{d-1}\tilde{\Xi}(u)\,du = -1.$$

Similar discretization yields the matrix equation $B\tilde{\Xi} = Y$ where the matrix:

$$B = A + \frac{\delta}{a}\overline{A}. \tag{19}$$

Here, $A = (a_{jk})$ is the matrix given by (14),

$$\overline{A} = \begin{bmatrix} (b-u_1)^{d-1} & (b-u_2)^{d-1} & \dots & (b-u_{N-1})^{d-1} & 0 \\ 0 & (b-u_2)^{d-1} & \dots & (b-u_{N-1})^{d-1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & (b-u_{N-1})^{d-1} & 0 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix},$$

$\tilde{\Xi} = [\tilde{\Xi}_1, \tilde{\Xi}_2, \dots, \tilde{\Xi}_N]^T$ and $Y = [0, \cdots, 0, 1]^T$. It follows from Lemma 1 that a sufficient condition for the invertibility of B is that $-(b-u_i) + \delta/a < 0$ for all $j = 1, \dots, N-1$, which is equivalent to the condition $\delta < a(b-u_{N-1})$. Although we do not know of more general conditions for the invertibility of B, it does not give rise to any problems in the numerical examples, as seen below.

Example 7. *Consider exponentially-distributed gains, where* $\overline{F}(x) = e^{-x}$. *Let* $\frac{c}{a} = b = 4$, $\frac{\lambda}{a} = d = 3.5$ *and* $N = 100$. *Figure 5 shows that the numerical approximation of* Φ *approaches the ruin probability as* $\delta \to 0$. *As expected, the Laplace transform of the ruin time is a decreasing function of* δ.

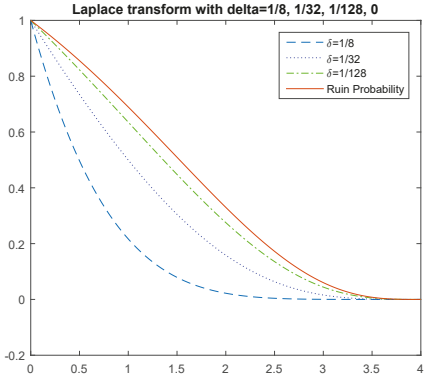


Figure 5. Laplace transform of the time of ruin for exponential mean one gains when $\delta = \frac{1}{8}, \frac{1}{32}, \frac{1}{128}, 0$.

6. Conclusions and Future Research

In summary, this paper provides a framework to study numerical approximations of the ruin probability in the dual risk model with a constant interest rate, as well as the Laplace transform of the time of ruin when the gain distribution is arbitrary. Using elementary analysis, the IDE satisfied by the ruin probability can be rewritten as an IE, which in turn can be approximated by a simple linear system.

One possible extension of this work is the numerical approximation for the Gerber–Shiu function. In the classical insurance model, the seminalwork by Gerber and Shiu (1998) builds around the study of the joint distribution of the time of ruin, the surplus immediately before ruin and the deficit at ruin. In the classical dual model, two of the three random variables here are identical. Since ruin is caused by continuous expenses/consumption, the surplus immediately before ruin and the deficit at ruin are both equal to zero. Nevertheless, for the dual risk model, one can study analogs of the Gerber–Shiu function by considering random variables such as the time of the last gain before ruin and its amount (see Yang and Sendova (2014)) or the time of the first gain after ruin and its amount (see Cheung (2012)).

Author Contributions: Both authors contributed equally to this work.

Acknowledgments: Support for Sooie-Hoe Loke from the School of Graduate Studies and Research at Central Washington University is gratefully acknowledged.

Conflicts of Interest: The authors declare no conflict of interest.

References

Avanzi, Benjamin, Hans U. Gerber, and Elias S.W. Shiu. 2007. Optimal dividends in the dual model. *Insurance: Mathematics and Economics* 41: 111–23. [CrossRef]

Cai, Jun, Runhuan Feng, and Gordon E. Willmot. 2009. The Compound Poisson Surplus Model with Interest and Liquid Reserves: Analysis of the Gerber–Shiu Discounted Penalty Function. *Methodol Comput Appl Probab* 11: 401–23. [CrossRef]

Chen, Li. 2010. Risk management for nonprofit organizations. Master's thesis, Oregon State University, Corvallis, OR, USA.

Cheung, Eric C. K. 2012. A unifying approach to the analysis of business with random gains. *Scandinavian Actuarial Journal* 2012: 153–82. [CrossRef]

Dong, Yinghui, and Guojing Wang. 2008. On a compounding assets model with positive jumps. *Applied Stochastic Models in Business and Industry* 24: 21–30. [CrossRef]

Gerber, Hans U., and Elias Shiu. 1998. On the time value of ruin. *North American Actuarial Journal* 2: 48–72. [CrossRef]

Kasozi, Juma, and Jostein Paulsen. 2005. Numerical Ultimate Ruin Probabilities under Interest Force. *Journal of Mathematics and Statistics* 1: 246–51. [CrossRef]

Konstantinides, Dimitrios G., Kai W. Ng, and Qihe Tang. 2010. The probabilities of absolute ruin in the renewal risk model with constant force of interest. *Journal of Applied Probability* 47: 323–34. [CrossRef]

Mishura, Yuliya, and Olena Ragulina. 2016. *Ruin Probabilities: Smoothness, Bounds, Supermartingale Approach*. New York: Elsevier.

Polyanin, Andrei D., and Valentin F. Zaitsev. 2002. *The Handbook of Exact Solutions for Ordinary Differential Equations*, 2nd ed. London: Chapman and Hall.

Schmidli, Hanspeter. 2015. Extended Gerber–Shiu functions in a risk model with interest. *Insurance: Mathematics and Economics* 61: 271–75. [CrossRef]

Segerdahl, Carl-Otto. 1942. Uber einige risikotheoretische Fragestellungen. *Scandinavian Actuarial Journal* 61: 43–83. [CrossRef]

Slater, Lucy Joan. 1960. *Confluent Hypergeometric Functions*. Cambridge: Cambridge University Press.

Sundt, Bjørn, and Jozef L. Teugels. 1995. Ruin estimates under interest force. *Insurance: Mathematics and Economics* 16: 7–22. [CrossRef]

Yang, Chen, and Kristina P. Sendova. 2014. The discounted moments of the surplus after the last innovation before ruin under the dual risk model. *Stochastic Models* 30: 99–124. [CrossRef]

Yang, Yang, and Yuebao Wang. 2010. Asymptotics for ruin probability of some negatively dependent risk models with a constant interest rate and dominatedly-varying-tailed claims. *Statistics and Probability Letters* 80: 143–54. [CrossRef]

Zeng, Fanzi, and Jisheng Xu. 2013. The Perturbed Dual Risk Model with Constant Interest and a Threshold Dividend Strategy. *Abstract and Applied Analysis* 2013: 981076. [CrossRef]

risks

MDPI

Article

Moments of Compound Renewal Sums with Dependent Risks Using Mixing Exponential Models

Fouad Marri [1,*], Franck Adékambi [2] and Khouzeima Moutanabbir [3]

[1] Department of Statistics and Actuarial Science, Institut National de Statistique et d'Economie Appliquée, INSEA, Rabat 10112, Morocco
[2] School of Economics, University of Johannesburg, Johannesburg 2006, South Africa; fadekambi@uj.ac.za
[3] Department of Mathematics and Actuarial Science, The American University in Cairo, New Cairo 11835, Egypt; k.moutanabbir@aucegypt.edu
* Correspondence: fmarri@insea.ac.ma

Received: 20 July 2018; Accepted: 22 August 2018; Published: 24 August 2018

Abstract: In this paper, we study the discounted renewal aggregate claims with a full dependence structure. Based on a mixing exponential model, the dependence among the inter-claim times, the claim sizes, as well as the dependence between the inter-claim times and the claim sizes are included. The main contribution of this paper is the derivation of the closed-form expressions for the higher moments of the discounted aggregate renewal claims. Then, explicit expressions of these moments are provided for specific copulas families and some numerical illustrations are given to analyze the impact of dependency on the moments of the discounted aggregate amount of claims.

Keywords: renewal process; discounted aggregate claims; copulas; archimedean copulas

1. Introduction

Over the past few years, extensive studies on the risk aggregation problem for insurance portfolios have appeared in the literature. Among these studies we find Albrecher and Boxma (2004), Albrecher and Teugels (2006) and Boudreault et al. (2006) which analyze ruin-related problems; Léveillé et al. (2010), Léveillé and Adékambi (2011, 2012), investigate the risk aggregation and the distribution of the discounted aggregate amount of claims; Léveillé and Garrido (2001a, 2001b) use the renewal theory to derive a closed expressions for the first two moments of the discounted aggregated claims; and Léveillé and Hamel (2013) study the aggregate discount payment and expenses process for medical malpractice insurance. Most recently, Jang et al. (2018) study the family of renewal shot-noise processes. Based on the piecewise deterministic Markov process theory and the martingale methodology, they obtained the Feynmann-Kac formula and then derived the Laplace transforms of the conditional moments and asymptotic moments of the processes.

For the risk management of non-life insurance portfolios, the mathematical expectation of the discounted aggregate claims plays an important role in determining the pure premium, in addition to giving a measure of the central tendency of its distribution. Moments centered at the 2nd, 3rd and 4th order average are the other moments usually considered, as they generally give a good indication of the pace of the distribution. The 2nd order centered moment gives us a measure of the dispersion around its mean, the 3rd order moment gives us a measure of the asymmetry of the distribution of and the 4th order moment gives us a measure of the flattening of the distribution of the discounted aggregate sums. Moments, whether simple, joint, or conditional, may be useful for constructing predictors, regression curves, or approximations of the distribution of the discounted aggregate claims.

The papers cited above assume that the inter-arrival times and the claim amounts are independent. Such an assumption is not supported by empirical observations which reduces the practicality of these works. For example, in non-life insurance, the same catastrophic event such as a flood or an earthquake

could lead to frequent and high losses. This means that in such context a positive dependence between the claim sizes and the inter-claim times should be observed.

During the last decade, few papers in the actuarial literature considered incorporating this type of dependence. For example, Barges et al. (2011) introduce the dependence between the claim sizes and the inter-claim times using a Farlie-Gumbel-Morgenstern (FGM) copula and derive a close-from expression for the moments of the discounted aggregate claims. Guo et al. (2013) incorporate time dependence in a mixed Poisson process to study loss models. Landriault et al. (2014) consider a non-homogeneous birth process for the claim counting process to study time dependent aggregate claims.

For a given portfolio, we consider the renewal risk process suggested by Andersen (1957) and described as follows. Let $\{N(t)\}_{t\geq 0}$ be a renewal process that counts the number of claims. The positive random variable (rv) W_k represents the time between the $(k-1)-$th and $k-$th claims, $k \in \mathbb{N}^\star = \{1, 2, \cdots\}$, and the amount of the k-th claim is given by the positive rv X_k. We also define $\{T_k,\, k \in \mathbb{N}^\star\}$ as a sequence of rvs such that $T_k = \sum_{i=1}^{k} W_i,\ T_0 = 0$. The rv T_k represents the occurrence time of the $k-$th received claim. For any given integer n and $t \geq 0$, we have $\{N(t) \geq n\} = \{T_n \leq t\}$. The main variable of interest in this paper is the discounted aggregate amount of claims up to a certain time $\mathcal{Z}(t)$ defined as follows

$$\mathcal{Z}(t) = \sum_{i=1}^{N(t)} e^{-\delta T_i} X_i, \quad t \geq 0, \tag{1}$$

with $\mathcal{Z}(t) = 0$ if $N(t) = 0$, where δ is the force of net interest (See e.g., Léveillé and Garrido 2001a). In the rest of the paper, it is assumed that

- $\{W_k, k \in \mathbb{N}^\star = \{1, 2, \cdots\}\}$ forms a sequence of continuous positive dependent and identically distributed rvs with a common cumulative distribution function (cdf) $F_W(.)$ and a survival function (sf) $\bar{F}_W(.) = 1 - F_W(.)$,
- The claim amounts $\{X_k, k \in \mathbb{N}^\star\}$ are positive dependent and identically distributed rvs with a common cdf $F_X(.)$ and a common sf $\bar{F}_X(.) = 1 - F_X(.)$, and
- $\{(W_k, X_k), k \in \mathbb{N}^\star\}$ forms a sequence of identically distributed random vectors distributed as the canonical random vector (W, X) in which the components may be dependent.

In this paper, we specify three sources of dependence: among the claims X_k, among the subsequent inter-claims time W_k, and a dependence between the subsequent inter-claims time W_k and the claims X_k. For the dependence between the inter-claim times $\{W_k, k \in \mathbb{N}^\star = \{1, 2, \cdots\}\}$, we assume the existence of a positive continuous rv Θ such that given $\Theta = \theta$ the rvs W_k are iid and exponentially distributed with a mean $\frac{1}{\theta}$. Similarly, we introduce the dependence between the amounts of claims $\{X_k, k \in \mathbb{N}^\star\}$ through a positive continuous rv Λ such that conditional on $\Lambda = \lambda$ the rvs X_k are iid and exponentially distributed with a mean $\frac{1}{\lambda}$. In other words, the conditional distributions of the components of W and X are only influenced by the rv Θ and Λ respectively. The rvs Θ and Λ represent the factors that introduce the dependence between risks (e.g., climate conditions, age, $\cdots$, etc.).

In what follows, let $F_{\Theta,\Lambda}$ be the joint cdf of the positive random vector (Θ, Λ) and the marginal cdfs are F_Θ and F_Λ. We also define the joint Laplace transform $f^\star_{\Theta,\Lambda}(s_1, s_2) = \int_0^\infty \int_0^\infty e^{-(\theta s_1 + \lambda s_2)} dF_{\Theta,\Lambda}(\theta, \lambda)$, for $s_1, s_2 \geq 0$, as well as the univariate Laplace transforms $f^\star_\Theta(s) = \int_0^\infty e^{-\theta s} dF_\Theta(\theta)$ and $f^\star_\Lambda(s) = \int_0^\infty e^{-\lambda s} dF_\Lambda(\lambda)$, for $s \geq 0$. Following the model's specifications, the univariate distributions of W_i and X_i are given as a mixture of exponential distributions with survival functions given by

$$\bar{F}_W(x) = \int_0^\infty e^{-\theta x} dF_\Theta(\theta) = f^\star_\Theta(x), \tag{2}$$

and

$$\bar{F}_X(x) = \int_0^\infty e^{-\lambda x} dF_\Lambda(\lambda) = f^\star_\Lambda(x), \tag{3}$$

for $x \geq 0$. This implies that the marginal distributions of W_i and X_i are completely monotone. We refer to Albrecher et al. (2011) for more details on the mixed exponential model and the completely monotone marginal distributions. The general mixed risk model that we consider in this paper is an extension of the risk model described in Albrecher et al. (2011).

This paper is structured as follows: In Section 2, we describe the dependence structure of our risk model. Moments of the aggregate discounted claims are derived in Section 3. Section 4 provides few examples of risk models for which explicit expressions for the moment are given. Numerical examples are provided to illustrate the impact of dependency on the moments of discounted aggregate claims. Section 5 concludes the paper.

2. The Dependence Structure

In this section, a description of the dependence between the different components of our model is provided. For a given n and under our conditional exponential model, the joint conditional survival function of $W_1, W_2, \cdots, W_n, X_1, X_2 \cdots, X_n$ is given by

$$\Pr\left(W_1 \geq t_1, \cdots, W_n \geq t_n, X_1 \geq s_1, \cdots, X_n \geq s_n \mid \Theta = \theta, \Lambda = \lambda\right) = e^{-\theta \sum_{i=1}^{n} t_i} e^{-\lambda \sum_{i=1}^{n} s_i},$$

for $n \in \{2, 3, \cdots\}$, $t_1, \cdots, t_n \geq 0$ and $s_1, \cdots, s_n \geq 0$. it is immediate that the multivariate survival function of $W_1, W_2, \cdots, W_n, X_1, X_2 \cdots, X_n$ could be expressed in terms of the bivariate Laplace transform $f^{\star}_{\Theta,\Lambda}$ such that

$$\begin{aligned} \bar{F}_{W_1,\cdots,W_n,X_1,\cdots,X_n}(t_1, \cdots, t_n, s_1, \cdots, s_n) &= \int_0^{\infty}\int_0^{\infty} e^{-\theta \sum_{i=1}^{n} t_i} e^{-\lambda \sum_{i=1}^{n} s_i} dF_{\Theta,\Lambda}(\theta, \lambda) \\ &= f^{\star}_{\Theta,\Lambda}\left(\sum_{i=1}^{n} t_i, \sum_{i=1}^{n} s_i\right). \end{aligned} \tag{4}$$

On the other hand, according to Sklar's theorem for survival functions, see e.g., Sklar (1959), the joint distribution of the tail of $W_1, \cdots, W_n, X_1, \cdots, X_n$ can be written as a function of the marginal survival functions $\bar{F}_{W_i}$, $\bar{F}_{X_i}$, $i = 1, \cdots, n$, and the copula C describing the dependence structure as follows

$$\bar{F}_{W_1,\cdots,W_n,X_1,\cdots,X_n}(t_1, \cdots, t_n, s_1, \cdots, s_n) = C\left(\bar{F}_{W_1}(t_1), \cdots, \bar{F}_{W_n}(t_n), \bar{F}_{X_1}(s_1), \cdots, \bar{F}_{X_n}(s_n)\right),$$

for $n \in \{2, 3, \cdots\}$, $t_1, \cdots, t_n \geq 0$ and $s_1, \cdots, s_n \geq 0$. By combining (2), (3) and (4) with the last expression, one deduces that for $(u_1, \cdots, u_n, v_1, \cdots, v_n) \in [0, 1]^{2n}$

$$C(u_1, \cdots, u_n, v_1, \cdots, v_n) = f^{\star}_{\Theta,\Lambda}\left(\sum_{i=1}^{n} f^{\star -1}_{\Theta}(u_i), \sum_{i=1}^{n} f^{\star -1}_{\Lambda}(v_i)\right). \tag{5}$$

According to (4), the bivariate survival function of (W_i, X_i), for $i = 1, \cdots, n$, is given by

$$\bar{F}_{W_i,X_i}(t, s) = f^{\star}_{\Theta,\Lambda}(t, s), \tag{6}$$

for $t \geq 0$ and $s \geq 0$. Hence, using Sklar's theorem, the dependency relation between W_i and X_i is generated by a copula C_{12} given by

$$C_{12}(u, v) = f^{\star}_{\Theta,\Lambda}\left(f^{\star -1}_{\Theta}(u), f^{\star -1}_{\Lambda}(v)\right), \tag{7}$$

for $(u,v) \in [0,1]^2$. Otherwise, it is clear from (4) that the multivariate survival function of $(W_1, \cdots, W_n)$ is given by

$$\bar{F}_{W_1,\cdots,W_n}(t_1,\cdots,t_n) = f^{\star}_{\Theta}\left(\sum_{i=1}^{n} t_i\right), \tag{8}$$

for $t_1, \cdots, t_n \geq 0$. Consequently, an application of Sklar's theorem shows that the joint distribution of the tail of $W_1, \cdots, W_n$ can be written as a function of the marginal survival functions $\bar{F}_{W_i}$, $i = 1, \cdots, n$, and a copula C_1 describing the dependence structure as follows

$$\bar{F}_{W_1,\cdots,W_n}(t_1,\cdots,t_n) = C_1\left(\bar{F}_{W_1}(t_1),\cdots,\bar{F}_{W_n}(t_n)\right).$$

An expression for C_1 is identified and for $(u_1, \cdots, u_n) \in [0,1]^n$, we obtain

$$C_1(u_1,\cdots,u_n) = f^{\star}_{\Theta}\left(\sum_{i=1}^{n} f^{\star\,-1}_{\Theta}(u_i)\right). \tag{9}$$

Similarly, the joint distribution of the tail of $X_1, \cdots, X_n$ is given by

$$\bar{F}_{X_1,\cdots,X_n}(t_1,\cdots,t_n) = f^{\star}_{\Lambda}\left(\sum_{i=1}^{n} t_i\right), \tag{10}$$

for $t_1, \cdots, t_n \geq 0$, and using Sklar's theorem yields the following survival copula for the Xs

$$C_2(u_1,\cdots,u_n) = f^{\star}_{\Lambda}\left(\sum_{i=1}^{n} f^{\star\,-1}_{\Lambda}(u_i)\right), \tag{11}$$

for $(u_1, \cdots, u_n) \in [0,1]^n$. From the expressions for the copulas C_1 and C_2 obtained above, one can identify that these two copulas belong to the large class of Archimedean copulas (e.g., Nelsen 1999) with the corresponding generators $f^{\star\,-1}_{\Theta}$ and $f^{\star\,-1}_{\Lambda}$. Note that although the dependence among the claim sizes and among the inter-claim times are described by Archimedean copulas. The dependence between W and X is not restricted to this family of copulas. Moreover, the mixture of exponentials model introduces a positive dependence between the inter-claim times Ws as well as a positive dependence between the amount Xs. First, we recall the following definition

Definition 1. *Let X and Y be random variables. X and Y are positively quadrant dependent (PQD) if for all (x,y) in $\mathbb{R}^2$,*

$$Pr\left[X \leq x, Y \leq y\right] \geq Pr\left[X \leq x\right] Pr\left[Y \leq y\right],$$

or equivalently

$$Pr\left[X > x, Y > y\right] \geq Pr\left[X > x\right] Pr\left[Y > y\right].$$

Proposition 2.1. *Consider the model described by (8) and (10). Then, W_i and W_j (X_i and X_j) are PQD for all $i, j = 1, 2, \cdots$.*

Proof. We refer the reader to Chapter 4 in Joe (1997) for the proof of this proposition. □

Combining (5), (7), (9)and (11), one gets

$$C(u_1,\cdots,u_n,v_1,\cdots,v_n) = C_{12}\Big(C_1(u_1,\cdots,u_n), C_2(v_1,\cdots,v_n)\Big),$$

for $(u_1, \cdots, u_n, v_1, \cdots, v_n) \in [0,1]^{2n}$. Throughout the paper, we suppose that the Laplace transform $f^{\star}_{\Theta,\Lambda}$ exists over a subset $K \times K \subset \mathbb{R}^2$ including a neighborhood of the origin. In the following section, the moments of the rv $\mathcal{Z}(t)$ are derived.

3. Moments of the Discounted Aggregate Claims

In order to find the moments of the discounted aggregate claims, we first derive an expression for the moments generating function (mgf) of the rv $\mathcal{Z}(t)$ under the dependent model introduced in the previous section.

Theorem 3.1. *Consider the discounted aggregate claims under the assumptions of the model in Section 2. Then, for any $t \geq 0$ and $\delta > 0$, the mgf of $\mathcal{Z}(t)$ is given by*

$$M_{\mathcal{Z}(t)}(s) \quad = \quad E\left[\frac{\Lambda - se^{-\delta t}}{\Lambda - s}\right]^{\frac{\Theta}{\delta}}. \tag{12}$$

Proof. Given $\Theta = \theta$ and $\Lambda = \lambda$, the aggregate discounted processes, $\mathcal{Z}(t)$ is a compound Poisson processes with independent subsequent inter-claim times. According to Léveillé et al. (2010), the mgf of $\mathcal{Z}(t)$ given $\Theta = \theta$ and $\Lambda = \lambda$ can be written as

$$\begin{aligned} M_{\mathcal{Z}(t)|\Theta=\theta,\Lambda=\lambda}(s) &= E\left[e^{s\mathcal{Z}(t)} \mid \Theta = \theta, \Lambda = \lambda\right] \\ &= e^{s\theta\int_0^t\left[\frac{e^{-\delta v}}{\lambda - se^{-\delta v}}\right]dv} = \left(\frac{\lambda - se^{-\delta t}}{\lambda - s}\right)^{\frac{\theta}{\delta}}. \end{aligned} \tag{13}$$

Otherwise $M_{\mathcal{Z}(t)}(s) = \int_0^\infty \int_0^\infty M_{\mathcal{Z}(t)|\Theta=\theta,\Lambda=\lambda}(s) dF_{\Theta,\Lambda}(\theta,\lambda)$. Substituting (13) into the last expression yields (12). □

The following theorem provides closed formulas for the higher moments of the discounted aggregate claims $\mathcal{Z}(t)$.

Theorem 3.2. *Consider the discounted aggregate claims under the assumptions of the model in Section 2. Then, for any $t \geq 0$, $n \in \mathbb{N}^{\star}$ and $\delta > 0$, the $n-$th moment of $\mathcal{Z}(t)$ is given by*

$$E\left[\mathcal{Z}^n(t)\right] \quad = \quad \sum \frac{n!}{k_1!k_2!\cdots k_n!}\bar{a}^k_{\overline{t}|\delta} E\left[\frac{\Theta(\Theta-\delta)\cdots(\Theta-\delta(k-1))}{\Lambda^n}\right], \tag{14}$$

where $\bar{a}_{\overline{t}|\delta} = \frac{1-e^{-t\delta}}{\delta}$ is the standard actuarial notation and the sum is over all nonnegative integer solutions of the Diophantine equation $k_1 + 2k_2 + \cdots + nk_n = n$, $k := k_1 + k_2 + \cdots + k_n$.

Proof. Conditional on the two rvs Θ and Λ, we have

$$E\left[\mathcal{Z}^n(t)\right] \quad = \quad \int_0^\infty \int_0^\infty E\left[\mathcal{Z}^n(t) \mid \Theta = \theta, \Lambda = \lambda\right] dF_{\Theta,\Lambda}(\theta,\lambda). \tag{15}$$

Taking the $n-$th order derivative of (13) with respect to s and using Faà di Bruno's rule (see Faa di Bruno 1855) yield

$$M^{(n)}_{\mathcal{Z}(t)|\Theta=\theta,\Lambda=\lambda}(s) = \sum \frac{n!}{k_1!k_2!\cdots k_n!} h^{(k)}\left(g(s)\right) \prod_{j=1}^{n}\left(\frac{g^{(j)}(s)}{j!}\right)^{k_j}, \tag{16}$$

where the sum is over all nonnegative integer solutions of the Diophantine equation $k_1 + 2k_2 + \cdots + nk_n = n$, $k := k_1 + k_2 + \cdots + k_n$, $g(s) = \frac{\lambda - se^{-\delta t}}{\lambda - s}$ and $h(s) = s^{\frac{\theta}{\delta}}$. Otherwise, the $k-$th derivatives of g and h are given respectively by

$$g^{(k)}(s) = \lambda(1 - e^{-\delta t})\frac{k!}{(\lambda - s)^{k+1}}, \tag{17}$$

and

$$h^{(k)}(s) = \frac{\Gamma(\frac{\theta}{\delta} + 1)}{\Gamma(\frac{\theta}{\delta} - k + 1)} s^{\frac{\theta}{\delta} - k}, \tag{18}$$

for $k = 1, \cdots, n$. By substituting (17) and (18) into (16) with $s = 0$, one concludes that

$$\begin{aligned} E\left[\mathcal{Z}^n(t) \mid \Theta = \theta, \Lambda = \lambda\right] &= \frac{1}{\lambda^n} \sum \frac{n!}{k_1! k_2! \cdots k_n!} \left(1 - e^{-\delta t}\right)^k \frac{\Gamma(\frac{\theta}{\delta} + 1)}{\Gamma(\frac{\theta}{\delta} - k + 1)} \\ &= \sum \frac{n!}{k_1! k_2! \cdots k_n!} \left(1 - e^{-\delta t}\right)^k \frac{\frac{\theta}{\delta}\left(\frac{\theta}{\delta} - 1\right) \cdots \left(\frac{\theta}{\delta} - (k-1)\right)}{\lambda^n} \\ &= \sum \frac{n!}{k_1! k_2! \cdots k_n!} \bar{a}_{\overline{t}|\delta}^k \frac{\theta(\theta - \delta) \cdots (\theta - \delta(k-1))}{\lambda^n}. \end{aligned} \tag{19}$$

Finally, substitution of (20) into (15) yields the required result. □

The moments of $\mathcal{Z}(t)$ given in (14) could be simplified and expressed in terms of the expected value of $E\left[\frac{\Theta^l}{\Lambda^n}\right]$. First, we write

$$\frac{\theta}{\delta}\left(\frac{\theta}{\delta} - 1\right) \cdots \left(\frac{\theta}{\delta} - (k-1)\right) = \left(\frac{\theta}{\delta}\right)_k,$$

where $(x)_k$ is the falling factorial. It is known that the falling factorial could be expanded as follows

$$(x)_k = \sum_{l=1}^{k} \begin{bmatrix} k \\ l \end{bmatrix} x^l, \tag{20}$$

where the coefficients $\begin{bmatrix} k \\ l \end{bmatrix}$ are the Stirling numbers of the first order (see e.g., Ginsburg 1928). Using (20), we find

$$\frac{\theta}{\delta}\left(\frac{\theta}{\delta} - 1\right) \cdots \left(\frac{\theta}{\delta} - (k-1)\right) = \sum_{l=1}^{k} \begin{bmatrix} k \\ l \end{bmatrix} \left(\frac{\theta}{\delta}\right)^l.$$

Thus,

$$E\left[\mathcal{Z}^n(t)\right] = \sum \frac{n!}{k_1! k_2! \cdots k_n!} \bar{a}_{\overline{t}|\delta}^k \sum_{l=1}^{k} \delta^{k-l} \begin{bmatrix} k \\ l \end{bmatrix} E\left[\frac{\Theta^l}{\Lambda^n}\right]. \tag{21}$$

In the rest of the paper, it is assumed that there exist an integer n such that the expected value of $\frac{\Theta^i}{\Lambda^j}$ is finite for positive integers i and j with $i, j \leq n$. Using the previous theorem, we give the explicit expressions of the first two moments of $\mathcal{Z}(t)$.

Corollary 3.1. *For a given time t and a positive constant forces of interest δ, we have*

$$E\left[\mathcal{Z}(t)\right] = \bar{a}_{\overline{t}|\delta} E\left[\frac{\Theta}{\Lambda}\right], \tag{22}$$

and

$$E\left[\mathcal{Z}^2(t)\right] = 2\bar{a}_{\bar{l}\bar{2}\delta}E\left[\frac{\Theta}{\Lambda^2}\right] + \bar{a}^2_{\bar{l}\bar{\delta}}E\left[\frac{\Theta^2}{\Lambda^2}\right]. \tag{23}$$

Proof. The results follow from Theorem (3.2). When $n = 1$, then $k_1 = k = 1$, which yields (22). When $n = 2$, we find that the nonnegative integer solutions of the equation $k_1 + 2k_2 = 2$ are $(k_1, k_2) = (2, 0)$ or $(0, 1)$ with corresponding values of k being 2 or 1 respectively, we get the required result. □

In the following corollary, we derive expressions for the first two moments of $\mathcal{Z}(t)$ when Θ and Λ are independent.

Corollary 3.2. *If the dependency relation between Θ and Λ is generated by the independence copula then*

$$E\left[\mathcal{Z}(t)\right] = \bar{a}_{\bar{l}\bar{\delta}}E\left[\Theta\right]E\left[\frac{1}{\Lambda}\right],$$

and

$$E\left[\mathcal{Z}^2(t)\right] = 2\bar{a}_{\bar{l}\bar{2}\delta}E\left[\Theta\right]E\left[\frac{1}{\Lambda^2}\right] + \bar{a}^2_{\bar{l}\bar{\delta}}E\left[\Theta^2\right]E\left[\frac{1}{\Lambda^2}\right].$$

Proof. The result follows easily from Corollary (3.1). □

Note that the moments of $\mathcal{Z}(t)$ are given in terms of the expected values of $\frac{\Theta^l}{\Lambda^n}$, for $l, n \in \mathbb{N}^\star \times \mathbb{N}^\star$. According to Cressie et al. (1981), the expression of $E\left[\frac{\Theta^l}{\Lambda^n}\right]$ can be derived from the $M_{\Theta,\Lambda}(t, s)$, the joint mgf of (Θ, Λ). We have

$$E\left[\frac{\Theta^l}{\Lambda^n}\right] = \frac{1}{\Gamma(n)}\int_0^\infty x^{n-1}\lim_{s\to 0}\frac{\partial^l M_{\Theta,\Lambda}(s,-x)}{\partial s^l}dx,$$

where the joint mgf $M_{\Theta,\Lambda}$ is given by

$$M_{\Theta,\Lambda}(s,x) = f^*_{\Theta,\Lambda}(-s,-x) = C_{12}\left(f^*_\Theta(-s), f^*_\Lambda(-x)\right).$$

It follows that

$$E\left[\frac{\Theta^l}{\Lambda^n}\right] = \frac{1}{\Gamma(n)}\int_0^\infty x^{n-1}\lim_{s\to 0}\frac{\partial^l f^*_{\Theta,\Lambda}(-s,x)}{\partial s^l}dx. \tag{24}$$

Application of Faà di Bruno's rule for the $l-$th derivative of $f^*_{\Theta,\Lambda}(-t, s)$ gives

$$\frac{\partial^l M_{\Theta,\Lambda}(s,-x)}{\partial s^l} = \sum\frac{l!}{m_1!m_2!\cdots m_l!}\frac{\partial^m C_{12}\left(f^*_\Theta(-s), f^*_\Lambda(x)\right)}{\partial u^m}\prod_{j=1}^{l}\left(\frac{\partial^j f^*_\Theta(-s)}{\partial s^j}\frac{1}{j!}\right)^{m_j},$$

where the sum is over all nonnegative integer solutions of the Diophantine equation $m_1 + 2m_2 + \cdots + lm_l = l$, $\quad m := m_1 + m_2 + \cdots + m_l$. It follows that

$$E\left[\frac{\Theta^l}{\Lambda^n}\right] = \frac{1}{\Gamma(n)}\sum\frac{l!}{m_1!m_2!\cdots m_l!}\prod_{j=1}^{l}\left(\frac{E\left[\Theta^j\right]}{j!}\right)^{m_j}\int_0^\infty x^{n-1}\frac{\partial^m C_{12}\left(1, f^*_\Lambda(x)\right)}{\partial u^m}dx. \tag{25}$$

4. Examples

In the previous section, a general formula for the moments of $\mathcal{Z}(t)$ is derived. In order to illustrate our findings and to discuss further features of our risk model, we provide some examples when additional assumptions on the marginal distributions and the copulas are added. For each example, first the joint Laplace distribution of the mixing distribution $F_{\Theta,\Lambda}$ is specified then the expressions of the copulas C_1, C_2 and C_{12} are identified. Applying our closed-form, the moments of $\mathcal{Z}(t)$ are given for these specific models. Some numerical illustrations are provided in order to stress the impact of dependence between different components of the risk models on the distribution of the discounted aggregated amount of claims.

4.1. Clayton Copula with Pareto Claims and Inter-Claim Times

Assume that the mixing random vector (Θ, Λ) has a bivariate Gamma distribution with a Laplace transform $f^{\star}_{\Theta,\Lambda}$ defined by

$$f^{\star}_{\Theta,\Lambda}(s,x) = \left[(1+as)^{\tilde{\alpha}_1} + (1+bx)^{\tilde{\alpha}_2} - 1\right]^{-\alpha}, \quad s \geq 0, \quad x \geq 0, \tag{26}$$

with $\alpha, a, b, \alpha_1, \alpha_2 > 0$ and $\tilde{\alpha}_i = \frac{\alpha_i}{\alpha}$, $i = 1, 2$. Then, the random variables Θ and Λ are distributed as gamma distributions, $\Theta \sim \mathcal{G}a(\alpha_1, \frac{1}{a})$ and $\Lambda \sim \mathcal{G}a(\alpha_2, \frac{1}{b})$. Also, from (2) and (3), the claim amounts X_i and the inter-claim times W_i, for $i = 1, 2, \cdots$, follow Pareto distributions $X \sim \mathcal{P}a(\alpha_2, \frac{1}{b})$ and $W \sim \mathcal{P}a(\alpha_1, \frac{1}{a})$. From (9) and (11), we identify the copulas C_1 and C_2 to be Clayton copulas with parameters $\frac{1}{\alpha_1}$ and $\frac{1}{\alpha_2}$, respectively. We have

$$C_1(u_1, \cdots, u_n) = \left[u_1^{\frac{-1}{\alpha_1}} + \cdots + u_n^{\frac{-1}{\alpha_1}} - (n-1)\right]^{-\alpha_1},$$

and

$$C_2(u_1, \cdots, u_n) = \left[u_1^{\frac{-1}{\alpha_2}} + \cdots + u_n^{\frac{-1}{\alpha_2}} - (n-1)\right]^{-\alpha_2},$$

for $(u_1, \cdots, u_n) \in [0,1]^n$. The Clayton copula is first introduced by Clayton (1978). The dependence between the Clayton copula parameter and Kendall's tau rank measure, τ_i, is given by (see e.g., Joe 1997 and Nelsen 1999):

$$\tau_i = \frac{1}{1+2\alpha_i}, \quad i = 1, 2. \tag{27}$$

This suggests that the Clayton copula does not allow for negative dependence. If $\alpha_i \to \infty$, $i = 1, 2$, then the marginal distributions become independent, when $\alpha_i = 0$, $i = 1, 2$, the Clayton copula approximates the Fréchet–Hoeffding upper bound.

From (7), the joint copula C_{12} is also a Clayton copula with a parameter $\frac{1}{\alpha}$ and we have

$$C_{12}(u,v) = \left[u^{\frac{-1}{\alpha}} + v^{\frac{-1}{\alpha}} - 1\right]^{-\alpha},$$

for $(u, v) \in [0,1]^2$. Let τ_{12} be the Kendall's tau dependence measure for the copula C_{12}. It follows that

$$\tau_{12} = \frac{1}{1+2\alpha}. \tag{28}$$

The following corollary gives the expressions of the first two moments of $\mathcal{Z}(t)$ for this model.

Corollary 4.1. *For a given horizon t and a positive constant forces of real interest δ, we have*

$$E\left[\mathcal{Z}(t)\right] = \frac{a\alpha_1}{b\Big(\tilde{\alpha}_2(\alpha+1)-1\Big)}\bar{a}_{\overline{t}|\delta},$$

for $\tilde{\alpha}_2 \geq \frac{1}{1+\alpha}$, *and*

$$E\left[\mathcal{Z}^2(t)\right] = \frac{2a\alpha_1}{b^2\Big(\tilde{\alpha}_2(\alpha+1)-1\Big)\Big(\tilde{\alpha}_2(\alpha+1)-2\Big)}\bar{a}_{\overline{t}|2\delta} + \frac{a^2}{b^2}\left[\frac{\alpha_1(1-\tilde{\alpha}_1)}{\Big(\tilde{\alpha}_2(\alpha+1)-1\Big)\Big(\tilde{\alpha}_2(\alpha+1)-2\Big)} + \frac{\alpha_1\tilde{\alpha}_1(1+\alpha)}{\Big(\tilde{\alpha}_2(\alpha+2)-1\Big)\Big(\tilde{\alpha}_2(\alpha+2)-2\Big)}\right]\bar{a}^2_{\overline{t}|\delta},$$

for $\tilde{\alpha}_1 \geq \frac{1}{1+\alpha}$.

Proof. We have from (4.1)

$$\lim_{s\to 0}\frac{\partial f^*_{\Theta,\Lambda}(-s,x)}{\partial s} = a\alpha_1\left[1+bx\right]^{-\tilde{\alpha}_2(1+\alpha)}, \tag{29}$$

and

$$\lim_{s\to 0}\frac{\partial^2 f^*_{\Theta,\Lambda}(-s,x)}{\partial s^2} = a^2\left[\alpha_1(1-\tilde{\alpha_1})\left(1+bx\right)^{-\tilde{\alpha_2}(1+\alpha)} + \alpha_1\tilde{\alpha_1}(1+\alpha)\left(1+bx\right)^{-\tilde{\alpha_2}(2+\alpha)}\right]. \tag{30}$$

Let $I(n,\alpha,b)$ be defined as

$$I(n,\alpha,b) = \int_0^\infty s^{n-1}(1+bs)^{-\alpha}ds, \qquad n\in\mathbb{N}^\star, \quad \alpha>0.$$

Set $x=(1+bs)^{-1}$, the integral becomes

$$I(n,\alpha,b) = \frac{1}{b^n}\int_0^1 x^{\alpha-n-1}(1-x)^{n-1}dx = \frac{\Gamma(n)\Gamma(\alpha-n)}{b^n\Gamma(\alpha)}, \tag{31}$$

for $\alpha > n$. Combination of (24), (29) and (31) yields

$$E\left[\frac{\Theta}{\Lambda}\right] = \frac{a\alpha_1}{\Gamma(1)}I\Big(1,\tilde{\alpha}_2(\alpha+1),b\Big) = \frac{a\alpha_1}{b\Big(\tilde{\alpha}_2(\alpha+1)-1\Big)}.$$

Substitution of (29) into (24) and use of (31) gives

$$E\left[\frac{\Theta}{\Lambda^2}\right] = \frac{a\alpha_1}{\Gamma(2)}I\Big(2,\tilde{\alpha}_2(\alpha+1),b\Big) = \frac{a\alpha_1}{b^2\Big(\tilde{\alpha}_2(\alpha+1)-1\Big)\Big(\tilde{\alpha}_2(\alpha+1)-2\Big)}.$$

Similarly, susbtitution of (30) into (24) and use of (31) gives

$$\begin{aligned} E\left[\frac{\Theta^2}{\Lambda^2}\right] &= \frac{a^2\alpha_1(1-\tilde{\alpha})}{\Gamma(2)}I\Big(2,\tilde{\alpha}_2(\alpha+1),b\Big) + \frac{a^2\alpha_1\tilde{\alpha}_1(1+\alpha)}{\Gamma(2)}I\Big(2,\tilde{\alpha}_2(\alpha+2),b\Big), \\ &= \frac{a^2}{b^2}\left[\frac{\alpha_1(1-\tilde{\alpha}_1)}{\Big(\tilde{\alpha}_2(\alpha+1)-1\Big)\Big(\tilde{\alpha}_2(\alpha+1)-2\Big)} + \frac{\alpha_1\tilde{\alpha}_1(1+\alpha)}{\Big(\tilde{\alpha}_2(\alpha+2)-1\Big)\Big(\tilde{\alpha}_2(\alpha+2)-2\Big)}\right]. \end{aligned}$$

Finally, we find the expressions for $E[\mathcal{Z}]$ and $E\left[\mathcal{Z}^2(t)\right]$ by applying the Corollary (3.1). □

Corollary 4.2. *For the special case $\alpha_1 = \alpha_2 = \alpha$, we have*

$$E[\mathcal{Z}(t)] = \frac{a}{b}\bar{a}_{\overline{t}|\delta}, \tag{32}$$

and

$$E\left[\mathcal{Z}^2(t)\right] = \frac{2a}{b^2(\alpha-1)}\bar{a}_{\overline{t}|2\delta} + \frac{a^2}{b^2}\bar{a}^2_{\overline{t}|\delta}. \tag{33}$$

Proof. The result follows directly from Corollary (4.1). □

4.2. Lomax Copula with Pareto Marginal Distributions

In the previous example and for the special case $\alpha_1 = \alpha_2 = \alpha$, we have

$$f^{\star}_{\Theta,\Lambda}(s,x) = (1+as+bx)^{-\alpha}, \quad s \geq 0, \quad x \geq 0.$$

This specification of the joint Laplace transform leads to the Clayton copula model with the same parameter for the copulas C_1, C_2 and C_{12}. It is possible to modify this model in order to include more flexibility in the model. In this example, it is assumed that the random vector (Θ, Λ) has a bivariate Gamma distribution with the following Laplace transform

$$f^{\star}_{\Theta,\Lambda}(s,x) = (1+as+bx+csx)^{-\alpha}, \quad s \geq 0, \quad x \geq 0, \tag{34}$$

with $c \geq 0$. The extra parameter c introduces more flexible dependence between the mixing distributions and between the Xs and Ws. For example, it is possible to obtain the independence between Θ and Λ which implies that W and X are independent when $c = ab$. The univariate Laplace transforms are given by

$$f^{\star}_{\Theta}(s) = (1+as)^{-\alpha},$$

and

$$f^{\star}_{\Lambda}(x) = (1+bx)^{-\alpha}.$$

It follows that the copulas C_1 and C_2 are Clayton copulas with dependence parameter $\frac{1}{\alpha}$. The joint survival copula of (W, X) is given by

$$\begin{aligned} C_{12}(u,v) &= f^{\star}_{\Theta,\Lambda}\left(a^{-1}(u^{\frac{-1}{\alpha}}-1), b^{-1}(v^{\frac{-1}{\alpha}}-1)\right) \\ &= \left(u^{\frac{-1}{\alpha}} + v^{\frac{-1}{\alpha}} - 1 + \frac{c}{ab}\left(u^{\frac{-1}{\alpha}}-1\right)\left(v^{\frac{-1}{\alpha}}-1\right)\right)^{-\alpha} \\ &= uv\left(u^{\frac{1}{\alpha}} + v^{\frac{1}{\alpha}} - u^{\frac{1}{\alpha}}v^{\frac{1}{\alpha}} + \frac{c}{ab}u^{\frac{1}{\alpha}}v^{\frac{1}{\alpha}}\left(u^{\frac{-1}{\alpha}}-1\right)\left(v^{\frac{-1}{\alpha}}-1\right)\right)^{-\alpha} \\ &= uv\left(1-\gamma(1-u^{\frac{1}{\alpha}})(1-v^{\frac{1}{\alpha}})\right)^{-\alpha}, \end{aligned} \tag{35}$$

which is the Lomax copula defined in Fang et al. (2000) with Kendall's tau, τ_{12}, given by (see e.g., Fang et al. 2000):

$$\tau_{12} = \frac{2\alpha\gamma}{(2\alpha+1)^2}\sum_{k=0}^{\infty}\frac{k!\gamma^k}{(2\alpha+2)_k}, \tag{36}$$

where $(a)_k = a(a+1)\cdots(a+k-1)$, and $(a)_0 = 1$ where a is a real number (See e.g., Erdélyi et al. 1953). Some properties of the family of copulas in (35) are the following:

- when $c = ab$, $(\gamma = 0)$, $C_{12}(uv) = uv$ corresponds to the case of independence.
- as $\alpha = 1$, C_{12} in (35) becomes $C_{12}(u,v) = \frac{uv}{1-\gamma(1-u)(1-v)}$, which is the Ali-Mikhail-Haq (AMH) copula.

- when $c = 0$, $(\gamma = 1)$, $C_{12}(u, v) = \left(u^{-\frac{1}{\alpha}} + v^{-\frac{1}{\alpha}} - 1\right)^{-\alpha}$ is the Clayton's copula.

Note that from (8) and (10), the joint survival function of $(W_1, W_2, \cdots, W_n)$ and $(X_1, X_2, \cdots, X_n)$ can then be written, for $x_i \geq 0$, $i = 1, \cdots, n$, as

$$\bar{F}_{W_1,\cdots,W_n}(s_1, \cdots, s_n) = \left(1 + a\sum_{i=1}^{n} s_i\right)^{-\alpha}, \tag{37}$$

and

$$\bar{F}_{X_1,\cdots,X_n}(x_1, \cdots, x_n) = \left(1 + b\sum_{i=1}^{n} x_i\right)^{-\alpha}, \tag{38}$$

which are the joint survival function of a Pareto II distribution proposed by Arnold (1983, 2015).

The following corollary gives the expressions of the first two moments of $\mathcal{Z}(t)$ for this model.

Corollary 4.3. *For a given time $t \geq 0$ and a positive constant forces of real interest δ, we have*

$$E\left[\mathcal{Z}(t)\right] = \left(\frac{a}{b} + \frac{c}{b^2(\alpha - 1)}\right)\bar{a}_{\overline{t}|\delta},$$

for $\alpha > 1$, and

$$E\left[\mathcal{Z}^2(t)\right] = 2\left(\frac{ab\alpha + 2(c - ab)}{b^3(\alpha - 1)(\alpha - 2)}\right)\bar{a}_{\overline{t}|2\delta} + \left(\frac{a^2}{b^2} + \frac{4ac}{b^3(\alpha - 1)} + \frac{6c^2}{b^4(\alpha - 1)(\alpha - 2)}\right)\bar{a}^2_{\overline{t}|\delta},$$

for $\alpha > 2$.

Proof. Use of (24) and (34), show that

$$\begin{aligned} E\left[\frac{\Theta^l}{\Lambda^n}\right] &= \frac{\Gamma(\alpha + l)}{\Gamma(n)\Gamma(\alpha)}\int_0^\infty x^{n-1}(a + cx)^l(1 + bx)^{-(\alpha+l)}dx \\ &= \frac{\Gamma(\alpha + l)}{\Gamma(n)\Gamma(\alpha)}\sum_{j=0}^{l}\binom{l}{j}a^{l-j}c^j I(n + j, \alpha + l, b), \end{aligned} \tag{39}$$

where $I(n, \alpha, b) = \int_0^\infty x^{n-1}(1 + bx)^{-\alpha}dx$. With the help of (31) and (39), one gets

$$E\left[\frac{\Theta}{\Lambda}\right] = \alpha\left[aI(1, \alpha + 1, b) + cI(2, \alpha + 1, b)\right] = \frac{a}{b} + \frac{c}{b^2(\alpha - 1)},$$

$$E\left[\frac{\Theta}{\Lambda^2}\right] = \alpha\left[aI(2, \alpha + 1, b) + cI(3, \alpha + 1, b)\right] = \frac{ab\alpha + 2(c - ab)}{b^3(\alpha - 1)(\alpha - 2)},$$

and

$$\begin{aligned} E\left[\frac{\Theta^2}{\Lambda^2}\right] &= \alpha(\alpha + 1)\left[a^2 I(2, \alpha + 2, b) + 2acI(3, \alpha + 2, b) + c^2 I(4, \alpha + 2, b)\right] \\ &= \frac{a^2}{b^2} + \frac{4ac}{b^3(\alpha - 1)} + \frac{6c^2}{b^4(\alpha - 1)(\alpha - 2)}. \end{aligned}$$

Applying corollary (3.1), we obtain expressions for the first two moments $E\left[\mathcal{Z}(t)\right]$ and $E\left[\mathcal{Z}^2(t)\right]$. □

4.3. Lomax Copulas and Mixed Exponential-Negative Binomial Marginal Distributions

The next model that we consider in our examples is the mixed exponential-Negative Binomial marginal distributions with Lomax copulas. For this purpose it is assumed that (Θ, Λ) has a bivariate shifted Negative Binomial distribution (see e.g., Marshall and Olkin 1988), the Laplace transform of (Θ, Λ) is defined by

$$f^{\star}_{\Theta,\Lambda}(s,x) = \left(\frac{p}{e^{s+x}-q}\right)^{\alpha}, \quad s,x \geq 0, \tag{40}$$

where $\alpha > 0$, $0 < p < 1$ and $q = 1-p$. Then, the random variables Θ and Λ are distributed as shifted Negative Binomial distributions $\Theta \sim \mathcal{NB}(p,\alpha)$ and $\Lambda \sim \mathcal{NB}(p,\alpha)$. With the help of (8), the multivariate survival function of $(W_1, W_2, \cdots, W_n)$ can be written, for $s_i \geq 0$, $i = 1, \cdots, n$, as

$$\bar{F}_{W_1,\cdots,W_n}(s_1,\cdots,s_n) = \left(\frac{p}{e^{\sum_{i=1}^{n} s_i}-q}\right)^{\alpha}. \tag{41}$$

Then, the marginal survival functions of W_i is given, for $s \geq 0$, by

$$\bar{F}_{W_i}(s) = \left(\frac{p}{e^{s}-q}\right)^{\alpha}, \quad i = 1, \cdots, n. \tag{42}$$

The corresponding copula takes the form

$$C_1(u_1,\cdots,u_n) = \left(\frac{p}{\prod_{i=1}^{n}\left(pu_i^{\frac{-1}{\alpha}}+q\right)-q}\right)^{\alpha}, \tag{43}$$

for $(u_1, \cdots, u_n) \in [0,1]^n$. Similarly, the joint survival function of $(X_1, X_2, \cdots, X_n)$ can be written, for $x_i \geq 0$, $i = 1, \cdots, n$, as

$$\bar{F}_{X_1,\cdots,X_n}(x_1,\cdots,x_n) = \left(\frac{p}{e^{\sum_{i=1}^{n} x_i}-q}\right)^{\alpha}. \tag{44}$$

The marginal survival functions of X_i is given by

$$\bar{F}_{X_i}(x) = \left(\frac{p}{e^{x}-q}\right)^{\alpha}, \quad i = 1, \cdots, n, \tag{45}$$

for $x \geq 0$ and $i = 1, \cdots, n$. The corresponding dependence structure takes the form

$$C_2(u_1,\cdots,u_n) = \left(\frac{p}{\prod_{i=1}^{n}\left(pu_i^{\frac{-1}{\alpha}}+q\right)-q}\right)^{\alpha}. \tag{46}$$

Note that the marginal survival functions of W_i and X_i, $i = 1, \cdots, n$, in (42) and (45) correspond to the survival function of the univariate mixed exponential-geometric distribution introduced in Adamidis and Loukas (1998). It is useful to note that the mixed exponential-geometric distribution is completely monotone (see Marshall and Olkin 1988). The copulas C_1 and C_2 in (43) and (46) are multivariate shifted negative binomial copulas presented in Joe (2014).

The joint survival function of the bivariate random vector (W_i, X_i) is given by

$$\bar{F}_{W_i,X_i}(s,x) = \left(\frac{p}{e^{s+x}-q}\right)^{\alpha}, \quad s, x \geq 0,$$

for $i = 1, \cdots, n$. Then, the corresponding dependence structure is the copula C_{12} given by

$$\begin{aligned} C_{12}(u_1,u_2) &= \left(\frac{p}{(q+pu_1^{-\frac{1}{\alpha}})(q+pu_2^{-\frac{1}{\alpha}})-q}\right)^{\alpha} \\ &= \left(\frac{pu_1^{\frac{1}{\alpha}}u_2^{\frac{1}{\alpha}}}{(qu_1^{\frac{1}{\alpha}}+p)(qu_2^{\frac{1}{\alpha}}+p)-qu_1^{\frac{1}{\alpha}}u_2^{\frac{1}{\alpha}}}\right)^{\alpha} \\ &= \frac{u_1u_2}{\left(1-q(1-u_1^{\frac{1}{\alpha}})(1-u_2^{\frac{1}{\alpha}})\right)^{\alpha}}, \end{aligned} \tag{47}$$

which corresponds to the Lomax copula.

We now state a Corollary for calculating the first an second moments of the discounted aggregate renewal claims.

Corollary 4.4. *For a positive constant forces of real interest δ:*

$$E\left[\mathcal{Z}(t)\right] = \bar{a}_{\overline{t}|\delta}, \tag{48}$$

and

$$E\left[\mathcal{Z}^2(t)\right] = \bar{a}^2_{\overline{t}|\delta} + 2\left(\frac{p}{q}\right)^{\alpha} B(q;\alpha,1-\alpha)\bar{a}_{\overline{t}|2\delta}, \tag{49}$$

where $B(z;\alpha,\beta) = \int_0^z u^{\alpha-1}(1-u)^{\beta-1}du$ is the incomplete Beta function.

Proof. From elementary calculus, one gets from (40)

$$\lim_{s\to 0}\frac{\partial f^{\star}_{\Theta,\Lambda}(-s,x)}{\partial s} = \alpha p^{\alpha}\frac{e^x}{(e^x-q)^{\alpha+1}}. \tag{50}$$

Substituting the last expression into (24) with ($n = l = 1$) yields $E\left[\frac{\Theta}{\Lambda}\right] = 1$. Combining this with Corollary (3.1), one gets (48). Otherwise, we get from (24) with ($n = 2$ and $l = 1$)

$$\begin{aligned} E\left[\frac{\Theta}{\Lambda^2}\right] &= \alpha p^{\alpha}\int_0^{\infty} x\frac{e^x}{(e^x-q)^{\alpha+1}}dx = p^{\alpha}\int_0^{\infty}\frac{1}{(e^x-q)^{\alpha}}dx \\ &= \left(\frac{p}{q}\right)^{\alpha}\int_0^q u^{\alpha-1}(1-u)^{-\alpha}du = \left(\frac{p}{q}\right)^{\alpha}B(q;\alpha,1-\alpha), \end{aligned} \tag{51}$$

where $B(z;\alpha,\beta) = \int_0^z u^{\alpha-1}(1-u)^{\beta-1}du$ is the incomplete Beta function. Otherwise, $\lim_{s\to 0}\frac{\partial^2 f^{\star}_{\Theta,\Lambda}(-s,x)}{\partial^2 s} = \alpha p^{\alpha}\frac{qe^x+\alpha e^{2x}}{(e^x-q)^{\alpha+2}}$. Substituting the last expression into (24) with ($n = 2$ and $l = 2$), one gets

$$E\left[\frac{\Theta^2}{\Lambda^2}\right] = \alpha q p^{\alpha}\int_0^{\infty}\frac{xe^x}{(e^x-q)^{\alpha+2}}dx + \alpha^2 p^{\alpha}\int_0^{\infty}\frac{xe^{2x}}{(e^x-q)^{\alpha+2}}dx. \tag{52}$$

Otherwise, integration by parts gives

$$\begin{aligned}\int_0^\infty \frac{xe^x}{(e^x-q)^{\alpha+2}}dx &= \frac{1}{\alpha+1}\int_0^\infty \frac{1}{(e^x-q)^{\alpha+1}}dx \\ &= \frac{1}{\alpha+1}\frac{1}{q^{\alpha+1}}B(q;\alpha+1,-\alpha).\end{aligned} \tag{53}$$

Similarly, integrating by parts

$$\begin{aligned}\int_0^\infty \frac{xe^{2x}}{(e^x-q)^{\alpha+2}}dx &= \frac{1}{\alpha+1}\int_0^\infty \frac{e^x+xe^x}{(e^x-q)^{\alpha+1}}dx \\ &= \frac{1}{\alpha+1}\left(\frac{1}{\alpha p^\alpha}+\frac{1}{\alpha}\frac{1}{q^\alpha}B(q;\alpha,-\alpha+1)\right).\end{aligned} \tag{54}$$

Hence, through (52), (53) and (54), we obtain

$$E\left[\frac{\Theta^2}{\Lambda^2}\right] = \frac{\alpha}{(\alpha+1)}+\frac{\alpha p^\alpha}{(\alpha+1)q^\alpha}\left(B(q;\alpha+1,-\alpha)+B(q;\alpha,1-\alpha)\right)=1.$$

Finally, we combine the last expression with (51) and Corollary (3.1) to obtain (49). □

Note that if $\alpha = 1$, the copula C_{12} in (48) reduces to the AMH copula with Kendall's tau, τ_{12}, given by (see e.g., Nelsen 1999)

$$\tau_{12} = \frac{3q-2}{3q}-\frac{2(1-q)^2 ln(1-q)}{3q^2}.$$

For this special case, we obtain $E\left[\mathcal{Z}(t)\right]=\bar{a}_{\overline{t}|\delta}$, and $E\left[\mathcal{Z}^2(t)\right]=\bar{a}^2_{\overline{t}|\delta}-2(\frac{p}{q})log(p)\bar{a}_{\overline{t}|2\delta}$.

4.4. Numerical Illustrations

In this subsection, we present numerical examples to illustrate how the distribution of the discounted renewal aggregate claims behaves when we change the dependency parameters. The computations provided are related to the general case of Clayton copulas. For the discounted aggregate amount of claims, as in Section 4.1, we assume that the force of interest is fixed at the value of $\delta = 5\%$ and we set $a = 1$ and $b = 0.2$. The sensitivity analysis is done by varying Kendall's tau dependence measures τ_i, $i = 1, 2$ and τ_{12} given by (27) and (28) respectively. In order to investigate the impact of the dependence structure on the distribution of $\mathcal{Z}(t)$, we compute the mean $E[\mathcal{Z}(t)]$, the standard deviation $SD[\mathcal{Z}(t)]$, the skewness $Skew[\mathcal{Z}(t)]$ and the kurtosis $Kurt[\mathcal{Z}(t)]$ using different values for the Kendall tau's of the copulas C_{12}, C_1 and C_2. Both the expressions of $E[\mathcal{Z}(t)]$ and $SD[\mathcal{Z}(t)]$ are given in Section 4.1. The third and the fourth moments are computed numerically. Using the software Matlab, we evaluate the integral in (25) then we use the closed form in (3.1) for $n = 3$ and 4. The results are presented using different time horizons where t is set to be 110, 100 and ∞.

Tables 1–3 display the obtained results. For a fixed t, τ_1 and τ_{12}, increasing the dependence between the claims leads to a higher level of risk, i.e., large values of $E\left[\mathcal{Z}(t)\right]$ and $SD\left[\mathcal{Z}(t)\right]$. On the other hand, increasing the dependence between the inter-claim times reduces the level of risk for the whole portfolio. We also notice that both the expected value and volatility of the aggregate discounted claims decrease as τ_{12} increases. A strong positive dependence between the inter-claim times and the claim sizes means that the portfolio generates either large and less frequent losses or small and very frequent losses. This leads to a small value of $E\left[\mathcal{Z}(t)\right]$ and less volatile $\mathcal{Z}(t)$. Increasing the dependence parameter τ_{12} or τ_1 generates longer and fatter right tails. Decreasing τ_2 has the same impact on the shape of the tails as increasing the Kendall's tau measures of the copulas C_{12} and C_1.

Table 1. Impact of changing τ_{12} on the distribution of $\mathcal{Z}(t)$ with $\tau_1 = 0.8$ and $\tau_2 = 0.3$.

$E[\mathcal{Z}(t)]$	τ_{12}	$t=1$	$t=10$	$t=100$	$t=\infty$
	0.45	0.2937	2.3694	5.9813	6.0219
	0.55	0.2020	1.6294	4.1132	4.1411
	0.65	0.1355	1.0930	2.7591	2.7778
	0.75	0.0851	0.6863	1.7324	1.7442
$SD[\mathcal{Z}(t)]$	τ_{12}	$t=1$	$t=10$	$t=100$	$t=\infty$
	0.45	0.8740	4.0281	9.2282	9.2864
	0.55	0.7306	3.4214	7.8775	7.9274
	0.65	0.5970	2.7841	6.4017	6.4422
	0.75	0.4650	2.0910	4.7519	4.7816
$Skew[\mathcal{Z}(t)]$	τ_{12}	$t=1$	$t=10$	$t=100$	$t=\infty$
	0.45	1.8969	1.2997	1.3984	1.3991
	0.55	2.5299	2.0445	2.2137	2.2148
	0.65	3.3014	3.0643	3.3997	3.4018
	0.75	4.3794	4.9922	5.8956	5.9010
$Kurt[\mathcal{Z}(t)]$	τ_{12}	$t=1$	$t=10$	$t=100$	$t=\infty$
	0.45	3.9901	2.1244	1.9211	1.9199
	0.55	6.7845	5.3719	5.7602	5.7624
	0.65	11.5174	13.0582	15.3542	15.3674
	0.75	21.2007	39.8072	52.3969	52.4717

Table 2. Impact of changing τ_1 on the distribution of $\mathcal{Z}(t)$ with $\tau_{12} = 0.4$ and $\tau_2 = 0.2$.

$E[\mathcal{Z}(t)]$	τ_{12}	$t=1$	$t=10$	$t=100$	$t=\infty$
	0.7	0.2850	2.2995	5.8048	5.8442
	0.75	0.2217	1.7885	4.5148	4.5455
	0.8	0.1663	1.3414	3.3861	3.4091
	0.85	0.1174	0.9469	2.3902	2.4064
$SD[\mathcal{Z}(t)]$	τ_{12}	$t=1$	$t=10$	$t=100$	$t=\infty$
	0.7	0.8683	4.0706	9.3752	9.4346
	0.75	0.7748	3.7135	8.6088	8.6637
	0.8	0.6777	3.3068	7.7043	7.7536
	0.85	0.5744	2.8438	6.6520	6.6947
$Skew[\mathcal{Z}(t)]$	τ_{12}	$t=1$	$t=10$	$t=100$	$t=\infty$
	0.7	1.9920	1.4973	1.6216	1.6225
	0.75	2.3856	1.7579	1.8353	1.8358
	0.8	2.8825	2.1144	2.1588	2.1592
	0.85	3.5647	2.6220	2.6404	2.6406
$Kurt[\mathcal{Z}(t)]$	τ_{12}	$t=1$	$t=10$	$t=100$	$t=\infty$
	0.7	4.1974	1.0229	0.1587	0.1537
	0.75	5.6748	1.3470	0.3437	0.3381
	0.8	7.9450	1.9977	0.7742	0.7675
	0.85	11.7835	3.2307	1.6279	1.6194

Table 3. Impact of changing τ_2 on the distribution of $\mathcal{Z}(t)$ with $\tau_{12} = 0.55$ and $\tau_1 = 0.85$.

$E[\mathcal{Z}(t)]$	τ_{12}	$t=1$	$t=10$	$t=100$	$t=\infty$
	0.05	0.0136	0.1094	0.2763	0.2781
	0.15	0.0491	0.3964	1.0006	1.0073
	0.25	0.1033	0.8332	2.1034	2.1176
	0.35	0.1957	1.5792	3.9865	4.0136
$SD[\mathcal{Z}(t)]$	τ_{12}	$t=1$	$t=10$	$t=100$	$t=\infty$
	0.05	0.1974	0.9952	2.3387	2.3537
	0.15	0.3730	1.8589	4.3553	4.3833
	0.25	0.5349	2.6167	6.1008	6.1399
	0.35	0.7224	3.4130	7.8788	7.9288
$Skew[\mathcal{Z}(t)]$	τ_{12}	$t=1$	$t=10$	$t=100$	$t=\infty$
	0.05	11.4633	9.3207	9.4760	9.4772
	0.15	5.8516	4.7009	4.8002	4.8009
	0.25	3.8518	3.0297	3.1204	3.1211
	0.35	2.5621	1.9254	2.0290	2.0298
$Kurt[\mathcal{Z}(t)]$	τ_{12}	$t=1$	$t=10$	$t=100$	$t=\infty$
	0.05	116.7367	44.8970	32.4352	32.3701
	0.15	31.1083	12.4559	9.3112	9.2946
	0.25	14.0675	6.5740	5.5232	5.5176
	0.35	6.9009	5.2676	5.6800	5.6824

5. Conclusions

In this paper, we derived explicit expressions for the higher moments of the discounted aggregate renewal claims with dependence. Closed expressions for the moments of the aggregate discounted claims are obtained when the claims and the subsequent inter-claim are distributed as Pareto and Mixed exponential-geometric distributions. Numerical examples are given to illustrate the impact of dependency on the moments of the discounted aggregate renewal mixed process.

Since the assumption of constant force of interest is quite restrictive, studying the discounted renewal aggregate claims with a stochastic force of interest would be interesting. A more challenging problem would be the extension of the mixed exponential risk model to incorporate other forms of dependence structure between the model components.

Author Contributions: All authors contributed equally to this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

Adamidis, Konstantinos, and Loukas Sotirios. 1998. A lifetime distribution with decreasing failure rate. *Statistics & Probability Letters* 39: 35–42.

Albrecher, Hansjörg, and Onno J. Boxma. 2004. A ruin model with dependence between claim sizes and claim intervals. *Insurance: Mathematics and Economics* 35: 245–54. [CrossRef]

Albrecher, Hansjörg, and Jef L. Teugels. 2006. Exponential behavior in the presence of dependence in risk theory. *Journal of Applied Probability* 43: 257–73. [CrossRef]

Albrecher, Hansjörg, Corina Constantinescu, and Stephane Loisel. 2011. Explicit ruin formulas for models with dependence among risks. *Insurance: Mathematics and Economics* 48: 265–70. [CrossRef]

Andersen, E. Sparre. 1957. On the collective theory of risk in case of contagion between claims. *Bulletin of the Institute of Mathematics and its Applications* 12: 275–79.

Arnold, Barry C. 1983. *Pareto Distributions*. Fairland: International Cooperative Publishing House.

Arnold, Barry C. 2015. *Pareto Distributions*. Boca Raton: Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

Barges, Mathieu, Hélène Cossette, Stéphane Loisel, and Etienne Marceau. 2011. On the moments of aggregate discounted claims with dependence introduced by a FGM copula. *ASTIN Bulletin: The Journal of the IAA* 41: 215–38.

Boudreault, Mathieu, Hélène Cossette, David Landriault, and Etienne Marceau. 2006. On a risk model with dependence between interclaim arrivals and claim sizes. *Scandinavian Actuarial Journal* 2006: 265–85. [CrossRef]

Clayton, David G. 1978. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* 65: 141–51. [CrossRef]

Cressie, Noel, Anne S. Davis, J. Leroy Folks, and J. Leroy Folks. 1981. The moment-generating function and negative integer moments. *The American Statistician* 35: 148–50.

Erdélyi, Arthur, Wilhelm Magnus, Fritz Oberhettinger, and Francesco G. Tricomi. 1953. *Higher Transcendental Functions (California Institute of Technology H. Bateman MS Project)*. New York: McGraw Hill, vol. 1.

Faa di Bruno, Francesco. 1855. Sullo sviluppo delle funzioni. *Annali di Scienze Matematiche e Fisiche* 6: 479–80.

Fang, Kai-Tai, Hong-Bin Fang, and Dietrich von Rosen. 2000. A family of bivariate distributions with non-elliptical contours. *Communications in Statistics-Theory and Methods* 29: 1885–98. [CrossRef]

Ginsburg, Jekuthiel. 1928. Discussions: Note on stirling's numbers. *The American Mathematical Monthly* 35: 77–80. [CrossRef]

Guo, Ling, David Landriault, and Gordon E. Willmot. 2013. On the analysis of a class of loss models incorporating time dependence. *European Actuarial Journal* 3: 273–94. [CrossRef]

Jang, Jiwook, Angelos Dassios, and Hongbiao Zhao. 2018. Moments of renewal shot-noise processes and their applications. *Scandinavian Actuarial Journal* 2018: 1–26.

Joe, Harry. 1997. *Multivariate Models and Multivariate Dependence Concepts*. London: Chapman & Hall.

Joe, Harry. 2014. *Dependence Modeling with Copulas*. Boca Raton: CRC Press.

Landriault, David, Gordon E. Willmot, and Di Xu. 2014. On the analysis of time dependent claims in a class of birth process claim count models. *Insurance: Mathematics and Economics* 58: 168–73. [CrossRef]

Léveillé, Ghislain, and Franck Adékambi. 2011. Covariance of discounted compound renewal sums with a stochastic interest rate. *Scandinavian Actuarial Journal* 2: 138–53. [CrossRef]

Léveillé, Ghislain, and Franck Adékambi. 2012. Joint moments of discounted compound renewal sums. *Scandinavian Actuarial Journal* 1: 40–55. [CrossRef]

Léveillé, Ghislain, and José Garrido. 2001a. Moments of compound renewal sums with discounted claims. *Insurance: Mathematics and Economics* 28: 217–31. [CrossRef]

Léveillé, Ghislain, and José Garrido. 2001b. Recursive moments of compound renewal sums with discounted claims. *Scandinavian Actuarial Journal* 2: 98–110. [CrossRef]

Léveillé, Ghislain, and Emmanuel Hamel. 2013. A compound renewal model for medical malpractice insurance. *European Actuarial Journal* 3: 471–90. [CrossRef]

Léveillé, Ghislain, José Garrido, and Ya Fang Wang. 2010. Moment generating functions of compound renewal sums with discounted claims. *Scandinavian Actuarial Journal* 3: 165–84.

Marshall, Albert W., and Ingram Olkin. 1988. Families of multivariate distributions. *Journal of the American Statistical Association* 83: 834–41. [CrossRef]

Nelsen, Roger B. 1999. *An Introduction to Copulas*. New York: Springer.

Sklar, Abe. 1959. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris* 8: 229–31.

risks

MDPI

Article

On a Multiplicative Multivariate Gamma Distribution with Applications in Insurance

Vadim Semenikhine [1], Edward Furman [1] and Jianxi Su [2,*]

[1] Department of Mathematics and Statistics, York University, Toronto, ON M3J 1P3, Canada; vadia@my.yorku.ca (V.S.); efurman@mathstat.yorku.ca (E.F.)
[2] Department of Statistics, Purdue University, West Lafayette, IN 47906, USA
* Correspondence: jianxi@purdue.edu; Tel.: +1-765-494-6037

Received: 13 July 2018; Accepted: 8 August 2018; Published: 12 August 2018

Abstract: One way to formulate a multivariate probability distribution with dependent univariate margins distributed gamma is by using the closure under convolutions property. This direction yields an additive background risk model, and it has been very well-studied. An alternative way to accomplish the same task is via an application of the Bernstein–Widder theorem with respect to a shifted inverse Beta probability density function. This way, which leads to an arguably equally popular multiplicative background risk model (MBRM), has been by far less investigated. In this paper, we reintroduce the multiplicative multivariate gamma (MMG) distribution in the most general form, and we explore its various properties thoroughly. Specifically, we study the links to the MBRM, employ the machinery of divided differences to derive the distribution of the aggregate risk random variable explicitly, look into the corresponding copula function and the measures of nonlinear correlation associated with it, and, last but not least, determine the measures of maximal tail dependence. Our main message is that the MMG distribution is (1) very intuitive and easy to communicate, (2) remarkably tractable, and (3) possesses rich dependence and tail dependence characteristics. Hence, the MMG distribution should be given serious considerations when modelling dependent risks.

Keywords: multivariate gamma distribution; multiplicative background risk model; aggregate risk; individual risk model; collective risk model

1. Introduction

Let $\mathcal{X}$ be a collection of actuarial risks, that is let it contain random variables (r.v.'s) $X : \Omega \to \mathbf{R}$ defined on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and interpreted as the financial risks an insurer is exposed to. Often, for applications in insurance, actuaries would consider those $X \in \mathcal{X}$, whose distributions are supported on the non-negative real half-line, have positive skewness, and allow for a certain degree of heavy-tailness. One such distribution, which has been of prominent importance in insurance applications, is gamma. We refer to Hürlimann (2001), Dornheim and Brazauskas (2007), Furman et al. (2018), and Zhou et al. (2018) for applications in solvency assessment, loss reserving, and aggregate risk approximation, respectively.

Furthermore, let $\gamma \in \mathbf{R}_+$ and $\sigma \in \mathbf{R}_+$ denote, correspondingly, the shape and scale parameters, then the r.v. X is said to be distributed gamma, succinctly $X \sim Ga(\gamma, \sigma)$, if it has the probability density function (p.d.f.)

$$f(x) = \frac{1}{\Gamma(\gamma)} e^{-x/\sigma} x^{\gamma-1} \sigma^{-\gamma} \text{ for all } x \in \mathbf{R}_+, \tag{1}$$

where $\Gamma(\cdot)$ stands for the complete gamma function. The popularity of the r.v.'s distributed gamma in insurance applications is not surprising: the p.d.f.'s of the (aggregate) insurance losses have as

a rule the same shape as p.d.f. (1), i.e., they are positively skewed, unimodal and have positive supports; p.d.f. (1) is log-convex for $\gamma \in (0,1)$ and so has decreasing failure rate, thus allowing for moderate heavy-tailness (Klugman et al. 2012); p.d.f. (1) has been very well studied and has turned out remarkably tractable.

When it comes to multivariate extensions of p.d.f. (1), there are an ample number of dependence structures with univariate margins distributed gamma to consider (e.g., Kotz et al. 2000; Balakrishnan and Ristić 2016, for a recent development and a comprehensive reference, respectively). However, irrespective of whether the two-steps copula approach or the more 'natural' stochastic representation approach to formulate the desired multivariate gamma distribution is pursued, the tractability of the end-result is often an issue. For the former approach, the cumulative distribution function (c.d.f.) of (1) cannot be written in a closed form, and consequently intensive numerical algorithms are often needed to implement copula-based multivariate gamma models (e.g., Cossette et al. 2018; Bahraoui et al. 2015). For the latter approach, consider the following example. Let $Y_j \sim Ga(\gamma_j, \sigma)$ for $\gamma_j \in \mathbf{R}_+$ and $j = 1, \ldots, n+1$ be mutually independent r.v.'s, and set $\mathbf{X} = (X_1, \ldots, X_n)' = (Y_1 + Y_{n+1}, \ldots, Y_n + Y_{n+1})'$. Then the distribution of the r.v. $\mathbf{X}$ is the multivariate gamma of Mathai and Moschopoulos (1991) (also, e.g., Avanzi et al. 2016; Furman and Landsman 2005, for recent applications in insurance). Consequently, for the p.d.f. of the r.v. $\mathbf{X}$, we have

$$f(x_1, \ldots, x_n) \propto \int_0^{\min_{i=1,\ldots,n} x_i} x^{\gamma_{n+1}-1} e^{-x/\sigma} \prod_{i=1}^{n} (x_i - x)^{\gamma_i - 1} e^{-(x_i - x)/\sigma} dx \tag{2}$$

for all $(x_1, \ldots, x_n)' \in \mathbf{R}_+^n$, which inconveniently takes distinct forms for each of the $n!$ orderings of $x_1, \ldots, x_n$.

Remark 1. *The r.v.'s $Y_1, \ldots, Y_n$ and Y_{n+1} are often interpreted as, respectively, the specific and systematic risk factors. The systematic risk factor, Y_{n+1}, has also been referred to as the background risk (Gollier and Pratt 1996), and so the distribution of the r.v. $\mathbf{X} = (X_1, \ldots, X_n)'$ can be associated with an Additive Background Risk Model with risk components distributed gamma (G-ABRM). Succinctly, for $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_n)'$, we write $\mathbf{X} \sim Ga_n^+(\boldsymbol{\gamma}, \gamma_{n+1}, \sigma)$, where γ_{n+1} serves as the dependence parameter.*

An alternative way to link the specific risk factors and the systematic (or background) risk factor is with the help of multiplication. Namely, in order to formulate a Multiplicative Background Risk Model with the risk components distributed gamma (G-MBRM), we must find a sequence of $(n+1)$ independent r.v.'s $Z_1, \ldots, Z_n$, Z_{n+1}, say, such that $\mathbf{X} = (X_1, \ldots, X_n)' = (Z_1 Z_{n+1}, \ldots, Z_n Z_{n+1})'$ results in the coordinates of the r.v. $\mathbf{X}$ being distributed gamma. One solution of this exercise, which is of pivotal importance for this paper, can be found in Feller (1968) (also, Albrecher et al. 2011; Sarabia et al. 2018). We organize the rest of the paper as follows: in Section 2, we explore the basic distributional properties of—what we call—the multiplicative multivariate gamma (MMG) distribution. Then, in Sections 3 and 4, respectively, we discuss in detail and elucidate with examples of actuarial interest the aggregation and (tail) dependence properties of the MMG distribution. Section 5 concludes the paper. All proofs are relegated to Appendix A to facilitate the reading.

2. Definition and Basic Properties

Multivariate distributions lay the very foundation of the successful (insurance) risk measurement—and thus of the consequent risk management—processes. However, the toolbox of the available stochastic dependencies that can be used to link stand-alone risk components into risk portfolios is somewhat overwhelming. Indeed, there are infinitely many ways to formulate the joint distribution of two dependent risk r.v.'s, whereas there is a single way only to write this distribution under the assumption of independence. The case of the multivariate distributions with the margins distributed gamma is of course not an exception (e.g., Kotz et al. 2000).

Nevertheless, real applications impose significant constraints on the model choice. Namely, practitioners often opt for those multivariate distributions that: (i) admit meaningful and relevant interpretations; (ii) allow for an adequate fit to the modelled data, be it in the 'tail', in the 'body', and/or in the dependence; and (iii) can be readily implemented. We feel that the multivariate distribution with the univariate margins distributed gamma that we put forward next (also, Albrecher et al. 2011; Sarabia et al. 2018) is exactly such.

Formally, let E_λ and Λ denote, respectively, an exponentially distributed r.v. with the rate parameter $\lambda \in \mathbf{R}_+$ and an arbitrarily distributed r.v. with the range $\Lambda \in \mathbf{R}_+$; assume that the r.v.'s E_Λ and Λ are independent. In addition, let '$*$' represent the mixture operator (e.g., Feller 1968; Su and Furman 2017a), such that, for '$\stackrel{d}{=}$' denoting equality in distribution, it holds that $E_\lambda * \Lambda \stackrel{d}{=} E_\Lambda$. We note in passing that the just-mentioned mixture operator is referred to as 'randomization' in Feller (1968), and is closely related—via the Bernstein–Widder theorem—to the notion of the Laplace transform of the p.d.f. of Λ. More specifically, if f_Λ and $\mathcal{L}\{f_\Lambda\}$ denote, correspondingly, the p.d.f. of Λ and its Laplace transform, which is

$$\mathcal{L}\{f_\Lambda\}(x) = \int_\Lambda e^{-x\lambda} f_\Lambda(\lambda) d\lambda, \tag{3}$$

then (3) establishes the decumulative distribution function (d.d.f.) of the r.v. E_Λ.

Recall that in this paper we are interested in formulating a multivariate distribution with the univariate margins distributed gamma and a dependence. To this end, we assume that the r.v. Λ is distributed as a special shifted inverse Beta, succinctly $\Lambda \sim IB(\gamma)$, with the p.d.f.

$$f_\Lambda(\lambda) \quad = \quad \frac{\lambda^{-1}(\lambda-1)^{-\gamma}}{\Gamma(1-\gamma)\Gamma(\gamma)} \text{ for all } \lambda > 1, \tag{4}$$

where $\gamma \in (0,1)$ is the shape parameter. In our context, the choice of p.d.f. (4) is unique, which readily follows from the Bernstein–Widder theorem. The next few facts are used frequently later on in the paper, and are hence formulated as a lemma. In the following, the k-th order derivative of the Laplace transform is denoted by $\psi^{(k)}$, $k \in \mathbf{N} := \{1, 2, \ldots\}$, also $\mathbf{R}_{0,+} := [0, \infty)$.

Lemma 1. *Let $\Lambda \sim IB(\gamma)$, $\gamma \in (0,1)$ with p.d.f. (4), then:*

(i) The Laplace transform of (4) is

$$\mathcal{L}\{f_\Lambda\}(x) = \Gamma(\gamma,\ x)/\Gamma(\gamma) \text{ for all } x \in \mathbf{R}_{0,+},$$

where $\Gamma(\gamma, x) := \int_x^\infty t^{\gamma-1} e^{-t} dt$ denotes the upper incomplete gamma function.

(ii) The negative k-th order moment of the r.v. Λ is

$$\mathbf{E}[\Lambda^{-k}] = \frac{\Gamma(\gamma+k)}{\Gamma(\gamma)\Gamma(k+1)} \text{ for all } k \in \mathbf{N}.$$

(iii) The alternating sign k-th order derivative of $\mathcal{L}\{f_\Lambda\}$ is

$$(-1)^k \psi_\Lambda^{(k)}(x) = \sum_{i=0}^{k-1} \binom{k-1}{i} \frac{\Gamma(i-\gamma+1)}{\Gamma(1-\gamma)\Gamma(\gamma)} e^{-x} x^{-(i-\gamma+1)} \text{ for all } k \in \mathbf{N}. \tag{5}$$

Let $E_{\lambda,1}, \ldots, E_{\lambda,n}$ denote independent copies of E_λ, and let $\Lambda \sim IB(\gamma)$, $\gamma \in (0,1)$. In addition, let $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)'$ denote a vector of positive parameters.

Definition 1. *Set* $X_j = \sigma_j E_{\lambda,j} * \Lambda$, $j = 1, \ldots, n$, *and then the r.v.* $\mathbf{X} = (X_1, \ldots, X_n)'$ *has a multiplicative multivariate distribution with univariate margins distributed gamma, and we succinctly write* $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$, *where* $\gamma \in (0,1)$ *and* $\boldsymbol{\sigma} \in \mathbf{R}_+^n$ *are parameters.*

Remark 2. *Let* $E_j := E_{1,j}$, $j = 1, \ldots, n$ *denote independent copies of a r.v. distributed exponentially with unit scale, then the joint distribution of the r.v.* $\mathbf{X} = (X_1, \ldots, X_n)'$ *in Definition 1 admits the following multiplicative background risk model representation (see, Asimit et al. 2016; Frank et al. 2006, for the corresponding economic implication and application)*

$$\mathbf{X} = (X_1, \ldots, X_n)' \stackrel{d}{=} (\sigma_1 E_1/\Lambda, \ldots, \sigma_n E_n/\Lambda)'. \tag{6}$$

Above, the r.v. Λ *can be interpreted as the systematic risk factor that endangers every risk component of the portfolio* $\mathbf{X} = (X_1, \ldots, X_n)'$ *in Equation (6). The Monte Carlo simulation of Equation (6) is immediate.*

Theorem 1. *Let* $\Lambda \sim IB(\gamma)$, $\gamma \in (0,1)$, *and let* $\sigma_1, \ldots, \sigma_n$ *be positive scale parameters, then the following assertions hold:*

(i) *The r.v.* $X = \sigma E_\lambda * \Lambda$ *has the d.d.f.*

$$\overline{F}(x) = \Gamma(\gamma, x/\sigma)/\Gamma(\gamma) \text{ for all } x \in \mathbf{R}_{0,+},$$

which is X *is distributed gamma with the shape and scale parameters equal to* $\gamma \in (0,1)$ *and* $\sigma \in \mathbf{R}_+$, *respectively.*

(ii) *The r.v.* $\mathbf{X} = (X_1, \ldots, X_n)'$ *with the j-th coordinate* $X_j = \sigma_j E_{\lambda,j} * \Lambda$, *has the joint d.d.f.*

$$\overline{F}(x_1, \ldots, x_n) = \frac{\Gamma(\gamma, x_1/\sigma_1 + \cdots + x_n/\sigma_n)}{\Gamma(\gamma)}, \tag{7}$$

for all $(x_1, \ldots, x_n)' \in \mathbf{R}_{0,+}^n$.

(iii) *The p.d.f. that corresponds to d.d.f. (7) is, for all* $(x_1, \ldots, x_n)' \in \mathbf{R}_+^n$,

$$f(x_1, \ldots, x_n) = \frac{1}{\prod_{i=1}^n \sigma_i} \sum_{i=0}^{n-1} \binom{n-1}{i} \frac{\Gamma(i-\gamma+1)}{\Gamma(1-\gamma)\Gamma(\gamma)} e^{-\sum_{j=1}^n \frac{x_j}{\sigma_j}} \left(\sum_{j=1}^n \frac{x_j}{\sigma_j}\right)^{-(i-\gamma+1)}. \tag{8}$$

The following facts are immediate from Theorem 1: (i) the distribution of $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$ is marginally closed, namely, $X_j \sim Ga(\gamma, \sigma_j)$, $j = 1, \ldots, n$; (ii) the mathematical expectation of the j-th coordinate is $\mathbf{E}[X_j] = \gamma\sigma_j$; and (iii) the variance of the j-th coordinate is $\mathbf{Var}[X_j] = \gamma\sigma_j^2$.

We further explore some less obvious properties of the MMG/G-MBRM and note with satisfaction that the risk portfolios with the joint distributions within this class are often more tractable than the portfolios having stochastically independent risk components distributed gamma. At the outset, we note in passing that the MMG distribution put forward in Definition 1 is a non-exchangeable generalization of the multivariate distributions having univariate margins distributed gamma that are discussed in Albrecher et al. (2011); Sarabia et al. (2018). As such, the MMG distribution requires a more delicate treatment when deriving the results below, which hinge crucially on the stochastic characteristics of the univariate margins of the r.v. $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$.

We look into the minima and maxima r.v.'s first; both are of evident importance in insurance. To this end, denote by $X_{\min} = \min_{i=1,\ldots,n} X_i \sim F_{\min}$ and by $X_{\max} = \max_{i=1,\ldots,n} X_i \sim F_{\max}$ the minima and maxima r.v.'s. Then we have – unlike in the independent case – that the coordinates of the r.v. $\mathbf{X} = (X_1, \ldots, X_n)'$ in Definition 1 are closed under minima.

Theorem 2. *Let* $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$, *then* $X_{\min}$ *is distributed gamma. More specifically, we have* $X_{\min} \sim Ga(\gamma, \sigma^*)$, *where* $\sigma^* = \left(\sum_{j=1}^n 1/\sigma_j\right)^{-1}$ *is the positive scale parameter, and* $\gamma \in (0,1)$ *is the shape parameter. In addition, the d.d.f. of* $X_{\max}$ *is a linear combination of the d.d.f.'s of the univariate r.v.'s distributed gamma, such that*

$$\overline{F}_{\max}(x) = \sum_{\mathcal{S} \subseteq \{1,\ldots,n\}} (-1)^{|\mathcal{S}|-1} \overline{F}_{X_{\mathcal{S}}}(x) \text{ for all } x \in \mathbf{R}_{0,+},$$

where $X_{\mathcal{S}} = \min_{s \in \mathcal{S} \subseteq \{1,\ldots,n\}} X_s$ *and* $X_{\mathcal{S}} \sim Ga(\gamma, \sigma_{\mathcal{S}}^*)$ *with* $\sigma_{\mathcal{S}}^* = (\sum_{j \in \mathcal{S}} 1/\sigma_j)^{-1}$.

Another r.v. of pivotal interest in insurance is the aggregate risk r.v. denoted by $X_+ = X_1 + \cdots + X_n$; in addition, let $X_+ \sim F_+$. It is well known that, if $X_1, \ldots, X_n$ are mutually independent and distributed gamma with arbitrary parameters, then F_+ admits an infinite sum representation (Moschopoulos 1985; Provost 1989). We further show that for $\mathbf{X} \sim G_n^{\times}(\gamma, \boldsymbol{\sigma})$ and when all the scale parameters are distinct, then F_+ is noticeably more elegant. The derivation of F_+ in the general case—i.e., for arbitrary (possibly equal) scale parameters—is more cumbersome and is presented in Section 3.

Let

$$w_i(\boldsymbol{\sigma}) = \prod_{j=1,\, j \neq i}^{n} \frac{1}{1 - \sigma_j/\sigma_i} \text{ for } i = 1, \ldots, n. \tag{9}$$

We often write w_i omitting the vector of scale parameters $\boldsymbol{\sigma}$ for the simplicity of notation.

Proposition 1. *Let* $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$ *and assume that all the scale parameters are distinct, which is* $\sigma_i \neq \sigma_j$ *for* $i \neq j \in \{1, \ldots, n\}$, *then the d.d.f. of* $X_+ = X_1 + \cdots + X_n$ *is*

$$\overline{F}_+(x) = \sum_{i=1}^{n} w_i \frac{\Gamma(\gamma, x/\sigma_i)}{\Gamma(\gamma)} \text{ for all } x \in \mathbf{R}_{0,+}. \tag{10}$$

The last result in this section provides an expression for the higher-order (product) moments of the r.v. $\mathbf{X} \sim G_n^{\times}(\gamma, \boldsymbol{\sigma})$. We employ a special form of this expression later on in Section 4 to derive the formula for the Pearson index of linear correlation.

Theorem 3. *Let* $\mathbf{X} \sim G_n^{\times}(\gamma, \boldsymbol{\sigma})$, *then, for* $h_1, \ldots, h_n \in \mathbf{N}$, *we have*

$$\mathbf{E}\left[\prod_{i=1}^{n} X_i^{h_i}\right] = \frac{\Gamma(\gamma + \sum_{i=1}^n h_i)}{\Gamma(\gamma)\Gamma(\sum_{i=1}^n h_i + 1)} \prod_{i=1}^{n} \sigma_i^{h_i} \Gamma(h_i + 1). \tag{11}$$

We conclude the discussion in the present section by noticing that joint p.d.f. (8) can be used to estimate the parameters of the MMG distribution via the (numerical) maximum likelihood approach, whereas expression (11) is of interest if the moment-based estimation is being pursued.

3. Aggregation Properties of the Multiplicative Multivariate Gamma Distribution

One of the key paradigms in the modern enterprise risk management requires that all risks are treated on a holistic basis. As a result, risk aggregation is of fundamental importance for the effective conglomerate-wide risk management, risk-sensitive supervision, and a great variety of other business decision making processes. In the context of the MMG distribution, when all scale parameters are distinct, the decumulative distribution function of the aggregate risk r.v. is given by (10). The situation with arbitrary (possibly equal) scales is more involved. We further show that, within the MMG/G-MBRM class, that is for $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$ with arbitrary scale parameters and $\gamma \in (0,1)$, the d.d.f. of the aggregate risk r.v. X_+ admits a finite sum representation. To this end, we employ

the well-studied machinery of divided differences (e.g., Milne-Thomson 2000, for a comprehensive treatment). The rest of the section is divided into two: theoretical considerations and applications.

3.1. Theoretical Considerations

We remind at the outset that the divided differences, denoted by $\omega(y_1,\ldots,y_m)$, on a grid $\Delta=\{y_1,\ldots,y_m\}$ for a function $\omega:\mathbf{R}\to\mathbf{R}$ can be written as (e.g., Milne-Thomson 2000)

$$\omega(y_1,\ldots,y_m)=\sum_{i=1}^{m}\prod_{1\le j\ne i\le m}\frac{\omega(y_i)}{y_i-y_j}. \tag{12}$$

Denote

$$g(y)=\frac{\Gamma(\gamma,y)}{y\Gamma(\gamma)}\text{ for all }y\in\mathbf{R}_{0,+}. \tag{13}$$

Then the following corollary is merely a rearrangement of Equation (10).

Corollary 1 (of Proposition 1)**.** *The d.d.f. of the r.v. X_+ can be formulated, for distinct $\sigma_j, j=1,\ldots,n$, as*

$$\overline{F}_+(x)=\frac{(-1)^{n-1}x^n}{\prod_{i=1}^{n}\sigma_i}g(x/\sigma_1,\ldots,x/\sigma_n)\text{ for all }x\in\mathbf{R}_{0,+},$$

where $g(x_1/\sigma_1,\ldots,x/\sigma_n)$ is the divided differences representation of $g(\cdot)$ defined as per Equation (13).

Obviously, Equation (12) does not yield sensible results when some of the scale parameters of the r.v. $\mathbf{X}\sim Ga_n^{\times}(\gamma,\boldsymbol{\sigma})$ are equal. To circumvent this inconvenience, we formulate and prove the following lemma.

Lemma 2. *Consider $\omega:\mathbf{R}\to\mathbf{R}$, and the grid $\Delta=\{y_1,\ldots,y_m\}$ as before. For $n_i\in\mathbf{N}$, $i=1,\ldots,m$, assume ω is at least $k=\max_{i=1,\ldots,m}n_i-1$ times differentiable; then, we have*

$$\begin{aligned}&\omega(\underbrace{y_1,\ldots,y_1}_{n_1},\ldots,\underbrace{y_m,\ldots,y_m}_{n_m})\\&\quad=\sum_{i=1}^{m}\frac{1}{\Gamma(n_i)}\sum_{h_1+\ldots+h_m=n_i-1}\binom{n_i-1}{h_1,\ldots,h_m}\frac{\partial^{h_i}}{\partial y_i^{h_i}}w(y_i)\prod_{1\le j\ne i\le m}(-n_j)_{h_j}(y_i-y_j)^{-n_j-h_j},\end{aligned}$$

where $(p)_n:=p(p-1)\ldots(p-n+1)$ for $n\in\mathbf{N}$ denotes the falling factorial, $(p)_0:=1$.

The next assertion establishes the distribution of the aggregate risk r.v. with arbitrary scale parameters. Its proof follows by rearranging d.d.f. (10) using the divided differences operator and consequently evoking Lemma 2.

Theorem 4. *Consider $\mathbf{X}\sim Ga_n^{\times}(\gamma,\boldsymbol{\alpha})$, where $\gamma\in(0,1)$ and $\boldsymbol{\sigma}=(\sigma_1,\ldots,\sigma_n)'$ with arbitrary coordinates in the latter vector of parameters. Let $\boldsymbol{\sigma}=(\underbrace{\sigma_1,\ldots,\sigma_1}_{n_1},\ldots,\underbrace{\sigma_m,\ldots,\sigma_m}_{n_m})'$ for $m\in\mathbf{N}$ and $n_1+\cdots+n_m=n$, then, for $x\in\mathbf{R}_{0,+}$, the d.d.f. of X_+ admits the following finite sum form:*

$$\overline{F}_+(x)\;=\;\frac{1}{\prod_{i=1}^{m}\sigma_i^{n_i}}\sum_{i=1}^{m}\sum_{h_1+\ldots+h_m=n_i-1}g^*\left(\frac{x}{\sigma_i}\right)\sigma_i^{h_i+1}\prod_{1\le j\ne i\le m}\frac{(-n_j)_{h_j}}{\Gamma(h_j+1)}\left(\frac{1}{\sigma_j}-\frac{1}{\sigma_i}\right)^{-n_j-h_j},$$

where

$$g^*(y) = \sum_{k=0}^{h_i} \frac{1}{\Gamma(k+1)} y^k (-1)^k \psi_\Lambda^{(k)}(y).$$

Remark 3. *A close look at Theorem 4 reveals that the distribution of the r.v. X_+ can be considered a finite mixture of the r.v.'s distributed Erlang with stochastic scale parameters. To see this, first note that*

$$\begin{aligned} g^*(x/\sigma_i) &= \sum_{k=0}^{h_i} \frac{1}{\Gamma(k+1)} (x/\sigma_i)^k (-1)^k \psi_\Lambda^{(k)}(x/\sigma_i) \\ &= \mathbf{E}\left[\sum_{k=0}^{h_i} \frac{1}{\Gamma(k+1)} (x\Lambda/\sigma_i)^k e^{-\Lambda x/\sigma_i} \right] \\ &= \overline{F}_{\mathbf{e}_{i,h_i}}(x) \text{ for all } x \in \mathbf{R}_{0,+}, \end{aligned}$$

in which $\mathbf{e}_{i,h_i}$ denotes the r.v. distributed Erlang with the shape parameter $h_i + 1$ and the random scale parameter σ_i/Λ. Then rewrite $\overline{F}_+$ as

$$\begin{aligned} \overline{F}_+(x) &= \frac{1}{\prod_{i=1}^m \sigma_i^{n_i}} \sum_{i=1}^m \sum_{h_i=0}^{n_i-1} \left[\sigma_i^{h_i+1} \sum_{h_1+\ldots+h_m=n_i-1} \prod_{1\le j\ne i\le m} \frac{(-n_j)_{h_j}}{\Gamma(h_j+1)} \left(\frac{1}{\sigma_j} - \frac{1}{\sigma_i} \right)^{-n_j-h_j} \right] \overline{F}_{\mathbf{e}_{i,h_i}}(x) \\ &= \sum_{i=1}^m \sum_{h_i=0}^{n_i-1} p_{i,h_i} \overline{F}_{\mathbf{e}_{i,h_i}}(x), \end{aligned}$$

where

$$p_{i,h_i} = \frac{\sigma_i^{h_i+1}}{\prod_{i=1}^m \sigma_i^{n_i}} \left[\sum_{h_1+\ldots+h_m=n_i-1} \prod_{1\le j\ne i\le m} \frac{(-n_j)_{h_j}}{\Gamma(h_j+1)} \left(\frac{1}{\sigma_j} - \frac{1}{\sigma_i} \right)^{-n_j-h_j} \right].$$

By setting $x = 0$, it is clear that p_{i,h_i} are generalized weights in the sense that $\sum_{i=1}^m \sum_{h_i=0}^{n_i-1} p_{i,h_i} = 1$. However, these weights are not necessarily positive. For an example, consider the bivariate case with $n = m = 2$ and $n_1 = n_2 = 1$. A simple calculation yields

$$\overline{F}_+(x) = p_{1,0}\overline{F}_{\mathbf{e}_{1,0}}(x) + p_{2,0}\overline{F}_{\mathbf{e}_{2,0}}(x) \text{ for all } x \in \mathbf{R}_{0,+},$$

where $e_{i,0} \sim Ga(1, \sigma_i/\Lambda)$, $i = 1, 2$, $p_{1,0} = \sigma_2^{-1}(1/\sigma_2 - 1/\sigma_1)^{-1}$, and $p_{2,0} = \sigma_1^{-1}(1/\sigma_1 - 1/\sigma_2)^{-1}$. Therefore, depending on the values of σ_1 and σ_2, one of the weights must be negative.

3.2. Applications

Herein we confine the discussion to the individual and collective risk models. In this respect, recall that we call the r.v. $S_n = X_1 + \cdots + X_n$, $n \in \mathbf{N}$ the individual risk model, where we let the severity r.v.'s X_j, $j = 1, \ldots, n$ be possibly non-homogeneous. In the collective risk model case, for $N \in \mathbf{Z}_{0,+} := \{0, 1, 2, \ldots\}$ denoting the frequency r.v., we are interested in exploring the distribution of the random sum $S_N = X_1 + \cdots + X_N$. In the context of the MMG/G-MBRM, we have

$$S_n = \sigma_1 \frac{E_1}{\Lambda} + \cdots + \sigma_n \frac{E_n}{\Lambda}, \tag{14}$$

and

$$S_N = \sigma \frac{E_1}{\Lambda} + \cdots + \sigma \frac{E_N}{\Lambda}. \tag{15}$$

P.d.f.'s—rather than d.d.f.'s—often play an important role in the individual/collective risk model contexts. Therefore, the p.d.f.'s of S_n and S_N engage us in the rest of this section. We start with the p.d.f. of the former r.v. in the following proposition. Recall that p_{i,h_i} are given in (14), and $\psi^{(k)}$ denotes the k-th order derivative of the Laplace transform.

Proposition 2. *Let* $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$ *with* $\boldsymbol{\sigma} = (\underbrace{\sigma_1, \ldots, \sigma_1}_{n_1}, \ldots, \underbrace{\sigma_m, \ldots, \sigma_m}_{n_m})'$ *for* $m \in \mathbf{N}$ *and* $n_1 + \cdots + n_m = n$, *then, for* $x \in \mathbf{R}_+$, *the p.d.f. of the r.v.* S_n *is given by*

$$f_{S_n}(x) = \sum_{i=1}^{m} \sum_{h_i=0}^{n_i-1} p_{i,h_i} \frac{(-\sigma_i)^{-(h_i+1)} x^{h_i}}{\Gamma(h_i+1)} \psi_{\Lambda}^{(h_i+1)}(x/\sigma_i).$$

The next corollary follows immediately from Proposition 2, by setting $m = 1$ and $n_1 = n$.

Corollary 2. *Let* $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$, *where* $\gamma \in (0,\ 1)$ *and* $\sigma_1 = \cdots = \sigma_n \equiv \sigma \in \mathbf{R}_+$. *Then, for* $x \in \mathbf{R}_+$, *the p.d.f. of the r.v.* S_n *is given by*

$$f_{S_n}(x) = \sum_{i=0}^{n-1} \frac{1}{\Gamma(i+1)\Gamma(n-i)} \frac{\Gamma(i-\gamma+1)}{\Gamma(1-\gamma)\Gamma(\gamma)} \sigma^{-(n-i+\gamma-1)} e^{-x/\sigma} x^{n-i+\gamma-2}. \tag{16}$$

Remark 4. *It is not difficult to see that p.d.f. (16) admits the following finite mixture representation*

$$f_{S_n}(x) = \sum_{i=0}^{n-1} p_i f_{\mathbf{e}_i}(x),$$

where $\mathbf{e}_i \sim Ga(n-i+\gamma-1, \sigma)$ *and the weights are given by*

$$p_i = \frac{\Gamma(n-i+\gamma-1)}{\Gamma(i+1)\Gamma(n-i)} \frac{\Gamma(i-\gamma+1)}{\Gamma(1-\gamma)\Gamma(\gamma)}$$

for $i = 0, \ldots, n-1$. *Remarkably, in this special case, the weights* p_i *are 'proper' in the sense that* $p_i > 0$ *and* $\sum_{i=0}^{n-1} p_i = 1$. *This observation complements Theorem 6 in Sarabia et al. (2018).*

We further derive the p.d.f. of the r.v. S_N.

Proposition 3. *Let* $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$ *with* $\sigma_1 = \cdots = \sigma_n \equiv \sigma \in \mathbf{R}_+$, *then, for* $x \in \mathbf{R}_+$, *the p.d.f. of the r.v.* S_N *is*

$$f_{S_N}(x) \quad = \quad \begin{cases} \frac{x^{\gamma-1}e^{-x/\sigma}}{\Gamma(\gamma)\sigma^{\gamma}} \sum_{i=0}^{\infty} \sum_{m=0}^{\infty} \frac{\langle 1-\gamma \rangle_i}{i!m!} \left(\frac{x}{\sigma}\right)^m \mathbf{P}[N = m+i+1], & \text{for } x > 0, \\ \mathbf{P}[N=0], & \text{for } x = 0, \end{cases} \tag{17}$$

where $\langle p \rangle_n = p(p+1)\ldots(p+n-1)$ *for* $n \in \mathbf{N}$ *denotes the rising factorial,* $\langle p \rangle_0 := 1$.

We conclude this section by specializing the p.d.f. of the r.v. S_N reported in Proposition 3 for particular choices of the frequency r.v. In actuarial science, some popular choices of the r.v. N are, e.g., the Poisson, negative binomial, and logarithmic (e.g., Klugman et al. 2012). Below, we first remind the reader in passing the probability mass functions (p.m.f.'s) of the just-mentioned r.v.'s, and we then present the p.d.f.'s of the aggregate r.v.'s within the framework of the corresponding collective risk models.

- If $N \sim Poisson(\lambda)$ with $\lambda \in \mathbf{R}_+$, then the p.m.f. is given by

$$\mathbf{P}[N = n] = \frac{\lambda^n}{n!} e^{-\lambda} \text{ for all } n \in \mathbf{Z}_{0,+}.$$

- If $N \sim NB(\beta, p)$, the negative binomial distribution with $\beta \in \mathbf{R}_+$ and $p \in (0,1)$, then

$$\mathbf{P}[N = n] = \frac{\Gamma(n+\beta)}{\Gamma(n+1)\Gamma(\beta)} p^{\beta}(1-p)^n \text{ for all } n \in \mathbf{Z}_{0,+}.$$

- If $N \sim Logm(\theta)$, the logarithmic distribution with $\theta \in (0,1)$, then the p.m.f. is

$$\mathbf{P}[N = n] = \frac{-\theta^n}{n \log(1-\theta)} \text{ for all } n \in \mathbf{Z}_+.$$

Let Φ_1 and Φ_3, respectively, denote the two-variable confluent hypergeometric series of the first and third kind (see, e.g., Srivastava and Karlsson 1985), that is with $a_1, a_2, a_3 \in \mathbf{R}_+$,

$$\Phi_1(a_1, a_2; a_3; x, y) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\langle a_1 \rangle_{(i+j)} \langle a_2 \rangle_{(j)}}{\langle a_3 \rangle_{(i+j)} i! j!} x^i y^j,$$

for $x \in \mathbf{R}$, $|y| < 1$, and

$$\Phi_3(a_1; a_2; x, y) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\langle a_1 \rangle_{(j)}}{\langle a_2 \rangle_{(i+j)} i! j!} x^i y^j,$$

for $x, y \in \mathbf{R}$. The following corollary follows readily.

Corollary 3 (of Proposition 3)**.** *In the context of the collective risk model, we have, for all $x \in \mathbf{R}_+$, $\gamma \in (0,1)$ and $\sigma \in \mathbf{R}_+$, that*

- *If $N \sim Poisson(\lambda)$, then*

$$f_{S_N}(x) = \lambda \frac{x^{\gamma-1} e^{-x/\sigma-\lambda}}{\Gamma(\gamma)\sigma^{\gamma}} \Phi_3(1-\gamma, 2, x\lambda/\sigma, \lambda).$$

- *If $N \sim NB(\beta, p)$, then*

$$f_{S_N}(x) = \beta p^{\beta}(1-p) \frac{x^{\gamma-1} e^{-x/\sigma}}{\Gamma(\gamma)\sigma^{\gamma}} \Phi_1(1+\beta, 1-\gamma, 2, x(1-p)/\sigma, 1-p).$$

- *If $N \sim Logm(\theta)$, then*

$$f_{S_N}(x) = -\frac{\theta}{\log(1-\theta)} \frac{x^{\gamma-1} e^{-x/\sigma}}{\Gamma(\gamma)\sigma^{\gamma}} \Phi_1(1, 1-\gamma, 2, x\theta/\sigma, \theta).$$

4. Dependence Properties of the Multiplicative Multivariate Gamma Distribution

At first sight, the dependence structure that underlies the MMG distribution—that is d.d.f. (7)—is not as versatile as the one behind the additive counterpart of Mathai and Moschopoulos (1991). This is because the Pearson correlation, ρ, for the former class of distributions does not attain every value in the interval $[0,1]$, whereas it does so in the context of the latter class of distributions (e.g., Das et al. 2007; Su and Furman 2017a, 2017b, for a similar constraint in the context of default risk). More formally, we have the following proposition, the proof of which is a direct application of Theorem 3 and is thus omitted.

Proposition 4. *Let $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$, then the Pearson correlation between any pair of X_i and X_j, for $i \neq j \in \{1, \ldots, n\}$ is*

$$\rho[X_i, X_j] = (1-\gamma)/2, \tag{18}$$

where $\gamma \in (0,1)$. *In addition, we have* $\rho[X_i, X_j] \in (0, 1/2)$ *and it is a decreasing function of* $\gamma \in (0,1)$.

In the rest of this section, we show that the just-mentioned seeming shortcoming should in fact be attributed to the Pearson index of correlation, ρ, itself, rather than to the dependence structure of the MMG distribution. As hitherto, we divide our observations herein into two subsections.

4.1. Theoretical Considerations

At the outset, we observe that the dependence structure that underlies the MMG/G-MBRM is not linear in the—common—background r.v. Λ. Therefore, the machinery of copulas lands itself very naturally to exploring the relevant dependence properties. The next theorem states the copula function (e.g., Joe 1997) of $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$.

Theorem 5. *Assume that* $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$, *then the copula function underlying the d.d.f. of* $\mathbf{X}$ *is given, for* $\gamma \in (0,1)$, *by*

$$C_\gamma(u_1, \ldots, u_n) = \frac{1}{\Gamma(\gamma)} \Gamma\left(\gamma, \sum_{i=1}^{n} \Gamma^{-1}(\gamma,\ u_i \Gamma(\gamma))\right), \tag{19}$$

where $(u_1, \ldots, u_n)' \in [0,1]^n$, *and* $\Gamma^{-1}(\cdot, s)$ *denotes the inverse incomplete gamma function evaluated at* $s \in \mathbf{R}_+$. *Moreover, the p.d.f. associated with* C_γ *is given by*

$$c_\gamma(u_1, \ldots, u_n) \quad = \quad (-1)^n \psi_\Lambda^{(n)}\left(\sum_{i=1}^{n} \Gamma^{-1}(\gamma, u_i \Gamma(\gamma))\right) \prod_{i=1}^{n} \frac{1}{f(\Gamma^{-1}(\gamma, u_i \Gamma(\gamma)))}, \tag{20}$$

where f *denotes the p.d.f. of* $Ga(\gamma, 1)$, *and* $(-1)^n \psi^{(n)}$ *is obtained in Lemma 1.*

Figure 1 depicts the simulated scatter plots of the copula function C_γ for varying values of the γ parameter.

Remark 5. *Copula function (19) is a member of the encompassing class of the Archimedean copulas. Specifically, set*

$$\phi(s) = \frac{\Gamma(\gamma, s)}{\Gamma(\gamma)} \text{ for all } s \in \mathbf{R}_{0,+},$$

and observe that (19) admits the following form, for $(u_1, \ldots, u_n)' \in [0,1]^n$,

$$C_\phi(u_1, \ldots, u_n) = \phi(\phi^{-1}(u_1) + \cdots + \phi^{-1}(u_n)),$$

where $\phi : [0, \infty) \to [0,1]$ *is a legitimate completely monotonic function—known as the Archimedean generator—and* ϕ^{-1} *is its inverse (e.g., McNeil and Nešlehová 2009). The MMG copula therefore enriches the encompassing toolbox of the distinct Archimedean copulas available to researchers and practitioners.*

We have already mentioned at the end of Section 2 that the maximum likelihood approach can be used in order to numerically estimate the parameters of the MMG distribution. An alternative way to estimate the parameters is via the two-step copula approach. That is, we first fit the MMG copula to the pseudo uniform samples based on the empirical c.d.f.'s of X_i, $i = 1, \ldots, n$ and estimate the γ parameter (e.g., Embrechts and Hofert 2013; Genest et al. 2011, and references therein), and we then estimate the σ_i parameters based on the univariate marginal distributions assuming that the γ parameter is known. Given the cumbersome form of p.d.f. (8), the copula-based approach is computationally simpler.

Besides the just-mentioned statistical inference, a useful contribution of copulas to the vast literature of multivariate modelling is that they have given rise to a number of indices of dependence that circumvent the known fallacies of the Pearson ρ. Such indices of dependence are, e.g., the Kendall τ

and Spearman ρ_S measures of rank correlation, and we derive these two in the next subsection in the context of the MMG copula function C_γ.

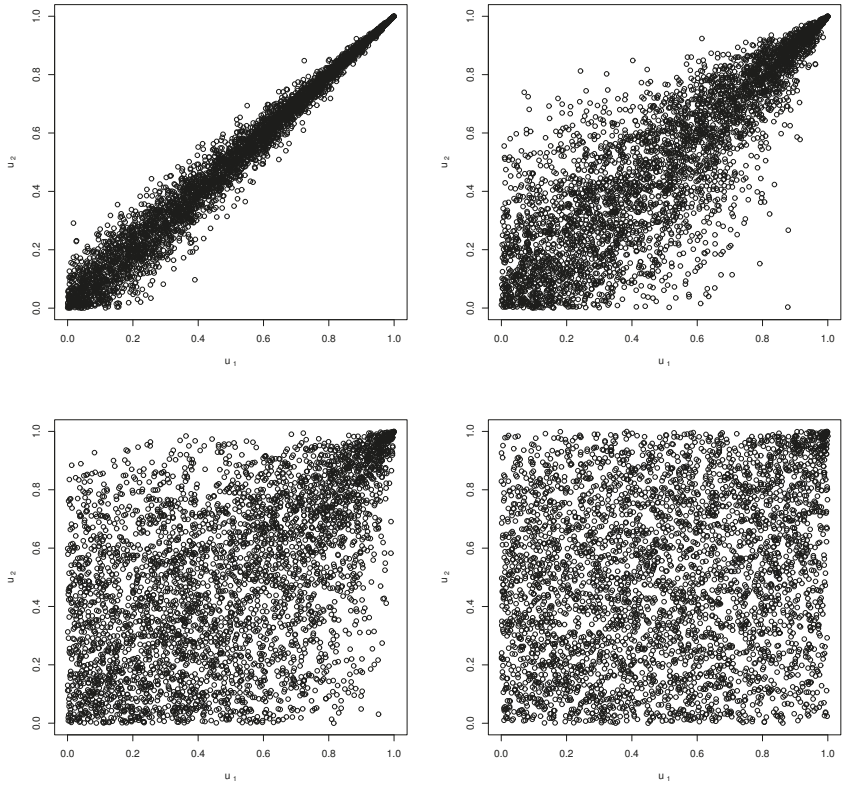


Figure 1. Scatter plots of the MMG copula (4000 simulation points) for varying values of the γ parameter: $\gamma = 0.05$ (**top left**), $\gamma = 0.2$ (**top right**), $\gamma = 0.5$ (**bottom left**), $\gamma = 0.8$ (**bottom right**).

In the rest of this subsection, we build up the theoretical groundwork necessary for exploring the tail dependence of C_γ. As tail dependence represents the co-movement of extreme risks, it is of particular importance in the era following the financial crisis of 2007–2009. We note in passing that, since the majority of the existing methods for quantifying tail dependence mainly aim at random pairs, we specialize the discussion in this part of the present report to the bivariate case only.

Let $\hat{C}$ denote the survival copula that corresponds to C, that is $\hat{C}(u_1, u_2) := u_1 + u_2 - 1 + C(1 - u_1, 1 - u_2)$, for $u_1, u_2 \in [0, 1]$. Then the first order lower and upper tail dependence parameters (e.g., Joe 1997) are given by

$$\lambda_L := \lim_{u \downarrow 0} \frac{C(u,u)}{u} \text{ and } \lambda_U := \lim_{u \downarrow 0} \frac{\hat{C}(u,u)}{u}, \tag{21}$$

whereas the second order tail dependence parameters (Coles et al. 1999) are given by

$$\chi_L := \lim_{u \downarrow 0} \frac{2 \log u}{\log C(u,u)} - 1 \text{ and } \chi_U := \lim_{u \downarrow 0} \frac{2 \log u}{\log \hat{C}(u,u)} - 1. \tag{22}$$

Recently, an argument has been put forward that tail dependence measures (21) and (22) may underestimate the extent of the tail dependence inherent in a copula. More specifically, Furman et al. (2015) claim and elucidate with numerous examples that as measures (21) and (22) are computed along the main diagonal (u, u), $u \in [0, 1]$, their values are not necessarily maximal when alternative paths in $[0, 1]^2$ are considered. This motivated the following definitions of the admissible paths and the paths of maximal dependence in ibid.

Definition 2. *A function $\varphi : [0, 1] \to [0, 1]$ is called admissible if it satisfies the following conditions:*

(C1) $\varphi(u) \in [u^2, 1]$ *for every* $u \in [0, 1]$; *and*
(C2) $\varphi(u)$ *and* $u^2/\varphi(u)$ *converge to* 0 *when* $u \downarrow 0$.

Then the path $(\varphi(u), u^2/\varphi(u))_{0\le u\le 1}$ is admissible whenever the function φ is admissible. In addition, we denote by $\mathcal{A}$ the set of all admissible functions φ.

Definition 3. *The path(s) $(\varphi(u), u^2/\varphi(u))_{0\le u\le 1}$ in $\mathcal{A}$ are called paths of maximal dependence if they maximize the probability*

$$\Pi_\varphi(u) = C\big(\varphi(u), u^2/\varphi(u)\big)$$

or, equivalently, the distance function

$$d_\varphi\left(C, C^\perp\right)(u) = C\big(\varphi(u), u^2/\varphi(u)\big) - C^\perp(\varphi(u), u^2/\varphi(u)),$$

where $C^\perp$ is the independence copula, i.e., $C^\perp(u_1, u_2) = u_1 u_2$ for all $0 \le u_1, u_2 \le 1$.

Obviously, the function $\varphi_0(u) = u$ is admissible and yields the representation of the diagonal path that serves as a building block for classical indices (21) and (22). For the Archimedean class of copulas, the following property of the maximal dependence path holds. The verification of the condition stated in Lemma 3 below is not trivial, and is carried out for the MMG copula C_γ in Theorem 6.

Lemma 3 (Furman et al. 2015)**.** *For an Archimedean copula with generator ϕ, if $x\frac{\partial}{\partial x}\phi^{-1}(x)$ is increasing on $x \in (0, 1)$, then the path of maximal dependence coincides with the main diagonal.*

The next lemma on a L'Hospital type rule for monotonicity, plays an importantly auxiliary role when deriving the maximal dependence path for C_γ.

Lemma 4 (Pinelis 2002)**.** *Let $-\infty \le a < b \le \infty$, also g_1 and g_2 be differentiable functions over the interval (a, b). Assume that $g_2'(s) < 0$ for $s \in (a, b)$, and $\lim_{s\downarrow a} g_1(s) = 0$ and $\lim_{s\downarrow a} g_2(s) = 0$. Then, g_1/g_2 is increasing on (a, b) if g_1'/g_2' is increasing.*

Our last result in this subsection implies that measures of tail dependence (21) and (22) are in fact maximal in the context of the MMG copula C_γ.

Theorem 6. *The maximal dependence path of the copula function C_γ in (19) is diagonal.*

4.2. Applications

The next assertion reports the Kendall tau and Spearman rho rank correlations, implied by the MMG copula (19). The hypergeometric function plays a pivotal role in deriving the Spearman rho correlation in the following proposition, and it is given in Gradshteyn and Ryzhik (2014)

$${}_{q+1}F_q(a_1, \ldots, a_{q+1}; b_1, \ldots, b_q; z) = \sum_{k=0}^{\infty} \frac{\langle a_1\rangle_k, \ldots, \langle a_{q+1}\rangle_k}{\langle b_1\rangle_k, \ldots, \langle b_q\rangle_k} \frac{z^k}{k!}. \tag{23}$$

For $a_1, \ldots, a_{q+1}$ all positive, and these are the cases of interest in the present report. The radius of convergence of the series is the open disk $|z| < 1$. On the boundary $|z| = 1$, the series converges absolutely if $d = b_1 + \cdots + b_q - a_1 - \cdots - a_{q+1} > 0$, and it converges except at $z = 1$ if $0 \geq d > -1$.

Proposition 5. *For the copula C_γ, the Kendall τ rank correlation is given by*

$$\tau(C_\gamma) = 1 - \frac{2\,\Gamma(\gamma + 1/2)}{\sqrt{\pi}\,\Gamma(\gamma)},$$

and the Spearman ρ_S rank correlation is given by

$$\rho_S(C_\gamma) = 6\left(\frac{8^{-\gamma}\Gamma(3\gamma)}{\Gamma(\gamma+1)\Gamma(2\gamma)}\,{}_2F_1(1, 3\gamma; 2\gamma+1; 1/2) - 1/2\right).$$

Figure 2 depicts the values for the Pearson ρ, Kendall τ and Spearman ρ_S indices of correlation with varying $\gamma \in (0, 1)$. The figure confirms that, while the Pearson ρ does not attain all values in $[0, 1]$ for the MMG/G-N goal.

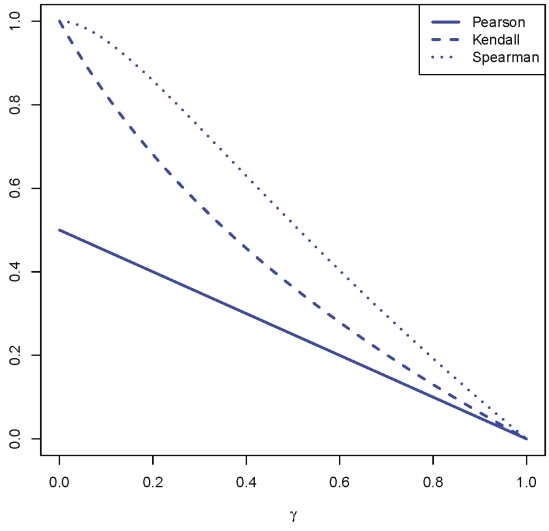


Figure 2. The plot of the Pearson rho, Kendall tau, and Spearman rho measures of correlation for varying values of $\gamma \in (0, 1)$.

Proposition 6. *Assume that $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$ has copula C_γ, the lower maximal tail dependence of C_γ is*

$$\lambda_L(C_\gamma) = \chi_L(C_\gamma) = 0.$$

The upper maximal tail dependence of C_γ is

$$\lambda_U(C_\gamma) = 2 - 2^\gamma, \text{ and } \chi_U(C_\gamma) = 1.$$

Proposition 6 readily implies—recall to this end that the copula C_γ is in fact a survival copula (by construction)—that the coordinates of $\mathbf{X} \sim Ga_n^{\times}(\gamma, \boldsymbol{\sigma})$ are asymptotically dependent in the lower tail, but independent in the upper tail. Speaking bluntly, this means that $\mathbf{X}$ is more likely to take smaller values simultaneously, but less likely to form a cluster of large values. This conforms to the already

made intuitive observation that the copula C_γ can serve as a reflected variant of the well-studied Clayton copula.

5. Conclusions

In the present report, we have systematically studied a class of multivariate multiplicative gamma distributions. We have demonstrated that the MMG distribution admits a very meaningful background risk model representation, where the interdependencies among risks are implied by a common systematic risk factor. Moreover, we have shown that the MMG distribution enjoys a remarkable level of analytical tractability, that is, the risk r.v.'s distributed MMG are straightforward to simulate, easy to aggregate and take maxima, closed under minima, and have attractive dependence and tail dependence characteristics. In view of the above, we think that the potential applications of the MMG distribution in actuarial science are vast, and we hope to draw the attention of the community to this class of distributions. In fact, reduced forms of the proposed MMG distribution have been recently heuristically adopted in the actuarial literature to model a variety of dependent insurance risks (e.g., Sarabia et al. 2018, also, Albrecher et al. 2011).

Author Contributions: All authors contributed equally to this work.

Funding: Vadim Semenikhine is grateful to the Natural Sciences and Engineering Research Council (NSERC) of Canada for its Undergraduate Summer Research Award. Edward Furman acknowledges the continuous support of his research by NSERC.

Acknowledgments: We are indebted to three anonymous referees for their very thorough reading of the paper, and the many suggestions that resulted in a clearer presentation.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proofs

Proof of Lemma 1. The proof of (i) is due to Equation 3.383(9) in Gradshteyn and Ryzhik (2014); (ii) follows readily via the integral representation of the Beta function. In order to check (iii), we have

$$\begin{aligned}(-1)^k \frac{d^k}{dx^k}\mathcal{L}\{f_\Lambda\}(x) &= \frac{1}{\Gamma(1-\gamma)\Gamma(\gamma)}\int_1^\infty \lambda^{k-1}e^{-\lambda x}(\lambda-1)^{-\gamma}d\lambda \\ &= \frac{1}{\Gamma(1-\gamma)\Gamma(\gamma)}\int_0^\infty (1+\lambda)^{k-1}e^{-(1+\lambda)x}\lambda^{-\gamma}d\lambda \\ &= \frac{1}{\Gamma(1-\gamma)\Gamma(\gamma)}\sum_{i=0}^{k-1}\binom{k-1}{i}e^{-x}\int_0^\infty e^{-\lambda x}\lambda^{i-\gamma}d\lambda.\end{aligned}$$

This completes the proof of the lemma. □

Proof of Theorem 1. The d.d.f.'s of the r.v.'s X and $\mathbf{X}$ follow immediately from Lemma 1, Statement (i) and Chapter 4 in Joe (1997). The joint p.d.f. follows from Lemma 1, Statement (iii) since

$$f(x_1,\ldots,x_n) = \frac{(-1)^n}{\prod_{i=1}^n \sigma_i}\psi_\Lambda^{(n)}\left(\sum_{i=1}^n \frac{x_i}{\sigma_i}\right) \text{ for all } (x_1,\ldots,x_n)\in\mathbf{R}_+^n.$$

This completes the proof of the theorem. □

Proof of Theorem 2. The closure under the minima operation is trivial by evoking Theorem 1, Statement (ii). The distribution of the r.v. $X_{\max}$ follows immediately (e.g., Corollary 2.2 in Su and Furman 2017a, for a similar result in the context of a multivariate Pareto distribution). This completes the proof of the theorem. □

Proof of Proposition 1. Recall (e.g., Akkouchi 2008) that, for the convolution of $\sigma_1 E_{\lambda,1}, \ldots, \sigma_n E_{\lambda,n}$ with $\sigma_i \neq \sigma_j,\ i \neq j \in \{1, \ldots, n\}$, we have

$$\overline{F}_+(x|\ \Lambda = \lambda) = \sum_{i=1}^{n} w_i \exp\{-x\lambda/\sigma_i\} \text{ for all } x \in \mathbf{R}_{0,+}.$$

Therefore we also have

$$\overline{F}_+(x) = \sum_{i=1}^{n} w_i \mathbf{E}\left[\exp\{-x\Lambda/\sigma_i\}\right] \text{ for all } x \in \mathbf{R}_{0,+},$$

and the assertion of the proposition follows evoking Lemma 1, Statement (i). □

Proof of Theorem 3. We immediately have

$$\begin{aligned}
\mathbf{E}\left[\prod_{i=1}^{n} X_i^{h_i}\right] &= \mathbf{E}\left[\mathbf{E}\left[\prod_{i=1}^{n} X_i^{h_i}\Big|\Lambda\right]\right] \\
&\overset{(1)}{=} \mathbf{E}\left[\prod_{i=1}^{n}\left(\frac{\sigma_i}{\Lambda}\right)^{h_i}\Gamma(h_i+1)\right] \\
&= \mathbf{E}\left[\Lambda^{-\sum_{i=1}^{n} h_i}\right]\prod_{i=1}^{n}\sigma_i^{h_i}\Gamma(h_i+1),
\end{aligned}$$

where '$\overset{(1)}{=}$' holds due to the moments' formula in the case of the exponentially distributed r.v.'s (see, e.g., Klugman et al. 2012). The proof is then completed by evoking Lemma 1, Statement (ii). □

Proof of Lemma 2. We start with Equation (6) of Kunz (1956) and have

$$\omega(\underbrace{y_1, \ldots, y_1}_{n_1}, \ldots, \underbrace{y_m, \ldots, y_m}_{n_m}) = \frac{1}{\prod_{i=1}^{m}\Gamma(n_i)}\frac{\partial^{n_1+\ldots+n_m-m}}{\prod_{i=1}^{m}\partial y_i^{n_i-1}}\sum_{i=1}^{m}\frac{\omega(y_i)}{\prod_{1\leq j\neq i\leq m}(y_i-y_j)}.$$

Then we differentiate term-by-term to obtain

$$\begin{aligned}
&\omega(\underbrace{y_1, \ldots, y_1}_{n_1}, \ldots, \underbrace{y_m, \ldots, y_m}_{n_m}) \\
= &\frac{1}{\prod_{i=1}^{m}\Gamma(n_i)}\sum_{i=1}^{m}\frac{\partial^{n_i-1}}{\partial y_i^{n_i-1}}\omega(y_i)\left(\frac{\partial^{\sum_{1\leq j\neq i\leq m} n_j-m+1}}{\prod_{1\leq j\neq i\leq m}\partial y_j^{n_j-1}}\prod_{1\leq j\neq i\leq m}(y_i-y_j)^{-1}\right) \\
= &\frac{1}{\prod_{i=1}^{m}\Gamma(n_i)}\sum_{i=1}^{m}\frac{\partial^{n_i-1}}{\partial y_i^{n_i-1}}\omega(y_i)\left(\prod_{1\leq j\neq i\leq m}\Gamma(n_j)(y_i-y_j)^{-n_j}\right) \\
= &\sum_{i=1}^{m}\frac{1}{\Gamma(n_i)}\frac{\partial^{n_i-1}}{\partial y_i^{n_i-1}}\omega(y_i)\prod_{1\leq j\neq i\leq m}(y_i-y_j)^{-n_j}.
\end{aligned}$$

Finally, we apply the Leibniz rule and readily have, for $i = 1, \ldots, m$,

$$\begin{aligned}
&\frac{\partial^{n_i-1}}{\partial y_i^{n_i-1}}\omega(y_i)\prod_{1\le j\neq i\le m}(y_i-y_j)^{-n_j}\\
=&\sum_{h_1+\ldots+h_m=n_i-1}\binom{n_i-1}{h_1,\ldots,h_m}\frac{\partial^{h_i}}{\partial y_i^{h_i}}w(y_i)\prod_{1\le j\neq i\le m}\frac{\partial^{h_j}}{\partial y_i^{h_j}}(y_i-y_j)^{-n_j}\\
=&\sum_{h_1+\ldots+h_m=n_i-1}\binom{n_i-1}{h_1,\ldots,h_m}\frac{\partial^{h_i}}{\partial y_i^{h_i}}w(y_i)\prod_{1\le j\neq i\le m}(-n_j)_{h_j}(y_i-y_j)^{-n_j-h_j}.
\end{aligned}$$

This concludes the proof of the lemma. □

Proof of Proposition 2. The proof of the proposition follows from Remark 3 that reports the mixture representation. Namely,

$$f_{S_n}(x) = \sum_{i=1}^{m}\sum_{h_i=0}^{n_i-1}p_{i,h_i}f_{\mathbf{e}_{i,h_i}}(x) = \sum_{i=1}^{m}\sum_{h_i=0}^{n_i-1}p_{i,h_i}\mathbf{E}\left[\frac{(\Lambda/\sigma_i)^{h_i+1}x^{h_i}}{\Gamma(h_i+1)}e^{-x\Lambda/\sigma_i}\right] \text{ for all } x\in\mathbf{R}_+.$$

This completes the proof. □

Proof of Proposition 3. We have the following string of equations, for all $x \in \mathbf{R}_+$,

$$\begin{aligned}
f_{S_N}(x) &= \sum_{n=1}^{\infty}f_{S_n}(x)\mathbf{P}[N=n]\\
&= \frac{x^{\gamma-1}e^{-x/\sigma}}{\Gamma(\gamma)\sigma^{\gamma}}\sum_{n=1}^{\infty}\sum_{i=0}^{n-1}\frac{\langle 1-\gamma\rangle_i}{i!(n-i-1)!}\left(\frac{x}{\sigma}\right)^{n-i-1}\mathbf{P}[N=n]\\
&= \frac{x^{\gamma-1}e^{-x/\sigma}}{\Gamma(\gamma)\sigma^{\gamma}}\sum_{i=0}^{\infty}\frac{\langle 1-\gamma\rangle_i}{i!}\sum_{n=i+1}^{\infty}\frac{(x/\sigma)^{n-i-1}}{(n-i-1)!}\mathbf{P}[N=n]\\
&= \frac{x^{\gamma-1}e^{-x/\sigma}}{\Gamma(\gamma)\sigma^{\gamma}}\sum_{i=0}^{\infty}\frac{\langle 1-\gamma\rangle_i}{i!}\sum_{m=0}^{\infty}\frac{(x/\sigma)^{m}}{m!}\mathbf{P}[N=m+i+1].
\end{aligned}$$

This completes the proof of the proposition. □

Proof of Theorem 5. The proof is a direct application of the Sklar's theorem. Namely, recall that

$$\mathbf{P}\left[X_i > x_i\right] = \frac{\Gamma(\gamma, x_i/\sigma_i)}{\Gamma(\gamma)} \text{ for all } x_i \in \mathbf{R}_{0,+},$$

and thus, for $i = 1, \ldots, n$, we have

$$X_i \overset{d}{=} \sigma_i\Gamma^{-1}(\gamma,\ U_i\Gamma(\gamma)),$$

where U_i denotes a r.v. distributed uniformly on $[0,1]$. Hence, the desired copula function is computed as

$$\begin{aligned}
C_\gamma(u_1,\ldots,u_n) &= \mathbf{P}[X_1 > \sigma_1\Gamma^{-1}(\gamma,u_1\Gamma(\gamma)),\ldots,X_n > \sigma_n\Gamma^{-1}(\gamma,u_n\Gamma(\gamma))]\\
&= \frac{1}{\Gamma(\gamma)}\Gamma\left(\gamma,\sum_{i=1}^{n}\Gamma^{-1}\left(\gamma,u_i\Gamma(\gamma)\right)\right),
\end{aligned}$$

where $(u_1,\ldots,u_n)' \in [0,1]^n$.

We next turn to study the p.d.f. of C_γ. By definition, we readily obtain

$$\begin{aligned} c_\gamma(u_1,\ldots,u_n) &= \frac{\partial^n}{\prod_{i=1}^n \partial u_i}\frac{1}{\Gamma(\gamma)}\Gamma\left(\gamma,\sum_{i=1}^n \Gamma^{-1}\left(\gamma,u_i\Gamma(\gamma)\right)\right) \\ &= \frac{1}{\Gamma(\gamma)}\Gamma^{(n)}\left(\gamma,\sum_{i=1}^n \Gamma^{-1}\left(\gamma,u_i\Gamma(\gamma)\right)\right)\prod_{i=1}^n \frac{\partial}{\partial u_i}\Gamma^{-1}(\gamma,u_i\Gamma(\gamma)) \\ &= \psi_\Lambda^{(n)}\left(\sum_{i=1}^n \Gamma^{-1}\left(\gamma,u_i\Gamma(\gamma)\right)\right)\prod_{i=1}^n \frac{\partial}{\partial s}\Gamma^{-1}(\gamma,s)\Big|_{s=u_i\Gamma(\gamma)}\Gamma(\gamma) \\ &= (-1)^n\psi_\Lambda^{(n)}\left(\sum_{i=1}^n \Gamma^{-1}\left(\gamma,u_i\Gamma(\gamma)\right)\right)\prod_{i=1}^n \frac{1}{f(\Gamma^{-1}(\gamma,u_i\Gamma(\gamma)))}, \end{aligned}$$

where f denotes the p.d.f. of $Ga(\gamma,1)$. This completes the proof of the theorem. □

Proof of Theorem 6. Let $\phi^{-1}(x)=\Gamma^{-1}(\gamma,x\Gamma(\gamma))$, for all $x\in(0,1)$, so

$$x\frac{\partial}{\partial x}\phi^{-1}(x)=\frac{x}{-f\left(\Gamma^{-1}(\gamma,x\Gamma(\gamma))\right)},$$

where $f(\cdot)$ is the p.d.f. of $Ga(\gamma,1)$. Note that, for $\gamma\in(0,1)$, which is exactly the case in the present report, $f(s)$ is decreasing for all $s\in\mathbf{R}_+$.

Now, set $g_1(x)=x$ and $g_2(x)=-f\left(\Gamma^{-1}(\gamma,x\Gamma(\gamma))\right)$. Clearly, $\lim_{x\downarrow 0}g_1(x)=0$, $\lim_{x\downarrow 0}g_2(x)=0$, and $g_2(x)$ is decreasing on $x\in(0,1)$. Moreover,

$$\frac{g_1'(x)}{g_2'(x)}=\frac{f(s)}{f'(s)}\Big|_{s=\Gamma^{-1}(\gamma,x\Gamma(\gamma))}=\frac{-1}{1+(1-\gamma)\left(\Gamma^{-1}(\gamma,x\Gamma(\gamma))\right)^{-1}}$$

is increasing on $x\in[0,1]$. Evoking Lemma 4, we conclude that $x\frac{\partial}{\partial x}\phi^{-1}(x)$ is increasing for $x\in(0,1)$. Finally, based on Lemma 3, the path of maximal dependence for C_γ is the diagonal, and the proof is completed. □

Proof of Proposition 5. Recall that copula function (19) is a special member of the Archimedean class of copulas having generator

$$\phi(s)=\frac{1}{\Gamma(\gamma)}\Gamma(\gamma,s)\ \text{ for all }\ s\in\mathbf{R}_{0,+}.$$

As a result, according to Theorem 4.3 in Joe (1997), we have

$$\begin{aligned} \tau(C_\gamma) &= 1-4\int_0^\infty s\left[\frac{\partial}{\partial s}\phi(s)\right]^2 ds \\ &= 1-\frac{4}{\Gamma(\gamma)^2}\int_0^\infty s^{2\gamma-1}e^{-2s}ds \\ &= 1-4^{1-\gamma}\frac{\Gamma(2\gamma)}{\Gamma(\gamma)^2}. \end{aligned} \tag{A1}$$

The expression for the Kendall τ is obtained by simplifying (A1).

We further proceed to the case of the Spearman ρ_S. For $i\neq j\in\{1,\ldots,n\}$, denote by f_i and f_j the marginal p.d.f.'s of the random pair $(X_i,X_j)'\subseteq\mathbf{X}$, then by definition (see, Section 2.1.9 in Joe 1997), we have

$$\rho_S(C_\gamma) = 12\int_0^\infty\int_0^\infty \overline{F}(x_i,x_j)f_i(x_i)f_j(x_j)dx_idx_j-3,$$

where

$$\begin{aligned}\int_0^\infty\int_0^\infty \overline{F}(x_i,x_j)f_i(x_i)f_j(x_j)dx_idx_j &= \frac{1}{\Gamma(\gamma)}\int_0^\infty\int_0^\infty \Gamma\left(\gamma,\frac{x_i}{\sigma_i}+\frac{x_j}{\sigma_j}\right)f_i(x_i)f_j(x_j)dx_idx_j\\ &= \frac{1}{\Gamma(\gamma)\Gamma(2\gamma)}\int_0^\infty \Gamma(\gamma,s)s^{2\gamma-1}e^{-s}ds\\ &\overset{(1)}{=} \frac{\Gamma(3\gamma)}{\Gamma(\gamma)\Gamma(2\gamma)}\frac{1}{2^{3\gamma+1}\gamma}{}_2F_1(1,3\gamma;2\gamma+1;1/2).\end{aligned}$$

Here, the equality '$\overset{(1)}{=}$' holds because of (6.455(1)) in Gradshteyn and Ryzhik (2014). This completes the proof of the proposition. □

Proof of Proposition 6. Let us first study the lower tail dependence of C_γ. The following string of equations holds:

$$\begin{aligned}\chi_L &= \lim_{u\downarrow 0}\frac{2\log\phi(\phi^{-1}(u))}{\log\phi(2\phi^{-1}(u))}-1\\ &= \lim_{t\to\infty}\frac{2\log\phi(t)}{\log\phi(2t)}-1\\ &= \lim_{t\to\infty}2\frac{-\log\Gamma(\gamma)+\log\Gamma(\gamma;t)}{-\log\Gamma(\gamma)+\log\Gamma(\gamma;2t)}-1.\end{aligned}$$

We know that, as $t\to\infty$, the following asymptotic expansion holds (Temme 1996):

$$\Gamma(\gamma;t)=t^{\gamma-1}e^{-t}(1+R(\gamma,t)),$$

with $R(\gamma,t)=\mathcal{O}(t^{-1})$. Then, we have

$$\begin{aligned}\lim_{t\to\infty}\frac{-\log\Gamma(\gamma)+\log\Gamma(\gamma;t)}{-\log\Gamma(\gamma)+\log\Gamma(\gamma;2t)} &= \lim_{t\to\infty}\frac{-\log\Gamma(\gamma)+(\gamma-1)\log t-t+\log((1+R(\gamma,t))}{-\log\Gamma(\gamma)+(\gamma-1)\log 2t-2t+\log((1+R(\gamma,2t))}\\ &= 1/2,\end{aligned}$$

and thus $\chi_L=0$, which automatically implies $\lambda_L=0$.

We now turn to study the upper tail dependence of C_γ. Note that the mixture r.v. Λ has d.d.f. $\overline{F}_\Lambda\in \mathrm{RV}_{-\gamma}$ that varies regularly at infinity with order $-\gamma$ (Bingham et al. 1987). The expressions for λ_U and χ_U are readily obtained by evoking Corollary 3.3 in Su and Hua (2017). This completes the proof of this proposition. □

References

Akkouchi, Mohamed. 2008. On the convolution of exponential distributions. *Journal of the Chungcheong Mathematical Society* 21: 501–10.

Albrecher, Hansjörg, Corina Constantinescu, and Stephane Loisel. 2011. Explicit ruin formulas for models with dependence among risks. *Insurance: Mathematics and Economics* 48: 265–70. [CrossRef]

Asimit, Alexandru V., Raluca Vernic, and Ricardas Zitikis. 2016. Background risk models and stepwise portfolio construction. *Methodology and Computing in Applied Probability* 18: 805–27. [CrossRef]

Avanzi, Benjamin, Greg Taylor, Phuong Anh Vu, and Bernard Wong. 2016. Stochastic loss reserving with dependence: A flexible multivariate Tweedie approach. *Insurance: Mathematics and Economics* 71: 63–78. [CrossRef]

Bahraoui, Zuhair, Catalina Bolancé, Elena Pelican, and Raluca Vernic. 2015. On the bivariate Sarmanov distribution and copula. An application on insurance data using truncated marginal distributions. *Statistics & Operations Research Transactions* 39: 209–30.

Balakrishnan, Narayanaswamy, and Miroslav M. Ristić. 2016. Multivariate families of gamma-generated distributions with finite or infinite support above or below the diagonal. *Journal of Multivariate Analysis* 143: 194–207. [CrossRef]

Bingham, Nicholas H., Charles M. Goldie, and Jef L. Teugels. 1987. *Regular Variation*. Cambridge: Cambridge University Press.

Coles, Stuart, Janet Heffernan, and Jonathan Tawn. 1999. Dependence measures for extreme value analyses. *Extremes* 2: 339–65. [CrossRef]

Cossette, Hélène, Etienne Marceau, Itre Mtalai, and Déry Veilleux. 2018. Dependent risk models with Archimedean copulas: A computational strategy based on common mixtures and applications. *Insurance: Mathematics and Economics* 78: 53–71. [CrossRef]

Das, Sanjiv R., Darrell Duffie, Nikunj Kapadia, and Leandro Saita. 2007. Common failings: How corporate defaults are correlated. *Journal of Finance* 62: 93–117. [CrossRef]

Dornheim, Harald, and Vytaras Brazauskas. 2007. Robust and efficient methods for credibility when claims are approximately gamma distributed. *North American Actuarial Journal* 11: 138–58. [CrossRef]

Embrechts, Paul, and Marius Hofert. 2013. Statistical inference for copulas in high dimensions: A simulation study. *ASTIN Bulletin* 43: 81–95. [CrossRef]

Feller, William. 1968. *An Introduction to Probability Theory and Its Applications*. New York: Wiley.

Frank, Günter, Harris Schlesinger, and Richard C. Stapleton. 2006. Multiplicative background risk. *Management Science* 52: 146–53. [CrossRef]

Furman, Edward, Daniel Hackmann, and Alexey Kuznetsov. 2018. *On Log-Normal Convolutions: An Analytical-Numerical Method with Applications to Economic Capital Determination*. Technical Report. Haifa: Actuarial Research Center—University of Haifa.

Furman, Edward, and Zinoviy Landsman. 2005. Risk capital decomposition for a multivariate dependent gamma portfolio. *Insurance: Mathematics and Economics* 37: 635–49. [CrossRef]

Furman, Edward, Jianxi Su, and Ričardas Zitikis. 2015. Paths and indices of maximal tail dependence. *ASTIN Bulletin* 45: 661–78. [CrossRef]

Genest, Christian, Johanna Nešlehová, and Johanna Ziegel. 2011. Inference in multivariate archimedean copula models. *Test* 20: 223–56. [CrossRef]

Gollier, Christian, and John W. Pratt. 1996. Risk vulnerability and the tempering effect of background risk. *Econometrica* 64: 1109–23. [CrossRef]

Gradshteyn, Izrail Solomonovich, and Iosif Moiseevich Ryzhik. 2014. *Table of Integrals, Series, and Products*, 8th ed. New York: Academic Press. [CrossRef]

Hürlimann, Werner. 2001. Analytical evaluation of economic risk capital for portfolios of gamma risks. *ASTIN Bulletin* 31: 107–22. [CrossRef]

Joe, Harry. 1997. *Multivariate Models and Dependence Concepts*. London: Chapman and Hall.

Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot. 2012. *Loss Models: From Data to Decisions*, 4th ed. Hoboken: Wiley.

Kotz, Samuel, Narayanaswamy Balakrishnan, and Norman L. Johnson. 2000. *Continuous Multivariate Distributions: Models and Applications*. New York: Wiley.

Kunz, Karls S. 1956. High accuracy quadrature formulas from divided differences with repeated arguments. *Mathematical Tables and Other Aids to Computation* 10: 87–90. [CrossRef]

Mathai, Arak M., and Panagis G. Moschopoulos. 1991. On a multivariate gamma. *Journal of Multivariate Analysis* 39: 135–53. [CrossRef]

McNeil, Alexander J., and Johanna Nešlehová. 2009. Multivariate Archimedean copulas, d-monotone functions and l1-norm symmetric distributions. *Annals of Statistics* 37: 3059–97. [CrossRef]

Milne-Thomson, Louis Melville. 2000. *The Calculus of Finite Differences*. Rhode Island: American Mathematical Society.

Moschopoulos, Peter G. 1985. The distribution of the sum of independent gamma random variables. *Annals of the Institute of Statistical Mathematics* 37: 541–44. [CrossRef]

Pinelis, Iosif. 2002. L'Hospital type rules for monotonicity: Applications to probability inequalities for sums of bounded random variables. *Journal of Inequalities in Pure and Applied Mathematics* 3: 7.

Provost, Serge. 1989. On sums of independent gamma eandom yariames. *Statistics* 20: 583–91. [CrossRef]

Sarabia, José María, Emilio Gómez-Déniz, Faustino Prieto, and Vanesa Jordá. 2018. Aggregation of dependent risks in mixtures of exponential distributions and extensions. *ASTIN Bulletin*. [CrossRef]

Srivastava, Hari M., and Per Wennerberg Karlsson. 1985. *Multiple Gaussian Hypergeometric Series*. Chichester: Ellis Horwood.

Su, Jianxi, and Edward Furman. 2017a. A form of multivariate Pareto distribution with applications to financial risk measurement. *ASTIN Bulletin* 47: 331–57. [CrossRef]

Su, Jianxi, and Edward Furman. 2017b. Multiple risk factor dependence structures: Distributional properties. *Insurance: Mathematics and Economics* 76: 56–68. [CrossRef]

Su, Jianxi, and Lei Hua. 2017. A general approach to full-range tail dependence copulas. *Insurance: Mathematics and Economics* 77: 49–64. [CrossRef]

Temme, Nico. 1996. *Special Functions: An Introduction to the Classical Functions of Mathematical Physics*. New York: Wiley.

Zhou, Ming, Jan Dhaene, and Jing Yao. 2018. An approximation method for risk aggregations and capital allocation rules based on additive risk factor models. *Insurance: Mathematics and Economics* 79: 92–100. [CrossRef]

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Risks Editorial Office
E-mail: risks@mdpi.com
www.mdpi.com/journal/risks

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34
Fax: +41 61 302 89 18

www.mdpi.com

ISBN 978-3-03928-517-4