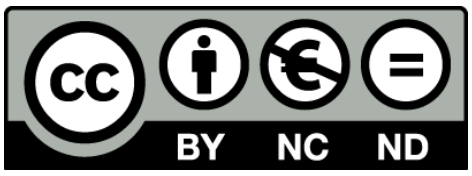
UNIVERSITAT DE BARCELONA

Quantitative risk assessment, aggregation functions and capital allocation problems

Jaume Belles Sampera



JAUME BELLES SAMPERA

QUANTITATIVE RISK ASSESSMENT, AGGREGATION FUNCTIONS AND CAPITAL ALLOCATION PROBLEMS

QUANTITATIVE RISK ASSESSMENT, AGGREGATION FUNCTIONS AND CAPITAL ALLOCATION PROBLEMS

JAUME BELLES SAMPERA

Research Group on Risk in Insurance and Finance

IREA - Institut de Recerca en Economia Aplicada Regional i Pública

Departament d'Econometria, Estadística i Economia Espanyola
Facultat d'Economia i Empresa
Universitat de Barcelona

May 2015

Jaume Belles Sampera: *Quantitative risk assessment, aggregation functions and capital allocation problems,* Riskcenter , May 2015.

Dissertation submitted in fulfillment of the requirements for the PhD in Business.

SUPERVISORS:
Dr. Miguel Ángel Santolino Prieto
Dr. Montserrat Guillén Estany

LOCATION:
Barcelona

Dedicated to Lidia and Marina.

AGRAÏMENTS

Aquesta tesi doctoral recull els resultats de la recerca acadèmica que he dut a terme els darrers quatre anys, però també és el fruit de tota la meva trajectòria professional, acadèmica i personal. Hi ha un munt de persones amb les qui he conviscut durant tots aquests anys i a qui estic molt agraït per diversos motius: perquè m'han instruït, recolzat i ajudat o, senzillament, perquè he tingut la sort de poder-les conèixer i compartir temps amb elles. Malgrat que no deixi constància en aquestes línies de tots els seus noms, les he tingut presents en molts moments.

Sóc incapaç d'imaginar com hagués estat aquesta tesi sense el guiatge d'en Miguel i de la Montse, els meus directors. Entre moltes de les seves virtuds en destacaria la passió que tots dos tenen per la recerca, la seva capacitat incansable de treball i la seva habilitat per fer entendor i senzill allò que a molts de nosaltres ens podria semblar intricat. La seva implicació en aquest projecte ha estat total des d'un bon inici, des del moment en el qual els vaig plantejar les idees que tenia més o menys embastades com a punt de partida. Immediatament em van ajudar a fixar objectius, a ampliar i enriquir el contingut, a centrar esforços i a esbossar un pla de treball que permetés, en algun moment dels quatre anys següents, donar per enllestida la meva tesi doctoral. I la seva feina ha estat tan encertada que aquest moment finalment ha arribat. Com en qualsevol procés de dissertació, hi ha hagut moments d'eufòria i moments de desencís: en Miguel i la Montse sempre m'han fer tocar de peus a terra en els primers i m'han animat i encoratjat en els segons. A més, tots dos han fet gala d'una inesgotable paciència, perquè he de reconèixer que sóc força tossut i, de vegades, empipadorament pessimista. Em podria estendre llargament sobre les coses que n'he après de tots dos, en àmbits purament acadèmics i de recerca però també en altres àmbits, com en la gestió de persones o en el treball en equip. Espero continuar gaudint de la seva companyia i aprenent d'ells tot allò que pugui: a part de l'evident satisfacció personal que comporta concloure una tesi doctoral, m'enduc la sensació d'haver fet dos nous amics d'aquells que ho són per tota la vida.

També vull donar les gràcies als meus companys al Riskcenter i a la Facultat d'Economia i Empresa de la Universitat de Barcelona. M'han envoltat d'un ambient de treball agradable i motivador, i m'han acollit amb els braços oberts des del primer moment. Manuela, Ana María, Helena, Óscar, Rosina, Tono, Maria,... trobaré a faltar les converses dels dinars (gairebé tant com els menjars que m'ha preparat

tantes vegades l'Emi, a qui li donc mil gràcies i una forta abraçada). Ramon, Merche i Cati, merci pel vostre suport, per poder comptar sempre amb els vostres consells. José María, moltes gràcies per introduir-me al món dels operadors d'agregació. Esther, em vas fer sentir molt a gust en el teu equip dedicat a la recerca en emprenedoria. Coloma i Dolors, un fort petó a totes dues.

Les meves experiències professionals en el sector financer i assegurador han contribuït decisivament en el resultat final d'aquest treball. Tinc molt bons records de diferents persones amb les que he tingut el gust de treballar a KPMG, als seus clients i, també, a Seguros Catalana Occidente, i els vull agrair a totes elles les vivències compartides.

No em vull oblidar de la meva colla d'amics. Fa molts anys que tinc la sort de compartir amb tots ells moments inoblidables, i vull aprofitar l'ocasió per agraïr-los que pugui formar part de cadascuna de les seves històries personals. A tots ells els duc ben aprop del cor.

La Lidia i la Marina omplen de comfort i d'amor la meva esfera més íntima. Són la meva llar i per mi és tot un orgull poder dir que són la meva dona i la meva filla. Tinc la impressió que tenim l'estrany privilegi d'haver-nos trobat i de poder i voler compartir les nostres vides. Senzillament, ens estimem. Moltíssimes gràcies Lidia per la teva paciència: ja sé que de vegades dius que no en tens gaire, però no totes les parelles poden dir que han superat la redacció d'una tesi doctoral!

Per acabar, un agraïment molt especial per tota la meva nombrosa i fantàstica família, particularment als meus pares Mario i Maria Dolors i als meus germans Màrius, Cristina, Damià i Lola. En aquests darrers quatre anys ens han passat un munt de coses, algunes de molt bones (la Muriel i la Marina en són els millors exemples) i altres de molt tristes. Sé que el pare estaria tan orgullós com tota la resta de poder veure impresa aquesta tesi. En nombroses ocasions al llarg de la meva vida he sentit que els *tetes* i els pares tenien una gran fe en les meves aptituds i capacitats. Sense el seu alè de suport hagués estat impossible arribar fins aquí.

ABSTRACT

This work is focused on the study of risk measures and solutions to capital allocation problems, their suitability to answer practical questions in the framework of insurance and financial institutions and their connection with a family of functions named aggregation operators. These operators are well-known among researchers from the information sciences or fuzzy sets and systems community.

The first contribution of this dissertation is the introduction of GlueVaR risk measures, a family belonging to the more general class of distortion risk measures. GlueVaR risk measures are simple to understand for risk managers in the financial and insurance sectors, because they are based on the most popular risk measures (VaR and TVaR) in both industries. For the same reason, they are almost as easy to compute as those common risk measures and, moreover, GlueVaR risk measures allow to capture more intricated managerial and regulatory attitudes towards risk.

The definition of the tail-subadditivity property for a pair of risks may be considered the second contribution. A distortion risk measure which satisfies this property has the ability to be subadditive in extremely adverse scenarios. In order to decide if a GlueVaR risk measure is a candidate to satisfy the tail-subadditivity property, conditions on its parameters are determined.

It is shown that distortion risk measures and several ordered weighted averaging operators in the discrete finite case are mathematically linked by means of the Choquet integral. It is shown that the overall aggregation preference of the expert may be measured by means of the local degree of orness of the distortion risk measure, which is a concept taken over from the information sciences community and brung into the quantitative risk management one.

New indicators for helping to characterize the discrete Choquet integral are also presented in this dissertation. The aim is complementing those already available, in order to be able to highlight particular features of this kind of aggregation function. Following this spirit, the degree of balance, the divergence, the variance indicator and Rényi entropies as indicators within the framework of the Choquet integral are here introduced.

A major contribution derived from the relationship between distortion risk measures and aggregation operators is the characterization of the risk attitude implicit into the choice of a distortion risk measure and a confidence or tolerance level. It is pointed out that the risk attitude implicit in a distortion risk measure is to some extent contained in its distortion function. In order to describe some relevant features of the distortion function, the degree of orness indicator and a quotient function are used. It is shown that these mathematical devices give insights on the implicit risk behavior involved in risk measures and entail the definitions of *overall*, *absolute* and *specific risk attitudes*.

Regarding capital allocation problems, a list of key elements to delimit these problems is provided and mainly two contributions are made. Firstly, it is shown that GlueVaR risk measures are as useful as other alternatives like VaR or TVaR to solve capital allocation problems. The second contribution is understanding capital allocation principles as compositional data. This interpretation of capital allocation principles allows the connection between aggregation operators and capital allocation problems, with an immediate practical application: Properly averaging several available solutions to the same capital allocation problem. This thesis contains some preliminary ideas on this connection, but it seems to be a promising research field.

RESUM

Aquest treball se centra en l'estudi de mesures de risc i de solucions a problemes d'assignació de capital, en la seva capacitat per respondre qüestions pràctiques en l'àmbit de les institucions asseguradores i financeres, i en la seva connexió amb una família de funcions anomenades operadors d'agregació. Aquests operadors són ben coneguts entre els investigadors de les comunitats de les ciències de la informació o dels conjunts i sistemes *fuzzy*.

La primera contribució d'aquesta tesi és la introducció de les mesures de risc GlueVaR, una família que pertany a la classe més general de les mesures de risc de distorsió. Les mesures de risc GlueVaR són senzilles d'entendre per als gestors de risc del sector financer i assegurador, perquè estan basades en les mesures de risc més populars (el VaR i el TVaR) en aquestes indústries. Per la mateixa raó, són quasi tan fàcils de calcular com aquestes mesures de risc més comuns però, a més, les mesures de risc GlueVaR permeten capturar actituds de gestió i regulatòries davant del risc més complicades.

La definició de la propietat de la subadditivitat en cues per a una parella de riscos es pot considerar la segona contribució. Una mesura de risc de distorsió que satisfà aquesta propietat té la capacitat de ser subadditiva en escenaris extremadament adversos. Per tal de decidir si una mesura de risc GlueVaR és candidata a satisfer la propietat de la subadditivitat en cues es determinen condicions sobre els seus paràmetres.

Es mostra que les mesures de risc de distorsió i diversos operadors de mitjanes ponderades ordenades en el cas finit i discret estan matemàticament relacionats a través de la integral de Choquet. Es mostra que la preferència global d'agregació de l'expert pot mesurar-se usant el nivell local d'*orness* de la mesura de risc de distorsió, que és un concepte traslladat de la comunitat de les ciències de la informació cap a la comunitat de la gestió quantitativa del risc.

Nous indicadors per a ajudar a caracteritzar les integrals de Choquet en el cas discret també es presenten en aquesta dissertació. Es pretén complementar-ne els ja existents, per tal de ser capaços de destacar característiques particulars d'aquest tipus de funcions d'agregació. Seguint aquest esperit, es presenten el nivell de balanç, la divergència, l'indicador de variança i les entropies de Rényi com a

indicadors en l'àmbit de la integral de Choquet.

Una contribució rellevant que es deriva de la relació entre les mesures de risc de distorsió i els operadors d'agregació és la caracterització de l'actitud davant del risc implícita en la tria d'una mesura de risc de distorsió i d'un nivell de confiança. S'assenyala que l'actitud davant del risc implícita en una mesura de risc de distorsió està continguda, fins a cert punt, en la seva funció de distorsió. Per tal de descriure alguns trets rellevants de la funció de distorsió s'usen l'indicador nivell d'*orness* i una funció quocient. Es mostra que aquests instruments matemàtics aporten informació relativa al comportament davant del risc implícit en les mesures de risc, i que d'ells se'n deriven les definicions de les *actituds davant del risc de tipus general*, *absolut* i *específic*.

Quant als problemes d'assignació de capital, es proporciona un llistat d'elements clau per a delimitar aquests problemes i es fan principalment dues contribucions. En primer lloc, es mostra que les mesures de risc GlueVaR són tan útils com altres alternatives com el VaR o el TVaR per resoldre problemes d'assignació de capital. La segona contribució consisteix en entendre els principis d'assignació de capital com a dades composicionals. Aquesta interpretació dels principis d'assignació de capital permet establir connexió entre els operadors d'agregació i els problemes d'assignació de capital, amb una aplicació pràctica immediata: calcular degudament la mitjana de diferents solucions disponibles per al mateix problema d'assignació de capital. Aquesta tesi conté algunes idees preliminars sobre aquesta connexió, però sembla un camp de recerca prometedor.

RESUMEN

Este trabajo se centra en el estudio de medidas de riesgo y de soluciones a problemas de asignación de capital, en su capacidad para responder cuestiones prácticas en el ámbito de las instituciones aseguradoras y financieras, y en su conexión con una familia de funciones denominadas operadores de agregación. Estos operadores son bien conocidos entre los investigadores de las comunidades de las ciencias de la información o de los conjuntos y sistemas *fuzzy*.

La primera contribución de esta tesis es la introducción de las medidas de riesgo GlueVaR, una familia que pertenece a la clase más general de las medidas de riesgo de distorsión. Las medidas de riesgo GlueVaR son sencillas de entender para los gestores de riesgo de los sectores financiero y asegurador, puesto que están basadas en las medidas de riesgo más populares (el VaR y el TVaR) de ambas industrias. Por el mismo motivo, son casi tan fáciles de calcular como estas medidas de riesgo más comunes pero, además, las medidas de riesgo GlueVaR permiten capturar actitudes de gestión y regulatorias ante el riesgo más complicadas.

La definición de la propiedad de la subadditividad en colas para un par de riesgos se puede considerar la segunda contribución. Una medida de riesgo de distorsión que cumple esta propiedad tiene la capacidad de ser subadditiva en escenarios extremadamente adversos. Con el propósito de decidir si una medida de riesgo GlueVaR es candidata a satisfacer la propiedad de la subadditividad en colas se determinan condiciones sobre sus parámetros.

Se muestra que las medidas de riesgo de distorsión y varios operadores de medias ponderadas ordenadas en el caso finito y discreto están matemáticamente relacionadas a través de la integral de Choquet. Se muestra que la preferencia global de agregación del experto puede medirse usando el nivel local de *orness* de la medida de riesgo de distorsión, que es un concepto trasladado des de la comunidad de las ciencias de la información hacia la comunidad de la gestión cuantitativa del riesgo.

Nuevos indicadores para ayudar a caracterizar las integrales de Choquet en el caso discreto también se presentan en esta disertación. Se pretende complementar a los existentes, con el fin de ser capaces de destacar características particulares de este tipo de funciones de agregación. Con este espíritu, se presentan el nivel de balance, la

divergencia, el indicador de varianza y las entropías de Rényi como indicadores en el ámbito de la integral de Choquet.

Una contribución relevante que se deriva de la relación entre las medidas de riesgo de distorsión y los operadores de agregación es la caracterización de la actitud ante el riesgo implícita en la elección de una medida de riesgo de distorsión y de un nivel de confianza. Se señala que la actitud ante el riesgo implícita en una medida de riesgo de distorsión está contenida, hasta cierto punto, en su función de distorsión. Para describir algunos rasgos relevantes de la función de distorsión se usan el indicador nivel de *orness* y una función cociente. Se muestra que estos instrumentos matemáticos aportan información relativa al comportamiento ante el riesgo implícito en las medidas de riesgo, y que de ellos se derivan las definiciones de les *actitudes ante el riego de tipo general*, *absoluto* y *específico*.

En cuanto a los problemas de asignación de capital, se proporciona un listado de elementos clave para delimitar estos problemas y se hacen principalmente dos contribuciones. En primer lugar, se muestra que las medidas de riesgo GlueVaR son tan útiles como otras alternativas tales como el VaR o el TVaR para resolver problemas de asignación de capital. La segunda contribución consiste en entender los principios de asignación de capital como datos composicionales. Esta interpretación de los principios de asignación de capital permite establecer conexión entre los operadores de agregación y los problemas de asignación de capital, con una aplicación práctica inmediata: calcular debidamente la media de diferentes soluciones disponibles para el mismo problema de asignación de capital. Esta tesis contiene algunas ideas preliminares sobre esta conexión, pero parece un campo de investigación prometedor.

PUBLICATIONS

Most ideas and figures in this dissertation have previously appeared in the following publications. Here is the list of chapters of this thesis to which they are related:

- Belles-Sampera, J., Guillén, M. and Santolino, M. [2014] , 'Beyond Value-at-Risk: GlueVaR distortion risk measures', *Risk Analysis* 34(1), 121–134. [Belles-Sampera et al., 2014a] — *Chapters 1 and 2*
- Belles-Sampera, J., Guillén, M. and Santolino, M. [2013], Generalizing some usual risk measures in financial and insurance applications, *in* M. Fernández-Izquierdo, M. Muñoz-Torres and R. León, eds, 'Modeling and Simulation in Engineering, Economics and Management. Proceedings of the MS 2013 International Conference', Vol. 145 of *Lecture Notes in Business Information Processing*, Springer-Verlag, pp. 75–82. [Belles-Sampera et al., 2013a] — *Chapter 2*
- Belles-Sampera, J., Guillén, M. and Santolino, M. [2014], 'The use of flexible quantile-based measures in risk assessment', *Communication in Statistics – Theory and Methods* (Accepted).[Belles-Sampera et al., 2014c] — *Chapters 2 and 3*
- Belles-Sampera, J., Merigó, J. M., Guillén, M. and Santolino, M. [2013],'The connection between distortion risk measures and ordered weighted averaging operators', *Insurance: Mathematics and Economics* 52(2), 411–420. [Belles-Sampera et al., 2013b] — *Chapter 4*
- Belles-Sampera, J., Merigó, J. M., Guillén, M. and Santolino, M. [2014], 'Indicators for the characterization of discrete Choquet integrals', *Information Sciences* 267, 201–216. [Belles-Sampera et al., 2014d] — *Chapter 5*
- Belles-Sampera, J., Merigó, J. M. and Santolino, M. [2013], Some new definitions of indicators for the Choquet integral, *in* H. Bustince, J. Fernández, T. Calvo and R. Mesiar, ed., 'Aggregation Functions in Theory and Practice. Proceedings of the 7th International Summer School on Aggregation Operators', Vol. 228 of *Advances in Intelligent Systems and Soft Computing*, Springer-Verlag, pp. 467–476. [Belles-Sampera et al., 2013d] — *Chapter 5*
- Belles-Sampera, J. and Santolino, M. [2013], Algunas reflexiones sobre los problemas de asignación de capital y la aplicación de ciertas medidas de riesgo, *in* E. Gómez Déniz, M. Guillén Estany and F. Vázquez Polo, eds, 'Investigaciones en Seguros y — *Chapters 7 and 8*

Gestión del Riesgo: Riesgo 2013', Vol. 194 of *Cuadernos de la Fundación*, Fundación MAPFRE, pp. 161–176. [Belles-Sampera and Santolino, 2013a]

Chapter 8

- Belles-Sampera, J., Guillén, M. and Santolino, M. [2014], 'Glue-VaR risk measures in capital allocation applications', *Insurance: Mathematics and Economics* 58, 132–137. [Belles-Sampera et al., 2014b]

In addition, the work done in the master thesis generated as an output another publication, which is also linked to some of the contents included in the present dissertation:

- Belles-Sampera, J. and Santolino, M. [2013], 'Asignación óptima de capital en base al perfil de riesgo de las instituciones de inversión colectiva: una aplicación de las medidas de riesgo distorsionadas', *Revista de Métodos Cuantitativos para la Economía y la Empresa* 15(2), 65–86. [Belles-Sampera and Santolino, 2013b]

I have been authorized by the coauthors of the previous publications, Dr. Montserrat Guillén, Dr. José M. Merigó and Dr. Miguel A. Santolino, to compile all those ideas and figures in the present work, and to be considered, to all effects, the author of this dissertation.

AWARDS

Chapters 1,2,3 and 8 were compiled in a monograph entitled *New risk measures for solvency purposes in insurance companies*, which received in October 2014 the 'Premi Ferran Armengol i Tubau' awarded by the Catalan Society of Economy.

Barcelona, May 2015

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Part I

INTRODUCTION

INTRODUCTION

BACKGROUND AND MOTIVATION

Since the appearance in 1988 of the first Basel Capital Accord for worldwide banking, quantitative risk management has become an essential instrument both in the financial and the insurance sectors. The Bank of International Settlements (BIS) was the leader of this agreement, which has been updated twice: once in year 2004 (Basel II agreement) and once more in 2010 after the worldwide spread of the last financial crisis (Basel III accord, see http://www.bis.org/list/bcbs/index.htm). The European Commission with the assistance of the European Insurance and Occupational Pensions Authority (EIOPA) has been developing, in parallel, the Solvency II regulatory framework for the insurance sector (see http://ec.europa.eu/internal_market/insurance/solvency/index_en.htm). The growing interest for quantitative risk management has possibly taken place because all the decision makers involved in the financial industry have endorsed it: from entities to local supervisory authorities, including audit, consultancy and advisory firms.

Quantitative risk management is one of the main concerns of management teams in the insurance and financial sectors. Among all the possible elements of interest that can be associated to this topic, it is my belief that risk measurement and the duality of risk aggregation and disaggregation are two key issues. This dissertation tries to shed some light on them.

This work is focused on the study of risk measures and solutions to capital allocation problems, and their suitability to answer practical questions in the framework of insurance and financial institutions.

The selection of appropriate risk measures and of solutions to capital allocation problems are issues under on-going discussion among researchers and practitioners. Insurance firms are subject to the capital requirements established by regulators' guidelines and directives. These requirements are typically equal to, or proportional to, a risk measure value that determines a minimum cushion of economic liquidity. The selection of such risk measures and tolerance levels is crucial therefore from the regulators' point of view.

Companies in the insurance sector prefer to minimize the level of capital reserves required by solvency regulations, because they must contend with many restrictions on how this capital can be invested and, as such, the return on their capital reserves is usually lower than that provided by other opportunities. For this reason, companies typically favor regulations that impose risk measures and tolerance levels that are not overly conservative. Managers also prefer simple, straightforward risk measures rather than more complicated alternatives, since they claim that the former are more easily communicated.

From the regulators' perspective, controlling the risk of insurance companies is fundamental in order to protect consumers and investors, which may have conflicting objectives. Strict solvency capital requirements may limit the capacity of firms, but they also reassure consumers and guarantee the position of the insurance industry in the economy. Thus, the debate as to what constitutes a suitable risk measure and what represents a suitable tolerance level is interminable, without apparently having been much investigation as to what might represent an appropriate compromise.

So with respect to risk measures, it seemed adequate to deepen in advantages and pitfalls of most commonly used risk measures in the actuarial and financial sectors, because the discussion could result attractive both to practitioners and supervisor authorities. This perspective allowed to list some of the additional proposals that can be found in the academic literature and, even, to devise some alternatives.

Capital allocation problems fall on the disaggregation side of risk management. These problems are associated to a wide variety of periodical management tasks inside the entities. In an insurance firm, for instance, risk capital allocation by business lines is a fundamental element for decision making from a risk management point of view. A sound implementation of capital allocation techniques may help insurance companies to improve their underwriting risk and to adjust the pricing of their policies, so to increase the value of the firm.

This piece of work is, to some extent, a theoretical study. Departing from an appropriate literature review, theoretical connections have been deduced. Nonetheless, the obtained theoretical results may be applied in practice. In order to support this statement, most of the theoretical developments have been illustrated with examples, which have been chosen as close to real practice as possible.

OUTLINE OF THE THESIS

This dissertation is structured in five parts. The first one is this introduction, while the fourth part is devoted to conclusions. The part number five contains the Appendix. The contributions of this thesis are exposed in the central parts, where they have been grouped as follows. The part number two consists of six chapters (from number 1 to 6) dedicated to investigate risk measures and their relationship with aggregation operators. The third part is devoted to provide an overview on capital allocation problems and to highlight how these problems may be related to aggregation functions. The last three chapters (from number 7 to 9) belong to this part. In what follows, some comments on the content of each chapter are presented, in order to get an overall picture of the work that has been done.

Chapter 1 - Beyond Value-at-Risk

A new family of risk measures, called GlueVaR, is defined within the class of distortion risk measures. The relationship between GlueVaR, Value-at-Risk (VaR) and Tail Value-at-Risk (TVaR) is explained. Analytical closed-form expressions are shown for the most frequently used distribution functions in financial and insurance applications, as well as first order Cornish-Fisher approximations for general skewed distribution functions. In addition, relationships between GlueVaR and Tail Distortion risk measures are shown to close this first chapter.

Chapter 2 - Tail-subadditivity for a pair of risks

This chapter is devoted to the definition of tail-subadditivity for a pair of risks given a confidence level, and to discuss the reasons for introducing this new property for risk measures in general. The tail-subadditivity for a pair of risks is, to some extent, a property of risk measures concerning their ability to indicate the benefits of diversification in extremely adverse situations. A proposition that states which GlueVaR risk measures are likely to meet that property is proven, and the chapter ends with a first discussion on the relationship between GlueVaR parameters and attitudes regarding risk measurement.

Chapter 3 - Risk measurement with GlueVaR

This chapter is devoted to illustrate different practical situations in which GlueVaR can be used. Two examples are presented. The same data set on insurance claim costs is used in both examples. One of them shows the values that different risk measures belonging to

the GlueVaR family give to a set of random variables. Different assumptions on the type of distribution function that characterize these random variables are considered. The second one illustrates the tail-subadditivity property and compares it with the subadditivity property.

Chapter 4 - Distortion risk measures and ordered averaging operators

Distortion risk measures summarize the risk of a loss distribution by means of a single value. In fuzzy systems, the ordered weighted averaging (OWA) and weighted ordered weighted averaging (WOWA) operators are used to aggregate a large number of fuzzy rules into a single value. In this chapter it is shown that these concepts can be derived from the Choquet integral, and then the mathematical relationship between distortion risk measures and the OWA and WOWA operators for discrete and finite random variables is presented. This connection offers a new interpretation of distortion risk measures and, in particular, Value-at-Risk and Tail Value-at-Risk can be understood from an aggregation operator perspective. The theoretical results are numerically illustrated and the local degree of orness concept is discussed.

Chapter 5 - Indicators for the discrete Choquet integrals

As it has already been mentioned in the previous chapter, ordered weighted averaging operators are powerful tools used in numerous decision-making problems and each OWA may be understood as a discrete Choquet integral. Aggregation operators are usually characterized by indicators. In this chapter four indicators usually associated with the OWA operator are extended to the discrete Choquet integral: namely, the degree of balance, the divergence, the variance indicator and Rényi entropies. All of these summarizing indicators are considered from a local and a global perspective. Linearity of indicators for linear combinations of capacities is investigated and, to illustrate the usefulness of results, indicators of the probabilistic ordered weighted averaging (POWA) operator are derived.

Chapter 6 - On the implicit risk attitude of a distortion risk measure

Understanding the risk attitude that is implicit in a risk assessment is crucial for decision makers. This chapter takes advantage of all the work made in the previous ones to characterize the underlying risk attitude involved in the choice of a risk measure, where it belongs to the family of distortion risk measures. A two-stage strategy is developed therein to reach this goal. First, it is shown that aggregation indicators defined for discrete Choquet integrals provide valuable information

related to the implicit risk attitude in aggregate terms, which leads to the definition of *overall risk attitude*. In the second stage a graphical analysis based on the distortion function is carried out to provide a local description of the underlying risk behavior. Here, the concepts of *absolute risk attitude* and *specific risk attitude* arise in a natural manner. The data set of insurance claim costs used in chapter 3 is used again to provide an example that shows the usefulness of this strategy in practice. This strategy is followed to solve a feasible problem that insurance companies under the Solvency II regulatory framework could face, and the impact of extreme observations is analyzed.

Chapter 7 - An overview on capital allocation problems

There is a strong relationship between risk measures and capital allocation problems. Briefly speaking, it may be assumed that each solution to a capital allocation problem is determined by a capital allocation criterion and a given risk measure. This chapter is intended to detect additional key elements involved in a solution of a capital allocation problem, in order to obtain a detailed initial picture on risk capital allocation proposals that can be found in the academic literature.
Personal notations and points of view are stated here and used from this point forward. Additionally, some particular solutions of interest are commented, trying to highlight both advantages and drawbacks of each one of them.

Chapter 8 - Contributions to capital allocation based on GlueVaR

This chapter is devoted to show how GlueVaR risk measures can be used for risk measurement purposes and to solve problems of proportional capital allocation through examples. The same data set is analyzed in both situations and, once more, it is the data set on insurance claim costs used in previous chapters. Regarding capital allocation, a first theoretical step is done for convenience: the capital allocation framework suggested by Dhaene et al. [2012b] is generalized to allow the application of the Value-at-Risk measure in combination with a stand-alone proportional allocation criterion (i.e., to accommodate the Haircut allocation principle). Afterwards, two new proportional capital allocation principles based on GlueVaR risk measures are defined an the example is presented, in which allocation solutions with tail-subadditive GlueVaR risk measures are discussed and compared with the solutions obtained when using the rest of alternatives.

Chapter 9 - Capital allocation principles as compositional data

In the last chapter, some connections between capital allocation problems and aggregation operators are emphasized. The approach is based on functions and operations defined in the standard simplex which, to best of my knowledge, remained an unexplored approach.

Appendix

Some of the largest or most intricate proofs have been gathered in the appendix to ease the reading.

PRELIMINARY DEFINITIONS

A few preliminary comments, notations and definitions should be helpful to keep this dissertation as self-contained as possible.

Definition 1 (Probability space)**.** *A probability space is defined by three elements* $(\Omega, \mathcal{A}, \mathcal{P})$*. The sample space* Ω *is a set of the possible events of a random experiment,* $\mathcal{A}$ *is a family of the set of all subsets of* Ω *(denoted as* $\mathcal{A} \in \wp(\Omega)$*) with a* σ*—algebra structure, and the probability* $\mathcal{P}$ *is a mapping from* $\mathcal{A}$ *to* $[0,1]$ *such that* $\mathcal{P}(\Omega) = 1$*,* $\mathcal{P}(\emptyset) = 0$ *and* $\mathcal{P}$ *satisfies the* $\sigma -$ additivity *property.*

A probability space is finite if the sample space is finite, i.e. $\Omega = \{\varpi_1, \varpi_2, ..., \varpi_n\}$. Then $\wp(\Omega)$ is the $\sigma - \text{algebra}$, which is denoted as 2^{Ω}. In the rest of this work, N instead of Ω and m instead of ϖ are used when referring to finite probability spaces. Hence, the notation is $\left(N, 2^N, \mathcal{P}\right)$, where $N = \{m_1, m_2, \ldots, m_n\}$.

Definition 2 (Random variable)**.** *Let* $(\Omega, \mathcal{A}, \mathcal{P})$ *be a probability space. A random variable* X *is a mapping from* Ω *to* $\mathbb{R}$ *such that* $X^{-1}((-\infty, x]) := \{\varpi \in \Omega : X(\varpi) \leqslant x\} \in \mathcal{A}$*,* $\forall x \in \mathbb{R}$*.*

A random variable X is discrete if $X(\Omega)$ is a finite set or a numerable set without cumulative points.

Definition 3 (Distribution function of a random variable)**.** *Let* X *be a random variable. The distribution function of* X*, denoted by* F_X*, is defined by* $F_X(x) := \mathcal{P}\left(X^{-1}((-\infty, x])\right) \equiv \mathcal{P}(X \leqslant x)$*.*

The distribution function F_X is non-decreasing, right-continuous and $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to +\infty} F_X(x) = 1$. The survival function of X, denoted by S_X, is defined by $S_X(x) := 1 - F_X(x)$, for all $x \in \mathbb{R}$. Note that the domain of the distribution function and the survival function is $\mathbb{R}$ even if X is a discrete random variable. In other words, F_X and S_X are defined for $X(\Omega) = \{x_1, x_2, ..., x_n, ...\}$ but also for any $x \in \mathbb{R}$.

Definition 4 (Risk measure)**.** *Let* Γ *be the set of all random variables defined for a given probability space* $(\Omega, \mathcal{A}, \mathcal{P})$*. A risk measure is a mapping* ρ *from* Γ *to* $\mathbb{R}$*, so* $\rho(X)$ *is a real value for each* $X \in \Gamma$*.*

Part II

ON RISK MEASURES AND AGGREGATION FUNCTIONS

1

BEYOND VALUE-AT-RISK

Value-at-Risk (VaR) has been adopted as a standard tool to assess the risk and to calculate capital requirements in the insurance industry. Value-at-Risk at level α is the α-quantile of a random variable X (which is often called loss), i.e.

$$\mathrm{VaR}_\alpha(X) = \inf\{x \mid F_X(x) \geqslant \alpha\} = F_X^{-1}(\alpha) \quad ,$$

where F_X is the cumulative distribution function (cdf) of X and α is the confidence or the tolerance level $0 \leqslant \alpha \leqslant 1$. However, VaR is known to present a number of pitfalls when applied in practice. A disadvantage when using VaR in the insurance or financial contexts is that the capital requirements for catastrophic losses based on this measure can be underestimated, i.e. the necessary reserves in adverse scenarios may well be less than they should be. The underestimation of capital requirements may be aggravated when fat-tailed losses are incorrectly modeled by mild-tailed distributions, such as the Normal distribution. There are attempts to overcome this kind of model risk when using VaR or, at least, to quantify the risk related to the modelling [Alexander and Sarabia, 2012]. But, in addition, a second drawback is that the VaR may fail the subadditivity property. A risk measure is subadditive when the aggregated risk is less than or equal to the sum of individual risks. Subadditvity is an appealing property when aggregating risks in order to preserve the benefits of diversification. VaR is subadditive for elliptically distributed losses [McNeil et al., 2005]. However, the subadditivity of VaR is not granted, as indicated in Artzner et al. [1999] and Acerbi and Tasche [2002].

Tail Value-at-Risk (TVaR) is defined as

$$\mathrm{TVaR}_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_\lambda(X)\, d\lambda.$$

Roughly speaking, the TVaR is understood as the mathematical expectation beyond VaR. The TVaR risk measure does not suffer the two drawbacks discussed above for VaR and, as such, would appear to be a more powerful measure for assessing the actual risks faced by insurance companies and financial institutions. However, TVaR has not been widely accepted by practitioners in the financial and insurance industry. VaR is currently the risk measure contemplated in the European solvency regulation for the insurance sector (Solvency II), and this is also the case of solvency regulation for the banking sector (Basel accords). The TVaR measures average losses in the most

adverse cases rather than just the minimum adverse loss, as the VaR does. Therefore, capital reserves based on the TVaR have to be considerably higher than those based on VaR and significant differences in the size of capital reserves can be obtained depending on which risk measure is adopted.

This chapter is motivated, therefore, by an attempt to respond to the following question. Can a risk measure be devised that would provide a risk assessment that lies somewhere between that offered by the VaR and the TVaR? To this end, a new family of risk measures (GlueVaR) is proposed, which forms part of a wider class referred to as distortion risk measures. The subadditivity properties of these GlueVaR risk measures are analyzed and it is shown that a subfamily of GlueVaR risk measures satisfies tail-subadditivity.

GlueVaR risk measures are defined by means of a four-parameter function. By calibrating the parameters, GlueVaR risk measures can be matched to a wide variety of contexts. Specifically, once a confidence level has been fixed, the new family contains risk measures that lie between those of VaR and TVaR and which may adequately reflect the risk of mild-tailed distributed losses without having to resort to VaR. In certain situations, however, more conservative risk measures even than TVaR may be preferred. It is shown that these highly conservative risk measures can also be defined by means of the GlueVaR family. Analytical closed-form expressions of GlueVaR for commonly used statistical distributions in the insurance context are derived. These closed-form expressions should enable practitioners to undertake an effortless transition from the use of VaR and TVaR to GlueVaR. First order Cornish-Fisher approximations to GlueVaR risk measures for general skewed distribution functions are also introduced in this chapter. Finally, the relationship between GlueVaR and Tail Distortion risk measures is shown.

1.1 OVERVIEW ON RISK MEASURES

Two main groups of axiom-based risk measures are *coherent risk measures*, as stated by Artzner et al. [1999], and *distortion risk measures*, as introduced by Wang [1996] and Wang et al. [1997]. Concavity of the distortion function is the key element to define risk measures that belong to both groups [Wang and Dhaene, 1998]. Suggestions on new desirable properties for distortion risk measures are proposed in Balbás et al. [2009], while generalizations of this kind of risk measures can be found, among others, in Hürlimann [2006] and Wu and Zhou [2006]. As shown in Goovaerts et al. [2012], it is possible to link distortion risk measures with other interesting families of risk measures

developed in the literature.

The axiomatic setting for risk measures has extensively been developed since seminal papers on coherent risk measures and distortion risk measures. Each set of axioms for risk measures corresponds to a particular behavior of decision makers under risk, as it has been shown, for instance, in Bleichrodt and Eeckhoudt [2006] and Denuit et al. [2006]. Most often, articles on axiom-based risk measurement present the link to a theoretical foundation of human behavior explicitly. For example, Wang [1996] shows the connection between distortion risk measures and Yaari's dual theory of choice under risk; Goovaerts et al. [2010b] investigate the additivity of risk measures in Quiggin's rank-dependent utility theory; and Kaluszka and Krzeszowiec [2012] introduce the generalized Choquet integral premium principle and relate it to Kahneman and Tversky's cumulative prospect theory.

Many articles have appeared in recent years that pay attention to risk measures based on distortion functions or on generalizations of the quantiles. An example of the first group is Zhu and Li [2012]. Bellini and Gianin [2012] and Bellini et al. [2014] fit to second group. An interplay between both groups is found in Dhaene et al. [2012a] and Goovaerts et al. [2012].

1.2 DISTORTION RISK MEASURES

Consider a probability space and the set of all random variables defined on this space. Any risk measure [see, for instance, Szëgo, 2002] ρ is a mapping from the set of random variables to the real line $\mathbb{R}$, $X \mapsto \rho(X) \in \mathbb{R}$. Distortion risk measures were introduced by Wang [Wang, 1995, 1996] and are closely related to the distortion expectation theory [Yaari, 1987]. A review on how risk measures can be interpreted from several perspectives is provided in Tsanakas and Desli [2005], and a clarifying explanation of the relationship between distortion risk measures and distortion expectation theory is included. A detailed literature review of distortion risk measures is available in [Denuit et al., 2005; Balbás et al., 2009]. There are two key elements to define a distortion risk measure: first, the associated distortion function; and, second, the concept of the Choquet [Choquet, 1954] Integral. The distortion function, Choquet Integral and the distortion risk measure can be defined as follows:

- **Distortion function.** Let $g : [0,1] \to [0,1]$ be a function such that $g(0) = 0$, $g(1) = 1$ and g is injective and non-decreasing. Then g is called a distortion function.

- **Choquet Integral** The (asymmetric) Choquet Integral with respect to a set function μ of a μ-measurable function $X : \Omega \to \overline{\mathbb{R}}$ is denoted as $\int X d\mu$ and is equal to

$$\int X d\mu = \int_{-\infty}^{0} \left[S_{\mu,X}(x) - \mu(\Omega) \right] dx + \int_{0}^{+\infty} S_{\mu,X}(x) dx,$$

  if $\mu(\Omega) < \infty$, where $S_{\mu,X}(x) = \mu(\{X > x\})$ denotes the *survival function* of X with respect to μ. Note that Ω denotes a set, which in financial and insurance applications is the sample space of a probability space. A set function μ in this context is a function defined from 2^{Ω} (the set of all subsets of Ω) to $\overline{\mathbb{R}}$. A μ-measurable function X is, widely speaking, a function defined on Ω such that expressions like $\mu(\{X > x\})$ or $\mu(\{X \leqslant x\})$ make sense. See Denneberg [1994] for more details.

- **Distortion risk measure.** Let g be a distortion function. Consider a random variable X and its survival function $S_X(x) = P(X > x)$. Function ρ_g defined by

$$\rho_g(X) = \int_{-\infty}^{0} \left[g(S_X(x)) - 1 \right] dx + \int_{0}^{+\infty} g(S_X(x)) dx$$

  is called a distortion risk measure.

From the previous definitions, it is straightforward to see that for any random variable X, $\rho_g(X)$ is the Choquet Integral of X with respect to the set function $\mu = g \circ P$, where P is the probability function associated with the probability space in which X is defined.

The mathematical expectation is a distortion risk measure whose distortion function is the identity function [Denuit et al., 2005], this is, $\rho_{id}(X) = \mathbb{E}(X)$. Therefore, a straightforward way to interpret a distortion risk measure is as follows: first, the survival function of the random variable is distorted $(g \circ S_X)$; second, the mathematical expectation of the distorted random variable is computed. From a theoretical point of view, note that this interpretation fits the discussion that risk may be defined as an expected value in many situations [Aven, 2012].

VaR and TVaR measures are in fact distortion risk measures. The associated distortion functions of these risk measures are shown in Table 1.1.

Based on the distortion functions shown in Table 1.1, once α is fixed it can be proved that $\text{VaR}_\alpha(X) \leqslant \text{TVaR}_\alpha(X)$ for any random variable X.

Table 1.1: VaR and TVaR distortion functions

Risk measure	Distortion function
VaR	$\psi_\alpha(u) = \begin{cases} 0 & \text{if} \quad 0 \leqslant u < 1-\alpha \\ 1 & \text{if} \quad 1-\alpha \leqslant u \leqslant 1 \end{cases}$
TVaR	$\gamma_\alpha(u) = \begin{cases} \dfrac{u}{1-\alpha} & \text{if} \quad 0 \leqslant u < 1-\alpha \\ 1 & \text{if} \quad 1-\alpha \leqslant u \leqslant 1 \end{cases}$

For a confidence level $\alpha \in (0,1)$.

Remark 1.2.1. *Let g and g^* be two distortion functions and let ρ_g and ρ_{g^*} be their respective distortion risk measures. Suppose that $g(u) \leqslant g^*(u)$ for all $u \in [0,1]$. Then $\rho_g(X) \leqslant \rho_{g^*}(X)$ for any random variable X.*
This result follows immediately from the definition of distortion risk measures, because

$$\rho_g(X) = \int_{-\infty}^{0} [g(S_X(x)) - 1]dx + \int_0^{+\infty} g(S_X(x))\,dx \leqslant$$
$$\int_{-\infty}^{0} [g^*(S_X(x)) - 1]dx + \int_0^{+\infty} g^*(S_X(x))\,dx = \rho_{g^*}(X).$$

1.3 A NEW FAMILY OF RISK MEASURES: GLUEVAR

A new family of risk measures, named GlueVaR, is here defined. Any GlueVaR risk measure can be described by means of its distortion function. Given a confidence level α, the distortion function for GlueVaR is:

$$\kappa_{\beta,\alpha}^{h_1,h_2}(u) = \begin{cases} \dfrac{h_1}{1-\beta} \cdot u, & \text{if} \quad 0 \leqslant u < 1-\beta \\ h_1 + \dfrac{h_2 - h_1}{\beta - \alpha} \cdot [u - (1-\beta)], & \\ \qquad \text{if} \quad 1-\beta \leqslant u < 1-\alpha & \\ 1, & \text{if} \quad 1-\alpha \leqslant u \leqslant 1 \end{cases} \tag{1.1}$$

where $\alpha, \beta \in [0,1]$ such that $\alpha \leqslant \beta$, $h_1 \in [0,1]$ and $h_2 \in [h_1, 1]$. Parameter β is the additional confidence level besides α. The shape of the GlueVaR distortion function is determined by the distorted survival probabilities h_1 and h_2 at levels $1-\beta$ and $1-\alpha$, respectively. We call parameters h_1 and h_2 the heights of the distortion function.

A wide range of risk measures may be defined under this framework. Note that VaR_α and TVaR_α are particular cases of this new family of risk measures. Namely, VaR_α and TVaR_α correspond to

distortion functions $\kappa^{0,0}_{\alpha,\alpha}$ and $\kappa^{1,1}_{\alpha,\alpha}$, respectively. By establishing suitable conditions on the heights h_1 and h_2, the GlueVaR family is very flexible. For example, risk managers might like to select α, β, h_1 and h_2 so that

$$\mathrm{VaR}_\alpha(X) \leqslant \mathrm{GlueVaR}^{h_1,h_2}_{\beta,\alpha}(X) \leqslant \mathrm{TVaR}_\alpha(X):$$

this can be achieved by selecting a set of parameters for their associated distortion functions to ensure that $\psi_\alpha(u) \leqslant \kappa^{h_1,h_2}_{\beta,\alpha}(u) \leqslant \gamma_\alpha(u)$ for any $u \in [0,1]$, following remark 1.2.1, i.e. by forcing condition $h_1 \leqslant \dfrac{1-\beta}{1-\alpha}$. An example of such a case is shown in Figure 1.1 (left-hand side).

The GlueVaR family also allows us to define a highly conservative risk measure $\mathrm{GlueVaR}^{h_1,h_2}_{\beta,\alpha}$, such that

$$\mathrm{TVaR}_\alpha(X) \leqslant \mathrm{GlueVaR}^{h_1,h_2}_{\beta,\alpha}(X) \leqslant \mathrm{TVaR}_\beta(X)$$

for any X and that the associated distortion function $\kappa^{h_1,h_2}_{\beta,\alpha}$ is concave in $[0,1]$. In this case, $\dfrac{1-\beta}{1-\alpha} \leqslant h_1$ and $h_2 = 1$ must be fulfilled, as occurs in the example shown in Figure 1.1 (right-hand side).

1.4 LINEAR COMBINATION OF RISK MEASURES

Given a random variable X and for fixed tolerance levels α and β so that $\alpha < \beta$, $\mathrm{GlueVaR}^{h_1,h_2}_{\beta,\alpha}(X)$ can be expressed as a linear combination of $\mathrm{TVaR}_\beta(X)$, $\mathrm{TVaR}_\alpha(X)$ and $\mathrm{VaR}_\alpha(X)$. This result allows us to translate the initial graphical-based construction of GlueVaR risk measures into an algebraic construction based on standard risk measures.

If the following notation is used,

$$\begin{cases} \omega_1 = & h_1 - \dfrac{(h_2-h_1)\cdot(1-\beta)}{\beta-\alpha} \\ \omega_2 = & \dfrac{h_2-h_1}{\beta-\alpha}\cdot(1-\alpha) \\ \omega_3 = & 1-\omega_1-\omega_2 \quad = 1-h_2, \end{cases} \tag{1.2}$$

then the distortion function $\kappa^{h_1,h_2}_{\beta,\alpha}(u)$ in (1.1) may be rewritten as (details can be found in the Appendix):

$$\kappa^{h_1,h_2}_{\beta,\alpha}(u) = \omega_1\cdot\gamma_\beta(u) + \omega_2\cdot\gamma_\alpha(u) + \omega_3\cdot\psi_\alpha(u) \tag{1.3}$$

where γ_β, γ_α, ψ_α are the distortion functions of TVaR at confidence levels β and α and of VaR at confidence level α, respectively (see Table 1.1). Therefore GlueVaR is a risk measure that can be expressed

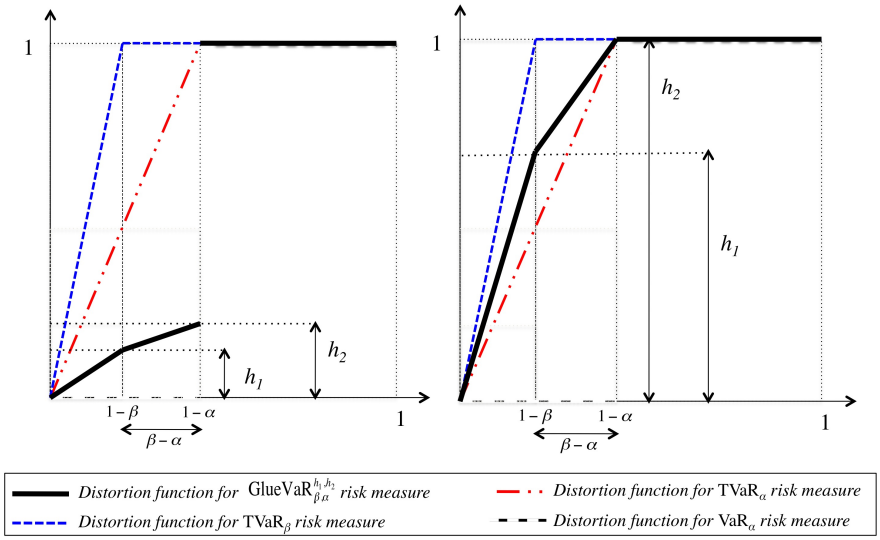


Figure 1.1: Examples of GlueVaR distortion functions.
Left. Distortion function is concave in $[0, 1-\alpha)$ and $\mathrm{VaR}_\alpha(X) \leqslant \mathrm{GlueVaR}^{h_1,h_2}_{\beta,\alpha}(X) \leqslant \mathrm{TVaR}_\alpha(X)$ for a random variable X;
Right. Distortion function is concave in the whole range $[0, 1]$ and $\mathrm{TVaR}_\alpha(X) \leqslant \mathrm{GlueVaR}^{h_1,h_2}_{\beta,\alpha}(X) \leqslant \mathrm{TVaR}_\beta(X)$ for a random variable X.

as a linear combination of three risk measures: TVaR at confidence levels β and α and VaR at confidence level α,

$$\begin{aligned} \mathrm{GlueVaR}_{\beta,\alpha}^{h_1,h_2}(X) = \omega_1 \cdot \mathrm{TVaR}_\beta(X) + \\ \omega_2 \cdot \mathrm{TVaR}_\alpha(X) + \omega_3 \cdot \mathrm{VaR}_\alpha(X). \end{aligned} \tag{1.4}$$

Given this relationship, some abuse of notation may be employed for $\mathrm{GlueVaR}_{\beta,\alpha}^{h_1,h_2}(X)$ and its related distortion function. The notation $\mathrm{GlueVaR}_{\beta,\alpha}^{\omega_1,\omega_2}(X)$ or $\kappa_{\beta,\alpha}^{\omega_1,\omega_2}(u)$ may, on occasions, be preferred to that based on heights h_1 and h_2. The bijective relationship between pairs (h_1, h_2) and (ω_1, ω_2) is also shown in the Appendix.

Specifically, in order to simplify the statement of Proposition 1.5.1, the expression of $\kappa_{\beta,\alpha}^{\omega_1,\omega_2}(u)$ is

$$\kappa_{\beta,\alpha}^{\omega_1,\omega_2}(u) = \begin{cases} \left[\dfrac{\omega_1}{1-\beta} + \dfrac{\omega_2}{1-\alpha}\right] \cdot u & \text{if} \quad 0 \leqslant u < 1-\beta \\ \omega_1 + \dfrac{\omega_2}{1-\alpha} \cdot u & \text{if} \quad 1-\beta \leqslant u < 1-\alpha \\ 1 & \text{if} \quad 1-\alpha \leqslant u \leqslant 1 \end{cases} \tag{1.5}$$

1.4.1 *Analytical closed-form expressions of GlueVaR*

A useful consequence of (1.4) is that when analytical closed-form expressions of $\mathrm{VaR}_\alpha(X)$ and $\mathrm{TVaR}_\alpha(X)$ are known for a random variable X, the closed-form expression of $\mathrm{GlueVaR}_{\beta,\alpha}^{h_1,h_2}(X)$ can automatically be derived without further complications. Otherwise, using the definition of GlueVaR as a distortion risk measure, the Choquet Integral of X with respect to the set function $\kappa_{\beta,\alpha}^{h_1,h_2} \circ P$ should be calculated.

1.4.1.1 *Illustration: GlueVaR expression for Student* t *distribution*

Let X be a random variable such that $\tilde{X} = \dfrac{X-\mu}{\sigma}$ is distributed as a Student t random variable with ν degrees of freedom (df). In such that case, X has μ mean and a standard deviation equal to $\sqrt{\frac{\nu \cdot \sigma^2}{\nu-2}}$. Then

$$\mathrm{VaR}_\alpha(X) = \mu + \sigma \cdot t_\alpha$$

$$\mathrm{TVaR}_\alpha(X) = \mu + \sigma \cdot \frac{\tau(t_\alpha)}{1-\alpha} \cdot \left(\frac{\nu + t_\alpha^2}{\nu - 1}\right),$$

where t_α is the α-quantile of a Student t distribution with ν df and τ is its density function.

Using (1.4) the GlueVaR of X random variable is

$$\begin{aligned}
\mathrm{GlueVaR}^{h_1,h_2}_{\beta,\alpha}(X) = \omega_1 \cdot \left[\mu + \sigma \cdot \frac{\tau(t_\beta)}{1-\beta} \cdot \left(\frac{\nu + t_\beta^2}{\nu - 1}\right)\right] + \\
\omega_2 \cdot \left[\mu + \sigma \cdot \frac{\tau(t_\alpha)}{1-\alpha} \cdot \left(\frac{\nu + t_\alpha^2}{\nu - 1}\right)\right] + \\
(1 - \omega_1 - \omega_2) \cdot (\mu + \sigma \cdot t_\alpha) = \\
= \mu + \sigma \cdot \left[\left(\frac{h_1}{1-\beta} - \frac{h_2 - h_1}{\beta - \alpha}\right) \cdot \tau(t_\beta) \cdot \left(\frac{\nu + t_\beta^2}{\nu - 1}\right) + \right. \\
\left. \frac{h_2 - h_1}{\beta - \alpha} \cdot \tau(t_\alpha) \cdot \left(\frac{\nu + t_\alpha^2}{\nu - 1}\right) + (1 - h_2) \cdot t_\alpha\right].
\end{aligned}$$

1.4.1.2 *Analytical expressions for other frequently used distributions*

Normal ($\mathcal{N}$), Lognormal ($\mathcal{LN}$) and Generalized Pareto ($\mathcal{GP}$) distributions have simple closed-form expressions of GlueVaR. Notation conventions are used. Namely, ϕ and Φ stand for the standard Normal pdf and cdf, respectively. The standard Normal distribution α and β quantiles are denoted as $q_\alpha = \Phi^{-1}(\alpha)$ and $q_\beta = \Phi^{-1}(\beta)$. For the $\mathcal{GP}$ distribution, the definition provided in Hosking and Wallis [1987] is considered, where the scale parameter is denoted by σ and k is the shape parameter. The $\mathcal{GP}$ distribution contains the Uniform ($k = 1$), the Exponential ($k = 0$), the Pareto ($k < 0$) and the type II Pareto ($k > 0$) distributions as special cases. Closed-form expressions of GlueVaR for several distributions are presented in Table **??**. Note that there are some exceptions to the general rule to deduce these closed-form expressions to be considered. When X follows a Pareto distribution with $k \leqslant 1$ and for any confidence level α, $\mathrm{TVaR}_\alpha(X) = +\infty$. But when $h_1 = 0$ $\mathrm{GlueVaR}^{h_1,h_2}_{\beta,\alpha}(X)$ is finite. There is a compensation effect between $\mathrm{TVaR}_\alpha(X)$ and $\mathrm{TVaR}_\beta(X)$. This is taken into account in Table **??**. This table is inspired by a similar one regarding VaR and TVaR that can be found in Sandström [2011].

Distribution	**GlueVaR$^{h_1,h_2}_{\beta,\alpha}$ expression**
Normal: $\mathcal{N}(\mu,\sigma^2)$	$\mu+\sigma\cdot q_\alpha\cdot(1-h_2)+\sigma\cdot\frac{h_2-h_1}{\beta-\alpha}\cdot[\phi(q_\alpha)-\phi(q_\beta)]+\sigma\cdot\frac{h_1}{1-\beta}\cdot\phi(q_\beta)$
Lognormal: $\mathcal{LN}(\mu,\sigma^2)$	$\exp(\mu+\sigma\cdot q_\alpha)\cdot(1-h_2)+\exp\left(\mu+\frac{\sigma^2}{2}\right)\cdot\frac{h_2-h_1}{\beta-\alpha}\cdot[\Phi(\sigma-q_\alpha)-\Phi(\sigma-q_\beta)]+\exp\left(\mu+\frac{\sigma^2}{2}\right)\cdot\frac{h_1}{1-\beta}\cdot\Phi(\sigma-q_\beta)$
Exponential: $\mathcal{GP}(k,\sigma)$, with $k=0$	$\sigma\cdot[h_2-\ln(1-\alpha)]+\sigma\cdot(1-\beta)\cdot\ln\left(\frac{1-\beta}{1-\alpha}\right)\cdot\left[\frac{h_2-h_1}{\beta-\alpha}-\frac{h_1}{1-\beta}\right]$
Pareto: $\mathcal{GP}(k,\sigma)$, with $k<0$	$\begin{cases} +\infty & \text{if} \quad k\leqslant -1, \quad h_1\neq 0 \\ \frac{\sigma}{k}\cdot\left[1-(1-\alpha)^k\right]+\frac{h_2-h_1}{\beta-\alpha}\cdot(1-\beta)\cdot\frac{\sigma}{k}\cdot\left[(1-\beta)^k-(1-\alpha)^k\right]+ & \\ +\frac{h_2-h_1}{\beta-\alpha}\cdot\frac{\sigma}{k+1}\cdot\left[(1-\alpha)^{k+1}-(1-\beta)^{k+1}\right] & \text{if} \quad k<-1, \quad h_1=0 \\ \sigma\cdot\left[\frac{1}{1-\alpha}-1\right]-\frac{h_2-h_1}{\beta-\alpha}\cdot(1-\beta)\cdot\sigma\cdot\left[\frac{1}{1-\beta}-\frac{1}{1-\alpha}\right]+ & \\ +\frac{h_2-h_1}{\beta-\alpha}\cdot\sigma\cdot\ln\left(\frac{1-\alpha}{1-\beta}\right) & \text{if} \quad k=-1, \quad h_1=0 \\ \frac{\sigma}{k}\cdot\left[1-(1-\alpha)^k\right]+\frac{\sigma}{k}\cdot\left(\frac{h_2-h_1}{\beta-\alpha}-\frac{h_1}{1-\beta}\right)\cdot\left[(1-\alpha)^k\cdot(1-\beta)\right]+ & \\ +\frac{h_2-h_1}{\beta-\alpha}\cdot\frac{\sigma}{k}\cdot\left[\frac{k\cdot(1-\alpha)^{k+1}}{k+1}\right]+\left(\frac{h_2-h_1}{\beta-\alpha}-\frac{h_1}{1-\beta}\right)\cdot\frac{\sigma}{k}\cdot\left[\frac{(1-\beta)^{k+1}}{k+1}\right] & \text{if} \quad k\in(-1,0) \end{cases}$
Type II Pareto: $\mathcal{GP}(k,\sigma)$, with $k>0$	$\frac{\sigma}{k}\cdot\left[1-(1-\alpha)^k\right]+\frac{\sigma}{k}\cdot\left(\frac{h_2-h_1}{\beta-\alpha}-\frac{h_1}{1-\beta}\right)\cdot\left[(1-\alpha)^k\cdot(1-\beta)\right]+$ $+\frac{h_2-h_1}{\beta-\alpha}\cdot\frac{\sigma}{k}\cdot\left[\frac{k\cdot(1-\alpha)^{k+1}}{k+1}\right]+\left(\frac{h_2-h_1}{\beta-\alpha}-\frac{h_1}{1-\beta}\right)\cdot\frac{\sigma}{k}\cdot\left[\frac{(1-\beta)^{k+1}}{k+1}\right]$

Table 1.2: Closed-form expressions of GlueVaR for some selected distributions

1.4.2 *The Cornish-Fisher approximation of GlueVaR*

Approximations to GlueVaR risk measures for general skewed distribution functions using a Cornish-Fisher expansion of their quantiles are provided in this section. In insurance applications managers often face to highly skewed random variables with right fat tails. In many of these situations, however, they do not know whether the underlying random variable of interest is distributed according to a known parametric distribution function. In those situations that the distribution is unknown, the value of the common quantile-based risk measures is routinely approximated by practitioners. It is shown that approximations of GlueVaR risk measures for general unknown skewed distribution functions can be straightforwardly obtained by means of the relationship of GlueVaR risk measures and the standard quantile-based risk measures.

The Cornish-Fisher expansion is widely used by practitioners to approximate the $\text{VaR}_\alpha(X)$ and $\text{TVaR}_\alpha(X)$ values when the random variable follows a skewed unknown distribution [see Cornish and Fisher, 1937; Fisher and Cornish, 1960; Johnson and Kotz, 1970; McCune and Gray, 1982]. The VaR and TVaR measure values can be approximated as $\text{VaR}_\alpha(X) \simeq \mu + q_{v,\alpha}\sigma$ and $\text{TVaR}_\alpha(X) \simeq \mu + q_{tv,\alpha}\sigma$, where $\mu = \mathbb{E}[X]$, $\sigma^2 = \mathbb{V}[X]$ and both $q_{v,\alpha}$ and $q_{tv,\alpha}$ are modified quantiles of the standard normal distribution that take into account the skewness of the distribution function of X.

Following Sandström [2007], the modified quantiles $q_{v,\alpha}$ and $q_{tv,\alpha}$ are computed as follows. Let us consider $\gamma = \mathbb{E}\left[(X-\mu)^3\right]/\sigma^3$ as a measure of the skewness of the distribution. If $q_\alpha = \Phi^{-1}(\alpha)$ and ϕ are the α-quantile and the density function of the standard normal distribution, respectively, then $q_{v,\alpha}$ and $q_{tv,\alpha}$ can be written as,

$$q_{v,\alpha} = \Phi^{-1}(\alpha) + \frac{\gamma}{6}\left[\left(\Phi^{-1}(\alpha)\right)^2 - 1\right] = q_\alpha + \frac{\gamma}{6}\left[q_\alpha^2 - 1\right],$$
$$q_{tv,\alpha} = \frac{\phi\left(\Phi^{-1}(\alpha)\right)}{1-\alpha}\left[1 + \frac{\gamma}{6}\left(\Phi^{-1}(\alpha)\right)^3\right] = \frac{\phi(q_\alpha)}{1-\alpha}\left[1 + \frac{\gamma}{6}q_\alpha^3\right].$$

Extensions of the Cornish-Fisher expansion that consider moments of higher order than γ have been provided in the literature [see, for instance, Giamouridis, 2006]. More details can be found in Appendix B of Sandström [2011].

According to the interpretation of GlueVaR measure as a linear combination of risk measures shown in (1.4), the approximation for the GlueVaR of X random variable following the Cornish-Fisher expansion can be obtained as

$$\text{GlueVaR}_{\beta,\alpha}^{h_1,h_2}(X) \simeq \mu + \sigma \left[\left(\frac{h_1}{1-\beta} - \frac{h_2 - h_1}{\beta - \alpha} \right) \phi(q_\beta) \left(1 + \frac{\gamma}{6} q_\beta^3 \right) + \left(\frac{h_2 - h_1}{\beta - \alpha} \right) \phi(q_\alpha) \left(1 + \frac{\gamma}{6} q_\alpha^3 \right) + (1 - h_2) \left(\frac{\gamma}{6} \left(q_\alpha^2 - 1 \right) + q_\alpha \right) \right].$$

The error of the approximation is upper bounded by the maximum error incurred when approximating $\text{VaR}_\alpha(X)$, $\text{TVaR}_\alpha(X)$ and $\text{TVaR}_\beta(X)$ using the equivalent Cornish-Fisher expansion for skewed distributions. This result is straightforwardly derived from the linear relationship shown in expression (1.4) and taking into account that weights ω_1, ω_2 and ω_3 are lower or equal than one, satisfying that $\omega_1 + \omega_2 + \omega_3 = 1$.

1.5 RELATIONSHIP BETWEEN GLUEVAR AND TAIL DISTORTION RISK MEASURES

As it has been aforementioned, different works that pay attention to risk measures based on distortion functions or based on several generalizations of quantiles have been appeared in recent years. See, for instance,Zhu and Li [2012]; Bellini and Gianin [2012]; Bellini et al. [2014]; Dhaene et al. [2012a] and Goovaerts et al. [2012].

Next paragraphs are devoted to reveal the connections between GlueVaR risk measures and Tail Distortion risk measures. To the best of my knowledge, Tail Distortion risk measures were introduced in Zhu and Li [2012]. Here the notation used for these family of risk measures is adapted from that in Lv et al. [2013]. Consider a distortion function g, this is, a non-decreasing and injective function g from $[0,1]$ to $[0,1]$ such that $g(0) = 0$ and $g(1) = 1$, and a confidence level $\alpha \in (0,1)$. The Tail Distortion Risk Measure $T_{g,\alpha}$ associated to g and α is defined as the distortion risk measure with distortion function g_α, where

$$g_\alpha(u) = g\left(\frac{u}{1-\alpha} \right) \cdot \mathbb{1}[0 \leqslant u < 1 - \alpha] + \mathbb{1}[1 - \alpha \leqslant u \leqslant 1].$$

In other words, if X is a random variable representing a loss in a probability space $(\Omega, \mathcal{A}, P)$ and its survival function is $S_X(x) = P(X > x)$, therefore

$$T_{g,\alpha}(X) = \int_{-\infty}^{0} \left[g_\alpha \left(S_X(x) \right) - 1 \right] dx + \int_0^{+\infty} g_\alpha \left(S_X(x) \right) dx \quad . \tag{1.6}$$

Note that g_α is continuous in $1 - \alpha$ or, alternatively, $g_\alpha(1 - \alpha) = 1$.

Proposition 1.5.1. *Consider a* $\text{GlueVaR}_{\beta,\alpha}^{\omega_1,\omega_2}$ *risk measure with parameters* α, β, ω_1 *and* ω_2*. This GlueVaR is equivalent to a Tail Distortion risk measure* $T_{g,\alpha}$ *if, and only if,* $\omega_2 = 1 - \omega_1$ *and*

$$g(t) = \left(\frac{\omega_1 \cdot (1-\alpha)}{1-\beta} + 1 - \omega_1\right) \cdot t \cdot \mathbb{1}\left[0 \leqslant t < (1-\alpha)^{-1} \cdot (1-\beta)\right] + (\omega_1 + (1-\omega_1) \cdot t) \cdot \mathbb{1}\left[(1-\alpha)^{-1} \cdot (1-\beta) \leqslant t \leqslant 1\right]. \tag{1.7}$$

The proof is provided in the Appendix.

It is worth noting that only GlueVaR risk measures with $\omega_3 = 0$ can be represented as Tail Distortion risk measures, because $\omega_1 + \omega_2 + \omega_3 = 1$ must hold as part of the definition of a $\text{GlueVaR}_{\beta,\alpha}^{\omega_1,\omega_2}$ risk measure. In other words, one can only represent as Tail Distortion risk measures those GlueVaR that do not give weight to VaR_α.

The origin of GlueVaR risk measures is in my master's thesis Belles-Sampera [2011]. As a curiosity, the definition of a parametric family of risk measures named PUp-TVaR can also be found therein, which are exactly the Tail Distortion risk measures linked to Proportional Hazards Distortion functions $g(u) = u^{\frac{1}{a}}, a \geqslant 1$ from the perspective of Zhu and Li [2012].

2

TAIL-SUBADDITIVITY FOR A PAIR OF RISKS

2.1 SUBADDITIVITY IN THE TAIL

This chapter is devoted to an analysis of the properties of the GlueVaR family of risk measures, with special attention to subadditivity. The main reason for defining these GlueVaR risk measures is a response to the concerns expressed by risk managers regarding the choice of risk measures in the case of regulatory capital requirements. However, an axiomatic approach to define or represent risk measures is more frequent in the literature [Artzner et al., 1999; Föllmer and Schied, 2002; Frittelli and Rosazza Gianin, 2002; Denuit et al., 2006; Song and Yan, 2009; Cerreia-Vioglio et al., 2011; Ekeland et al., 2012; Goovaerts et al., 2012; Grechuk et al., 2012].

In a seminal article [Artzner et al., 1999] the following set of axioms that a risk measure should satisfy was established: positive homogeneity, translation invariance, monotonicity and subadditivity. Authors referred to such risk measures as *coherent risk measures*. Distortion risk measures always satisfy the first three properties, but subadditivity is only guaranteed when the distortion function is concave [Denneberg, 1994; Wang and Dhaene, 1998; Wirch and Hardy, 2002]. Therefore, VaR, unlike TVaR, is not coherent. In some situations, coherence of risk measures is a requirement [Cox, 2012] but, nonetheless, some criticisms can be found [Dhaene et al., 2008]. Additional properties for distortion risk measures are provided in [Jiang, 2008; Balbás et al., 2009].

As shown in the previous chapter, GlueVaR risk measures may be interpreted as a linear combination of VaR and TVaR risk measures. Therefore, a GlueVaR risk measure is coherent when the weight assigned to VaR is zero and the weights of TVaR are non-negative. In terms of the parameters of the distortion function, GlueVaR is subadditive (and thus coherent) if $h_2 = 1$ and $\dfrac{1-\beta}{1-\alpha} \leqslant h_1$. More generally, any property satisfied by TVaR but not by VaR will be inherited by GlueVaR if $\omega_1 \geqslant 0$ and $\omega_3 = 0$ in expression (1.2).

Subaddtitivity in the whole domain is a strong condition. When dealing with fat tail losses (i.e. low-frequency and large-loss events), risk managers are especially interested in the tail region. Fat right-tails have been extensively studied in insurance and finance [Wang, 1998; Embrechts et al., 2009a,b; Degen et al., 2010; Nam et al., 2011;

Chen et al., 2012] and the behavior of aggregate risks in the tail region has received huge attention by researchers in last years [Cheung, 2009; Song and Yan, 2009; Hua and Joe, 2012]. To the best of my knowledge, however, previous studies of the subadditivity of risk measures in the tail region are scarce [Danielsson et al., 2005; Hua and Joe, 2012]. The milder condition of subadditivity in the tail region is investigated here.

The concept of subadditivity in the right tail for a pair of risks is introduced. Note that if interested in the left -as opposed to the right- tail, a simple change of sign in the random variable suffices. Subadditivity in the right tail is defined in this discussion for distortion risk measures. Consider a probability space with sample space Ω. Let $s_\alpha(Z)$ the α-quantile of random variable Z, $s_\alpha(Z) = \inf\{z \mid S_Z(z) \leqslant 1-\alpha\}$. Let $\Omega_{\alpha,Z}$ be defined by

$$\Omega_{\alpha,Z} := \{\omega \mid Z(\omega) > s_\alpha(Z)\} \subseteq \Omega,$$

so $\Omega_{\alpha,Z}$ means here the tail region of random variable Z given a confidence level α. Let X, Y be two risks defined on the same probability space. When aggregating two risks, the common tail for both risks must be taken into account. This common tail region is defined here as follows: $\Omega_{\alpha,X,Y} := \Omega_{\alpha,X} \cap \Omega_{\alpha,Y} \cap \Omega_{\alpha,X+Y}$.

Definition 2.1.1. *Given a confidence level $\alpha \in [0,1]$, a distortion risk measure ρ_g is subadditive in the tail for the pair X, Y if $\Omega_{\alpha,X,Y} \neq \emptyset$ and*

$$\int_{\Omega_{\alpha,X,Y}} (X+Y)\, d(g \circ P) \leqslant \int_{\Omega_{\alpha,X,Y}} X d(g \circ P) + \int_{\Omega_{\alpha,X,Y}} Y d(g \circ P),$$

where the integral symbol stands for Choquet Integrals with respect to the set function $g \circ P$.

When there is no ambiguity as to which confidence level α and random variables X, Y are taken into account, *tail-subadditivity* is used to refer to this property. If notation $m_{\alpha,Z} = \inf\{z \in Z(\Omega_{\alpha,X,Y})\}$, is introduced, the integral condition used in the definition can be rewritten, in terms of survival functions, as

$$\begin{aligned}
&\int_{\inf\{0,m_{\alpha,X+Y}\}}^{0} [g(S_{X+Y}(z)) - 1]\, dz + \\
\int_{\sup\{0,m_{\alpha,X+Y}\}}^{+\infty} g(S_{X+Y}(z))\, dz \leqslant &\int_{\inf\{0,m_{\alpha,X}\}}^{0} [g(S_X(x)) - 1]\, dx + \\
\int_{\sup\{0,m_{\alpha,X}\}}^{+\infty} g(S_X(x))\, dx + &\int_{\inf\{0,m_{\alpha,Y}\}}^{0} [g(S_Y(y)) - 1]\, dy + \\
&\int_{\sup\{0,m_{\alpha,Y}\}}^{+\infty} g(S_Y(y))\, dy.
\end{aligned}$$

Theorem 2.1.1. *Given a confidence level α and a pair of risks X and Y so that $\mathcal{Q}_{\alpha,X,Y} \neq \emptyset$, a GlueVaR risk measure is tail-subadditive if its associated distortion function $\kappa_{\beta,\alpha}^{h_1,h_2}$ is concave in $[0, 1-\alpha)$.*

The proof is contained in the Appendix.

Tail-subadditivity is a desirable property, because it implies that the benefits of diversification may not be valid in every situation but, at least, they hold in extreme cases.

Note that, in terms of parameters h_1 and h_2, a GlueVaR risk measure may be tail-subadditive if, and only if, $h_2 \leqslant h_1 \cdot \dfrac{1-\alpha}{1-\beta}$, as a corollary of Theorem 2.1.1.

2.2 RISK ATTITUDES IN GLUEVAR

An interesting interpretation in the context of decision making and risk management is that GlueVaR risk measures arise as a linear combination of three possible scenarios. So, two levels of severity can be fixed, namely α and β, with $\alpha < \beta$. Then, the risk can be measured in the highly conservative scenario with TVaR at level β; in the conservative scenario with TVaR at level α; and in the less conservative scenario with VaR at level α.

Each combination of these risk scenarios reflects a concrete risk attitude. Therefore, it can be said that the combination of these risk scenarios in this context is something that is directly identified by an explicit GlueVaR risk measure. To some extent, these risk attitudes could be related to risk appetite [Aven, 2013].

From the practitioner's point of view, four parameters must be fixed in order to define the GlueVaR risk measure. The α and β values correspond to the confidence levels used for bad and very bad scenarios, respectively. For instance, $\alpha = 95\%$ and $\beta = 99.5\%$ could be selected, which are equivalent to one bad event every twenty years or one bad event every two hundred, respectively. The other two parameters are directly related to the weights given to these scenarios. For instance, it could be said that the three components of GlueVaR in expression (1.4) are equally important. This would imply $\omega_1 = \omega_2 = \omega_3 = 1/3$, so the corresponding h_1 and h_2 parameters could be found. When $\omega_1 = \omega_2 = \omega_3 = 1/3$ and $\alpha = 95\%$, $\beta = 99.5\%$, these parameters are $h_1 = 11/30$ and $h_2 = 2/3$.

2.3 GEOMETRICAL DISCUSSION ON RISK ATTITUDES

Given α and β, the shaded areas in Figure 2.1 delimit feasible weights (ω_1, ω_2) for $\text{GlueVaR}_{\beta,\alpha}^{\omega_1,\omega_2}$. The point $(1/3, 1/3)$ corresponds to a balanced risk attitude on the part of risk managers when faced by the three components shown in (1.4). The corresponding distortion function $\kappa_{\beta,\alpha}^{\omega_1,\omega_2}$ is concave on $[0, 1-\alpha)$ in the lightly shaded area and, thus, the associated GlueVaR risk measure can be tail-subadditive. Yet, the distortion function is not concave on $[0, 1-\alpha)$ in the darkly shaded area and, thus, the associated GlueVaR risk measure cannot be tail-subadditive. The distortion function is concave in $[0, 1]$ in the boldest continuous segment and, thus, the associated GlueVaR risk measure is subadditive.

If $\omega_1 < 0$, risk managers are optimistic regarding the impossibility of the occurrence of the worst case scenario, and so attach a negative weight to it.

Note that any pair of weights (ω_1, ω_2) on the boldest line in Figure 2.1 leads to $\omega_3 = 0$. This means that a zero weight is allocated to the less conservative scenario, i.e. the one associated with the $\text{VaR}_\alpha(X)$. This is indicative of the decision makers' conservative approach. Nonetheless, differences in just how restrictive this conservative attitude is can be found among the weights lying on this line: the nearer to $(\omega_1, \omega_2) = \left(\frac{\beta-1}{\beta-\alpha}, \frac{1-\alpha}{\beta-\alpha}\right)$, the less restrictive it is, while the nearer to $(\omega_1, \omega_2) = (1, 0)$, the more conservative it is.

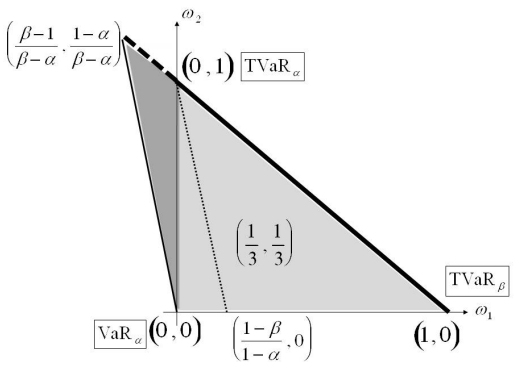


Figure 2.1: Given α and β, the shaded areas delimits feasible weights (ω_1, ω_2) for $\text{GlueVaR}_{\beta,\alpha}^{\omega_1,\omega_2}$.

3

RISK MEASUREMENT WITH GLUEVAR

3.1 AN EXAMPLE OF RISK MEASUREMENT ON CLAIM COSTS

Data for the cost of claims involving property damages and medical expenses from a major Spanish motor insurer are used to illustrate the application of GlueVaR measures in risk measurement. The sample consists of $n = 518$ observations of the cost of individual claims in thousands of euros. These data were previously analyzed in Bolancé et al. [2008] and Guillén et al. [2011].

In Table 3.1 a set of quantile-based risk measures including three different GlueVaR are displayed. The table is divided into three blocks, each block representing the corresponding risk figures for the cost of claims for property damage (X_1), the cost of claims of medical expenses (X_2) and the aggregate cost of claims ($X_1 + X_2$). Risk measure values using the empirical distribution (first row) are compared with outcomes when Normal, Lognormal, Student t with 4 df and Generalized Pareto distributions are fitted to data. In the last two rows of each block outcome results are shown when risk measure values are approximated by the Cornish-Fisher expansion shown in Chapter 1. The sample mean ($\hat{\mu} = \bar{z}$), the sample deviation ($\hat{\sigma}^2 = \sum_i (Z_i - \bar{z})^2 / (n-1)$) and the sample skewness (calculated as $\hat{\gamma} = \hat{\sigma}^{-3}(\sum_i (Z_i - \bar{z})^3 / n)$) are considered as estimators of μ, σ and γ when Z is one of the three random variables $X_1, X_2, X_1 + X_2$. Sample statistics were computed using observations that fall below the 99.5% quantile in order to exclude the effect of extreme losses on estimates (first Cornish-Fisher approximation). That means, a subsample of the first 516 increasingly ordered elements of the random variable were used to estimate parameters. Let remind that the sample size is 518 observations. Therefore, the two highest values were considered as extreme losses and were not included. Outcome values of risk measures were compared with the risk measure approximations when all the observations are included on sample estimates (second Cornish-Fisher approximation). All the calculations were made in R and MS Excel.

The selection of the three GlueVaR risk measures included in Table 3.1 deserves further explanation. The two confidence levels considered are $\alpha = 95\%$ and $\beta = 99.5\%$. The heights (h_1, h_2) are $(11/30, 2/3)$, $(0, 1)$ and $(1/20, 1/8)$ respectively. Different attitudes in front of the three scenarios of risk assessment are represented.

$\text{GlueVaR}^{11/30,2/3}_{99.5\%,95\%}$ corresponds to a balanced attitude because the three quantile-based risk measures $\text{TVaR}_{99.5\%}$, $\text{TVaR}_{95\%}$ and $\text{VaR}_{95\%}$ are equally important, i.e. $\omega_1 = \omega_2 = \omega_3 = 1/3$. A different attitude is symbolized by $\text{GlueVaR}^{0,1}_{99.5\%,95\%}$ with associated weights $\omega_1 = -1/9$, $\omega_2 = 10/9$ and $\omega_3 = 0$. It corresponds to a scenario in which the manager overweights $\text{TVaR}_{95\%}$ and allocates the lowest feasible weight to $\text{TVaR}_{99.5\%}$ given that a zero weight is allocated to $\text{VaR}_{95\%}$. Finally, $\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$ reflects a risk measurement attitude just a little bit more conservative than the one represented by using $\text{VaR}_{95\%}$, assigning low weights to $\text{TVaR}_{99.5\%}$ and $\text{TVaR}_{95\%}$ ($\omega_1 = 1/24$ and $\omega_2 = 1/12$).

As it is shown in Table 3.1, $\text{GlueVaR}^{11/30,2/3}_{99.5\%,95\%}$ is more conservative than the other two selected GlueVaR measures. This result can be generalized to all situations because the associated distortion function of $\text{GlueVaR}^{11/30,2/3}_{99.5\%,95\%}$ is greater than the other two distortion functions in the whole domain. Note that it is also observed in Table 3.1 that $\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%} \leqslant \text{GlueVaR}^{0,1}_{99.5\%,95\%}$. It is only valid to these data and an ordering between them can not be generalized. However, a relationship between these two GlueVaR risk measures and quantile-based risk measures can be established. In Chapter 1 it has been shown that $\text{VaR}_\alpha \leqslant \text{GlueVaR}^{h_1,h_2}_{\beta,\alpha} \leqslant \text{TVaR}_\alpha$ if $h_1 \leqslant (1-\beta)/(1-\alpha)$. That means, $\text{VaR}_{95\%} \leqslant \text{GlueVaR}^{0,1}_{99.5\%,95\%} \leqslant \text{TVaR}_{95\%}$, because $0 \leqslant 0.1$, and $\text{VaR}_{95\%} \leqslant \text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%} \leqslant \text{TVaR}_{95\%}$, because $0.05 \leqslant 0.1$. Although results in Table 3.1 invite to deduce that $\text{TVaR}_{95\%} \leqslant \text{GlueVaR}^{11/30,2/3}_{99.5\%,95\%} \leqslant \text{TVaR}_{99.5\%}$, it can not be asserted because conditions on the parameters of the GlueVaR risk measure to satisfy $\text{TVaR}_\alpha \leqslant \text{GlueVaR}^{h_1,h_2}_{\beta,\alpha} \leqslant \text{TVaR}_\beta$ are $h_1 \geqslant (1-\beta)/(1-\alpha)$ and $h_2 = 1$. In this case it holds $0.37 \geqslant 0.1$ but $h_2 \neq 1$.

Some comments related to outcome values for the Cornish-Fisher approximation of the quantile-based risk measures should be made. According to obtained results, it seems that this kind of risk measurement corresponds to a conservative attitude for the two types of approximations shown in Table 3.1. Relevant differences are observed depending on the approximation finally used. If the first Cornish-Fisher approximation is considered, i.e. when sample statistics were estimated excluding extreme losses, it is observed that the outcome values for this approximation are in most of the cases larger than those values associated with the empirical or the parametric distributions. It happens in thirteen cases among the sixteen examples. Although conservative values are obtained with this approximation, results are in general comparable with those computed with the empirical and parametric distributions. Unlike values of this first Cornish-Fisher approximation, outcome values related to the second Cornish-Fisher approximation are drastically larger than the rest in

Table 3.1: Examples of risk measurement of costs of insurance claims using quantile-based risk measures

					$\mathbf{GlueVaR}^{h_1,h_2}_{99.5\%,95\%}$		
	Model	$\mathbf{VaR}_{95\%}$	$\mathbf{TVaR}_{95\%}$	$\mathbf{TVaR}_{99.5\%}$	$\left(\frac{11}{30}, \frac{2}{3}\right)$	$(0,1)$	$\left(\frac{1}{20}, \frac{1}{8}\right)$
X_1	Empirical	38.8	112.5	440.0	197.1	76.1	61.7
	Normal	78.9	96.1	130.4	101.8	92.3	82.5
	Lognormal	42.5	110.1	388.3	180.3	79.2	62.5
	Student t (4 d.f.)	99.0	143.2	272.1	171.4	128.9	109.9
	Pareto	38.3	82.4	264.5	128.4	62.2	51.4
	Cornish-Fisher$^{(1a)}$	61.3	169.2	724.3	318.3	107.5	98.0
	Cornish-Fisher$^{(1b)}$	262.1	1,081.9	5,437.9	2,260.6	597.9	546.1
X_2	Empirical	6.4	18.4	54.2	26.3	14.4	9.4
	Normal	10.2	12.4	16.7	13.1	11.9	10.7
	Lognormal	6.6	15.4	50.1	24.0	11.5	9.1
	Student t (4 d.f.)	12.8	18.3	34.5	21.9	16.5	14.2
	Pareto	5.9	12.4	38.5	18.9	9.5	7.8
	Cornish-Fisher$^{(2a)}$	14.3	45.4	207.3	89.0	27.4	24.9
	Cornish-Fisher$^{(2b)}$	22.1	76.1	359.4	152.5	44.6	40.6
$X_1 + X_2$	Empirical	47.6	125.5	479.0	217.4	86.2	72.1
	Normal	87.0	105.9	143.4	112.1	101.7	90.9
	Lognormal	49.1	124.1	428.8	200.7	90.2	71.2
	Student t (4 d.f.)	109.0	157.5	298.6	188.4	141.8	120.9
	Pareto	44.2	94.6	301.4	146.7	71.6	59.1
	Cornish-Fisher$^{(3a)}$	71.3	198.0	850.7	373.3	125.4	114.4
	Cornish-Fisher$^{(3b)}$	283.6	1,164.0	5,840.3	2,429.3	644.4	588.5

$^{(1a)}$ $\hat{\mu} = 9.0$, $\hat{\sigma} = 17.9$ and $\hat{\gamma} = 4.5$. Subsample without extreme losses.
The two largest values of X_1 are excluded.

$^{(1b)}$ $\hat{\mu} = 11.0$, $\hat{\sigma} = 41.3$ and $\hat{\gamma} = 15.6$. Full sample.

$^{(2a)}$ $\hat{\mu} = 1.5$, $\hat{\sigma} = 3.7$ and $\hat{\gamma} = 6.4$. Subsample without extreme losses.
The two largest values of X_2 are excluded.

$^{(2b)}$ $\hat{\mu} = 1.7$, $\hat{\sigma} = 5.2$ and $\hat{\gamma} = 8.0$. Full sample.

$^{(3a)}$ $\hat{\mu} = 10.5$, $\hat{\sigma} = 20.6$ and $\hat{\gamma} = 4.6$. Subsample without extreme losses.
The two largest values of $X_1 + X_2$ are excluded.

$^{(3b)}$ $\hat{\mu} = 12.7$, $\hat{\sigma} = 45.2$ and $\hat{\gamma} = 15.3$. Full sample.

all the examples. These outcome values would be associated to a excessively conservative (unrealistic) attitude. Let remind that only the two largest losses are not included in the sample estimates involving the first approximation. In other words, the Cornish-Fisher approximation shows a poor performance when the data are severely right skewed distributed, as in this case. However, the performance of this approximation seems to be improved when extreme losses are excluded for the sample estimates of parameters.

An important issue that arises from results is the model risk. Even when the same risk measure is used, huge differences are observed depending on the hypothesis about the underlying distribution of the claim cost random variables. Let us assume that the regulator is focused on the $\text{VaR}_{95\%}$ for the aggregate cost $X_1 + X_2$ as a measure of pure underwriting risk (without taking into account the premium paid by the policyholders). If it is supposed that the random variable is Pareto distributed, then the institution will need 44.2 thousands of euros for regulatory solvency purposes. The company should set aside almost 2.5 times this economic amount whether the underlying distribution is Student-t with 4 degrees of freedom. This topic is out of the scope of this dissertation. The interested reader is addressed, for instance, to the study of Alexander and Sarabia [2012] which deals with VaR model risk.

3.2 INSIGHTS ON THE TAIL-SUBADDITIVITY PROPERTY

In order to preserve the benefits of diversification when aggregating risks, an appealing property of a risk measure is subadditivity. The subadditivity property ensures that the risk measure value of the aggregated risk is lower than or equal to the sum of individual risk measure values. The subadditivity characteristic is guaranteed for the TVaR but not for the VaR risk measure.

It has been argued in Chapter 2 that, in practice, main concerns of managers are related to the performance of aggregated risks in the tail region. The properties of GlueVaR risk measures in tails have been investigated from a theoretical point of view, where foundations of the tail-subadditivity were established. It has been shown that a subfamily of GlueVaR risk measures may satisfy this property.

Implications of tail-subadditivy are now investigated not only from a theoretical perspective: implications of this property for insurance institutions in comparison to subadditivity in the whole domain are analyzed. To reach this goal, the subadditivity and tail-subadditivity properties of GlueVaR risk measures in the aggregation

of risks are empirically examined, and both illustrated with a numerical example based on real insurance claim data.

3.3 AN EXAMPLE OF TAIL-SUBADDITIVITY ON CLAIM COSTS

Although subadditivity in the whole domain is in general not satisfied by GlueVaR risk measures, it has been shown in Chapter 2 that a subfamily of GlueVaR measures may satisfy the subadditivity property in the tail region. Recall the definition of tail-subadditivity for a pair of risks, following the notation used in Section 2.1.

The idea is that the risk of a sum is smaller than or equal to the sum of risks when focusing on the extreme region, without worrying about whatever relationship exists between the risk of the sum and the sum of risks outside the extreme region. Remember that, given a confidence level α, a GlueVaR risk measure only can be tail-subadditive if its associated distortion function is concave in $[0, 1-\alpha)$. In this chapter it is shown through an example that tail-subadditivity is a convenient property to preserve the benefits of diversification in extremely adverse cases.

Now the example described at the beginning of this chapter is followed to investigate the tail-subadditivity property of GlueVaR risk measures. In that example $\text{GlueVaR}^{11/30,2/3}_{99.5\%,95\%}$ and $\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$ are candidates to satisfy subadditivity in tails for a pair of risks at confidence level $\alpha = 95\%$. Note that it holds in both cases that $h_2 \leqslant h_1 (1-\alpha)/(1-\beta)$ ($2/3 \leqslant 11/3$ and $1/8 \leqslant 1/2$, respectively). However, this inequality is not fulfilled by $\text{GlueVaR}^{0,1}_{99.5\%,95\%}$ and, then, this GlueVaR risk measure does not satisfy the tail-subadditivity property. In fact, Table 3.1 seems to reflect subadditivity of the two risk measures $\text{GlueVaR}^{11/30,2/3}_{99.5\%,95\%}$ and $\text{GlueVaR}^{0,1}_{99.5\%,95\%}$. Indeed, the risk measure outcomes for the aggregate risk are lower than the sum of individual risk values in all of the models with the exception of the outcomes associated to the second Cornish-Fisher approximation considered. It must be emphasized that this result is strongly related to these data but the subadditivity property can not be generalized to all the circumstances.

A comment on the subadditivity of risk measures when the Cornish-Fisher approximation is used should be made. Unlike the VaR risk measure, it was previously discussed that the TVaR risk measure satisfies the subadditivity property. In the example, however, the second Cornish-Fisher approximation of the TVaR risk measure value fails subadditivity. Note that it is deduced from the results displayed in Table 3.1 that $\text{TVaR}_\beta(X_1) + \text{TVaR}_\beta(X_2) < \text{TVaR}_\beta(X_1 + X_2)$ (this is,

$5,797.3 < 5,840.3$) and $\text{TVaR}_\alpha(X_1) + \text{TVaR}_\alpha(X_2) < \text{TVaR}_\alpha(X_1 + X_2)$ (this is, $1,158.0 < 1,164.0$). Therefore, the subadditivity property of the TVaR measure is not ensured when the risk measure value is approximated by the second Cornish-Fisher approximation in the example. This result supports the statement that the second Cornish-Fisher approximation shown is not adequate to estimate quantile-based risk measure values for highly right skewed data.

Let us focus on the outcomes for the $\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$ when the empirical distribution is considered. Table 3.1 shows that, in this case, $\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$ fails to be subadditive for X_1 and X_2, since $61.7 + 9.4 \leqslant 72.1$. In order to analyze the tail-subadditivity property for this GlueVaR risk measure, the common right tail of the empirical distribution has to be firstly isolated. The common 5%-right tail for the empirical distribution is separated as follows. A subsample is selected which satisfies the criterion that each individual risk values are above its respective 95%-quantile given that the values of the aggregate random variable fall above its 95%-quantile and the values of the other individual random risk fall above its respective 95%-quantile as well. Risk measure values are then computed for this subsample, where the survival probabilities associated to the observations of this subsample have not been changed. Table 3.2 displays the values of their common 5%-right tail for the individual random variables and the aggregate random variable.

An illustration of tail-subadditivity of the $\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$ risk measure is provided in Table 3.3, where results obtained for the risk measure $\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$ when aggregating risks in the whole domain are compared with those $\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$ outcomes in the common 5%-right tail. In the second block, risk measure values are computed for the three random variables in the common 5%-right tail, i.e. using data shown in Table 3.2. Outcome results of the $\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$ are in bold type to highlight differences between subadditivity in the whole range and subadditivity in tails. The last row of the second block illustrates numerically the 95% tail-subadditivity property of $\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$ for the pair of risks X_1 and X_2, where diversification benefit is computed as the difference between the sum of GlueVaR outcome values for individual risks and the outcome value of the aggregate risk. On the common 5%-right tail, a benefit of diversification of 2.2 thousands of euros is observed for the $\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$ risk measure. In other words, the aggregate risk $X_1 + X_2$ is preferable than these risks individually taken in simultaneously adverse events for X_1 and X_2, according to the results of the $\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$. However, it does not hold whether the whole domain of the random variables is considered. Last row in the first block shows a negative value for the diversification benefit associated

Table 3.2: Common 5%-right tail for X_1, X_2 and $X_1 + X_2$

$\mathbf{i}$	$\mathbf{x_{1,i}}$	$\mathbf{x_{2,i}}$	$\mathbf{s_i = x_{1,i} + x_{2,i}}$
1	829.0	71.3	900.3
12	108.2	23.7	131.9
32	55.0	44.3	99.3
185	121.6	32.5	154.1
189	74.2	13.2	87.4
198	88.8	30.1	118.9
213	57.5	10.0	67.5
214	148.7	10.2	158.9
289	145.4	42.2	187.6
294	44.8	7.5	52.3
297	221.5	8.3	229.8

A discrete finite probability space $\Omega = \{\varpi_1, \varpi_2, ..., \varpi_{518}\}$ is considered.
Each $\mathtt{ith}$ observation $(x_{1,i}, x_{2,i}, s_i)$ corresponds to a realization
of random event ϖ_i.
Note that all values in last three columns are greater or equal than
their empirical quantiles at 95% level, where
$\mathrm{VaR}_{95\%}(X_1)$=38.8; $\mathrm{VaR}_{95\%}(X_2) = 6.4$;$\mathrm{VaR}_{95\%}(X_1 + X_2)$=47.6.

to the $\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$ on the whole domain.

The underlying idea in the assessment of the incurred risk in this context is that the benefit of diversification in simultaneously adverse events is balanced by the cost of diversification in the rest of cases. Therefore, divergence decisions would be taken by managers depending on where the attention is paid, all the scenarios or highly adverse scenarios. As it is shown in Table 3.3, this phenomenon is due to the lack of subadditivity of $\text{VaR}_{95\%}$ on the whole domain. By considering the common 5%-right tail, the effect of $\text{VaR}_{95\%}$ on the whole domain is blurred on the tail.

Table 3.3: Subadditivity and tail-subadditivity

		$\mathbf{X_1}$	$\mathbf{X_2}$	$\mathbf{X_1 + X_2}$	**Difference**$^{(*)}$
		(a)	(b)	(c)	(a+b-c)
Whole range	$\text{VaR}_{95\%}$	38.8	6.4	47.6	-2.4
	$\text{TVaR}_{95\%}$	112.5	18.4	125.5	5.4
	$\text{TVaR}_{99.5\%}$	440.0	54.2	479.0	15.2
	$\mathbf{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$	**61.7**	**9.4**	**72.1**	**-1.0**
Common 5%-right tail$^{(**)}$	$\text{VaR}_{95\%}$	0.0	0.0	0.0	0.0
	$\text{TVaR}_{95\%}$	75.3	12.5	76.8	11.0
	$\text{TVaR}_{99.5\%}$	411.3	46.7	426.7	31.3
	$\mathbf{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$	**23.4**	**3.0**	**24.2**	**2.2**

$^{(*)}$ Benefit of diversification.

$^{(**)}$ The figures represent the contributions to the overall value of each risk measure that are linked to the common 5% right tail.

4

THE CONNECTION BETWEEN DISTORTION RISK MEASURES AND ORDERED WEIGHTED AVERAGING OPERATORS

GlueVaR risk measures have been introduced in the previous chapters. In this one, some relationships between two different worlds, namely risk measurement and fuzzy systems, are investigated. Risk measurement evaluates potential losses and is useful for decision making under probabilistic uncertainty. Broadly speaking, fuzzy logic is a form of reasoning based on the 'degree of truth' rather than on the binary true-false principle. But risk measurement and fuzzy systems share a common core theoretical background.

Both fields are related to the human behavior under risk, ambiguity or uncertainty. The expected utility theory by von Neumann and Morgenstern [1947] was one of the first attempts to provide a theoretical foundation to human behavior in decision-making, mainly based on setting up axiomatic preference relations of the decision maker. Similar theoretical approaches are, for instance, the certainty-equivalence theory [Handa, 1977], the cumulative prospect theory [Kahneman and Tversky, 1979; Tversky and Kahneman, 1992], the rank-dependent utility theory [Quiggin, 1982], the dual theory of choice under risk [Yaari, 1987] and the expected utility without subadditivity [Schmeidler, 1989], where the respective axioms reflect possible human behaviors or preference relations in decision-making.

The study of the relationship between risk measurement and fuzzy systems is a topic of ongoing research from both fields. Goovaerts et al. [2010a], for instance, discuss the hierarchical order between risk measures and decision principles, while Aliev et al. [2012] propose a decision theory under imperfect information from the perspective of fuzzy systems.

Previous attempts to link risk management and fuzzy logic approaches are mainly found in the literature on fuzzy systems. Most authors have focused on the application of fuzzy criteria to financial decision making [Engemann et al., 1996; Gil-Lafuente, 2005; Merigó and Casanovas, 2011], and some have smoothed financial series under fuzzy logic for prediction purposes [Yager and Filev, 1999; Yager, 2008]. In the literature on risk management, contributions made by Shapiro [2002, 2004, 2009] regarding the application of fuzzy logic in

the insurance context must be remarked.

In this chapter the mathematical relationship between risk measurement and some aggregation instruments used in fuzzy systems for discrete random variables is analyzed. It is known that a risk measure quantifies the complexity of a random loss in one value that reflects the amount at risk. A key concept in fuzzy systems applications is the aggregation operator, which also allows to combine data into a single value. The relationship between the well-known distortion risk measures introduced by Wang [1996] -presented in this dissertation at Section 1.2 of Chapter 1- and two specific aggregation operators, the Ordered Weighted Averaging (OWA) operator introduced by Yager [1988] and the Weighted Ordered Weighted Averaging (WOWA) operator introduced by Torra [1997] is shown in this chapter.

Distortion risk measures, OWA and WOWA operators can be analyzed using the theory of measure. Classical measure functions are additive, and linked to the Lebesgue integral. When the additivity is relaxed, alternative measure functions and, hence, associated integrals are derived. This is the case of non-additive measure functions [see, for instance, Denneberg, 1994], often called capacities as it was the name coined by Choquet [1954]. It is shown that the link between distortion risk measures and OWA and WOWA operators is derived by means of the integral linked to capacities, i.e. the Choquet integral. I present the concept of local degree of orness for distortion risk measures and illustrate its usefulness.

4.1 THE OWA AND WOWA OPERATORS AND THE CHOQUET INTEGRAL

Aggregation operators (or aggregation functions) have been extensively used as a natural form to combine inputs into a single value. These inputs may be understood as degrees of preference, membership or likelihood, or as support of a hypothesis. Let us denote by $\overline{\mathbb{R}} = [-\infty, +\infty]$ the extended real line, and by $\mathbb{I}$ any type of interval in $\overline{\mathbb{R}}$ (open, closed, with extremes being $\mp\infty$,...). Following Grabisch et al. [2011], an aggregation operator is defined.

Definition 4.1.1 (Aggregation operator)**.** *An aggregation operator in* $\mathbb{I}^n$ *is a function* $F^{(n)}$ *from* $\mathbb{I}^n$ *to* $\mathbb{I}$*, that is non-decreasing in each variable; fulfills the following boundary conditions,* $\inf_{\vec{x}\in\mathbb{I}^n} F^{(n)}(\vec{x}) = \inf \mathbb{I}$, $\sup_{\vec{x}\in\mathbb{I}^n} F^{(n)}(\vec{x}) = \sup \mathbb{I}$*; and* $F^{(1)}(x) = x$ *for all* $x \in \mathbb{I}$.

Some basic aggregation operators are displayed in Table 4.1.

Table 4.1: Basic aggregation operators.

Name	Mathematical expression	Type of interval $\mathbb{I}$
Arithmetic mean	$AM(\vec{x}) = \frac{1}{n}\sum_{i=1}^{n} x_i$	Arbitrary $\mathbb{I}$. If $\mathbb{I} = \overline{\mathbb{R}}$, the convention $+\infty + (-\infty) = -\infty$ is often considered.
Product	$\Pi(\vec{x}) = \prod_{i=1}^{n} x_i$	$\mathbb{I} \in \{\|0,1\|, \|0,+\infty\|, \|1,+\infty\|\}$, where $\|a,b\|$ means any kind of interval, with boundary points a and b, and with the convention $0 \cdot (+\infty) = 0$.
Geometric mean	$GM(\vec{x}) = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$	$\mathbb{I} \subseteq [0,+\infty]$, with the convention $0 \cdot (+\infty) = 0$.
Minimum function	$Min(\vec{x}) = \min\{x_1, x_2, ..., x_n\}$	Arbitrary $\mathbb{I}$.
Maximum function	$Max(\vec{x}) = \max\{x_1, x_2, ..., x_n\}$	Arbitrary $\mathbb{I}$.
Sum function	$\sum(\vec{x}) = \sum_{i=1}^{n} x_i$	$\mathbb{I} \in \{\|0,+\infty\|, \|-\infty,0\|, \|-\infty,+\infty\|\}$, with the convention $+\infty + (-\infty) = -\infty$.
k-order statistics	$OS_k(\vec{x}) = x_j$, $k \in \{1,...,n\}$ where x_j is such that $\#\{i \mid x_i \leqslant x_j\} \geqslant k$ and $\#\{i \mid x_i > x_j\} < n-k$	Arbitrary $\mathbb{I}$.
kth projection	$P_k(\vec{x}) = x_k$, $k \in \{1,...,n\}$	Arbitrary $\mathbb{I}$.

$\vec{x}$ denotes $(x_1, x_2, ..., x_n)$.

Source: Grabisch et al. [2011].

There is a huge amount of literature on aggregation operators and its applications. See, among others, Beliakov et al. [2007], Torra and Narukawa [2007] and Grabisch et al. [2009, 2011]. Despite the large number of aggregation operators, I focus on the OWA operator and on the WOWA operator. Several reasons lead me to this selection. The OWA operator has been extensively applied in the context of decision making under uncertainty because it provides a unified formulation for the optimistic, the pessimistic, the Laplace and the Hurwicz criteria [Yager, 1993], and there are also some interesting generalizations [Yager et al., 2011]. The WOWA operator combines the OWA operator with the concept of weighted average, where weights are a mechanism to include expert opinion on the accuracy of information. This operator is closely linked to distorted probabilities.

Ordered Weighted Averaging operator

The OWA operator is an aggregation operator that provides a parameterized family of aggregation operators offering a compromise between the minimum and the maximum aggregation functions. It was introduced in Yager [1988] and it can be defined as follows.

Definition 4.1.2 (OWA operator). *Let* $\vec{w} = (w_1, w_2, ..., w_n) \in [0,1]^n$ *such that* $\sum_{i=1}^{n} w_i = 1$. *The Ordered Weighted Averaging (OWA) operator with respect to* $\vec{w}$ *is a mapping from* $\mathbb{R}^n$ *to* $\mathbb{R}$ *defined by*

$$OWA_{\vec{w}}(x_1, x_2, ..., x_n) := \sum_{i=1}^{n} x_{\sigma(i)} \cdot w_i,$$

where σ *is a permutation of* $(1, 2, ..., n)$ *such that* $x_{\sigma(1)} \leqslant x_{\sigma(2)} \leqslant ... \leqslant x_{\sigma(n)}$, *i.e.* $x_{\sigma(i)}$ *is the* i*th smallest value of* $x_1, x_2, ..., x_n$.

Unlike the original definition, an ascending order in $\vec{x}$ instead of a decreasing one is here considered. This definition is convenient from the risk management perspective since $\vec{x}$ may be a set of losses in ascending order. The relationship between the ascending OWA and the descending OWA operators is already provided by Yager [1993].

The OWA operator is commutative, monotonic and idempotent, and it is lower-bounded by the minimum and upper-bounded by the maximum operators. Commutativity is referred to any permutation of the components of $\vec{x}$. That is, if the $OWA_{\vec{w}}$ operator is applied to any $\vec{y}$ such that $y_i = x_{r(i)}$ for all i, and r is any permutation of $(1, ..., n)$, then $OWA_{\vec{w}}(\vec{y}) = OWA_{\vec{w}}(\vec{x})$. Monotonicity means that if $x_i \geqslant y_i$ for all i, then $OWA_{\vec{w}}(\vec{x}) \geqslant OWA_{\vec{w}}(\vec{y})$. Idempotency assures that if $x_i = a$ for all i, then $OWA_{\vec{w}}(\vec{x}) = a$. The OWA operator accomplishes the boundary conditions because it is delimited by the minimum and the maximum functions, i.e. $\min_{i=1,...,n}\{x_i\} \leqslant OWA_{\vec{w}}(\vec{x}) \leqslant \max_{i=1,...,n}\{x_i\}$.

The $OWA_{\vec{w}}$ is unique with respect to the vector $\vec{w}$ (the proof is provided in the Appendix). The characterization of the weighting vector $\vec{w}$ is often made by means of the *degree of orness* indicator [Yager, 1988].

Definition 4.1.3 (Degree of orness of an OWA operator)**.** *Let* $\vec{w} \in [0,1]^n$ *such that* $\sum_{i=1}^{n} w_i = 1$*, the degree of orness of* $OWA_{\vec{w}}$ *is defined by*

$$\boldsymbol{\omega}\left(OWA_{\vec{w}}\right) := \sum_{i=1}^{n} \left(\frac{i-1}{n-1}\right) \cdot w_i.$$

Note that the degree of orness represents the level of aggregation preference between the minimum and the maximum when $\vec{w}$ is fixed. The degree of orness can be understood as the value that the OWA operator returns when it is applied to $\vec{x^*} = \left(\frac{0}{n-1}, \frac{1}{n-1}, ..., \frac{n-2}{n-1}, \frac{n-1}{n-1}\right)$. In other words, $\boldsymbol{\omega}\left(OWA_{\vec{w}}\right) = OWA_{\vec{w}}\left(\vec{x^*}\right)$. It is straightforward to see that $\boldsymbol{\omega}\left(OWA_{\vec{w}}\right) \in [0,1]$, because $\vec{x^*}, \vec{w} \in [0,1]^n$. If $\vec{w} = (1,0,...,0)$, then $OWA_{\vec{w}} \equiv Min$ and $\boldsymbol{\omega}\left(Min\right) = 0$. Conversely, if $\vec{w} = (0,0,...,1)$, then $OWA_{\vec{w}} \equiv Max$ and $\boldsymbol{\omega}\left(Max\right) = 1$. And when $\vec{w}$ is such that $w_i = \frac{1}{n}$ for all i, then $OWA_{\vec{w}}$ is the arithmetic mean and its degree of orness is 0.5. As it is seen later, orness is closely related to the α level chosen in risk measures.

Alternatively to the degree of orness, other measures can be used to characterize the weighting vector, such as the *entropy of dispersion* [Yager, 1988] based on the Shannon entropy [Shannon, 1948] and the *divergence of the weighting vector* [Yager, 2002], as it is shown in Chapter 5.

The OWA operator has been extended and generalized in many ways. For example, Xu and Da [2002] introduced the uncertain OWA (UOWA) operator in order to deal with imprecise information, Merigó and Gil-Lafuente [2009] developed a generalization by using induced aggregation operators and quasi-arithmetic means called the induced quasi-OWA (Quasi-IOWA) operator and Yager [2010] introduced a new approach to cope with norms in the OWA operator. Although it is out of the scope of this chapter, the OWA operator is also related to the linguistic quantifiers introduced by Zadeh [1985], and a subset of them may be interpreted as distortion functions.

Weighted Ordered Weighted Averaging operator

The WOWA operator is the aggregation function introduced by Torra [1997]. This operator unifies in the same formulation two other functions, the weighted mean function and the OWA operator. It is defined in the following way.

Definition 4.1.4 (WOWA operator)**.** *Let* $\vec{v} = (v_1, v_2, ..., v_n) \in [0,1]^n$ *and* $\vec{q} = (q_1, q_2, ..., q_n) \in [0,1]^n$ *such that* $\sum_{i=1}^{n} v_i = 1$ *and* $\sum_{i=1}^{n} q_i = 1$.

The Weighted Ordered Weighted Averaging (WOWA) operator with respect to $\vec{v}$ *and* $\vec{q}$ *is a mapping from* $\mathbb{R}^n$ *to* $\mathbb{R}$ *defined by*

$$WOWA_{h,\vec{v},\vec{q}}(x_1, x_2, ..., x_n) := \sum_{i=1}^{n} x_{\sigma(i)} \cdot \left[h\left(\sum_{j \in A_{\sigma,i}} q_j \right) - h\left(\sum_{j \in A_{\sigma,i+1}} q_j \right) \right],$$

where σ *is a permutation of* $(1, 2, ..., n)$ *such that* $x_{\sigma(1)} \leqslant x_{\sigma(2)} \leqslant ... \leqslant x_{\sigma(n)}$, $A_{\sigma,i} = \{\sigma(i), ..., \sigma(n)\}$ *and* $h : [0,1] \to [0,1]$ *is a non-decreasing function such that* $h(0) := 0$ *and* $h\left(\frac{i}{n}\right) := \sum_{j=n-i+1}^{n} v_j$*; and* h *is linear if the points* $\left(\frac{i}{n}, \sum_{j=n-i+1}^{n} v_j \right)$ *lie on a straight line.*

Note that this definition implies that weights v_i can be expressed as $v_i = h\left(\frac{n-i+1}{n}\right) - h\left(\frac{n-i}{n}\right)$ and that $h(1) = 1$.
It has to be mentioned that in the original definition $\vec{x}$ components are in descending order, while the ascending order is used in this dissertation. An additional subindex to emphasize dependence on function h is also introduced here.

Remark 1

The WOWA operator generalizes the OWA operator. Given a $WOWA_{h,\vec{v},\vec{q}}$ operator on $\mathbb{R}^n$, if weights w_i are defined by

$$w_i := h\left(\sum_{j \in A_{\sigma,i}} q_j \right) - h\left(\sum_{j \in A_{\sigma,i+1}} q_j \right),$$

and $OWA_{\vec{w}}$ where $\vec{w} = (w_1, ..., w_n)$, then the following equality holds $WOWA_{h,\vec{v},\vec{q}} = OWA_{\vec{w}}$. As it can easily be shown, vector $\vec{w}$ satisfies the following conditions:
(i) $\vec{w} \in [0,1]^n$;
(ii) $\sum_{i=1}^{n} w_i = 1$.

Condition *(i)* is straightforward. Let us denote $s_i = \sum_{j \in A_{\sigma,i}} q_j$ and $s_{n+1} := 0$. Hence, $s_i \geqslant s_{i+1}$ for all i due to the fact that $A_{\sigma,i} \supseteq A_{\sigma,i+1}$ and that $q_j \geqslant 0$. Then $h(s_i) \geqslant h(s_{i+1})$ since h is a non-decreasing function. Finally, as $s_i \in [0,1]$ and $h(s) \in [0,1]$ for all $s \in [0,1]$, then it follows that $w_i = h(s_i) - h(s_{i+1}) \in [0,1]$ for all i.

To prove condition *(ii)*, note that $A_{\sigma,1} = N$, $\sum_{j \in N} q_j = 1$ and that $h(1) = 1$ and $h(0) = 0$, then $\sum_{i=1}^{n} w_i = \sum_{i=1}^{n} (h(s_i) - h(s_{i+1})) = h(s_1) - h(s_{n+1}) = 1 - 0 = 1$.

Remark 2

Let me analyze the particular case when OWA and WOWA operators provide the expectation of random variables. Suppose that X is a discrete random variable that takes n different values and $\vec{x} \in \mathbb{R}^n$ is the vector of values, where the components are in ascending order. Let $\vec{p} \in [0,1]^n$ be a vector consisting on the probabilities of the components of $\vec{x}$. Obviously, it holds that $OWA_{\vec{p}}(\vec{x}) = \mathbb{E}(X)$. Besides,

$$\begin{aligned} WOWA_{h,\vec{v},\vec{p}}(\vec{x}) &= \sum_{i=1}^{n} x_i \cdot \left[h\left(\sum_{j=i}^{n} p_j \right) - h\left(\sum_{j=i+1}^{n} p_j \right) \right] \\ &= \sum_{i=1}^{n} x_i \cdot \left[h\left(S_X(x_{i-1}) \right) - h\left(S_X(x_i) \right) \right]. \end{aligned}$$

If h is the identity function then $WOWA_{h,\vec{v},\vec{p}}(\vec{x}) = \mathbb{E}(X)$ since $S_X(x_{i-1}) - S_X(x_i) = p_i$ for all i (with the convention $x_0 := -\infty$).

Remark 3

Note that if X is discrete and uniformly distributed then $S_X(x_{i-1}) = \dfrac{n-i+1}{n}$ for all $i = 2, ..., n+1$, and hence

$$h\left(S_X(x_{i-1}) \right) = h\left(\frac{n-i+1}{n} \right) = \sum_{j=i}^{n} v_j.$$

This remark is helpful to interpret the WOWA operator from the perspective of risk measurement. In the WOWA operator the subjective opinion of experts may be represented by vector $\vec{v}$. Let us suppose that no information regarding the distribution function of a discrete and finite random variable X is available. If it is assumed that X is discrete and uniformly distributed, then vector $\vec{v}$ directly consists of the subjective probabilities of occurrence of the components x_i according to the expert opinion. Another possible point of view in this case is that $\vec{v}$ represents the subjective importance that the expert gives to each x_i.

Remark 4

Since the domain of the survival function is $\mathbb{R}$, then the selected function h is crucial from the risk measurement point of view, especially for a small n.

The Choquet integral in the finite and discrete case

The Choquet integral presented in Section 1.2 of Chapter 1 has become a familiar concept to risk management experts since it was

introduced by Wang [1996] in the definition of distortion risk measures. OWA and WOWA operators can also be defined based on the concept of the Choquet integral. In this subsection Grabisch et al. [2011] is followed to provide several definitions which are needed in Section 4.2.

Definition 4.1.5 (Capacity). *Let* $N = \{m_1, ..., m_n\}$ *be a finite set and* $2^N = \wp(N)$ *be the set of all subsets of* N. *A capacity or a fuzzy measure on* N *is a mapping from* 2^N *to* $[0, 1]$ *which satisfies*

(I) $\mu(\emptyset) = 0$;

(II) $A \subseteq B \Rightarrow \mu(A) \leqslant \mu(B)$, *for any* $A, B \in 2^N$ *(monotonicity).*

If $\mu(N) = 1$, *then it is said that* μ *satisfies normalization, which is a frequently required property.*
A capacity μ *is additive if* $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$ *for any* $A, B \subseteq N$.
A capacity μ *is symmetric if* $\mu(A) = \mu(B)$ *for all* A, B *with the same cardinality (i.e.,* $|A| = |B|$*).*

Definition 4.1.6 (Dual capacity). *Let* μ *be a capacity on* N. *Its dual or conjugate capacity* $\bar{\mu}$ *is a capacity on* N *defined by*

$$\bar{\mu}(A) := \mu(N) - \mu\left(\bar{A}\right),$$

where $\bar{A} = N \backslash A$ *(i.e.,* $\bar{A}$ *is the set of all the elements in* N *that do not belong to* A*).*

If a finite probability space $\left(N, 2^N, \mathcal{P}\right)$ is considered, note that the probability $\mathcal{P}$ is a capacity (or a fuzzy measure) on N that satisfies normalization. In addition, $\mathcal{P}$ is its own dual capacity.

Definition 4.1.7 (Choquet integral for discrete positive functions). *Let* μ *be a capacity on* N, *and* $f : N \to [0, +\infty)$ *be a function. Let* σ *be a permutation of* $(1, ..., n)$, *such that* $f\left(m_{\sigma(1)}\right) \leqslant f\left(m_{\sigma(2)}\right) \leqslant ... \leqslant f\left(m_{\sigma(n)}\right)$, *and* $A_{\sigma,i} = \left\{m_{\sigma(i)}, ..., m_{\sigma(n)}\right\}$, *with* $A_{\sigma,n+1} = \emptyset$. *The Choquet integral of* f *with respect to* μ *is defined by*

$$\mathcal{C}_\mu(f) := \sum_{i=1}^{n} f\left(m_{\sigma(i)}\right)\left(\mu(A_{\sigma,i}) - \mu(A_{\sigma,i+1})\right).$$

If $f\left(m_{\sigma(0)}\right) := 0$, then an equivalent expression for the definition of the Choquet integral is

$$\mathcal{C}_\mu(f) = \sum_{i=1}^{n} \left[f\left(m_{\sigma(i)}\right) - f\left(m_{\sigma(i-1)}\right)\right] \mu(A_{\sigma,i}).$$

The concept of degree of orness introduced for the OWA operator may be extended to the case of the Choquet integral for positive functions as

$$\omega\left(\mathcal{C}_\mu\right) := \sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)\cdot\left(\mu\left(A_{id,i}\right)-\mu\left(A_{id,i+1}\right)\right). \tag{4.1}$$

This is not the only way to define this concept, as it is discussed in next chapters. In order to distinguish definition linked to expression 4.1 from other alternatives, I refer to it as *local degree of orness*. Let me illustrate the degree of orness for three simple capacities. The first one, denoted as μ_*, is such that $\mu_*(A) = 0$ if $A \neq N$ and $\mu_*(N) = 1$. In this case, $\mathcal{C}_{\mu_*} \equiv Min$ and it is deduced through expression (4.1) that $\omega(Min) = 0$. The second case, denoted as μ^*, is such that $\mu^*(\emptyset) = 0$ and $\mu^*(A) = 1$ if $A \neq \emptyset$. In this situation, $\mathcal{C}_{\mu^*} \equiv Max$ and, as expected, it is obtained that $\omega(Max) = 1$. Finally, consider capacity $\mu^\#$ such that $\mu^\#(A)$ solely depends on the cardinality of A for all $A \subseteq N$. Then $\mu^\#(A_{\sigma,i}) - \mu^\#(A_{\sigma,i+1})$ is defined by i. If notation $w_i = \mu^\#(A_{\sigma,i}) - \mu^\#(A_{\sigma,i+1})$ is used for all i, it follows that $\mathcal{C}_{\mu^\#}$ is equal to $OWA_{\vec{w}}$. In the particular case where $\mu^\#(A) = \frac{\#A}{n}$ for any $A \subseteq N$, then $w_i = \frac{n-(i-1)}{n} - \frac{n-i}{n} = \frac{1}{n}$. So, in this situation $\mathcal{C}_{\mu^\#}$ is the arithmetic mean, and it can be easily verified that $\omega\left(\mathcal{C}_{\mu^\#}\right) = 0.5$:

$$\begin{aligned}\omega\left(\mathcal{C}_{\mu^\#}\right) &= \sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)\cdot\left(\mu^\#\left(A_{id,i}\right)-\mu^\#\left(A_{id,i+1}\right)\right)\\ &= \sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)\cdot\frac{1}{n} = \frac{1}{2}.\end{aligned} \tag{4.2}$$

In order to be able to work with negative functions, the Choquet integral of such functions needs to be defined also for them. Below the asymmetric Choquet integral is defined, which is the classical extension from real-valued positive functions to negative functions. Note that symmetric extensions have gained an increasing interest [Greco et al., 2011; Mesiar et al., 2011], but I am not going to use them in this dissertation.

Definition 4.1.8 (Asymmetric Choquet integral for discrete negative functions)**.** *Let* $f : N \to (-\infty, 0]$ *be a function,* μ *a capacity on* N *and* $\bar{\mu}$ *its dual capacity. The asymmetric Choquet integral of* f *with respect to* μ *is defined by* $\mathcal{C}_\mu(f) := -\mathcal{C}_{\bar{\mu}}(-f)$.

Given the previous definition, the definition of the Choquet integral can now be extended to any function f from N to $\mathbb{R}$.

Definition 4.1.9 (Choquet integral for discrete functions)**.** *Let* μ *be a capacity on* N *and* f *a function from* N *to* $\mathbb{R}$*. It is denoted by* $f^+(m_i) = \max\{f(m_i), 0\}$ *and* $f^-(m_i) = \min\{f(m_i), 0\}$*. Then the Choquet integral of* f *with respect to* μ *is defined by*

$$\mathcal{C}_\mu(f) := \mathcal{C}_\mu\left(f^+\right) + \mathcal{C}_\mu\left(f^-\right) = \mathcal{C}_\mu\left(f^+\right) - \mathcal{C}_{\bar{\mu}}\left(-f^-\right). \tag{4.3}$$

4.2 THE RELATIONSHIP BETWEEN DISTORTION RISK MEASURES, OWA AND WOWA OPERATORS

Three results for discrete random variables are presented in this section. First, the equivalence between the Choquet integral and a distortion risk measure is shown, when the distortion risk measure is fixed on a finite probability space. Second, the link between this distortion risk measure and OWA operators is provided. And, third, the relationship between the fixed distortion risk measure and WOWA operators is given. Finally, it is shown that the local degree of orness of the VaR_α and $TVaR_\alpha$ risk measures may be defined as a function of the confidence level when the random variable is given. To my knowledge, some of these results provide a new insight into the way classical risk quantification is understood, which can now naturally be viewed as a weighted aggregation.

The link between the Choquet integral and distortion risk measures for arbitrary random variables is well-known since the inception of distortion risk measures [Wang, 1996], and has lead to many interesting results. For example, the concept of Choquet pricing and its associated equilibrium conditions [De Waegenaere et al., 2003]; the study of stochastic comparison of distorted variability measures [Sordo and Suarez-Llorens, 2011]; or the conditions for optimal behavioral insurance [Sung et al., 2011] and the analysis of competitive insurance markets in the presence of ambiguity [Anwar and Zheng, 2012]. Here the discrete version is presented, which is useful for the following presentation.

The relationship between the WOWA operator and the Choquet integral is also known by the fuzzy systems community [Torra, 1998], as well as the relationship between distorted probabilities and aggregation operators [Honda and Okazaki, 2005]. However, the results shown in this section provide a comprehensive presentation that allows for a connection to risk measurement.

Proposition 4.2.1. *Let $(N, 2^N, \mathcal{P})$ be a finite probability space, and let X be a discrete finite random variable defined on this space. Let $g : [0,1] \to [0,1]$ be a distortion function, and let ρ_g be the associated distortion risk measure. Then, it follows that*

$$\mathcal{C}_{g\circ\mathcal{P}}(X) = \rho_g(X).$$

Proof. Let $N = \{m_1, ..., m_n\}$ for some $n \geqslant 1$ and let me suppose that I can write $X(N) = \{x_1, ..., x_n\}$, with $X(\{m_i\}) = x_i$, and such that $x_i < x_j$ if $i < j$; additionally, let $k \in \{1, ..., n\}$ be such that $x_i < 0$ if $i = \{1, ..., k-1\}$ and $x_i \geqslant 0$ if $i = \{k, , ..., n\}$. In order to obtain the Choquet integral of X, a capacity μ defined on N is needed. As previously

indicated, $\mathcal{P}$ is a capacity on N that satisfies normalization, although it is not the one needed.

Since g is a distortion function, $\mu := g \circ \mathcal{P}$ is another capacity on N that satisfies normalization: $\mu(\emptyset) = g(\mathcal{P}(\emptyset)) = g(0) = 0$, $\mu(N) = g(\mathcal{P}(N)) = g(1) = 1$, and if $A \subseteq B$, the fact that $\mathcal{P}(A) \leqslant \mathcal{P}(B)$ and the fact that g is non-decreasing imply that $\mu(A) \leqslant \mu(B)$.

Regarding X^+, the permutation $\sigma = id$ on $(1, ..., k-1, k, ..., n)$ is such that $x^+_{\sigma(i)} \leqslant x^+_{\sigma(i+1)}$ for all i or, in other words, $x^+_1 \leqslant x^+_2 \leqslant ... \leqslant x^+_{k-1} \leqslant x^+_k \leqslant x^+_{k+1} \leqslant ... \leqslant x^+_n$. Then, $A_{\sigma,i} = \{m_i, ..., m_n\}$ and taking into account $x^+_i = 0 \; \forall i < k$, I can write $\mathcal{C}_{g\circ\mathcal{P}}(X^+)$ as

$$\begin{aligned} \mathcal{C}_{g\circ\mathcal{P}}(X^+) &= \sum_{i=1}^{n} \left(x^+_i - x^+_{i-1}\right) (g \circ \mathcal{P})(A_{\sigma,i}) \\ &= \sum_{i=k}^{n} \left(x^+_i - x^+_{i-1}\right) g\left(\sum_{j=i}^{n} p_j\right). \end{aligned} \tag{4.4}$$

Additionally, the permutation s on $(1, ..., k-1, k, ..., n)$ such that $s(i) = n + 1 - i$, satisfies $-x^-_{s(i)} \leqslant -x^-_{s(i+1)}$ for all i, so $-x^-_n \leqslant -x^-_{n-1} \leqslant ... \leqslant -x^-_k \leqslant -x^-_{k-1} \leqslant -x^-_{k-2} \leqslant ... \leqslant -x^-_1$.
I have $A_{s,i} = \left\{m_{s(i)}, ..., m_{s(n)}\right\} = \{m_{n+1-i}, ..., m_1\}$ and, therefore, $\bar{A}_{s,i} = \{m_{n+2-i}, ..., m_n\}$. Taking into account that $x^-_i = 0 \; \forall i \geqslant k$, I can write $\mathcal{C}_{\overline{g\circ\mathcal{P}}}(-X^-)$ as

$$\begin{aligned} \mathcal{C}_{\overline{g\circ\mathcal{P}}}(-X^-) &= \sum_{i=1}^{n} \left(-x^-_{s(i)} + x^-_{s(i-1)}\right) \left(\overline{g \circ \mathcal{P}}\right)(A_{s,i}) \\ &= \sum_{i=1}^{n} \left(-x^-_{n+1-i} + x^-_{n+2-i}\right) \left(\overline{g \circ \mathcal{P}}\right)(A_{s,i}) \\ &= \sum_{i=1}^{n} \left(-x^-_i + x^-_{i+1}\right) \left(\overline{g \circ \mathcal{P}}\right)(A_{s,n+1-i}) \\ &= \sum_{i=1}^{n} \left(-x^-_i + x^-_{i+1}\right) \left[1 - (g \circ \mathcal{P})\left(\bar{A}_{s,n+1-i}\right)\right] \\ &= \sum_{i=1}^{n} \left(-x^-_i + x^-_{i+1}\right) \left[1 - (g \circ \mathcal{P})(\{m_{i+1}, ..., m_n\})\right] \\ &= \sum_{i=1}^{k-1} \left(x^-_{i+1} - x^-_i\right) \left[1 - g\left(\sum_{j=i+1}^{n} p_j\right)\right]. \end{aligned} \tag{4.5}$$

Expressions (4.4) and (4.5) lead to

$$\begin{aligned}
\mathcal{C}_{g\circ\mathcal{P}}(X) = \;& \mathcal{C}_{g\circ\mathcal{P}}(X^+) - \mathcal{C}_{\overline{g\circ\mathcal{P}}}(-X^-) \\
= \;& -\sum_{i=1}^{k-1}\left(x_{i+1}^- - x_i^-\right)\left[1 - g\left(\sum_{j=i+1}^{n} p_j\right)\right] \\
& + \sum_{i=k}^{n}\left(x_i^+ - x_{i-1}^+\right) g\left(\sum_{j=i}^{n} p_j\right) \\
= \;& -\sum_{i=2}^{k}(x_i - x_{i-1})\left[1 - g\left(\sum_{j=i}^{n} p_j\right)\right] \\
& + x_k\left[1 - g\left(\sum_{j=k}^{n} p_j\right)\right] \\
& + \sum_{i=k+1}^{n}(x_i - x_{i-1})\, g\left(\sum_{j=i}^{n} p_j\right) + x_k g\left(\sum_{j=k}^{n} p_j\right) \\
= \;& -\sum_{i=2}^{k}(x_i - x_{i-1})\left[1 - g\left(\sum_{j=i}^{n} p_j\right)\right] \\
& + x_k + \sum_{i=k+1}^{n}(x_i - x_{i-1})\, g\left(\sum_{j=i}^{n} p_j\right).
\end{aligned} \tag{4.6}$$

Now consider $\rho_g(X)$ as in Section 1.2 of Chapter 1, and note that the random variable X is defined on the probability space $(N, 2^N, \mathcal{P})$. Given the properties of Riemann's integral, if I define $x_0 := -\infty$ and $x_{n+1} := +\infty$, then the distortion risk measure can be written as

$$\begin{aligned}
\rho_g(X) = \;& -\left[\sum_{i=1}^{k}\int_{x_{i-1}}^{x_i}[1 - g(S_X(x))]dx - \int_0^{x_k}[1 - g(S_X(x))]dx\right] \\
& + \int_0^{x_k} g(S_X(x))dx + \sum_{i=k+1}^{n+1}\int_{x_{i-1}}^{x_i} g(S_X(x))dx.
\end{aligned} \tag{4.7}$$

If $x \in [x_{i-1}, x_i)$ is considered, then $F_X(x) = \sum_{j=1}^{i-1} p_j$, since $F_X(x) = \mathcal{P}(X \leqslant x)$ and $S_X(x) = 1 - \sum_{j=1}^{i-1} p_j = \sum_{j=i}^{n} p_j$. Given that the distortion

function g is such that $g(0) = 0$ and $g(1) = 1$, expression (4.7) can be rewritten as

$$\begin{aligned}
\rho_g(X) = & -\sum_{i=1}^{k}\int_{x_{i-1}}^{x_i}\left[1-g\left(\sum_{j=i}^{n}p_j\right)\right]dx+\int_0^{x_k}\left[1-g\left(\sum_{j=k}^{n}p_j\right)\right]dx \\
& +\int_0^{x_0} g\left(\sum_{j=k}^{n}p_j\right)dx+\sum_{i=k+1}^{n+1}\int_{x_{i-1}}^{x_i} g\left(\sum_{j=i}^{n}p_j\right)dx \\
= & -\int_{-\infty}^{x_1}[1-g(1)]\,dx-\sum_{i=2}^{k}\int_{x_{i-1}}^{x_i}\left[1-g\left(\sum_{j=i}^{n}p_j\right)\right]dx \\
& +\int_0^{x_k}\left[1-g\left(\sum_{j=k}^{n}p_j\right)\right]dx+\int_0^{x_k} g\left(\sum_{j=k}^{n}p_j\right)dx \\
& +\sum_{i=k+1}^{n}\int_{x_{i-1}}^{x_i} g\left(\sum_{j=i}^{n}p_j\right)dx+\int_{x_n}^{+\infty} g(0)\,dx \\
= & -\sum_{i=2}^{k}(x_i-x_{i-1})\left[1-g\left(\sum_{j=i}^{n}p_j\right)\right] \\
& +x_k\left[1-g\left(\textstyle\sum_{j=k}^{n}p_j\right)+g\left(\textstyle\sum_{j=k}^{n}p_j\right)\right] \\
& +\sum_{i=k+1}^{n}(x_i-x_{i-1})\,g\left(\sum_{j=i}^{n}p_j\right) \\
= & -\sum_{i=2}^{k}(x_i-x_{i-1})\left[1-g\left(\sum_{j=i}^{n}p_j\right)\right] \\
& +x_k+\sum_{i=k+1}^{n}(x_i-x_{i-1})\,g\left(\sum_{j=i}^{n}p_j\right).
\end{aligned} \tag{4.8}$$

And then the proof is finished because $\rho_g(X) = \mathcal{C}_{g\circ\mathcal{P}}(X)$ using (4.8) and (4.6). □

Let me present $\mathcal{C}_{g\circ\mathcal{P}}(X)$ in a more compact form. I denote $F_{i-1} = 1-g\left(\sum_{j=i}^{n}p_j\right)$ and $S_{i-1} = g\left(\sum_{j=i}^{n}p_j\right)$ for $i = 1,\ldots,n+1$, so $F_{i-1} = 1-S_{i-1}$. Note that $F_0 = 0$ and $S_n = 0$, so

$$\sum_{i=2}^{k}(x_{i-1}-x_i)\,F_{i-1} = \sum_{i=1}^{k-1}x_i\,(F_i-F_{i-1})-x_kF_{k-1},$$

and

$$\sum_{i=k+1}^{n}(x_i-x_{i-1})\,S_{i-1} = \sum_{i=k+1}^{n}x_i\,(S_{i-1}-S_i)-x_kS_k.$$

The previous expressions applied to $\mathcal{C}_{g\circ\mathcal{P}}(X)$ lead to

$$\begin{aligned}\mathcal{C}_{g\circ\mathcal{P}}(X) = & \sum_{i=1}^{k-1} x_i (F_i - F_{i-1}) - x_k F_{k-1} + x_k \\ & + \sum_{i=k+1}^{n} x_i (S_{i-1} - S_i) - x_k S_k \\ = & \sum_{i=1}^{n} x_i (S_{i-1} - S_i) \\ = & \sum_{i=1}^{n} x_i \left[g\left(\sum_{j=i}^{n} p_j\right) - g\left(\sum_{j=i+1}^{n} p_j\right)\right].\end{aligned} \tag{4.9}$$

It has to be noted that a similar expression is used by Kim [2010] as an empirical estimate of the distortion risk measure, where the probabilities are obtained from the empirical distribution function. If $g = id$, then $\rho_{id}(X) = \mathbb{E}(X)$. The same result for a continuous random variable is easy to prove using the definition of distortion risk measure and Fubinni's theorem. Expression (4.9) is useful to prove the following two propositions.

Proposition 4.2.2 (OWA equivalence to distortion risk measures)**.** *Let X be a discrete finite random variable and $\left(N, 2^N, \mathcal{P}\right)$ be a probability space as defined in proposition 4.2.1. Let ρ_g be a distortion risk measure defined in this probability space, and let p_j be the probability of x_j for all j. Then there exist a unique $OWA_{\vec{w}}$ operator such that $\rho_g(X) = OWA_{\vec{w}}(\vec{x})$. The OWA operator is defined by weights*

$$w_i = g\left(\sum_{j=i}^{n} p_j\right) - g\left(\sum_{j=i+1}^{n} p_j\right). \tag{4.10}$$

The proof is straightforward. From proposition 4.2.2 it follows that a finite and discrete random variable X must be fixed to obtain a one-to-one equivalence between a distortion risk measure and an OWA operator.

Proposition 4.2.3 (WOWA equivalence to distortion risk measures)**.** *Let X be a discrete finite random variable and $\left(N, 2^N, \mathcal{P}\right)$ be a probability space as in proposition 4.2.1. If ρ_g is a distortion risk measure defined on this probability space, and p_j is the probability of x_j for all j, consider the WOWA operator such that $h = g$, $\vec{q} = \vec{p}$ and $v_i = g\left(\frac{n-i+1}{n}\right) - g\left(\frac{n-i}{n}\right)$ for all $i = 1, \ldots, n$. Then*

$$\rho_g(X) = WOWA_{g,\vec{v},\vec{p}}(\vec{x}). \tag{4.11}$$

Proof. Using proposition 4.2.2 it is known that there exists a unique $\vec{w} \in [0,1]^n$ such that $OWA_{\vec{w}}(\vec{x}) = \rho_g(X)$:

$$w_i = g\left(\sum_{j=i}^{n} p_j\right) - g\left(\sum_{j=i+1}^{n} p_j\right) = g(S_X(x_{i-1})) - g(S_X(x_i)). \tag{4.12}$$

In addition, there exists an $OWA_{\vec{u}}$ operator such that $OWA_{\vec{u}} = WOWA_{g,\vec{v},\vec{p}}$ defined by

$$\begin{aligned} u_i &= g\left(\sum_{\Omega_j \in A_{id,i}} p_j\right) - g\left(\sum_{\Omega_j \in A_{id,i+1}} p_j\right) \\ &= g(S_X(x_{i-1})) - g(S_X(x_i)). \end{aligned} \tag{4.13}$$

Expressions (4.12) and (4.13) show that $\vec{w} = \vec{u}$ and, due to the uniqueness of the OWA operator, it is concluded that

$$\rho_g(X) = OWA_{\vec{w}}(\vec{x}) = WOWA_{g,\vec{v},\vec{p}}(\vec{x}),$$

where $v_i = g\left(\frac{n-i+1}{n}\right) - g\left(\frac{n-i}{n}\right)$. □

Again, the one-to-one equivalence between a distortion risk measure and a WOWA operator is obtained given that the discrete and finite random variable is fixed.

To summarize the results, for a given distortion function g and a discrete and finite random variable X, there are three alternative ways to calculate the distortion risk measure that lead to the same result than using the definition provided in Section 1.2 of Chapter 1:

1. By means of the Choquet integral of X with respect to $\mu = g \circ \mathcal{P}$ using expression (4.9).
2. Applying the $OWA_{\vec{w}}$ operator to $\vec{x}$, following definition 4.1.2 with $w_i = g\left(\sum_{j=i}^{n} p_j\right) - g\left(\sum_{j=i+1}^{n} p_j\right)$, $i = 1, ..., n$, and p_j the probability of x_j for all j.
3. And, finally, applying the $WOWA_{g,\vec{v},\vec{p}}$ operator to $\vec{x}$, following definition 4.1.4, where $v_i = g\left(\frac{n-i+1}{n}\right) - g\left(\frac{n-i}{n}\right)$ and p_j the probability of x_j for all j.

4.3 FIRST INTERPRETATIONS OF THE LOCAL DEGREE OF ORNESS

An interesting application from expression (4.9) can be derived. In particular, the concept of degree of orness introduced for the OWA operator may be formally extended to the case of $\mathcal{C}_{g\circ\mathcal{P}}(X)$, as:

$$\boldsymbol{\omega}\left(\mathcal{C}_{g\circ\mathcal{P}}(X)\right) := \sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)\cdot\left[g\left(S_X\left(x_{i-1}\right)\right)-g\left(S_X\left(x_i\right)\right)\right]. \quad (4.14)$$

This will be called *local degree of orness* of $\mathcal{C}_{g\circ\mathcal{P}}(X)$. Note that this expression is similar to (4.1). This result is now applicable to both positive and negative values and only the distorted probabilities are considered among capacities.

Let me show risk management applications of the local degree of orness of the distortion risk measures. Note, for instance, that the regulatory requirements on risk measurement based on distortion risk measures may be reinterpreted by means of the local degree of orness. Given a finite and discrete random variable X, when the value $\rho_g(X)$ of the distortion risk measure applied to it is required there is an implicit *preference weighting rule* with respect to the values of X, which takes into account probabilities. This preference weighting rule can be summarized by $\boldsymbol{\omega}(OWA_{\vec{w}})$, where $\vec{w}$ is such that $w_i = g\left(S_X\left(x_{i-1}\right)\right)-g\left(S_X\left(x_i\right)\right)$.

There are some cases of special interest, such as the mathematical expectation, the VaR_α and $TVaR_\alpha$ risk measures:

- If $g = id$, then $\mathcal{C}_{g\circ\mathcal{P}} \equiv \mathbb{E}$ and

$$\begin{aligned}\boldsymbol{\omega}\left(\mathbb{E}(X)\right) &= \sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)\cdot\left[S_X\left(x_{i-1}\right)-S_X\left(x_i\right)\right]\\ &= \sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)\cdot p_i.\end{aligned} \quad (4.15)$$

  In particular, if the random variable X is discrete and uniform, i.e. $p_i = \frac{1}{n}$, then expression (4.15) equals $1/2$.

Given a confidence level $\alpha \in (0,1)$, let $k_\alpha \in \{1,2,...,n\}$ be such that $x_{k_\alpha} = \inf\{x_i | F_X(x_i) \geqslant \alpha\} = \inf\{x_i | S_X(x_i) \leqslant 1-\alpha\}$, i.e. x_{k_α} is the $\alpha-$quantile of X.

- Regarding VaR_α, from Table 1.1 it is known that $\psi_\alpha(S_X(x_i)) = \mathbb{1}[1-\alpha \leqslant S_X(x_i) \leqslant 1]$. Since $\psi_\alpha(S_X(x_{i-1})) - \psi_\alpha(S_X(x_i)) = \mathbb{1}[i = k_\alpha]$, the local degree of orness of VaR_α is obtained as

$$\begin{aligned}\omega(VaR_\alpha(X)) &= \sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)\cdot[\psi_\alpha(S_X(x_{i-1})) - \psi_\alpha(S_X(x_i))] \\ &= \frac{k_\alpha - 1}{n-1}.\end{aligned} \tag{4.16}$$

- In the case of $TVaR_\alpha$, recall from Table 1.1 that $\gamma_\alpha(S_X(x_i)) = \min\left\{\frac{S_X(x_i)}{1-\alpha}, 1\right\}$. Taking into account that

$$\gamma_\alpha(S_X(x_{i-1})) - \gamma_\alpha(S_X(x_i)) = \begin{cases} 0 & i < k_\alpha \\ 1 - \dfrac{1}{1-\alpha}\displaystyle\sum_{j=k_\alpha+1}^{n} p_j & i = k_\alpha \\ \dfrac{p_i}{1-\alpha} & i > k_\alpha. \end{cases},$$

therefore

$$\begin{aligned}\omega(TVaR_\alpha(X)) &= \sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)\cdot[\gamma_\alpha(S_X(x_{i-1})) - \gamma_\alpha(S_X(x_i))] \\ &= \left(\frac{k_\alpha - 1}{n-1}\right)\cdot\left[1 - \frac{1}{1-\alpha}\sum_{j=k_\alpha+1}^{n} p_j\right] \\ &\quad + \sum_{i=k_\alpha+1}^{n}\left(\frac{i-1}{n-1}\right)\cdot\frac{p_i}{1-\alpha} \\ &= \frac{k_\alpha - 1}{n-1} + \frac{1}{1-\alpha}\cdot\sum_{i=k_\alpha+1}^{n}\left(\frac{i-k_\alpha}{n-1}\right)p_i.\end{aligned} \tag{4.17}$$

Note that for VaR_α and $TVaR_\alpha$, the local degree of orness is directly connected to the α level chosen for the risk measure, i.e. the value of the distribution function at the point given by the quantile. In the following example an application of the local degree of orness in the context of risk measurement is presented.

4.4 AN ILLUSTRATION

A numerical example taken from Wang [2002] is provided. This example is selected as a particular case where common risk measures show drawbacks in the comparison of two random variables, X and Y. Table 4.2 summarizes the probabilities, distribution functions and

Table 4.2: Example of loss random variables X and Y.

Loss	p_x	F_X	S_X	p_y	F_Y	S_Y
0	0.6	0.6	0.4	0.6	0.6	0.4
1	0.375	0.975	0.025	0.39	0.99	0.01
5	0.025	1	0			
11				0.01	1	0

survival functions of both random variables.

Distortion risk measures for X and Y using aggregation operators can be calculated. In particular, I am interested in $\mathbb{E}$, VaR_α and $TVaR_\alpha$ for $\alpha = 95\%$, which follow from expression (4.9) and ψ_α and γ_α as in Table 1.1. In this example $\mathbb{E}$, $VaR_{95\%}$ and $TVaR_{95\%}$ have the same value for the two random variables.

The weighting vectors linked to the OWA operators (see expression 4.10) for $\mathbb{E}$, $VaR_{95\%}$ and $TVaR_{95\%}$ are displayed in Table 4.3. The values of the distortion risk measures for each random variable and the associated local degree of orness are shown in Table 4.4. In addition, the weighting vectors linked to the WOWA operators (see expression 4.11) are listed in Table 4.5.

Table 4.3: Distorted probabilities in the ordered weighted averaging operators for X and Y ($\vec{w}$).

	$\mathbb{E}(X)$	$\mathbb{E}(Y)$	$VaR_{95\%}(X)$	$VaR_{95\%}(Y)$	$TVaR_{95\%}(X)$	$TVaR_{95\%}(Y)$
Loss	$\vec{w}$	$\vec{w}$	$\vec{w}$	$\vec{w}$	$\vec{w}$	$\vec{w}$
0	0.6	0.6	0	0	0	0
1	0.375	0.39	1	1	0.5	0.8
5	0.025		0		0.5	
11		0.01		0		0.2

First, note that point probabilities are distorted and a weighted average of the random values with respect to this distortion ($OWA_{\vec{w}}$) is calculated to obtain the distortion risk measures. Second, the results show that weights $\vec{v}$ for the WOWA represent, to some extent, a risk attitude. It is taken into account how the random variable is distributed by means of weights $\vec{p}$. In this example, the decision makers are only worried about the maximum loss when they consider $VaR_{95\%}$ and $TVaR_{95\%}$. All values have the same importance in the case of the mathematical expectation.

Table 4.4: Distortion risk measures and the associated local degree of orness for X and Y.

	$\mathbb{E}(X)$	$\mathbb{E}(Y)$	$VaR_{95\%}(X)$	$VaR_{95\%}(Y)$	$TVaR_{95\%}(X)$	$TVaR_{95\%}(Y)$
Risk value	0.5	0.5	1	1	3	3
Degree of orness	0.2125	0.205	0.5	0.5	0.75	0.6

Table 4.5: Weighted ordered weighted averaging operator vectors linked to distortion risk measures for X and Y.

	$\mathbb{E}(X)$		$\mathbb{E}(Y)$		$VaR_{95\%}(X)$		$VaR_{95\%}(Y)$	
Loss	$\vec{p}$	$\vec{v}$	$\vec{p}$	$\vec{v}$	$\vec{p}$	$\vec{v}$	$\vec{p}$	$\vec{v}$
0	0.6	1/3	0.6	1/3	0.6	0	0.6	0
1	0.375	1/3	0.39	1/3	0.375	0	0.39	0
5	0.025	1/3			0.025	1		
11			0.01	1/3			0.01	1

	$TVaR_{95\%}(X)$		$TVaR_{95\%}(Y)$	
Loss	$\vec{p}$	$\vec{v}$	$\vec{p}$	$\vec{v}$
0	0.6	0	0.6	0
1	0.375	0	0.39	0
5	0.025	1		
11			0.01	1

Note that $VaR_{95\%}$ and $TVaR_{95\%}$ have equal $\vec{v}$ and $\vec{p}$ for each random variable, although the distortion risk measures have different values. It is due to the fact that function h in WOWA plays an important role to determine the particular distortion risk measure that is calculated, since function h is the distortion function for VaR_α and $TVaR_\alpha$.

Finally, it is interesting to note that the local degree of orness of a distortion risk measure can be understood as another risk measure for the random variable, with a value that belongs to $[0, 1]$. In this particular illustration, the additional riskiness information provided by the local degree of orness can be summarized as follows:

- It is shown that $\omega(\mathbb{E}(X)) \neq \omega(\mathbb{E}(Y))$, and both are less than 0.5. Note that 0.5 is the local degree of orness of the mathematical expectation of an uniform random variable. The greater the difference (in absolute value) between the local degree of orness of the mathematical expectation and 0.5, the greater the difference between the random variable and an uniform. In the example, both random variables are far from a discrete uniform, but Y is farther than X;
- The $\omega(VaR_{95\%}(X))$ is equal to $\omega(VaR_{95\%}(Y))$, because the number of observations is the same and $VaR_{95\%}$ is located at the same position for both variables;
- The local degree of orness of $TVaR_{95\%}$ is different for both random variables, although they have the same value for the $TVaR_{95\%}$. Given these two random variables with the same number of observations, $VaR_{95\%}$, local orness of $VaR_{95\%}$ and $TVaR_{95\%}$, more extreme losses are associated to the random variable with the lower local degree of orness of $TVaR_{95\%}$. Therefore, this additional information provided by the local degree of orness may be useful to compare X and Y, given that they are indistinguishable in terms of $\mathbb{E}$, $VaR_{95\%}$ and $TVaR_{95\%}$.

More details on the risk information that local degree of orness provides are given in Chapter 6.

5

INDICATORS FOR DISCRETE CHOQUET INTEGRALS

Aggregation operators are very useful functions for summarizing information and have been widely used in recent decades [Beliakov et al., 2007; Grabisch et al., 2009; Torra and Narukawa, 2007]. In this context, the Choquet integral [Choquet, 1954], a class of integral linked to non-additive measures, has taken a leading role. Integrals are used to aggregate values of functions, and as such can be understood as aggregation operators. The Choquet integral includes a wide range of aggregation operators as particular cases. Over the last few years, the Choquet integral has received much attention from researchers, and this has generated new extensions and generalizations of this class of integral. For instance, Greco et al. [2011] proposed an extension of the Choquet integral in which the capacity depends on the values to be aggregated. Similarly, Yager [2004a] presented new induced aggregation operators inspired by the Choquet integral and Xu [2010] introduced some intuitionistic fuzzy aggregation functions also based on the Choquet integral. Klement et al. [2010] presented a universal integral that covers the Choquet and the Sugeno integral for non-negative functions, while Torra and Narukawa [2010] studied a generalization of the Choquet integral inspired by the Losonczi mean. Bolton et al. [2008] connected the Choquet integral with distance metrics and, more recently, Torra and Narukawa [2012] introduced an operator that generalizes the Choquet integral and the Mahalanobis distance.

Two particular cases of aggregation operators that can be understood as Choquet integrals are the weighted arithmetic mean (WAM) and the ordered weighted averaging (OWA) operator [Yager, 1988]. As pointed out in Chapter 4, several authors have turned their attention to the study of the OWA operator [Yager et al., 2011], since it serves to provide a parameterized family of aggregation operators between the minimum and the maximum. In order to assess OWA operators appropriately, indicators for characterizing the weighting vector are required. Initially, Yager [1988] introduced the orness/andness indicators and the entropy of dispersion for just this purpose. Later, he proposed complementary indicators, including the balance indicator [Yager, 1996] and the divergence [Yager, 2002], to be used in exceptional situations. Meanwhile, Fullér and Majlender [2003] suggested the use of a variance indicator and Majlender [2005] introduced the Rényi entropy [Rényi, 1961] as a generalization of the Shannon en-

tropy [Shannon, 1948] in the framework of the OWA operator. Some of these indicators have been extended for the Choquet integral. For example, Marichal [2004] and Grabisch et al. [2009] presented several types of degree of orness indicators: Marichal [2004] for the Choquet integral, while Grabisch et al. [2009] for general aggregation functions. Likewise, Yager [2000], Marichal [2002] and Kojadinovic et al. [2005] studied the entropy of dispersion in the framework of the Choquet integral. Marichal and Roubens [2000] analyzed the relationship between the alternative definitions of the entropy of dispersion indicator introduced by Yager [2000] and Marichal [2002]. However, to the best of my knowledge, additional indicators have yet to be defined for the Choquet integral.

The aim of this chapter is to further enrich the present set of indicators for the Choquet integral, by incorporating new ones to earlier contributions and by presenting an unified compilation of indicators for describing its aggregation features. Hopefully, these indicators may be helpful to enrich our vision on distortion risk measures. Four indicators commonly used for the OWA operator -that is, the degree of balance, the divergence, the variance indicator and Rényi entropies- are extended to the discrete Choquet integral. The advantage of incorporating these additional indicators is that they can help to cover a wide range of situations, including exceptional types of aggregation that cannot be correctly identified by means of the degree of orness or the entropy of dispersion. Two different perspectives are considered so as to allow both local and global indicators to be defined.

The linearity of indicators is investigated when dealing with linear combinations of capacities. Indicators are presented for the probabilistic OWA (POWA) operator [Merigó, 2011, 2012], which deals with a linear combination of two particular cases of the Choquet integral (the OWA and the WAM) in order to obtain more complex aggregations. The importance of these two aggregation operators is determined by the particular weight assigned to them in the linear combination.

5.1 AVERAGING OPERATORS AND CAPACITIES

Let $\vec{w}$ and $\vec{p}$ be two vectors with components belonging to $[0,1]$ such that $\sum_{i=1}^{n} w_i = 1$ and $\sum_{i=1}^{n} p_i = 1$, and consider the aggregation operators $\mathrm{OWA}_{\vec{w}}$ and $\mathrm{WAM}_{\vec{p}}$. The representation of OWA and WAM operators as Choquet integrals has been shown in the literature [Grabisch, 1995; Grabisch et al., 2011]. Proposition 10(v) and Proposition 10(vi) in Grabisch et al. [2011] imply that OWA and WAM operators can be understood as Choquet integrals with respect to normal-

ized capacities μ and $\mathcal{P}$ respectively, $\text{OWA}_{\vec{w}} = \mathcal{C}_\mu$ and $\text{WAM}_{\vec{p}} = \mathcal{C}_\mathcal{P}$. These capacities are such that:

- $\mu(A) = \sum_{u=0}^{i-1} w_{n-u}$, for all $A \in N$ with cardinality i ($|A| = i$), $i = 1, ..., n$. Because of $|A_{\sigma,i}| = n - i + 1$, then $\mu(A_{\sigma,i}) = \sum_{k=i}^{n} w_k$ for all $i = 1, ..., n$, being σ a permutation as in the definition of $\text{OWA}_{\vec{w}}$;
- $\mathcal{P}(\{m_i\}) = p_i$ for all $i = 1, ..., n$, being $\mathcal{P}$ additive. That is, the probability $\mathcal{P}$ understood as an additive capacity on N.

5.2 INDICATORS FOR AGGREGATION OPERATORS

5.2.1 *Indicators associated with OWA operators*

Various indicators associated with OWA operators can be found in the literature and the main ones are briefly explained here. A summary of these indicators, their analytical expressions and references are shown in Table 5.1.

Table 5.1: Summary of indicators associated with ordered weighted averaging operators

Indicator	Analytical expression	Reference
Degree of orness	$\omega(\vec{w}) = \sum_{i=1}^{n} \left(\frac{i-1}{n-1}\right) \cdot w_i$	Yager [1988]
Dispersion (Shannon entropy)[1]	$\text{Disp}(\vec{w}) = -\sum_{i=1}^{n} \ln(w_i) \cdot w_i$	Yager [1988]
Degree of balance	$\text{Bal}(\vec{w}) = \sum_{i=1}^{n} \left(\frac{2 \cdot i - (n+1)}{n-1}\right) \cdot w_i$	Yager [1996]
Divergence	$\text{Div}(\vec{w}) = \sum_{i=1}^{n} \left(\frac{i-1}{n-1} - \omega(\vec{w})\right)^2 \cdot w_i$	Yager [2002]
Variance indicator	$D^2(\vec{w}) = \frac{1}{n} \cdot \sum_{i=1}^{n} w_i^2 - \frac{1}{n^2}$	Fullér and Majlender [2003]
Rényi entropy ($\alpha \neq 1$)	$H_\alpha(\vec{w}) = \frac{1}{1-\alpha} \cdot \log_2 \left(\sum_{i=1}^{n} w_i^\alpha\right)$	Majlender [2005]

[1] If the Shannon entropy of $\vec{w}$ is denoted by $H_S(\vec{w}) = -\sum_{i=1}^{n} \log_2(w_i) \cdot w_i$, then $\text{Disp}(\vec{w}) = \ln(2) \cdot H_S(\vec{w})$.

Degree of orness

The definition of this indicator has been previously provided in Chapter 4 (see Definition 4.1.3). The degree of orness of an $OWA_{\vec{w}}$ operator was defined in Yager [1988] as representing the level of aggregation preference between the minimum and the maximum operators given by $\vec{w} \in [0,1]^n$. Therefore, the degree of orness highlights how much disjunctive is the aggregation function. The degree of orness of $OWA_{\vec{w}}$ can be understood as the value that the $OWA_{\vec{w}}$ operator returns when it is applied to $\vec{x^*} = \left(\frac{0}{n-1}, \frac{1}{n-1}, ..., \frac{n-2}{n-1}, \frac{n-1}{n-1}\right)$ or, alternatively, as the value of $WAM_{\vec{w}}\left(\vec{x^*}\right)$.

Dispersion (Shannon entropy)

The dispersion indicator of an $OWA_{\vec{w}}$ operator was introduced by Yager [1988] to measure the amount of information given by $\vec{x}$ that is used when $OWA_{\vec{w}}(\vec{x})$ is computed. This indicator provides the same information as the entropy introduced by Shannon [1948] but at a different scale, as shown in Table 5.1.

Degree of balance

The concept of degree of balance of an $OWA_{\vec{w}}$ operator was introduced by Yager [1996] and it is closely related to the degree of orness, providing the same information but at a different scale. It is defined as $Bal(\vec{w}) = 2 \cdot \omega(\vec{w}) - 1$.

Divergence indicator

The divergence indicator of an $OWA_{\vec{w}}$ operator was introduced by Yager [2002]. The definition is provided in Table 5.1. Note that it can be understood to be the value of the $WAM_{\vec{w}}$ applied to $\vec{z} = (z_1, z_2, ..., z_n)$ where $z_i = \left(\frac{i-1}{n-1} - \omega(\vec{w})\right)^2$ for all $i = 1, ..., n$. In general, $z_i \leqslant z_j$ does not hold if $i \leqslant j$. That is, components of $\vec{z}$ are not in increasing order, so divergence indicator $Div(\vec{w})$ cannot be expressed as $OWA_{\vec{w}}(\vec{z})$.

Recall that the variance of a random variable X with respect to a probability $\mathcal{P}$ is $Var_{\mathcal{P}}(X) := \mathbb{E}_{\mathcal{P}}\left[(X - \mathbb{E}_{\mathcal{P}}(X))^2\right]$, where $\mathbb{E}_{\mathcal{P}}(X)$ denotes the mathematical expectation of random variable X with respect to probability $\mathcal{P}$. In the discrete and finite case, $\mathbb{E}_{\mathcal{P}}(X) = \sum_{i=1}^{n} x_i \cdot p_i$ and $Var_{\mathcal{P}}(X) = \sum_{i=1}^{n} \left(x_i - \sum_{i=1}^{n} x_i \cdot p_i\right)^2 \cdot p_i$. Bearing this in mind and from a statistical viewpoint, if the random variable X^* is considered with $x_i^* = \frac{i-1}{n-1}$ and the probabilities $\mathcal{P}\left(X^* = x_i^*\right)$ are equal to w_i for all $i = 1, ..., n$, then $Div(\vec{w})$ is just the variance of the random variable X^* with respect to the probabilities $\vec{p}$ when the lat-

ter are equal to the weights $\vec{w}$, i.e. $\vec{p} = \vec{w}$. In other words, the divergence indicator can be understood as $Div(\vec{w}) = Var_{\vec{w}}(X^*) = \mathbb{E}_{\vec{w}}\left[(X^*)^2\right] - (\mathbb{E}_{\vec{w}}[X^*])^2$.

The divergence indicator describes how much scattered and dispersed the evaluation supplied by the aggregation function is. Therefore, the divergence indicator highlight different aspects of the aggregation function at hand than those aspects pointed out by the degree of orness and the dispersion indicators. As Yager [2002] claimed when analyzing the OWA operator, the degree of orness and the dispersion indicator are insufficient for characterizing a weighting vector $\vec{w}$. For example, OWA operators with different weighting vectors can share the values of the degree of orness and the dispersion indicator. Let us consider two vectors in $\mathbb{R}^9$, $\vec{w} = (0, 0.5, 0, 0, 0, 0, 0, 0.5, 0)$ and $\vec{w^*} = (0, 0, 0, 0.5, 0, 0.5, 0, 0, 0)$. An analysis of the degree of orness and the dispersion of $OWA_{\vec{w}}$ and $OWA_{\vec{w^*}}$ provides the same results: $\omega(\vec{w}) = \omega(\vec{w^*}) = 0.5$ and $Disp(\vec{w}) = Disp(\vec{w^*}) = 0.693$. Thus, in order to distinguish between $OWA_{\vec{w}}$ and $OWA_{\vec{w^*}}$ operators, an additional measure is required. By using the divergence indicator (Table 5.1), such a distinction can be achieved. In this particular example, $Div(\vec{w}) = 0.140625$ and $Div(\vec{w^*}) = 0.015625$. Thus, although $OWA_{\vec{w}}$ and $OWA_{\vec{w^*}}$ present identical degrees of orness and dispersion, the latter has a lower divergence than the former.

Variance indicator

The variance indicator is defined as $D^2(\vec{w}) = \frac{1}{n} \cdot \sum_{i=1}^{n} w_i^2 - \frac{1}{n^2}$. This indicator computes the variance of the weighting vector $\vec{w}$ where each component is considered equally probable. It has been used, for instance, in Fullér and Majlender [2003] to determine the analytical expression of a minimum variability $OWA_{\vec{w}}$ operator.

Rényi entropies

Entropy measures other than dispersion can be used to characterize the weighting vector. Generalizations of the Shannon entropy that could be used include Rényi entropies [Majlender, 2005; Rényi, 1961]. Recall that the Rényi entropy of $\vec{w} \in \mathbb{R}^n$ with degree $\alpha \in \mathbb{R}\backslash\{1\}$ is defined as $H_\alpha(\vec{w}) = \frac{1}{1-\alpha} \cdot \log_2\left(\sum_{i=1}^{n} w_i^\alpha\right)$. Thus, given the $OWA_{\vec{w}}$, $H_\alpha(\vec{w})$ can be considered as the Rényi entropy of degree α of this OWA operator. The Shannon entropy can be obtained from Rényi entropies as $H_S(\vec{w}) = \lim_{\alpha \to 1} H_\alpha(\vec{w})$: this result can be proved using l'Hôpital rule.

The parameter α, when positive, can be related to the underlying $\mathcal{L}^\alpha$-norm linked to each H_α, as follows $H_\alpha(\vec{w}) = \frac{\alpha}{1-\alpha} \cdot \log_2(\|\vec{w}\|_\alpha)$. Therefore, the parameter α can be interpreted to distort the distance between the vector $\vec{w}$ and $\vec{0} \in \mathbb{R}^n$, a distance required to compute the entropy of $\vec{w}$. Rényi entropies are not only a mathematical generalization of the Shannon entropy. Applications of Rényi entropies are found in several fields of knowledge, including coding theory, statistical physics and multifractal systems [Csiszár, 1995; Lenzi et al., 2000; Jizba and Arimitsu, 2004].

5.2.2 *Existing indicators extended to the Choquet integral*

Some of the indicators described above have already been generalized for the discrete Choquet integral, particularly, the degree of orness and the dispersion indicator (Shannon entropy). The purpose of this chapter is to propose indicators that have not yet been defined for the Choquet integral. However, the existing indicators are here described in order to provide a complete compilation of indicators for identifying features linked to the Choquet integral. Hereinafter, the indicators are considered from two perspectives, the global and the local. Global indicators involve the computation of the $n!$ permutations of $(1, 2, ..., n)$ while local indicators take into account only one of the $n!$ permutations. Broadly speaking, a global indicator does not depend on the assumption of ordering on the input data set to be aggregated by the Choquet integral while a local one does. Local indicators are an interesting alternative to global indicators, particularly appealing for applications in which the input data set to be aggregated is increasingly ordered. This is the case, for instance, of statistical analysis of empirical distribution functions. A practical consequence is that the computation of local indicators is in general easier, something especially important whether the number of elements of N is large.

The terminology is inspired by Dujmović [2006], who proposes a classification of orness indicators by means of a three-letter code. I extend this categorization to all of the indicators but my classification is exclusively based on whether ordering in the input data set is assumed or not. This being the case, the categories can be determined solely by one letter. So here the categories are seen as global and local, respectively, and denoted as G and L. In other words, the last two letters of the codes used in the classification of Dujmović are common to both categories in this context, which means that I only consider direct indicators that depend on the number of variables.

A similar terminology is also adopted by Kolesárová and Mesiar [2009] who explain the meaning that they provide to both global and

local perspectives and, in addition, introduce a generalized characterization which they refer to as the 'mixed approach'.

Degree of orness for the Choquet integral

A generalization of the global degree of orness for the Choquet integral has been proposed by Marichal [2004]. As shown in expression (1.2), if $\mathcal{C}_\mu$ is the Choquet integral with respect to μ, then

$$\omega_G(\mathcal{C}_\mu) = \frac{1}{n-1} \cdot \sum_{i=1}^{n-1} \left[\binom{n}{i}^{-1} \cdot \sum_{\substack{A \subseteq N \\ |A|=i}} \mu(A) \right]. \tag{5.1}$$

Likewise, a local degree of orness for a Choquet integral has been suggested in chapter 4 by means of expression (4.1),

$$\omega_L(\mathcal{C}_\mu) = \sum_{i=1}^{n} \left(\frac{i-1}{n-1} \right) \cdot (\mu(A_{id,i}) - \mu(A_{id,i+1})).$$

This expression can be rearranged to obtain

$$\omega_L(\mathcal{C}_\mu) = \frac{1}{n-1} \cdot \sum_{i=1}^{n-1} [\mu(A_{id,n-i+1})]. \tag{5.2}$$

Differences between expression (5.1) and (5.2) are highlighted when noting that the number of elements of the set $A_{id,n-i+1}$ is equal to i, for all $i = 1, ..., n-1$. Expression (5.1) takes into account all permutations on $(1, 2, ..., n)$, so an ordering on function values to be aggregated through the Choquet integral is not assumed when computing $\omega_G(\mathcal{C}_\mu)$. Therefore, $\omega_G(\mathcal{C}_\mu)$ is called global. On the other hand, it is implicitly assumed that function values to be aggregated through the Choquet integral are increasingly ordered when computing $\omega_L(\mathcal{C}_\mu)$ and, consequently, $\omega_L(\mathcal{C}_\mu)$ is called local.

The idea underpinning this local generalization is to transfer to the Choquet integral the fact that the degree of orness of $OWA_{\vec{w}}$ can be understood as the value that the $OWA_{\vec{w}}$ operator returns when it is applied to $\vec{x^*} = \left(\frac{0}{n-1}, \frac{1}{n-1}, ..., \frac{n-2}{n-1}, \frac{n-1}{n-1}\right)$ and, at the same time, as the value of $WAM_{\vec{w}}(\vec{x^*})$. When considering a Choquet integral with respect to a normalized symmetric capacity μ (that is, when dealing with OWA operators), the local and global degrees of orness are equal, i.e. $\omega_G(\mathcal{C}_\mu) = \omega_L(\mathcal{C}_\mu)$. On the other hand, if μ is normalized and additive ($\mathcal{C}_\mu = WAM_{\vec{p}}$ with $p_i = \mu(\{m_i\})$ for all $i = 1, ..., n$), it is straightforward to prove that $\omega_G(\mathcal{C}_\mu) \neq \omega_L(\mathcal{C}_\mu)$. The difference derives from the fact that $\omega_L(\mathcal{C}_\mu)$ only takes into account one of the $n!$ feasible permutations of $(1, 2, ..., n)$ - the identity permutation - while $\omega_G(\mathcal{C}_\mu)$ considers them all. In order to simplify the notation,

hereinafter, $\omega_L(\mu)$ and $\omega_G(\mu)$ will be used instead of $\omega_L(\mathcal{C}_\mu)$ and $\omega_G(\mathcal{C}_\mu)$, respectively.

Alternative generalizations of the degree of orness for the Choquet integral and other aggregation functions can be found in Grabisch et al. [2009].

Dispersion (Shannon entropy) for the Choquet integral

The dispersion indicator (Shannon entropy) associated with the OWA operator has been analyzed and generalized in several studies [Yager, 2000; Dukhovny, 2002; Marichal, 2002; Kojadinovic et al., 2005]. Unlike the degree of orness, the Shannon entropy is always a global indicator because the value of $\mathrm{Disp}(\vec{w})$ is not modified if $w_{\sigma(i)}$ instead of w_i is used for all i (see Table 5.1). The analytical expression of the generalization proposed in Yager [2000] is shown in Table 5.2.

Table 5.2: Summary of existing indicators extended to the Choquet integral

Indicator	Analytical expression	Reference
Global degree of orness[1]	$\omega_G(\mu) = \frac{1}{n-1} \cdot \sum_{i=1}^{n-1} \left[\binom{n}{i}^{-1} \cdot \sum_{\substack{A \subseteq N \\ \lvert A \rvert = i}} \mu(A) \right]$	Marichal [2004]
Local degree of orness	$\omega_L(\mu) = \sum_{i=1}^{n} \left(\frac{i-1}{n-1} \right) \cdot (\mu(A_{id,i}) - \mu(A_{id,i+1}))$	Belles-Sampera et al. [2013b], chapter 4 in this dissertation
Dispersion (Yager's Shannon entropy)[2,3]	$H_Y(\mu) = -\sum_{i=1}^{n} \phi_i(\mu) \cdot \ln[\phi_i(\mu)]$	Yager [2000]

[1] Other degrees of orness can be found in Grabisch et al. [2009].

[2] Following notation used in Kojadinovic et al. [2005], where $\phi_i(\mu)$ stands for the ith component of the Shapley value of μ.

[3] Alternative entropy measures can be found in Dukhovny [2002]; Marichal [2002] and Kojadinovic et al. [2005].

5.3 NEW INDICATORS EXTENDED TO THE CHOQUET INTEGRAL

Generalizations of the degree of balance, the divergence, the variance indicator and Rényi entropies for the Choquet integral are proposed in this section. Each of these generalizations satisfies the fol-

lowing property: when the capacity μ linked to the Choquet integral $\mathcal{C}_\mu$ is symmetric and normalized (implying that a weighting vector $\vec{w}$ exists such that $\mathcal{C}_\mu = \mathrm{OWA}_{\vec{w}}$), then the indicators for $\mathcal{C}_\mu$ coincide with the respective indicators for $\mathrm{OWA}_{\vec{w}}$. That is, the information provided by indicators associated to the Choquet integral $\mathcal{C}_\mu$ with symmetric and normalized capacity μ corresponds to the information provided by indicators of the $\mathrm{OWA}_{\vec{w}}$ operator, although former indicators cover a wider range of aggregation operators.

Degree of balance for the Choquet integral

Expressions (5.3) for the global and local degrees of balance indicators associated with Choquet integrals are proposed. Note that the degree of balance introduced by Yager [1996] was in the range $[-1, 1]$, where values of the degree of orness from $[0, 1]$ were rescaled. Here, the degree of balance is defined for any interval $[a, b] \subseteq \mathbb{R}$ where $b > a$.

$$\begin{aligned} \mathrm{Bal}_{G,[a,b]}(\mathcal{C}_\mu) &:= (b-a)\cdot \omega_G(\mu) + \mu(N)\cdot a, \\ \mathrm{Bal}_{L,[a,b]}(\mathcal{C}_\mu) &:= (b-a)\cdot \omega_L(\mu) + \mu(N)\cdot a. \end{aligned} \tag{5.3}$$

Note that definitions (5.3) are linear transformations of the degree of orness. If μ is not normalized, the values of the degree of balance belong to the interval $[a\cdot\mu(N), b - a\cdot(1-\mu(N))]$. These definitions fulfill linearity conditions with respect to capacities, as shown in Section 5.4.

It is straightforward to check that when μ is symmetric then the following holds $\mathrm{Bal}_{L,[a,b]}(\mathcal{C}_\mu) = \mathrm{Bal}_{G,[a,b]}(\mathcal{C}_\mu)$. If, in addition, μ is normalized and $a = -1$ and $b = 1$ then $\mathrm{Bal}(\vec{w}) = \mathrm{Bal}_{L,[-1,1]}(\mathcal{C}_\mu) = \mathrm{Bal}_{G,[-1,1]}(\mathcal{C}_\mu)$.

As in the case of the degree of orness, if μ is additive and normalized, then in general $\mathrm{Bal}_{L,[a,b]}(\mathcal{C}_\mu) \neq \mathrm{Bal}_{G,[a,b]}(\mathcal{C}_\mu)$. In particular, $\mathrm{Bal}_{G,[a,b]}(\mathcal{C}_\mu) = \dfrac{a+b}{2}$ and $\mathrm{Bal}_{L,[a,b]}(\mathcal{C}_\mu) = b + a\cdot\sum_{i=1}^{n}\left(\dfrac{n-i}{n-1}\right)\cdot p_i$ are both satisfied.

Divergence indicator for the Choquet integral

Extensions of the divergence indicator to the Choquet integral level are provided in this section. As mentioned previously in the context of $\mathrm{OWA}_{\vec{w}}$ operators, situations exist in which the degree of orness and the dispersion indicator are insufficient for characterizing a weighting vector $\vec{w}$. In such instances, a supplementary measure providing additional information is required. The divergence indicator

is a good candidate to fill this gap.

The divergence indicator of a Choquet integral is defined from a global and a local perspective.

Definition 5.3.1. *Let $\mathcal{C}_\mu$ be the Choquet integral with respect to a capacity μ defined on N. The global divergence indicator of $\mathcal{C}_\mu$ is defined as*

$$\mathrm{Div}_G(\mathcal{C}_\mu) := \sum_{i=1}^{n} \left(\frac{i-1}{n-1} - \omega_G(\mu) \right)^2 \cdot \left[\frac{\sum_{\substack{A \subseteq N \\ |A|=n-i+1}} \mu(A)}{\binom{n}{n-i+1}} - \frac{\sum_{\substack{A \subseteq N \\ |A|=n-i}} \mu(A)}{\binom{n}{n-i}} \right]. \tag{5.4}$$

It can be interpreted that the global divergence provides a mean variability around the degree of orness of the aggregation function. In other words, the value of the global divergence is associated with the scattering of the aggregation function around the global degree of orness. So, as the global divergence increases, the global degree of orness (the level of disjunction) becomes less important in the aggregation process.

It should be pointed out that the divergence indicator of the OWA operator was interpreted as $\mathrm{Div}(\vec{w}) = \mathbb{E}_{\vec{w}}[(X^*)^2] - (\mathbb{E}_{\vec{w}}[X^*])^2$ in Section 5.2. So, the divergence indicator of the OWA operator can be understood as a function of statistical moments of random variable X^*. Expression (5.4) can be interpreted in a similar way. In order to achieve this parallelism for the divergence indicator of the Choquet integral, the introduction of an additional concept is required. With this purpose, I propose a new capacity linked to μ.

Definition 5.3.2 (AQWA capacity)**.** *Let μ be a capacity defined on the set $N = \{m_1, ..., m_n\}$, where the subindex indicates an ordering on the elements of N. The ascending quadratic weighted additive (AQWA) capacity linked to μ is an additive capacity η on N defined by*

(i) $$\eta(\{m_j\}) := 2 \cdot \left(\frac{j-1}{n-1} \right)^2 \cdot \left[\frac{\sum_{\substack{A \subseteq N \\ |A|=n-i+1}} \mu(A)}{\binom{n}{n-i+1}} - \frac{\sum_{\substack{A \subseteq N \\ |A|=n-i}} \mu(A)}{\binom{n}{n-i}} \right],$$

for all $j = 1, .., n$;

(ii) $$\eta(A) := \sum_{m_k \in A} \eta(\{m_k\});$$ *and* $\eta(\emptyset) := 0$.

A proof that η is a capacity on N is provided in the Appendix. Intuitively, the AQWA capacity can be interpreted as an additive capacity which assigns a part of the global orness of $\mathcal{C}_\mu$ to each element

m_j multiplied by a convenient factor. Note that the definition of the AQWA capacity is based on the capacity μ. It has to be mentioned that, although the underlying idea is the same, this definition slightly differs from the one provided in Belles-Sampera et al. [2013d]. The goal in both cases is to devise an additive capacity linked to μ which includes in its definition the weights $((j-1)/(n-1))^2$ for $j = 1, ..., n$. Then, a compact expression for (5.4) may be obtained as shown in the Appendix. In Belles-Sampera et al. [2013d] capacities were restricted to take values in $[0, 1]$. Here, this restriction is relaxed and capacities are permitted to take any non-negative value. My opinion is that this definition is more elegant and appealing expressions can be derived. Taking advantage of this capacity linked to μ, it is possible to transform expression (5.4) into a more compact expression in the following proposition.

Proposition 5.3.1. *The global divergence indicator for a discrete Choquet integral given by expression* (5.4) *is equivalent to*

$$\mathrm{Div}_G\left(\mathcal{C}_\mu\right) = \boldsymbol{\omega}_G(\eta) - [2 - \mu(N)] \cdot \boldsymbol{\omega}_G^2(\mu) \quad , \tag{5.5}$$

where η *is the AQWA capacity linked to* μ.

A proof of Proposition 5.3.1 is given in the Appendix. Let me emphasize that expression (5.5) reminds $\mathrm{Div}(\vec{w}) = \mathbb{E}_{\vec{w}}[(X^*)^2] - (\mathbb{E}_{\vec{w}}[X^*])^2$ where the role that the mathematical expectation was playing for $\mathrm{Div}(\vec{w})$ is now the role of the global degree of orness, using η and μ instead of X^{*2} and X^*. Proposition 5.3.1 allows a straightforward computation of the value of the divergence indicator when $\boldsymbol{\omega}_G(\eta)$ and $\boldsymbol{\omega}_G^2(\mu)$ are known.

The definition of the local divergence indicator is as follows:

$$\mathrm{Div}_L\left(\mathcal{C}_\mu\right) := \sum_{i=1}^{n} \left(\frac{i-1}{n-1} - \boldsymbol{\omega}_L(\mu)\right)^2 \cdot \left(\mu\left(A_{id,i}\right) - \mu\left(A_{id,i+1}\right)\right). \tag{5.6}$$

This definition is inspired by the fact that $\mathrm{Div}(\vec{w}) = \mathrm{WAM}_{\vec{w}}(\vec{z})$ in the case of OWA operators (see Section 5.2.1) and corresponds to the local perspective.

When μ is symmetric and normalized, it holds that

$$\mathrm{Div}_L\left(\mathcal{C}_\mu\right) = \mathrm{Div}_G\left(\mathcal{C}_\mu\right) = \mathrm{Div}\left(\vec{w}\right). \tag{5.7}$$

If μ is additive and normalized, $\boldsymbol{\omega}_G(\mu) = \frac{1}{2}$ and $\boldsymbol{\omega}_G(\eta) = \frac{2 \cdot n - 1}{6 \cdot (n-1)}$, and hence

$$\mathrm{Div}_G\left(\mathcal{C}_\mu\right) = \frac{2 \cdot n - 1}{6 \cdot (n-1)} - (2-1) \cdot \frac{1}{4} = \frac{1}{12} \cdot \frac{n+1}{n-1}, \tag{5.8}$$

so, in general, it is easy to observe that $\mathrm{Div}_G\left(\mathcal{C}_\mu\right) \neq \mathrm{Div}_L\left(\mathcal{C}_\mu\right)$ when μ is additive and normalized. The proofs of (5.7) and (5.8) are given in the Appendix.

Variance indicator and Rényi entropies for the Choquet integral

In Section 5.2.1 above two additional indicators for OWA operators were shown, namely the variance indicator of the weighting vector and the Rényi entropy of degree α. The generalized definitions of these global indicators for the Choquet integral can be provided but the local perspective for this indicators is not considered. The reason for this being that the two indicators are only defined in terms of the weighting vector in the case of OWA operators, but not in terms of $\frac{i-1}{n-1}$ or $\frac{2\cdot i-(n+1)}{n-1}$.

The global variance indicator of a capacity linked to a Choquet integral may be defined as

$$D_G^2(\mathcal{C}_\mu) = \frac{1}{n}\cdot\sum_{i=1}^{n}\left[\frac{\sum\limits_{\substack{A\subseteq N\\ |A|=n-i+1}}\mu(A)}{\binom{n}{n-i+1}} - \frac{\sum\limits_{\substack{A\subseteq N\\ |A|=n-i}}\mu(A)}{\binom{n}{n-i}}\right]^2 - \frac{\mu(N)^2}{n^2} \quad . \tag{5.9}$$

The global Rényi entropies of degree $\alpha \in \mathbb{R}\backslash\{1\}$ for a Choquet integral with respect to μ may be defined as

$$H_{G,\alpha}(\mathcal{C}_\mu) = \frac{1}{1-\alpha}\cdot\log_2\left(\sum_{i=1}^{n}\left[\frac{\sum\limits_{\substack{A\subseteq N\\ |A|=n-i+1}}\mu(A)}{\binom{n}{n-i+1}} - \frac{\sum\limits_{\substack{A\subseteq N\\ |A|=n-i}}\mu(A)}{\binom{n}{n-i}}\right]^\alpha\right). \tag{5.10}$$

A summary of the set of indicators extended to the Choquet integral is presented in Table 5.3. As it is shown, global indicators involve the computation of all permutations of elements from N. Note that it can be a cumbersome task whether the number of elements of N is large. Unlike global indicators, local indicators are computationally tractable even in the case of a large N.

5.4 INDICATORS WITH RESPECT TO A LINEAR COMBINATION OF CAPACITIES

5.4.1 *Linearity features of the extended indicators*

Let me denote any global or local indicator associated with a Choquet integral with respect to a capacity μ as $F(\mu)$. My intention is

Table 5.3: Summary of new indicators extended to the Choquet integral.

Indicator	Analytical expression
Global degree of balance	$Bal_{G,[a,b]}(\mathcal{C}_\mu) = (b-a)\cdot\omega_G(\mu) + \mu(N)\cdot a$
Local degree of balance	$Bal_{L,[a,b]}(\mathcal{C}_\mu) = (b-a)\cdot\omega_L(\mu) + \mu(N)\cdot a$
Global divergence	$Div_G(\mathcal{C}_\mu) = \sum_{i=1}^{n}\left(\frac{i-1}{n-1} - \omega_G(\mu)\right)^2 \cdot \left[\frac{\sum_{\substack{A\subseteq N \\ \lvert A\rvert = n-i+1}} \mu(A)}{\binom{n}{n-i+1}} - \frac{\sum_{\substack{A\subseteq N \\ \lvert A\rvert = n-i}} \mu(A)}{\binom{n}{n-i}}\right]$
Local divergence	$Div_L(\mathcal{C}_\mu) = \sum_{i=1}^{n}\left(\frac{i-1}{n-1} - \omega_L(\mu)\right)^2 \cdot (\mu(A_{id,i}) - \mu(A_{id,i+1}))$
Variance indicator	$D_G^2(\mathcal{C}_\mu) = \frac{1}{n}\cdot\sum_{i=1}^{n}\left[\frac{\sum_{\substack{A\subseteq N \\ \lvert A\rvert = n-i+1}} \mu(A)}{\binom{n}{n-i+1}} - \frac{\sum_{\substack{A\subseteq N \\ \lvert A\rvert = n-i}} \mu(A)}{\binom{n}{n-i}}\right]^2 - \frac{\mu(N)^2}{n^2}$
Rényi entropy ($\alpha \neq 1$)	$H_{G,\alpha}(\mathcal{C}_\mu) = \frac{1}{1-\alpha}\cdot\log_2\left(\sum_{i=1}^{n}\left[\frac{\sum_{\substack{A\subseteq N \\ \lvert A\rvert = n-i+1}} \mu(A)}{\binom{n}{n-i+1}} - \frac{\sum_{\substack{A\subseteq N \\ \lvert A\rvert = n-i}} \mu(A)}{\binom{n}{n-i}}\right]^\alpha\right)$

to assess the expressions of $F(\lambda_1\cdot\mu_1 + \lambda_2\cdot\mu_2)$, where $\lambda_1, \lambda_2 \in [0,1]$ and μ_1, μ_2 are capacities defined on N. If the indicator is linear with respect to capacities then $F(\lambda_1\cdot\mu_1 + \lambda_2\cdot\mu_2) = \lambda_1\cdot F(\mu_1) + \lambda_2\cdot F(\mu_2)$ must hold.

Linearity of the degree of orness and the degree of balance

The global and the local degrees of orness are both linear with respect to capacities. From expressions (5.1) and (5.2) with $\mu = \lambda_1\cdot\mu_1 + \lambda_2\cdot\mu_2$, and noting that $(\lambda_1\cdot\mu_1 + \lambda_2\cdot\mu_2)(A) = \lambda_1\cdot\mu_1(A) + \lambda_2\cdot\mu_2(A)$ for any $A \in 2^N$, then it is deduced that $\omega_G(\lambda_1\cdot\mu_1 + \lambda_2\cdot\mu_2) = \lambda_1\cdot\omega_G(\mu_1) + \lambda_2\cdot\omega_G(\mu_2)$ and $\omega_L(\lambda_1\cdot\mu_1 + \lambda_2\cdot\mu_2) = \lambda_1\cdot\omega_L(\mu_1) + \lambda_2\cdot\omega_L(\mu_2)$.

The linearity of the degree of balance (global and local) with respect to capacities can be assessed using the above expressions and the fact

that this indicator is a linear transformation of the degree of orness (as shown in Section 5.3). The expression

$$\begin{aligned} &\mathrm{Bal}_{*,[a,b]}\left(\mathcal{C}_{\lambda_1\cdot\mu_1+\lambda_2\cdot\mu_2}\right) \\ &= (b-a)\cdot\boldsymbol{\omega}_*(\lambda_1\cdot\mu_1+\lambda_2\cdot\mu_2)+(\lambda_1\cdot\mu_1+\lambda_2\cdot\mu_2)(N)\cdot a \\ &\quad = \lambda_1\cdot(b-a)\cdot\boldsymbol{\omega}_*(\mu_1)+\lambda_1\cdot\mu_1(N)\cdot a \\ &\quad +\lambda_2\cdot(b-a)\cdot\boldsymbol{\omega}_*(\mu_2)+\lambda_2\cdot\mu_2(N)\cdot a \\ &\quad = \lambda_1\cdot\mathrm{Bal}_{*,[a,b]}\left(\mathcal{C}_{\mu_1}\right)+\lambda_2\cdot\mathrm{Bal}_{*,[a,b]}\left(\mathcal{C}_{\mu_2}\right), \end{aligned} \tag{5.11}$$

holds for global and local indicators (i.e., either if $* = G$ or $* = L$). Thus, the degree of balance is linear with respect to capacities.

Non-linearity of the divergence, the dispersion, the variance indicator and Rényi entropies

The divergence indicator is not linear with respect to capacities in the general case, as can be deduced from expressions (5.5) and (5.6). Nonetheless, a result that characterizes the geometric locus where the divergence indicator satisfies linearity is presented in Belles-Sampera et al. [2013c]. Although not explicitly proved, the lack of linearity of the dispersion, the variance indicator and Rényi entropies is evident due to the lack of linearity (in the general case) of functions $\ln(x)$, x^2 and $\log_2(x)$, respectively.

5.4.2 *Application: inherited indicators of POWA operators*

Indicators for the Probabilistic Ordered Weighted Averaging (POWA) operator are derived. The POWA operator was introduced in Merigó [2011], Merigó and Wei [2011] and Merigó [2012]. Let a vector $\vec{w}$, $\vec{w} = (w_1, w_2, ..., w_n) \in [0,1]^n$, be such that $\sum_{i=1}^n w_i = 1$ and let $\vec{p} = (p_1, p_2, ..., p_n) \in [0,1]^n$ be such that $\sum_{i=1}^n p_i = 1$. In addition, consider $\beta \in [0,1]$. The POWA operator with respect to $\vec{w}, \vec{p}$ and β is a mapping from $\mathbb{R}^n$ to $\mathbb{R}$ defined by

$$\mathrm{POWA}_{\vec{w},\vec{p},\beta}(x_1,...,x_n) := \beta\cdot\sum_{i=1}^n x_{\sigma(i)}\cdot w_i + (1-\beta)\cdot\sum_{i=1}^n x_{\sigma(i)}\cdot p_{\sigma(i)}, \tag{5.12}$$

where σ is a permutation of $(1,2,...,n)$ such that $x_{\sigma(1)} \leqslant x_{\sigma(2)} \leqslant ... \leqslant x_{\sigma(n)}$, i.e. $x_{\sigma(i)}$ is the ith smallest value of $x_1, x_2, ..., x_n$.

An alternative expression to (5.12) is

$$\mathrm{POWA}_{\vec{w},\vec{p},\beta}(x_1,...,x_n) = \sum_{i=1}^n x_{\sigma(i)}\cdot v_{i,\sigma}, \tag{5.13}$$

where $v_{i,\sigma} = \beta \cdot w_i + (1-\beta) \cdot p_{\sigma(i)}$ for all $i = 1,..n$. It is straightforward to see that $v_{i,\sigma} \in [0,1]$ and $\sum_{i=1}^{n} v_{i,\sigma} = 1$. Note that the POWA operator can be understood as a weighted average between an OWA operator and a WAM. When a random variable X that can take n different values denoted by $\{x_i\}_{i=1,...,n}$ is such that $\mathcal{P}(X = x_i) = p_i$ for all $i = 1,...,n$, then the POWA operator can also be understood as a weighted average between an OWA operator and the mathematical expectation of the random variable X:

$$\begin{aligned} \text{POWA}_{\vec{w},\vec{p},\beta}(\vec{x}) &= \beta \cdot \text{OWA}_{\vec{w}}(\vec{x}) + (1-\beta) \cdot \text{WAM}_{\vec{p}}(\vec{x}) \\ &= \beta \cdot \text{OWA}_{\vec{w}}(\vec{x}) + (1-\beta) \cdot \mathbb{E}(X). \end{aligned} \tag{5.14}$$

Alternatively, POWA operators may be viewed as convex combinations of projections and order statistics. In other words, note that expression (5.12) is equivalent to

$$\text{POWA}_{\vec{w},\vec{p},\beta}(\vec{x}) = \beta \cdot \sum_{i=1}^{n} w_i \cdot \text{OS}_i(\vec{x}) + (1-\beta) \cdot \sum_{i=1}^{n} p_i \cdot \Pi_i(\vec{x}) \ ,$$

where $\text{OS}_i(\vec{x})$ denotes the ith order statistic of $\vec{x}$ and $\Pi_i(\vec{x})$ the ith projection of $\vec{x}$. Whether only projections have non zero weights, then WAM are recovered; if only order statistics have non zero weights, then POWA is actually an OWA operator.

Taking into account the relationship between OWA operators and Choquet integrals (Section 5.1), expression (5.14) may be formulated as $\text{POWA}_{\vec{w},\vec{p},\beta} = \beta \cdot \mathcal{C}_\mu + (1-\beta) \cdot \mathcal{C}_\mathcal{P}$. The capacities μ and $\mathcal{P}$ are normalized, where the former is symmetric and the latter a probability. This expression is a convex combination of two Choquet integrals that combines an OWA and a probabilistic perspective. Considering now the linearity of Choquet integrals with respect to the capacity (see Proposition 9(i) in [Grabisch et al., 2011]), the representation of the POWA operator as a Choquet integral is directly derived as

$$\text{POWA}_{\vec{w},\vec{p},\beta} = \mathcal{C}_{\beta \cdot \mu + (1-\beta) \cdot \mathcal{P}} \ . \tag{5.15}$$

Therefore indicators for the POWA operator may be defined as follows:

$$\begin{aligned}
\omega_*(\mathrm{POWA}_{\vec{w},\vec{p},\beta}) &:= \omega_*(\beta\cdot\mu+(1-\beta)\cdot\mathcal{P})\\
&= \beta\cdot\omega_*(\mu)+(1-\beta)\cdot\omega_*(\mathcal{P}),\\
\mathrm{H_Y}\left(\mathrm{POWA}_{\vec{w},\vec{p},\beta}\right) &:= -\beta\cdot\sum_{i=1}^{n}\phi_i(\mu)\cdot\ln\left[\beta\cdot\phi_i(\mu)+(1-\beta)\cdot\phi_i(\mathcal{P})\right]\\
&-(1-\beta)\cdot\sum_{i=1}^{n}\phi_i(\mathcal{P})\cdot\ln\left[\beta\cdot\phi_i(\mu)+(1-\beta)\cdot\phi_i(\mathcal{P})\right],\\
\mathrm{Bal}_{*,[a,b]}(\mathrm{POWA}_{\vec{w},\vec{p},\beta}) &:= \mathrm{Bal}_{*,[a,b]}(\mathcal{C}_{\beta\cdot\mu+(1-\beta)\cdot\mathcal{P}})\\
&= \beta\cdot\mathrm{Bal}_{*,[a,b]}(\mathcal{C}_\mu)+(1-\beta)\cdot\mathrm{Bal}_{*,[a,b]}(\mathcal{C}_\mathcal{P}),\\
\mathrm{Div}_*(\mathrm{POWA}_{\vec{w},\vec{p},\beta}) &:= \mathrm{Div}_*(\mathcal{C}_{\beta\cdot\mu+(1-\beta)\cdot\mathcal{P}}),\\
\mathrm{D}^2_\mathrm{G}(\mathrm{POWA}_{\vec{w},\vec{p},\beta}) &:= \mathrm{D}^2_\mathrm{G}(\mathcal{C}_{\beta\cdot\mu+(1-\beta)\cdot\mathcal{P}}),\\
\mathrm{H}_{\mathrm{G},\alpha}(\mathrm{POWA}_{\vec{w},\vec{p},\beta}) &:= \mathrm{H}_{\mathrm{G},\alpha}(\mathcal{C}_{\beta\cdot\mu+(1-\beta)\cdot\mathcal{P}}).
\end{aligned}$$

Note that the linearity properties of the degree of orness and the degree of balance allow the degree of orness and the degree of balance indicators to be defined for the POWA operator as linear combinations of the indicators associated to the underlying OWA and WAM operators. The dispersion, the divergence, the variance indicator and Rényi entropies of the POWA operator are not linear combinations of the dispersion, the divergence, the variance indicator and Rényi entropies of the underlying OWA and WAM operators. Only in special cases, such as those derived in Belles-Sampera et al. [2013c] for the divergence, is linearity satisfied. Note that the dispersion and the divergence for the POWA operator introduced in Merigó [2012] represents an alternative approach. There, the author proposes a linear combination of these indicators for the underlying OWA and WAM operators.

To conclude, I have derived indicators for the POWA operator. However, the POWA operator is only one of a set of possible examples. For instance, the weighted ordered weighted averaging (WOWA) operator introduced by Torra [1997] might also be considered and the inherited indicators shown [Belles-Sampera et al., 2013d], due to the relationship between the WOWA operator and the Choquet integral with respect to particular capacities (see Theorem 4 in Torra [1998]).

6

ON THE IMPLICIT RISK ATTITUDE OF A DISTORTION RISK MEASURE

Up to my knowledge scarce attention has been paid to the implicit risk attitude that the decision maker is assuming when using the VaR and TVaR risk measures. Given a α-confidence level, the $\text{TVaR}_\alpha(X)$ is always more conservative than $\text{VaR}_\alpha(X)$. This direct comparison between the two risk measures can not be longer made if a higher confidence level is associated to the VaR risk measure. For instance, let suppose that the decision maker wants to compare the implicit risk attitude between $\text{TVaR}_{90\%}(X)$ and $\text{VaR}_{95\%}(X)$. In that case it is not easy to select the risk measure that involves the most conservative attitude. Even more, what does it happen if the implicit risk attitude of other risk measures different to these two quantile-based risk measures is studied? In this context a risk attitude comparison between measures seems to be even less intuitive.

This chapter pursues to contribute into the study of the underlying risk behavior in risk assessment. The study focus the attention on the analysis of the risk attitude implicit into distortion risk measures, where the VaR and TVaR are particular cases of this class of risk measures. The characterization of the implicit risk attitude of a distortion risk measure is carried out by means of the computation of aggregation indicators, as shown in Chapters 4 and 5, and a graphical analysis based on the distortion function. It is argued that the combination of these two tools provide a precise portrait of the underlying risk behavior of a decision maker when using a particular risk measure.

As it has been shown in previous chapters of this dissertation, distortion risk measures can be represented as Choquet integrals. In Chapter 5 it has been argued that a way to describe the characteristics of a discrete Choquet integral is using a set of aggregation indicators. These indicators provide information regarding the features of the underlying aggregation operator. Now, a risk-based interpretation of aggregation indicators is provided in the context of risk assessment. Particularly, the degree of orness associated to each risk measure is estimated, and it is interpreted in terms of the risk attitude of managers when they are using a risk measure in a particular context. The quantitative information related to the aggregate risk attitude linked to the risk measure provided by these indicators is investigated. It is claimed that these indicators may be useful to characterize the *overall*

risk behavior implicit in the risk measure.

The overall risk attitude only provides a partial portrait of the underlying risk attitude implicit into the risk measure. It is reasonable to suppose that decision makers are not worried about all random events in the same way. Decision makers frequently give different treatments to different random events. Note that some of these events can represent benefits or affordable losses. Therefore, the *overall risk behavior* of a risk measure has to be completed with local information about the implicit risk attitude in front of a particular random event. A quotient function based on the distortion function associated to the risk measure is defined with this purpose. The quotient function is graphically analyzed to investigate the risk attitude of the agent at any point of the survival function when using a certain risk measure. The graphical evaluation of the risk-appetite pattern of the manager in the range of feasible values is the basis of the definition of two concepts, *absolute risk behavior* and *specific risk behavior*.

An illustrative example of the risk behavior characterization implicit in a distortion risk measure is included. The example focuses on regulatory based risk quantification in the European insurance market. The risk measure of reference in this framework is the $\text{VaR}_{99.5\%}$. Therefore, it is supposed that the insurance risk manager pursues a risk assessment equivalent to the regulatory reference returned by the $\text{VaR}_{99.5\%}$. Two scenarios involving different worst losses are proposed to investigate the performance of the VaR and the implicit risk attitude that is associated to this risk measure. Using a real dataset, reserves are estimated by means of the $\text{VaR}_{99.5\%}$ and the same amounts are replicated with a set of equivalent GlueVaR. In the example it is illustrated how the calibration process of GlueVaR parameters can be performed to obtain equivalent values to the amount returned by the $\text{VaR}_{99.5\%}$. The additional risk attitude information provided by the GlueVaR measures in comparison to the VaR and its usefulness for decision makers is examined. As a consequence, in this chapter the applicability of GlueVaR risk measures on a real environment and the fact that these risk measures may be useful complementary tools to the VaR for evaluating risk in the European insurance context are highlighted.

6.1 RISK ATTITUDE OF AGENTS

The risk behavior of agents when they select a particular distortion risk measure to assess the risk is analyzed here. The degree of orness and a graphical analysis of the relative distortion associated to the risk measure are two useful tools to characterize the risk attitude of agents.

The former gives an aggregated indicator of the risk attitude of the agent towards losses. The risk behavior is frequently not constant over the whole range of losses. The agent's risk behavior can be examined in any interval of losses by a graphical inspection of the relative distortion associated to the risk measure.

6.1.1 *Degree of orness and risk attitude*

The definitions of the degree of orness given in Section 5.2.2 of Chapter 5 is here reminded. Let introduce expression (6.1) to provide compact notation,

$$\mathbb{S}_{i,*}^{\mu} = \begin{cases} \mu\left(A_{id,n-i+1}\right) & \text{if } * = L \\ \binom{n}{i}^{-1} \cdot \sum\limits_{\substack{A \subseteq N \\ |A|=i}} \mu\left(A\right) & \text{if } * = G \end{cases}, \tag{6.1}$$

where L stands for local and G for global indicators. Remember from previous definitions that $A_{id,n-i+1} = \{m_{n-i+1}, \cdots, m_n\}$, where $N = \{m_1, \ldots, m_n\}$.

The degree of orness pursues to summarize the importance of each ith order statistic, $i = 1, ..., n$, in the aggregation process associated to the Choquet integral with respect to capacity μ (sometimes noted as $\mathcal{C}_\mu$). The degree of orness is computed as follows:

$$\omega_*\left(\mathcal{C}_\mu\right) = \sum_{i=1}^{n} \left(\frac{i-1}{n-1}\right) \cdot \left[\mathbb{S}_{n-i+1,*}^{\mu} - \mathbb{S}_{n-i,*}^{\mu}\right]. \tag{6.2}$$

Therefore, the degree of orness provides some kind of *level of preference* inherent to such an aggregation function in a $[0, 1]$ scale, where 0 represents the minimum and 1 the maximum order statistic. To simplify notation, the difference $\mathbb{S}_{n-i+1,*}^{\mu} - \mathbb{S}_{n-i,*}^{\mu}$ is denoted as $w_{i,*}^{\mu}$, which can be understood as the weight indicating the importance given to the ith order statistic represented by the fraction $\frac{i-1}{n-1}$, $i = 1 \ldots n$.

The distinction between global and local indicators is related to the ordering of input data, as it was shown in Chapter 5. In the case of the local degree of orness, this *level of preference* related to order statistics is computed assuming that input data $(x_1, x_2, ..., x_n)$ is increasingly ordered, while global indicators do not make any assumption on the ordering of input data. Note that the assumption of ordered data associated to local indicators may be necessary when n is large. The calculation of $\mathbb{S}_{i,G}^{\mu}$ involves combinatorial numbers and it may have

high computational costs if n is large.

To give some additional insights about the information on risk attitude provided by the degree of orness, let me rewrite expression (6.2) in the following way:

$$\omega_* (\mathcal{C}_\mu) = \sum_{i=1}^{n-1} \frac{1}{n-1} \cdot \mathbb{S}^\mu_{n-i,*}. \tag{6.3}$$

If $\mu = g \circ P$, it is straightforward to check that next expressions hold:

$$\begin{aligned} \mathbb{S}^\mu_{n-i,L} &= g\left(S_X\left(x_i\right)\right) = g(\sum_{j=i+1}^{n} p_j) \\ \mathbb{S}^\mu_{n-i,G} &= \binom{n}{n-i}^{-1} \cdot \sum_{\substack{A \subseteq N \\ |A|=n-i}} g(\sum_{x_j \in A} p_j) \quad . \end{aligned} \tag{6.4}$$

Let me put the attention on the local degree of orness. When combining expressions (6.3) and (6.4) one can deduce that $\omega_L\left(\mathcal{C}_{g\circ P}\right)$ provides an approximation to the area under the function g. The level of accuracy of the approximation would increase as larger is n, where n is the number of different values. Figure 6.1.1 illustrates how the local degree of orness approximates the area under the function g for the mathematical expectation and the VaR_α. Recall that the mathematical expectation can be understood as a distortion risk measure involving the identity function id as associated distortion function g. The area under the id function is one half, so the local degree of orness associated to the mathematical expectation should be approximately one half. In the case of the VaR_α risk measure, the local degree of orness linked to VaR_α must be close to α for an enough large value of n, since α is the value of the area under the graph of $\psi_\alpha(u) = \mathbb{1}\left[1-\alpha \leqslant u \leqslant 1\right]$. In Figure 6.1.1 it is shown that the accuracy of the approximation depends on the size of n. The area under the function g is computed by means of n rectangles, where each rectangle has a width equal to $\frac{1}{n-1}$. So, a more accurate approximation of the area under g would be given as larger is n conditioned that a larger set of distinct values of X is involved when n increases.

As far as the local degree of orness provides an accurate approximation to the area under g, this quantity can be understood as an indicator of the *overall risk behavior* of an agent. Decision-makers are frequently classified as either risk averse, risk neutral or risk loving agents. Note that a risk neutral agent would not distort the survival function, so the associated distortion function linked to overall risk neutrality should be the id function. The area under the id function is one half and this value could be used as a benchmark of a neutral

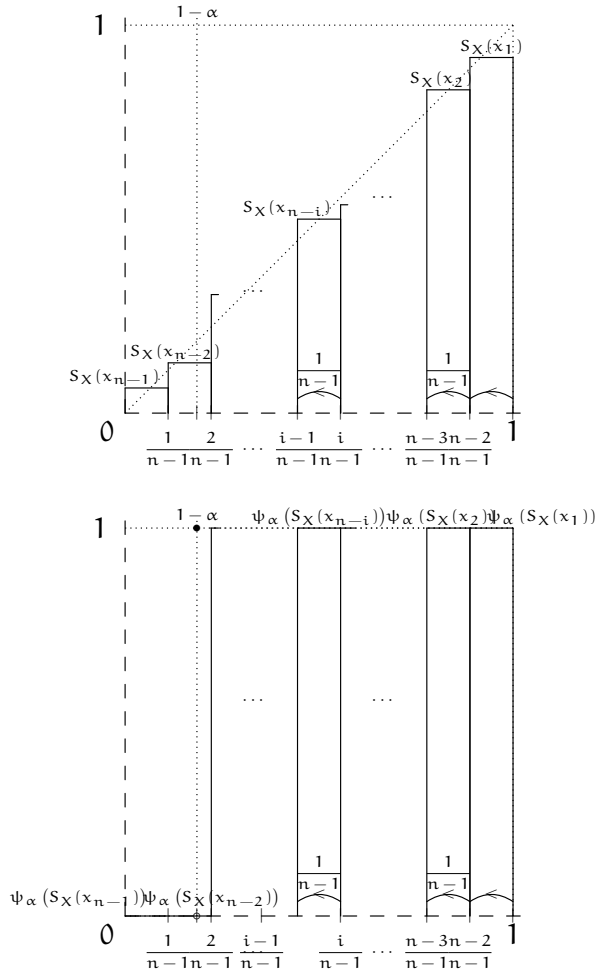


Figure 6.1: Approximation to the area under the distortion function given by the local degree of orness. Up: the identity function id in the case of the mathematical expectation. Down: distortion function ψ_α of the VaR_α risk measure ($\psi_\alpha(u) = \mathbb{1}[1-\alpha \leqslant u \leqslant 1]$).

overall risk attitude or behavior. An overall risk averse agent would make an upper distortion of the survival function in accumulated terms. In consequence, the area under g of an overall risk averse agent should be larger than one half. Similarly, an agent would have overall risk appetite if the area under g is lower than one half.

The interpretation of the global degree of orness is less intuitive from a risk management perspective. Until some extent, one can think that it returns a different approximation of the area under the graph of g, which differs from the previous one in the fact that none ordering on the set $\{p_1, p_2, ..., p_n\}$ is assumed or, equivalently, on the input data set $\{x_1, x_2, ..., x_n\}$. At each node $\frac{i-1}{n-1}, i = 1, ..., n$ the survival value is approximated as the average of $\binom{n}{n-i}$ feasible survival

values linked to the set $\{p_1, p_2, ..., p_n\}$.

An alternative interpretation can be provided. Under the perspective of aggregation operators, the global degree of orness is interpreted as an index that computes the similarity of the Choquet integral with the maximum function. Remember that the global degree of orness of an aggregation function is ranged between zero and one. The global degree of orness takes one whether the weight $w^{\mu}_{n,G}$ associated to the maximum is equal to one and the remaining weights are zero. It takes zero when the weight of the minimum $w^{\mu}_{1,G}$ is equal to one and the rest are zero. So, the value of global degree of orness indicates how the aggregation performed by the Choquet integral is ranked between these two boundaries [see, for instance, Torra and Narukawa, 2007; Fernández Salido and Murakami, 2003].

Adapting this interpretation to the risk measurement context, the global degree of orness would provide a *level of similarity* between the distorted survival function associated to the risk measure and the theoretical distorted survival function which returns the maximum loss. Note that the maximum value is returned in the worst-case scenario, i.e. the distorted survival function is equal to one in the range $(0, 1]$ and zero in zero (to be a distortion function). As lower is the global degree of orness as more dissimilar is the distorted survival function to the worst-case survival function. Therefore, an agent would be less risk averse in aggregated terms as lower is the global degree of orness.

6.1.2 *Graphical analysis of risk attitude*

A graphical analysis is proposed to complement the examination of risk attitudes of decision makers when using distortion risk measures. It was previously shown that the local and global degree of orness could be interpreted as indicators of the overall risk behavior. These indicators evaluate the accumulated distortion performed all over the survival function. Unfortunately, these measures do not have into account which part of the survival function was distorted. From the point of view of the manager, of course, it is not equivalent to distort the survival probability associated to the right tail of the random variable linked to losses or to distort the left tail. Additionally, all distortion functions with an area equal to one half would be linked to overall risk neutrality, where the id function is only a particular case. For instance, the median (the 50%-quantile or the $\mathrm{VaR}_{50\%}$) is another risk measure whose distortion function $\psi_{50\%}(u) = \mathbb{1}[0.5 \leqslant u \leqslant 1]$ satisfies that condition.

An option is to define the risk behavior in absolute terms. An *absolute risk neutral agent* should be the decision-maker who does not distort any survival probability and, then, who use the id function as associated distortion function, i.e. $g(u) = id(u) = u$ for all $0 < u < 1$. An *absolute risk averse agent* should have associated a distortion function g such that $g(u) > u$, for all $0 < u < 1$. And, in the same line, an *absolute risk loving agent* should have a distortion function g such that $g(u) < u$, for all $0 < u < 1$. This definition of risk behavior is in absolute terms in the sense that the relationship of ordering between $g(u)$ and u must be fulfilled in the whole range $(0, 1)$. Note that these considerations lead to a more restrictive definition of risk behavior than the previous *overall risk behavior*. Under the definition of overall risk behavior it is only required to fulfill the classification criterion in aggregated terms. So, all absolute risk averse agents are overall risk averse agents but the opposite does not hold.

The absolute risk behavior definition implies that the implicit risk attitude of an agent is invariable over the range of values. Nevertheless, there are no reasons to have an unique risk attitude on the whole range. An agent could have a different risk behavior depending on the interval of values under consideration. The risk attitude implicit in frequently used risk measures is not invariant. It is the case, for instance, of VaR_α. When using VaR_α, a risk averse behavior is involved in the interval $[1-\alpha, 1)$, but a risk loving attitude is associated to the interval $(0, 1-\alpha)$. So, an absolute risk behavior can not be linked to the VaR_α risk measure. In those situations that the implicit risk attitude is not similar all over $(0, 1)$, the absolute risk behavior criterion would not be an adequate classification.

The study of the risk behavior of agents when used a particular risk measure is often more explanatory when it is locally investigated in each particular area. Let me define the function $Q_g(u)$ on $u \in (0, 1]$ as the quotient between the distortion function g and the identity function, $Q_g(u) = \frac{g(u)}{u}$ for all $0 < u \leqslant 1$. The $Q_g(u)$ allows the analysis of the agent's risk behavior at any survival probability point. It takes non-negative values. This quotient function provides a function on survival probabilities u which describes the distortion factor applied at each u level by g.

The quotient Q_g is a quantifier of the *specific risk behavior* of the agent at any point. The quotient value represents the specific risk attitude of the decision maker in comparison to the risk neutral attitude of the agent who is confident with the survival probability. An agent is risk neutral, risk averse or risk loving at point u if $Q_g(u)$ is equal, higher or lower than one, respectively. The graphical analysis consists of plotting $Q_g(u)$ against the identity function, for all

$0 < u \leqslant 1$. Two examples are shown in Figure 6.2, where quotient functions associated to VaR_α and to TVaR_α are displayed.

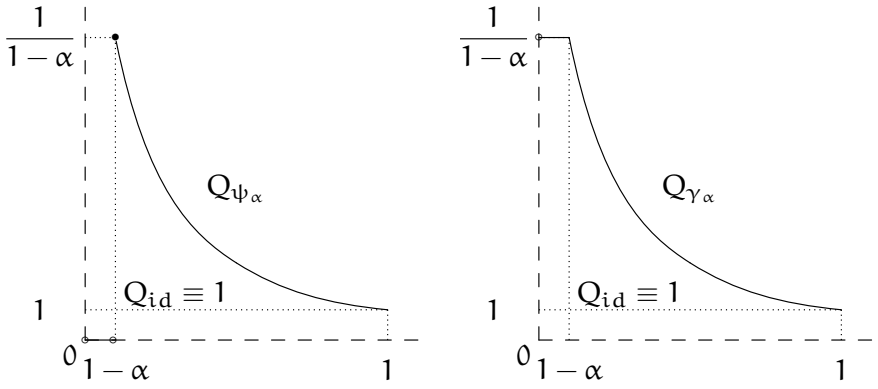


Figure 6.2: Quotient functions for VaR_α (left) and TVaR_α (right). Note that they only differ on the interval $(0, 1-\alpha)$. Q_{id} is the quotient function of the mathematical expectation $\mathbb{E}$.

Let me emphasize that the quotient $\frac{1}{u}$, for all $0 < u \leqslant 1$, marks out the maximum risk aversion frontier (as it can be deduced from Figure 6.2). The function $Q_g(u)$ computes the ratio between the distorted survival probability and the survival probability, so $\frac{1}{u}$ is the maximum value attainable by $Q_g(u)$. Note that the function $Q_g(u)$ is equal to 1 in the whole range for a completely risk neutral agent in absolute terms (*absolute risk neutral agent*). So, values $0 < Q_g(u) < 1$ indicate a risk loving behavior of the agent at point u.

If the attention is focused on the $Q_g(u)$ associated to the VaR_α, it is shown that a radical risk attitude is implicit in the interval $[1-\alpha, 1)$ which varies to the opposite extreme position in the interval $(0, 1-\alpha)$. Indeed, maximum risk aversion is involved in $[1-\alpha, 1)$ and maximum risk loving attitude in $(0, 1-\alpha)$. Some similarities are found when $Q_g(u)$ associated to the the TVaR_α is examined. Two ranges involving a different risk attitude are distinguished as well. Maximum risk aversion is involved in the interval $[1-\alpha, 1)$ and a constant (non-boundary) risk aversion attitude is involved in $(0, 1-\alpha)$. Unlike the VaR_α, an absolute risk averse behavior is associated to the TVaR_α since $Q_g(u) > 1$ or, equivalently, $g(u) > u$, for all $0 < u < 1$.

To conclude, it is shown that the quotient function Q_g can be used to characterize the specific risk behavior of an agent at any point. Note that the area under the quotient function Q_g provides the same information than the area under g, but now it is expressed on the basis of risk neutrality. This area can also be interpreted as a quantifier

of the overall risk behavior, where an area equal to one indicates overall risk neutrality, an area larger than one overall risk aversion and an area lower than one overall risk appetite. Remember that in the case of the distortion function g, the classification criterion of overall risk behavior was an area equal, larger or lower than one half.

In the case of a $\text{GlueVaR}^{h_1,h_2}_{\beta,\alpha}$, recall expression (1.3) in Chapter 1 for the distortion function $\kappa^{h_1,h_2}_{\beta,\alpha}$, which informs that it is the weighted sum of the distortion functions of a TVaR_β, a TVaR_α and a VaR_α. Recall also the abuse of notation $\kappa^{\omega_1,\omega_2}_{\beta,\alpha}$ used in expression (1.5). Therefore,

$$\begin{aligned}
&\text{Area}(\kappa^{\omega_1,\omega_2}_{\beta,\alpha}) \\
&= \omega_1 \cdot \text{Area}(\gamma_\beta) + \omega_2 \cdot \text{Area}(\gamma_\alpha) + (1-\omega_1-\omega_2) \cdot \text{Area}(\psi_\alpha) \\
&= \omega_1 \cdot \left(\frac{1+\beta-2\alpha}{2}\right) + \omega_2 \cdot \left(\frac{1-\alpha}{2}\right) + \alpha \quad .
\end{aligned}$$

An example of the graphical information provided by $\text{Q}_{\kappa^{\omega_1,\omega_2}_{\beta,\alpha}}$ is reproduced in Figure 6.3.

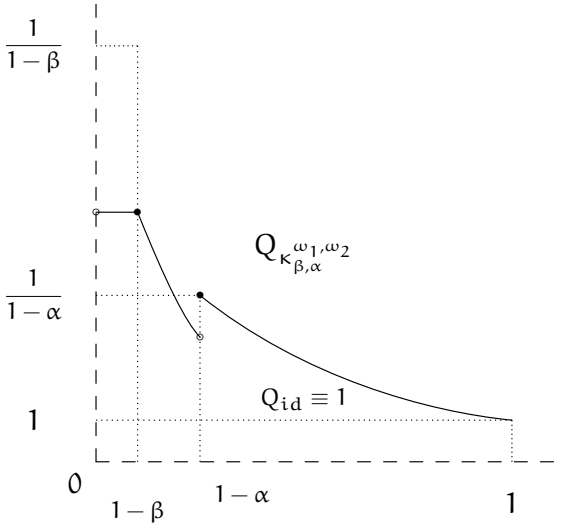


Figure 6.3: Graphical information provided by $\text{Q}_{\kappa^{\omega_1,\omega_2}_{\beta,\alpha}}$.

6.2 AN EXAMPLE ON CLAIM COSTS

The present section is devoted to illustrate how all previous findings can be applied to characterize the underlying risk attitude in practice. The illustration pursues to show real challenges faced by European insurance undertakings. Particularly, it is devised to highlight situations in which the implicit risk attitude linked to VaR is not able to detect sensitive changes in potential worst losses. It is

argued that the use of equivalent GlueVaR risk measures can be helpful to overcome this drawback. The section is structured as follows. First, the risk assessment scenario is described. Next, the calibration process to obtain equivalent GlueVaR risk measures is defined, and, finally, outcome results obtained with a particular data set are shown.

6.2.1 *Risk assessment in a scenario involving changes in worst losses*

Imagine a risk management team of an European insurance company working under the rules of Directive 2009/138/EC of 25 November 2009 (known as Solvency II Directive). In this framework the VaR with a confidence level $\alpha = 99.5\%$ is required to assess the regulatory capital.

Main concerns of the risk management team can be related to the use of VaR as a risk measure, because of the lack of risk-based information on worst losses intrinsic to this risk measure. Two firms with remarkable differences in the amount of potential losses in adverse events could be associated to the same risk value, although they are not exposed to the same level of risk, so their disparities would be unobservable by decision makers when their risks are evaluated by the VaR measure. Alternatives to $\text{VaR}_{99.5\%}$ that take into account worst losses can be considered by risk managers. Traditional approaches frequently lead to severely higher economic reserves. The management team needs to find out a risk measure that generates similar economic reserves than $\text{VaR}_{99.5\%}$ for the overall risk faced by the insurance company and, additionally, the management team needs that this alternative risk measure provides risk-based information on worst losses and that, hopefully, it satisfies appealing subadditivity properties. This is a real practical problem because the Solvency II Directive allows the insurance undertakings to look for such alternatives (see Article 122(1) of the Directive).

6.2.2 *Calibration of GlueVaR parameters*

The four-parameter $\text{GlueVaR}^{\omega_1,\omega_2}_{\beta,\alpha}$ risk measures can be used as an alternative to the $\text{VaR}_{99.5\%}$. Note that here the weights (ω_1, ω_2) instead of the heights (h_1, h_2) are going to be used or, in other words, expression (1.5) instead of expression (1.1) of the distortion function of GlueVaR is going to be taken into account from this point forward. To apply these measures in practice, it is needed to assign values to the parameters that define the GlueVaR risk measure. All steps to calibrate GlueVaR risk measures are here developed. The criterion that is followed in the calibration procedure is to obtain the same amount at risk with GlueVaR risk measures than the amount at risk obtained with $\text{VaR}_{99.5\%}$. Moreover, the selection of the risk measure

is restricted to the subfamily of GlueVaR risk measures which may satisfy tail-subadditivity at a given confidence level.

The strategy to calibrate the parameters of the GlueVaR risk measure is as follows:

- A minimum and maximum admissible value of α and β confidence levels should be determined, α_{min} and β_{max}. The selection implies a trade-off between protection and competitiveness. The level of policyholders' protection could be reduced with a low α_{min} but it could entail higher tail-subadditivity implications in practice. An increase of the β_{max} could involve higher economic reserves and, therefore, policyholders protection would be increased but it could affect the insurer's competitiveness;

- Let assume that Z random variable represents the overall risk. A set of $d \times d$ constrained optimization problems is defined at this step:

$$P_{i,j}: \quad \min_{\omega_1,\omega_2} \left| \text{GlueVaR}^{\omega_1,\omega_2}_{\beta_j,\alpha_i}(Z) - \text{VaR}_{99.5\%}(Z) \right|, \quad \text{subject to} \quad \begin{cases} 0 \leqslant \omega_1 \leqslant 1, \\ \omega_2 \geqslant 0, \\ \omega_1 + \omega_2 \leqslant 1 \end{cases} \tag{6.5}$$

  where $i, j = 1, \ldots d$, $\alpha_i = \alpha_{min} + \frac{i-1}{d-1}(\beta_{max} - \alpha_{min})$, $\beta_j = \alpha_i + \frac{j-1}{d-1}(\beta_{max} - \alpha_i)$. Flexibility rises with d but computational costs too. Constraints are fixed to guarantee that the GlueVaR satisfies α_i-tail subadditivity.;

- An optimization algorithm should be used to solve that set of problems. Note that $P_{i,j}$ problems could have not solutions. If this was the case then optimization criteria should be revised, involving a lower α_{min}, a higher β_{max} and/or a larger d. But in general, if P_{i^*,j^*} represents the problem for which the minimum value of the objective function is reached and (ω_1^*, ω_2^*) is the associated solution, then a $\text{GlueVaR}^{\omega_1^*,\omega_2^*}_{\beta_{j^*},\alpha_{i^*}}$ is found that is a candidate to satisfy the α_{i^*}-tail subadditivity property and gives similar reserve values to those obtained with $\text{VaR}_{99.5\%}$ when applied to the overall risk of the company.

- More than one GlueVaR solution would be frequently found. Alternative combinations of parameters' values would return the same value of the objective function, or with insignificant differences. In this situation, solutions could be ranked in accordance

with the underlying risk attitude involved. Aggregation indicators are useful to characterize the underlying risk behavior of the agent when using each risk measure.
Here, it is proposed to rank solutions depending on the value of aggregation indicators associated to each optimal risk measure. With this aim, local and global degrees of orness are computed for (multiple) optimal $\text{GlueVaR}^{\omega_1^*,\omega_2^*}_{\beta_{j^*},\alpha_{i^*}}$ solutions. Two particular GlueVaR measures among the set of solutions are of especial interest:

LOWER-LIMIT SOLUTION. Selection of the GlueVaR risk measure with the associated minimum degree of orness;

UPPER-LIMIT SOLUTION. Selection of the GlueVaR risk measure with the associated maximum degree of orness.

In other words, boundaries of the local degree of orness and the global degree of orness are detected or, as it has been shown in this chapter, solutions with boundary overall risk behaviors are identified. Optimal GlueVaR risk measures linked to boundaries would reflect extreme risk attitudes of agents when the random variable Z is analyzed.

The calibration procedure is not the primary topic in this study. It is not my intention to cover all possible calibration criteria. Other options could have been followed in the selection procedure. The $\text{VaR}_{99.5\%}$ is chosen as risk measure of reference due to its real application in practice, but the analysis could be easily extended to other distortion risk measures, and confidence levels. Similarly, indicators of the Choquet integral different to the degree of orness could be contemplated. Some examples of such alternatives are the divergence and the variance indicators shown in Chapter 5.

6.2.3 *Results*

An illustration of the analysis of implicit risk behavior is provided under the framework previously described. The example involves the calibration of GlueVaR risk measures for a real dataset. In the example two loss random variables X_1 and X_2 affect a motor insurance company. The aggregate risk faced by the insurer is the sum of both random variables, $Z = X_1 + X_2$. The dataset used in Chapter 3 is chosen to illustrate the calibration procedure. Recall that it is data for the cost of motor claims provided by a major Spanish motor insurer. The sample consists of $n = 518$ observations of the cost of individual claims in thousands of euros. The dataset contains X_1 and X_2, which collect the cost of property damages and the cost of bodily injury damages respectively. Total claim costs are the sum of both, $Z = X_1 + X_2$. It is assumed that the insurer uses the $\text{VaR}_{99.5\%}$ as risk

measure to manage these risks.

One of the main objectives is to illustrate how information related to the risk attitude of the agent when using $VaR_{99.5\%}$ may be enriched by means of a set of comparable GlueVaR risk measures. Before dealing with the calibration of GlueVaR risk measures, first the $VaR_{99.5\%}(Z)$ and associated local and global degrees of orness are computed. Due to the size of n the calculation of the global degree of orness is a cumbersome task. For this reason, the value of the indicator is approximated grouping data in n_0 mass points, where $n_0 = 20$. Taking this remark into account, the risk measure value is equal to $VaR_{99.5\%}(Z) = 187.6119$, the local degree of orness is equal to $\omega_L\,(VaR_{99.5\%}) = 0.9958$and the global degree of orness is $\omega_G\,(VaR_{99.5\%}) \simeq 0.9974$.Let me now modify the original dataset in the following manner. The two pairs of bivariate losses (x_1, x_2) that lead to the highest aggregated loss z are multiplied each of them by 10. That is, worst losses are artificially increased. Total claim costs in the modified dataset are represented by the random variable Z^*, and the value of the $VaR_{99.5\%}(Z^*)$ and associated degrees of orness are computed. It is obtained that values of $VaR_{99.5\%}$, ω_L and ω_G are equal in both scenarios, for the original and modified datasets. These results are in accordance with the theoretical discussion developed in Section 6.1. Note that probabilities and n remain unchanged. Only Z values located on the *specific risk loving* zone of $VaR_{99.5\%}$ have been modified. Based on this information, a risk manager comfortable with the implicit risk behavior linked to $VaR_{99.5\%}$ is not aware that he faces a sensitive increase of worst case risks. Or, until some extent, one could think that a less risk averse agent is involved in the risk assessment of the second scenario.

Let me now focus on the strategy to calibrate GlueVaR parameters. The next steps are performed to obtain GlueVaR risk measures comparable to $VaR_{99.5\%}(Z)$: a) the minimum and maximum values of confidence levels are fixed at 90% and 99.9%, i.e. $\alpha_{min} = 90\%$ and $\beta_{max} = 99.9\%$; b) the number of partitions is stipulated in $d = 25$, so 625 optimization problems are considered; c) the empirical distribution function of total claim costs is used for the risk quantification, and, finally, d) outcome GlueVaR solutions are obtained using `constrOptim` function from `rootSolve` library in R. The same phases are repeated for the modified random variable Z^*.

A set of optimal GlueVaR risk measures which return the same risk value than $VaR_{99.5\%}$ is found for each of these two particular contexts. In particular, 341 optimal solutions were found for the random variable Z and 605 in the case of Z^*. Once a set of GlueVaR risk measures is given, aggregation indicators are computed to character-

ize the underlying overall risk attitude of agents when they applied these risk measures. The attention is first focused on the local degree of orness. Boundary values of the indicator and the associated GlueVaR risk measures are reported in Table 6.1. Outcomes results are shown for the original and modified datasets. Let me emphasize that the maximum local degree of orness is equal to the local degree of orness of the $\text{VaR}_{99.5\%}$ in both scenarios. In fact, the optimal $\text{GlueVaR}^{\omega_1,\omega_2}_{\beta,\alpha}$ solution with the maximum local degree of orness is the $\text{VaR}_{99.5\%}$. Note that if $\omega_1 = \omega_2 = 0$ then $\omega_3 = 1$, so it holds $\text{GlueVaR}^{0,0}_{99.9\%,99.5\%} = \text{VaR}_{99.5\%}$. This result makes all the sense. It has been argued in Section 6.1.1 that the local degree of orness is interpreted as an approximation of the area under the distortion function g. The distortion function associated to the VaR_α assigns one to survival values higher than $(1-\alpha)$ and zero to the rest, so it has the highest possible area. In other words, given a certain risk value, the VaR_α is the risk measure with the highest area under the associated distortion function among all the distortion risk measures which return this value.

Table 6.1: Boundary values of local degree of orness of equivalent GlueVaR risk measures to $\text{VaR}_{99.5\%}$ for both Z (original dataset) and $Z*$ (modified dataset). Parameters' values of the associated GlueVaR measure are displayed.

	Minimum orness		Maximum orness	
	Original dataset	Modified dataset	Original dataset	Modified dataset
Degree of orness	**0.9122644**	**0.8960470**	**0.9957627**	**0.9957627**
α	90%	90%	99.5%	99.5%
β	99.9%	99.4875%	99.9%	99.5%
ω_1	0.188	0.0386	0	0
ω_2	$1.21 \cdot 10^{-6}$	$8.12 \cdot 10^{-7}$	0	0

Information related to the underlying overall risk behavior of the agent can be improved with the minimum degree of orness. Table 6.1 shows that, for the original dataset, there exists an optimal GlueVaR risk measure for which the area of the associated distortion function is approximately 0.9122644. So, this GlueVaR risk measure gives an equivalent reserve than $\text{VaR}_{99.5\%}$ when it is applied to Z, but, in aggregated terms, it involves a more moderate distortion of the original survival distribution function. If the local degree of orness is understood as an indicator of the overall risk aversion, it can be concluded

that the agent could be less risk averse than he is supposed to be using the $VaR_{99.5\%}$ as risk measure. Even more, the minimum local degree of orness falls to 0.8960470 when the risk value is quantified for total claim losses in the second scenario. This result would indicate that the underlying overall risk aversion of the agent could be even lower whether the $VaR_{99.5\%}$ risk measure is applied to $Z*$ instead of Z. This finding is more consistent with prior expectations because the amount of reserves does not change but the modified-data scenario is a more risky scenario involving larger worst losses.

The local degree of orness is an indicator of the overall risk behavior linked to the distortion risk measure. It has been discussed in Section 6.1.2 that this indicator evaluates the accumulated distortion made over the survival distribution function. Additional elements are required for a full understanding of the underlying risk behavior of the agent. It is claimed that the degree of orness should be complemented with the examination of the quotient function which allows the analysis of the specific risk behavior at any point of the survival function. Quotient functions associated to the optimal GlueVaR risk measures reported in Table 6.1 are examined. Remember that GlueVaR risk measures can be understood as a linear combination of a $TVaR_\beta$, a $TVaR_\alpha$ and VaR_α (see Section 1.4 in Chapter 1). Therefore, note that all quotient functions are located in the upper risk-aversion frontier in the range $[0.10, 1)$, where $Q_g(u) = \frac{1}{u}$ for $u \geqslant 0.10$. Plotting quotient functions in $[0.10, 1)$ do not provide information about differences in the implicit risk behavior. To ease comparisons, quotient functions are rescaled and their left-tails in range $(0, 0.10]$ are plotted in Figure 6.4.

Remarkable differences are observed on the specific risk attitude implicit on the left-tail of quotient functions (Figure 6.4). Let me first examine quotient functions of GlueVaR risk measures with the highest local degree of orness. In both scenarios the Q_g is the quotient function associated to the $VaR_{99.5\%}$. The agent is maximum risk averse at any point of the interval $[0.5\%, 1)$ and maximum risk lover at $(0, 0.5\%)$. That means, the quotient function is located in the upper frontier at range $[0.5\%, 1)$ and in the lower frontier at $(0, 0.5\%)$.

Distinct underlying risk behaviors are inferred when GlueVaR measures with minimum degree of orness are analyzed. Shaped patterns of left-tails of quotient functions are undoubtedly different to that linked to $VaR_{99.5\%}$. An interesting finding is that the Q_g is not located in boundaries at any point of the interval $(0, 0.10)$. That means, the risk averse attitude is not maximized in the range $[0.5\%, 0.10)$ but, in return, the agent is more risk averse to worst losses in $(0, 0.5\%)$ than when using $VaR_{99.5\%}$. Differences in the implicit risk attitude

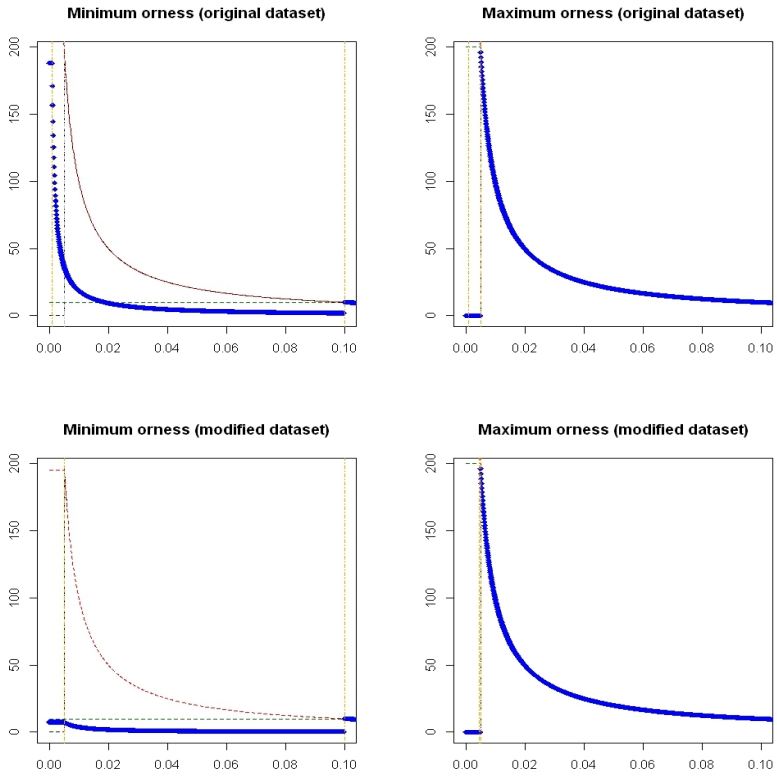


Figure 6.4: Quotient functions of optimal solutions with boundary local degrees of orness in both scenarios.

between scenarios deserve some attention. In accordance with my prior beliefs, the specific risk aversion of the agent is lower in the scenario involving higher worst losses. It can be interpreted as an alert that either a change on the risk behavior of the agent or a change on faced risk is detected. Focusing on the original sample scenario, the specific risk behavior implicit in the risk measure changes for survival probabilities in the range $(0, 0.10]$ as shown in Figure 6.4. Precisely, a risk loving attitude underlies in the interval $[0.02, 0.10)$, where quotient function values stands below one (risk neutrality frontier). On the opposite, a risk averse behavior is attributed in the range $(0, 0.02)$, although maximum risk aversion is not attained at any point. Let emphasize that a risk loving attitude is implicit in the whole tail range for the modified sample scenario. Here, two different risk patterns are also derived. A no-constant risk loving attitude is observed in the range $[0.5\%, 0.10)$ while a constant risk loving behavior in $(0, 0.5\%)$.

Before concluding, it is convenient to appoint that the same analysis of the underlying risk behavior was performed for the global degree of orness. The global degree of orness was computed to the set of optimal GlueVaR risk measures and quotient functions of risk measures with boundary indicators were displayed. Findings were qualitatively the same than results obtained from the analysis of the local degree of orness. To avoid an unnecessary extending of the study they are not reproduced in the text. Interested reader may find main results about the global degree of orness in the Appendix.

Part III

ON CAPITAL ALLOCATION PROBLEMS AND AGGREGATION FUNCTIONS

7

AN OVERWIEW ON CAPITAL ALLOCATION PROBLEMS

Capital allocation problems of insurance and financial institutions arise when a management unit must distribute an amount among different business units. This amount may be the total costs faced by the company, its solvency capital requirement or the total variable economic compensation to be shared across business units, among other examples. This kind of problems are frequent and relevant from an Enterprise Risk Management (ERM) perspective, mainly if the risk that each business unit faces is, somehow, taken into account for the final allocation. Sometimes the capital allocation is merely notional, as pointed out in Dhaene et al. [2012b]. This does not diminish the importance of studying these problems at all, because the allocation may be useful to conduct different business analyses in order to improve the risk management.

There is a large number of academic works related to capital allocation problems. An extensive literature can be found discussing solutions to capital allocation problems [see, among others. Denault, 2001; Kalkbrener, 2005; Tsanakas, 2009; Buch et al., 2011; van Gulick et al., 2012]. Some recent literature focuses on specific probability distributions of losses Cossette et al. [2012, 2013], risk dependence structures Cai and Wei [2014], asymptotic of capital allocations based on commonly used risk measures Asimit et al. [2011] or modifications of the optimization function to overcome limitations of allocations based on minimizing the loss function [Xu and Hu, 2012; Xu and Mao, 2013]. To more precisely detail some recent contributions, You and Li [2014] analyze capital allocation problems concerning mutually interdependent risks, mainly where they are tied through an Archimedean copula. Wang [2014] investigates the usefulness of the Tail Covariance Premium Adjusted principle in the case of two business lines with exponentially distributed losses, where their dependence structure corresponds to a Farlie-Gumbel-Morgenstern copula. Zaks and Tsanakas [2014] generalize the framework proposed in Dhaene et al. [2012b], allowing the inclusion of different hierarchical levels of preferences about risk in the final solution. And in Urbina and Guillén [2014] several principles are examined to solve a capital allocation problem related to operational risk. This list of academic contributions on capital allocation problems is not exhaustive. In fact, this topic is object of ongoing research. Two very recent examples are Tsanakas and Mil-

lossovich [2014] and Li and You [2015].

Main concepts and notations used all along the dissertation regarding capital allocation problems are introduced in next section. These are the building blocks on which the rest of the discussion is based and represent a necessary starting point to go further in next chapters. The overview on capital allocation problems is completed with a description of some particularly interesting solutions and with a list of properties that particular solutions to these problems may satisfy.

7.1 MAIN CONCEPTS AND NOTATION

In general terms, a capital allocation problem may be understood in the following way:

> "An amount $K > 0$ has to be distributed across $n \in \mathbb{N}$ agents, and the allocation must be a *full allocation*."

Described in such a way, it seems reasonable to think about these problems as *disaggregation problems* and, to some extent, as the opposite questions that are answered through aggregation operators. Nevertheless, connections between capital allocation problems and aggregation functions can be found, and such connections are shown in next chapters.

Returning to capital allocation problems, several comments must be made. First of all, on the risk management framework in which this kind of problems arises. From my point of view, even the name given to these problems is strongly related to the fact that risk managers from the insurance and banking industries must determine, at different levels of granularity, the contributions of *agents* to the risk-based regulatory capital required to companies. In that sense, the concept of agent must be understood in a broad way: it may be a commercial agent, a business unit, a branch of the overall business or even a particular guarantee included in a set of contracts. Nonetheless, it has to be noted that similar risk management problems are faced by asset management firms when planning investment strategies or when assessing performance of their investment portfolios. In such contexts it is more usual to refer to these problems as *risk attribution or risk budgeting problems* [see, for instance, Grégoire, 2007; Rahl, 2012].

Secondly, it is important to list the main elements that play a role in a capital allocation problem. My opinion on this particular issue is that each one of these problems may be described by means of the following elements:

- The capital $K > 0$ to be distributed;
- The agents, indexed by $i = 1, \ldots, n$;
- Random variables linked to each agent, $\{X_i\}_{i=1,\ldots,n}$;
- Functions $f_i, i = 1, \ldots, n$ used to simplify the information provided by each X_i;
- A distribution criterion;
- Capitals $K_i, i = 1, \ldots, n$ assigned to each agent as a solution to the problem;
- The goal. Some examples are *cost of risk* allocation, *reward to riskless* allocation or *reward on risk&return* allocation.

In words, a solution to a capital allocation problem is a set of n capitals $\{K_i\}_{i=1,\ldots,n}$ which add up to K, this is, $\sum_{j=1}^{n} K_j = K$ (*full allocation*). Capital K_i is the one assigned to the ith agent and it is related to the risk X_i faced by that agent. Commonly risks $X_i, i = 1, \ldots, n$ are random variables representing losses. Usually a solution to a capital allocation problem is also known as a capital allocation principle. One of the fundamental elements characterizing a capital allocation principle is the distribution criterion, which drives the allocation. Distribution criteria may be classified under several perspectives, but a relevant one is the proportionality perspective. Proportional allocation criteria are such that each capital $K_i, i = 1, \ldots, n$ may be expressed as the product of capital K times a proportion of the form $\frac{f_i(X_i)}{\sum_{j=1}^{n} f_j(X_j)}$, where f_i functions simplify all the information provided by risk X_i, either in a *stand-alone* way or in a *marginal* way with respect to the rest of risks $\{X_j\}_{j \neq i}$. Therefore, the general expression for a proportional allocation principle is

$$K_i = K \cdot \frac{f_i(X_i)}{\sum_{j=1}^{n} f_j(X_j)} \quad , i = 1, \ldots, n. \tag{7.1}$$

Frequently, functions f_i are risk measures or partial contributions to the value that a risk measure assigns to the whole random loss understood as $S = \sum_{j=1}^{n} X_j$. If f_i are simply risk measures, the proportional allocation principle is classified as an *stand-alone* one. On the other hand, when dealing with $\{f_i\}_{i=1,\ldots,n}$ which represent partial contributions to $\rho(S)$ for a given risk measure ρ, the proportional allocation principle is known to be based on *marginal or partial contributions*. The name is inherited by the fact that expression $\rho(S) = \sum_{j=1}^{n} f_j(X_j)$ holds. In such those cases, notation $f_i(X_i) = \rho(X_i \mid S)$ is going to

be used and, therefore, the general expression for proportional allocation principles based on partial contributions is

$$K_i = K \cdot \frac{\rho\left(X_i \mid S\right)}{\rho\left(S\right)} \quad , i = 1, \ldots, n. \tag{7.2}$$

Among proportional capital allocation principles, the main difference between stand-alone principles and the ones based on partial contributions is related to diversification effects. Stand-alone principles do not take into account neither benefits nor penalizations on risk of each ith agent due to fact that the agent belongs to a set of agents, while principles based on partial contributions do.

Non-proportional allocation principles are such that an expression like (7.1) for each of the assigned capitals K_i, $i = 1, \ldots, n$ cannot be achieved. An example of this kind of principles is the *excess based allocation principle* shown in section 7.2.3. An important number of such principles can be devised when using the so-called *quadratic optimization criterion* in the framework provided by Dhaene et al. [2012b]. For instance, principles like

$$K_i = \rho_i(X_i) + \nu_i \cdot \left[K - \sum_{j=1}^{n} \rho_j(X_j) \right] \quad , \tag{7.3}$$

where ρ_i, $i = 1, \ldots, n$ are risk measures and ν_i are weights such that $\sum_{j=1}^{n} \nu_j = 1$, which satisfy that at least one of the ν_i, say ν_{i_0}, is not equal to $\rho_i\left(X_i\right) / \left(\sum_{j=1}^{n} \rho_j\left(X_j\right)\right)$. In other words, there is an $i_0 \in \{1, \ldots, n\}$ such that $\nu_{i_0} \neq \rho_{i_0}\left(X_{i_0}\right) / \left(\sum_{j=1}^{n} \rho_j\left(X_j\right)\right)$.

Last but not least, a major feature of a capital allocation problem is its goal. Three sort of purposes are listed before: *cost of risk*, *reward to riskless* and *reward on risk&return*. A brief description of each one follows. Other alternative goals may be considered. The aim of a capital allocation problem with a *cost of risk* goal is to distribute a cost among the agents by taking into account some measure of the risk faced by each one of them. An example of such a *cost of risk* allocation should be the disaggregation of the Solvency Capital Requirement (SCR) of the whole business of an European insurance company under the Solvency II regime among its lines of business. On the opposite side, when the management team wants to stimulate a risk averse attitude among the business units it has in charge, this management team may adopt a compensation scheme based on the following idea: the riskier the business unit is, the lesser the reward it receives. In such a situation, a capital allocation problem with a *reward to riskless* objective is conducted. It is my opinion that this kind of problems rarely appears in practice, but one can think about them theoretically. For instance, it could be used to notionally distribute the contribution of

each agent to the overall diversification benefit, where only there is information about a final *cost of risk* allocation and the overall diversification benefit. That is, where there is not information about each individual diversification benefit. A much more usual situation is to take into account *reward on risk&return* allocations, in order to better reward those agents whose trade-off between return obtained and risk faced is higher. Therefore, some return-on-risk measure seems to be the natural choice of functions $\{f_i\}_{i=1,\dots,n}$ to assign rewards under this perspective.

Frequently I am going to denote a principle by $\vec{K} = (K_1, K_2, \dots, K_n)$. An abuse of notation is made because K is used both to denote the vector $\vec{K}$ whose components are K_i and the amount to be distributed among agents. Given an (absolute) capital allocation principle $\vec{K}$ with $K = \sum_{j=1}^{n} K_j$, its relative counterpart is defined as $\vec{x}$, where components are $x_i = K_i/K$ and satisfy that $\sum_{j=1}^{n} x_j = 1$. If there is no room for confusion, upper-case letters mean absolute principles while lower-case letters mean relative ones. This notation is going to be extensively used in chapter 9.

7.2 REVIEW OF SOME PARTICULAR PRINCIPLES

A collection of particular principles is commented in this section. This section does not pursue to be a deep review of principles found in the literature. Both the selection and discussion of principles have been purely driven by subjective criteria. In most cases, the principles have been chosen in order to have a better understand of them. That is, explaining them as a way to achieve this goal. The attention has been paid to specific issues of each capital allocation principle, so an irregular extension is found in the discussion of each principle.

7.2.1 *The gradient allocation principle*

This principle is also known as Euler allocation principle [McNeil et al., 2005] or, from a game-theoretic perspective, as Aumann-Shapley allocation principle [Denault, 2001]. According to Tasche [1999, 2004, 2007] capital allocation principles based on the gradient are the most appropriate allocation principles to deal with risk adjusted returns. As long as I believe that *reward on risk&return* allocations could be specially useful in sound ERM systems, key elements of gradient allocation principles and their usefulness as *reward on risk&return* allocations are discussed hereinafter. In my opinion, the basic idea that must be remarked is that the gradient allocation principle takes advantage of the Euler's theorem on homogeneous functions applied to positively homogeneous risk measures. Let me discuss this point,

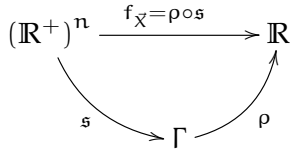
beginning with a necessary definition and the statement of a theorem.

Definition 7.2.1 (Homogeneous function of degree r)**.** *Let* f *be a function from* $\mathbb{R}^n$ *to* $\mathbb{R}$, $n > 0$.
f *is homogeneous of degree* $r \Leftrightarrow \forall \lambda \in \mathbb{R} \quad f(\lambda \cdot \vec{u}) = \lambda^r \cdot f(\vec{u}) \quad .$

Theorem 7.2.1 (Euler's theorem on homogeneous functions)**.** *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be a differentiable function on* $\mathbb{R}^n$. *Then,*
f *is an homogeneous function of degree* $r \Leftrightarrow \sum_{i=1}^{n} u_i \cdot \frac{\partial f}{\partial u_i}(\vec{u}) = r \cdot f(\vec{u}) \quad .$

It has to be noted that a differentiable function f defined from $\mathbb{R}^n$ to $\mathbb{R}$ has a gradient equal to $\nabla f(\vec{u}) = \left(\frac{\partial f}{\partial u_1}(\vec{u}), \frac{\partial f}{\partial u_2}(\vec{u}), \ldots, \frac{\partial f}{\partial u_n}(\vec{u}) \right)$ and, therefore, the right-hand side of the equivalence in Theorem 7.2.1 can be also written as $< \vec{u}, \nabla f(\vec{u}) >= r \cdot f(\vec{u})$, where $<,>$ stands for the interior product in $\mathbb{R}^n$.

Consider now a positively homogeneous risk measure ρ. This means that $\rho(\lambda \cdot X) = \lambda \cdot \rho(X)$ for all $\lambda \geqslant 0$ and for all $X \in \Gamma$. Now, given a random vector $\vec{X} = (X_1, X_2, \ldots, X_n) \in \Gamma^n$, consider the following function $f_{\vec{X}}$ as well:



$$\vec{u} \quad \mapsto \quad s(\vec{u}) = \sum_{i=1}^{n} u_i X_i \quad \mapsto \quad \rho(s(\vec{u})) = \rho\left(\sum_{i=1}^{n} u_i X_i \right) = f_{\vec{X}}(\vec{u}) \tag{7.4}$$

Taking into account definition 7.2.1 restricted to $(\mathbb{R}^+)^n$; that ρ is a positively homogeneous risk measure; and that $\sum_{i=1}^{n} u_i X_i \in \Gamma$ for all $\vec{u} \in (\mathbb{R}^+)^n$ if $\vec{X} \in \Gamma^n$, then the fact that $f_{\vec{X}}$ is an *homogeneous function of degree* $r = 1$ is deduced. In such a case, applying Theorem 7.2.1 the following expression holds:

$$f_{\vec{X}}(\vec{u}) = \sum_{i=1}^{n} u_i \cdot \frac{\partial f_{\vec{X}}}{\partial u_i}(\vec{u}) =< \vec{u}, \nabla f_{\vec{X}}(\vec{u}) > \quad . \tag{7.5}$$

If $\vec{u} = (1, 1, \ldots, 1)$ then the sum $\sum_{i=1}^{n} u_i X_i$ is the sum of all the components of the random vector $\vec{X}$. From this point forward, this sum is denoted as S, so $S = \sum_{i=1}^{n} X_i$. If the following abuse of notation is used

$$\frac{\partial \rho}{\partial u_i}(S) = \frac{\partial f_{\vec{X}}}{\partial u_i}(\vec{u})_{|\vec{u}=(1,1,\ldots,1)} \quad ,$$

therefore expression (7.5) becomes

$$\rho(S) = \sum_{i=1}^{n} \frac{\partial \rho}{\partial u_i}(S) \quad , \tag{7.6}$$

which is the formula usually related to the underlying idea of the Euler allocation principle.

Regarding the idea that the gradient allocation principle is *reward on risk&return* compatible, let me first consider a particular *Return on Risk Adjusted Capital* (RORAC) measure, and some concepts and notations taken from Tasche [2007]:

Definition 7.2.2. *The total RORAC of portfolio* $S = \sum_{i=1}^{n} X_i$ *is defined by*

$$\mathrm{RORAC}(S) = \frac{-\mathbb{E}(S)}{\rho(S)} \quad ,$$

where ρ *is a risk measure and each random variable* X_i, $i = 1, \ldots, n$ *represents a loss.*

Definition 7.2.3. *Given a portfolio* $S = \sum_{i=1}^{n} X_i$ *and a set of contributions* $\rho(X_i \mid S)$, $i = 1, \ldots, n$ *to the value of the risk of the portfolio measured by* ρ, *i.e.* $\rho(S) = \sum_{i=1}^{n} \rho(X_i \mid S)$, *the portfolio-related RORAC of each loss* X_i *is defined by*

$$\mathrm{RORAC}(X_i \mid S) = \frac{-\mathbb{E}(X_i)}{\rho(X_i \mid S)}, \quad \forall i = 1, \ldots, n \quad .$$

First thing to be noted is that, in common situations, numerators in definitions 7.2.2 and 7.2.3 are positive, because the mathematical expectations of S and X_i, $i = 1, \ldots, n$ will be, hopefully, negative: it may be assumed that ith business unit does not expect losses, so $\mathbb{E}(X_i) < 0$ due to the fact that positive values of X_i mean losses. Secondly, it has to be remarked that definition 7.2.3 depends on both portfolio S and partial contributions to $\rho(S)$. Bearing these two previous definitions in mind, let me now present what the RORAC compatibility is as defined in Tasche [2007]:

Definition 7.2.4 (RORAC compatible risk contributions)**.** *Risk contributions* $\rho(X_i \mid S), i = 1, \ldots, n$ *are RORAC compatible if there are some* $\epsilon_i > 0, i = 1, \ldots, n$ *such that*

$$\mathrm{RORAC}(X_i \mid S) > \mathrm{RORAC}(S) \Rightarrow$$
$$\mathrm{RORAC}(S + hX_i) > \mathrm{RORAC}(S) \quad \textit{for all} \quad 0 < h < \epsilon_i$$

In words, this means that if the partial return&risk performance of ith agent given by definition 7.2.3 is greater than the return&risk performance of the overall portfolio given by definition 7.2.2 then the return&risk performance of the overall portfolio is improved by slightly

increasing the position of ith agent in the portfolio. Alternatively, if a $\text{RORAC}(X_i \mid S)$ greater than $\text{RORAC}(S)$ is found and contribution $\rho(X_i \mid S)$ is RORAC compatible, therefore the position on ith agent should be increased in order to improve the overall performance of the portfolio.

Assuming that it is possible to slightly increase the position of ith agent in the portfolio, going from X_i to $X_i \cdot (1+h)$ with $h \in (0, \epsilon_i)$, the necessary condition expressed in definition 7.2.4 can be understood as equivalent to the following one

$$\frac{\partial \text{RORAC}}{\partial u_i}(\mathfrak{s}(\vec{u}))_{|\vec{u}=(1,1,\ldots,1)} > 0 \quad , \tag{7.7}$$

simply by computing

$$\lim_{h\to 0} \frac{1}{h}\left[\text{RORAC}(S + hX_i) - \text{RORAC}(S)\right].$$

Taking advantage of expression (7.7) the RORAC compatibility of the gradient allocation principle can be proved.

Proposition 7.2.1. *Suppose that* $\rho(\mathfrak{s}(\vec{u}))$ *and* $\frac{\partial f_{\vec{X}}}{\partial u_i}(\vec{u})$ *for all* $i = 1, \ldots, n$ *are strictly positive. A gradient allocation principle* $\vec{K} \in \mathbb{R}^n$ *of the form* $K_i = K \cdot \frac{\rho(X_i \mid S)}{\rho(S)}$ *where risk contributions are* $\rho(X_i \mid S) = \frac{\partial \rho}{\partial u_i}(S)$ *for all* $i = 1, \ldots, n$*, is such that all the risk contributions are RORAC compatible.*

Proof of Proposition 7.2.1. Let me show that expression (7.7) holds for each $i = 1, \ldots, n$:

$$\begin{aligned}
\frac{\partial \text{RORAC}}{\partial u_i}(\mathfrak{s}(\vec{u})) &= \frac{\partial}{\partial u_i}\left[\frac{-\mathbb{E}\left(\sum_{j=1}^n u_j \cdot X_j\right)}{\rho\left(\sum_{j=1}^n u_j \cdot X_j\right)}\right] \\
&= \frac{\frac{\partial}{\partial u_i}\left[-\mathbb{E}\left(\sum_{j=1}^n u_j \cdot X_j\right)\right] \cdot \rho\left(\sum_{j=1}^n u_j \cdot X_j\right)}{\left[\rho\left(\sum_{j=1}^n u_j \cdot X_j\right)\right]^2} \\
&\quad - \frac{\left[-\mathbb{E}\left(\sum_{j=1}^n u_j \cdot X_j\right)\right] \cdot \frac{\partial}{\partial u_i}\rho\left(\sum_{j=1}^n u_j \cdot X_j\right)}{\left[\rho\left(\sum_{j=1}^n u_j \cdot X_j\right)\right]^2} \\
&= \frac{\frac{\partial}{\partial u_i}\left[-\mathbb{E}\left(\sum_{j=1}^n u_j \cdot X_j\right)\right] \cdot \rho(\mathfrak{s}(\vec{u})) - [-\mathbb{E}(\mathfrak{s}(\vec{u}))] \cdot \frac{\partial f_{\vec{X}}}{\partial u_i}(\vec{u})}{[\rho(\mathfrak{s}(\vec{u}))]^2} \\
&= \frac{-\mathbb{E}(X_i) \cdot \rho(\mathfrak{s}(\vec{u})) + [\mathbb{E}(\mathfrak{s}(\vec{u}))] \cdot \frac{\partial f_{\vec{X}}}{\partial u_i}(\vec{u})}{[\rho(\mathfrak{s}(\vec{u}))]^2} \quad .
\end{aligned}$$

As long as the denominator of the previous expression is always positive, therefore it is deduced that $\frac{\partial \text{RORAC}}{\partial u_i}(\mathfrak{s}(\vec{u})) > 0$ if and only

if $-\mathbb{E}(X_i) \cdot \rho(\mathfrak{s}(\vec{u})) + \mathbb{E}(\mathfrak{s}(\vec{u})) \cdot \frac{\partial f_{\vec{X}}}{\partial u_i}(\vec{u}) > 0$. Consider that both $\rho(\mathfrak{s}(\vec{u}))$ and $\frac{\partial f_{\vec{X}}}{\partial u_i}(\vec{u})$ are strictly positive: these conditions may usually hold, because of dealing with risk values or risk contributions of a portfolio of risky positions. Being this the case, this last expression may be written as

$$\frac{-\mathbb{E}(X_i)}{\frac{\partial f_{\vec{X}}}{\partial u_i}(\vec{u})} > \frac{-\mathbb{E}(\mathfrak{s}(\vec{u}))}{\rho(\mathfrak{s}(\vec{u}))} \quad .$$

Moreover, where restricted to $\vec{u} = (1, 1, \ldots,)$ this last expression is providing the following information:

$$\frac{\partial \text{RORAC}}{\partial u_i}(\mathfrak{s}(\vec{u}))_{|\vec{u}=(1,1,\ldots,1)} > 0, \text{if and only if}$$

$$\frac{-\mathbb{E}(X_i)}{\frac{\partial \rho}{\partial u_i}(S)} > \frac{-\mathbb{E}(S)}{\rho(S)} \Leftrightarrow \frac{-\mathbb{E}(X_i)}{\rho(X_i \mid S)} > \frac{-\mathbb{E}(S)}{\rho(S)}$$

$$\Leftrightarrow \text{RORAC}(X_i \mid S) > \text{RORAC}(S) \quad .\square$$

Some final comments on the gradient allocation principle. It is an elegant approach to proportional capital allocation principles based on partial contributions, but it has, from my point of view, two main drawbacks. On the one hand, the assumption that infinitesimal (or very small) perturbations on the risky position of an agent can be made is often not realistic in practice. Frequently, it is feasible that limited small changes can be made but not arbitrarily small changes. Being this the case, the compatibility of RORAC contributions should be barely satisfied even for the risk contributions linked to the gradient allocation principle. The second limitation is related to the computation of risk contributions $\frac{\partial f_{\vec{X}}}{\partial u_i}(\vec{u})$, where the value of the risk measure ρ for sums $\mathfrak{s}(\vec{u})$, $\vec{u} \in \mathbb{R}^n$ can not be expressed in an analytic closed-form expression. This is, probably, the most frequent practical situation. In most of those cases, some decisions must be taken in order to do an allocation (nearly) based on the gradient. This drawback is known in Tasche [2007], for instance, where the author shows how the risk contributions of X_i to the VaR_α of the portfolio calculated using kernel estimators do not add up to the natural estimators of VaR_α of the portfolio S. Although differences tend to be small, this kind of issues break the theoretical elegance of these principles when they are applied in a real context.

7.2.2 *Other proportional capital allocation principles based on partial contributions*

There are other examples of proportional capital allocation principles based on partial contributions fitting expression (7.2). Two exam-

ples are given here, one from a probabilistic perspective and another one from a game-theoretic perspective.

7.2.2.1 *The covariance allocation principle*

This principle is proposed, for instance, in Overbeck [2000]. It takes into account the variance as the risk measure for the whole portfolio: $\rho(S) = \mathbb{V}(S)$. The partial contribution of the ith agent X_i is the covariance of X_i with respect to S, so $\rho(X_i \mid S) = \mathrm{Cov}(X_i, S)$. Therefore, this principle is expressed as

$$K_i = K \cdot \frac{\mathrm{Cov}(X_i, S)}{\mathbb{V}(S)}, \quad \forall i = 1, \ldots, n. \tag{7.8}$$

Note that $\rho(S) = \sum_{j=1}^{n} \rho(X_i \mid S)$ because of the (bi)linearity of the covariance:

$$\begin{aligned} \rho(S) = \mathbb{V}(S) = \mathrm{Cov}(S, S) = \mathrm{Cov}(\textstyle\sum_{j=1}^{n} X_j, S) \\ = \sum_{j=1}^{n} \mathrm{Cov}(X_j, S) = \sum_{j=1}^{n} \rho(X_i \mid S) \quad . \end{aligned}$$

This is an interesting case because, from the perspective of the Euler's Theorem on homogeneous functions, it can be understood in two different (but related) ways. First interpretation is as follows: take as risk measure ρ the variance in expression (7.4), in order to interpret the covariance principle similarly to a gradient principle. The resulting function $f_{\vec{X}} = \mathbb{V} \circ \mathfrak{s}$ is not an homogeneous function of degree $r = 1$ but an *homogeneous function of degree* $r = 2$, because the variance is not a positively homogeneous risk measure but satisfy the following relationship: for all $\lambda \in \mathbb{R}$ and for all $X \in \Gamma$, $\mathbb{V}(\lambda \cdot X) = \lambda^2 \cdot \mathbb{V}(X)$. From Theorem (7.2.1) this means that expression

$$2 \cdot \mathbb{V}\Big(\sum_{j=1}^{n} u_j \cdot X_j\Big) = \sum_{i=1}^{n} u_i \cdot \frac{\partial \mathbb{V}(\sum_{k=1}^{n} u_k \cdot X_k)}{\partial u_i} \tag{7.9}$$

holds or, in other words, that

$$\left[\frac{1}{2} \cdot \frac{\partial \mathbb{V}(\sum_{k=1}^{n} u_k \cdot X_k)}{\partial u_i}\right]_{|\vec{u}=(1,1,\ldots,1)} = \mathrm{Cov}(X_i, S) \, .$$

Let me check this last equivalence:

$$\begin{aligned}
&\frac{\partial}{\partial u_i}\mathbb{V}\left(\sum_{j=1}^n u_j \cdot X_j\right) \\
&= \frac{\partial}{\partial u_i}\mathrm{Cov}\left(\sum_{j=1}^n u_j \cdot X_j, \sum_{k=1}^n u_k \cdot X_k\right) \\
&= \frac{\partial}{\partial u_i}\left[\sum_{j=1}^n u_j \cdot \mathrm{Cov}\left(X_j, \sum_{k=1}^n u_k \cdot X_k\right)\right] \\
&= \frac{\partial}{\partial u_i}\left[\sum_{j=1}^n \sum_{k=1}^n u_j \cdot u_k \cdot \mathrm{Cov}\left(X_j, X_k\right)\right] \\
&= \frac{\partial}{\partial u_i}\left[\sum_{k=1}^n u_i \cdot u_k \cdot \mathrm{Cov}\left(X_i, X_k\right)\right. \\
&\quad \left. + \sum_{j\neq i}\sum_{k=1}^n u_j \cdot u_k \cdot \mathrm{Cov}\left(X_j, X_k\right)\right] \\
&= \frac{\partial}{\partial u_i}\left[u_i^2 \cdot \mathrm{Cov}\left(X_i, X_i\right) + \sum_{k\neq i} u_i \cdot u_k \cdot \mathrm{Cov}\left(X_i, X_k\right)\right. \\
&\quad + \sum_{j\neq i}\sum_{k\neq i} u_j \cdot u_k \cdot \mathrm{Cov}\left(X_j, X_k\right) \\
&\quad \left. + \sum_{j\neq i} u_j \cdot u_i \cdot \mathrm{Cov}\left(X_j, X_i\right)\right] \\
&= 2 \cdot u_i \cdot \mathrm{Cov}\left(X_i, X_i\right) + 2 \cdot \sum_{k\neq i} u_k \cdot \mathrm{Cov}\left(X_i, X_k\right) \\
&= 2 \cdot \mathrm{Cov}\left(X_i, \mathfrak{s}\left(\vec{u}\right)\right)
\end{aligned}$$

If last expression is evaluated at $\vec{u} = (1, 1, \ldots, 1)$ then the desired result is found.

The second interpretation allows to understand the covariance allocation principle as a pure gradient allocation principle as explained in section 7.2.1. The key is to consider as risk measure ρ in (7.4) the covariance of a random variable with respect to the sum S of the components of $\vec{X}$. In other words, take as function $f_{\vec{X}}$ the following one: $f_{\vec{X}} = \mathrm{Cov}(\cdot, S) \circ \mathfrak{s}$. As long as $\mathrm{Cov}(\lambda \cdot X, S) = \lambda \cdot \mathrm{Cov}(X, S)$ for all $\lambda \in \mathbb{R}$ and for all $X \in \Gamma$, $f_{\vec{X}}$ is an homogeneous function of degree $r = 1$ and Theorem 7.2.1 may be applied in this case as in Proposition 7.2.1.

Finally, some comments on strengths and weaknesses of the covariance principle may be pointed out. As a strength in front of other gradient allocation principles, estimators of both $\mathbb{V}(S)$ and $\mathrm{Cov}(X_i, S)$ for all $i = 1, \ldots, n$ can be found satisfying that the sum of the estimated covariances add up to the estimated variance of the overall portfolio, whatever set of random variables $\{X_i\}_{i=1,\ldots,n}$ is in place. In this sense, the covariance principle skips the second drawback commented at the end of the previous section. As a weakness, the allocation only takes care of linear dependence structures between random variables X_i, $i = 1, \ldots, n$, and may lead to negative allocated capitals K_i.

It has to be mentioned that recently published research [Wang, 2014] is inspired both by the covariance allocation principle and the tail variance risk measure presented in Furman and Landsman [2006]. In there, the authors define what they call capital allocation principles based on the Tail Covariance Premium Adjusted.

7.2.2.2 *The Shapley value principle and one of its simplifications*

Another proportional allocation principle based on partial contributions can be derived by understanding the capital allocation problem as a cooperative game in which capital K has to be *fairly* shared by the agents, taking into account that the cost of a coalition is linked to the risk that this coalition assumes. It is not the aim of this discussion to go deeper in this interpretation, but it must be mentioned that the key concept to find such a fair allocation is the Shapley value (sometimes referred to as Bondareva-Shapley value).

Let me use the following notations: $N = \{1, \ldots, n\}$, $A \subseteq N$ denotes a subset of N with cardinality $a = |A|$ and $\mathfrak{rho}(A) = \rho\left(\sum_{k \in A} X_k\right)$. A capital allocation principle based on the Shapley value is of the form (7.2), where

$$\rho(X_i \mid S) = \sum_{A \subseteq N \setminus \{i\}} \frac{a! \cdot (n - a - 1)!}{n!} \cdot [\mathfrak{rho}(A \cup \{i\}) - \mathfrak{rho}(A)] \quad . \tag{7.10}$$

Note that $\mathfrak{rho}(N) = \rho(S)$. Additionally, it can be proved that $\rho(S) = \sum_{i=1}^{n} \rho(X_i \mid S)$ using the properties of the Shapley value. The contribution of ith agent to the overall risk is, basically, a weighted average of all the marginal contributions that ith agent makes on the risk of each of the coalitions that can be obtained without ith agent. As it can be deduced, this principle can be unfeasible where n grows due to the high computational demand of each $\rho(X_i \mid S)$. This is, for sure, the main concern when trying to use this principle.

In order to avoid this drawback, some authors propose an alternative that is a simplification of this principle. In Balog [2010] this alternative is called *incremental principle*. It is built reducing the terms added up in expression (7.10) only to the one linked to the set $N \setminus \{i\}$. In other words, the incremental principle is of the form (7.1) where

$$f_i(X_i) = \mathfrak{rho}(N) - \mathfrak{rho}(N \setminus \{i\}) = \rho(S) - \rho(\sum_{j \neq i} X_j) \quad , \forall i = 1, \ldots, n.$$

This principle assigns as partial contribution of ith agent the difference between the overall risk and the risk quantified in absence of the ith agent. Until some extent, this principle can be considered as an hybrid between a stand-alone proportional principle and a proportional principle based on partial contributions: it seems clear that $\sum_{j=1}^{n} f_j(X_j) \neq \rho(S)$ and, therefore, can not be considered a pure proportional principle based on partial contributions. But, at the same time, it also seems clear that some relationship between ith agent and the rest of participants is taken into account by f_i, so it can not be considered a pure stand-alone proportional principle. In my opinion, this hybridization and the loss of certain information is the price that

must be paid in order to skip the computational deal of the Shapley value for large n.

7.2.3 *The excess based allocation principle*

The last principle commented in this chapter is the one proposed in van Gulick et al. [2012] and the reason is twofold. On the one hand, because of its originality and, on the other hand, because of its non-proportionality. Taken the authors' own words [cf. page 29]

"The allocation rule that we propose determines the allocation that lexicographically minimizes the portfolio's excesses among a set of allocations that satisfies two basic properties. First, no portfolio is allocated more risk capital than the amount of risk capital that it would need to withhold if it were on its own. Second, a portfolio is not allocated less than the minimum loss it can incur".

To better understand this principle, the following definition from van Gulick et al. [2012] must be presented.

Definition 7.2.5 (Lexicographical ordering)**.** *For $m \in \mathbb{N}$ and any two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$, $\vec{x}$ is lexicographically strictly smaller than $\vec{y}$, denoted as $\vec{x} <_{lex} \vec{y}$, if there exists an $i \leqslant m$ such that $x_i < y_i$, and for all $j < i$ it holds that $x_j = y_j$. Moreover, $\vec{x}$ is lexicographically smaller than $\vec{y}$, denoted by $\vec{x} \leqslant_{lex} \vec{y}$, if $\vec{x} = \vec{y}$ or $\vec{x} <_{lex} \vec{y}$.*

The authors are always considering a coherent risk measure ρ and that the capital K to be shared among the agents is, in fact, equal to $\rho(S)$. They use notation $N = \{1, \ldots, n\}$. Being these preliminaries established, the idea of the excess based allocation principle may be outlined in four steps:

(i) Consider any capital allocation principle $\vec{K}$ such that $\sum_{j=1}^{n} K_j = \rho(S)$ and such that the following boundary conditions are satisfied for all $i \in N$: $\max\{0, \min_{\omega\in\Omega} X_i(\omega)\} \leqslant K_i \leqslant \rho(X_i)$. The set of all the principles satisfying these conditions is called the *set of feasible principles*. Let me note it as $\mathfrak{F}$.

(ii) Compute, for each feasible principle, the vector of dimension 2^n consisting in $\bar{e}(\vec{K}) = (\mathbb{E}[\left(\sum_{j\in A}(X_j - K_j)\right)_+])_{A\subseteq N}$. So, there is a component for each subset $A \subseteq N$, and each component is the mathematical expectation of the random variable that represents the non-negative excess of capital that principle $\vec{K}$ assigns to coalition A.

(iii) For each feasible principle $\vec{K}$, order the components of $\bar{e}(\vec{K})$ in a decreasing way. The ordered resulting vector in $\mathbb{R}^n$ is denoted by $\theta[\bar{e}(\vec{K})]$.

(iv) The excess based allocation principle, denoted by $\vec{K}_{EBA}$, is the feasible principle which lexicographically minimizes the $\theta[\bar{e}(\vec{K})]$. In other words, $\vec{K}_{EBA}$ is chosen among all feasible principles as the principle associated to the first position in set of the ordered $\theta[\bar{e}(\vec{K})]$, supposing that this order is similar to the one provided by a librarian who was increasingly ordering vectors $\theta[\bar{e}(\vec{K})]$ alphabetically.

Obviously, last comment on step (iv) is not formal. A more precise way to present the excess based allocation principle is by

$$\vec{K}_{EBA} = \left\{ \vec{K} \in \mathfrak{F} \mid \quad \theta[\bar{e}(\vec{K})] \leqslant_{lex} \theta[\bar{e}(\vec{C})] \quad \forall \vec{C} \in \mathfrak{F} \right\} \quad , \qquad (7.11)$$

taking into account that the set at the right-hand side of expression (7.11) is a single value set [as proved in van Gulick et al., 2012] and that, therefore, there is an abuse of notation identifying a set consisting on a single element with that element.

Although this perspective on the allocation procedure is quite interesting, its computational cost when dealing with a large n should make the calculation of the principle not affordable. With this respect, this is a similar concern as the one stated for the capital allocation principle based on the Shapley value.

Last comment regards properties that capital allocation principles may satisfy. It is not my intention to go deep in this interesting question. The interested reader is referred, for instance, to van Gulick et al. [2012]. In there, several properties are defined for capital allocation principles such that $K = \rho(S)$ for a particular risk measure ρ which is assumed to be coherent in the sense of Artzner et al. [1999]. That list of properties for capital allocation principles is the following: no diversification; riskless portfolio; symmetry; translation invariance; scale invariance; continuity; and monotonicity with respect to the concordance order.

8

CONTRIBUTIONS TO CAPITAL ALLOCATION BASED ON GLUEVAR

In Section 7.1 of the previous chapter a set of elements to fully describe a capital allocation problem were identified. Nonetheless, two of those elements are of main importance: the assignment criterion and the functions used to simplify the information provided by each random loss. So one could think that guidelines about how capital should be shared among firm's units are basically defined in terms of two components: (1) a capital allocation criterion and (2) a risk measure. The choice of the specific form that each component takes is essential insofar as different capital allocation solutions result from the combinations selected.

The Haircut allocation principle, for instance, combines a standalone proportional capital allocation criterion with the classical Value-at-Risk (VaR) measure; however, this principle was not originally included in the general theoretical framework provided by Dhaene et al. [2012b] in which most of the capital allocation principles that can be found in the academic literature are accommodated. In this chapter it is shown how the Haircut allocation principle also fits in this framework.

Two new proportional capital allocation principles based on GlueVaR risk measures are proposed in next sections. A discussion follows on how allocation principles based on GlueVaR risk measures are applied in practice and some of the implications of tail-subadditivity are described.

8.1 RISK CAPITAL ALLOCATION FOLLOWING THE HAIRCUT PRINCIPLE

In this section we consider the framework suggested by Dhaene et al. [2012b]. This is a unifying framework in which a capital allocation problem is represented by means of three elements: a non-negative function (usually linked to a norm), a set of weights, and a set of auxiliary random variables. However, the Haircut allocation principle could not be fitted into this framework despite its simplicity.

Here, we propose a slight modification of the framework forwarded by Dhaene et al. [2012b] by relaxing some of the conditions so as to

include the Haircut capital allocation principle.

Assume that a capital $K > 0$ has to be allocated across n business units denoted by $i = 1, ..., n$. Following Dhaene et al. [2012b], any capital allocation problem can be described as the optimization problem given by

$$\min_{K_1, K_2, ..., K_n} \sum_{j=1}^{n} \nu_j \cdot \mathbb{E}\left[\zeta_j \cdot D\left(\frac{X_j - K_j}{\nu_j}\right)\right] \quad \text{s.t.} \quad \sum_{j=1}^{n} K_j = K, \quad (8.1)$$

with the following characterizing elements:

(a) a function $D : \mathbb{R} \to \mathbb{R}^+$;

(b) a set of positive weights ν_i, $i = 1, ..., n$, such that $\sum_{i=1}^{n} \nu_i = 1$; and

(c) a set of random variables ζ_i, $i = 1, ..., n$, with $\mathbb{E}[\zeta_i] < +\infty$.

Unlike the original framework provided by Dhaene et al. [2012b]), a distinction is made in (c) so that each ζ_i is now no longer forced to be positive with each $\mathbb{E}[\zeta_i]$ equal to 1. Following this modification, the Haircut capital allocation solution can be obtained from the minimization problem (8.1). If a capital $K > 0$ has to be allocated across n business units, the Haircut allocation principle states that the capital K_i to be assigned to each business unit must be

$$K_i = K \cdot \frac{F_{X_i}^{-1}(\alpha)}{\sum_{j=1}^{n} F_{X_j}^{-1}(\alpha)} \quad \forall i = 1, ..., n, \quad (8.2)$$

where X_i is the random loss linked to the ith business unit, $F_{X_i}^{-1}$ is the inverse of the cumulative distribution function of X_i and $\alpha \in (0, 1)$ is a given confidence level.

Let us consider $d_i = \min\left\{d \geqslant 1 \mid 0 < |M^d[X_i]| < +\infty\right\}$ for all $i = 1, ..., n$, where $M^d[X_i] = \mathbb{E}\left[X_i^d\right]$ is the moment of order $d > 0$ of random variable X_i. Note that $d_i \geqslant 1$ for each i to face a feasible capital allocation problem. In other words, if a business unit presents a random loss with no finite moments, then the risk taken by that business unit is not insurable.

The approach for fitting the Haircut allocation principle in the framework linked to the optimization problem (8.1) can be summarized as follows: if a constant r_i must be expressed as $r_i = \mathbb{E}[\zeta_i \cdot X_i]$, then using $\zeta_i = \left(X_i^{d_i - 1} / M^{d_i}[X_i]\right) \cdot r_i$, a solution is found because $\mathbb{E}[\zeta_i \cdot X_i] = \mathbb{E}\left[\left(X_i^{d_i} / M^{d_i}[X_i]\right)\right] \cdot r_i = r_i$. Although this is an elegant approach, the interpretation of the transformation made by ζ_i on X_i

is intricate. I only recommend to follow this strategy when there is no available alternative involving an interpretable ζ_i.

Proposition 8.1.1. *Let us consider a confidence level $\alpha \in (0,1)$. Then, the three characterizing elements required to represent the Haircut allocation principle in the general framework defined by (8.1) are:*

(a) $D(x) = x^2$,

(b) $\nu_i = \dfrac{\mathbb{E}[\zeta_i \cdot X_i]}{\sum_{j=1}^{n} \mathbb{E}[\zeta_j \cdot X_j]}$, $i = 1, ..., n$; *and*

(c) $\zeta_i = \dfrac{X_i^{d_i - 1}}{M^{d_i}[X_i]} \cdot F_{X_i}^{-1}(\alpha)$, $i = 1, ..., n$.

Proof of Proposition 8.1.1. In this setting it is straightforward to show that the solution $\vec{K} = (K_1, K_2, ..., K_n)$ to the minimization problem (8.1) is the Haircut allocation solution expressed by (8.2). Dhaene et al. [2012b] show that, if function D is the squared Euclidean norm ($D(x) = x^2$), then any solution to (8.1) can be written as

$$K_i = \mathbb{E}[\zeta_i \cdot X_i] + \nu_i \cdot \left(K - \sum_{j=1}^{n} \mathbb{E}[\zeta_j \cdot X_j] \right), \quad \text{for all} \quad i = 1, ..., n. \tag{8.3}$$

In this setting, $\nu_i = \mathbb{E}[\zeta_i \cdot X_i] / \sum_{j=1}^{n} \mathbb{E}[\zeta_j \cdot X_j]$ for each i, so

$$K_i = \mathbb{E}[\zeta_i \cdot X_i] + K \cdot \frac{\mathbb{E}[\zeta_i \cdot X_i]}{\sum_{j=1}^{n} \mathbb{E}[\zeta_j \cdot X_j]} - \mathbb{E}[\zeta_i \cdot X_i] = K \cdot \frac{\mathbb{E}[\zeta_i \cdot X_i]}{\sum_{j=1}^{n} \mathbb{E}[\zeta_j \cdot X_j]}.$$

And, finally, for all i it is true that $\mathbb{E}[\zeta_i \cdot X_i] = F_{X_i}^{-1}(\alpha)$ because of (c). Therefore, each K_i in the solution $\vec{K}$ is given by

$$K_i = K \cdot \frac{F_{X_i}^{-1}(\alpha)}{\sum_{j=1}^{n} F_{X_j}^{-1}(\alpha)}. \square$$

Some comments on ν_i weights and ζ_i auxiliary random variables follow. These ideas concern expression (8.3), namely the general solution of the optimization problem (8.1) when the squared Euclidean norm is used as D function in the reference framework. Capital allocation principles driven by (8.3) can be thought of as two step allocation procedures: in a first step, a particular amount ($C_i = \mathbb{E}[\zeta_i \cdot X_i]$) is allocated to each business unit and, as the sum of all these amounts

does not necessarily equal K (i.e., $\sum_{j=1}^{n} C_j \neq K$), in the second step the difference $\left(K - \sum_{j=1}^{n} C_j\right)$ is allocated to the business units considering weights ν_i. From this perspective, C_i capitals are expected values of X_i losses restricted to particular events of interest and, therefore, ζ_i auxiliary random variables are used to select those events of interest for each business unit. On the other hand, ν_i weights are related to the second step of the procedure, indicating how the difference between K and $\sum_{j=1}^{n} C_j$ must be distributed among business units. For a deeper interpretation of ν_i weights and ζ_i auxiliary random variables in more general cases, the interested reader is referred to Dhaene et al. [2012b].

8.2 PROPORTIONAL RISK CAPITAL ALLOCATION PRINCIPLES USING GLUEVAR

Most of the proportional allocation principles found in the literature can be described in the framework suggested by Dhaene et al. [2012b], where the three characteristic elements are the Euclidean norm, weights $\nu_i = \mathbb{E}[\zeta_i \cdot X_i] / \left(\sum_{j=1}^{n} \mathbb{E}\left[\zeta_j \cdot X_j\right]\right)$, and a set of appropriate ζ_i, for all $i = 1, ..., n$. Following the notation used by these authors, we deal with business unit driven proportional allocation principles when ζ_i depends on X_i. If ζ_i depends on $S = \sum_{i=1}^{n} X_i$ then we have aggregate portfolio driven proportional allocation principles. In the former case, the marginal risk contributions of business units to the overall risk of the portfolio are not taken into account; in the latter, they are. Adopting the notation introduced in the previous chapter, principles belonging to the first category are here denoted as stand-alone proportional allocation principles while principles in the second category are denoted as proportional allocation principles based on partial contributions.

In this chapter two new proportional capital allocation principles are proposed using GlueVaR risk measures. Both principles share the expressions for two of the three characterizing elements: $D(x) = x^2$ and $\nu_i = \mathbb{E}[\zeta_i \cdot X_i] / \left(\sum_{j=1}^{n} \mathbb{E}\left[\zeta_j \cdot X_j\right]\right)$, for all $i = 1, ..., n$. They differ in the set of random variables ζ_i, $i = 1, ..., n$, which are presented below for the case of continuous random variables X_i.

8.2.1 *Stand-alone proportional allocation principles using GlueVaR*

Given two confidence levels α and β in $(0,1)$, $\alpha < \beta$, and two distorted survival probabilities h_1 and h_2, if ζ_i is fixed as

$$\begin{aligned}\zeta_i &= \omega_1 \cdot \frac{\mathbb{1}\left[X_i \geqslant F_{X_i}^{-1}(\beta)\right]}{1-\beta} + \omega_2 \cdot \frac{\mathbb{1}\left[X_i \geqslant F_{X_i}^{-1}(\alpha)\right]}{1-\alpha} \\ &+ \omega_3 \cdot \frac{X_i^{d_i-1}}{M^{d_i}[X_i]} \cdot F_{X_i}^{-1}(\alpha), \quad \text{for all } i = 1,...,n,\end{aligned} \tag{8.4}$$

then the stand-alone proportional allocation principle using as risk measure the $\text{GlueVaR}_{\beta,\alpha}^{h_1,h_2}$ can be represented in the modified capital allocation framework. Components of the solution $(K_1, K_2, ..., K_n)$ are expressed as

$$K_i = K \cdot \frac{\text{GlueVaR}_{\beta,\alpha}^{h_1,h_2}(X_i)}{\sum_{j=1}^{n} \text{GlueVaR}_{\beta,\alpha}^{h_1,h_2}(X_j)}, \quad \text{for all } i = 1,...,n. \tag{8.5}$$

8.2.2 *Proportional allocation principles based on partial contributions using GlueVaR*

Similarly, if there exists a confidence level $\alpha^* \in (0,1)$ such that $F_S^{-1}(\alpha) = \sum_{j=1}^{n} F_{X_j}^{-1}(\alpha^*)$, the proportional allocation principle based on partial contributions using $\text{GlueVaR}_{\beta,\alpha}^{h_1,h_2}$ can be fitted to the modified capital allocation framework. In this case, ζ_i has to be equal to

$$\begin{aligned}\zeta_i &= \omega_1 \cdot \frac{\mathbb{1}\left[S \geqslant F_S^{-1}(\beta)\right]}{1-\beta} + \omega_2 \cdot \frac{\mathbb{1}\left[S \geqslant F_S^{-1}(\alpha)\right]}{1-\alpha} \\ &+ \omega_3 \cdot \frac{X_i^{d_i-1}}{M^{d_i}[X_i]} \cdot F_{X_i}^{-1}(\alpha^*), \quad \text{for all } i = 1,...,n.\end{aligned} \tag{8.6}$$

Each component of the solution $(K_1, K_2, ..., K_n)$ is then obtained as

$$\begin{aligned}K_i = K \cdot \Bigg[&\omega_1 \cdot \frac{\mathbb{E}\left[X_i \mid S \geqslant F_S^{-1}(\beta)\right]}{\text{GlueVaR}_{\beta,\alpha}^{h_1,h_2}(S)} + \omega_2 \cdot \frac{\mathbb{E}\left[X_i \mid S \geqslant F_S^{-1}(\alpha)\right]}{\text{GlueVaR}_{\beta,\alpha}^{h_1,h_2}(S)} \\ &+ \omega_3 \cdot \frac{F_{X_i}^{-1}(\alpha^*)}{\text{GlueVaR}_{\beta,\alpha}^{h_1,h_2}(S)}\Bigg].\end{aligned} \tag{8.7}$$

Alternatively, another approach can be considered. There exists a set of confidence levels $\alpha_j \in (0,1)$ for all $j = 1,...,n$ such that $F_S^{-1}(\alpha) = \sum_{j=1}^{n} F_{X_j}^{-1}(\alpha_j)$. Therefore, the proportional allocation principle based on partial contributions using $\text{GlueVaR}_{\beta,\alpha}^{h_1,h_2}$ can also be

fitted to the modified capital allocation framework. In this case, ζ_i has to be equal to

$$\begin{aligned}\zeta_i &= \omega_1 \cdot \frac{\mathbb{1}\left[S \geqslant F_S^{-1}(\beta)\right]}{1-\beta} + \omega_2 \cdot \frac{\mathbb{1}\left[S \geqslant F_S^{-1}(\alpha)\right]}{1-\alpha} \\ &\quad + \omega_3 \cdot \frac{X_i^{d_i-1}}{M^{d_i}[X_i]} \cdot F_{X_i}^{-1}(\alpha_i), \quad \text{for all } i = 1, ..., n.\end{aligned} \tag{8.8}$$

Each component of the solution $(K_1, K_2, ..., K_n)$ is then obtained as

$$\begin{aligned}K_i = K \cdot \Bigg[&\omega_1 \cdot \frac{\mathbb{E}\left[X_i \mid S \geqslant F_S^{-1}(\beta)\right]}{\text{GlueVaR}_{\beta,\alpha}^{h_1,h_2}(S)} + \omega_2 \cdot \frac{\mathbb{E}\left[X_i \mid S \geqslant F_S^{-1}(\alpha)\right]}{\text{GlueVaR}_{\beta,\alpha}^{h_1,h_2}(S)} \\ &+ \omega_3 \cdot \frac{F_{X_i}^{-1}(\alpha_i)}{\text{GlueVaR}_{\beta,\alpha}^{h_1,h_2}(S)} \Bigg].\end{aligned} \tag{8.9}$$

8.3 AN EXAMPLE OF INSURANCE RISK CAPITAL ALLOCATION USING GLUEVAR ON CLAIM COSTS

The same dataset linked to the illustrative examples in Chapters 3 and 7 is used here. Three GlueVaR measures are shown in Table 8.1, corresponding to different risk attitudes. $\text{GlueVaR}_{99.5\%,95\%}^{11/30,2/3}$ reflects a balanced attitude, weighting $\text{TVaR}_{99.5\%}$, $\text{TVaR}_{95\%}$ and $\text{VaR}_{95\%}$ equally. $\text{GlueVaR}_{99.5\%,95\%}^{0,1}$ corresponds to a scenario in which a zero weight is allocated to $\text{VaR}_{95\%}$, the $\text{TVaR}_{95\%}$ is overweighted and the lowest feasible weight is allocated to $\text{TVaR}_{99.5\%}$. Finally, the third risk measure considered, $\text{GlueVaR}_{99.5\%,95\%}^{1/20,1/8}$, reflects a more conservative attitude than that represented by using $\text{VaR}_{95\%}$ on its own. Table 8.1 is divided into two blocks. In the first, risk was calculated for the whole data set and in the second, contributions to the risk shown in the first block coming only from the 5%-common tail were computed. Recall the definition of the α-common tail provided in Chapter 2: thus, in this second block, only the observations that lie simultaneously to the right of the 95% quantile of X_1, X_2 and $X_1 + X_2$ were considered. The last column presents the concentration index, which is the ratio of the risk of $X_1 + X_2$ divided by the sum of the risk of X_1 plus the risk of X_2. A concentration index smaller than one indicates subadditivity and, hence, a diversification effect.

In this example, $\text{VaR}_{95\%}$ and one of the GlueVaR measures are not subadditive in the whole domain, because their associated distortion functions are not concave in the whole $[0, 1]$ interval. However, $\text{GlueVaR}_{99.5\%,95\%}^{11/30,2/3}$, $\text{GlueVaR}_{99.5\%,95\%}^{0,1}$ and $\text{GlueVaR}_{99.5\%,95\%}^{1/20,1/8}$ satisfy tail-subadditivity at confidence level $\alpha = 95\%$. Note that the concentration indexes smaller than one reveal that all the measures

Table 8.1: Risk assessment of claim costs using GlueVaR risk measures

	$\mathbf{X_1}$	$\mathbf{X_2}$	$\mathbf{X_1 + X_2}$	**Difference**$^{(*)}$	**Concentration index**
	(a)	(b)	(c)	(a)+(b)-(c)	(c)/((a)+(b))
Whole domain					
$\text{VaR}_{95\%}$	38.8	6.4	47.6	-2.4	1.05
$\text{TVaR}_{95\%}$	112.5	18.4	125.5	5.4	0.96
$\text{TVaR}_{99.5\%}$	440.0	54.2	479.0	15.2	0.97
$\text{GlueVaR}^{11/30,2/3}_{99.5\%,95\%}$	197.1	26.3	217.4	6.0	0.97
$\text{GlueVaR}^{0,1}_{99.5\%,95\%}$	76.1	14.4	86.2	4.3	0.95
$\text{GlueVaR}^{1/20,2/8}_{99.5\%,95\%}$	61.7	9.4	72.1	-1.0	1.01
Common 5%-right tail$^{(**)}$					
$\text{VaR}_{95\%}$	0.0	0.0	0.0	0.0	–
$\text{TVaR}_{95\%}$	75.3	12.5	76.8	11.0	0.88
$\text{TVaR}_{99.5\%}$	411.3	46.7	426.7	31.3	0.93
$\text{GlueVaR}^{11/30,2/3}_{99.5\%,95\%}$	162.2	19.7	167.8	14.1	0.92
$\text{GlueVaR}^{0,1}_{99.5\%,95\%}$	37.9	8.7	37.9	8.7	0.81
$\text{GlueVaR}^{1/20,2/8}_{99.5\%,95\%}$	23.4	3.0	24.2	2.2	0.92

$^{(*)}$ Benefit of diversification.

$^{(**)}$ Part of the risk measure arising from the intersection of 5%-right tails of X_1, X_2 and $X_1 + X_2$.

are subadditive in the tail.

Next, a capital allocation application is illustrated where total capital has to be allocated between the two units of risk, X_1 and X_2. Table 8.2 shows particular allocation solutions for two proportional risk capital allocation principles using GlueVaR.

A similar behavior is observed for the three GlueVaR risk measures. The capital is allocated primarily to risk X_1 regardless of the allocation criterion. Note that the percentages of capital allocated to X_1 are higher when the partial contribution criterion is used and a confidence level $\alpha^* = 95.37\%$ is set such that $F_S^{-1}(95\%) = F_{X_1}^{-1}(95.37\%) + F_{X_2}^{-1}(95.37\%)$. This is an expected result, because the right tail of X_1 is fatter than that of X_2.

Let me focus on capital allocation solutions involving the partial contribution criterion in which confidence levels α_j, $j = 1, 2$ are not forced to be equal across the risk units. A notable fall in the risk allocated to X_1 is observed if a partial contribution criterion with no constant level α^* and $\text{GlueVaR}_{99.5\%,95\%}^{1/20,2/8}$ is chosen.

This result is obtained because the impact on the quantile of X_1 is the opposite of that on X_2 when α_j, $j = 1, 2$, are estimated as $F_S^{-1}(95\%) = F_{X_1}^{-1}(\alpha_1) + F_{X_2}^{-1}(\alpha_2)$. These confidence levels are equal to $\alpha_1 = 94.78\%$ and $\alpha_2 = 97.49\%$. This particular risk measure is not subadditive in the whole domain and is tail-subadditive for these data. In fact, the associated quantiles for individual variables are $\text{VaR}_{94.78\%}(X_1)$ and $\text{VaR}_{97.49\%}(X_2)$, so the risk contribution of X_1 is underweighted compared to the risk of X_2.

8.4 FURTHER COMMENTS ON THESE CONTRIBUTIONS

Three comments are included to close this chapter. They go from particular notes to general observations. First of all, let me present a comment related to the gradient allocation principle. Using the notation introduced in this chapter and in the previous ones, the gradient allocation principle can be fitted into the framework of reference following a similar strategy than the one in Proposition 8.1.1, but changing $F_{X_i}^{-1}(\alpha)$ by $\frac{\partial \rho}{\partial u_i}(S)$ for all $i = 1, \ldots, n$. Or, in other words, letting be $D(x) = x^2$, $\nu_i = \mathbb{E}[\zeta_i \cdot X_i] / \left(\sum_{j=1}^n \mathbb{E}[\zeta_j \cdot X_j]\right)$ and

$$\zeta_i = \left(\frac{X_i^{d_i - 1}}{M^{d_i}[X_i]}\right) \cdot \frac{\partial \rho}{\partial u_i}(S)$$

Table 8.2: Proportional capital allocation solutions using GlueVaR for the claim costs data

	Proportion allocated to X_1	**Proportion allocated to X_2**
Stand-alone criterion		
$\text{GlueVaR}^{11/30,2/3}_{99.5\%,95\%}$	88.21%	11.79%
$\text{GlueVaR}^{0,1}_{99.5\%,95\%}$	84.07%	15.93%
$\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$	86.79%	13.21%
Partial contribution criterion with constant[(a)] α^*		
$\text{GlueVaR}^{11/30,2/3}_{99.5\%,95\%}$ [(a)]	90.75%	9.25%
$\text{GlueVaR}^{0,1}_{99.5\%,95\%}$ [(a)]	87.83%	12.17%
$\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$ [(a)]	88.06%	11.94%
Partial contribution criterion with non constant[(b)] α_j		
$\text{GlueVaR}^{11/30,2/3}_{99.5\%,95\%}$ [(b)]	89.93%	10.07%
$\text{GlueVaR}^{0,1}_{99.5\%,95\%}$ [(b)]	87.83%	12.17%
$\text{GlueVaR}^{1/20,1/8}_{99.5\%,95\%}$ [(b)]	81.55%	18.45%

[(a)] A confidence level α^* such that $F_S^{-1}(95\%) = F_{X_1}^{-1}(\alpha^*) + F_{X_2}^{-1}(\alpha^*)$. In this case $\alpha^* = 95.37\%$.

[(b)] Confidence levels $\alpha_j \in (0,1)$ are selected to satisfy $F_S^{-1}(95\%) = F_{X_1}^{-1}(\alpha_1) + F_{X_2}^{-1}(\alpha_2)$. In this case $\alpha_1 = 94.78\%$ and $\alpha_2 = 97.49\%$.

for all $i = 1, \ldots, n$. Therefore, it may be assumed that the gradient allocation principle is a proportional principle based on partial contributions, although the proposed fitting in the framework of reference scarcely informs about the driving idea of the gradient criterion.

The second comment regards non-proportional capital allocation principles using GlueVaR. It has to be mentioned that it is possible to effortless obtain non-proportional principles using any of the auxiliary random variables ζ_i described in expressions (8.4),(8.6) or (8.8). If function $D(x) = x^2$, therefore the only thing that must be taken into account is that at least one of the weights $\nu_i, i = 1, \ldots, n$ must be different of $\mathbb{E}[\zeta_i \cdot X_i] / \left(\sum_{j=1}^{n} \mathbb{E}\left[\zeta_j \cdot X_j\right]\right)$. Under these restrictions, whatever set of auxiliary random variables $\vec{\zeta}$ is chosen among expressions (8.4),(8.6) or (8.8), non-proportional capital allocation principles $\vec{K}$ using GlueVaR are obtained through expression (8.3).

Third and last comment is about links between aggregation functions and the framework proposed by Dhaene et al. [2012b]. As it has been noted several times before in this dissertation, capital allocation problems are disaggregation problems and therefore, until some extent, the goal of capital allocation principles is the opposite of the goal of aggregation functions, which is a summarizing purpose. Nonetheless, the optimization perspective taken into account in expression (8.1) involves aggregation operators in the objective function. For instance, one can thought of the function

$$\mathbb{E}\left[\sum_{j=1}^{n} \nu_j \cdot \zeta_j \cdot D\left(\frac{X_j - K_j}{\nu_j}\right)\right]$$

to be minimized in (8.1) as the composition of two main aggregation operators: one aggregation operator is given by expression

$$\sum_{j=1}^{n} \nu_j \cdot \zeta_j \cdot D\left(\frac{X_j - K_j}{\nu_j}\right)$$

and the other one is the mathematical expectation $\mathbb{E}$. It has to be noted that a similar perspective is proposed in Xu and Hu [2012], where the first aggregation function may be represented as

$$\Phi(L(\vec{K})) = \Phi(\sum_{j=1}^{n} \phi(X_j - K_j)) \quad ,$$

being ϕ a function usually linked to a distance and Φ an increasing function (which could be the identity function, for instance). Thinking about capital allocation principles as solutions to optimization problems has become usual in the recent literature [see, for instance You and Li, 2014; Zaks and Tsanakas, 2014]. Aggregation functions

may be defined as solutions to optimization problems, as proposed in De Baets [2013]. Therefore, some relationships between capital allocation principles and particular aggregation functions may be explored on further research.

9

CAPITAL ALLOCATION PRINCIPLES AS COMPOSITIONAL DATA

In Chapter 7 it was shown that given an (absolute) capital allocation principle $\vec{K}$ with $K = \sum_{j=1}^{n} K_j$, its relative counterpart is defined as $\vec{x}$, where components are $x_i = K_i/K$. This chapter is devoted to show the main relationship between capital allocation problems and aggregation operators introduced in this dissertation, which is based on thinking about the relative capital allocation principles as belonging to the (standard) simplex $\mathcal{S}^n = \{\vec{z} \in \mathbb{R}^n \mid z_j \geqslant 0, j = 1, \ldots, n, \sum_{j=1}^{n} z_j = 1\}$ provided with a particular structure of vector and metric space. Following a nomenclature often used by geologists, any vector $\vec{z} \in \mathcal{S}^n$ is called a *composition* and any set of vectors in the simplex is called *compositional data*. First thing to do is to present this vector space and the particular distance chosen to give $\mathcal{S}^n$ a metric space structure. Secondly, some comments on how to move forward and backward from relative capital allocation principles to compositions are provided. Afterwards, some applications to exploit the established relationship are discussed and illustrated with an example extracted from the capital allocation literature. To close this chapter, ideas on further lines of research linked to understand a relative capital allocation principle as a composition are presented.

9.1 THE SIMPLEX AND ITS VECTORIAL AND METRIC STRUCTURE

Recall that a set of vectors needs two operations (often called vector addition and scalar multiplication) to be considered a vector space over $\mathbb{R}$. Moreover, these operations must satisfy particular properties. The vector addition must be commutative, associative, and must have a neutral element and, for each vector, its additive inverse. The scalar multiplication for a vector space over $\mathbb{R}$ combines a real number with a vector and this combination must belong again to the set of vectors. Additionally, a neutral element for the scalar multiplication must exist, and the distributivity of the scalar multiplication with respect to the vector addition and, on the other side, the distributivity of the vector addition with respect to the scalar multiplication must be both satisfied.

Following the notation used in Aitchinson and Egozcue [2005], in the case of the set $\mathcal{S}^n$ the vector addition is called *perturbation* (denoted by $\oplus$) and the scalar multiplication is called *powering* (denoted

by $\odot$). These operations are defined by expressions (9.1) and (9.2), respectively, where $\vec{x}, \vec{y} \in \mathcal{S}^n$ and $\lambda \in \mathbb{R}$:

$$\vec{x} \oplus \vec{y} = \left(\frac{x_1 \cdot y_1}{\sum_{j=1}^{n} x_j \cdot y_j}, \ldots, \frac{x_n \cdot y_n}{\sum_{j=1}^{n} x_j \cdot y_j} \right) \quad , \tag{9.1}$$

$$\lambda \odot \vec{x} = \left(\frac{x_1^\lambda}{\sum_{j=1}^{n} x_j^\lambda}, \ldots, \frac{x_n^\lambda}{\sum_{j=1}^{n} x_j^\lambda} \right) \quad . \tag{9.2}$$

It is not my purpose to prove that the simplex $\mathcal{S}^n$ provided with operations $\oplus$ and $\odot$ has a linear vector space structure of dimension $n-1$ but to take advantage of this fact. Before going on, let me introduce another commonly used concept regarding compositional data, the *closure function* $\mathcal{C}$. Although notation $\mathcal{C}$ has been used in previous chapters to refer to the Choquet integral, it is my believe that there is no room for confusion in this context. The closure function applied to a vector in $\mathbb{R}^n$ returns another vector whose components are the components of the original vector divided by the sum of all the components of the original vector. Keeping this in mind, the following expressions hold:

$$\vec{x} \oplus \vec{y} = \mathcal{C}\left[(x_1 \cdot y_1, \ldots, x_n \cdot y_n)\right] \quad , \quad \lambda \odot \vec{x} = \mathcal{C}\left[(x_1^\lambda, \ldots, x_n^\lambda)\right] .$$

Moreover, assuming the vector space structure of $(\mathcal{S}^n, \oplus, \odot)$, the neutral element $\vec{0}$ of $\oplus$ can be deduced. Given a vector $\vec{x}$ such that $x_i > 0$ for all i, the relationship $\vec{x} \oplus \vec{r} = \vec{0}$ informs that $\vec{r}$ is the inverse of $\vec{x}$ with respect to the perturbation operation, so it should be written as $\vec{r} = (-1) \odot \vec{x}$. In other words,

$$\vec{r} = \left(\frac{1/x_1}{\sum_{j=1}^{n}(1/x_j)}, \ldots, \frac{1/x_n}{\sum_{j=1}^{n}(1/x_j)} \right) \quad .$$

Then, using this last expression and (9.1),

$$\vec{0} = \vec{x} \oplus \vec{r} = \mathcal{C}\left[\left(\frac{1}{\sum_{j=1}^{n}(1/x_j)}, \ldots, \frac{1}{\sum_{j=1}^{n}(1/x_j)} \right) \right] = \left(\frac{1}{n}, \ldots, \frac{1}{n} \right),$$

so the neutral element $\vec{0}$ of the perturbation operation is the composition with all of its n elements equal to $1/n$. I would like to remark that, as far as I know, there is still some discussion on how compositions with null components should be considered under this framework. This is something to remind in real world applications.

Finally, in order to consider the vector space $(\mathcal{S}^n, \oplus, \odot)$ as a metric space, a distance is needed. The *simplicial metric* defined in Aitchinson [1983] is considered. Given two compositions $\vec{x}, \vec{y}$, the distance between them from the point of view of the simplicial metric is

$$\Delta(\vec{x}, \vec{y}) = \left[\sum_{i=1}^{n} \left[\ln\left(\frac{x_i}{GM(\vec{x})} \right) - \ln\left(\frac{y_i}{GM(\vec{y})} \right) \right]^2 \right]^{1/2} \quad , \tag{9.3}$$

where $GM(\vec{z})$ denotes the geometric mean of the components of $\vec{z}$ vector as in Table 4.1 from Chapter 4.
An equivalent expression for $\Delta(\vec{x},\vec{y})$ is the following:

$$\Delta(\vec{x},\vec{y}) = \left[\frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\ln\left(\frac{x_i}{x_j}\right)-\ln\left(\frac{y_i}{y_j}\right)\right]^2\right]^{1/2} \quad . \tag{9.4}$$

This simplicial metric is linked to a norm $\|\cdot\|_\Delta$ and to an inner product $<,>_\Delta$ in a usual way: given two vectors $\vec{x},\vec{y}\in\mathcal{S}^n$,

$$\Delta(\vec{x},\vec{y}) = \| \vec{x}\ominus\vec{y} \|_\Delta = \sqrt{< \vec{x}\ominus\vec{y}, \vec{x}\ominus\vec{y} >_\Delta},$$

where $\vec{x}\ominus\vec{y} = \vec{x}\oplus[(-1)\odot\vec{y}]$, and

$$< \vec{u},\vec{v} >_\Delta = \frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\ln\left(\frac{u_i}{u_j}\right)\cdot\ln\left(\frac{v_i}{v_j}\right)\right] \quad . \tag{9.5}$$

Under all this framework, as it was shown in De Baets [2013], the *simplicial arithmetic mean* of the compositional data $\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_m$ may be understood as a solution of a minimization problem, in the following way:

$$AM_\Delta(\vec{x}_1,\ldots,\vec{x}_m) = \frac{1}{m}\odot\bigoplus_{k=1}^{m}\vec{x}_k = \arg\min_{\vec{z}}\sum_{k=1}^{m}\| \vec{z}\ominus\vec{x}_k \|_\Delta^2 \quad , \tag{9.6}$$

where $\odot\bigoplus_{k=1}^{m}\vec{x}_k$ means the perturbation of the set of m compositions $\{\vec{x}_k\}_{k=1,\ldots,m}$.
At first sight, this is an equivalent expression of that of the arithmetic mean of m real numbers $u_1,u_2,\ldots,u_m$:

$$AM(u_1,\ldots,u_m) = \frac{1}{m}\cdot\sum_{k=1}^{m}u_k = \arg\min_{v}\sum_{k=1}^{m}\| v-u_k \|_2^2 \quad , \tag{9.7}$$

so the simplicial metric presented in this section is the natural one if computation of (simplicial) arithmetic means are going to be conducted or, in other words, expression (9.6) contains the proper definition of the arithmetic mean of $\vec{x}_1\ldots,\vec{x}_m$ in the metric space that has been denoted by $(\mathcal{S}^n,\oplus,\odot,\Delta)$. From the definitions of both perturbation and powering operations, an explicit expression for the simplicial arithmetic mean presented in (9.6) is

$$\begin{aligned} &AM_\Delta(\vec{x}_1,\ldots,\vec{x}_m) = \mathcal{C}[(G_1,\ldots,G_n)],\\ &\text{where } G_k = GM(x_{1,k},x_{2,k},\ldots,x_{m,k}),\ \text{i.e.}\\ &G_k = [\Pi_{i=1}^{m}x_{i,k}]^{1/m}\ \forall k=1,\ldots,n \quad . \end{aligned} \tag{9.8}$$

9.1.1 *From capital allocation principles to compositional data and backwards*

Recall again that an absolute capital allocation $\vec{K}$ has its relative counterpart $\vec{x}$, satisfying that $x_i = K_i/K$ for all $i = 1, \ldots, n$. Note that it is satisfied that $\sum_{j=1}^{n} x_j = 1$. It could happen that some component would be negative and then $\vec{x} \notin \mathcal{S}^n$. For the rest of the chapter it is assumed that $\vec{x}$ has strictly positive components. This assumption allows to avoid negative or zero values on components of $\vec{x}$, which are an inconvenient for practitioners (the negative ones) and for operating in the simplex (the zero ones). Briefly, I assume that $\vec{x}$ is a composition with non-zero components.

At this point, some concepts introduced in Chapter 7 to classify absolute capital allocation problems may be related to some of the concepts introduced in this chapter. For instance, think about proportional capital allocation principles as stated in expression (7.1). The relative counterpart $\vec{y}$ of the absolute principle $\vec{K} = (K_1, \ldots, K_n)$ may be interpreted as the closure of the vector with components equal to $f_i(X_i), i = 1, \ldots, n$:

$$\begin{aligned} &\vec{K} \text{ s.t. } K_i = K \cdot \frac{f_i(X_i)}{\sum_{j=1}^{n} f_j(X_j)}, \forall i = 1, \ldots, n \\ &\Leftrightarrow \vec{y} = \mathcal{C}\left[(f_1(X_1), \ldots, f_n(X_n))\right] \quad . \end{aligned} \tag{9.9}$$

Moreover, when dealing with stand-alone proportional principles, last expression helps to clarify why none dependence structure between random variables $\{X_i\}_{i=1,\ldots,n}$ is taken into account for the allocation: In a first step, it is assigned to agents the amount of risk faced by each one of them, which is summarized in $f_i(X_i)$, $i = 1, \ldots, n$. Subsequently, this amount of risk is scaled by K.

9.2 PERTURBATION INVERSE, SIMPLICIAL DISTANCE AND SIMPLICIAL ARITHMETIC MEAN APPLIED TO CAPITAL ALLOCATION PROBLEMS

In the previous section it has been shown that relative capital allocation principles and compositions may be naturally linked. Once this relationship is established, the idea is to take advantage of the geometric structure of the simplex to enrich the description of each capital allocation principle.

9.2.1 *Cost of risk, reward to riskless and reward on return&risk principles*

Given a relative capital allocation principle $\vec{x}$ linked to a cost of risk goal, one could think to depart from it to allocate rewards instead of costs, so to fulfill an allocation with a reward to riskless objetive. An intuitive idea is to invert each of the relative components, in order to reflect the inverse nature of the allocation (a relative low cost allocated to ith agent should mean a relative high reward assigned to it). But one must normalize the sum of all $1/x_i$ in order to provide a full allocation of reward amount K. This normalization is nothing else than using the closure function. Note that all components of $\vec{x}$ must be different than 0 in order to be invertible.

The above idea has a natural interpretation in the simplex $\mathcal{S}^n$. Given a relative capital allocation principle $\vec{x}$, let $\vec{r}$ be the closure of the vector with components $1/x_i$ for $i = 1, \ldots, n$. As it has been shown in Section 9.1, $\vec{r}$ is the inverse of $\vec{x}$ with respect to the perturbation operation: $\vec{r} = (-1) \odot \vec{x}$.

Using risk based capital allocation principles to determine penalizations or rewards may lead to undesirable behaviors of the agents. Basically, agents would be compelled to conservative business decisions because taking less risk results in a better reward. In order to prevent it, some return-on-risk measure seems to be preferable to assign rewards. It has been emphasized several times that reward on return&risk allocations may be of great relevance for a sound ERM system. But at this point this kind of allocation has only been briefly discussed in Section 7.2.1 of Chapter 7, when presenting the RORAC compatibility of the gradient allocation principles. Note now that, bearing in mind all of the building blocks put in place, there are some direct absolute reward on return&risk capital allocation principles that may be considered:

$$\begin{gathered}\text{Depart from a given } \vec{x} = \mathcal{C}[\vec{y}] \text{ where} \\ y_i = \frac{\text{RORAC}(X_i \mid S)}{\text{RORAC}(S)} \quad \forall i = 1, \ldots, n \\ \text{and then obtain the absolute principle } \vec{K} \text{ by} \\ K_i = K \cdot x_i \quad \forall i = 1, \ldots, n.\end{gathered} \tag{9.10}$$

The idea is to better reward those agents whose relative RORAC with respect to the overall RORAC of the portfolio is higher. Note that different definitions of return-on-risk measures than expressions (7.2.3) and (7.2.2) in Chapter 7 for $\text{RORAC}(X_i \mid S)$ and $\text{RORAC}(S)$ may be considered in (9.10), and the objective of the allocation would not change.

9.2.2 *Ranking principles using the simplicial distance*

It has to be reminded that $\vec{0} \in \mathcal{S}^n$ is the composition with all of its n components equal to $1/n$. As there is a simplicial metric or distance Δ which helps to constitute $\mathcal{S}^n$ as a metric space, the distance between any relative capital allocation principle $\vec{x}$ and $\vec{0}$ can be computed, or even the distance between any pair of relative capital allocation principles belonging to $\mathcal{S}^n$. It is argued that both uses of the simplicial distance can be useful to compare different capital allocation principles in a quantitative way.

For instance, when comparing to $\vec{0}$, a quantitative indicator of how far is the allocation principle from a neutral assignment is obtained, because $\vec{0} \in \mathcal{S}^n$ is linked to a capital allocation principle in which no matter how much risk is each agent assuming: each one of them has assigned the amount K/n. On the other hand, if an allocation principle is taken as a reference (for instance, a gradient allocation principle as explained in Section 7.2.1 of Chapter 7), the distance between the composition linked to this principle and any other composition quantifies how far is the principle linked to it from the allocation of reference. In other words, imagine that four allocation principles are in hand for the same amount K and the same n agents: a haircut allocation principle (8.2), $\vec{K}_h$; a covariance allocation principle (7.8), $\vec{K}_c$; a stand-alone proportional allocation principle based on GlueVaR (8.5), $\vec{K}_s$; and a gradient allocation principle related to (7.5), $\vec{K}_g$. If their respective relative allocation principles $\vec{x}_h, \vec{x}_c, \vec{x}_s$ and $\vec{x}_g$ are in $\mathcal{S}^n$ and each of the components of $\vec{x}_t, t \in \{h, c, s, g\}$ is strictly positive, then it is possible to rank them in, for instance, two different ways:

1) Compute $\Delta(\vec{x}_t, \vec{0})$ for $t \in \{h, c, s, g\}$ and increasingly order these values. In this case, the higher the order position the further the allocation is from a neutral allocation;

2) Choose one of the principles as reference (for instance, the gradient allocation principle). Compute $\Delta(\vec{x}_t, \vec{x}_g)$ for $t \in \{h, c, s\}$. These three values are quantifying, until some extent, how far is each principle from the allocation of reference.

As long as the simplicial distance Δ is an aggregation function, these are direct applications of an aggregation function to capital allocation problems. The idea of using Δ to quantitatively rank capital allocation principles is one of the contributions of this dissertation with regards to the relationship between capital allocation problems and aggregation functions.

9.2.3 *Averaging principles using the simplicial arithmetic mean*

Sometimes several management teams provide different allocations regarding the same assignment problem. The situation commented at the end of last section could be one of such examples. In general, let us imagine m management teams providing m absolute capital allocation principles $\vec{K}_k$ of amount K to the same n agents, and let $\vec{x}_k$ be the relative capital allocation principles linked to $\vec{K}_k$, $k = 1, \ldots, m$. Once again, taking advantage of the geometric structure of $\mathcal{S}^n$, the concept of *averaging* the m points of view on the same allocation problem is easily derived. In other words, the expression

$$\vec{z} = \mathrm{AM}_\Delta\left(\vec{x}_1, \ldots, \vec{x}_m\right)$$

is the proper definition of the arithmetic mean of $\vec{x}_1 \ldots, \vec{x}_m$ in the metric space $(\mathcal{S}^n, \oplus, \odot, \Delta)$ as it was shown in (9.6). Once the relative arithmetic mean is obtained, what remains to do is assigning an amount of $\bar{K}_i = K \cdot z_i$ monetary units to each ith agent, $i = 1, \ldots, n$, in order to provide an allocation principle which balance the opinions of all the involved management teams.

The arithmetic mean is one of the very first examples given when talking about aggregation operators. In my opinion, it is interesting to have found a way of aggregating different capital allocation principles through an arithmetic mean, because it may be the base to extend such a summarizing idea to other aggregation operators as the WAM or the OWA shown in Chapter 4, for instance.

9.2.4 *An illustration*

In order to illustrate the applications described in this section with an example, some values from the numerical illustration in van Gulick et al. [2012] are taken into account. The authors consider an insurance company that holds portfolios of three types of life insurance:

- a (deferred) single life annuity that yields a yearly payment in every year that the insured is alive and older than 65;
- a survivor annuity that yields a yearly payment in every year that the spouse outlives the insured, if the insured dies before age 65;
- a death benefit insurance that yields a single payment in the year the insured dies, if the insured dies before age 65.

Without going deep into the details, the number and gender of the insureds in each of the portfolios is 45,000 males, 15,000 males and 15,000 males, respectively. Four capital allocation principles are shown in van Gulick et al. [2012], while only three of them are going to be

used here because one of them leads to negative allocated values for some of the agents. Using the notation introduced in Chapter 7, the authors consider $n = 3$ agents, being each one of them the three aforementioned portfolios of life insurance. The random variables linked to each agent are the present values of the liability payments to all insureds in each portfolio, and are represented by X_{sl}, X_{surv} and X_{db}, respectively. The capital amount to be distributed among the agents is $K = \text{TVaR}_{99\%}(S)$, being $S = X_{sl} + X_{surv} + X_{db}$, and two proportional allocation criteria and one non-proportional criterion are considered in this illustration. This leads to three principles: a proportional allocation principle (7.1) based on the standard deviation as risk measure; the gradient allocation principle, which may be also considered proportional as it has been shown in Section 7.2.1 of Chapter 7; and the non-proportional excess based allocation principle proposed by the authors. The figures published in van Gulick et al. [2012] are used to going forward. In that sense, $K = \text{TVaR}_{99\%}(X_{sl} + X_{surv} + X_{db}) = 376,356$ and the allocated capitals by each principle are shown in Table 9.1. Relative capital allocation principles linked to the absolute ones are also presented in Table 9.1. All the displayed relative principles belong to the simplex $\mathcal{S}^3$ and have non-zero components.

Table 9.1: Absolute capital allocation principles taken from van Gulick et al. [2012] and their relative counterparts.

	Single life annuity (X_{sl})	Survivor annuity (X_{surv})	Death benefit (X_{db})
	Relative principles		
Proportional principle based on the st.dev.	89.20%	6.57%	4.23%
Gradient allocation principle	96.84%	2.12%	1.04%
Excess based allocation principle	95.74%	2.79%	1.47%
	Absolute principles		
Proportional principle based on the st.dev.	335,724	24,725	15,907
Gradient allocation principle	364,477	7,979	3,900
Excess based allocation principle	360,324	10,495	5,537

To begin with, the inverse of these capital allocation principles are deduced following Section 9.2.1. Recall that the idea was to build allocation principles with a reward to riskless goal from allocation principles whose original objectives were cost of risk ones. As it has been argued before, probably the most suitable situation in which risk managers may be interested in allocation principles with reward to riskless objectives is when assigning risk diversification benefits to the agents, mainly where the only available information is a capital allocation principle with a cost of risk goal and an overall amount of diversification benefit. As in van Gulick et al. [2012] the following values $\text{TVaR}_{99\%}(X_{sl}) = 364,477$, $\text{TVaR}_{99\%}(X_{surv}) = 11,657$ and $\text{TVaR}_{99\%}(X_{db}) = 6,346$ are provided, the diversification benefit can be computed as

$$(364,477 + 11,657 + 6,346) - 376,356 = 382,480 - 376,356 = 6,124.$$

Therefore, in order to put in value the perturbation inverse allocation, absolute principles shown in Table 9.2 are such that the shared reward is the diversification benefit, i.e. $K = 6,124$. Note that, as long as in this case the individual diversification benefits could be calculated for all principles, the results shown in Table 9.2 do not match with the ones that could be deduced in a direct way. For instance, individual diversification benefits linked to the gradient allocation principle are, from Table 9.1 and values of $\text{TVaR}_{99\%}(X_{sl})$, $\text{TVaR}_{99\%}(X_{surv})$ and $\text{TVaR}_{99\%}(X_{db})$, the following ones: for the single life annuity 364,477-364,477=0; for the survivor annuity,11,657-7,979=3,678; and for the death benefit, 6,346-3,900=2,446. These values may be interpreted directly as an absolute allocation principle of the overall diversification benefit, which leads to the relative principle $(0.00\%, 60.06\%, 39.94\%)$. There are significant differences between this relative principle and the one associated to the gradient criterion in Table 9.2, $(0.71\%, 32.60\%, 66.69\%)$. The first one is mainly focused on the absolute diversification benefit of each agent, while it is my belief that the second one takes also into account the relative riskiness of each agent with respect to the rest of agents.

Using the relative principles displayed in Table 9.1 two rankings of principles are deduced, following the proposal in Section 9.2.2. If symbols $\sigma, \nabla, \text{EBA}$ and 0 are now used to refer to the proportional principle based on the standard deviation, the gradient allocation principle, the excess based allocation principle and the neutral allocation principle, respectively, then the following simplicial distances may be calculated from expression (9.3):

$$\begin{aligned} \Delta_{\sigma,0} &= \sqrt{3.557 + 0.522 + 1.354} = 2.3308 \quad , \\ \Delta_{\nabla,0} &= \sqrt{7.764 + 1.072 + 3.066} = 3.4499 \quad , \\ \Delta_{\text{EBA},0} &= \sqrt{6.608 + 0.932 + 2.576} = 3.1806 \quad , \end{aligned} \tag{9.11}$$

Table 9.2: Perturbation inverse relative capital allocation principles and their absolute counterparts assuming that the amount shared is 6,124.

	Single life annuity (X_{sl})	Survivor annuity (X_{surv})	Death benefit (X_{db})
		Relative principles	
Proportional principle based on the st.dev.	2.80%	38.05%	59.15%
Gradient allocation principle	0.71%	32.60%	66.69%
Excess based allocation principle	1.00%	34.19%	64.81%
		Absolute principles	
Proportional principle based on the st.dev.	172	2,330	3,622
Gradient allocation principle	44	1,996	4,084
Excess based allocation principle	61	2,094	3,969

and

$$\begin{aligned} \Delta_{\sigma,\nabla} &= \sqrt{0.8107 + 0.0978 + 0.3452} = 1.1197 \quad , \\ \Delta_{\text{EBA},\nabla} &= \sqrt{0.0466 + 0.0049 + 0.0213} = 0.2698 \quad . \end{aligned} \tag{9.12}$$

Isolated values are not too much informative but allows to rank the principles with respect to one principle of reference, as it is graphically shown in Figure 9.2.4. From results (9.11), it can be deduced that in this example the gradient allocation principle is the one furthest from the neutral allocation, followed by the EBA allocation and the σ one in this order. Looking at Table 9.1 this ranking seems reasonable. On the other hand, imagine that you do not have the information (9.11) and decide to take as principle of reference the ∇ one. From results (9.12), you can find that the EBA one is much similar to the reference principle than the σ principle. Again, looking at Table 9.1 it seems reasonable. Last comment on results (9.11) and (9.12) is that the rankings deduced from them are not contradictory.

As a final application, let us *average* the three relative allocations displayed in Table 9.1 using the strategy explained in Section 9.2.3. In

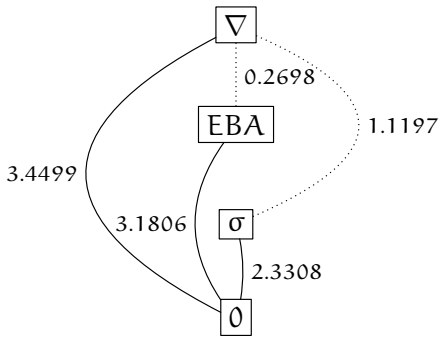


Figure 9.1: Ranking of capital allocation principles using the simplicial distance in the worked example. Solid lines are linked to the resulting classification where comparing against the neutral allocation principle 0. Dotted lines are linked to the distances where the principle of reference is the gradient principle ∇.

other words, let $\vec{x}_\sigma$, $\vec{x}_\nabla$ and $\vec{x}_{\text{EBA}}$ denote each of the three relative allocations displayed in Table 9.1, this is:

$$\begin{aligned}
\vec{x}_\sigma &= (89.20\%, 6.57\%, 4.23\%),\\
\vec{x}_\nabla &= (96.84\%, 2.12\%, 1.04\%) \text{ and}\\
\vec{x}_{\text{EBA}} &= (95.74\%, 2.79\%, 1.47\%).
\end{aligned}$$

Let me compute the three ($n = 3$) geometric means of the three ($m = 3$) components of these relative allocations and call them G_1, G_2 and G_3, respectively. Their values are

$$\begin{aligned}
G_1 &= (89.20\% \cdot 96.84\% \cdot 95.74\%)^{1/3} = 93.87\%,\\
G_2 &= (6.57\% \cdot 2.12\% \cdot 2.79\%)^{1/3} = 3.39\% \text{ and}\\
G_3 &= (4.23\% \cdot 1.04\% \cdot 1.47\%)^{1/3} = 1.88\%.
\end{aligned}$$

Following expression (9.8), last thing to be done is to calculate the value of $\mathcal{C}[(G_1, G_2, G_3)]$, the closure of the vector with components being the previous geometric means G_1,G_2 and G_3. Recall that with this last step the value of $\text{AM}_\Delta\left(\vec{x}_\sigma, \vec{x}_\nabla, \vec{x}_{\text{EBA}}\right)$ is obtained. This simplicial average is a relative allocation principle. Both relative and absolute results are shown in Table 9.3.

As it can be checked, the components of the simplicial average are not equal to the arithmetic mean of the components of the original absolute (or relative) principles. Actually, the components of the simplicial average are linked to the geometric mean of the components of the original relative principles.

	Single life annuity (X_{sl})	Survivor annuity (X_{surv})	Death benefit (X_{db})
AM_{Δ} (relative principle)	94.71%	3.42%	1.88%
Average (as an absolute principle)	356,431	12,859	7,066

Table 9.3: Average of the three capital allocation principles from Table 9.1.

Part IV

CONCLUSIONS

DISCUSSION AND FUTURE CHALLENGES

The first contribution of this dissertation is the introduction of GlueVaR risk measures, a family belonging to the more general class of distortion risk measures. One of the advantages of GlueVaR over other alternatives that can be found in the actuarial literature is its simplicity. This new family combines the most popular risk measures in both insurance and financial sectors and considers more than just one parameter to capture managerial and regulatory attitudes towards risk. Additionally, it is possible to work out analytical closed-form expressions for many statistical distributions that are frequently used in insurance and financial applications. As it is discussed in this thesis, the GlueVaR family should enhance the way in which regulatory capital requirements are calculated, as GlueVaR can incorporate more information about agents' attitudes to risk and can be useful in helping regulators and practitioners reach a consensus. The incorporation of qualitative information in decision making tools is essential for risk managers and, as such, the GlueVaR risk measures can play a key role in achieving this goal.

The definition of the tail-subadditivity property for a pair of risks may be considered the second contribution. It has been argued that this property is related to the ability of risk measures of become subadditive in extremely adverse scenarios. Going deeper in its implications should be a line of further research.

It has been shown that distortion risk measures, OWA and WOWA operators in the discrete finite case are mathematically linked by means of the Choquet integral. Aggregation operators are used as tools to summarize human subjectivity in decision making and have a direct connection to risk measurement of discrete random variables. From the risk management point of view, the main contribution is that it is shown how distortion risk measures may be derived -and then computed- from ordered weighted averaging operators. The mathematical links presented in Chapter 4 may help to interpret distortion risk measures under the information sciences' perspective. It is shown that the aggregation preference of the expert may be measured by means of the local degree of orness of the distortion risk measure. Regulatory capital requirements and provisions may then be associated to the aggregation attitude of the regulator and the risk managers, respectively. The mathematical link between risk measurement and information sciences' concepts presented in this thesis of-

fers an extra perspective in quantitative risk management.

New indicators for helping to characterize the discrete Choquet integral have also been presented in this dissertation. The aim is complementing those already available, so that a more complete formulation might be provided which covers a wider range of situations. This need arises because at times it is necessary to highlight different features of the aggregation function than those provided by the degree of orness and the entropy of dispersion. The degree of balance, the divergence, the variance indicator and Rényi entropies as indicators within the framework of the Choquet integral have been introduced in Chapter 5, following this spirit. It has been shown that these four indicators, which are commonly used for the OWA operator, can also be considered for the Choquet aggregation, and we have discussed the potential of these indicators to provide supplementary information to decision makers.

The indicators for characterizing the discrete Choquet integral were defined from a local and a global perspective. The local perspective proposed in the definition of these indicators may be preferred to the global one in certain applications, those in which the number of values to be aggregated is large and an increasing ordering of input values can be assumed, like in some statistical or actuarial applications. It is the case, for instance, of the example shown in Chapter 6.

Despite the fact that, in practice, risk management decisions are usually taken in the discrete and finite world, some comments must be made on the possibility to extend the results to the context of countable or continuous random variables. Countable and continuous cases have received much less attention in information systems literature in comparison to the discrete and finite case. Up to the best of my knowledge, proposals of aggregation functions with countable [Grabisch et al., 2009] or continuous [Yager, 2004b; Yager and Xu, 2006] arguments are scarcely used by fuzzy experts. A next natural step in this research might be the analysis of countable probability spaces. Considering convenient aggregation operators with countable arguments and setting additional conditions regarding convergence of series, it seems that results shown in Chapters 4 and 5 might be extended to the countable case.

A major contribution derived from the relationship between distortion risk measures and aggregation operators is the characterization of the risk attitude implicit into the choice of a risk measure and a tolerance level. It has been pointed out that the risk value returned by a distortion risk measure basically depends on the characteristics of the random variable, which are collected into the survival distri-

bution, and the associated distortion function. These two parts are disaggregated in Chapter 6 and the attention has been focused on the analysis of the distortion function. The distortion function can be understood as a weighting function of the survival probabilities. Therefore, the risk attitude implicit in a distortion risk measure is to some extent contained in its distortion function. Two aggregation indicators and a quotient function are used to describe alternative features of the distortion function. It is shown that these mathematical devices give insights of the implicit risk behavior involved in risk measures.

In short, distortion risk measures may be understood as belonging to a class of aggregation functions (Choquet integrals) and a set of indicators is developed in the literature for helping to characterize aggregation functions (as shown in Chapter 5). The local and global degrees of orness are then interpreted under a risk assessment perspective and it is shown that these indicators provide approximations of the area under the distortion function of the distortion risk measure. The size of the area reveals the accumulated distortion made over the survival distribution. So, these indicators provide valuable information about the overall risk attitude when a particular distortion risk measure is applied. A conclusion drawn from the analysis is that the degree of orness and the area under the distortion function can be interchangeably used to study the overall risk behavior.

Risk attitude implicit in a risk measure is not only determined by the area size of the distortion function. As important as the size of the overall distortion is how this aggregated distortion is distributed through the range of values. The risk information provided by aggregation indicators should be complemented by the quotient function defined in Chapter 6, which is the fraction between the distortion function and the identity function. It is argued that a graphical analysis of this quotient function gives local information of the specific risk behavior associated to the distortion risk measure at any point of the range of values of the loss. In addition, the area under the quotient function provides similar overall risk information than the area under the distortion function, so this quotient function could replace the distortion function in the analysis of overall risk behavior.

As an application of these findings, the risk behavior linked to the Value-at-Risk is analyzed. It is illustrated that in some situations the VaR_α is not sensitive to changes into more risky scenarios. The reason is that this risk measure involves two extreme risk attitudes, i.e. maximum risk aversion in $[1-\alpha, 1)$ and maximum risk loving in $(0, 1-\alpha)$. It is emphasized that additional risk information may be obtained by means of comparable GlueVaR risk measures, which are calibrated

to return the same risk value than VaR_α in a particular context. The distortion functions of these GlueVaR risk measures are examined. The study of the risk behavior implicit in the choice of these GlueVaR measures gives a more complex characterization of the risk framework providing, for instance, boundary risk attitudes.

The example shown in Chapter 6 provides an illustration of how GlueVaR risk measures can be used in practice. It is claimed that GlueVaR measures should be used by regulators and risk managers to enrich risk information provided by the standard quantile. Regulators may have a better knowledge of risks faced by regulated firms and risk managers a deeper control of undertaken risks. Note that the calibration procedure depends on the risk measure of reference and also on the random variable. Then, the set of comparable GlueVaR measures can be different if the random variable changes. This is not necessarily an unbridgeable drawback when asking the supervisory authorities to change the risk measure if the insurance undertaking is able to justify the GlueVaR selection process. For instance, these measures may be used to analyze variations from one year to the next in the implicit risk aversion of boundary cases when the VaR_α is applied to assess the annual risk.

In the third part of this thesis, a particular vision on capital allocation problems has been proposed. It has been argued that capital allocation principles with a return&risk-reward objective are of special interest, because insurance risk management practitioners typically have to deal with two opposing demands: on the one hand, they want business units to achieve or outperform the objectives fixed by the firm's executive committee, yet, on the other, they are responsible for controlling their economic risks. Finding a trade-off between these two demands is the challenging task that risk managers face on a daily basis. Examples of capital allocation principles that cover this necessity are commented, as it is the case of gradient allocation principles and of some of the proposals investigated in Chapter 9.

Two main contributions on capital allocation has been made. On the one hand, it has been shown that GlueVaR risk measures are as useful as other alternatives like VaR or TVaR to solve capital allocation problems, because both stand-alone and based on risk contributions proportional allocation principles can be obtained using GlueVaR measures. Moreover, it has been indicated how to obtain non-proportional allocation principles using GlueVaR with the help of the framework proposed in Dhaene et al. [2012b]. The second contribution is understanding capital allocation principles as compositional data. This interpretation of capital allocation principles allows the connection between aggregation operators and capital allocation

problems in an appealing way.

This thesis contains some preliminary ideas on this connection, such as taking advantage of the simplicial distance or the simplicial arithmetic mean. Most of these ideas were inspired by the opening plenary session given by professor DeBaets in the AGOP 2013 conference that took place in Pamplona, and later from the lecture of Aitchinson and Egozcue [2005]. It is my belief that there is room for further and promising research on this field. Nonetheless, some concerns on this perspective on capital allocation principles should be addressed, as its economic, financial or actuarial interpretation.

I would like to acknowledge some limitations of this thesis and to share some ideas that were initially explored. Regarding limitations, all the theoretical and practical developments both for risk measures and capital allocation problems have been focused on a static setting (only one period at a time). Therefore, dynamic perspectives have not been taken into account. The interested reader in dynamic capital allocation is referred to Hamada et al. [2006]; Froot [2007]; Diers [2011]; Ai et al. [2012], for instance; the one interested in risk measures in a dynamic context is referred to Tsanakas [2004]; Riedel [2004]; Weber [2006]; Artzner et al. [2007] and the references therein. With respect to GlueVaR, it is an interesting family of risk measures for practitioners, but this is not a guarantee on their future use in the insurance and financial sectors. The tail-subadditivity property seems to be difficult to be satisfied when increasing the number of random variables to aggregate, because each additional random variable involves additional intersections to find out the common tail. Regarding some relationships between capital allocation problems and aggregation operators that were initially considered, it was my prior belief that a relationship between mixture operators [Marques Pereira and Ribeiro, 2003; Calvo et al., 2004; Mesiar and Spirkova, 2006; Mesiar et al., 2008] and objective functions in the framework proposed by Dhaene et al. [2012b] should lead to interesting results, but this line of research was finally not followed.

The theoretical research presented in this thesis regarding risk measures and capital allocation problems has important practical implications. Only as an example, in the particular case of insurance companies in the European Union, both the assessment of their solvency and capital allocation exercises are required by Directive 2009/138/EC of 25 November 2009 (known as Solvency II directive). Concretely, insurance companies are:

- forced to calculate standard Solvency Capital Requirements (SCR, the minimum cushion of economic liquidity required to the in-

Article 101(3) of the Directive

stitution) using Value-at-Risk (VaR) risk measure at 99.5% confidence level;

Article 122 of the Directive

- allowed to apply for the use of total or partial internal models to calculate SCR and, in particular, to select a different risk measure than VaR and/or a different confidence level than 99.5%;

Article 45 of the Directive

- forced to perform Own Risk and Solvency Assessment (ORSA), for wich freedom in the risk measure and the confidence level used is allowed;

Article 123 of the Directive

- forced to deliver, at least annually, a capital allocation exercise linked to the Attribution of Profit and Loss Report.

All the previous arguments lead to the conclusion that the theoretical developments shown in this dissertation can be of particular practical relevance for insurance companies, not only regarding the selection of a risk measure reflecting multiple risk attitudes but also when solving capital allocation problems.

Part V

APPENDIX

A

APPENDIX

A.1 EQUIVALENT EXPRESSION FOR THE GLUEVAR DISTORTION FUNCTION

Details on the definition of the GlueVaR distortion function $\kappa_{\beta,\alpha}^{h_1,h_2}(u)$ as a linear combination of the distortion functions of TVaR at confidence levels β and α, and VaR at confidence level α are provided, i.e. an explanation of how to obtain expression (1.3) can be found here. Expression (1.1) of the distortion function $\kappa_{\beta,\alpha}^{h_1,h_2}(u)$ can be rewritten as,

$$\begin{aligned}\kappa_{\beta,\alpha}^{h_1,h_2}(u) = \; & h_1 \cdot \gamma_\beta(u) \cdot \mathbb{1}[0 \leqslant u < 1-\beta] + \\ & + \left(h_1 + \frac{h_2 - h_1}{\beta - \alpha} \cdot (1-\alpha) \cdot \gamma_\alpha(u) - \right. \\ & \left. \frac{h_2 - h_1}{\beta - \alpha} \cdot (1-\beta) \right) \cdot \mathbb{1}[1-\beta \leqslant u < 1-\alpha] + \\ & + \psi_\alpha(u),\end{aligned} \tag{A.1}$$

where $\mathbb{1}[x_1 \leqslant u < x_2]$ is an indicator function, so it takes a value of 1 if $u \in [x_1, x_2)$ and 0 otherwise.

Note that

$$\gamma_\beta(u) \cdot \mathbb{1}[0 \leqslant u < 1-\beta] = \gamma_\beta(u) - \psi_\beta(u), \tag{A.2}$$

$$\mathbb{1}[1-\beta \leqslant u < 1-\alpha] = \psi_\beta(u) - \psi_\alpha(u), \tag{A.3}$$

$$\begin{aligned}&\gamma_\alpha(u) \cdot \mathbb{1}[1-\beta \leqslant u < 1-\alpha] = \\ &\gamma_\alpha(u) - \psi_\alpha(u) - \left(\frac{1-\beta}{1-\alpha}\right) \cdot \left[\gamma_\beta(u) - \psi_\beta(u)\right].\end{aligned} \tag{A.4}$$

Taking into account expressions (A.2), (A.3) and (A.4), expression (A.1) may be rewritten as,

$$\begin{aligned}\kappa_{\beta,\alpha}^{h_1,h_2}(u) = &\left[h_1 - \frac{(h_2-h_1)\cdot(1-\beta)}{\beta-\alpha}\right]\cdot\gamma_\beta(u) + \\ &\left[-h_1 + h_1 - \frac{(h_2-h_1)\cdot(1-\beta)}{\beta-\alpha} + \right. \\ &\left.\frac{(h_2-h_1)\cdot(1-\beta)}{\beta-\alpha}\right]\cdot\psi_\beta(u) + \\ &\frac{h_2-h_1}{\beta-\alpha}\cdot(1-\alpha)\cdot\gamma_\alpha(u) + \\ &\left[1 - h_1 + \frac{(h_2-h_1)\cdot(1-\beta)}{\beta-\alpha} - \right. \\ &\left.\frac{h_2-h_1}{\beta-\alpha}\cdot(1-\alpha)\right]\cdot\psi_\alpha(u).\end{aligned} \tag{A.5}$$

Given that $\omega_1 = h_1 - \frac{(h_2-h_1)\cdot(1-\beta)}{\beta-\alpha}$, $\omega_2 = \frac{h_2-h_1}{\beta-\alpha}\cdot(1-\alpha)$ and $\omega_3 = 1-h_2$, expression (1.3) follows directly from (A.5).□

A.2 BIJECTIVE RELATIONSHIP BETWEEN HEIGHTS AND WEIGHTS AS PARAMETERS FOR GLUEVAR RISK MEASURES

Pairs of GlueVaR heights (h_1, h_2) and weights (ω_1, ω_2) are linearly related to each other. The parameter relationships are $(h_1, h_2)' = H\cdot(\omega_1,\omega_2)'$ and, inversely, $(\omega_1,\omega_2)' = H^{-1}\cdot(h_1,h_2)'$, where H and H^{-1} matrices are $H = \begin{pmatrix} 1 & \frac{1-\beta}{1-\alpha} \\ 1 & 1 \end{pmatrix}$ and $H^{-1} = \begin{pmatrix} \frac{1-\alpha}{\beta-\alpha} & \frac{\beta-1}{\beta-\alpha} \\ \frac{\alpha-1}{\beta-\alpha} & \frac{1-\alpha}{\beta-\alpha} \end{pmatrix}$, respectively.

A.3 TAIL-SUBADDITIVITY FOR GLUEVAR RISK MEASURES

This appendix is devoted to the proof of Theorem 2.1.1. Given a confidence level α and a pair of random variables X and Y so that $\mathcal{Q}_{\alpha,X,Y} \neq \emptyset$, a GlueVaR risk measure is tail-subadditive if its associated distortion function $\kappa_{\beta,\alpha}^{h_1,h_2}$ is concave in $[0, 1-\alpha)$.

The subadditivity theorem and the integration on subsets of Ω are defined as in Denneberg [1994]:

- **Subadditivity theorem.** Let $\mu : 2^\Omega \to \overline{\mathbb{R}}_+$ be a monotone, submodular set function. Then for functions $X, Y : \Omega \to \overline{\mathbb{R}}$ being μ-essentially$> -\infty$

$$\int (X+Y)\,d\mu \leqslant \int X d\mu + \int Y d\mu.$$

If μ is continuous from below the assumption on X, Y being μ-essentially $> -\infty$ can be dropped.

- **Integration on subsets.** Let μ be a monotone set function on a set system $\mathcal{S} \subset 2^{\Omega}$ with $\Omega \in \mathcal{S}$ and closed under intersection. For $A \in \mathcal{S}$ define $\mu_A(B) := \mu(B \cap A)$, $B \in \mathcal{S}$. Then μ_A is a monotone set function on $\mathcal{S}$ and we define $\int_A X d\mu := \int X d\mu_A$.

A set system is, generally speaking, a collection of sets. Definitions of monotone, modular or submodular set functions, as well as the definition of continuity from below, are given in next paragraphs. A proper definition of a function X μ-essentially $> -\infty$ is not needed in the proof and, thus, not provided. Interested readers can find this definition in Denneberg [1994].

According to definition 2.1.1, given a confidence level α and taking into account that $\mathcal{Q}_{\alpha,X,Y} \neq \emptyset$ for the fixed pair of random variables, i.e. $X, Y : \Omega \to \overline{\mathbb{R}}$, the tail-subadditivity property is satisfied by a distortion risk measure ρ_g if the subadditivity theorem can be applied to the set function $(g \circ P)_{\mathcal{Q}_{\alpha,X,Y}}$, i.e. the set function so that for any $B \in 2^{\Omega}$, $(g \circ P)_{\mathcal{Q}_{\alpha,X,Y}}(B) = g(P(B \cap \mathcal{Q}_{\alpha,X,Y}))$.

Therefore, subadditivity in the tail for a pair of risks is proven if $(g \circ P)_{\mathcal{Q}_{\alpha,X,Y}}$ is submodular and continuous from below.

Let me provide at this stage of the proof some definitions regarding the modularity of set functions. A set function μ is modular if $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$, and it is submodular if $\mu(A \cup B) + \mu(A \cap B) \leqslant \mu(A) + \mu(B)$. A set function μ is monotone if $\mu(A) \leqslant \mu(B)$ for any $A \subseteq B$ in 2^{Ω}.

Now, if ρ_g is a distortion risk measure such that its associated distortion function g is concave in $[0, 1-\alpha)$, then it is shown that $(g \circ P)_{\mathcal{Q}_{\alpha,X,Y}}$ is submodular. Consider the set function ν defined by $\nu(B) := P(B \cap \mathcal{Q}_{\alpha,X,Y})$, for any $B \in 2^{\Omega}$. Note that $\nu(B) \in [0, 1-\alpha)$ because $P(\mathcal{Q}_{\alpha,X,Y}) < 1-\alpha$ and P is a monotone set function. The set function ν is modular because P is modular, i.e. $\nu(A \cup B) + \nu(A \cap B) = \nu(A) + \nu(B)$ for any $A, B \in 2^{\Omega}$. Given $A, B \in 2^{\Omega}$ suppose, without loss of generality, that $A \subseteq B$. Let me rename $a := \nu(A)$, $b := \nu(B)$, $i := \nu(A \cap B)$ and $u := \nu(A \cup B)$. Because ν is monotone then it holds that $i \leqslant a \leqslant b \leqslant u$ due to $A \cap B \subseteq A \subseteq B \subseteq A \cup B$. The modularity of ν implies that $i + u = a + b$, i.e. $[i, u]$ and $[a, b]$ have common centers, $\frac{i+u}{2} = \frac{a+b}{2}$. Then, because g is concave in $[i, u]$ it can be concluded that $g(u) + g(i) \leqslant g(a) + g(b)$ or, equivalently, that $g \circ \nu = (g \circ P)_{\mathcal{Q}_{\alpha,X,Y}}$ is submodular.

The property of continuity from below of $g \circ \nu = (g \circ P)_{\mathcal{Q}_{\alpha,X,Y}}$ must also be satisfied to use the subadditivity theorem. An arbitrary set function μ is continuous from below if for any increasing collection of subsets in the set system ($A_n \in \mathcal{S}$, $A_n \subseteq A_{n+1}$ for $n \in \mathbb{N}$) so that $A := \cup_{n=1}^{\infty} A_n \in \mathcal{S}$ then equality $\lim_{n\to\infty} \mu(A_n) = \mu(A)$ holds. So $\mu =$

$(g \circ P)_{\mathcal{Q}_{\alpha,X,Y}}$ is continuous from below because $(g \circ P)_{\mathcal{Q}_{\alpha,X,Y}} : 2^{\Omega} \longrightarrow g([0, 1-\alpha))$, P is a probability (hence, continuous from below) and g is continuous on $[0, 1-\alpha)$.

Given that $(g \circ P)_{\mathcal{Q}_{\alpha,X,Y}}$ is submodular and continuous from below, applying the subadditivity theorem and using integration on subsets, it is true that, given X and Y:

$$\int_{\mathcal{Q}_{\alpha,X,Y}} (X+Y)\, d(g \circ P) \leqslant \int_{\mathcal{Q}_{\alpha,X,Y}} X d(g \circ P) + \int_{\mathcal{Q}_{\alpha,X,Y}} Y d(g \circ P),$$

which prove that the associated risk measure ρ_g is tail-subadditive.

Consider a GlueVaR risk measure so that weights (ω_1, ω_2) belong to the lightly shaded area in Figure 2.1. This is a sufficient condition to guarantee concavity of the distortion function on $[0, 1-\alpha)$. Therefore, these GlueVaR risk measures are candidates to satisfy the tail-subadditivity property. □

A.4 RELATIONSHIP BETWEEN GLUEVAR AND TAIL DISTORTION RISK MEASURES

This section of the appendix is intended to present the proof of Proposition 1.5.1. Following the notation introduced along this work, as for any random variable X it holds that $\text{GlueVaR}^{\omega_1,\omega_2}_{\beta,\alpha}(X) = \int X d\mu$ with $\mu = \kappa^{\omega_1,\omega_2}_{\beta,\alpha} \circ P$ and $T_{g,\alpha}(X) = \int X d\eta$ with $\eta = g_\alpha \circ P$, proving proposition 1.5.1 is equivalent to proving that $\kappa^{\omega_1,\omega_2}_{\beta,\alpha} = g_\alpha$ under the proper conditions on ω_1, ω_2 and g.

On one hand, suppose that $\omega_2 = 1 - \omega_1$ and that g is given by expression (1.7). First of all, let me rewrite g as

$$g(t) = \left(\frac{\omega_1 \cdot (1-\alpha)}{1-\beta} + \omega_2 \right) \cdot t \cdot \mathbb{1}\left[0 \leqslant t < (1-\alpha)^{-1} \cdot (1-\beta) \right] + (\omega_1 + \omega_2 \cdot t) \cdot \mathbb{1}\left[(1-\alpha)^{-1} \cdot (1-\beta) \leqslant t \leqslant 1 \right] \tag{A.6}$$

Recall that g_α is built as $g\left(\frac{u}{1-\alpha}\right) \cdot \mathbb{1}[0 \leqslant u < 1-\alpha] + \mathbb{1}[1-\alpha \leqslant u \leqslant 1]$. If u is less than $1-\beta$ therefore $t = \frac{u}{1-\alpha}$ is less than $(1-\alpha)^{-1} \cdot (1-\beta)$; if u is comprised between $1-\beta$ and $1-\alpha$, then $t = \frac{u}{1-\alpha}$ satisfies that $(1-\alpha)^{-1} \cdot (1-\beta) \leqslant t \leqslant 1$. Summarizing,

$$g_\alpha(u) = \begin{cases} \left[\frac{\omega_1}{1-\beta} + \frac{\omega_2}{1-\alpha} \right] \cdot u & \text{if} \quad 0 \leqslant u < 1-\beta \\ \omega_1 + \frac{\omega_2}{1-\alpha} \cdot u & \text{if} \quad 1-\beta \leqslant u < 1-\alpha \\ 1 & \text{if} \quad 1-\alpha \leqslant u \leqslant 1 \end{cases} \tag{A.7}$$

which is the definition of distortion function $\kappa^{\omega_1,\omega_2}_{\beta,\alpha}$ as shown in (1.5).

On the other hand, consider as starting point the aforementioned expression (1.5) of $\kappa_{\beta,\alpha}^{\omega_1,\omega_2}$. As pointed out, g_α is always continuous in $1-\alpha$. Consequently parameters of $\kappa_{\beta,\alpha}^{\omega_1,\omega_2}$ must be such that guaranty continuity of the equivalent g_α in $1-\alpha$. In other words, $\lim_{u\uparrow(1-\alpha)} \kappa_{\beta,\alpha}^{\omega_1,\omega_2}(u) = \omega_1 + \omega_2 = 1 = \lim_{u\downarrow(1-\alpha)} \kappa_{\beta,\alpha}^{\omega_1,\omega_2}(u)$. This is exactly condition $\omega_2 = 1-\omega_1$. Now, forcing $g_\alpha = \kappa_{\beta,\alpha}^{\omega_1,\omega_2}$, it is straightforward to go backwards from expression (A.7) to expression (1.7) to complete the proof. □

A.5 Proof of the ordered weighted averaging operator uniqueness

Given two different vectors $\vec{w}$ and $\vec{u}$ from $[0,1]^n$ we wonder if $OWA_{\vec{w}} = OWA_{\vec{u}}$, i.e. if the respective OWA operators on $\mathbb{R}^n$ are the same. It is shown that this is not possible. Suppose that, for all $\vec{x} \in \mathbb{R}^n$, $OWA_{\vec{w}}(\vec{x}) = OWA_{\vec{u}}(\vec{x})$. Let vectors $\vec{z}_k \in \mathbb{R}^n, k = 1,...,n$ be defined by

$$\vec{z}_{k,i} = \begin{cases} 0 & \text{if} \quad i < k \\ 1/(n-i+1) & \text{if} \quad i \geqslant k \end{cases}.$$

Then, iterating from $k = n$ to $k = 1$:

- **Step** $k = n$. We have $\vec{z}_n = (0,0,...,0,1)$, and permutation $\sigma = id$ is useful to calculate $OWA_{\vec{w}}(\vec{z}_n)$ and $OWA_{\vec{u}}(\vec{z}_n)$. Precisely, $OWA_{\vec{w}}(\vec{z}_n) = 1 \cdot w_n$ and $OWA_{\vec{u}}(\vec{z}_n) = 1 \cdot u_n$. If $OWA_{\vec{w}} = OWA_{\vec{u}}$, then $u_n = w_n$.
- **Step** $k = n-1$. We have $\vec{z}_{n-1} = \left(0,0,...,\frac{1}{2},1\right)$, and permutation $\sigma = id$ is still useful. So $OWA_{\vec{w}}(\vec{z}_{n-1}) = \frac{1}{2} \cdot w_{n-1} + 1 \cdot w_n$ and, taking into account the previous step, $OWA_{\vec{u}}(\vec{z}_{n-1}) = \frac{1}{2} \cdot u_{n-1} + 1 \cdot w_n$. If the hypothesis $OWA_{\vec{w}} = OWA_{\vec{u}}$ holds, then $u_{n-1} = w_{n-1}$.
- **Step** $k = i$. From previous steps we have that $u_j = w_j$, $j = i+1,...,n$ and in this step we obtain $u_i = w_i$.
- **Step** $k = 1$. Finally, supposing again that $OWA_{\vec{w}} = OWA_{\vec{u}}$, we obtain that $u_j = w_j$ for all $j = 1,...,n$. But this is a contradiction with the fact that $\vec{w} \neq \vec{u}$.

A.6 An ascending quadratic weighted additive set function is a capacity defined on N

To prove that η is a capacity on N it is necessary to see that $\eta(A) \in [0, K_\eta]$ for some $K_\eta > 0$ for all $A \subseteq N$, $\eta(\emptyset) = 0$ and that $\eta(A) \leqslant$

$\eta(B)$ if $A \subseteq B$. By definition of AQWA capacity, $\eta(\emptyset) = 0$. If $\eta(\{m_k\}) \geqslant 0$ for all $m_k \in N$, then $\eta(A) \leqslant \eta(B)$ if $A \subseteq B$ because $\eta(A) := \sum_{m_k \in A} \eta(\{m_k\})$. So let us see that $\eta(\{m_k\}) \geqslant 0$ for all $m_k \in N$.
Recall that by definition of AQWA capacity, for all $j = 1, ..., n$

$$\eta\left(\{m_j\}\right) := 2 \cdot \left(\frac{j-1}{n-1}\right)^2 \cdot \left[\binom{n}{n-j+1}^{-1} \sum_{\substack{A \subseteq N \\ |A|=n-j+1}} \mu(A) - \binom{n}{n-j}^{-1} \sum_{\substack{A \subseteq N \\ |A|=n-j}} \mu(A) \right].$$

The first factor is less than or equal to 2, because $\frac{j-1}{n-1} \leqslant 1$. This is, the first factor is greater than or equal to 0.
For the second factor, and for each j, two situations are considered: whether $s_{n-j+1} = \#\{A \quad \text{s.t.} \quad |A| = n-j+1\} = \binom{n}{n-j+1}$ is greater or less than $s_{n-j} = \#\{A \quad \text{s.t.} \quad |A| = n-j\} = \binom{n}{n-j}$. Once this notation is introduced, this second factor may be rewritten as

$$s_{n-j+1}^{-1} \sum_{\substack{A \subseteq N \\ |A|=n-j+1}} \mu(A) - s_{n-j}^{-1} \sum_{\substack{A \subseteq N \\ |A|=n-j}} \mu(A).$$

So, supposing j is fixed:

- If $s_{n-j+1} \geqslant s_{n-j}$, then

$$\begin{aligned} & s_{n-j+1}^{-1} \sum_{\substack{A \subseteq N \\ |A|=n-j+1}} \mu(A) - s_{n-j}^{-1} \sum_{\substack{A \subseteq N \\ |A|=n-j}} \mu(A) \\ & \geqslant s_{n-j+1}^{-1} \sum_{\substack{A \subseteq N \\ |A|=n-j}} \mu(A) - s_{n-j}^{-1} \sum_{\substack{A \subseteq N \\ |A|=n-j}} \mu(A) \\ & = \left(\frac{(n-j+1)!(j-1)!}{n!} - \frac{(n-j)!(j)!}{n!} \right) \cdot \sum_{\substack{A \subseteq N \\ |A|=n-j}} \mu(A) \\ & = \frac{(j-1)!(n-j)!(n+1)}{n!} \cdot \sum_{\substack{A \subseteq N \\ |A|=n-j}} \mu(A) \geqslant 0. \end{aligned} \tag{A.8}$$

  The hypothesis is used to ensure that the first inequality holds, because $\sum_{\substack{A \subseteq N \\ |A|=n-j+1}} \mu(A) \geqslant \sum_{\substack{A \subseteq N \\ |A|=n-j}} \mu(A)$ under the hypothesis. This is true due to the fact that there are fewer summands on the right-hand side ($s_{n-j} \leqslant s_{n-j+1}$) and, in addition, each summand on the right is less than or equal to one on the left (μ is

monotone);

- If $s_{n-j+1} < s_{n-j}$, then it can be shown that

$$
\begin{aligned}
& s_{n-j}^{-1}\left[\sum_{\substack{A\subseteq N\\ |A|=n-j+1}}\mu(A)+\frac{s_{n-j}-s_{n-j+1}}{s_{n-j+1}}\sum_{\substack{A\subseteq N\\ |A|=n-j+1}}\mu(A)\right]\\
& -s_{n-j}^{-1}\sum_{\substack{A\subseteq N\\ |A|=n-j}}\mu(A)\\
& \geqslant s_{n-j}^{-1}\left[\sum_{\substack{A\subseteq N\\ |A|=n-j+1}}\mu(A)+\frac{s_{n-j}-s_{n-j+1}}{s_{n-j+1}}\sum_{\substack{A\subseteq N\\ |A|=n-j+1}}\mu(A)\right]\\
& -s_{n-j}^{-1}\sum_{\substack{B\subseteq N\\ |B|=n-j}}\left(\sum_{\substack{A\subseteq N\\ |A|=n-j+1}}s_{n-j+1}^{-1}\,\mu(A)\right)=0
\end{aligned}
\tag{A.9}
$$

  In this case, the hypothesis is used to prove inequality: for any $B \subseteq N$ such that $|B| = n-j$, $\mu(B) \leqslant \sum_{\substack{A\subseteq N\\ |A|=n-j+1}} s_{n-j+1}^{-1}\,\mu(A)$ under the hypothesis. Otherwise, a contradiction with the fact that μ is monotone arises.

As this result implies $\eta(A) \leqslant \eta(B)$ if $A \subseteq B$, if $\eta(N) = K_\eta > 0$ is shown then $\eta(A) \in [0, K_\eta]$ will hold for each $A \subseteq N$. To see that $\eta(N) = K_\eta > 0$, note first that for all $j = 1, ..., n$,

$$
\begin{aligned}
\eta\left(\{m_j\}\right) &\leqslant 2\cdot\left(\frac{j-1}{n-1}\right)^2\cdot\left[\binom{n}{n-j+1}^{-1}\sum_{\substack{A\subseteq N\\ |A|=n-j+1}}\mu(A)\right]\\
&\leqslant 2\cdot\left(\frac{j-1}{n-1}\right)^2\cdot K_\mu.
\end{aligned}
$$

Given previous inequalities,

$$
\begin{aligned}
\eta(N) &= \sum_{m_j\in N}\eta\left(\{m_j\}\right)\leqslant\sum_{j=1}^{n}2\cdot\left(\frac{j-1}{n-1}\right)^2\cdot K_\mu\\
&= \frac{n(2n-1)}{3(n-1)}\cdot K_\mu := K_\eta > 0\,,
\end{aligned}
$$

because $K_\mu > 0$ taking into account that μ is a capacity. Hence, the fact that η is a capacity has been proved.

A.6.1 *Two particular cases of AQWA capacities*

Two special cases of AQWA capacities are those linked to symmetric capacities and those linked to additive capacities:

- If μ is symmetric then $\mu(A)$ depends only on the number of elements of A and not on the particular elements of A. Considering the weights w_j for all $j \in \{1,...,n\}$, such that $\mathcal{C}_\mu = \text{OWA}_{\vec{w}}$, it is known that $\mu(A) = \sum_{u=0}^{n-j} w_{n-u}$ for a set $A \subseteq N$ with $n-j+1$ elements (as shown in Section 5.1). If index u is changed by $k = n-u$, then $\mu(A) = \sum_{k=n}^{j} w_k = \sum_{k=j}^{n} w_k$ since $k = n$ and $k = j$ when $u = 0$ and $u = n-j$, respectively.

  In that case, $\sum_{\substack{A\subseteq N \\ |A|=n-j+1}} \mu(A) = \sum_{\substack{A\subseteq N \\ |A|=n-j+1}} \sum_{k=j}^{n} w_k$. Given that weights do not depend on the particular elements of set A and there are $\binom{n}{n-j+1}$ sets of cardinality $n-j+1$ in N, then $\sum_{\substack{A\subseteq N \\ |A|=n-j+1}} \sum_{k=j}^{n} w_k = \binom{n}{n-j+1} \cdot \sum_{k=j}^{n} w_k$. Bearing these comments in mind, if μ is symmetric, then for all $j = 1,...,n$

$$
\begin{aligned}
\eta(\{m_j\}) &= 2 \cdot \left(\frac{j-1}{n-1}\right)^2 \cdot \left[\frac{\sum_{\substack{A\subseteq N \\ |A|=n-j+1}} \mu(A)}{\binom{n}{n-j+1}} - \frac{\sum_{\substack{A\subseteq N \\ |A|=n-j}} \mu(A)}{\binom{n}{n-j}} \right] \\
&= 2 \cdot \left(\frac{j-1}{n-1}\right)^2 \cdot \left[\frac{\sum_{\substack{A\subseteq N \\ |A|=n-j+1}} \sum_{k=j}^{n} w_k}{\binom{n}{n-j+1}} - \frac{\sum_{\substack{A\subseteq N \\ |A|=n-j}} \sum_{k=j+1}^{n} w_k}{\binom{n}{n-j}} \right] \\
&= 2 \cdot \left(\frac{j-1}{n-1}\right)^2 \cdot \left[\sum_{k=j}^{n} w_k - \sum_{k=j+1}^{n} w_k \right] = 2 \cdot \left(\frac{j-1}{n-1}\right)^2 \cdot w_j.
\end{aligned}
\tag{A.10}
$$

- If μ is additive, then for all $j = 1, ..., n$

$$\eta\left(\{m_j\}\right) = 2 \cdot \left(\frac{j-1}{n-1}\right)^2 \cdot \left[\frac{\sum\limits_{\substack{A \subseteq N \\ |A|=n-j+1}} \sum\limits_{m_k \in A} \mu(\{m_k\})}{\binom{n}{n-j+1}} - \frac{\sum\limits_{\substack{A \subseteq N \\ |A|=n-j}} \sum\limits_{m_k \in A} \mu(\{m_k\})}{\binom{n}{n-j}} \right] \tag{A.11}$$

$$= 2 \cdot \left(\frac{j-1}{n-1}\right)^2 \cdot \left[\sum_{k=1}^{n} \frac{\binom{n-1}{n-j}}{\binom{n}{n-j+1}} w_k - \sum_{k=1}^{n} \frac{\binom{n-1}{n-j-1}}{\binom{n}{n-j}} w_k \right]$$

$$= 2 \cdot \left(\frac{j-1}{n-1}\right)^2 \cdot \left[\sum_{k=1}^{n} \frac{n-j+1}{n} w_k - \sum_{k=1}^{n} \frac{n-j}{n} w_k \right]$$

$$= \frac{2}{n} \cdot \left(\frac{j-1}{n-1}\right)^2 \cdot \sum_{k=1}^{n} w_k.$$

A proof of the equality $Div_L\left(\mathcal{C}_\mu\right) = Div_G\left(\mathcal{C}_\mu\right) = Div\left(\vec{w}\right)$ in (5.7) is as follows. First, note that the global degree of orness of the AQWA capacity linked to a symmetric capacity is equal to,

$$\begin{aligned}
\omega_G(\eta) = & \frac{1}{(n-1)} \sum_{i=1}^{n-1} \binom{n}{i}^{-1} \sum_{\substack{A \subseteq N \\ |A|=i}} \eta(A) \\
= & \frac{1}{(n-1)} \sum_{i=1}^{n-1} \binom{n}{i}^{-1} \sum_{\substack{A \subseteq N \\ |A|=i}} \sum_{m_j \in A} 2 \cdot \left(\frac{j-1}{n-1}\right)^2 \cdot w_j \\
= & \frac{2}{(n-1)} \sum_{i=1}^{n-1} \binom{n}{i}^{-1} \sum_{j=1}^{n} \binom{n-1}{i-1} \left(\frac{j-1}{n-1}\right)^2 \cdot w_j \\
= & \frac{2}{(n-1)} \sum_{i=1}^{n-1} \frac{i}{n} \sum_{j=1}^{n} \left(\frac{j-1}{n-1}\right)^2 \cdot w_j \\
= & \frac{2}{(n-1)} \frac{n-1}{2} \sum_{j=1}^{n} \left(\frac{j-1}{n-1}\right)^2 \cdot w_j \\
= & \sum_{j=1}^{n} \left(\frac{j-1}{n-1}\right)^2 \cdot w_j.
\end{aligned} \tag{A.12}$$

If μ is symmetric, $\omega_G(\eta)$ in expression (5.5) can be replaced by (A.12) and it holds that

$$Div_G(\mathcal{C}_\mu) = \sum_{j=1}^{n} \left(\frac{j-1}{n-1}\right)^2 \cdot w_j - [2-\mu(N)] \cdot \omega_G^2(\mu). \tag{A.13}$$

Given that $\omega_G^2(\mu) = \omega_L^2(\mu)$ when μ is symmetric, expression (A.13) is equivalent to $Div_L(\mathcal{C}_\mu)$ because

$$\begin{aligned} Div_L(\mathcal{C}_\mu) = & \sum_{i=1}^{n} \left[\left(\frac{i-1}{n-1}\right)^2 - 2 \cdot \left(\frac{i-1}{n-1}\right) \cdot \omega_L(\mu) + \omega_L^2(\mu)\right] \\ & \cdot (\mu(A_{id,i}) - \mu(A_{id,i+1})) \\ = & \sum_{i=1}^{n} \left(\frac{i-1}{n-1}\right)^2 \cdot w_i - [2-\mu(N)] \cdot \omega_L^2(\mu). \end{aligned} \tag{A.14}$$

From expressions (A.13) and (A.14), it follows that $Div_L(\mathcal{C}_\mu) = Div_G(\mathcal{C}_\mu)$. Finally, when μ is a symmetric and normalized capacity such that $\mathcal{C}_\mu = OWA_{\vec{w}}$, it is easy to check that any of these two previous expressions is equal to $Div(\vec{w})$ because $\omega_G(\mu) = \omega_L(\mu) = \omega(\vec{w})$. The proof of (5.7) is finished.

In order to proof (5.8) it has to be mentioned that, when μ is additive and normalized, expression (5.5) can be simplified. Note that η is additive but not normalized. If μ is additive and normalized then, from expression (A.11), it follows that $\eta(N) = \dfrac{2n-1}{3(n-1)}$. Furthermore, it is known that

$$\begin{aligned} \omega_G(\eta) = & \frac{1}{n-1} \sum_{i=1}^{n-1} \left[\binom{n}{i}^{-1} \cdot \sum_{\substack{A \subseteq N \\ |A|=i}} \sum_{m_j \in A} \eta(\{m_j\})\right] \\ = & \frac{1}{n-1} \sum_{i=1}^{n-1} \left[\binom{n}{i}^{-1} \cdot \sum_{j=1}^{n} \binom{n-1}{i-1} \eta(\{m_j\})\right] \\ = & \frac{1}{n-1} \sum_{j=1}^{n} \sum_{i=1}^{n-1} \left(\frac{i}{n}\right) \frac{2}{n} \cdot \left(\frac{j-1}{n-1}\right)^2 \\ = & \frac{1}{n-1} \sum_{j=1}^{n} \left[\frac{n-1}{2} \cdot \frac{2}{n} \cdot \left(\frac{j-1}{n-1}\right)^2\right] = \frac{2n-1}{6(n-1)}. \end{aligned}$$

Moreover, since it is known that if μ is additive and normalized, then $\omega_G(\mu) = \dfrac{1}{2}$, from expression (5.5) we obtain (5.8):

$$Div_G(\mathcal{C}_\mu) = \frac{2n-1}{6(n-1)} - (2-1)\frac{1}{4} = \frac{1}{12} \cdot \frac{n+1}{n-1}.$$

A.7 PROOF OF PROPOSITION 5.3.1

Note that expression (5.4) can be written as

$$\begin{aligned}\mathrm{Div}(\mathcal{C}_\mu) = &\sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)^2\left[\frac{\sum_{\substack{A\subseteq N\\ |A|=n-i+1}}\mu(A)}{\binom{n}{n-i+1}} - \frac{\sum_{\substack{A\subseteq N\\ |A|=n-i}}\mu(A)}{\binom{n}{n-i}}\right]\\ &-2\omega(\mathcal{C}_\mu)\sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)\left[\frac{\sum_{\substack{A\subseteq N\\ |A|=n-i+1}}\mu(A)}{\binom{n}{n-i+1}} - \frac{\sum_{\substack{A\subseteq N\\ |A|=n-i}}\mu(A)}{\binom{n}{n-i}}\right]\\ &+\omega(\mathcal{C}_\mu)^2\mu(N)\,.\end{aligned} \tag{A.15}$$

In order to prove the proposition, it is enough to check the following

(i) $$\omega(\mathcal{C}_\eta) = \sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)^2\left[\frac{\sum_{\substack{A\subseteq N\\ |A|=n-i+1}}\mu(A)}{\binom{n}{n-i+1}} - \frac{\sum_{\substack{A\subseteq N\\ |A|=n-i}}\mu(A)}{\binom{n}{n-i}}\right];$$

and

(ii) $$\omega(\mathcal{C}_\mu) = \sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)\left[\frac{\sum_{\substack{A\subseteq N\\ |A|=n-i+1}}\mu(A)}{\binom{n}{n-i+1}} - \frac{\sum_{\substack{A\subseteq N\\ |A|=n-i}}\mu(A)}{\binom{n}{n-i}}\right],$$

because using (i) and (ii), expression (A.15) is equivalent to

$$\mathrm{Div}(\mathcal{C}_\mu) = \omega(\mathcal{C}_\eta) - [2-\mu(N)]\,\omega(\mathcal{C}_\mu)^2.$$

Item (i) is satisfied:

$$\omega(\mathcal{C}_\eta) = \frac{1}{n-1}\sum_{i=1}^{n-1}\left[\binom{n}{i}^{-1}\sum_{\substack{A\subseteq N\\|A|=i}}\eta(A)\right]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n-1}\binom{n}{i}^{-1}\sum_{\substack{A\subseteq N\\|A|=i}}\sum_{m_j\in A}\eta(\{m_j\})$$

$$= \frac{1}{n-1}\sum_{i=1}^{n-1}\binom{n}{i}^{-1}\sum_{\substack{A\subseteq N\\|A|=i}}\sum_{m_j\in A} 2\left(\frac{j-1}{n-1}\right)^2\left[\frac{\sum_{\substack{B\subseteq N\\|B|=n-j+1}}\mu(B)}{\binom{n}{n-j+1}} - \frac{\sum_{\substack{B\subseteq N\\|B|=n-j}}\mu(B)}{\binom{n}{n-j}}\right]$$

$$= \frac{2}{n-1}\sum_{i=1}^{n-1}\sum_{j=1}^{n}\left(\frac{j-1}{n-1}\right)^2\left[\frac{\binom{n-1}{i-1}\sum_{\substack{B\subseteq N\\|B|=n-j+1}}\mu(B)}{\binom{n}{i}\binom{n}{n-j+1}} - \frac{\binom{n-1}{i-1}\sum_{\substack{B\subseteq N\\|B|=n-j}}\mu(B)}{\binom{n}{i}\binom{n}{n-j}}\right]$$

$$= \frac{2}{n-1}\sum_{i=1}^{n-1}\frac{i}{n}\sum_{j=1}^{n}\left(\frac{j-1}{n-1}\right)^2\left[\frac{\sum_{\substack{B\subseteq N\\|B|=n-j+1}}\mu(B)}{\binom{n}{n-j+1}} - \frac{\sum_{\substack{B\subseteq N\\|B|=n-j}}\mu(B)}{\binom{n}{n-j}}\right]$$

$$= \sum_{j=1}^{n}\left(\frac{j-1}{n-1}\right)^2\left[\frac{\sum_{\substack{B\subseteq N\\|B|=n-j+1}}\mu(B)}{\binom{n}{n-j+1}} - \frac{\sum_{\substack{B\subseteq N\\|B|=n-j}}\mu(B)}{\binom{n}{n-j}}\right].$$

Secondly, item (ii) is also true:

$$\begin{aligned}
\omega(\mathcal{C}_\mu) &= \frac{1}{n-1}\sum_{i=1}^{n-1}\left[\binom{n}{i}^{-1}\sum_{\substack{A\subseteq N\\|A|=i}}\mu(A)\right]\\
&= \frac{1}{n-1}\sum_{j=2}^{n}\left[\binom{n}{n-j+1}^{-1}\sum_{\substack{A\subseteq N\\|A|=n-j+1}}\mu(A)\right]\\
&= \sum_{j=2}^{n}\left(\frac{j-1}{n-1}-\frac{j-2}{n-1}\right)\left[\binom{n}{n-j+1}^{-1}\sum_{\substack{A\subseteq N\\|A|=n-j+1}}\mu(A)\right]\\
&= \sum_{j=1}^{n}\left(\frac{j-1}{n-1}\right)\left[\frac{\sum_{\substack{A\subseteq N\\|A|=n-j+1}}\mu(A)}{\binom{n}{n-j+1}}-\frac{\sum_{\substack{A\subseteq N\\|A|=n-j}}\mu(A)}{\binom{n}{n-j}}\right].
\end{aligned}$$

A.8 Results for global degree of orness in the numerical example of Chapter 6

The selection criterion of optimal GlueVaR risk measures may be based on the global degree of orness. Information required in the analysis of the risk attitude implicit in optimal GlueVaR risk measures with boundary solutions is reported in Table (A.1) and Figure (??). Findings are in essence very similar than results obtained from the analysis of local degree of orness. A result that deserves some attention is that the GlueVaR with maximum global degree orness does not exactly match with the $\text{VaR}_{99.5\%}$. This result is due to the use of $n_0 = 20 < 518 = n$ for computational issues. This reduction in the size of n involves a loss in the precision of the approximation.

Table A.1: Boundary values of global degree of orness of equivalent Glue-VaR risk measures to $VaR_{99.5\%}$ for both Z (original dataset) and Z∗ (modified dataset). Parameters' values of the associated Glue-VaR measure are displayed.

	Minimum orness		Maximum orness	
	Original dataset	Modified dataset	Original dataset	Modified dataset
Degree of orness	**0.9312850**	**0.9185759**	**0.9977952**	**0.9974090**
α	90%	90%	95.54167%	95.54167%
β	99.9%	99.4875%	99.9%	99.9%
ω_1	0.1880	0.0386	0.1622	0.0154
ω_2	$1.21 \cdot 10^{-6}$	$8.12 \cdot 10^{-7}$	$1.60 \cdot 10^{-6}$	$4.09 \cdot 10^{-7}$

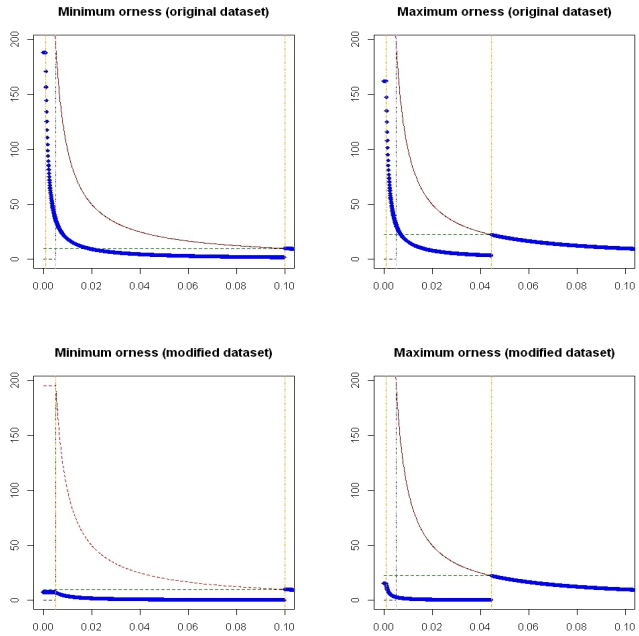


Figure A.1: Quotient functions of optimal solutions with boundary global degrees of orness in both scenarios.

BIBLIOGRAPHY

Acerbi, C. and Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503. (Cited on page 13.)

Ai, J., Brockett, P. L., Cooper, W. W., and Golden, L. L. (2012). Enterprise Risk Management Through Strategic Allocation of Capital. *Journal of Risk and Insurance*, 79(1):29–56. (Cited on page 139.)

Aitchinson, J. (1983). Principal component analysis of compositional data. *Biometrika*, 70(1):57–65. (Cited on page 122.)

Aitchinson, J. and Egozcue, J. (2005). Compositional data analysis: Where are ae and where should we be heading? *Mathematical Geology*, 37(7):829–850. (Cited on pages 121 and 139.)

Alexander, C. and Sarabia, J. M. (2012). Quantile uncertainty and Value-at-Risk model risk. *Risk Analysis*, 32(8):1293–1308. (Cited on pages 13 and 34.)

Aliev, R., Pedrycz, W., Fazlollahi, B., Huseynov, O., Alizadeh, A., and Guirimov, B. (2012). Fuzzy logic-based generalized decision theory with imperfect information. *Information Sciences*, 189:18–42. (Cited on page 39.)

Anwar, S. and Zheng, M. (2012). Competitive insurance market in the presence of ambiguity. *Insurance: Mathematics and Economics*, 50(1):79–84. (Cited on page 48.)

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228. (Cited on pages 13, 14, 27, and 108.)

Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., and Ku, H. (2007). Coherent multiperiod risk adjusted values and Bellman's principle. *Annals of Operations Research*, 152:5–22. (Cited on page 139.)

Asimit, A. V., Furman, E., Tang, Q., and Vernic, R. (2011). Asymptotics for risk capital allocations based on Conditional Tail Expectation. *Insurance: Mathematics and Economics*, 49(3):310–324. (Cited on page 95.)

Aven, T. (2012). Foundational issues in risk assessment and risk management. *Risk Analysis*, 32(10):1647–1656. (Cited on page 16.)

Aven, T. (2013). On the meaning and use of the risk appetite concept. *Risk Analysis*, 33(3):462–468. (Cited on page 29.)

Balbás, A., Garrido, J., and Mayoral, S. (2009). Properties of distortion risk measures. *Methodology and Computing in Applied Probability*, 11(3, SI):385–399. (Cited on pages 14, 15, and 27.)

Balog, D. (2010). Risk based capital allocation. In *Proceedings of FIKUSZ '10 Symposium for Young Researchers*, pages 17–26. (Cited on page 106.)

Beliakov, G., Pradera, A., and Calvo, T. (2007). *Aggregation Functions: A Guide to Practitioners*. Springer, Berlin. (Cited on pages 42 and 59.)

Belles-Sampera, J. (2011). Capital allocation and distortion risk measures. Master's thesis, Department of Econometrics - University of Barcelona. (Cited on page 25.)

Belles-Sampera, J., Guillén, M., and Santolino, M. (2013a). Generalizing some usual risk measures in financial and insurance applications. In Fernández-Izquierdo, M., Muñoz-Torres, M., and León, R., editors, *Modeling and Simulation in Engineering, Economics and Management. Proceedings of the MS 2013 International Conference*, volume 145 of *Lecture Notes in Business Information Processing*, pages 75–82. Springer-Verlag. (Cited on page xv.)

Belles-Sampera, J., Guillén, M., and Santolino, M. (2014a). Beyond Value-at-Risk: GlueVaR distortion risk measures. *Risk Analysis*, 34(1):121–134. (Cited on page xv.)

Belles-Sampera, J., Guillén, M., and Santolino, M. (2014b). GlueVaR risk measures in capital allocation applications. *Insurance: Mathematics and Economics*, 58:132–137. (Cited on page xvi.)

Belles-Sampera, J., Guillén, M., and Santolino, M. (2014c). The use of flexible quantile-based measures in risk assessment. *Communication in Statistics – Theory and Methods*, (Accepted). (Cited on page xv.)

Belles-Sampera, J., Merigó, J. M., Guillén, M., and Santolino, M. (2013b). The connection between distortion risk measures and ordered weighted averaging operators. *Insurance: Mathematics and Economics*, 52(2):411–420. (Cited on pages xv and 66.)

Belles-Sampera, J., Merigó, J. M., Guillén, M., and Santolino, M. (2014d). Indicators for the characterization of discrete Choquet integrals. *Information Sciences*, 267:201–216. (Cited on page xv.)

Belles-Sampera, J., Merigó, J. M., and Santolino, M. (2013c). Indicators for the characterization of discrete Choquet integrals. IREA Working Papers 201311, University of Barcelona, Research Institute of Applied Economics. (Cited on pages 72 and 74.)

Belles-Sampera, J., Merigó, J. M., and Santolino, M. (2013d). Some new definitions of indicators for the Choquet integral. In H. Bustince, J. Fernández, T. Calvo and R. Mesiar, editor, *Aggregation Functions in Theory and Practice. Proceedings of the 7th International Summer School on Aggregation Operators*, volume 228 of *Advances in Intelligent Systems and Soft Computing*, pages 467–476. Springer-Verlag. (Cited on pages xv, 69, and 74.)

Belles-Sampera, J. and Santolino, M. (2013a). Algunas reflexiones sobre los problemas de asignación de capital y la aplicación de ciertas medidas de riesgo. In Gómez Déniz, E., Guillén Estany, M., and Vázquez Polo, F., editors, *Investigaciones en Seguros y Gestión del Riesgo: Riesgo 2013*, volume 194 of *Cuadernos de la Fundación*, pages 161–176. Fundación MAPFRE. (Cited on page xvi.)

Belles-Sampera, J. and Santolino, M. (2013b). Asignación óptima de capital en base al perfil de riesgo de las instituciones de inversión colectiva: una aplicación de las medidas de riesgo distorsionadas. *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 15(2):65–86. (Cited on page xvi.)

Bellini, F. and Gianin, E. R. (2012). Haezendonck-Goovaerts risk measures and Orlicz quantiles. *Insurance: Mathematics and Economics*, 51(1):107–114. (Cited on pages 15 and 24.)

Bellini, F., Klar, B., Mueller, A., and Gianin, E. R. (2014). Generalized quantiles as risk measures. *Insurance: Mathematics and Economics*, 54:41–48. (Cited on pages 15 and 24.)

Bleichrodt, H. and Eeckhoudt, L. (2006). Survival risks, intertemporal consumption, and insurance: The case of distorted probabilities. *Insurance: Mathematics and Economics*, 38(2):335–346. (Cited on page 15.)

Bolancé, C., Guillén, M., Pelican, E., and Vernic, R. (2008). Skewed bivariate models and nonparametric estimation for the CTE risk measure. *Insurance: Mathematics and Economics*, 43(3):386–393. (Cited on page 31.)

Bolton, J., Gader, P., and Wilson, J. N. (2008). Discrete Choquet integral as a distance metric. *IEEE Transactions on Fuzzy Systems*, 16(4):1107–1110. (Cited on page 59.)

Buch, A., Dorfleitner, G., and Wimmer, M. (2011). Risk capital allocation for RORAC optimization. *Journal of Banking and Finance*, 35(11):3001–3009. (Cited on page 95.)

Cai, J. and Wei, W. (2014). Some new notions of dependence with applications in optimal allocations problems. *Insurance: Mathematics and Economics*, 55:200–209. (Cited on page 95.)

Calvo, T., Mesiar, R., and Yager, R. R. (2004). Quantitative weights and aggregation. *IEEE Transactions on Fuzzy Systems*, 12(1):62–69. (Cited on page 139.)

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., and Montrucchio, L. (2011). Risk measures: Rationality and diversification. *Mathematical Finance*, 21(4):743–774. (Cited on page 27.)

Chen, D., Mao, T., Pan, X., and Hu, T. (2012). Extreme value behavior of aggregate dependent risks. *Insurance: Mathematics and Economics*, 50(1):99–108. (Cited on page 28.)

Cheung, K. C. (2009). Upper comonotonicity. *Insurance: Mathematics and Economics*, 45(1):35 – 40. (Cited on page 28.)

Choquet, G. (1954). Theory of Capacities. *Annales de l'Institute Fourier*, 5:131–295. (Cited on pages 15, 40, and 59.)

Cornish, E. A. and Fisher, R. A. (1937). Moments and cumulants in the specification of distributions. *Revue de l'Institut International de Statistique*, 4:307–320. (Cited on page 23.)

Cossette, H., Côté, M., Marceau, E., and Moutanabbir, K. (2013). Multivariate distribution defined with Farlie-Gumbel-Morgenstern copula and mixed Erland marginals: Aggregation and capital allocation. *Insurance: Mathematics and Economics*, 52:560–572. (Cited on page 95.)

Cossette, H., Mailhot, M., and Marceau, E. (2012). TVaR-based capital allocation for multivariate compound distributions with positive continuous claim amounts. *Insurance: Mathematics and Economics*, 50:247–256. (Cited on page 95.)

Cox, L. A. T. (2012). Confronting deep uncertainties in risk analysis. *Risk Analysis*, 32(10):1607–1629. (Cited on page 27.)

Csiszár, I. (1995). Generalized cutoff rates and Rényi's information measures. *IEEE Transactions on Information Theory*, 41(1):26–34. (Cited on page 64.)

Danielsson, J., Jorgensen, B. J., Sarma, M., and de Vries, C. G. (2005). Subadditivity re-examined: the case for Value-at-Risk. Technical report, CiteSeerx. (Cited on page 28.)

De Baets, B. (2013). Aggregation 2.0. Opening plenary session of the AGOP 2013 conference, Pamplona, Spain. (Cited on pages 119 and 123.)

De Waegenaere, A., Kast, R., and Lapied, A. (2003). Choquet pricing and equilibrium. *Insurance: Mathematics and Economics*, 32(3):359–370. (Cited on page 48.)

Degen, M., Lambrigger, D. D., and Segers, J. (2010). Risk concentration and diversification: Second-order properties. *Insurance: Mathematics and Economics*, 46(3):541–546. (Cited on page 27.)

Denault, M. (2001). Coherent allocation of risk capital. *Journal of Risk*, 4(1):1–34. (Cited on pages 95 and 99.)

Denneberg, D. (1994). *Non-Additive Measure and Integral*. Kluwer Academic Publishers, Dordrecht. (Cited on pages 16, 27, 40, 144, and 145.)

Denuit, M., Dhaene, J., Goovaerts, M., and Kaas, R. (2005). *Actuarial Theory for Dependent Risks. Measures, Orders and Models*. John Wiley & Sons Ltd, Chichester. (Cited on pages 15 and 16.)

Denuit, M., Dhaene, J., Goovaerts, M., Kaas, R., and Laeven, R. (2006). Risk measurement with equivalent utility principles. *Statistics & Decisions*, 24(1):1–25. (Cited on pages 15 and 27.)

Dhaene, J., Kukush, A., Linders, D., and Tang, Q. (2012a). Remarks on quantiles and distortion risk measures. *European Actuarial Journal*, 2(2):319–328. (Cited on pages 15 and 24.)

Dhaene, J., Laeven, R. J. A., Vanduffel, S., Darkiewicz, G., and Goovaerts, M. J. (2008). Can a coherent risk measure be too subadditive? *Journal of Risk and Insurance*, 75(2):365–386. (Cited on page 27.)

Dhaene, J., Tsanakas, A., Valdez, E. A., and Vanduffel, S. (2012b). Optimal capital allocation principles. *Journal of Risk and Insurance*, 79(1):1–28. (Cited on pages 7, 95, 98, 109, 110, 111, 112, 118, 138, and 139.)

Diers, D. (2011). Management strategies in multi-year Enterprise Risk Management. *Geneva Papers on Risk and Insurance - Issues and Practice*, 36(1):107–125. (Cited on page 139.)

Dujmović, J. J. (2006). A comparison of andness/orness indicators. In *Proceedings of the 11th Information Processing and Management of Uncertainty international (IPMU 2006)*, pages 691–698. (Cited on page 64.)

Dukhovny, A. (2002). General entropy of general measures. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(3):213–225. (Cited on page 66.)

Ekeland, I., Galichon, A., and Henry, M. (2012). Comonotonic measures of multivariate risks. *Mathematical Finance*, 22(1):109–132. (Cited on page 27.)

Embrechts, P., Lambrigger, D. D., and Wuethrich, M. V. (2009a). Multivariate extremes and the aggregation of dependent risks: examples and counter-examples. *Extremes*, 12(2):107–127. (Cited on page 27.)

Embrechts, P., Neslehova, J., and Wuethrich, M. V. (2009b). Additivity properties for Value-at-Risk under Archimedean dependence and heavy-tailedness. *Insurance: Mathematics and Economics*, 44(2, SI):164–169. (Cited on page 27.)

Engemann, K. J., Miller, H. E., and Yager, R. R. (1996). Decision making with belief structures: An application in risk management. *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems*, 4(1):1–25. (Cited on page 39.)

Fernández Salido, J. and Murakami, S. (2003). Extending Yager's orness concept for the OWA aggregators to other mean operators. *Fuzzy Sets and Systems*, 139(3):515–542. (Cited on page 80.)

Fisher, R. A. and Cornish, E. A. (1960). The percentile points of distributions having known cumulants. *Technometrics*, 2(2):209–225. (Cited on page 23.)

Föllmer, H. and Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447. (Cited on page 27.)

Frittelli, M. and Rosazza Gianin, E. (2002). Putting order in risk measures. *Journal of Banking & Finance*, 26(7):1473–1486. (Cited on page 27.)

Froot, K. A. (2007). Risk management, capital budgeting, and capital structure policy for insurers and reinsurers. *Journal of Risk and Insurance*, 74(2):273–299. (Cited on page 139.)

Fullér, R. and Majlender, P. (2003). On obtaining minimal variability OWA operator weights. *Fuzzy Sets and Systems*, 136(2):203–215. (Cited on pages 59, 61, and 63.)

Furman, E. and Landsman, Z. (2006). Tail variance premium with applications for elliptical portfolio of risks. *ASTIN Bulletin*, 36(2):433–462. (Cited on page 105.)

Giamouridis, D. (2006). Estimation risk in financial risk management: A correction. *Journal of Risk*, 8(4):121–125. (Cited on page 23.)

Gil-Lafuente, A. M. (2005). *Fuzzy Logic in Financial Analysis*. Springer, Berlin. (Cited on page 39.)

Goovaerts, M., Linders, D., Van Weert, K., and Tank, F. (2012). On the interplay between distortion, mean value and Haezendonck-Goovaerts risk measures. *Insurance: Mathematics and Economics*, 51(1):10–18. (Cited on pages 14, 15, 24, and 27.)

Goovaerts, M. J., Kaas, R., and Laeven, R. J. (2010a). Decision principles derived from risk measures. *Insurance: Mathematics and Economics*, 47(3):294–302. (Cited on page 39.)

Goovaerts, M. J., Kaas, R., and Laeven, R. J. (2010b). A note on additive risk measures in rank-dependent utility. *Insurance: Mathematics and Economics*, 47(2):187–189. (Cited on page 15.)

Grabisch, M. (1995). On equivalence classes of fuzzy-connectives: The case of fuzzy integrals . *IEEE Transactions on Fuzzy Systems*, 3(1):96–109. (Cited on page 60.)

Grabisch, M., Marichal, J.-L., Mesiar, R., and Pap, E. (2009). *Aggregation functions*, volume 127 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press. (Cited on pages 42, 59, 60, 66, and 136.)

Grabisch, M., Marichal, J.-L., Mesiar, R., and Pap, E. (2011). Aggregation functions: Means. *Information Sciences*, 181(1):1–22. (Cited on pages 40, 41, 42, 46, 60, and 73.)

Grechuk, B., Molyboha, A., and Zabarankin, M. (2012). Mean-deviation analysis in the theory of choice. *Risk Analysis*, 32(8):1277–1292. (Cited on page 27.)

Greco, S., Matarazzo, B., and Giove, S. (2011). The Choquet integral with respect to a level dependent capacity. *Fuzzy Sets and Systems*, 175(1):1–35. (Cited on pages 47 and 59.)

Grégoire, P. (2007). *Advanced Portfolio Attribution Analysis: New Approaches to Return and Risk*, chapter 12 - Risk Attribution, pages 309–329. Riskbooks, Incisive Media. (Cited on page 96.)

Guillén, M., Prieto, F., and Sarabia, J. M. (2011). Modelling losses and locating the tail with the Pareto Positive Stable distribution. *Insurance: Mathematics and Economics*, 49(3):454–461. (Cited on page 31.)

Hamada, M., Sherris, M., and Van Der Hoek, J. (2006). Dynamic portfolio allocation, the dual theory of choice and probability distortion functions. *ASTIN Bulletin*, 36(1):187–217. (Cited on page 139.)

Handa, J. (1977). Risk, probabilities, and a new theory of cardinal utility. *Journal of Political Economy*, 85(1):97–122. (Cited on page 39.)

Honda, A. and Okazaki, Y. (2005). Identification of fuzzy measures with distorted probability measures. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 9(5):467–476. (Cited on page 48.)

Hosking, J. and Wallis, J. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, 29(3):339–349. (Cited on page 21.)

Hürlimann, W. (2006). A note on generalized distortion risk measures. *Finance Reseach Letters*, 3(4):267–272. (Cited on page 14.)

Hua, L. and Joe, H. (2012). Tail comonotonicity: Properties, constructions, and asymptotic additivity of risk measures. *Insurance: Mathematics and Economics*, 51(2):492–503. (Cited on page 28.)

Jiang, L. (2008). Convexity, translation invariance and subadditivity for g-expectations and related risk measures. *Annals of Applied Probability*, 18(1):245–258. (Cited on page 27.)

Jizba, P. and Arimitsu, T. (2004). The world according to Rényi: Thermodynamics of multifractal systems. *Annals of Physics*, 312(1):17–59. (Cited on page 64.)

Johnson, N. L. and Kotz, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions - 1*. John Wiley & Sons, New York. (Cited on page 23.)

Kahneman, D. and Tversky, A. (1979). Prospect theory - Analysis of decision under risk. *Econometrica*, 47(2):263–291. (Cited on page 39.)

Kalkbrener, M. (2005). An axiomatic approach to capital allocation. *Mathematical Finance*, 15(3):425–437. (Cited on page 95.)

Kaluszka, M. and Krzeszowiec, M. (2012). Pricing insurance contracts under cumulative prospect theory. *Insurance: Mathematics and Economics*, 50(1):159–166. (Cited on page 15.)

Kim, J. H. T. (2010). Bias correction for estimated distortion risk measure using the bootstrap. *Insurance: Mathematics and Economics*, 47(2):198–205. (Cited on page 52.)

Klement, E. P., Mesiar, R., and Pap, E. (2010). A universal integral as common frame for Choquet and Sugeno integral. *IEEE Transactions on Fuzzy Systems*, 18(1):178–187. (Cited on page 59.)

Kojadinovic, I., Marichal, J.-L., and Roubens, M. (2005). An axiomatic approach to the definition of the entropy of a discrete Choquet capacity. *Information Sciences*, 172(1-2):131–153. (Cited on pages 60 and 66.)

Kolesárová, A. and Mesiar, R. (2009). Parametric characterization of aggregation functions. *Fuzzy Sets and Systems*, 160(6):816–831. (Cited on page 64.)

Lenzi, E. K., Mendes, R. S., and da Silva, L. R. (2000). Statistical mechanics based on Rényi entropy. *Physica A*, 280(3-4):337–345. (Cited on page 64.)

Li, X. and You, Y. (2015). Permutation monotone functions of random vectors with applications in financial and actuarial risk management. *Advances in Applied Probability*, 47(1):270–291. (Cited on page 96.)

Lv, W., Pan, X., and Hu, T. (2013). Asymptotics of the risk concentration based on the tail distortion risk measure. *Statistics and Probability Letters*, 83(12):2703–2710. (Cited on page 24.)

Majlender, P. (2005). OWA operators with maximal Rényi entropy. *Fuzzy Sets and Systems*, 155(3):340–360. (Cited on pages 59, 61, and 63.)

Marichal, J.-L. (2002). Entropy of discrete Choquet capacities. *European Journal of Operational Research*, 137(3):612–624. (Cited on pages 60 and 66.)

Marichal, J.-L. (2004). Tolerant or intolerant character of interacting criteria in aggregation by the Choquet integral. *European Journal of Operational Research*, 155(3):771–791. (Cited on pages 60, 65, and 66.)

Marichal, J.-L. and Roubens, M. (2000). Entropy of discrete fuzzy measures. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 8(6):625–640. (Cited on page 60.)

Marques Pereira, R. A. and Ribeiro, R. A. (2003). Aggregation with generalized mixture operators using weighting functions. *Fuzzy Sets and Systems*, 137(1):43–58. (Cited on page 139.)

McCune, E. and Gray, H. (1982). Cornish-Fisher and Edgeworth expansions. In Samuel Kotz and Norman L. Johnson, editor, *Encyclopedia of Statistical Sciences*, volume 2. John Wiley & Sons, New York. (Cited on page 23.)

McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management*. Princeton Series in Finance. Princeton University Press, New York. (Cited on pages 13 and 99.)

Merigó, J. M. (2011). Fuzzy multi-person decision making with fuzzy probabilistic aggregation operators. *International Journal of Fuzzy Systems*, 13(3):163–174. (Cited on pages 60 and 72.)

Merigó, J. M. (2012). Probabilities in the OWA operator. *Expert Systems With Applications* , 39(13):11456–11467. (Cited on pages 60, 72, and 74.)

Merigó, J. M. and Casanovas, M. (2011). The uncertain induced quasi-arithmetic OWA operator. *International Journal of Intelligent Systems*, 26(1):1–24. (Cited on page 39.)

Merigó, J. M. and Gil-Lafuente, A. M. (2009). The induced generalized OWA operator. *Information Sciences*, 179(6):729–741. (Cited on page 43.)

Merigó, J. M. and Wei, G. (2011). Probabilistic aggregation operators and their application in uncertain multi-person decision-making. *Technological and Economic Development of Economy*, 17(2):335–351. (Cited on page 72.)

Mesiar, R., Mesiarová-Zemánková, A., and Ahmad, K. (2011). Discrete Choquet integral and some of its symmetric extensions. *Fuzzy Sets and Systems*, 184(1):148–155. (Cited on page 47.)

Mesiar, R. and Spirkova, J. (2006). Weighted means and weighting functions. *Kybernetika*, 42(2):151–160. (Cited on page 139.)

Mesiar, R., Spirkova, J., and Vavrikova, L. (2008). Weighted aggregation operators based on minimization. *Information Sciences*, 178(4):1133–1140. (Cited on page 139.)

Nam, H. S., Tang, Q., and Yang, F. (2011). Characterization of upper comonotonicity via tail convex order. *Insurance: Mathematics and Economics*, 48(3):368–373. (Cited on page 27.)

Overbeck, L. (2000). *Measuring Risk in Complex Systems*, chapter Allocation of Economic Capital in Loan Portfolios. Springer. (Cited on page 104.)

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behaviour & Organization*, 3(4):323–343. (Cited on page 39.)

Rahl, L., editor (2012). *Risk Budgeting: Risk Appetite and Governance in the Wake of the Financial Crisis*. Riskbooks, Incisive Media, 2nd edition. (Cited on page 96.)

Rényi, A. (1961). On measures of entropy and information. *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, pages 547–561. (Cited on pages 59 and 63.)

Riedel, F. (2004). Dynamic coherent risk measures. *Stochastic Processes and their Applications*, 112(2):185–200. (Cited on page 139.)

Sandström, A. (2007). Solvency II: Calibration for skewness. *Scandinavian Actuarial Journal*, (2):126–134. (Cited on page 23.)

Sandström, A. (2011). *Handbook of Solvency for Actuaries and Risk Managers. Theory and Practice*. CRC Finance Series. Chapman & Hall / CRC , Boca Raton. (Cited on pages 21 and 23.)

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587. (Cited on page 39.)

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423. (Cited on pages 43, 60, and 62.)

Shapiro, A. F. (2002). The merging of neural networks, fuzzy logic, and genetic algorithms. *Insurance: Mathematics and Economics*, 31(1):115–131. (Cited on page 39.)

Shapiro, A. F. (2004). Fuzzy logic in insurance. *Insurance: Mathematics and Economics*, 35(2):399–424. (Cited on page 39.)

Shapiro, A. F. (2009). Fuzzy random variables. *Insurance: Mathematics and Economics*, 44(2):307–314. (Cited on page 39.)

Song, Y. and Yan, J.-A. (2009). Risk measures with comonotonic subadditivity or convexity and respecting stochastic orders. *Insurance: Mathematics and Economics*, 45(3):459–465. (Cited on pages 27 and 28.)

Sordo, M. A. and Suarez-Llorens, A. (2011). Stochastic comparisons of distorted variability measures. *Insurance: Mathematics and Economics*, 49(1):11–17. (Cited on page 48.)

Sung, K., Yam, S., Yung, S., and Zhou, J. (2011). Behavioral optimal insurance. *Insurance: Mathematics and Economics*, 49(3):418–428. (Cited on page 48.)

Szëgo, G. (2002). Measures of risk. *Journal of Banking and Finance*, 26(7):1253–1272. (Cited on page 15.)

Tasche, D. (1999). Risk contributions and performance measurement. Working paper, Lehrstuhl für Mathematische Statistik, TU München. (Cited on page 99.)

Tasche, D. (2004). *Economic Capital: A Practitioner's Guide*, chapter 13 - Allocating portfolio economic capital to sub-portfolios, pages 275–302. Risk Books, IncisiveMedia. (Cited on page 99.)

Tasche, D. (2007). Euler allocation: Theory and practice. Working paper. (Cited on pages 99, 101, and 103.)

Torra, V. (1997). The weighted OWA operator. *International Journal of Intelligent Systems*, 12(2):153–166. (Cited on pages 40, 43, and 74.)

Torra, V. (1998). On some relationships between the WOWA operator and the Choquet integral. In *Proceedings of the IPMU 1998 Conference, Paris, France*, pages 818–824. (Cited on pages 48 and 74.)

Torra, V. and Narukawa, Y. (2007). *Modeling Decisions: Information Fusion and Aggregation Operators*. Springer, Berlin. (Cited on pages 42, 59, and 80.)

Torra, V. and Narukawa, Y. (2010). Some relationships between Losonczi's based OWA generalizations and the Choquet-Stieltjes integral. *Soft Computing*, 14(5):465–472. (Cited on page 59.)

Torra, V. and Narukawa, Y. (2012). On a comparison between Mahalanobis distance and Choquet integral: The Choquet-Mahalanobis operator. *Information Sciences*, 190(1):56–63. (Cited on page 59.)

Tsanakas, A. (2004). Dynamic capital allocation with distortion risk measures. *Insurance: Mathematics and Economics*, 35(2):223–243. (Cited on page 139.)

Tsanakas, A. (2009). To split or not to split: Capital allocation with convex risk measures. *Insurance: Mathematics and Economics*, 44(2):268–277. (Cited on page 95.)

Tsanakas, A. and Desli, E. (2005). Measurement and pricing of risk in insurance markets. *Risk Analysis*, 25(6):1653–1668. (Cited on page 15.)

Tsanakas, A. and Millossovich, P. (2014). Sensitivity analysis using risk measures. Working paper, Cass Business School, City University London. (Cited on page 95.)

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323. (Cited on page 39.)

Urbina, J. and Guillén, M. (2014). An application of capital allocation principles to operational risk and the cost of fraud. *Expert Systems with Applications*, 41(16):7023–7031. (Cited on page 95.)

van Gulick, G., De Waegenaere, A., and Norde, H. (2012). Excess based allocation of risk capital. *Insurance: Mathematics and Economics*, 50(1):26–42. (Cited on pages 95, 107, 108, 127, 128, and 129.)

von Neumann, J. and Morgenstern, O. (1947). *Theory of Games and Economic Behaviour*. Princeton University Press, Princeton, NJ. (Cited on page 39.)

Wang, M. (2014). Capital allocation based on the tail covariance premium adjusted. *Insurance: Mathematics and Economics*, 57:125–131. (Cited on pages 95 and 105.)

Wang, S. S. (1995). Insurance pricing and increased limits ratemaking by proportional hazard transforms. *Insurance: Mathematics and Economics*, 17(1):43–54. (Cited on page 15.)

Wang, S. S. (1996). Premium calculation by transforming the layer premium density. *ASTIN Bullletin*, 26(1):71–92. (Cited on pages 14, 15, 40, 46, and 48.)

Wang, S. S. (1998). An actuarial index of the right-tail risk. *North American Actuarial Journal*, 2(2):88–101. (Cited on page 27.)

Wang, S. S. (2002). A risk measure that goes beyond coherence. In *Proceedings of the 2002 AFIR (Actuarial approach to financial risks)*. (Cited on page 55.)

Wang, S. S. and Dhaene, J. (1998). Comonotonicity, correlation order and premium principles. *Insurance: Mathematics and Economics*, 22(3):235–242. (Cited on pages 14 and 27.)

Wang, S. S., Young, V. R., and Panjer, H. H. (1997). Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics*, 21(2):173–183. (Cited on page 14.)

Weber, S. (2006). Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance*, 16(2):419–441. (Cited on page 139.)

Wirch, J. L. and Hardy, M. R. (2002). Distortion risk measures: Coherence and stochastic dominance. IME Conference, Lisbon. (Cited on page 27.)

Wu, X. Y. and Zhou, X. (2006). A new characterization of distortion premiums via countable additivity for comonotonic risks. *Insurance: Mathematics and Economics*, 38(2):324–334. (Cited on page 14.)

Xu, M. and Hu, T. (2012). Stochastic comparisons of capital allocations with applications. *Insurance: Mathematics and Economics*, 50(3):293–298. (Cited on pages 95 and 118.)

Xu, M. and Mao, T. (2013). Optimal capital allocation based on the Tail Mean-Variance model. *Insurance: Mathematics and Economics*, 53:533–543. (Cited on page 95.)

Xu, Z. (2010). Choquet integrals of weighted intuitionistic fuzzy information. *Information Sciences*, 180(5):726–736. (Cited on page 59.)

Xu, Z. S. and Da, Q. L. (2002). The uncertain OWA operator. *International Journal of Intelligent Systems*, 17(6):569–575. (Cited on page 43.)

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, 55(1):95–115. (Cited on pages 15 and 39.)

Yager, R. R. (1988). On ordered weighted averaging operators in multicriteria decision-making. *IEEE Transactions on Systems, Man and Cybernetics*, 18(1):183–190. (Cited on pages 40, 42, 43, 59, 61, and 62.)

Yager, R. R. (1993). Families of OWA operators. *Fuzzy Sets and Systems*, 59(2):125–148. (Cited on page 42.)

Yager, R. R. (1996). Constrained OWA aggregation. *Fuzzy Sets and Systems*, 81(1):89–101. (Cited on pages 59, 61, 62, and 67.)

Yager, R. R. (2000). On the entropy of fuzzy measures. *IEEE Transactions on Fuzzy Systems*, 8(4):453–461. (Cited on pages 60 and 66.)

Yager, R. R. (2002). Heavy OWA operators. *Fuzzy Optimization and Decision Making*, 1:379–397. (Cited on pages 43, 59, 61, 62, and 63.)

Yager, R. R. (2004a). Choquet aggregation using order inducing variables. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 12(1):69–88. (Cited on page 59.)

Yager, R. R. (2004b). OWA aggregation over a continuous interval argument with applications to decision making. *IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics*, 34(5):1952–1963. (Cited on page 136.)

Yager, R. R. (2008). Time series smoothing and OWA aggregation. *IEEE Transactions on Fuzzy Systems*, 16(4):994–1007. (Cited on page 39.)

Yager, R. R. (2010). Norms induced from OWA operators. *IEEE Transactions on Fuzzy Systems*, 18(1):57–66. (Cited on page 43.)

Yager, R. R. and Filev, D. P. (1999). Induced ordered weighted averaging operators. *IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics*, 29(2):141–150. (Cited on page 39.)

Yager, R. R., Kacprzyk, J., and Beliakov, G. (2011). *Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice*. Springer, Berlin. (Cited on pages 42 and 59.)

Yager, R. R. and Xu, Z. (2006). The continuous ordered weighted geometric operator and its application to decision making. *Fuzzy Sets and Systems*, 157:1393–1402. (Cited on page 136.)

You, Y. and Li, X. (2014). Optimal capital allocations to interdependent actuarial risks. *Insurance: Mathematics and Economics*, 57:104–113. (Cited on pages 95 and 118.)

Zadeh, L. A. (1985). Syllogistic reasoning in fuzzy-logic and its application to usuality and reasoning with dispositions. *IEEE Transactions on Systems, Man and Cybernetics*, 15(6):754–763. (Cited on page 43.)

Zaks, Y. and Tsanakas, A. (2014). Optimal capital allocation in a hierarchical corporate structure. *Insurance: Mathematics and Economics*, 56:48–55. (Cited on pages 95 and 118.)

Zhu, L. and Li, H. (2012). Tail distortion risk and its asymptotic analysis. *Insurance: Mathematics and Economics*, 51(1):115–121. (Cited on pages 15, 24, and 25.)

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

http://code.google.com/p/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured here:

http://postcards.miede.de/

DECLARATION

I herewith declare that I have produced this work without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This work has not previously been presented in identical or similar form to any examination board.

The dissertation work was conducted from October 2011 to April 2015 under the supervision of Dr. Miguel Ángel Santolino Prieto and Dr. Montserrat Guillén Estany at the Universitat de Barcelona.

DECLARACIÓ

Declaro que he dut a terme aquesta tesi sense l'assistència prohibida de terceres persones i sense fer ús d'altres ajudes que les especificades. Les nocions preses de forma directa o indirecta d'altres fonts han estat identificades com a tals. Aquest treball no ha estat presentat prèviament de manera idèntica o similar.

El treball de tesi es va dur a terme entre l'octubre de 2011 i l'abril de 2015 sota la supervisió del Dr. Miguel Ángel Santolino Prieto i de la Dra. Montserrat Guillén Estany a la Universitat de Barcelona.

DECLARACIÓN

Por la presente declaro que he realizado esta tesis sin la asistencia prohibida de terceros y sin hacer uso de otras ayudas a las especificadas. Las nociones tomadas de forma directa o indirecta de otras fuentes han sido identificadas como tales. Este trabajo no ha sido previamente presentado en forma idéntica o similar.

El trabajo de tesis se llevó a cabo entre octubre de 2011 y abril de 2015 bajo la supervisión del Dr. Miguel Ángel Santolino Prieto y de la Dra. Montserrat Guillén Estany en la Universitat de Barcelona.

Barcelona, May 2015

Jaume Belles Sampera