Game-theoretic approaches to optimal risk sharing

Game-theoretic approaches to optimal risk sharing

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit van Tilburg op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op woensdag 15 januari 2014 om 16.15 uur door

TIM JAIJ BOONEN

geboren op 17 juni 1987 te Amsterdam

Promotiecommissie:

Promotores: prof. dr. Anja De Waegenaere
prof. dr. Henk Norde

Overige leden: dr. Enrico Biffis
prof. dr. Hans Schumacher
prof. dr. Dolf Talman
prof. dr. ir. Michel Vellekoop

Acknowledgements

With this dissertation, I conclude three years of work, stress and, above all, a pleasant time at Tilburg University. I started this Ph.D. project more than three years ago, not having an idea how to start this trajectory. Now looking back, I had very productive days in the first months. Quickly, we had an idea how to proceed with my both master theses. At that time, I didn't realize how fortunate I was being able to start proper. By now, I know that finding a proper research question is not as easy as I thought.

This is, among many others, one of the reasons I want to emphasize that I am grateful with both my advisors, Anja De Waegenaere and Henk Norde. They both guided me through writing my master thesis. I was always stimulated to continue into academia. Moreover, they helped me through the research master phase, where they also supervised my second master thesis. In this way, I was fully prepared to start writing my dissertation. Also, they very actively helped me finding well-specified research topics.

I also want to thank Bas Werker for his help with the application for my Ph.D. position. Not knowing me that well at that time, he was dedicated to help me with this. I am very grateful to Netspar for funding me to write this dissertation. Also, I would like to thank Peter Borm for his ideas about research topics. Moreover, he first introduced me as a master's student to my advisors.

I want to thank my committee members for their thorough feedback. The comments of Enrico Biffis, Hans Schumacher, Dolf Talman and Michel Vellekoop were very useful. They gave me scope to think about for further projects. I especially want to thank Enrico Biffis. Not only for flying over to Tilburg for my defense, but also for hosting me at Imperial College London in June, 2012. We had very fruitful discussions, also via phone, and I am confident those talks will lead to a good project in the future. Also, I want to thank Michel Vellekoop for the useful discussions about my dissertation; not only during the defense, but also for the feedback during my application at the University of Amsterdam. I hope we can work together in the next years in Amsterdam.

This dissertation strongly benefitted from many discussions I had and presentations I did in the last three years. This would not have been possible without the financial help of Netspar, the department of Econometrics and Operations Research and the NWO grant of Anja De

Waegenaere. I was able to travel to conferences in Amsterdam (three times), Turin, Paris (twice), San Diego (twice), Frankfurt (twice), Barcelona, Budapest, Istanbul, and Minneapolis. I am also grateful to my hosts at the University of New South Wales, John Piggott and Michael Sherris; my stay of three months was very pleasant.

During the last three year, Jochem was my office mate. It was a pleasant time, not in the last place due to our many (informal) talks. Together with a variety of “members” of our lunch group, we always had casual and lengthy talk of the day.

Despite my friends from academia, I also spent some some time with my friends outside academia. I am fortunate that during the full phase of writing this dissertation, my football team Ajax became national champion. This was three years ago not as self-evident as it nowadays may sound. I spent a great time with my friends from afcalibra. I also spent a lot of my spare time with my former “study” mates and friends from Maarheeze.

Last, but not least, I’d like to thank my parents Karel and Bertie, my sister Cris, and Elske for their support in writing this dissertation. Without you, I probably would not have even started this dissertation.

Tim Boonen
Amsterdam, October 2013

Contents

Chapter 1

Introduction

This dissertation addresses topics in optimal risk sharing and capital allocation. Risk sharing and capital allocation are both important for the risk management of a firm. Risk management is the practice of optimizing the exposure to risk within a firm to increase its economic value by for example using financial instruments. Examples of risks are credit risk, inflation risk, liquidity risk and longevity risk. Risk management is important for financial and insurance companies and pension funds. This is to protect the risk for its shareholders and customers by preventing moral hazard of the board of a firm so that it is not overexposed to risk. For insurers, the regulation will be given by Solvency II, for which the content is more or less known. Pension funds in the Netherlands are regulated by the Financial Assessment Framework (FTK) of the Dutch central bank (DNB) and the Netherlands Authority for the Financial Markets (AFM). The regulatory requirements prescribe the buffer that a firm should withhold in order to cope with future downside shocks. It serves as a tool to prevent bankruptcy of a firm. Holding this buffer is costly for the firm, as it needs to be safely invested in liquid products and, so, it cannot be used for operational or financial investments.

The regulation prescribes the required buffer for a firm as a whole. Whereas the buffer needs to be withheld at firm-level, the operations of the firms are allocated over several divisions. Think about a big insurance firm, that sells insurance products dealing with a variety of different classes of risk. For instance, one division sells damage related insurance, like bike theft and car insurance. Another division sells fire and earthquake insurance of real instate. The third division sells death benefit products, which are mainly sold to individuals who buy a mortgage. The interaction of the classes of risk is important for the firm. If earthquakes are very lethal, the earthquake risk is very severe in the worst-case scenarios of the total firm. The consequences of such a "stroke of bad luck" are regulated by means of holding a sufficiently large buffer. Therefore, the firm tries to mitigate a part of its risk by holding several classes of insurance risk. For risk management, a question that arises is which divisions are most influential for the total buffer that the firm should withhold by for example the Solvency II regulations. To answer this question, a firm could allocate the buffer to the divisions. This problem is called the *risk capital allocation problem*. It is not immediately clear what a “fair” allocation should be.

There are several reasons why this problem is important. First, allocating the buffer is important for performance measurement. It is not uncommon that the managers of the divisions are evaluated on the basis of the return earned on the amount of risk capital to be withheld for their portfolio. This requires an allocation of buffer to divisions that is perceived as “fair” by the managers. Second, the allocation is important for decisions regarding whether to increase or decrease the engagement in the operational activities of certain divisions. The attractiveness of a specific risk (e.g., a specific financial investment) is typically evaluated by means of a risk-return trade-off. Evaluating the performance of a division’s activities in isolation, however, can be very misleading. So, the allocation indicates the riskiness of a division’s portfolio for the total firm.

The second chapter of this dissertation proposes to use the Aumann-Shapley value as a solution of this problem. This allocation rule received considerable attention in the literature (see, e.g., Tasche, 1999; Denault, 2001; Tsanakas and Barnett, 2003; Kalkbrener, 2005) and is given by a gradient of a specific function. I derive a generalization of the Aumann-Shapley value that is also well-defined if this gradient does not exist. Aumann and Shapley (1972) show in their seminal book that a very general asymptotic approach leads towards the Aumann-Shapley value for many functions, but not the one that corresponds to the risk capital allocation problem. As a solution, I propose a weaker asymptotic approach. I, however, show existence in case of the risk capital allocation problem. Moreover, the approach that I use to characterize the allocation rule allows us to give an explicit formula for the corresponding capital allocations. The specific formula has a geometric interpretation. It still satisfies some properties that were known to be valid for the Aumann-Shapley value if it exists.

The third and fourth chapter analyze optimal risk sharing for firms. First, I study risk sharing in the context of longevity risk. Longevity risk is the risk of populations living longer or shorter than expected, for example, through medical advances or declining health risks such as smoking. It is the systematic risk involved in life-contingent liabilities that arises from the fact that death rates change in an unpredictable way. This risk is a major concern for Defined Benefit pension funds, since they typically promise their participants a retirement payment until death. An increase in longevity leads to an increase in the present value of these pension liabilities. On the other hand, death benefit insurers offer a pay-off in case of (early) death. If people live longer than expected, this leads to either later payment, or a larger chance of not paying out this insurance. This implies that an increase of longevity leads to a decrease of the present value of the death benefit insurer’s liabilities. Exposure to longevity risk can be rather substantial for pension funds and death benefit insurers (see, e.g., Jones, 2013).

One of the few examples of capital market transactions of longevity risk is Swiss Re, who issued an index-based instrument to protect against adverse longevity trends called VITA I in 2004.

"Until recently, virtually all Longevity Risk Transfers (LRT) activity had occurred in the United Kingdom, but 2012 saw three large non-UK transactions, a $26 billion pension buy-out deal between General Motors and Prudential Insurance, a €12 billion longevity swap between Aegon and Deutsche Bank, and a $7 billion pension buy-out between Verizon Communications and Prudential. However, as impressive as these volumes are, they represent only a small fraction of the aforementioned multi-trillion dollar potential market size." (Basel Committee on Banking Supervision, 2013).

So, reinsurance contracts do exist, but the capacity of reinsurance is limited (see, e.g., OECD, 2005). Attracting investors to a long-duration asset class is a challenge that still requires further development.

Moreover, the market for these instruments is still limited as there is still considerable uncertainty regarding the price of longevity risk. As a consequence, longevity-linked contracts are mainly traded Over-the-Counter (OTC). As pointed out by Dowd, Blake, Cairns and Dawson (2006), the market for Over-the-Counter survivor swaps is expected to grow fast. Via Over-The-Counter trade, the firms could mitigate longevity risk by exploiting natural hedge potential. Existing literature regarding the natural hedge potential that arises from combining liabilities with different sensitivities focuses on the optimal liability mix, but does not address the question whether and how changes in the liability mix can be obtained. When a firm does not succeed to attract participants to fulfill an optimal natural hedge of longevity risk, it can trade Over-The-Counter with pension funds and death benefit insurers. In Chapter 3, I analyze opportunities for pension funds and death benefit insurers to benefit from mutual redistribution of risk via such a trade. I allow for heterogeneous beliefs regarding the underlying probability distribution of longevity. I model the risk redistribution as the outcome of a bargaining process in which the involved parties bargain for a reallocation of risk and a price that benefits all. In a bargaining process, firms will cooperate with each other, and try to find a social optimum that is beneficial for the firms altogether. This social optimum corresponds with a unique redistribution of risk. Such a solution is characterized by Nash (1950) as the unique solution satisfying a set of four properties. Moreover, when more than two parties are involved, they will only reach a mutual agreement if no subset of parties can be better off by splitting off and instead redistributing the risk amongst each other.

I consider the case where pension funds and insurers redistribute risk in order to reduce the volatility of their liabilities at a prespecified future date. Firms determine a risk redistribution via cooperative bargaining. I show in a calibrated example with one pension fund and one death benefit insurer that the welfare gains are substantial. Both the pension fund as the death benefit insurer can reduce their risk by trading with each other.

In Chapter 4, I study optimal risk sharing if firms or investors have alternative preferences. Historically, one often assumes that investors optimize a trade-off between return (expected value) and risk (measured by the variance). In this trade-off, investors aim for a high return and a low risk of their (financial) investments. The preference for a low variance compared to a high expected return is called the risk aversion, and is different across investors. In this way, for instance, the classical Capital Asset Pricing Model (CAPM) is derived from earlier work of Markowitz (1952) via an equilibrium argument. The CAPM a theoretical model that is used to determine an expected rate of return of an asset, if investors optimally construct a portfolio via the risk-reward trade-off. The expected rate of return only depends on the sensitivity to the systematic risk in the market, often represented by a β.

Historically, variance was often used as a risk measure; it measures the volatility of a stochastic variable. It can however be shown that using the variance to calculate risk can lead to some strange outcomes for the preference relation. For instance, if a risky asset has a higher return than the risk-free asset in every scenario, you would expect the risky asset to be preferable. In fact, arbitrage opportunities arise in this case; one can construct a portfolio with price zero and a non-negative pay-off in every scenario and a strict positive pay-off in at least one scenario. Using the variance, however, might lead to still buying the risk-free asset. More recently, focus has shifted to the use of coherent risk measures. Distortion risk measures capture almost the whole class of coherent risk measures. In Chapter 4, I replace the variance to calculate risk by a distortion risk measure introduced by Yaari (1987). This resolves the problem mentioned above; the preference relation with distortion risk measures satisfies a set of reasonable properties. This follows from the fact that a risk measure weighs downside and upside risk differently. Moreover, the use of risk measures is in line with Basel II regulations for banks as a popular example is Expected Shortfall. Expected Shortfall has been gaining practitioner interest.

I study optimal risk sharing in the case where the objective of firms is to optimize a risk-reward trade-off with a distortion risk measure. I allow firms to use heterogeneous risk measures to evaluate risk. In Pareto optimal solutions, we often observe that there is a tranching (or layering) of the aggregate risk as shown by Ludkovski and Young (2009). Every tranche is allocated to a specific firm. Under mild conditions, I show that this specific Pareto optimal risk redistribution is unique. The firms only have to bargain about a specific price of this risk redistribution. Particularly, in trades of insurance products, tranching of the aggregate risk is empirically observed. Typically, the non-systematic part of insurance risk can be hedged via pooling contracts of clients. This creates incentives for insurance in the first place. Multiple small firms can benefit from risk redistributions by pooling their risk with other firms. The systematic part of insurance risk cannot be hedged by pooling risk. However, firms can still benefit from trading the systematic part of insurance risk with other firms which face insurance risk in other insurance risk classes. The systematic risk can be shared and, therefore, firms can benefit.

I also analyze prices of the tranched Pareto optimal risk redistribution. I do so in two settings. The first setting is in line with Chapter 3, where there is trading in not well-functioning markets.

I derive a characterization of specific prices based on four properties. I hereby focus on stability of a risk redistribution; no subset of firms should be better off if the members of this subset only cooperate with each other. The corresponding cooperative game extends the game of Denault (2001) by allowing for heterogeneous risk measures. The characterized prices coincide with the ones that I find in the second setting, where I focus on perfect markets and analyze competitive equilibria. The prices are characterized via the competitive equilibrium. I show necessary and sufficient conditions to guarantee uniqueness of the equilibrium prices and corresponding risk redistribution. I also derive a Capital Asset Pricing Model for the case where firms use risk measures to evaluate risk via these equilibrium prices. This model is tested by De Giorgi and Post (2008) in case firms are endowed with the same risk measure, and they show that the CAPM model with risk measures fits asset returns from the S&P 500 better than when firms use the variance to calculate risk.

Chapter 2

A generalization of the Aumann-Shapley value for risk capital allocation problems

This chapter is based on Boonen, De Waegenaere and Norde (2012b).

2.1 Introduction

This chapter proposes a rule to allocate *risk capital* among divisions within a firm. Regulators require that financial institutions withhold a level of capital that is invested safely in order to mitigate the effects of adverse events such as, for example, a financial crisis. This amount of capital is referred to as risk capital. Regulatory requirements focus at the level of risk capital to be withheld at firm level. Our focus is on how this amount of risk capital is allocated to different business divisions within the firm.[1] This problem is called the *risk capital allocation problem*.

There are several reasons why firms want to allocate risk capital to divisions. First, allocating risk capital is important for performance evaluation. Investment activities of financial institutions are typically divided into different portfolios, with different divisions within the firm being responsible for different portfolios. It is not uncommon that the managers of these divisions are evaluated on the basis of the return earned on the amount of risk capital to be withheld for their portfolio. This requires an allocation of risk capital to divisions that is perceived as "fair" by the managers. Second, allocating risk capital to business divisions is important for decisions regarding whether to increase or decrease the engagement in the activities of certain divisions. The attractiveness of a specific risky activity (e.g., a specific financial investment) is typically

[1] Alternatively, one can interpret a division as a financial portfolio.

evaluated by means of a risk-return trade-off. Evaluating the performance of a division's activities in isolation, however, can be very misleading. For example, the activity might seem highly risky in isolation, but may be useful in hedging risk in other divisions's activities.[2] One approach to evaluate the attractiveness of increasing the engagement in the activities of a specific division taking into account potential hedge effects is to determine the effect of increasing the level of the activities on the allocation of risk capital to all divisions.

The allocation problem is non-trivial because whenever a coherent risk measure (Artzner, Delbaen, Eber and Heath, 1999) is used to determine risk capital, the amount of risk capital to be withheld for the firm as a whole would typically be lower than the sum of the amounts of risk capital that would need to be withheld for each division in isolation. The reason is that the individual risks associated with the divisions are typically not perfectly correlated, and, hence, there can be some hedge potential from combining the risks. The allocation rule then determines how the benefits of this hedge potential are allocated to the divisions.

There is a large literature on capital allocation rules, with approaches based on finance (see, e.g., Tasche, 1999; Myers and Read, 2001; Sherris, 2006), optimization (see, e.g., Dhaene, Goovaerts and Kaas, 2003; Laeven and Goovaerts, 2006), and game theory (see, e.g., Denault, 2001; Tsanakas and Barnett, 2003; Tsanakas, 2004 and 2009). Our focus in this chapter is on game-theoretic approaches to allocating risk capital. A game-theoretic approach that has received considerable attention is the one of Denault (2001). He models the risk capital allocation problem as a *fuzzy game*. Specifically, he defines a risk capital allocation function as a function that assigns an amount of risk capital to every collection of fractions of divisions. The fraction of a division included in a collection is referred to as the *participation level* of that division. He then considers risk capital allocations that satisfy the stability condition that requires that, for any given collection of fractions of divisions, the amount of risk capital allocated to that collection is weakly lower than the amount of risk capital that they would need to withhold if they would separate from the firm. In game-theoretic terms, this condition means that the allocation is an element of the *fuzzy core*. Denault specifies a number of other desirable properties of a risk capital allocation rule, and shows that the *Aumann-Shapley value* (Aumann and Shapley, 1974) is the only allocation rule that is in the fuzzy core and satisfies these additional properties.[3] Moreover, Kalkbrener (2005) imposes a diversification axiom that requires the risk capital allocation of a division not to exceed its corresponding stand-alone risk capital. The Aumann-Shapley value is then characterized as the only allocation rule that satisfies this condition and two more technical conditions.

The Aumann-Shapley value as a risk capital allocation rule has received considerable attention

[2]An example would be an insurance company that holds both annuities and death benefit insurance. Both types of liabilities are sensitive to longevity risk (the risk associated with unpredictable changes in survival rates in a population). In isolation, each of these liabilities could be evaluated as relatively risky. However, the death benefit insurance provides hedge potential for the annuity portfolio. Van Gulick, De Waegenaere and Norde (2012) show the impact of this hedge potential on the allocation of risk capital.

[3]For general production functions, the Aumann-Shapley value is characterized by, e.g., Aumann and Shapley (1974), Aubin (1981), Billera and Heath (1982), and Mirman and Tauman (1982).

in the literature. Financial and economic arguments in favor of the Aumann-Shapley value are provided by, e.g., Tasche (1999) and Myers and Read (2001). One of the drawbacks, however, of the Aumann-Shapley value is that it requires *partial differentiability* of the fuzzy risk capital allocation function at the level of full participation of each division. It is well-known that the fuzzy risk capital function is generally not differentiable everywhere when the probability distributions of the risks associated with the divisions are not continuous (see, e.g., Tasche, 1999). We propose a generalization of the Aumann-Shapley value that is well-defined even when the risk capital function is not differentiable. The rule that we propose is inspired by the idea underlying the *Shapley value* (Shapley, 1953) for non-fuzzy cooperative games. We first discretize the participation levels of divisions by considering a finite grid of participation levels. Then, for any given discrete path on the grid starting from no participation (the participation profile where the participation level of each division is zero) and ending at full participation (the participation profile with full participation of each division), we determine the corresponding path-based allocation. Specifically, in each step of the path, the participation level of exactly one division is increased, and the corresponding difference in risk capital is allocated to that division. Proceeding in this way along the path, the total risk capital will be allocated once the path reaches the level of full participation. This procedure yields a risk capital allocation for every possible path. Moreover, the average of the corresponding risk capital allocations over all possible paths is also a risk capital allocation.[4] We show that when the grid size converges to zero, this average converges as well. The allocation rule that we propose in this chapter equals this asymptotic value. We refer to it as the *Weighted Aumann-Shapley* value. For risk capital allocation problems for which the corresponding risk capital function is differentiable at the level of full participation, the Weighted Aumann-Shapley value coincides with the Aumann-Shapley value. In contrast to the Aumann-Shapley value, however, the Weighted Aumann-Shapley value is well-defined even when the risk capital allocation function is non-differentiable.

In the seminal book of Aumann and Shapley (1974) the Aumann-Shapley value is introduced and characterized for special classes of games with a continuum of players. Roughly speaking, the characteristic functions of these games are obtained as differentiable functions of a finite number of non-atomic probability measures. Aumann and Shapley moreover provide the well-known "diagonal formula" for their value. Mertens (1980 and 1988) extends the Aumann-Shapley value and its axiomatic characterization to a much larger class of vector measure games by dropping the differentiability assumption. An overview of the Mertens value for vector measure games is given by Neyman (2002). Since fuzzy games can be considered as special examples of vector measure games, the Aumann Shapley value can be computed for these games as well under some differentiability assumptions and the Mertens value under much milder assumptions. In fact, Mertens shows that in order to compute the Mertens value the Aumann-Shapley diagonal formula should be generalized to a diagonal formula where an expectation of partial derivatives

[4] The construction of a rule as average over paths is in line with, e.g., Moulin (1995) and Sprumont (2005), who both consider a discrete production problem.

along random, small perturbations around the diagonal should be integrated.

Aumann and Shapley (1974) show that under very strong assumptions their value (and hence the Mertens value) can be obtained via an asymptotic approach. However, in Example 19.2 of their book they show that fuzzy games, corresponding to convex, piecewise affine functions (like the fuzzy games related to risk capital allocation problems which we consider in this chapter) do not satisfy this strong assumption (also pointed out by Neyman and Smorodinsky, 2004). In this chapter, we provide an allocation rule that follows a much weaker asymptotic approach than the one used by Aumann and Shapley (1974). In return, we get that our approach is convergent for all fuzzy games related to risk capital allocation problems. It is still an open question whether our value coincides with the Mertens value. An axiomatization of the Mertens value on the class of piece-wise linear fuzzy games is provided by Haimanko (2001).

We also show that the corresponding risk capital allocation rule satisfies a number of desirable properties. Some of these properties are known to be satisfied by the regular Aumann-Shapley value on the class of risk capital allocation problems for which the Aumann-Shapley value is well-defined. Moreover, the approach that we use to characterize the allocation rule allows us to give an explicit formula for the corresponding capital allocations. The specific formula has a geometric interpretation.

This chapter is set out as follows. In Section 2.2, we define risk capital and risk capital allocation problems. Two of the most prominent game-theoretic solution concepts for allocation problems are discussed in Section 2.3, namely the Shapley value and the Aumann-Shapley value. In Section 2.4, we provide the structure of the risk capital function. In Section 2.5 we define a class of path-based allocation rules. In Section 2.6, we introduce an allocation rule based on the average of path-based allocations. We show that the corresponding allocation rule can be seen as a generalization of the Aumann-Shapley value, and that it satisfies some desirable properties. We also provide a closed form expression with a geometric interpretation. Finally, Section 2.7 concludes.

2.2 Risk measures and risk capital allocation problems

In this chapter, we propose a rule to allocate risk capital among divisions. The firm uses a risk measure to determine this capital. In this section, we briefly introduce risk measures and risk capital allocation problems.

2.2.1 Risk measures

In this subsection, we discuss risk measures as in Artzner et al. (1999) and Delbaen (2000). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a fixed probability space, i.e., Ω is the state space, $\mathcal{F}$ is a σ-algebra on Ω and $\mathbb{P}$ is the physical probability measure on $(\Omega, \mathcal{F})$. We denote $\mathcal{P}(\Omega, \mathcal{F})$ as the set of all probability measures on $(\Omega, \mathcal{F})$ and $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ for the space of all bounded, measurable, real valued stochastic variables. If there is no confusion possible, we write $L^\infty = L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. We interpret a realization of a stochastic variable as a future loss.

A risk measure is a function $\rho : L^\infty \to \mathbb{R}$.[5] So, a risk measure maps stochastic variables into real numbers. It serves as a measure to determine the cash reserve for holding risk. The purpose of this reserve is to make the risk acceptable to the regulator. In this chapter, we only focus on *coherent* risk measures. Coherence is first introduced by Artzner et al. (1999). A risk measure ρ is called coherent if it satisfies the following four properties:

- *Sub-additivity*: For all $X, Y \in L^\infty$, we have

  $$\rho(X + Y) \leq \rho(X) + \rho(Y).$$

- *Monotonicity*: For all $X, Y \in L^\infty$ such that $X(\omega) \geq Y(\omega)$ holds almost surely for $\omega \in \Omega$ with respect to the measure $\mathbb{P}$, we have

  $$\rho(X) \geq \rho(Y).$$

- *Positive Homogeneity*: For every $X \in L^\infty$ and every $c > 0$, we have

  $$\rho(cX) = c\rho(X).$$

- *Translation Invariance*: For every $X \in L^\infty$ and every $c \in \mathbb{R}$, we have

  $$\rho(X + c \cdot e_\Omega) = \rho(X) + c,$$

  where $e_\Omega \in L^\infty$ is the risk with realization one in every state $\omega \in \Omega$.

The relevance of these properties is widely discussed by Artzner et al. (1999). Furthermore, the following property of a risk measure ρ is defined by, e.g., Delbaen (2000):

- *Comonotonic Additivity*: For all $X, Y \in L^\infty$ such that X and Y are comonotone, we have that

  $$\rho(X + Y) = \rho(X) + \rho(Y).$$

Random variables X and Y are comonotone if the inequality $[X(\omega_1) - X(\omega_2)] \cdot [Y(\omega_1) - Y(\omega_2)] \geq 0$ holds almost surely for all $(\omega_1, \omega_2) \in \Omega \times \Omega$ with respect to the product measure $\mathbb{P} \times \mathbb{P}$ (see, e.g., Delbaen, 2000). If a risk measure satisfies *Comonotonic Additivity*, it means that if stochastic variables are "perfectly" dependent, there is no benefit from pooling.

[5] Here, we assume that ρ is only defined on L^∞. For a discussion of risk measures on the class of all stochastic variables, we refer to Delbaen (2000).

Artzner et al. (1999) and Delbaen (2000) show that a risk measure ρ is coherent if and only if there exists a set of probability measures $Q \subset \mathcal{P}(\Omega, \mathcal{F})$ such that[6]

$$\rho(X) = \sup \{E_{\mathbb{Q}}[X] : \mathbb{Q} \in Q\}, \qquad \text{for all } X \in L^\infty. \tag{2.1}$$

The set Q need not be unique. We refer to a set Q that satisfies (2.1) as a generating probability measure set of ρ. Moreover, Delbaen (2000) shows that for every coherent risk measure ρ satisfying *Comonotonic Additivity*, there is a submodular function $\upsilon^\rho : \mathcal{F} \to \mathbb{R}_+$ with $\upsilon^\rho(\emptyset) = 0$ and $\upsilon^\rho(\Omega) = 1$ such that the following set Q is generating ρ:[7,8]

$$Q = \{\mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F}) : \mathbb{Q}(A) \geq \upsilon^\rho(A) \text{ for all } A \in \mathcal{F}\}. \tag{2.2}$$

In all examples of this chapter, we focus on a special class of coherent risk measures satisfying *Comonotonic Additivity*. This class is the class distortion risk measures (Wang, 1995) with a distortion function g^ρ.[9] It can be shown that, subject to a technical condition, any coherent risk measure satisfying *Comonotonic Additivity* can be represented by a distortion risk measure (Wang, Panjer and Young, 1997). In Chapter 4 of this dissertation, we discuss distortion risk measures in more detail in the context of optimal risk sharing.

Example 2.2.1 Based on Denneberg (1994), it holds for distortion risk measures that a function υ^ρ satisfying (2.2) is given by

$$\upsilon^\rho(A) = 1 - g^\rho(1 - \mathbb{P}(A)), \qquad \text{for all } A \in \mathcal{F}, \tag{2.3}$$

where g^ρ is the distortion function.[10] So, a function υ^ρ as in (2.2) has a known functional form that is only dependent on the function g^ρ and the probability space. Substituting (2.3) in (2.2) yields the following generating probability measure set of ρ:

$$Q = \{\mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F}) : \mathbb{Q}(A) \leq g^\rho(\mathbb{P}(A)) \text{ for all } A \in \mathcal{F}\}. \tag{2.4}$$

Next, we discuss a well-known coherent risk measure. We use this measure in all examples in this chapter. The risk measure *Expected Shortfall* (see, e.g., Acerbi and Tasche, 2002) is defined

[6] This result is shown by Artzner et al. (1999) in case of a finite state space and generalized by Delbaen (2000) to stochastic variables on L^∞.

[7] This result is deduced by Delbaen (2000) from earlier results of Denneberg (1994) and Schmeidler (1986).

[8] A function $\upsilon : \mathcal{F} \to \mathbb{R}$ is submodular if $\upsilon(S \cup T) + \upsilon(S \cap T) \geq \upsilon(S) + \upsilon(T)$ for all $S, T \in \mathcal{F}$.

[9] Distortion risk measures are given by $\rho(X) = \int_0^\infty g^\rho(\mathbb{P}(X > x))\, dx + \int_{-\infty}^0 (g^\rho(\mathbb{P}(X > x)) - 1)\, dx$ for all $X \in L^\infty$, where g^ρ is a continuous, concave and increasing function such that $g^\rho(0) = 0$ and $g^\rho(1) = 1$. Here, convergence is guaranteed by boundedness of X.

[10] This result is shown by Denneberg (1994) for Choquet integrals. Distortion risk measures are examples of Choquet integrals.

as follows.[11] Let $\alpha \in (0,1)$. The $(1-\alpha)$-quantile is given by

$$q_{1-\alpha}(X) = \sup\{x \in \mathbb{R} : \mathbb{P}(\{\omega \in \Omega : X(\omega) \geq x\}) > \alpha\}, \qquad \text{for all } X \in L^{\infty}. \tag{2.5}$$

Then, the risk measure Expected Shortfall with significance level $\alpha \in (0,1)$ is defined as

$$\rho_{\alpha}^{ES}(X) = \alpha^{-1}(E_{\mathbb{P}}\left[X \cdot \mathbb{1}_{X \geq q_{1-\alpha}(X)}\right] - q_{1-\alpha}(X)(\mathbb{P}[X \geq q_{1-\alpha}(X)] - \alpha)), \qquad \text{for all } X \in L^{\infty}, \tag{2.6}$$

where $\mathbb{1}_{X \geq q_{1-\alpha}(X)} \in L^{\infty}$ is such that $\mathbb{1}_{X \geq q_{1-\alpha}(X)}(\omega) = 1$ if $X(\omega) \geq q_{1-\alpha}(X)$ and $\mathbb{1}_{X \geq q_{1-\alpha}(X)}(\omega) = 0$ otherwise. Note that if X is continuously distributed, we have $\rho_{\alpha}^{ES}(X) = E\left[X : X \geq q_{1-\alpha}(X)\right]$. Tasche (2002) shows that this risk measure ρ is coherent and, moreover, that it satisfies *Comonotonic Additivity*.

Dhaene et al. (2006) show that Expected Shortfall with significance level $\alpha \in (0,1)$ is a distortion risk measure and its distortion function is given by

$$g^{\rho_{\alpha}^{ES}}(x) = \min\left\{\frac{x}{\alpha}, 1\right\}. \tag{2.7}$$

If the state space Ω is finite and the σ-algebra $\mathcal{F}$ equals its power set, we can replace the event $A \in \mathcal{F}$ in expression (2.4) by state $\omega \in \Omega$. This holds since $g(x)$ is linear for $x \leq \alpha$ and $\mathbb{Q}(A) \leq 1$ for all $\mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F})$ and for all $A \in \mathcal{F}$. Hence, in this case, the generating probability measure set of ρ_{α}^{ES} from (2.4) is given by

$$Q = \left\{\mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F}) : \mathbb{Q}(\{\omega\}) \leq \frac{\mathbb{P}(\{\omega\})}{\alpha} \text{ for all } \omega \in \Omega\right\}. \tag{2.8}$$

In Example 2.2.4, we discuss this probability measure set in more detail. $\triangledown$

Next, we introduce a special class of risk measures. In this chapter, we consider risk measures that are finitely generated, i.e., the risk measure has a finite generating probability measure set Q.

Definition 2.2.2 A coherent risk measure $\rho : L^{\infty}(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ is *finitely generated* if there exists a finite generating probability measure set $Q \subset \mathcal{P}(\Omega, \mathcal{F})$, i.e.,

$$\rho(X) = \max\left\{E_{\mathbb{Q}}\left[X\right] : \mathbb{Q} \in Q\right\}, \qquad \text{for all } X \in L^{\infty}. \tag{2.9}$$

Again, we note that this set Q need not be uniquely determined.

The condition in Definition 2.2.2 on coherent risk measures seems quite restrictive. However, we show that all coherent risk measures satisfying *Comonotonic Additivity* satisfy this property

[11] The risk measure Expected Shortfall is also referred by other authors as Worst Conditional Expectation (see, e.g., Artzner et al., 1999), Conditional VaR or Tail VaR (see, e.g., Dhaene et al., 2006).

in case the state space is finite.

Proposition 2.2.3 *If the state space Ω is finite and the risk measure ρ is coherent and satisfies Comonotonic Additivity, then ρ is finitely generated.*

A finitely generated risk measure, however, does not need to satisfy *Comonotonic Additivity.* As a main example, Expected Shortfall belongs to the class of finitely generated risk measures if Ω is finite.

In the next example, we provide an explicit expression for a specific finite generating probability measure set corresponding to Proposition 2.2.3.

Example 2.2.4 Let the state space Ω be finite, the σ-algebra $\mathcal{F}$ equal its power set, and the risk measure ρ be a coherent risk measure satisfying *Comonotonic Additivity.* Recall the generating probability measure set in (2.2). Let σ be an order on Ω, i.e., $\sigma : \{1, \ldots, |\Omega|\} \to \Omega$ is the bijective function that corresponds with a permutation on the state space. The state at position $j \in \{1, \ldots, |\Omega|\}$ in the order σ is denoted by $\sigma(j) \in \Omega$ and the set of all orders on the state space Ω is denoted by $\Pi(\Omega)$. We define for every order $\sigma \in \Pi(\Omega)$ the following stochastic variable:

$$m^\sigma(\sigma(j)) = \upsilon^\rho\left(\bigcup_{k=1}^{j}\{\sigma(k)\}\right) - \upsilon^\rho\left(\bigcup_{k=1}^{j-1}\{\sigma(k)\}\right), \text{ for all } j \in \{1, \ldots, |\Omega|\}. \tag{2.10}$$

The stochastic variable m^σ is a probability measure on $(\Omega, \mathcal{F})$ due to submodularity and non-negativity of υ^ρ, $\upsilon^\rho(\emptyset) = 0$, and $\upsilon^\rho(\Omega) = 1$. For distortion risk measures, we can simplify (2.10) to

$$m^\sigma(\sigma(j)) = g^\rho\left(\sum_{k=j}^{|\Omega|}\mathbb{P}(\{\sigma(k)\})\right) - g^\rho\left(\sum_{k=j+1}^{|\Omega|}\mathbb{P}(\{\sigma(k)\})\right), \text{ for all } j \in \{1, \ldots, |\Omega|\}, \tag{2.11}$$

and for all $\sigma \in \Pi(\Omega)$. Then, the following set is a generating probability measure set of ρ:

$$Q = \{m^\sigma : \sigma \in \Pi(\Omega)\}. \tag{2.12}$$

This result follows almost directly from the proof of Proposition 2.2.3.[12]

Next, we provide an example of the construction of (2.12). Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$. If the state space is finite, we write probability measures as row vectors. We let the physical probability measure given by $\mathbb{P} = \left(\frac{1}{20}, \frac{9}{20}, \frac{1}{2}\right)$, and $\rho = \rho^{ES}_{0.1}$. Using (2.11) with $g^{\rho^{ES}_{0.1}}(x) = \min\left\{\frac{x}{0.1}, 1\right\}$ for all $x \in [0, 1]$, we obtain the following outcomes of $(m^\sigma(\omega_1), m^\sigma(\omega_2), m^\sigma(\omega_3))$ for all $\sigma \in \Pi(\Omega)$:

[12] Suppose (Ω, υ^ρ) is a Transferable Utility game, where Ω is the corresponding "player" set. In game-theoretical terms, the representation (2.2) of a finite generating probability measure set coincides with the core of the game (Ω, υ^ρ). Then, submodularity of the function υ^ρ is equivalent with convexity of the corresponding game (Shapley, 1971). Shapley (1971) shows that the core of convex games coincides with the convex hull of the marginal vectors. A marginal vector corresponds with a vector m^σ for some $\sigma \in \Pi(\Omega)$. All extreme points of the core of (Ω, υ^ρ) are in (2.12).

σ	ω_1	ω_2	ω_3
$\omega_1\,\omega_2\,\omega_3$	0	0	1
$\omega_1\,\omega_3\,\omega_2$	0	1	0
$\omega_2\,\omega_1\,\omega_3$	0	0	1
$\omega_2\,\omega_3\,\omega_1$	$\frac{1}{2}$	0	$\frac{1}{2}$
$\omega_3\,\omega_2\,\omega_1$	$\frac{1}{2}$	$\frac{1}{2}$	0
$\omega_3\,\omega_1\,\omega_2$	0	1	0

According to (2.12), a finite generating probability measure set of $\rho^{ES}_{0.1}$ is given by

$$Q = \{\mathbb{Q}_1, \mathbb{Q}_2, \mathbb{Q}_3, \mathbb{Q}_4\}, \tag{2.13}$$

where $\mathbb{Q}_1 = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$, $\mathbb{Q}_2 = (0, 0, 1)$, $\mathbb{Q}_3 = (0, 1, 0)$, and $\mathbb{Q}_4 = \left(\frac{1}{2}, 0, \frac{1}{2}\right)$. Recall that the elements of Q are the extreme points of the set in (2.8). $\triangledown$

Throughout the sequel of this chapter, we fix for a given finitely generated risk measure ρ a finite generating probability measure set, which we denote by $Q(\rho)$. This assumption is made without loss of generality.

2.2.2 Risk capital allocation problems

In this subsection, we discuss risk capital allocation problems as in, e.g., Denault (2001). Consider a firm, for example a pension fund or an insurance company. This firm consists of multiple divisions that face risk. The risk of a division is summarized by a stochastic loss variable at a common future time. The problem is to allocate the total risk among all divisions.

The finite set of all divisions within a firm is denoted by N. Throughout this chapter, we fix N and the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote for each division $i \in N$ the stochastic loss as $X_i \in L^\infty$. The total loss of the firm is given by $\sum_{i\in N} X_i$. We assume that the risk capital of the firm is measured using a finitely generated risk measure ρ. In the following definition, we summarize the risk capital allocation problem.

Definition 2.2.5 A *risk capital allocation problem* is a tuple $((X_i)_{i\in N}, \rho)$, where $X_i \in L^\infty$ for all $i \in N$, and ρ is a finitely generated risk measure. The class of all risk capital allocation problems is denoted by $\mathcal{R}$.

In the following definition, we state the concept of *risk capital allocations* and *risk capital allocation rules.*

Definition 2.2.6 A vector $a \in \mathbb{R}^N$ is a *risk capital allocation* for $R = ((X_i)_{i\in N}, \rho) \in \mathcal{R}$ if $\sum_{i\in N} a_i = \rho\left(\sum_{i\in N} X_i\right)$. For $\widetilde{\mathcal{R}} \subseteq \mathcal{R}$, a risk capital allocation rule is a function $K : \widetilde{\mathcal{R}} \to \mathbb{R}^N$ that assigns to every risk capital allocation problem $R \in \widetilde{\mathcal{R}}$ a unique risk capital allocation $K(R)$.

It holds that

$$\sum_{i\in N} K_i(R) = \rho\left(\sum_{i\in N} X_i\right), \qquad \text{for all } R \in \widetilde{\mathcal{R}}.$$

The *Sub-additivity* property of coherent risk measures implies that there can be benefits from pooling risks. Specifically, it implies that

$$\rho\left(\sum_{i\in N} X_i\right) \leq \sum_{i\in N} \rho(X_i). \tag{2.14}$$

This property implies that allocating risk capital among divisions is generally non-trivial. The aim is to allocate the gains from pooling risk in a fair way.

Based on Denault (2001), we define the following properties of a risk capital allocation rule $K : \widetilde{\mathcal{R}} \to \mathbb{R}^N$:

- *Translation Invariance*: For all $R = ((X_i)_{i\in N}, \rho) \in \widetilde{\mathcal{R}}$, it holds that if $\hat{R} = ((\hat{X}_i)_{i\in N}, \rho) \in \widetilde{\mathcal{R}}$ where $(\hat{X}_i)_{i\in N} = (X_j + c \cdot e_\Omega, X_{-j})$ for some $c \in \mathbb{R}$ and $j \in N$, then

  $$K(\hat{R}) = K(R) + c \cdot e_j,$$

  where e_j is the j-th unit vector in $\mathbb{R}^N$.

- *Scale Invariance*: For all $R = ((X_i)_{i\in N}, \rho) \in \widetilde{\mathcal{R}}$, it holds that if $\hat{R} = ((\hat{X}_i)_{i\in N}, \rho) \in \widetilde{\mathcal{R}}$ where $(\hat{X}_i)_{i\in N} = (cX_i)_{i\in N}$ for some $c > 0$, then

  $$K(\hat{R}) = cK(R).$$

- *Monotonicity*: For all $R = ((X_i)_{i\in N}, \rho) \in \widetilde{\mathcal{R}}$ where ρ is non-decreasing in the sense that $\rho\left(\sum_{i\in N} \lambda_i X_i\right) \leq \rho\left(\sum_{i\in N} \lambda_i^* X_i\right)$ whenever $\lambda, \lambda^* \in [0,1]^N$ and $\lambda \leq \lambda^*$, we have

  $$K(R) \geq 0.$$

- *Fuzzy Core Selection*: For all $R = ((X_i)_{i\in N}, \rho) \in \widetilde{R}$, we have

  $$K(R) \in FCore(R),$$

  where $FCore(R)$ denotes the fuzzy core (Aubin, 1979), which is defined as:

$$FCore(R) = \left\{ a \in \mathbb{R}^N : \sum_{i\in N} \lambda_i a_i \leq \rho\left(\sum_{i\in N} \lambda_i X_i\right) \text{ for all } \lambda \in [0,1]^N, \sum_{i\in N} a_i = \rho\left(\sum_{i\in N} X_i\right) \right\}. \tag{2.15}$$

In particular, the property *Fuzzy Core Selection* is widely discussed in the game-theoretic literature on risk capital allocation problems (see, e.g., Denault, 2001; Tsanakas and Barnett, 2003). For an allocation in the fuzzy core, no portfolio of fractional risks would have a lower stand-alone risk capital than the corresponding risk capital allocation. This property ensures a stable allocation.

2.3 The Shapley value and the Aumann-Shapley value

In this section, we discuss cooperative game-theoretic solution concepts for risk capital allocation problems. There is one particular solution concept that has received considerable attention in cooperative game theory, namely the Shapley value (Shapley, 1953). The following definition states the Shapley value for risk capital allocation problems.

Definition 2.3.1 The *Shapley value* for risk capital allocation problems, denoted by $S : \mathcal{R} \to \mathbb{R}^N$, is given by

$$S_i(R) = \sum_{S \subseteq N \setminus \{i\}} w(|S|) \left(\rho \left(\sum_{j \in S \cup \{i\}} X_j \right) - \rho \left(\sum_{j \in S} X_j \right) \right),$$

for all $R \in \mathcal{R}$ and $i \in N$, where $|S|$ denotes the number of divisions in $S \subseteq N$ and

$$w(|S|) = \frac{|S|!(|N| - |S| - 1)!}{|N|!}.$$

The Shapley value is originally defined for Transferable Utility games, and is here applied to the atomic risk capital cost game of Denault (2001).[13] The Shapley value can be interpreted as the average of all marginal vectors. Given an ordering of divisions, a marginal vector is created by assigning to every division its marginal contribution if they enter the coalition one-by-one according to the order. The weight $w(|S|)$ assigned to a marginal contribution of division $i \in N$ to a coalition S represents the "probability" that in a uniformly random ordering of divisions, all divisions in S are on the first positions, and thereafter is division i.

It is easy to show that the Shapley value is a risk capital allocation rule. Denault (2001) shows that the Shapley value does not satisfy the stability criterium that, in game-theoretical terms, means that the Shapley value may not yield a core element. This is seen as a major drawback of the Shapley value for risk capital allocation problems.

Aumann and Shapley (1974) developed an allocation rule for general cost functions. This is given in the following definition.

[13] Denault (2001) defines the atomic risk capital game (N, c) as $c(S) = \rho(\sum_{i \in S} X_i)$ for all $S \subseteq N$.

Definition 2.3.2 The *Aumann-Shapley value* of a function $r : [0,1]^N \to \mathbb{R}$ is given by $a \in \mathbb{R}^N$ such that

$$a_i = \int_0^1 \frac{\partial r}{\partial \lambda_i}(\gamma e_N)\, d\gamma, \qquad \text{for all } i \in N, \tag{2.16}$$

whenever these integrals exist, and where e_N is the vector of ones in $\mathbb{R}^N$.

The Aumann-Shapley value can be interpreted as the average of the marginal changes of the risk capital function r if the participation of all divisions increases simultaneously.

Denault (2001) applies the Aumann-Shapley value to risk capital allocation problems. The *risk capital function* $r : [0,1]^N \to \mathbb{R}$ of a risk capital allocation problem $R = ((X_i)_{i\in N}, \rho) \in \mathcal{R}$ is defined as

$$r(\lambda) = \rho\left(\sum_{i\in N} \lambda_i X_i\right), \qquad \text{for all } \lambda \in [0,1]^N. \tag{2.17}$$

The interpretation is as follows. Here, divisions are allowed to participate fractionally. The participation level for division $i \in N$, denoted by $\lambda_i \in [0,1]$, can be seen as the fractional involvement of division i in a coalition. Here, the maximal participation level of every division is normalized to one.[14] Moreover, we define a participation profile $\lambda \in [0,1]^N$ as a collection of participation levels of all divisions in N. Note that the risk capital of the firm is given by $r(e_N) = \rho(\sum_{i\in N} X_i)$.

Let $\mathcal{R}' \subset \mathcal{R}$ be the set of all risk capital allocation problems for which r is partially differentiable at $\lambda = e_N$. Since the risk measure ρ satisfies *Positive Homogeneity*, we can simplify expression (2.16) as in the following corollary of, e.g., Denault (2001).

Corollary 2.3.3 *The Aumann-Shapley value for risk capital allocation problems, denoted by* $AS : \mathcal{R}' \to \mathbb{R}^N$*, is given by*

$$AS_i(R) = \frac{\partial r}{\partial \lambda_i}(e_N), \qquad \text{for all } i \in N \text{ and } R \in \mathcal{R}', \tag{2.18}$$

where the risk capital function r is as in (2.17).

The Aumann-Shapley value for risk capital allocation problems is the gradient of the risk capital function evaluated in the vector of full participation. As we focus on risk capital allocation problems in this chapter, we continue by referring to AS as the Aumann-Shapley value.

We discuss the main game-theoretic argument why the Aumann-Shapley value is widely supported as allocation rule for risk capital. This property involves the fuzzy core that is defined in (2.15). Allocations in the fuzzy core are derived on the premise that every fuzzy portfolio

[14] This normalization is contrary to the approach of Denault (2001), but it is without loss of generality.

should be allocated less than its stand-alone risk capital, i.e., no fuzzy coalition λ has an incentive to split off from the firm. The following theorem about the relationship between the fuzzy core and the Aumann-Shapley value is due to Denault (2001), which is based on an earlier result of Aubin (1979).

Theorem 2.3.4 (Denault, 2001, Theorem 7, page 20) *For all $R \in \mathcal{R}'$, the fuzzy core $FCore(R)$ consists of only one element, which is the Aumann-Shapley value.*

The set of worst-case probability measures for the firm is given by

$$Q^*(\rho) = \left\{ \mathbb{Q} \in Q(\rho) : r(e_N) = E_{\mathbb{Q}}\left[\sum_{i \in N} X_i\right] \right\}. \tag{2.19}$$

The following theorem provides an expression of the fuzzy core. This is based on Aubin (1979, Proposition 4, page 343).

Theorem 2.3.5 *Let $R \in \mathcal{R}$. Then, we have*

$$FCore(R) = conv\{(E_{\mathbb{Q}}[X_i])_{i \in N} : \mathbb{Q} \in Q^*(\rho)\},$$

where conv denotes the convex hull operator, and $Q^*(\rho)$ is defined in (2.19).

Next, we can reformulate the Aumann-Shapley value from (2.18). It follows from Theorem 2.3.4 and Theorem 2.3.5 that the Aumann-Shapley value for division $i \in N$, if well-defined, is given by the expectation of X_i under any worst-case probability measure $\mathbb{Q} \in Q^*(\rho)$, i.e.,

$$AS_i(R) = E_{\mathbb{Q}}[X_i], \qquad \text{for all } R \in \mathcal{R}' \text{ and } i \in N \text{ and for any} \mathbb{Q} \in Q^*(\rho). \tag{2.20}$$

If there exists a division $i \in N$ and $\mathbb{Q}, \widetilde{\mathbb{Q}} \in Q^*(\rho)$ such that $E_{\mathbb{Q}}[X_i] \neq E_{\widetilde{\mathbb{Q}}}[X_i]$, then the Aumann-Shapley value does not exist. Moreover, the following result is based on a characterization of the Aumann-Shapley value of Denault (2001).

Theorem 2.3.6 (Denault, 2001, Corollary 1, page 20) *The Aumann-Shapley value satisfies the properties Translation Invariance, Scale Invariance and Monotonicity on $\mathcal{R}'$.*

Next, we provide an example about the construction of the Aumann-Shapley value.

Example 2.3.7 Recall the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ from Example 2.2.4. Let $N = \{1, 2\}$. If the state space is finite, we write stochastic variables as column vectors. Let the risk capital allocation problem $R = ((X_i)_{i \in N}, \rho)$ be given by

$$X_1 = \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}, X_2 = \begin{bmatrix} 6 \\ 2 \\ -2 \end{bmatrix}, \text{ and, again, } \rho = \rho_\alpha^{ES}.$$

The corresponding risk capital function r, which is defined in (2.17), is given by

$$r(\lambda_1, \lambda_2) = \begin{cases} \lambda_1 + 4\lambda_2 & \text{if } 2\lambda_2 \geq \lambda_1, \\ 4\lambda_1 - 2\lambda_2 & \text{otherwise,} \end{cases} \tag{2.21}$$

for all $\lambda \in [0, 1]^N$. This fuzzy risk capital function is displayed in Figure 2.1. Note that the

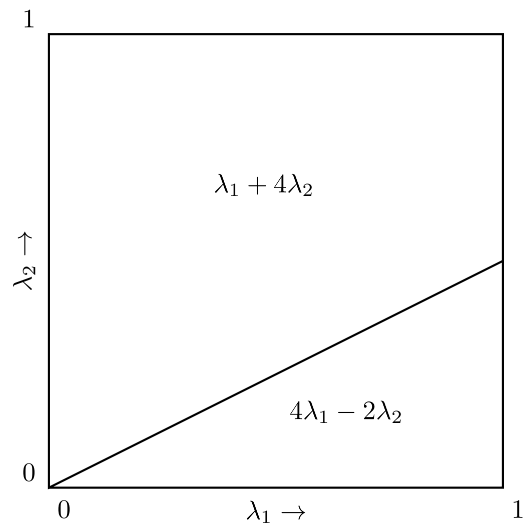


Figure 2.1: Plot of $r(\lambda_1, \lambda_2)$ corresponding to (2.21).

participation profiles λ where the risk capital function r is not partially differentiable are located on a straight line through the origin. It can be verified that $\mathbb{Q} = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$ is a probability measure in $Q^*(\rho_{0.1}^{ES})$ from (2.19). Hence, using (2.20), the Aumann-Shapley value is given by

$$AS(R) = (E_{\mathbb{Q}}[X_1], E_{\mathbb{Q}}[X_2]) = (1, 4).$$

$\bigtriangledown$

The main drawback of the Aumann-Shapley value is that it requires partial differentiability of the risk capital function. In Section 2.6, we extend the Aumann-Shapley value such that it is always well-defined.

2.4 Fuzzy risk capital allocation functions

In this section, we analyze the structure of the risk capital function r defined in (2.17). We show that it is piecewise linear and almost everywhere partially differentiable. We also introduce some notation that we need in Section 2.6, where we define an allocation rule.

Definition 2.4.1 Let $R = ((X_i)_{i\in N}, \rho) \in \mathcal{R}$. Then, for all $\mathbb{Q} \in Q(\rho)$ the linear function $f_{\mathbb{Q}} : \mathbb{R}^N \to \mathbb{R}$ is given by

$$f_{\mathbb{Q}}(\lambda) = \sum_{i\in N} \lambda_i E_{\mathbb{Q}}[X_i], \qquad \text{for all } \lambda \in \mathbb{R}^N.$$

Note that it is possible that different measures $\mathbb{Q} \in Q(\rho)$ yield the same function $f_{\mathbb{Q}}$.

Proposition 2.4.2 *For all $R = ((X_i)_{i\in N}, \rho) \in \mathcal{R}$, the risk capital function r is piecewise linear on $[0,1]^N$.*

From (2.42)-(2.44) in Appendix 2.A it follows that for every $\lambda \in [0,1]^N$, there exists at least one $\mathbb{Q} \in Q(\rho)$ such that $r(\lambda) = f_{\mathbb{Q}}(\lambda)$.

Next, we define the set of participation profiles corresponding to a probability measure in $Q(\rho)$ where this measure is the worst-case probability measure.

Definition 2.4.3 Let $R = ((X_i)_{i\in N}, \rho) \in \mathcal{R}$. Then, for all $\mathbb{Q} \in Q(\rho)$ the participation profile set $A_{\mathbb{Q}} \subseteq [0,1]^N$ is given by

$$A_{\mathbb{Q}} = \left\{\lambda \in [0,1]^N : r(\lambda) = f_{\mathbb{Q}}(\lambda)\right\}.$$

Moreover, we determine $K_1, \ldots, K_p$, $p^* \leq p$ and $\mathbb{Q}_1, \ldots, \mathbb{Q}_p \in Q(\rho)$ such that:

- $K_1, \ldots, K_p$ is an exhaustive list of the elements of $\{A_{\mathbb{Q}} : \mathbb{Q} \in Q(\rho)\}$ without repetitions;
- $K_m = A_{\mathbb{Q}_m}$ for all $m \in \{1, \ldots, p\}$;
- $e_N \in K_m$ for $m \in \{1, \ldots, p^*\}$, and $e_N \notin K_m$ otherwise.

It is straightforward to show that K_m is a polytope. Define $e_\emptyset$ as the zero vector in $\mathbb{R}^N$. The participation profile $\lambda = e_\emptyset$ is an element of K_m for all $m \in \{1, \ldots, p\}$. Moreover, we have

$$\bigcup_{m=1}^{p} K_m = [0,1]^N, \tag{2.22}$$

and $p \leq |Q(\rho)|$.

Example 2.4.4 In this example, we illustrate Definition 2.4.3. Recall the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbb{Q}_1, \ldots, \mathbb{Q}_4$ from Example 2.2.4, and the risk capital allocation problem from Example 2.3.7. The ordering of $\mathbb{Q}_1, \ldots, \mathbb{Q}_4$ corresponds with Definition 2.4.3. Based on this definition, we get $K_1 = A_{\mathbb{Q}_1} = \left\{\lambda \in [0,1]^N : \lambda_1 \leq 2\lambda_2\right\}$, $K_2 = A_{\mathbb{Q}_2} = \left\{\lambda \in [0,1]^N : \lambda_1 \geq 2\lambda_2\right\}$, and $K_3 = A_{\mathbb{Q}_3} = A_{\mathbb{Q}_4} = \left\{\lambda \in [0,1]^N : \lambda_1 = 2\lambda_2\right\}$. Note that $f_{\mathbb{Q}_3} = f_{\mathbb{Q}_4}$ and, so, it is without loss of generality to drop $\mathbb{Q}_4$. $\triangledown$

Next, we focus on partial differentiability of the risk capital function r. Partial differentiability is a key issue for existence of the Aumann-Shapley value (see Corollary 2.3.3) and, moreover, it is key in the risk capital allocation rule that we define in Section 2.6. For every λ such that there exists a unique $m \in \{1,\ldots,p\}$ such that $\lambda \in K_m$, there exists a neighborhood $U \subset [0,1]^N$ of λ such that $r(\hat{\lambda}) = f_{\mathbb{Q}_m}(\hat{\lambda})$ for all $\hat{\lambda} \in U$ and, so,

$$\frac{\partial r}{\partial \lambda_i}(\lambda) = E_{\mathbb{Q}_m}[X_i], \qquad \text{for all } i \in N. \tag{2.23}$$

Hence, uniqueness of $m \in \{1,\ldots,p\}$ such that $\lambda \in K_m$ is a sufficient condition for the risk capital function r to be partially differentiable in λ.

Lemma 2.4.5 *For all $R = ((X_i)_{i\in N}, \rho) \in \mathcal{R}$, the collection of profiles where the risk capital function r is not partially differentiable is a subset of a collection of a finite number of hyperplanes passing trough $\lambda = e_\emptyset$.*

From Lemma 2.4.5, we immediately obtain the following corollary.

Corollary 2.4.6 *For all $R \in \mathcal{R}$, the risk capital function r is almost everywhere partially differentiable.*

In the sequel of this chapter, we use the risk capital function r for defining an allocation rule in Section 2.6. We need the results in Lemma 2.4.5 and Corollary 2.4.6 in Subsection 2.6.2.

2.5 Path based allocation rules

In this section, we discuss *path based allocation rules* as introduced by Wang (1999). We construct an allocation based on the idea of the marginal vectors of the Shapley value. We extend this idea to a problem where divisions can participate fractionally via a finite set of participation levels. We first describe this allocation rule informally and, thereafter, we provide a formal definition.

Let $n \in \mathbb{N}$, and define the grid on $[0,1]^N$ with grid size $\frac{1}{n}$ by

$$G^n = \left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\right\}^N. \tag{2.24}$$

The starting point on grid G^n is the participation profile $e_\emptyset$ in which the participation level of each division is zero. In the first step the participation level of some division i is increased by $\frac{1}{n}$, and the corresponding difference in risk capital, $r\left((1/n)\cdot e_i\right) - r(e_\emptyset)$, is allocated to division i. In the second step of the path again the participation level of some division (not necessarily the same as the one in the first step) is increased by $\frac{1}{n}$, and the risk change is allocated to this division. Proceeding in this way, we will end up after $|N|n$ steps in e_N, and risk capital $r(e_N)$ has been allocated to the divisions by then.

Formally, we define a path in the following way.

Definition 2.5.1 Let $n \in \mathbb{N}$. A path on the grid G^n is a map $P : \{0, 1, \ldots, |N|n\} \to G^n$ satisfying:

1. $P(0) = e_\emptyset$ and $P(|N|n) = e_N$;
2. for every $k \in \{0, \ldots, |N|n - 1\}$ there exists a unique $i \in N$ such that
$$P(k+1) - P(k) = \frac{1}{n} e_i. \tag{2.25}$$
This unique division i is denoted as $i(P, k)$.

An example of a path P on the grid G^n is given in Figure 2.2. We denote the collection of all paths over the grid G^n by $\mathcal{P}^n$. For any path $P \in \mathcal{P}^n$, we define an allocation rule A^P on $\mathcal{R}$ by

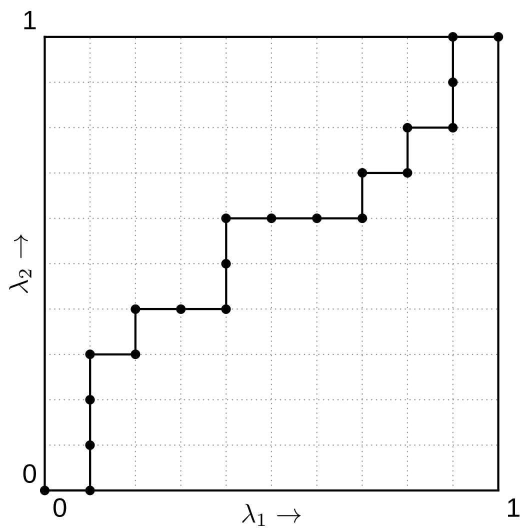


Figure 2.2: Example of a path $P \in \mathcal{P}^n$ for $|N| = 2$ with $n = 10$. We connected succeeding elements of the path as illustration.

allocating the risk changes along the path to the corresponding divisions. Formally, we define this allocation rule as follows.

Definition 2.5.2 For a given path $P \in \mathcal{P}^n$, the path based allocation rule $A^P : \mathcal{R} \to \mathbb{R}^N$ is given by

$$A^P(R) = \sum_{k=0}^{|N|n-1} [r(P(k+1)) - r(P(k))] \cdot e_{i(P,k)}, \qquad \text{for all } R \in \mathcal{R},$$

where the risk capital function r is defined in (2.17).

Proposition 2.5.3 *Let $n \in \mathbb{N}$. For any $P \in \mathcal{P}^n$, A^P is an allocation rule on $\mathcal{R}$.*

We refer to A^P as a path based allocation rule. Next, we show some general properties of path based allocation rules.

Theorem 2.5.4 *Let $n \in \mathbb{N}$. For every $P \in \mathcal{P}^n$, the allocation rule A^P satisfies the properties Translation Invariance, Scale Invariance and Monotonicity on $\mathcal{R}$.*

Compare this result with Theorem 2.3.6. Denault (2001) shows that there exists an allocation rule satisfying the above-mentioned properties, and that the Aumann-Shapley value is such a rule. Using Theorem 2.5.4, we extend this result for all allocation rules based on a path. Moreover, one can verify using *Positive Homogeneity* and *Sub-additivity* of ρ that a path based allocation rule always satisfies the following bounds:

$$\min\{E_{\mathbb{Q}}[X_i] : \mathbb{Q} \in Q(\rho)\} \leq A_i^P(R) \leq \rho(X_i),$$

for all $P \in \mathcal{P}^n$, $R = ((X_i)_{i \in N}, \rho) \in \mathcal{R}$ and $i \in N$. So, every path based allocation is individually rational, and is always weakly more than the expectation of X_i under the best-case probability measure.

We can approximate expression (2.16) of the Aumann-Shapley value (if existent) using a very small grid and a path close to the diagonal. Note that the Aumann-Shapley value is not defined if the risk capital function r is not partially differentiable along the diagonal. As a solution, we propose a generalization based on paths that is not prone to this problem in the next section.

2.6 The Weighted Aumann-Shapley value

2.6.1 A sequence of discrete rules

In this section we introduce a generalization of the well-known Aumann-Shapley value. In line with the Shapley value (1953), we define an allocation rule based on the average of all path based allocation rules corresponding to paths on a finite grid, i.e., average of A^P for all $P \in \mathcal{P}^n$ for some $n \in \mathbb{N}$. Then, we let the grid size converge to zero. We show that the limit of the corresponding allocation rules exists, and that it is a well-defined value also in case of non-differentiability of the risk capital function r. We show that in case of differentiability the outcome is the standard Aumann-Shapley value. In case of non-differentiability however, the outcome can be regarded as a weighted average of standard Aumann-Shapley values for "nearby" risk capital allocation problems with a differentiable risk capital function. Moreover, we provide a geometric interpretation of the corresponding weights. We refer to this value as the *Weighted Aumann-Shapley value.*

In Definition 2.5.2 we introduced an allocation rule based on a path $P \in \mathcal{P}^n$. Since A^P is a risk capital allocation rule (see Proposition 2.5.3), the average of all path based risk capital allocation rules is a risk capital allocation rule itself. This allocation rule is defined as follows.

Definition 2.6.1 Let $n \in \mathbb{N}$. Then, $K^n : \mathcal{R} \to \mathbb{R}^N$ is given by

$$K^n(R) = \frac{1}{|\mathcal{P}^n|} \sum_{P \in \mathcal{P}^n} A^P(R), \qquad \text{for all } R \in \mathcal{R}, \tag{2.26}$$

where A^P is defined in Definition 2.5.2.

For a given n, one can rewrite the definition of K^n to the standard Aumann-Shapley method, which is proposed by Moulin (1995) for a given discrete production problem with a continuously differentiable production function. If $n = 1$, this allocation rule equals the Shapley value, i.e., $K^1 = S$. The asymptotic behavior of $K^n(R)$ when $n \to \infty$ (or, equivalently, when the grid size converges to 0) is a central topic of this chapter. Next, we rewrite $K^n(R)$ as a weighted sum of marginal contributions over all participation profiles on the grid G^n.

Proposition 2.6.2 *Let $R \in \mathcal{R}$ and $n \in \mathbb{N}$. Then, we have for all $i \in N$ that*

$$K_i^n(R) = \sum_{\lambda \in G^n : \lambda_i < 1} t^n(\lambda) p_i^n(\lambda) [r(\lambda + (1/n) \cdot e_i) - r(\lambda)], \tag{2.27}$$

where

$$t^n(\lambda) = \frac{\prod_{j \in N} \begin{pmatrix} n \\ n\lambda_j \end{pmatrix}}{\begin{pmatrix} |N|n \\ |N|n\bar{\lambda} \end{pmatrix}} \tag{2.28}$$

and

$$p_i^n(\lambda) = \frac{1 - \lambda_i}{\sum_{j \in N}(1 - \lambda_j)}, \tag{2.29}$$

for all $\lambda \in G^n \backslash \{e_N\}$,

$$\bar{\lambda} = \frac{1}{|N|} \sum_{i \in N} \lambda_i, \qquad \text{for all } \lambda \in \mathbb{R}^N, \tag{2.30}$$

and where the risk capital function r is defined in (2.17).

The function $t^n(\lambda)$ represents the probability that λ lies on a path, if we randomly select a path from $\mathcal{P}^n$ according to the discrete uniform distribution. Moreover, $p_i^n(\lambda)$ is the conditional probability that $\lambda + (1/n) \cdot e_i$ lies on a path, provided that the path passes through λ. So, in order to compute $K_i^n(R)$, each marginal contribution $r(\lambda + (1/n)\cdot e_i) - r(\lambda)$ is multiplied by the probability that both λ and $\lambda + (1/n)\cdot e_i$ are on a path.

We show in the sequel that $\lim_{n\to\infty} K^n(R)$ exists if $R \in \mathcal{R}$. This enables us to define the Weighted Aumann-Shapley value $WAS : \mathcal{R} \to \mathbb{R}^N$ by

$$WAS(R) = \lim_{n\to\infty} K^n(R), \qquad \text{for all } R \in \mathcal{R}. \tag{2.31}$$

Moreover, we show that this allocation rule satisfies $WAS(R) = AS(R)$ in case the risk capital function r is partially differentiable in e_N. So, K^1 is the Shapley value, and $\lim_{n\to\infty} K^n$ the Aumann-Shapley value, if existent.

Moreover, in case of non-differentiability, we show that the Weighted Aumann-Shapley value is a weighted average of standard Aumann-Shapley values of "nearby" risk capital allocation problems where the risk capital function is partially differentiable.

2.6.2 Convergence

In this subsection, we provide our main result of this chapter. We show that the Weighted Aumann-Shapley value exists, and provide a closed form solution. We use the following notation throughout this subsection:

- We use the Bachmann-Landau notation. Let $f, g : \mathbb{N} \to \mathbb{R}$ be two real-valued functions. Then, we write $f(n) = \mathcal{O}(g(n))$ if there is a $K > 0$ such that $|f(n)| \le K|g(n)|$ for every $n \in \mathbb{N}$. If $f : \mathbb{N} \to \mathbb{R}$ is such that $f(n) = \mathcal{O}(n^{-p})$ for every $p > 0$, we write $f(n) = \mathcal{O}(n^{-\infty})$. Moreover, if $g : \mathbb{R}_{++} \to \mathbb{R}$ is such that there is a $K > 0$ such that $|g(\varepsilon)| \le K\varepsilon$ for every $\varepsilon > 0$, we write $g(\varepsilon) = \mathcal{O}(\varepsilon)$. Here, $\mathbb{R}_{++} = (0, \infty)$ is the set of all positive, real numbers.
- Let $f : \mathbb{R}_{++} \times \mathbb{N} \to \mathbb{R}$ and $g : \mathbb{N} \to \mathbb{R}$. Then, we write $f(\varepsilon, n) = \mathcal{O}^\varepsilon(g(n))$ if for every $\varepsilon > 0$, there is a $K_\varepsilon > 0$ such that $|f(\varepsilon, n)| \le K_\varepsilon |g(n)|$ for all $n \in \mathbb{N}$. This notation is an extension of the standard Bachmann-Landau notation.
- For all $\lambda \in \mathbb{R}^N$, we write $\|\lambda\| = \sqrt{\sum_{i\in N} \lambda_i^2}$ as the Euclidean norm of λ.
- We define the set of participation profiles that are not nearby $\lambda = e_\emptyset$ and e_N as follows. For all $n \in \mathbb{N}$ and $\varepsilon > 0$, we define

  $$G_\varepsilon = \left\{\lambda \in [0,1]^N : \varepsilon \le \bar{\lambda} \le 1 - \varepsilon\right\},$$

  and

  $$G_\varepsilon^n = G^n \cap G_\varepsilon,$$

  where $\bar{\lambda}$ is defined in (2.30).

- We define D^d as the set of participation profiles in the d-environment of the diagonal, i.e., for all $d > 0$, we have

$$D^d = \{\lambda \in [0,1]^N : \|\lambda - \bar{\lambda} \cdot e_N\| < d\}.$$

Moreover, we define for all $n \in \mathbb{N}$ the set

$$D(n) = D^{d_n}, \qquad \text{where } d_n = n^{-\frac{1}{2}+\frac{1}{8|N|}}.$$

In Figure 2.3, we provide an illustration of the sets G_ε and D^d in case of two divisions. We only consider participation profiles in G_ε for an arbitrary choice of $\varepsilon > 0$. In the following proposition, we approximate the weight functions t^n and p_i^n that are defined in Proposition 2.6.2. For large n, we obtain expressions with a nice interpretation, which are obtained using, among others, several Taylor and Stirling approximations. We get that $t^n(\lambda)$ is exponentially small for λ away from the diagonal.

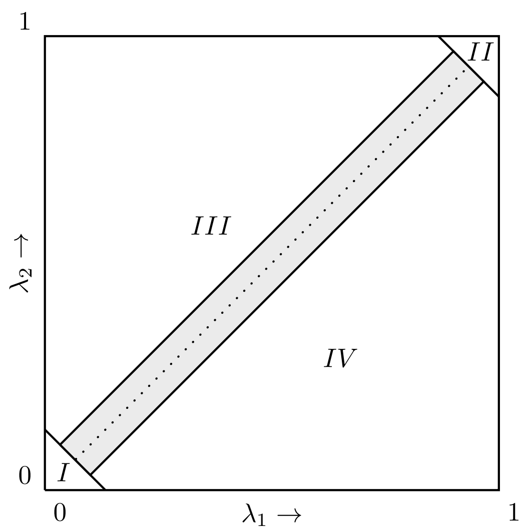


Figure 2.3: The shaded set is a set of participation profiles with substantial aggregate contribution to the Weighted Aumann-Shapley value (see Lemma 2.A.11 and Lemma 2.A.12) in case $|N| = 2$. Here, $I \cup II = [0,1]^N \backslash G_\varepsilon$ and $III \cup IV = G_\varepsilon \backslash D(n)$ for an arbitrary choice of $\varepsilon > 0$ and $n \in \mathbb{N}$.

Proposition 2.6.3 *Let $i \in N$, and define $Dom = \{(\varepsilon, n, \lambda) : \varepsilon > 0, n \in \mathbb{N}, \lambda \in G_\varepsilon^n\}$. Then, we have*

$$t^n(\lambda) = \begin{cases} \left(e^{-c(\bar{\lambda})n\|\lambda-\bar{\lambda}\cdot e_N\|^2}\right) b(n,\bar{\lambda})[1+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}})], & \text{if } \lambda \in D(n), \quad (2.32)\\ \mathcal{O}^\varepsilon(n^{-\infty}), & \text{if } \lambda \notin D(n), \quad (2.33)\end{cases}$$

and

$$p_i^n(\lambda) = \begin{cases} \frac{1}{|N|}[1+\mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}})], & \textit{if } \lambda \in D(n), \quad (2.34) \\ \mathcal{O}(1), & \textit{if } \lambda \notin D(n), \quad (2.35) \end{cases}$$

for all $(\varepsilon, n, \lambda) \in Dom$*, where*

$$c(\bar{\lambda}) = \frac{1}{2\bar{\lambda}(1-\bar{\lambda})} > 0 \tag{2.36}$$

and

$$b(n,\bar{\lambda}) = (2\pi n)^{\frac{1}{2}(1-|N|)}\sqrt{|N|}(\bar{\lambda}(1-\bar{\lambda}))^{\frac{1}{2}(1-|N|)}. \tag{2.37}$$

For large n, we get that $t^n(\lambda)$ only depends on λ via $\bar{\lambda}$ and $\|\lambda - \bar{\lambda}\cdot e_N\|$ and that $p_i^n(\lambda)$ is symmetric close to the diagonal. For a given $n \in \mathbb{N}$ and $\bar{\lambda} \in \{0, \frac{1}{n}, \ldots, 1\}$, the function $b(n,\bar{\lambda})$ is approximately the probability that a path goes through the diagonal (i.e., through $\bar{\lambda}\cdot e_N$) and $c(\bar{\lambda})$ indicates a speed at which $t^n(\lambda)$ converges to zero for participation profiles away from the diagonal. The function $t^n(\lambda)$ is exponentially small in n if λ is not nearby to the diagonal, i.e., $\lambda \notin D(n)$. Moreover, $p_i^n(\lambda)$ is bounded. Therefore, only participation profiles very close to the diagonal are relevant for K^n if n converges to infinity.

The function $h^n : [0,1]^N \backslash \{e_\emptyset, e_N\} \to \mathbb{R}_{++}$ is given by

$$h^n(\lambda) = \left(e^{-c(\bar{\lambda})n\|\lambda-\bar{\lambda}\cdot e_N\|^2}\right) b(n,\bar{\lambda})\frac{1}{|N|}, \tag{2.38}$$

for all $\lambda \in [0,1]^N \backslash \{e_\emptyset, e_N\}$ and $n \in \mathbb{N}$, where $c(\bar{\lambda})$ is defined in (2.36) and $b(n,\bar{\lambda})$ in (2.37). From Proposition 2.6.3, we get that

$$t^n(\lambda)p_i^n(\lambda) = h^n(\lambda)[1+\mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}})], \tag{2.39}$$

for all $(\varepsilon, n, \lambda) \in Dom$ such that $\lambda \in D(n)$.

In Lemma 2.A.11 and Lemma 2.A.12, we show that all participation profiles in G^n that are not close to the diagonal or that are nearby $e_\emptyset$ or e_N have a negligible aggregate contribution to K^n if n converges to infinity. In case of two divisions, we illustrate these participation profiles in Figure 2.3. We obtain that all participation profiles in K_m for $m \notin \{1, \ldots, p^*\}$ have a negligible aggregate contribution.

From Corollary 2.4.6, we get that the risk capital function r is almost everywhere partially differentiable. In Lemma 2.A.16 in Appendix 2.A, we extent this result by showing that participation profiles in a $\frac{1}{n}$-environment of participation profiles where r is non-differentiable have a negligible aggregate contribution for large n as well. For all other risk profiles λ, we obtain from

(2.23) that

$$r(\lambda + (1/n) \cdot e_i) - r(\lambda) = \frac{1}{n} E_{\mathbb{Q}_m}[X_i],$$

for all $i \in N$ and $\lambda \in G^n \cap K_m$ such that $\lambda_i < 1$ and $m \in \{1, \ldots, p\}$. So, for these risk profiles, we know the marginal effect exactly.

Proposition 2.6.4 *Let $R \in \mathcal{R}$. Then, for all $i \in N$ we have*

$$K_i^n(R) = \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \phi_m^{n,\varepsilon} + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}}),$$

where

$$\phi_m^{n,\varepsilon} = \frac{1}{n} \sum_{\lambda \in G_\varepsilon^n \cap D(n) \cap K_m} h^n(\lambda), \tag{2.40}$$

and p^*, $\mathbb{Q}_m$ and K_m are defined in Definition 2.4.3.

The expression $\phi_m^{n,\varepsilon}$ is a weight for a gradient of the risk capital function r "nearby" the diagonal, namely $(E_{\mathbb{Q}_m}[X_i])_{i \in N}$. Next, we show that we can replace this weight by an expression that has a geometric interpretation, and is not dependent on n or ε anymore. This result is obtained by replacing the sum in (2.40) by an integral (see Lemma 2.A.18 and Lemma 2.A.19) and, thereafter, solving this integral.

Proposition 2.6.5 *For all $R \in \mathcal{R}$, it holds that*

$$K_i^n(R) = \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \phi_m + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}}), \qquad \text{for all } i \in N,$$

where we define

$$S = \left\{ z \in \mathbb{R}^N : \sum_{i \in N} z_i = 0, \|z\| = 1 \right\}$$
$$S_m = \left\{ z \in S : f_{\mathbb{Q}_m}(z) = \max_{\ell \in \{1,\ldots,p^*\}} f_{\mathbb{Q}_\ell}(z) \right\},$$

and

$$\phi_m = \frac{\mu(S_m)}{\mu(S)},$$

for all $m \in \{1, \ldots, p^*\}$, where μ is the surface area measure on S and the function $f_{\mathbb{Q}}$ is defined in Definition 2.4.1.

Remark that from Lemma 2.4.5, we get

$$\sum_{m=1}^{p^*} \phi_m = 1. \tag{2.41}$$

Recall the definition of the Weighted Aumann-Shapley value in (2.31). Using Proposition 2.6.5, we next show that this value exists and, moreover, we provide a closed form expression. This result is given in the following theorem.

Theorem 2.6.6 *For all $R \in \mathcal{R}$, we have that $WAS(R) = \lim_{n\to\infty} K^n(R)$ exists and, moreover, it is given by*

$$WAS_i(R) = \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i]\phi_m, \qquad \text{for all } i \in N,$$

with ϕ_m as defined in Proposition 2.6.5.

Theorem 2.6.6 shows that the Weighted Aumann-Shapley value is a convex combination of regular Aumann-Shapley values of "nearby" risk capital allocation problems. The weight ϕ_m has a geometric interpretation, as we show in the next constructive example.

Example 2.6.7 In this example, we discuss a case with three divisions. Let $N = \{1, 2, 3\}$, $\Omega = \{\omega_1, \ldots, \omega_5\}$, and $\mathbb{P}(\{\omega\}) = \frac{1}{5}$ for all $\omega \in \Omega$. Moreover, let the risk capital allocation problem given by $R = ((X_i)_{i\in N}, \rho^{ES}_{0.1}) \in \mathcal{R}$ such that

$$X_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, X_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 2 \end{bmatrix}, \text{ and } X_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

Next, we show the value of ϕ_m. We pick out the shaded triangle $T = \{\lambda \in [0, 1]^N : \bar{\lambda} = \frac{1}{3}\}$ as in Figure 2.4.

Let $\mathbb{Q}_m = (1_{\omega_m}, 0_{-\omega_m})$ for $m \in \{1, \ldots, 5\}$. We display the set $(T \cap K_m)_{m\in\{1,\ldots,p\}}$ in Figure 2.5.a, and we immediately see that $p^* = 4$. Then, the fraction ϕ_m corresponds with the normalized angle that the set $T \cap K_m$ forms in point $\lambda = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. We depict a sufficiently small environment $T \cap D^d$ around this point (as shown in Figure 2.5.a), and display this set in Figure 2.5.b. We obtain $\phi_1 = \frac{1}{6}$, $\phi_2 = \frac{1}{6}$, $\phi_3 = \frac{1}{3}$, and $\phi_4 = \frac{1}{3}$. Note that, instead of the triangle T with $\bar{\lambda} = \frac{1}{3}$, we could have depicted every triangle where $0 < \bar{\lambda} < 1$. Then, we obtain from Theorem 2.6.6 that

$$WAS(R) = \left(\frac{1}{2}, 0, \frac{1}{2}\right).$$

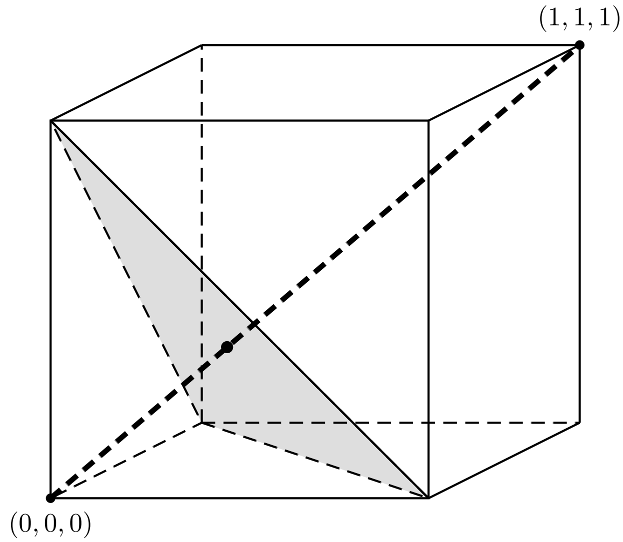
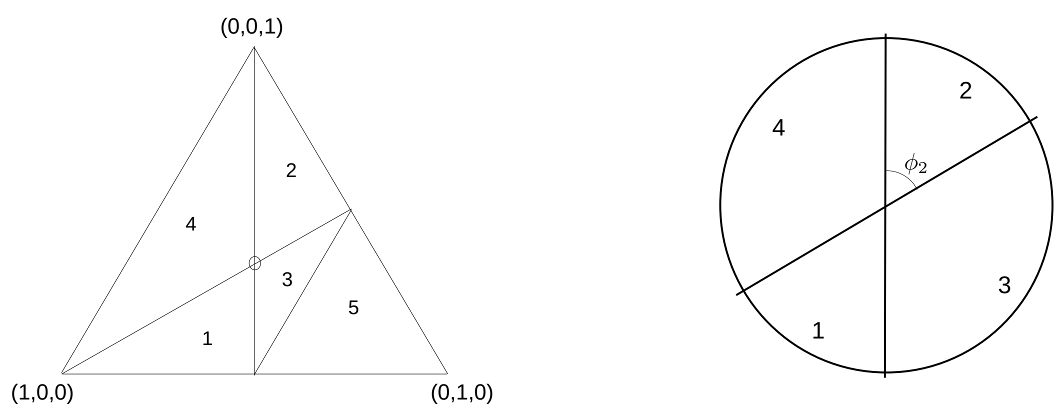


Figure 2.4: A representation of the path via the diagonal for $N = \{1, 2, 3\}$ in $[0, 1]^N$. The shaded triangle is the set $T = \left\{\lambda \in [0, 1]^N : \bar{\lambda} = \frac{1}{3}\right\}$.



(a) Simplex T of three-division unit-cube. The number m reflects the area $T \cap K_m$ for all $m \in \{1, 2, 3, 4, 5\}$.

(b) A sufficiently small environment of the diagonal, where ϕ_2 corresponds with the normalized angle of $T \cap K_2$ in $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$.

Figure 2.5: Illustration of the Weighted Aumann-Shapley value corresponding to Example 2.6.7.

So, Division 2, which holds the portfolio with the highest expected loss, gets assigned the lowest risk capital allocation. This is due to large hedge benefits as the loss in state ω_5 is high for this portfolio, while this state is the best-case scenario for the firm. $\triangledown$

Remark We can prove Theorem 2.6.6 using a diagonal width $d_n = n^{-\frac{1}{2}+\delta}$ for all $\delta \in \left(0, \frac{1}{2(|N|+2)}\right)$.

Then, we can adjust Proposition 2.6.5 as follows:

$$K_i^n(R) = \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i]\phi_m + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{2}+\delta(|N|+2)}), \qquad \text{for all } R \in \mathcal{R} \text{ and } i \in N.$$

In this subsection, we depicted $\delta = \frac{1}{8|N|}$ to avoid tedious results.

2.6.3 Properties of the Weighted Aumann-Shapley value

In this subsection, we show some properties of the Weighted Aumann-Shapley value. As stated earlier, the Weighted Aumann-Shapley value equals the Aumann-Shapley if the risk capital function r is differentiable in e_N. This follows directly from Theorem 2.6.6, and is shown in the following corollary.

Corollary 2.6.8 *For all $R \in \mathcal{R}'$, it holds that*

$$WAS(R) = AS(R).$$

Corollary 2.6.8 is stated for the Aumann-Shapley mechanism in cost sharing by Moulin (1995). Moulin, however, typically requires the function r to be continuously differentiable, whereas this typically does not need to be satisfied in case that ρ is a coherent risk measure.

We next show a range of values of ϕ_m.

Proposition 2.6.9 *For all $R \in \mathcal{R}$ and $m \in \{1, \ldots, p^*\}$, we have*

$$\phi_m \in \left[0, \frac{1}{2}\right] \cup \{1\}.$$

The range of ϕ_m in Proposition 2.6.9 is tight for $|N| > 2$, i.e., for every $c \in [0, \frac{1}{2}] \cup \{1\}$, one can construct a risk capital allocation problem such that $\phi_m = c$ for an $m \in \{1, \ldots, p^*\}$.

Next, we generalize the property that the Aumann-Shapley value, if existent, is in the single-valued fuzzy core (see Theorem 2.3.4). We namely show that the Weighted Aumann-Shapley is always in the fuzzy core.

Proposition 2.6.10 *The Weighted Aumann-Shapley value satisfies Fuzzy Core Selection on $\mathcal{R}$.*

All properties that are satisfied by path based allocation rules are also satisfied by the Weighted Aumann-Shapley value. This holds as this rule is the average over allocation rules based on a path. Hence, according to Theorem 2.5.4, the Weighted Aumann-Shapley value satisfies *Translation Invariance*, *Scale Invariance* and *Monotonicity* on $\mathcal{R}$.

Let $\nabla^+ r(e_N)$ be the right derivative of the risk capital function r in e_N and $\nabla^- r(e_N)$ the corresponding left derivative. We straightforwardly obtain from Theorem 2.6.6 the following result for risk capital allocation problems with two divisions.

Corollary 2.6.11 *If $|N| = 2$ and $R \in \mathcal{R}$, we have*

$$WAS(R) = \frac{1}{2}\nabla^{+}r(e_N) + \frac{1}{2}\nabla^{-}r(e_N).$$

One can easily verify that $\nabla^{+}r(e_N) = (\max\{E_{\mathbb{Q}}[X_i] : \mathbb{Q} \in Q^*(\rho)\})_{i\in N}$ and $\nabla^{-}r(e_N) = (\min\{E_{\mathbb{Q}}[X_i] : \mathbb{Q} \in Q^*(\rho)\})_{i\in N}$, where $Q^*(\rho)$ is defined in (2.19).

Remark We note that we only need positive homogeneity and piecewise linearity of the risk capital function r for Theorem 2.6.6 to hold, where $E_{\mathbb{Q}_m}[X_i]$ is then replaced by $\frac{\partial r}{\partial \lambda_i}(\lambda)$ for $\lambda \in K_m$ and for $i \in N$ with obvious meaning of notation. On this class of allocation problems, also Proposition 2.6.9, Proposition 2.6.10 and Corollary 2.6.11 hold. For only a piecewise linear function r, one can also deduce a closed form expression of K^n in Theorem 2.6.6 based on derivatives of r in participation profiles along the diagonal.

2.7 Conclusion

This chapter considers the allocation problem that arises when the total risk capital withheld by a firm needs to be divided over several portfolios within the firm. We propose a generalization of the Aumann-Shapley value for risk capital allocation problems. The Aumann-Shapley value receives considerable attention in the literature. It is well-known that this allocation rule requires differentiability of the fuzzy game for existence.

Our rule is also well-defined in case a differentiability condition is not satisfied. We introduce this allocation rule inspired by the Shapley value in a fuzzy setting. It follows a much weaker asymptotic approach than the one proposed by Aumann and Shapley (1974). The asymptotic approach of Aumann and Shapley (1974) is not valid for fuzzy games related to risk capital allocation problems. We take a grid on a fuzzy participation set, define paths on this grid, and construct an allocation rule based on each path. Then, we show that the limit of the average over these allocations exists, when the grid size converges to zero. We define the Weighted Aumann-Shapley value as this limit. We provide an explicit formula for this allocation rule, which has a geometric interpretation. Moreover, it satisfies some properties which are known to hold for the Aumann-Shapley value. It is still an open question whether our value coincides with the Mertens value.

2.A Proofs

Proof of Proposition 2.2.3: Let Q be the generating probability measure set of ρ that is defined in (2.2), i.e.,

$$Q = \{\mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F}) : \mathbb{Q}(A) \geq v^{\rho}(A) \text{ for all } A \in \mathcal{F}\},$$

where $\upsilon^\rho : \mathcal{F} \to \mathbb{R}_+$ is submodular, $\upsilon^\rho(\emptyset) = 0$ and $\upsilon^\rho(\Omega) = 1$. Note that as the state space Ω is finite, the σ-algebra $\mathcal{F}$ is finite as well. Because $\mathcal{F}$ is finite, Q is defined via a finite number of linear inequalities on $[0,1]^\Omega$. So, Q is a polytope. Let $\widetilde{Q}$ be the finite collection of extreme points of this polytope. Because $\mathbb{Q} \to E_\mathbb{Q}[X]$ is a linear map on Q for every $X \in L^\infty$, (2.1) is a linear programming problem and, therefore, we have

$$\rho(X) = \sup\left\{E_\mathbb{Q}[X] : \mathbb{Q} \in Q\right\} = \max\left\{E_\mathbb{Q}[X] : \mathbb{Q} \in \widetilde{Q}\right\}, \qquad \text{for all } X \in L^\infty.$$

Hence, $\rho(X)$ equals the maximum of all expectations of X under the probability measures in $\widetilde{Q}$. Hence, $\widetilde{Q}$ is a generating probability measure set. This concludes the proof. □

Proof of Proposition 2.4.2: For all $R \in \mathcal{R}$, we have

$$r(\lambda) = \max\left\{E_\mathbb{Q}\left[\sum_{i\in N}\lambda_i X_i\right] : \mathbb{Q} \in Q(\rho)\right\} \tag{2.42}$$

$$= \max\left\{\sum_{i\in N}\lambda_i E_\mathbb{Q}[X_i] : \mathbb{Q} \in Q(\rho)\right\} \tag{2.43}$$

$$= \max\{f_\mathbb{Q}(\lambda) : \mathbb{Q} \in Q(\rho)\}, \tag{2.44}$$

for all $\lambda \in [0,1]^N$. This concludes the proof. □

Definition 2.A.1 Let $R \in \mathcal{R}$. Then, the set $L(R)$ is given by

$$L(R) = \left\{\lambda \in [0,1]^N : \text{there exists a unique } m \in \{1,\ldots,p\} \text{ such that } \lambda \in K_m\right\},$$

where $K_1,\ldots,K_p$ are defined in Definition 2.4.3.

The set $L(R)$ is open on $[0,1]^N$ and, therefore, $L(R)$ is a the set of participation profiles where risk capital function r is locally linear.

Proof of Lemma 2.4.5: We obtain for all $\ell, m \in \{1,\ldots,p\}$ that

$$K_\ell \cap K_m = \left\{\lambda \in [0,1]^N : r(\lambda) = f_{\mathbb{Q}_\ell}(\lambda) = f_{\mathbb{Q}_m}(\lambda)\right\} \tag{2.45}$$

$$\subseteq \left\{\lambda \in [0,1]^N : f_{\mathbb{Q}_\ell}(\lambda) = f_{\mathbb{Q}_m}(\lambda)\right\} \tag{2.46}$$

$$= \left\{\lambda \in [0,1]^N : \sum_{i\in N}\lambda_i\left(E_{\mathbb{Q}_\ell}[X_i] - E_{\mathbb{Q}_m}[X_i]\right) = 0\right\}. \tag{2.47}$$

If $E_{\mathbb{Q}_\ell}[X_i] = E_{\mathbb{Q}_m}[X_i]$ for all $i \in N$, we have $K_\ell = K_m$ which implies $\ell = m$. So, the set $K_\ell \cap K_m$ is a (possibly empty) subset of a hyperplane passing through $\lambda = e_\emptyset$ for all $\ell, m \in \{1, \ldots, p\}$ such that $\ell \neq m$. We have by construction that

$$[0,1]^N \backslash L(R) = \bigcup_{\ell,m\in\{1,\ldots,p\}:\ell\neq m} (K_\ell \cap K_m), \qquad \text{for all } R \in \mathcal{R}. \tag{2.48}$$

From this it follows that the collection of profiles where the risk capital function r is not partially differentiable is a subset of the collection of a finite number of hyperplanes passing trough $\lambda = e_\emptyset$. □

Proof of Proposition 2.5.3: Let $R \in \mathcal{R}$. The result follows directly from

$$\sum_{i\in N} A_i^P(R) = \sum_{i\in N} \sum_{k=0}^{|N|n-1} [r(P(k+1)) - r(P(k))] \cdot \mathbb{1}_{i(P,k)=i} \tag{2.49}$$

$$= \sum_{k=0}^{|N|n-1} [r(P(k+1)) - r(P(k))] \sum_{i\in N} \mathbb{1}_{i(P,k)=i} \tag{2.50}$$

$$= \sum_{k=0}^{|N|n-1} [r(P(k+1)) - r(P(k))] \tag{2.51}$$

$$= r(P(|N|n)) - r(P(0))$$

$$= r(e_N), \tag{2.52}$$

where $\mathbb{1}_{i(P,k)=i} = 1$ if $i(P,k) = i$ and $\mathbb{1}_{i(P,k)=i} = 0$ otherwise. Here, (2.49) follows from Definition 2.5.2, (2.50) follows by interchanging the summations, (2.51) follows from the fact that there is precisely one $i \in N$ such that $i(P,k) = i$ for all $k \in \{0, \ldots, |N|n-1\}$, and (2.52) follows from Definition 2.5.1.1. This concludes the proof. □

Proof of Theorem 2.5.4: We start with showing the property *Translation Invariance*. Let $P \in \mathcal{P}^n$, $j \in N$, $R = ((X_i)_{i\in N}, \rho) \in \mathcal{R}$ and $\hat{R} = ((\hat{X}_i)_{i\in N}, \rho) \in \mathcal{R}$ such that $(\hat{X}_i)_{i\in N} = (X_j + ce_\Omega, X_{-j})$ for some $c \in \mathbb{R}$. Moreover, we let the risk capital function r (resp. $\hat{r}$), defined in (2.17), correspond with R (resp. $\hat{R}$). Then, we get

$$\hat{r}(\lambda) = \rho\left(\sum_{i\in N} \lambda_i \hat{X}_i\right)$$

$$= \rho\left(\sum_{i\in N} \lambda_i X_i + c\lambda_j \cdot e_\Omega\right)$$

$$= \rho\left(\sum_{i\in N} \lambda_i X_i\right) + c\lambda_j \tag{2.53}$$

$$= r(\lambda) + c\lambda_j, \tag{2.54}$$

for all $\lambda \in [0,1]^N$, where (2.53) follows from *Translation Invariance* of ρ. We get

$$A^P(\hat{R}) = \sum_{k=0}^{|N|n-1} [\hat{r}(P(k+1)) - \hat{r}(P(k))] \cdot e_{i(P,k)} \tag{2.55}$$

$$= \sum_{k=0}^{|N|n-1} [r(P(k+1)) + cP_j(k+1) - r(P(k)) - cP_j(k)] \cdot e_{i(P,k)} \tag{2.56}$$

$$= A^P(R) + c \sum_{k=0}^{|N|n-1} [P_j(k+1) - P_j(k)] \cdot e_{i(P,k)} \tag{2.57}$$

$$= A^P(R) + c \sum_{k=0}^{|N|n-1} [P_j(k+1) - P_j(k)] \cdot e_j \tag{2.58}$$

$$= A^P(R) + c[P_j(|N|n) - P_j(0)] \cdot e_j$$

$$= A^P(R) + c \cdot e_j, \tag{2.59}$$

where $P_j(k)$ is the j-th element of $P(k)$. Here, (2.55) follows from Definition 2.5.2, (2.56) follows from (2.54), (2.57) follows from Definition 2.5.2, (2.58) follows from $P_j(k+1) - P_j(k) = 0$ if $i(P,k) \neq j$ (see (2.25)), and (2.59) follows from Definition 2.5.1.1. This concludes the proof of *Translation Invariance*.

The proof of *Scale Invariance* is similar to the proof of *Translation Invariance*.

Next, we show *Monotonicity*. Let the risk measure ρ be non-decreasing. This implies $r(P(k+1)) - r(P(k)) \geq 0$ for all $k \in \{0,\ldots,|N|n-1\}$. Then, from Definition 2.5.2 it follows directly that $A^P(R) \geq 0$. This concludes the proof of *Monotonicity*. □

Proof of Proposition 2.6.2: In this proof, we use the following notation. The set $\widetilde{G}_k^n$ is given by

$$\widetilde{G}_k^n = \left\{\lambda \in G^n : \sum_{i\in N} \lambda_i = \frac{k}{n}\right\}, \qquad \text{for all } n \in \mathbb{N} \text{ and } k \in \{0,\ldots,|N|n\}. \tag{2.60}$$

The set $\widetilde{G}_k^n$ consists of all participation profiles on the grid where the sum of the coordinates is constant. Note that we have

$$\widetilde{G}_k^n = \{P(k) : P \in \mathcal{P}^n\}, \qquad \text{for all } n \in \mathbb{N} \text{ and } k \in \{0,\ldots,|N|n\}. \tag{2.61}$$

Next, we show (2.27). Expression (2.26) of $K^n(R)$ can be rewritten as

$$\begin{aligned}
K^n(R) &= \frac{1}{|\mathcal{P}^n|}\sum_{P\in\mathcal{P}^n} A^P(R) && (2.62)\\
&= \frac{1}{|\mathcal{P}^n|}\sum_{P\in\mathcal{P}^n}\sum_{k=0}^{|N|n-1}[r(P(k+1))-r(P(k))]\cdot e_{i(P,k)} && (2.63)\\
&= \sum_{k=0}^{|N|n-1}\sum_{P\in\mathcal{P}^n}\frac{1}{|\mathcal{P}^n|}[r(P(k+1))-r(P(k))]\cdot e_{i(P,k)}, && (2.64)
\end{aligned}$$

where (2.62) follows from Definition 2.6.1, and (2.63) follows from Definition 2.5.2. Let $i \in N$. Then, we obtain

$$\begin{aligned}
K_i^n(R) &= \sum_{k=0}^{|N|n-1}\sum_{P\in\mathcal{P}^n:i(P,k)=i}\frac{1}{|\mathcal{P}^n|}[r(P(k+1))-r(P(k))] && (2.65)\\
&= \sum_{k=0}^{|N|n-1}\sum_{P\in\mathcal{P}^n:i(P,k)=i}\frac{1}{|\mathcal{P}^n|}[r(P(k)+(1/n)\cdot e_i)-r(P(k))] && (2.66)\\
&= \sum_{k=0}^{|N|n-1}\sum_{\lambda\in\widetilde{G}_k^n}\sum_{\substack{P\in\mathcal{P}^n:\\ i(P,k)=i,P(k)=\lambda}}\frac{1}{|\mathcal{P}^n|}[r(\lambda+(1/n)\cdot e_i)-r(\lambda)] && (2.67)\\
&= \sum_{k=0}^{|N|n-1}\sum_{\lambda\in\widetilde{G}_k^n:\lambda_i<1}[r(\lambda+(1/n)\cdot e_i)-r(\lambda)]\sum_{\substack{P\in\mathcal{P}^n:\\ i(P,k)=i,P(k)=\lambda}}\frac{1}{|\mathcal{P}^n|} && (2.68)\\
&= \sum_{k=0}^{|N|n-1}\sum_{\lambda\in\widetilde{G}_k^n:\lambda_i<1}[r(\lambda+(1/n)\cdot e_i)-r(\lambda)]\,t^n(\lambda)p_i^n(\lambda) && (2.69)\\
&= \sum_{\lambda\in G^n:\lambda_i<1}[r(\lambda+(1/n)\cdot e_i)-r(\lambda)]\,t^n(\lambda)p_i^n(\lambda), && (2.70)
\end{aligned}$$

where we define

$$t^n(\lambda) = \frac{\left|\left\{P\in\mathcal{P}^n : P\left(|N|n\bar{\lambda}\right)=\lambda\right\}\right|}{|\mathcal{P}^n|}$$

as the fraction of paths in $\mathcal{P}^n$ that pass through λ, and

$$p_i^n(\lambda) = \frac{\left|\left\{P\in\mathcal{P}^n : P\left(|N|n\bar{\lambda}\right)=\lambda, i\left(P,|N|n\bar{\lambda}\right)=i\right\}\right|}{\left|\left\{P\in\mathcal{P}^n : P\left(|N|n\bar{\lambda}\right)=\lambda\right\}\right|}$$

as the fraction of the paths in $\mathcal{P}^n$ passing through λ, that pass through $\lambda+\frac{1}{n}\cdot e_i$ as well. Here, (2.65) follows from (2.64), (2.66) follows from (2.25), (2.67) follows from (2.61), (2.68) follows from the fact that if $k\in\{0,\ldots,|N|n-1\}$ and $\lambda\in\widetilde{G}_k^n$ are such that $\lambda_i=1$ then no path $P\in\mathcal{P}^n$

exists with $i(P,k) = i$ and $P(k) = \lambda$, (2.69) follows from the fact that if $k \in \{0, \ldots, |N|n-1\}$ and $\lambda \in \widetilde{G}^n_k$ are such that $P(k) = \lambda$ then $k = |N|n\bar{\lambda}$, and (2.70) follows from the fact that $\bigcup_{k=1}^{|N|n-1} G^n_k = G^n$ and $G^n_{k_1} \cap G^n_{k_2} = \emptyset$ if $k_1 \neq k_2$.

Next, we show (2.28). Any path can be regarded as an ordered sequence of $|N|n$ steps, where for every division $i \in N$ precisely n steps are made in the direction of division i. Hence,

$$|\mathcal{P}^n| = \frac{(|N|n)!}{(n!)^{|N|}}. \tag{2.71}$$

Let $\lambda \in G^n \backslash \{e_N\}$. The number of paths P in $\mathcal{P}^n$ such that $P\left(|N|n\bar{\lambda}\right) = \lambda$ is given by

$$\left|\left\{P \in \mathcal{P}^n : P\left(|N|n\bar{\lambda}\right) = \lambda\right\}\right| = \frac{\left(|N|n\bar{\lambda}\right)!\left(|N|n(1-\bar{\lambda})\right)!}{\prod_{j \in N}((n\lambda_j)!(n(1-\lambda_j))!)}. \tag{2.72}$$

Hence, one can verify that dividing (2.72) by (2.71) yields (2.28). Note that, keeping $\bar{\lambda}$ constant, the various values of $t^n(\lambda)$ constitute a density function of some multivariate hypergeometric distribution.

Finally, we show (2.29). The number of paths P in $\mathcal{P}^n$ with $P(|N|n\bar{\lambda}) = \lambda$ and $i(P, |N|n\bar{\lambda}) = i$ (i.e. passing through λ and $\lambda + (1/n) \cdot e_i$) is given by:

$$\left|\left\{P \in \mathcal{P}^n : P\left(|N|n\bar{\lambda}\right) = \lambda, i\left(P, |N|n\bar{\lambda}\right) = i\right\}\right| = \frac{(|N|n\bar{\lambda})!(|N|n(1-\bar{\lambda})-1)!}{\prod_{j \in N}(n\lambda_j)! \prod_{j \in N\backslash\{i\}}(n(1-\lambda_j))! \cdot (n(1-\lambda_i)-1)!}. \tag{2.73}$$

Dividing (2.73) by (2.72) yields (2.29) in a straightforward way. □

In the sequel of this appendix, we use the following definitions, notation, and properties:

- The function $g : \mathbb{R}_+ \to \mathbb{R}$ is given by

$$g(x) = \begin{cases} x\ln(x) & \text{if } x > 0, \\ 0 & \text{if } x = 0. \end{cases}$$

- The function $G : [0,1]^N \to \mathbb{R}$ is given by

$$G(\lambda) = |N| g\left(\bar{\lambda}\right) - \sum_{i \in N} g(\lambda_i) + |N| g\left(1-\bar{\lambda}\right) - \sum_{i \in N} g(1-\lambda_i), \qquad \text{for all } \lambda \in [0,1]^N. \tag{2.74}$$

- For all $\lambda \in [0,1]^N$, we define

$$N_1^\lambda = \{i \in N : \lambda_i > 0\} \text{ and } N_2^\lambda = \{i \in N : \lambda_i < 1\}. \tag{2.75}$$

- For $x, y \in \mathbb{R}$ we denote $[x; y]$ as the interval $[\min\{x, y\}, \max\{x, y\}]$, i.e., $[x; y] = [x, y]$ if $x \leq y$ and $[x; y] = [y, x]$ if $x > y$.

- Some arithmetic rules of the Bachmann-Landau notation are given by:

$$\begin{aligned} f(n) = \mathcal{O}(n^a), g(n) = \mathcal{O}(n^b) &\to f(n) + g(n) = \mathcal{O}(n^a), && \text{for all } a \geq b, \\ f(n) = \mathcal{O}(n^a), g(n) = \mathcal{O}(n^{-\infty}) &\to f(n) + g(n) = \mathcal{O}(n^a), && \text{for all } a \in \mathbb{R}, \\ f(n) = \mathcal{O}(n^a), g(n) = \mathcal{O}(n^b) &\to f(n)g(n) = \mathcal{O}(n^{a+b}), && \text{for all } a, b \in \mathbb{R}, \\ f(n) = \mathcal{O}(n^a) &\to f(n) = \mathcal{O}(n^b), && \text{for all } a \leq b. \end{aligned}$$

  Moreover, we have

$$f(n) = \mathcal{O}(n^a), g(n) = \mathcal{O}^{\varepsilon}(n^b) \to f(n) + g(n) = \mathcal{O}^{\varepsilon}(n^a), \qquad \text{for all } a \geq b.$$

- It is well known that for any $k \in \mathbb{R}$, $\delta > 0$ and $c \in (0, 1)$ the function $f : \mathbb{N} \to \mathbb{R}_{++}$, defined by $f(n) = n^k c^{n^\delta}$, is such that $f(n) = \mathcal{O}(n^{-\infty})$.

Lemma 2.A.2 *The function g is continuous and strictly convex, i.e., if $x, y \in \mathbb{R}_+$, $x \neq y$ and $\lambda \in (0, 1)$, then $g(\lambda x + (1 - \lambda)y) < \lambda g(x) + (1 - \lambda)g(y)$.*

Proof: Continuity of f follows from continuity of $x \to x \ln(x)$ for $x > 0$ and the fact that $\lim_{x \downarrow 0} x \ln(x) = 0$. Strict convexity follows from $g''(x) = \frac{1}{x} > 0$ for every $x > 0$. □

Lemma 2.A.3 *For the function G the following holds:*

1. *G is continuous;*

2. *$G(\lambda) \leq 0$ for all $\lambda \in [0, 1]^N$; moreover, $G(\lambda) = 0$ if and only if $\lambda_1 = \lambda_2 = \cdots = \lambda_{|N|}$;*

3. *for all $\lambda \in (0, 1)^N$, we have*

$$G(\lambda) = -c(\bar{\lambda})\|\lambda - \bar{\lambda} \cdot e_N\|^2 + R,$$

   where $|R| \leq \frac{1}{3}|N| \min\{\lambda_1, \ldots, \lambda_{|N|}, 1 - \lambda_1, \ldots, 1 - \lambda_{|N|}\}^{-2}\|\lambda - \bar{\lambda} \cdot e_N\|^3$ and c is defined in (2.36).

Proof:

1. This follows from continuity of g (Lemma 2.A.2).

2. This follows from strict convexity of g (Lemma 2.A.2).

3. Let $\lambda \in (0, 1)^N$ and $i \in N$. Then, there exists a $\xi_{i,1} \in [\lambda_i; \bar{\lambda}]$ such that

$$g(\lambda_i) = g(\bar{\lambda}) + g'(\bar{\lambda})(\lambda_i - \bar{\lambda}) + \frac{g''(\bar{\lambda})}{2}(\lambda_i - \bar{\lambda})^2 + \frac{g'''(\xi_{i,1})}{6}(\lambda_i - \bar{\lambda})^3 \tag{2.76}$$

$$= g(\bar{\lambda}) + (\ln(\bar{\lambda}) + 1)(\lambda_i - \bar{\lambda}) + \frac{1}{2\bar{\lambda}}(\lambda_i - \bar{\lambda})^2 - \frac{1}{6\xi_{i,1}^2}(\lambda_i - \bar{\lambda})^3, \tag{2.77}$$

where (2.76) follows from Taylor's theorem. Note that

$$\sum_{i \in N}(\lambda_i - \bar{\lambda}) = 0. \tag{2.78}$$

Then, summing the expression (2.77) of $g(\lambda_i)$ over all $i \in N$ yields

$$\sum_{i \in N} g(\lambda_i) = |N| g(\bar{\lambda}) + \frac{1}{2\bar{\lambda}}\|\lambda - \bar{\lambda} \cdot e_N\|^2 - \sum_{i \in N} \frac{1}{6\xi_{i,1}^2}(\lambda_i - \bar{\lambda})^3.$$

Similarly, we obtain

$$\sum_{i \in N} g(1 - \lambda_i) = |N| g(1 - \bar{\lambda}) + \frac{1}{2(1 - \bar{\lambda})}\|\lambda - \bar{\lambda} \cdot e_N\|^2 + \sum_{i \in N} \frac{1}{6\xi_{i,2}^2}(\lambda_i - \bar{\lambda})^3,$$

for some $\xi_{i,2} \in [1 - \lambda_i; 1 - \bar{\lambda}]$ for all $i \in N$. Now the upperbound of $|R|$ follows from $\xi_{i,1} \geq \min\{\lambda_1, \dots, \lambda_{|N|}\}$, $\xi_{i,2} \geq \min\{1 - \lambda_1, \dots, 1 - \lambda_{|N|}\}$ and $|(\lambda_i - \bar{\lambda})^3| \leq \|\lambda - \bar{\lambda} \cdot e_N\|^3$ for all $i \in N$.

□

Lemma 2.A.4 *Let $d, \varepsilon > 0$. Then, for all $\lambda \in G_\varepsilon \cap D^d$, we have*

$$\min\{\lambda_1, \dots, \lambda_{|N|}, 1 - \lambda_1, \dots, 1 - \lambda_{|N|}\} > \varepsilon - d.$$

Proof: Let $\lambda \in G_\varepsilon \cap D^d$. Since $|\lambda_i - \bar{\lambda}| \leq \|\lambda - \bar{\lambda} \cdot e_N\| < d$, we obtain $\lambda_i > \bar{\lambda} - d$ and $1 - \lambda_i > 1 - \bar{\lambda} - d$ for all $i \in N$. Moreover, we have $\varepsilon \leq \bar{\lambda} \leq 1 - \varepsilon$. Hence, we obtain $\lambda_i > \varepsilon - d$ and $1 - \lambda_i > \varepsilon - d$ for all $i \in N$. This concludes the proof. □

Lemma 2.A.5 *For all (n, λ) such that $n \in \mathbb{N}$ and $\lambda \in G^n \backslash \{e_\emptyset, e_N\}$, we have*

$$t^n(\lambda) = \left(e^{G(\lambda)}\right)^n (2\pi n)^{\frac{1}{2}(1 + |N| - |N_1^\lambda| - |N_2^\lambda|)} \sqrt{|N|} \frac{(\bar{\lambda}(1 - \bar{\lambda}))^{\frac{1}{2}}}{\prod_{i \in N_1^\lambda} \sqrt{\lambda_i} \prod_{i \in N_2^\lambda} \sqrt{1 - \lambda_i}}$$
$$\cdot \left[1 + \mathcal{O}\left(\frac{1}{n \min(\{\lambda_j : j \in N_1^\lambda\} \cup \{1 - \lambda_j : j \in N_2^\lambda\})}\right)\right], \tag{2.79}$$

where N_1^λ and N_2^λ are defined in (2.75).

Proof: Using (2.28), we obtain for all (n, λ) such that $n \in \mathbb{N}$ and $\lambda \in G^n \backslash \{e_\emptyset, e_N\}$ that

$$\begin{aligned} t^n(\lambda) &= \frac{\prod_{i \in N} \begin{pmatrix} n \\ n\lambda_i \end{pmatrix}}{\begin{pmatrix} |N|n \\ |N|n\bar{\lambda} \end{pmatrix}} \\ &= \frac{(n!)^{|N|}(|N|n\bar{\lambda})!(|N|n(1-\bar{\lambda}))!}{(|N|n)! \prod_{i \in N}[(n\lambda_i)!(n(1-\lambda_i))!]} \\ &= \frac{(n!)^{|N|}(|N|n\bar{\lambda})!(|N|n(1-\bar{\lambda}))!}{(|N|n)! \prod_{i \in N_1^\lambda}(n\lambda_i)! \prod_{i \in N_2^\lambda}(n(1-\lambda_i))!}. \end{aligned}$$

Taking the logarithm yields

$$\begin{aligned} \ln(t^n(\lambda)) = {} & |N| \ln(n!) + \ln((|N|n\bar{\lambda})!) + \ln((|N|n(1-\bar{\lambda}))!) - \ln((|N|n)!) \\ & - \sum_{i \in N_1^\lambda} \ln((n\lambda_i)!) - \sum_{i \in N_2^\lambda} \ln((n(1-\lambda_i))!). \end{aligned} \tag{2.80}$$

Now, using Stirling's approximation, which is given by

$$\ln(n!) = g(n) - n + \frac{1}{2}\ln(2\pi n) + \mathcal{O}\left(\frac{1}{n}\right), \qquad \text{for all } n \in \mathbb{N},$$

formula (2.80) can be written as

$$\begin{aligned} \ln(t^n(\lambda)) = {} & |N|g(n) - |N|n + \frac{1}{2}|N|\ln(2\pi n) + \mathcal{O}\left(\frac{1}{n}\right) \\ & + g(|N|n\bar{\lambda}) - |N|n\bar{\lambda} + \frac{1}{2}\ln(2\pi|N|n\bar{\lambda}) + \mathcal{O}\left(\frac{1}{|N|n\bar{\lambda}}\right) \\ & + g(|N|n(1-\bar{\lambda})) - |N|(n(1-\bar{\lambda})) + \frac{1}{2}\ln(2\pi|N|n(1-\bar{\lambda})) + \mathcal{O}\left(\frac{1}{|N|n(1-\bar{\lambda})}\right) \\ & - \left[g(|N|n) - |N|n + \frac{1}{2}\ln(2\pi|N|n) + \mathcal{O}\left(\frac{1}{|N|n}\right)\right] \\ & - \sum_{i \in N_1^\lambda}\left[g(n\lambda_i) - n\lambda_i + \frac{1}{2}\ln(2\pi n\lambda_i) + \mathcal{O}\left(\frac{1}{n\lambda_i}\right)\right] \\ & - \sum_{i \in N_2^\lambda}\left[g(n(1-\lambda_i)) - n(1-\lambda_i) + \frac{1}{2}\ln(2\pi n(1-\lambda_i)) + \mathcal{O}\left(\frac{1}{n(1-\lambda_i)}\right)\right]. \end{aligned}$$

Now, using that $g(xy) = xg(y) + yg(x)$ for all $x, y \geq 0$, $g(0) = 0$, $\sum_{i \in N_1^\lambda} \lambda_i = |N|\bar{\lambda}$ and $\sum_{i \in N_2^\lambda}(1-\lambda_i) = |N|(1-\bar{\lambda})$, we get

$$\ln(t^n(\lambda)) = |N|g(n) - |N|n + \frac{1}{2}|N|\ln(2\pi n) + \mathcal{O}\left(\frac{1}{n}\right)$$

$$
\begin{aligned}
&+ \bar{\lambda} g(|N|n) + |N|ng(\bar{\lambda}) - |N|n\bar{\lambda} + \frac{1}{2}\ln(2\pi n) + \frac{1}{2}\ln(|N|) + \frac{1}{2}\ln(\bar{\lambda}) + \mathcal{O}\left(\frac{1}{n\bar{\lambda}}\right) \\
&+ (1-\bar{\lambda})g(|N|n) + |N|ng(1-\bar{\lambda}) - |N|n(1-\bar{\lambda}) + \frac{1}{2}\ln(2\pi n) + \frac{1}{2}\ln(|N|) + \frac{1}{2}\ln(1-\bar{\lambda}) \\
&+ \mathcal{O}\left(\frac{1}{n(1-\bar{\lambda})}\right) \\
&- g(|N|n) + |N|n - \frac{1}{2}\ln(2\pi n) - \frac{1}{2}\ln(|N|) + \mathcal{O}\left(\frac{1}{n}\right) \\
&- |N|\bar{\lambda}g(n) - \sum_{i\in N} ng(\lambda_i) + |N|n\bar{\lambda} - \frac{1}{2}|N_1^\lambda|\ln(2\pi n) - \frac{1}{2}\sum_{i\in N_1^\lambda}\ln(\lambda_i) + \sum_{i\in N_1^\lambda}\mathcal{O}\left(\frac{1}{n\lambda_i}\right) \\
&- |N|(1-\bar{\lambda})g(n) - \sum_{i\in N} ng(1-\lambda_i) + |N|n(1-\bar{\lambda}) - \frac{1}{2}|N_2^\lambda|\ln(2\pi n) - \frac{1}{2}\sum_{i\in N_2^\lambda}\ln(1-\lambda_i) \\
&+ \sum_{i\in N_2^\lambda}\mathcal{O}\left(\frac{1}{n(1-\lambda_i)}\right).
\end{aligned}
$$

From $|N|g(n) - |N|\bar{\lambda}g(n) - |N|(1-\bar{\lambda})g(n) = 0$, $-|N|n - |N|n\bar{\lambda} - |N|(n(1-\bar{\lambda})) + |N|n + |N|n\bar{\lambda} + |N|n(1-\bar{\lambda}) = 0$, $\bar{\lambda}g(|N|n) + (1-\bar{\lambda})g(|N|n) - g(|N|n) = 0$, and rearranging and collecting some terms it follows that

$$
\begin{aligned}
\ln(t^n(\lambda)) = n&\left[|N|g(\bar{\lambda}) - \sum_{i\in N} g(\lambda_i) + |N|g(1-\bar{\lambda}) - \sum_{i\in N} g(1-\lambda_i)\right] \\
&+ \left[\frac{1}{2}(1+|N|-|N_1^\lambda|-|N_2^\lambda|)\right]\ln(2\pi n) + \frac{1}{2}\ln(|N|) \\
&+ \frac{1}{2}\ln(\bar{\lambda}) + \frac{1}{2}\ln(1-\bar{\lambda}) - \frac{1}{2}\sum_{i\in N_1^\lambda}\ln(\lambda_i) - \frac{1}{2}\sum_{i\in N_2^\lambda}\ln(1-\lambda_i) \\
&+ \mathcal{O}\left(\frac{1}{n}\right) + \mathcal{O}\left(\frac{1}{n\bar{\lambda}}\right) + \mathcal{O}\left(\frac{1}{n(1-\bar{\lambda})}\right) + \sum_{i\in N_1^\lambda}\mathcal{O}\left(\frac{1}{n\lambda_i}\right) + \sum_{i\in N_2^\lambda}\mathcal{O}\left(\frac{1}{n(1-\lambda_i)}\right).
\end{aligned}
$$

Then, recall the function G from (2.74). We get

$$
\begin{aligned}
\ln(t^n(\lambda)) = nG(\lambda) &+ \left[\frac{1}{2}(1+|N|-|N_1^\lambda|-|N_2^\lambda|)\right]\ln(2\pi n) + \frac{1}{2}\ln(|N|) + \frac{1}{2}\ln(\bar{\lambda}(1-\bar{\lambda})) \\
&- \frac{1}{2}\sum_{i\in N_1^\lambda}\ln(\lambda_i) - \frac{1}{2}\sum_{i\in N_2^\lambda}\ln(1-\lambda_i) + \mathcal{O}\left(\frac{1}{n}\right) + \mathcal{O}\left(\frac{1}{n\bar{\lambda}}\right) \\
&+ \mathcal{O}\left(\frac{1}{n(1-\bar{\lambda})}\right) + \sum_{i\in N_1^\lambda}\mathcal{O}\left(\frac{1}{n\lambda_i}\right) + \sum_{i\in N_2^\lambda}\mathcal{O}\left(\frac{1}{n(1-\lambda_i)}\right).
\end{aligned}
$$

So, taking the exponent and using the fact that $e^x = 1 + \mathcal{O}(x)$ if $x \in [0, K]$ for some constant

$K > 0$, yields

$$t^n(\lambda) = \left(e^{G(\lambda)}\right)^n (2\pi n)^{\frac{1}{2}(1+|N|-|N_1^\lambda|-|N_2^\lambda|)} \sqrt{|N|} \frac{(\bar{\lambda}(1-\bar{\lambda}))^{\frac{1}{2}}}{\prod_{i\in N_1^\lambda}\sqrt{\lambda_i}\prod_{i\in N_2^\lambda}\sqrt{1-\lambda_i}} \cdot \left[1+\mathcal{O}\left(\frac{1}{n}\right)\right]$$
$$\cdot \left[1+\mathcal{O}\left(\frac{1}{n\bar{\lambda}}\right)\right] \cdot \left[1+\mathcal{O}\left(\frac{1}{n(1-\bar{\lambda})}\right)\right] \cdot \prod_{i\in N_1^\lambda}\left[1+\mathcal{O}\left(\frac{1}{n\lambda_i}\right)\right] \cdot \prod_{i\in N_2^\lambda}\left[1+\mathcal{O}\left(\frac{1}{n(1-\lambda_i)}\right)\right].$$

Then, as $\lambda_i \geq \min\{\lambda_j : j \in N_1^\lambda\}$ for all $i \in N_1^\lambda$, $1-\lambda_i \geq \min\{1-\lambda_j : j \in N_2^\lambda\}$ for all $i \in N_2^\lambda$, $\bar{\lambda} \geq \frac{1}{|N|}\min\{\lambda_j : j \in N_1^\lambda\}$, and $1-\bar{\lambda} \geq \frac{1}{|N|}\min\{1-\lambda_j : j \in N_2^\lambda\}$, the result follows in a straightforward way. □

Lemma 2.A.6 *We have for all $(\varepsilon, n, \lambda) \in Dom$ with $d_n < \frac{1}{2}\varepsilon$ and $\lambda \in D(n)$ that*

$$\frac{(\bar{\lambda}(1-\bar{\lambda}))^{\frac{1}{2}|N|}}{\prod_{i\in N}\sqrt{\lambda_i}\prod_{i\in N}\sqrt{1-\lambda_i}} = 1 + \mathcal{O}^\varepsilon(n^{-1+\frac{1}{4|N|}}). \tag{2.81}$$

Proof: According to Lemma 2.A.4 we have $\lambda_i \geq \frac{1}{2}\varepsilon$ and $1-\lambda_i \geq \frac{1}{2}\varepsilon$ for all $i \in N$. Consequently, we have $\bar{\lambda} \geq \frac{1}{2}\varepsilon$ and $1-\bar{\lambda} \geq \frac{1}{2}\varepsilon$. According to Taylor's theorem, we have

$$\ln(\lambda_i) = \ln(\bar{\lambda}) + \frac{1}{\bar{\lambda}}(\lambda_i - \bar{\lambda}) - \frac{1}{2\xi_{i,1}^2}(\lambda_i-\bar{\lambda})^2, \tag{2.82}$$

for some $\xi_{i,1} \in [\lambda_i; \bar{\lambda}]$ and for all $i \in N$. From (2.78) and (2.82) it follows that

$$\frac{1}{2}\sum_{i\in N}\ln(\lambda_i) = \frac{1}{2}|N|\ln(\bar{\lambda}) - \sum_{i\in N}\frac{1}{4\xi_{i,1}^2}(\lambda_i-\bar{\lambda})^2. \tag{2.83}$$

Similarly, we obtain

$$\frac{1}{2}\sum_{i\in N}\ln(1-\lambda_i) = \frac{1}{2}|N|\ln(1-\bar{\lambda}) - \sum_{i\in N}\frac{1}{4\xi_{i,2}^2}(\bar{\lambda}-\lambda_i)^2, \tag{2.84}$$

for some $\xi_{i,2} \in [1-\lambda_i; 1-\bar{\lambda}]$ and for all $i \in N$. Since $\xi_{i,1} \geq \frac{1}{2}\varepsilon$, $\xi_{i,2} \geq \frac{1}{2}\varepsilon$ and $(\lambda_i - \bar{\lambda})^2 \leq \|\lambda - \bar{\lambda}\cdot e_N\|^2$ for all $i \in N$, we get

$$\begin{aligned}\sum_{i\in N}\frac{1}{4\xi_{i,1}^2}(\lambda_i-\bar{\lambda})^2 + \sum_{i\in N}\frac{1}{4\xi_{i,2}^2}(\bar{\lambda}-\lambda_i)^2 &\leq 2|N|\varepsilon^{-2}\|\lambda-\bar{\lambda}\cdot e_N\|^2 \\ &\leq 2|N|\varepsilon^{-2}d_n^2 \\ &= 2|N|\varepsilon^{-2}n^{-1+\frac{1}{4|N|}} \\ &= \mathcal{O}^\varepsilon(n^{-1+\frac{1}{4|N|}}).\end{aligned}$$

Using the fact that $e^x = 1 + \mathcal{O}(x)$ if $x \in [0, K]$ for some constant $K > 0$ yields

$$e^{\mathcal{O}^\varepsilon(n^{-1+\frac{1}{4|N|}})} = 1 + \mathcal{O}^\varepsilon(n^{-1+\frac{1}{4|N|}}).$$

Now taking the exponent in (2.83) and (2.84) yields the desired result. □

Proof of Proposition 2.6.3: We prove the result step-by-step: (2.32) is shown in Lemma 2.A.7, (2.33) in Lemma 2.A.8, (2.34) in Lemma 2.A.9, and (2.35) in Lemma 2.A.10. We implicitly use in the statement of this proposition that if $g(n) = \mathcal{O}(n^c)$ for some $c \leq -\frac{1}{4}$, we have $g(n) = \mathcal{O}(n^{-\frac{1}{4}})$. Note that the result follows directly if $|N| = 1$, so we let $|N| \geq 2$.

Lemma 2.A.7 *We have for all $(\varepsilon, n, \lambda) \in Dom$ that*

$$t^n(\lambda) = \left(e^{-c(\bar{\lambda})n\|\lambda - \bar{\lambda}\cdot e_N\|^2}\right) b(n, \bar{\lambda}) \left[1 + \mathcal{O}^\varepsilon\left(n^{-\frac{1}{2}+\frac{3}{8|N|}}\right)\right], \qquad \text{if } \lambda \in D(n).$$

Proof: It is sufficient to show this result for all $n \in \mathbb{N}$ such that $d_n < \frac{1}{2}\varepsilon$. From Lemma 2.A.4, we then get

$$\min\{\lambda_1, \ldots, \lambda_{|N|}, 1 - \lambda_1, \ldots, 1 - \lambda_{|N|}\} > \frac{1}{2}\varepsilon, \tag{2.85}$$

and, so,

$$N_1^\lambda = N_2^\lambda = N. \tag{2.86}$$

Using Lemma 2.A.3.3 and the fact that $\|\lambda - \bar{\lambda} \cdot e_N\|^3 = \mathcal{O}(n^{-\frac{3}{2}+\frac{3}{8|N|}})$, we get that

$$G(\bar{\lambda}) = -c(\bar{\lambda})\|\lambda - \bar{\lambda} \cdot e_N\|^2 + \mathcal{O}^\varepsilon(n^{-\frac{3}{2}+\frac{3}{8|N|}}).$$

Hence,

$$e^{nG(\lambda)} = e^{-c(\bar{\lambda})n\|\lambda-\bar{\lambda}\cdot e_N\|^2} \cdot e^{\mathcal{O}^\varepsilon(n^{-\frac{3}{2}+\frac{3}{8|N|}})} \tag{2.87}$$

$$= e^{-c(\bar{\lambda})n\|\lambda-\bar{\lambda}\cdot e_N\|^2} \left[1 + \mathcal{O}^\varepsilon\left(n^{-\frac{1}{2}+\frac{3}{8|N|}}\right)\right], \tag{2.88}$$

where (2.88) follows from the fact that $e^x = 1 + \mathcal{O}(x)$ if $x \in [0, K]$ for some constant $K > 0$. Substituting (2.81), (2.85), (2.86) and (2.87)-(2.88) in (2.79) yields the desired result. □

Lemma 2.A.8 *We have for all $(\varepsilon, n, \lambda) \in Dom$ that*

$$t^n(\lambda) = \mathcal{O}^\varepsilon(n^{-\infty}), \qquad \text{if } \lambda \notin D(n).$$

Proof: Let $\varepsilon \in (0, 1)$, denote $d = \frac{1}{3|N|}\varepsilon^2$, and recall the function G in (2.74). The set $G_\varepsilon \backslash D^d$ is compact. Moreover, the function G is continuous (Lemma 2.A.3.1). Hence, the function G

takes a maximum value m_ε on $G_\varepsilon \backslash D^d$. As $\lambda \in D^d$ if $\lambda_1 = \cdots = \lambda_{|N|}$, we obtain from Lemma 2.A.3.2 that $m_\varepsilon < 0$. Let (n, λ) be such that $n \in \mathbb{N}$ and $\lambda \in G_\varepsilon^n \backslash D^d$. Since $\lambda_i \geq \frac{1}{n}$ for all $i \in N_1^\lambda$, $1 - \lambda_i \geq \frac{1}{n}$ for all $i \in N_2^\lambda$ and $\bar{\lambda}(1 - \bar{\lambda}) < 1$, we get from Lemma 2.A.5 that

$$t^n(\lambda) = \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1+|N|)}(e^{m_\varepsilon})^n).$$

Since $e^{m_\varepsilon} \in (0, 1)$ and $\lim_{n\to\infty} c^n n^d = 0$ for $c \in (0, 1)$ and $d \in \mathbb{R}$, we have for all $(\varepsilon, n, \lambda) \in Dom$ that

$$t^n(\lambda) = \mathcal{O}^\varepsilon(n^{-\infty}), \qquad \text{if } \lambda \notin D^d.$$

Next, we show this result for all (n, λ) such that $n \in \mathbb{N}$ and $\lambda \in (G_\varepsilon^n \cap D^d)\backslash D(n)$. We obtain from Lemma 2.A.3.3 that

$$G(\lambda) = -c(\bar{\lambda})\|\lambda - \bar{\lambda} \cdot e_N\|^2 + R \tag{2.89}$$

$$= -c(\bar{\lambda})\|\lambda - \bar{\lambda} \cdot e_N\|^2 \left[1 - R(c(\bar{\lambda}))^{-1}\|\lambda - \bar{\lambda} \cdot e_N\|^{-2}\right], \tag{2.90}$$

where $|R| \leq \frac{1}{3}|N| \min\{\lambda_1, \ldots, \lambda_{|N|}, 1 - \lambda_1, \ldots, 1 - \lambda_{|N|}\}^{-2}\|\lambda - \bar{\lambda} \cdot e_N\|^3$. From Lemma 2.A.4, we get

$$\min\{\lambda_1, \ldots, \lambda_{|N|}, 1 - \lambda_1, \ldots, 1 - \lambda_{|N|}\} > \varepsilon - d > \frac{3|N| - 1}{3|N|}\varepsilon > \frac{1}{2}\varepsilon.$$

Moreover, we have $(c(\bar{\lambda}))^{-1} = 2\bar{\lambda}(1 - \bar{\lambda}) \leq \frac{1}{2}$ and $\|\lambda - \bar{\lambda} \cdot e_N\| < d$. Therefore, we have

$$|R(c(\bar{\lambda}))^{-1}\|\lambda - \bar{\lambda} \cdot e_N\|^{-2}| \leq |R|(c(\bar{\lambda}))^{-1}\|\lambda - \bar{\lambda} \cdot e_N\|^{-2} \leq \frac{1}{6}|N|\left(\frac{1}{2}\varepsilon\right)^{-2} d < \frac{1}{2}. \tag{2.91}$$

So, from (2.89)-(2.90) and (2.91) it follows that

$$nG(\lambda) < -c(\bar{\lambda})n\|\lambda - \bar{\lambda} \cdot e_N\|^2 \frac{1}{2} \leq -n\|\lambda - \bar{\lambda} \cdot e_N\|^2 \leq -n^{\frac{1}{4|N|}},$$

where the second inequality follows from $c(\bar{\lambda}) \geq 2$, and the third inequality from $\|\lambda - \bar{\lambda} \cdot e_N\| \geq n^{-\frac{1}{2}+\frac{1}{8|N|}}$. Hence,

$$e^{nG(\lambda)} < e^{-n^{\frac{1}{4|N|}}}. \tag{2.92}$$

We get

$$t^n(\lambda) = \mathcal{O}^\varepsilon(e^{nG(\lambda)} n^{\frac{1}{2}(1-|N|)}) \tag{2.93}$$

$$= \mathcal{O}^\varepsilon((e^{-1})^{n^{\frac{1}{4|N|}}} n^{\frac{1}{2}(1-|N|)}) \tag{2.94}$$

$$= \mathcal{O}^\varepsilon(n^{-\infty}), \tag{2.95}$$

where (2.93) follows from Lemma 2.A.5, (2.94) follows from (2.92) and (2.95) follows from the fact that $\lim_{n\to\infty} n^k c^{n^\delta} = 0$ for all $k \in \mathbb{R}$, $c \in (0,1)$ and $\delta > 0$. □

Lemma 2.A.9 *We have for all $i \in N$ and $(\varepsilon, n, \lambda) \in Dom$ that*

$$p_i^n(\lambda) = \frac{1}{|N|}\left[1 + \mathcal{O}^\varepsilon\left(n^{-\frac{1}{2}+\frac{1}{8|N|}}\right)\right], \qquad \text{if } \lambda \in D(n).$$

Proof: Note that from $\lambda \in G_\varepsilon^n$ it follows that $\lambda \neq e_N$, so $\bar{\lambda} < 1$. Then, the result follows directly from

$$\begin{aligned}
\left|\frac{1-\lambda_i}{\sum_{j\in N}(1-\lambda_j)} - \frac{1}{|N|}\right| &= \left|\frac{1-\lambda_i}{(1-\bar{\lambda})|N|} - \frac{1-\bar{\lambda}}{(1-\bar{\lambda})|N|}\right| \\
&= \frac{|\bar{\lambda}-\lambda_i|}{(1-\bar{\lambda})|N|} \\
&< \frac{n^{-\frac{1}{2}+\frac{1}{8|N|}}}{(1-\bar{\lambda})|N|} \qquad (2.96) \\
&\leq \frac{n^{-\frac{1}{2}+\frac{1}{8|N|}}}{\varepsilon|N|}, \qquad (2.97)
\end{aligned}$$

for all $(\varepsilon, n, \lambda) \in Dom$ such that $\lambda \in D(n)$. Here, (2.96) follows from $|\bar{\lambda} - \lambda_i| \leq \|\lambda - \bar{\lambda}\cdot e_N\| < d_n = n^{-\frac{1}{2}+\frac{1}{8|N|}}$, and (2.97) follows from $1 - \bar{\lambda} \geq \varepsilon$. This concludes the proof. □

Lemma 2.A.10 *We have for all $i \in N$ and $(\varepsilon, n, \lambda) \in Dom$ that*

$$p_i^n(\lambda) = \mathcal{O}(1).$$

Proof: This follows directly from $0 \leq p_i^n(\lambda) \leq 1$. □

This concludes the proof of Proposition 2.6.3. □

In this sequel of this appendix, we use the following notation:

- For all $x \in \mathbb{R}$, we write $\lfloor x \rfloor$ as the largest integer not greater than x, and $\lceil x \rceil$ as the smallest integer not less than x.
- For all $n \in \mathbb{N}$ and $\lambda \in G^n$, the set $C^n(\lambda)$ is given by

$$C^n(\lambda) = \left\{\lambda + \frac{1}{n}x : x \in [0,1]^N\right\}. \qquad (2.98)$$

- The set $D'(n)$ is given by

$$D'(n) = D^{d'_n}, \qquad \text{where } d'_n = d_n + (\sqrt{|N|}/n) = n^{-\frac{1}{2}+\frac{1}{8|N|}} + \frac{\sqrt{|N|}}{n}. \qquad (2.99)$$

- If there might be confusion about the notation $|\cdot|$ for the absolute value of a real number and the cardinality of a set, we sometimes write $\sharp(A)$ as the cardinality of the set A.

- We write $\nu(B)$ as the Lebesgue measure of the set B. Note that

$$\nu(C^n(\lambda)) = n^{-|N|}, \qquad \text{for all } \lambda \in G^n, \tag{2.100}$$

and

$$\nu(D'(n)) = \mathcal{O}(d_n'^{|N|-1}) = \mathcal{O}(n^{-\frac{1}{2}+\frac{1}{8|N|}}), \qquad \text{for all } n \in \mathbb{N}. \tag{2.101}$$

- Let $R \in \mathcal{R}$. We define the set $B(R,n)$ by

$$B(R,n) = \left\{\lambda \in [0,1]^N : \exists \hat{\lambda} \in [0,1]^N \backslash L(R) : \|\lambda - \hat{\lambda}\| < \frac{1}{n}\right\}, \tag{2.102}$$

for all $R \in \mathcal{R}$ and $n \in \mathbb{N}$, where $L(R)$ is defined in Definition 2.A.1. This is the set of all participation profiles close to a participation profile that is an element of multiple sets K_m. As the risk capital allocation problem is always clear from the context, we write $B(n) = B(R,n)$.

First, we show that only the participation profiles in G_ε^n have a non-negligible aggregate contribution.

Lemma 2.A.11 *For all $i \in N$, we have*

$$n^{-1} \sum_{\lambda \in G^n \backslash G_\varepsilon^n : \lambda_i < 1} t^n(\lambda) p_i^n(\lambda) = \mathcal{O}(\varepsilon) + \mathcal{O}(n^{-1}).$$

Proof: Recall (2.60) for the definition of $\widetilde{G}_k^n$. Then, we obtain

$$\sum_{\lambda \in G^n \backslash G_\varepsilon^n : \lambda_i < 1} t^n(\lambda) p_i^n(\lambda) = \sum_{k=0}^{\lceil \varepsilon |N| n \rceil - 1} \sum_{\lambda \in \widetilde{G}_k^n : \lambda_i < 1} t^n(\lambda) p_i^n(\lambda) + \sum_{k=\lfloor (1-\varepsilon)|N|n \rfloor + 1}^{|N|n-1} \sum_{\lambda \in \widetilde{G}_k^n : \lambda_i < 1} t^n(\lambda) p_i^n(\lambda) \tag{2.103}$$

$$\leq \sum_{k=0}^{\lceil \varepsilon |N| n \rceil - 1} \sum_{\lambda \in \widetilde{G}_k^n : \lambda_i < 1} t^n(\lambda) + \sum_{k=\lfloor (1-\varepsilon)|N|n \rfloor + 1}^{|N|n-1} \sum_{\lambda \in \widetilde{G}_k^n : \lambda_i < 1} t^n(\lambda) \tag{2.104}$$

$$\leq \sum_{k=0}^{\lceil \varepsilon |N| n \rceil - 1} 1 + \sum_{k=\lfloor (1-\varepsilon)|N|n \rfloor + 1}^{|N|n-1} 1 \tag{2.105}$$

$$= \lceil \varepsilon |N| n \rceil + \lceil \varepsilon |N| n \rceil - 1$$

$$< 2\varepsilon |N| n + 1 \tag{2.106}$$

$$= \mathcal{O}(\varepsilon)n + \mathcal{O}(1).$$

Here, (2.103) follows from (2.60) and (2.69), (2.104) follows from $0 \leq p_i^n(\lambda) \leq 1$ for all $\lambda \in G^n \backslash \{e_N\}$, (2.105) follows from $\sum_{\lambda \in \widetilde{G}_k^n} t^n(\lambda) = 1$ for all $k \in \{0, \ldots, |N|n - 1\}$, and (2.106) follows from the fact that $\lceil x \rceil < x + 1$ for all $x \in \mathbb{R}$. □

The following result follows almost directly from Proposition 2.6.3.

Lemma 2.A.12 *For all $i \in N$, we have*

$$\sum_{\lambda \in [G_\varepsilon^n \backslash D(n)]:\lambda_i < 1} t^n(\lambda) p_i^n(\lambda) = \mathcal{O}^\varepsilon(n^{-\infty}).$$

Proof: This result follows directly from

$$\sum_{\lambda \in [G_\varepsilon^n \backslash D(n)]:\lambda_i < 1} t^n(\lambda) p_i^n(\lambda) = \sum_{\lambda \in [G_\varepsilon^n \backslash D(n)]:\lambda_i < 1} \mathcal{O}^\varepsilon(n^{-\infty}) \tag{2.107}$$

$$< (n+1)^{|N|} \mathcal{O}^\varepsilon(n^{-\infty}) \tag{2.108}$$

$$= \mathcal{O}^\varepsilon(n^{-\infty}),$$

where (2.107) follows from Proposition 2.6.3, and (2.108) follows from $\sharp(\{\lambda \in [G_\varepsilon^n \backslash D(n)] : \lambda_i < 1\}) < \sharp(G^n) = (n+1)^{|N|}$. □

Lemma 2.A.13 *Let $R \in \mathcal{R}$. Then, we have*

$$r\left(\lambda + (1/n) \cdot e_i\right) - r(\lambda) = \mathcal{O}\left(n^{-1}\right),$$

for all $i \in N$ and (n, λ) such that $n \in \mathbb{N}$, $\lambda \in G^n$ and $\lambda_i < 1$.

Proof: Denote $c = \max\{|f_{\mathbb{Q}}(e_j)| : \mathbb{Q} \in Q(\rho), j \in N\}$. Let $\mathbb{Q}_1, \mathbb{Q}_2 \in Q(\rho)$ be such that $r\left(\lambda + (1/n) \cdot e_i\right) = f_{\mathbb{Q}_1}\left(\lambda + (1/n) \cdot e_i\right)$ and $r(\lambda) = f_{\mathbb{Q}_2}(\lambda)$. Then, we have

$$\begin{aligned} r\left(\lambda + (1/n) \cdot e_i\right) - r(\lambda) &= f_{\mathbb{Q}_1}\left(\lambda + (1/n) \cdot e_i\right) - f_{\mathbb{Q}_2}\left(\lambda\right) \\ &\leq f_{\mathbb{Q}_1}\left(\lambda + (1/n) \cdot e_i\right) - f_{\mathbb{Q}_1}\left(\lambda\right) \\ &= \frac{1}{n} f_{\mathbb{Q}_1}\left(e_i\right) \\ &\leq \frac{1}{n} c \end{aligned}$$

and

$$\begin{aligned} r\left(\lambda + (1/n) \cdot e_i\right) - r(\lambda) &= f_{\mathbb{Q}_1}\left(\lambda + (1/n) \cdot e_i\right) - f_{\mathbb{Q}_2}\left(\lambda\right) \\ &\geq f_{\mathbb{Q}_2}\left(\lambda + (1/n) \cdot e_i\right) - f_{\mathbb{Q}_2}\left(\lambda\right) \\ &= \frac{1}{n} f_{\mathbb{Q}_2}\left(e_i\right) \end{aligned}$$

$$\geq -\frac{1}{n}c.$$

This concludes the proof. □

Lemma 2.A.14 *For all $i \in N$, we have*

$$\sum_{\lambda \in G_\varepsilon^n \cap D(n)} |t^n(\lambda)p_i^n(\lambda) - h^n(\lambda)| = \mathcal{O}^\varepsilon(n^{\frac{3}{4}}).$$

Proof: It is sufficient to show this result only for $n \in \mathbb{N}$ such that $d_n < \frac{1}{2}\varepsilon$. If $|N| = 1$ the result is trivial as $t^n(\lambda)p_i^n(\lambda) = h^n(\lambda) = 1$ for all $\lambda \in G_\varepsilon^n$. Next, we let $|N| \geq 2$. For all $\lambda \in G_\varepsilon^n \cap D(n)$, we have

$$\begin{aligned} |t^n(\lambda)p_i^n(\lambda) - h^n(\lambda)| &= \left| h^n(\lambda)[1 + \mathcal{O}^\varepsilon(n^{-\frac{1}{2}+\frac{3}{8|N|}})] \cdot [1 + \mathcal{O}^\varepsilon(n^{-\frac{1}{2}+\frac{1}{8|N|}})] - h^n(\lambda) \right| && (2.109) \\ &= \left| h^n(\lambda)\mathcal{O}^\varepsilon(n^{-\frac{1}{2}+\frac{3}{8|N|}}) \right| \\ &= \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+\frac{3}{8|N|}}), && (2.110) \end{aligned}$$

where (2.109) follows from Lemma 2.A.7 and Lemma 2.A.9, and (2.110) follows from $h^n(\lambda) = \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)})$. If $y \in C^n(\lambda)$ for some $\lambda \in G_\varepsilon^n \cap D(n)$, we have

$$\begin{aligned} \|y - \bar{y} \cdot e_N\| &\leq \|y - \bar{\lambda} \cdot e_N\| && (2.111) \\ &\leq \|y - \lambda\| + \|\lambda - \bar{\lambda} \cdot e_N\| && (2.112) \\ &< (\sqrt{|N|}/n) + n^{-\frac{1}{2}+\frac{1}{8|N|}}, && (2.113) \end{aligned}$$

where (2.111) and (2.112) follow from the triangular inequality, and (2.113) follows from the fact that $\|y - \lambda\| \leq (\sqrt{|N|}/n)$ for all $y \in C^n(\lambda)$. So, we get

$$\bigcup_{\lambda \in G_\varepsilon^n \cap D(n)} C^n(\lambda) \subset D'(n). \quad (2.114)$$

From this and (2.100), we get

$$\begin{aligned} n^{-|N|}\sharp(G_\varepsilon^n \cap D(n)) &\leq \nu\left(D'(n)\right) && (2.115) \\ &= \mathcal{O}(d_n'^{|N|-1}) && (2.116) \\ &= \mathcal{O}\left(\left(n^{-\frac{1}{2}+\frac{1}{8|N|}} + (\sqrt{|N|}/n)\right)^{(|N|-1)}\right) && (2.117) \\ &= \mathcal{O}(n^{-\frac{1}{2}|N|+\frac{5}{8}-\frac{1}{8|N|}}), && (2.118) \end{aligned}$$

where (2.115) follows from (2.100) and (2.114), and (2.116) follows from (2.101). From this, we get

$$\sum_{\lambda\in G_\varepsilon^n\cap D(n)} |t^n(\lambda)p_i^n(\lambda) - h^n(\lambda)| \leq \sharp(G_\varepsilon^n \cap D(n))\mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+\frac{3}{8|N|}}) = \mathcal{O}^\varepsilon(n^{\frac{5}{8}+\frac{2}{8|N|}}).$$

As $|N| \geq 2$, this concludes the proof. □

Lemma 2.A.15 *Let $R \in \mathcal{R}$. Then, for all $\varepsilon > 0$ and all $m \in \{p^*+1,\ldots,p\}$, we have for sufficiently large n that*

$$G_\varepsilon \cap D(n) \cap K_m = \emptyset.$$

Proof: If $p^* = p$, the result follows directly, and, so, we let $p^* < p$. Denote

$$\alpha = r(e_N) - \max_{m'\in\{p^*+1,\ldots,p\}} f_{\mathbb{Q}_{m'}}(e_N) > 0,$$

and let $\ell \in \{1,\ldots,p^*\}$ and $m \in \{p^*+1,\ldots,p\}$. Then, we have

$$f_{\mathbb{Q}_\ell}(e_N) \geq f_{\mathbb{Q}_m}(e_N) + \alpha.$$

By linearity of $f_{\mathbb{Q}_\ell}$, we have

$$f_{\mathbb{Q}_\ell}(t\cdot e_N) - f_{\mathbb{Q}_m}(t\cdot e_N) = t(f_{\mathbb{Q}_\ell}(e_N) - f_{\mathbb{Q}_m}(e_N)) \geq t\alpha, \qquad \text{for all } t\in[0,1]. \tag{2.119}$$

If $f_{\mathbb{Q}_{m'}}(e_i) = 0$ for all $m' \in \{1,\ldots,p\}$ and for all $i \in N$, we have $p = p^* = 1$, which contradicts the assumption that $p^* < p$. So, let $M = \max_{m'\in\{1,\ldots,p\}} \|(f_{\mathbb{Q}_{m'}}(e_i))_{i\in N}\| > 0$ and $\varepsilon > 0$. Then, define $N_\varepsilon = \left(\frac{2M}{\alpha\varepsilon}\right)^4$ and let $n > N_\varepsilon$. Then, we obtain for every $\lambda \in G_\varepsilon \cap D(n)$ that

$$f_{\mathbb{Q}_\ell}(\lambda) - f_{\mathbb{Q}_m}(\lambda) = f_{\mathbb{Q}_\ell}(\bar\lambda\cdot e_N) - f_{\mathbb{Q}_m}(\bar\lambda\cdot e_N) + f_{\mathbb{Q}_\ell}(\lambda-\bar\lambda\cdot e_N) - f_{\mathbb{Q}_m}(\lambda-\bar\lambda\cdot e_N) \tag{2.120}$$

$$\geq \bar\lambda\alpha + f_{\mathbb{Q}_\ell}(\lambda-\bar\lambda\cdot e_N) - f_{\mathbb{Q}_m}(\lambda-\bar\lambda\cdot e_N), \tag{2.121}$$

where (2.120) follows from linearity of $f_{\mathbb{Q}_\ell}$ and $f_{\mathbb{Q}_m}$, and (2.121) follows from (2.119). Moreover, we obtain that

$$|f_{\mathbb{Q}_{m'}}(\lambda-\bar\lambda\cdot e_N)| \leq \|(f_{\mathbb{Q}_{m'}}(e_i))_{i\in N}\|\cdot\|\lambda-\bar\lambda\cdot e_N\| \tag{2.122}$$

$$\leq Mn^{-\frac{1}{4}} \tag{2.123}$$

$$< MN_\varepsilon^{-\frac{1}{4}} \tag{2.124}$$

$$= \frac{1}{2}\varepsilon\alpha \tag{2.125}$$

$$\leq \frac{1}{2}\bar\lambda\alpha, \tag{2.126}$$

for all $m' \in \{1, \ldots, p\}$, where (2.122) follows from the Cauchy-Schwartz inequality applied to $\sum_{i \in N} f_{\mathbb{Q}_m}(e_i)(\lambda_i - \bar{\lambda})$, (2.123) follows from $m' \in \{1, \ldots, p\}$ and $\lambda \in D(n)$, (2.124) follows from $n > N_\varepsilon$, (2.125) follows from substituting the definition of N_ε, and (2.126) follows from $\lambda \in G_\varepsilon$. Hence, substituting (2.122)-(2.126) in (2.120)-(2.121) yields that $f_{\mathbb{Q}_\ell}(\lambda) - f_{\mathbb{Q}_m}(\lambda) > 0$. Therefore, we have $\lambda \notin K_m$ for every $\lambda \in G_\varepsilon \cap D(n)$ and, hence,

$$G_\varepsilon \cap D(n) \cap K_m = \emptyset. \tag{2.127}$$

□

Note that from (2.22) and Lemma 2.A.15 it follows for every $\varepsilon > 0$ that

$$G_\varepsilon \cap D(n) \subset \bigcup_{m \in \{1,\ldots,p^*\}} K_m, \qquad \text{for large } n.$$

We next show that we can neglect participation profiles close to profiles where the function r is non-differentiable. Note that $B(n)$, as defined in (2.102), is the set of participation profiles close to a participation profile where the function r is non-differentiable. For all $n \in \mathbb{N}$ we have that if $\lambda \in K_m \backslash B(n)$ for some $m \in \{1, \ldots, p\}$, then $\lambda + (1/n) \cdot e_i \in K_m$ for all $i \in N$ and, by linearity of $f_{\mathbb{Q}_m}$, $r(\lambda + (1/n) \cdot e_i) - r(\lambda) = \frac{1}{n} E_{\mathbb{Q}_m}[X_i]$.

Lemma 2.A.16 *Let $R \in \mathcal{R}$. Then, we have*

$$\sum_{\lambda \in G_\varepsilon^n \cap D(n) \cap B(n)} h^n(\lambda) = \mathcal{O}^\varepsilon(n^{\frac{5}{8}}).$$

Proof: If $p = 1$, we have that $B(n) = \emptyset$ for all $n \in \mathbb{N}$ and, so, the result follows directly. Next, let $p > 1$. Recall (2.48), i.e.,

$$[0,1]^N \backslash L(R) = \bigcup_{\ell, m \in \{1,\ldots,p\}: \ell \neq m} K_\ell \cap K_m.$$

Let $\varepsilon > 0$, $\ell, m \in \{1, \ldots, p\}$, $\ell \neq m$ and $n > \frac{2}{\varepsilon}$. We define

$$H^n(\ell, m) = \left\{ \lambda \in G_\varepsilon^n \cap D(n) : \exists \hat{\lambda} \in K_\ell \cap K_m : \|\lambda - \hat{\lambda}\| \leq \frac{1}{n} \right\}$$

and $D_\varepsilon = \{\lambda \in G_\varepsilon : \lambda = \bar{\lambda} \cdot e_N\}$. According to Lemma 2.A.15 we have for all $m \in \{p^* + 1, \ldots, p\}$ that $D_\varepsilon \cap K_m = \emptyset$. Since D_ε and K_m are both compact we can define $\alpha_{\varepsilon,m} = \text{dist}(D_\varepsilon, K_m) = \min\{\|x - y\| : x \in D_\varepsilon, y \in K_m\}$. Obviously, $\alpha_{\varepsilon,m} > 0$. So, if $\ell \notin \{1, \ldots, p^*\}$ or $m \notin \{1, \ldots, p^*\}$ we get $H^n(\ell, m) = \emptyset$ for large n. If $p^* = 1$ it follows from this that $H^n(\ell, m) = \emptyset$ for all $\ell, m \in \{1, \ldots, p\}$. Next, let $p^* > 1$, and $\ell, m \in \{1, \ldots, p^*\}$. Recall (2.45)-(2.47) from Lemma

2.4.5, i.e.,

$$K_\ell \cap K_m \subset \left\{ \lambda \in \mathbb{R}^N : \sum_{i\in N} \lambda_i (E_{\mathbb{Q}_\ell}[X_i] - E_{\mathbb{Q}_m}[X_i]) = 0 \right\} := V(\ell, m).$$

Note that $V(\ell,m)$ is an $(|N|-1)$-dimensional linear space where $\{t \cdot e_N : t \in \mathbb{R}\} \subset V(\ell,m)$. To obtain an upperbound of the cardinality of $H^n(\ell,m)$, we first derive the Lebesgue measure of the following Euclidean set

$$\widetilde{H}^n(\ell,m) = \left\{ \lambda \in G_{\frac{1}{2}\varepsilon} \cap D'(n) : \exists \hat{\lambda} \in V(\ell,m) : \|\lambda - \hat{\lambda}\| \le \frac{\sqrt{|N|}}{n} + \frac{1}{n} \right\}.$$

We describe this set via the Gram-Schmidt process. Choose an orthonormal basis $u_1, \ldots, u_{|N|}$ of $\mathbb{R}^N$ such that $u_1 = \frac{e_N}{\sqrt{|N|}}$, $u_1, \ldots, u_{|N|-1}$ is an orthonormal basis of the $(|N|-1)$-dimensional space $V(\ell,m)$, and $u_{|N|}$ is a unit normal vector of the $(|N|-1)$-dimensional space $V(\ell,m)$. So $u_{|N|}$ is a multiple of the vector $(E_{\mathbb{Q}_\ell}[X_i] - E_{\mathbb{Q}_m}[X_i])_{i\in N}$. Now let $\lambda \in \widetilde{H}^n(\ell,m)$. Let λ_1 be the unique element in $V(\ell,m)$ that is closest to λ. Obviously $\|\lambda - \lambda_1\| \le \frac{\sqrt{|N|}}{n} + \frac{1}{n}$. Let $\lambda_2 = \bar{\lambda}_1 \cdot e_N$ $(= \bar{\lambda} \cdot e_N)$ be the unique element in $\{t \cdot e_N : t \in \mathbb{R}\}$ that is closest to λ_1 (and hence closest to λ). We provide an overview of the construction of λ_1 and λ_2 in Figure 2.6. Obviously

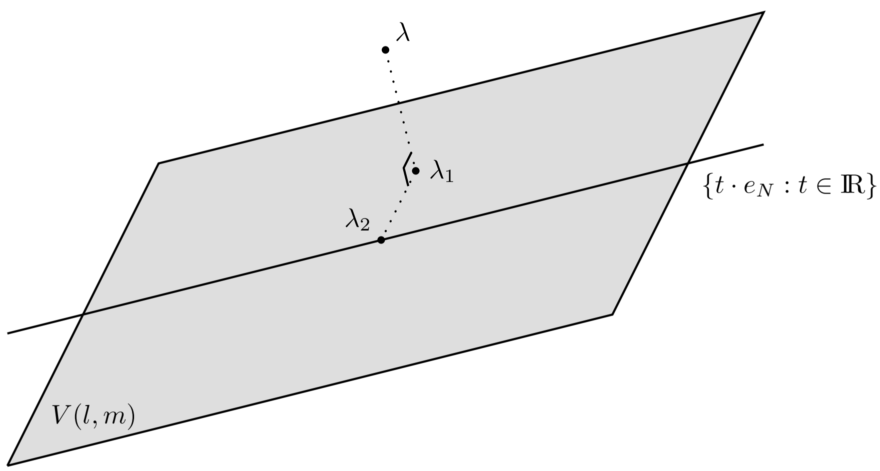


Figure 2.6: Illustration of λ_1 and λ_2 corresponding to the proof of Lemma 2.A.16.

$\|\lambda-\lambda_2\|^2 = \|\lambda-\lambda_1\|^2 + \|\lambda_1-\lambda_2\|^2$, and, hence, $\|\lambda_1 - \lambda_2\| \le \|\lambda - \lambda_2\| = \|\lambda - \bar{\lambda}\cdot e_N\| < d'_n$. Now we can write $\lambda = \alpha_1 u_1 + \cdots + \alpha_{|N|} u_{|N|}$ where $\lambda_2 = \alpha_1 u_1$, $\lambda_1 - \lambda_2 = \alpha_2 u_2 + \cdots + \alpha_{|N|-1} u_{|N|-1}$ and $\lambda - \lambda_1 = \alpha_{|N|} u_{|N|}$. From this it follows that $|\alpha_1| = \|\lambda_2\| = \bar{\lambda}\sqrt{|N|} < \sqrt{|N|}$, $|\alpha_k| \le \sqrt{\alpha_2^2 + \cdots + \alpha_{|N|-1}^2} = \|\lambda_1 - \lambda_2\| < d'_n = \mathcal{O}(n^{-\frac{1}{2}+\frac{1}{8|N|}})$ for all $k \in \{2, \ldots, |N|-1\}$, and $|\alpha_{|N|}| = \|\lambda - \lambda_1\| \le \frac{1}{n} + \frac{\sqrt{|N|}}{n} = \mathcal{O}(n^{-1})$. Hence,

$$\nu(\widetilde{H}^n(\ell,m)) = \mathcal{O}(1)\mathcal{O}(n^{(-\frac{1}{2}+\frac{1}{8|N|})(|N|-2)})\mathcal{O}(n^{-1}) \tag{2.128}$$

$$= \mathcal{O}(n^{-\frac{1}{2}|N|+\frac{1}{8}}). \tag{2.129}$$

For all $\lambda \in G_\varepsilon^n$ and $y \in C^n(\lambda)$, we get from

$$\bar{y} = \bar{\lambda} + (\bar{y} - \bar{\lambda}) \begin{cases} \geq \varepsilon - (1/n) > \frac{1}{2}\varepsilon, \\ \leq 1 - \varepsilon + (1/n) < 1 - \frac{1}{2}\varepsilon, \end{cases} \tag{2.130}$$

that $y \in G_{\frac{1}{2}\varepsilon}$. Moreover, we get

$$\min_{\hat{\lambda} \in V(\ell,m)} \|y - \hat{\lambda}\| \leq \|y - \lambda\| + \min_{\hat{\lambda} \in V(\ell,m)} \|\lambda - \hat{\lambda}\| \leq \frac{\sqrt{|N|}}{n} + \frac{1}{n}, \qquad \text{for all } \lambda \in H^n(\ell, m) \text{ and } y \in C^n(\lambda).$$

From this, (2.114) and (2.130), we get

$$\bigcup_{\lambda \in H^n(\ell,m)} C^n(\lambda) \subset \widetilde{H}^n(\ell, m), \qquad \text{for all } n \in \mathbb{N} \text{ such that } n > \frac{2}{\varepsilon}. \tag{2.131}$$

From (2.100) and (2.131) we get

$$n^{-|N|} \sharp(H^n(\ell, m)) \leq \nu(\widetilde{H}^n(\ell, m)). \tag{2.132}$$

Substituting (2.128)-(2.129) in (2.132) yields

$$\sharp(H^n(\ell, m)) = \mathcal{O}(n^{\frac{1}{2}|N|+\frac{1}{8}}). \tag{2.133}$$

Then, we obtain

$$\begin{aligned} \sum_{\lambda \in G_\varepsilon^n \cap D(n) \cap B(n)} h^n(\lambda) &\leq \sum_{\ell, m \in \{1,\ldots,p\}: \ell \neq m} \ \sum_{\lambda \in H^n(\ell,m)} h^n(\lambda) && (2.134) \\ &\leq \begin{pmatrix} p \\ 2 \end{pmatrix} \max_{\ell, m \in \{1,\ldots,p\}: \ell \neq m} \sharp(H^n(\ell, m)) \max_{\lambda \in G_\varepsilon} h^n(\lambda) && (2.135) \\ &= \begin{pmatrix} p \\ 2 \end{pmatrix} \mathcal{O}(n^{\frac{1}{2}|N|+\frac{1}{8}}) \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)}) && (2.136) \\ &= \mathcal{O}^\varepsilon(n^{\frac{5}{8}}), \end{aligned}$$

where (2.134) follows from (2.48), (2.135) follows from $\sharp(\{\ell, m \in \{1, \ldots, p\} : \ell \neq m\}) = \begin{pmatrix} p \\ 2 \end{pmatrix}$, and (2.136) follows from (2.133) and $h^n(\lambda) = \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)})$ for all $\lambda \in G_\varepsilon$. This concludes the proof. □

Proof of Proposition 2.6.4: It is sufficient to show this result for sufficiently large n. We get

$$K_i^n(R) = \sum_{\lambda \in G^n : \lambda_i < 1} t^n(\lambda) p_i^n(\lambda) \left[r\left(\lambda + (1/n) \cdot e_i\right) - r(\lambda) \right] \tag{2.137}$$

$$= \sum_{\lambda \in G_\varepsilon^n : \lambda_i < 1} t^n(\lambda) p_i^n(\lambda) \left[r\left(\lambda + (1/n) \cdot e_i\right) - r(\lambda) \right] + \mathcal{O}(\varepsilon) + \mathcal{O}(n^{-1}) \tag{2.138}$$

$$= \sum_{\lambda \in G_\varepsilon^n \cap D(n)} t^n(\lambda) p_i^n(\lambda) \left[r\left(\lambda + (1/n) \cdot e_i\right) - r(\lambda) \right] + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon\left(n^{-1}\right) \tag{2.139}$$

$$= \sum_{\lambda \in G_\varepsilon^n \cap D(n)} h^n(\lambda) \left[r\left(\lambda + (1/n) \cdot e_i\right) - r(\lambda) \right] + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{2.140}$$

$$= \sum_{\lambda \in [G_\varepsilon^n \cap D(n)] \setminus B(n)} h^n(\lambda) \left[r\left(\lambda + (1/n) \cdot e_i\right) - r(\lambda) \right] + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{2.141}$$

$$= \sum_{m=1}^{p} \sum_{\lambda \in [G_\varepsilon^n \cap D(n) \cap K_m] \setminus B(n)} h^n(\lambda) \frac{1}{n} E_{\mathbb{Q}_m}[X_i] + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{2.142}$$

$$= \sum_{m=1}^{p^*} \sum_{\lambda \in [G_\varepsilon^n \cap D(n) \cap K_m] \setminus B(n)} h^n(\lambda) \frac{1}{n} E_{\mathbb{Q}_m}[X_i] + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{2.143}$$

$$= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \frac{1}{n} \sum_{\lambda \in [G_\varepsilon^n \cap D(n) \cap K_m] \setminus B(n)} h^n(\lambda) + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}})$$

$$= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \frac{1}{n} \sum_{\lambda \in G_\varepsilon^n \cap D(n) \cap K_m} h^n(\lambda) + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}), \tag{2.144}$$

where (2.137) follows from Proposition 2.6.2, (2.138) follows from Lemma 2.A.11 and Lemma 2.A.13, (2.139) follows from Lemma 2.A.12 and Lemma 2.A.13, (2.140) follows from Lemma 2.A.13 and Lemma 2.A.14, (2.141) follows from Lemma 2.A.13 and Lemma 2.A.16, (2.142) follows from $[0,1]^N \setminus L(R) \subset B(n)$, (2.143) follows from Lemma 2.A.15, and (2.144) follows from Lemma 2.A.16. This concludes the proof. □

Lemma 2.A.17 *The function h^n is differentiable for any fixed $n \in \mathbb{N}$, and, moreover, we have for all $i \in N$ and $(\varepsilon, n, \lambda) \in Dom$ that*

$$\frac{\partial h^n}{\partial \lambda_i}(\lambda) = \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+1+\frac{1}{8|N|}}), \qquad \text{if } \lambda \in D'(n),$$

where $D'(n)$ is defined in (2.99).

Proof: Define the functions $f^n(\lambda) = -c(\bar{\lambda})n\|\lambda - \bar{\lambda} \cdot e_N\|^2$ and $g(\lambda) = (\bar{\lambda}(1-\bar{\lambda}))^{\frac{1}{2}(1-|N|)}$ for all $\lambda \in [0,1]^N$. Then, we obtain

$$\frac{\partial h^n}{\partial \lambda_i}(\lambda) = \frac{\partial f^n}{\partial \lambda_i}(\lambda) \cdot h^n(\lambda) + \frac{\partial g}{\partial \lambda_i}(\lambda) \cdot \frac{h^n(\lambda)}{g(\lambda)}, \qquad \text{for all } \lambda \in [0,1]^N \setminus \{e_\emptyset, e_N\}. \tag{2.145}$$

Moreover, we obtain the following approximations for all $\lambda \in G_\varepsilon \cap D'(n)$:

$$\begin{aligned}
\frac{\partial f^n}{\partial \lambda_i}(\lambda) &= -c(\bar\lambda)n\left[\sum_{k\neq i} 2\left(\lambda_k-\bar\lambda\right)\cdot -\frac{1}{|N|} + 2(\lambda_i-\bar\lambda)\left(1-\frac{1}{|N|}\right)\right] + \frac{1-2\bar\lambda}{2|N|(\bar\lambda(1-\bar\lambda))^2}n\|\lambda-\bar\lambda\cdot e_N\|^2\\
&= -c(\bar\lambda)n2\left(\lambda_i-\bar\lambda\right) + \frac{1-2\bar\lambda}{2|N|(\bar\lambda(1-\bar\lambda))^2}n\|\lambda-\bar\lambda\cdot e_N\|^2\\
&= \mathcal{O}^\varepsilon(n^{\frac{1}{2}+\frac{1}{8|N|}}) + \mathcal{O}^\varepsilon(n^{\frac{1}{4|N|}}) \qquad (2.146)\\
&= \mathcal{O}^\varepsilon(n^{\frac{1}{2}+\frac{1}{8|N|}}),\\
\frac{\partial g}{\partial \lambda_i}(\lambda) &= \mathcal{O}^\varepsilon(1),\\
(g(\lambda))^{-1} &= \mathcal{O}(1),\\
h^n(\lambda) &= \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)}),
\end{aligned}$$

where (2.146) follows from $|\lambda_i-\bar\lambda| \le \|\lambda-\bar\lambda\cdot e_N\| \le d'_n = \mathcal{O}(n^{-\frac{1}{2}+\frac{1}{8|N|}})$. Then, the result follows from substituting these equations in (2.145). □

Lemma 2.A.18 *Let $R \in \mathcal{R}$. Then, we have for all $m \in \{1,\dots,p\}$ that*

$$\sum_{\lambda\in G^n_\varepsilon\cap D(n)\cap K_m} h^n(\lambda) = n^{|N|}\sum_{\lambda\in G^n_\varepsilon\cap D(n)\cap K_m}\int_{C^n(\lambda)} h^n(\lambda^*)\,d\lambda^* + \mathcal{O}^\varepsilon(n^{\frac{5}{8}}),$$

where $C^n(\lambda)$ is defined in (2.98)

Proof: Let $\varepsilon > 0$. It is sufficient to show this result for all $n \in \mathbb{N}$ such that $n > \frac{2}{\varepsilon}$. Let $\lambda \in G^n_\varepsilon \cap D(n)$. From (2.114) and (2.130) it follows that

$$C^n(\lambda) \subset G_{\frac{1}{2}\varepsilon} \cap D'(n). \qquad (2.147)$$

We get from (2.147) and Lemma 2.A.17 that h^n is differentiable in λ^* for all $\lambda^* \in C^n(\lambda)$. Applying Taylor's theorem yields that

$$h^n(\lambda)-h^n(\lambda^*) = \sum_{i\in N}\frac{\partial h}{\partial\lambda_i}(\chi)(\lambda_i-\lambda_i^*),\ \text{for all } \lambda^*\in C^n(\lambda),\ \text{for some } \chi\in\operatorname{conv}\{\lambda,\lambda^*\}. \qquad (2.148)$$

Here, as $\chi \in C^n(\lambda)$, we get from Lemma 2.A.17 that

$$\frac{\partial h}{\partial \lambda_i}(\chi) = \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+1+\frac{1}{8|N|}}), \qquad \text{for all } i \in N. \qquad (2.149)$$

So, as $|\lambda_i-\lambda_i^*| \le n^{-1}$ for all $\lambda^* \in C^n(\lambda)$ and $i \in N$, we get from (2.148) and (2.149) that

$$h^n(\lambda) - h^n(\lambda^*) = |N|\mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+1+\frac{1}{8|N|}})n^{-1}$$

$$= \mathcal{O}^{\varepsilon}(n^{-\frac{1}{2}|N|+\frac{1}{8|N|}}),$$

for all $\lambda^* \in C^n(\lambda)$. From this, we directly get

$$h^n(\lambda) - n^{|N|} \int_{C^n(\lambda)} h^n(\lambda^*)\, d\lambda^* = \mathcal{O}^{\varepsilon}(n^{-\frac{1}{2}|N|+\frac{1}{8|N|}}), \qquad \text{for all } \lambda \in G^n_{\varepsilon} \cap D(n). \tag{2.150}$$

Moreover, from (2.115)-(2.118) we get

$$\sharp(G^n_{\varepsilon} \cap D(n) \cap K_m) \leq \sharp(G^n_{\varepsilon} \cap D(n)) \tag{2.151}$$
$$= \mathcal{O}(n^{\frac{1}{2}|N|+\frac{5}{8}-\frac{1}{8|N|}}). \tag{2.152}$$

Hence, from (2.150) and (2.151)-(2.152) it follows that

$$\left| \sum_{\lambda \in G^n_{\varepsilon} \cap D(n) \cap K_m} \left(h^n(\lambda) - n^{|N|} \int_{C^n(\lambda)} h^n(\lambda^*)\, d\lambda^* \right) \right| \leq \sharp(G^n_{\varepsilon} \cap D(n) \cap K_m) \mathcal{O}^{\varepsilon}(n^{-\frac{1}{2}|N|+\frac{1}{8|N|}})$$
$$= \mathcal{O}^{\varepsilon}(n^{\frac{5}{8}}).$$

This concludes the result. □

Lemma 2.A.19 *Let $R \in \mathcal{R}$. Then, we have for all $m \in \{1, \dots, p\}$ that*

$$\sum_{\lambda^* \in G^n_{\varepsilon} \cap D(n) \cap K_m} \int_{C^n(\lambda^*)} h^n(\lambda)\, d\lambda = \int_{G_{\varepsilon} \cap D(n) \cap K_m} h^n(\lambda)\, d\lambda + \mathcal{O}^{\varepsilon}(n^{-|N|+\frac{5}{8}}).$$

Proof: Let $\varepsilon > 0$, and define $D''(n) = D^{d''_n}$, where $d''_n = d_n - (\sqrt{|N|}/n)$. It is sufficient to show this result for all $n \in \mathbb{N}$ such that $n > \frac{2}{\varepsilon}$. Define $A = \bigcup_{\lambda^* \in G^n_{\varepsilon} \cap D(n) \cap K_m} C^n(\lambda^*)$ and $B = G_{\varepsilon} \cap D(n) \cap K_m$. Moreover, define

$$E^n_1 = B(n/(\sqrt{|N|}+1)) \cap G_{\frac{1}{2}\varepsilon}$$
$$E^n_2 = [G_{\varepsilon-(1/n)} \cap D'(n)] \backslash D''(n)$$
$$E^n_3 = [D'(n) \cap G_{\varepsilon-(1/n)}] \backslash G_{\varepsilon+(1/n)},$$

where the set $B(n)$ is defined in (2.102). We first show

$$(A \backslash B) \cup (B \backslash A) \subset E^n_1 \cup E^n_2 \cup E^n_3. \tag{2.153}$$

Let $y_1 \in A \backslash B$, so we have $y_1 \in C^n(\lambda)$ for some $\lambda \in G^n_{\varepsilon} \cap D(n) \cap K_m$. If $y_1 \notin K_m$, there is a $\lambda' \in [0,1]^N \backslash L(R)$ such that $\lambda' \in \text{conv}\{\lambda, y_1\}$ and, so, $y_1 \in E^n_1$. If $y_1 \notin D(n)$, we have according to (2.113) that $\|y_1 - \bar{y}_1 \cdot e_N\| < (\sqrt{|N|}/n) + d_n = d'_n$ and, so, $y_1 \in E^n_2$. If $y_1 \notin G^n_{\varepsilon}$, then $\bar{y}_1 < \varepsilon$ or $\bar{y}_1 > 1-\varepsilon$, and, so, we have according to (2.130) that $\varepsilon - (1/n) \leq \bar{y}_1 \leq 1 - (\varepsilon - (1/n))$. Hence, we have $y_1 \in E_3^n$. Now, let $y_2 \in B\backslash A$, so we have $y_2 \in G_\varepsilon \cap D(n) \cap K_m$, and there does not exist a $\lambda \in G_\varepsilon^n \cap D(n) \cap K_m$ such that $y_2 \in C^n(\lambda)$. Let λ be such that $y_2 \in C^n(\lambda)$. If $\lambda \notin K_m$, there exists a $\lambda' \in [0,1]^N \backslash L(R)$ such that $\lambda' \in \text{conv}\{\lambda, y_2\}$ and, so, $y_2 \in E_1^n$. If $\lambda \notin D(n)$, we get from the triangle inequality that $\|y_2 - \bar{y}_2 \cdot e_N\| \geq \|\lambda - \bar{\lambda} \cdot e_N\| - \|y_2 - \lambda\| \geq d_n - (\sqrt{|N|}/n) = d_n''$ and, so, $y_2 \notin D''(n)$. So, $y_2 \in E_2^n$. If $\lambda \notin G_\varepsilon^n$, then $\bar{\lambda} < \varepsilon$ or $\bar{\lambda} > 1-\varepsilon$, and, so, $\bar{y}_2 = \bar{\lambda} + (\bar{y}_2 - \bar{\lambda}) < \varepsilon + (1/n)$ or $\bar{y}_2 < 1 - (\varepsilon + (1/n))$. Hence, we have $y_2 \notin G_{\varepsilon+(1/n)}$ and, so, $y_2 \in E_3^n$. Hence, we have shown (2.153). Then, we get

$$\begin{aligned}\left|\int_A h^n(\lambda) d\lambda - \int_B h^n(\lambda)\, d\lambda\right| &\leq \int_{A\backslash B} h^n(\lambda) d\lambda + \int_{B\backslash A} h^n(\lambda)\, d\lambda && (2.154)\\ &\leq \int_{E_1^n \cup E_2^n \cup E_3^n} h^n(\lambda)\, d\lambda && (2.155)\\ &\leq \sum_{k=1}^{3} \int_{E_k^n} h^n(\lambda)\, d\lambda && (2.156)\\ &\leq \sum_{k=1}^{3} \nu(E_k^n) \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)}) && (2.157)\\ &= \mathcal{O}^\varepsilon(n^{-|N|+\frac{5}{8}}). && (2.158)\end{aligned}$$

Here, (2.154) follows from $\int_A h^n(\lambda) d\lambda - \int_B h^n(\lambda)\, d\lambda = \int_{A\backslash B} h^n(\lambda)\, d\lambda - \int_{B\backslash A} h^n(\lambda)\, d\lambda$, (2.155) follows from (2.153), (2.156) is a standard rule of integration, (2.157) follows from $h^n(\lambda) = \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)})$ for all $\lambda \in G_{\frac{1}{2}\varepsilon}$, and (2.158) follows from $\nu(E_1^n) = \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+\frac{1}{8}})$ (see (2.128)-(2.129)), and we get in a similar fashion as for (2.128)-(2.129) via a Gram-Schmidt process that

$$\begin{aligned}\nu(E_2^n) &= \mathcal{O}\left(\left(n^{-\frac{1}{2}+\frac{1}{8|N|}} + (\sqrt{|N|}/n)\right)^{|N|-1} - \left(n^{-\frac{1}{2}+\frac{1}{8|N|}} - (\sqrt{|N|}/n)\right)^{|N|-1}\right)\\ &= \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+\frac{1}{8}}),\\ \nu(E_3^n) &= \mathcal{O}((n^{-\frac{1}{2}+\frac{1}{8|N|}} + (\sqrt{|N|}/n))^{|N|-1} n^{-1})\\ &= \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|-\frac{3}{8}}).\end{aligned}$$

This concludes the proof. □

Lemma 2.A.20 *For all $t \in (0,1)$ it holds that*

$$\int_0^{n^{\frac{1}{4|N|}}/2t(1-t)} e^{-s} s^{\frac{1}{2}(|N|-3)}\, ds = \Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right) + \mathcal{O}(n^{-\infty}),$$

where Γ is the Gamma function:

$$\Gamma(\kappa) = \int_0^\infty e^{-t} t^{\kappa-1}\, dt, \qquad \textit{for all } \kappa > 0. \tag{2.159}$$

Proof: We get

$$\Gamma\left(\frac{1}{2}|N|-\frac{1}{2}\right)-\int_0^{n^{\frac{1}{4|N|}}/2t(1-t)} e^{-s}s^{\frac{1}{2}(|N|-3)}\,ds=\int_{n^{\frac{1}{4|N|}}/2t(1-t)}^{\infty} e^{-s}s^{\frac{1}{2}(|N|-3)}\,ds \tag{2.160}$$

$$\leq K\int_{n^{\frac{1}{4|N|}}/2t(1-t)}^{\infty} e^{-\frac{1}{2}s}\,ds \tag{2.161}$$

$$= K2e^{-n^{\frac{1}{4|N|}}/4t(1-t)} \tag{2.162}$$

$$\leq K2e^{-n^{\frac{1}{4|N|}}} \tag{2.163}$$

$$= \mathcal{O}(n^{-\infty}), \tag{2.164}$$

where $K>0$. Here, (2.160) is a standard integration rule, (2.161) follows from that there exists a constant $K>0$ such that $e^{-s}s^{\frac{1}{2}(|N|-3)}<Ke^{-\frac{1}{2}s}$ for all $s>1$, (2.162) follows from $\int_a^b e^{-\frac{1}{2}s},ds=-2(e^{-\frac{1}{2}b}-e^{-\frac{1}{2}a})$ for all $a\leq b$, (2.163) follows from $4t(1-t)\leq 1$ for all $t\in(0,1)$, and (2.164) follows from the fact that $(e^{-1})^{n^{\frac{1}{4|N|}}}=\mathcal{O}(n^{-\infty})$. This concludes the proof. □

Proof of Proposition 2.6.5: We get

$$K_i^n(R)=\sum_{m=1}^{p^*}E_{\mathbb{Q}_m}[X_i]\frac{1}{n}\sum_{\lambda\in G_\varepsilon^n\cap D(n)\cap K_m}h^n(\lambda)+\mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{2.165}$$

$$=\sum_{m=1}^{p^*}E_{\mathbb{Q}_m}[X_i]n^{|N|-1}\sum_{\lambda\in G_\varepsilon^n\cap D(n)\cap K_m}\int_{C^n(\lambda)}h^n(\lambda^*)\,d\lambda^*+\mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{2.166}$$

$$=\sum_{m=1}^{p^*}E_{\mathbb{Q}_m}[X_i]n^{|N|-1}\int_{G_\varepsilon\cap D(n)\cap K_m}h^n(\lambda)\,d\lambda+\mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{2.167}$$

$$=\sum_{m=1}^{p^*}E_{\mathbb{Q}_m}[X_i]n^{\frac{1}{2}(|N|-1)}(2\pi)^{\frac{1}{2}(1-|N|)}|N|^{-\frac{1}{2}} \tag{2.168}$$
$$\cdot\int_{G_\varepsilon\cap D(n)\cap K_m}\left(e^{-\frac{1}{2\bar\lambda(1-\bar\lambda)}n\|\lambda-\bar\lambda\cdot e_N\|^2}\right)(\bar\lambda(1-\bar\lambda))^{\frac{1}{2}(1-|N|)}\,d\lambda+\mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}})$$

$$=\sum_{m=1}^{p^*}E_{\mathbb{Q}_m}[X_i]n^{\frac{1}{2}(|N|-1)}(2\pi)^{\frac{1}{2}(1-|N|)} \tag{2.169}$$
$$\cdot\int_\varepsilon^{1-\varepsilon}\int_0^{d_n}\int_{S_m}e^{-\frac{1}{2t(1-t)}r^2n}(t(1-t))^{\frac{1}{2}(1-|N|)}r^{|N|-2}\,d\omega\,dr\,dt+\mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}})$$

$$=\sum_{m=1}^{p^*}E_{\mathbb{Q}_m}[X_i]n^{\frac{1}{2}(|N|-1)}(2\pi)^{\frac{1}{2}(1-|N|)}\mu(S_m) \tag{2.170}$$
$$\cdot\int_\varepsilon^{1-\varepsilon}\int_0^{d_n}e^{-\frac{1}{2t(1-t)}r^2n}(t(1-t))^{\frac{1}{2}(1-|N|)}r^{|N|-2}\,dr\,dt+\mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}})$$

$$= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] n^{\frac{1}{2}(|N|-1)} (2\pi)^{\frac{1}{2}(1-|N|)} \phi_m 2 \frac{\pi^{-\frac{1}{2}(1-|N|)}}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)} \tag{2.171}$$

$$\cdot \int_{\varepsilon}^{1-\varepsilon} \int_0^{d_n} e^{-\frac{1}{2t(1-t)} r^2 n} (t(1-t))^{\frac{1}{2}(1-|N|)} r^{|N|-2}\, dr\, dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}})$$

$$= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \phi_m n^{\frac{1}{2}(|N|-1)} 2^{1\frac{1}{2} - \frac{1}{2}|N|} \frac{1}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)} \left(\frac{2}{n}\right)^{\frac{1}{2}(|N|-2)} \tag{2.172}$$

$$\cdot \int_{\varepsilon}^{1-\varepsilon} \int_0^{n^{\frac{1}{4|N|}}/2t(1-t)} e^{-s} s^{\frac{1}{2}(|N|-2)} (t(1-t))^{-\frac{1}{2}} \sqrt{\frac{t(1-t)}{2ns}}\, ds\, dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}})$$

$$= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \phi_m \frac{1}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)} \tag{2.173}$$

$$\cdot \int_{\varepsilon}^{1-\varepsilon} \int_0^{n^{\frac{1}{4|N|}}/2t(1-t)} e^{-s} s^{\frac{1}{2}(|N|-3)}\, ds\, dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}})$$

$$= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \phi_m \frac{1}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)} \int_{\varepsilon}^{1-\varepsilon} \left(\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right) + \mathcal{O}(n^{-\infty}) \right) dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}}) \tag{2.174}$$

$$= \sum_{m=1}^{p^*} \phi_m E_{\mathbb{Q}_m}[X_i] + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}}). \tag{2.175}$$

Here, (2.165) follows from Proposition 2.6.4, (2.166) follows from Lemma 2.A.18, (2.167) follows from Lemma 2.A.19, (2.168) follows from (2.38), (2.169) follows from the polar coordinate transformation $\lambda = t \cdot e_N + r\omega$ and $d\lambda = r^{|N|-2}|N|^{\frac{1}{2}} d(t, r, \omega)$, (2.170) follows from the fact that $\int_{S_m} d\omega = \mu(S_m)$, (2.171) follows from the well-known result

$$\mu(S_m) = \phi_m \mu(S) = \phi_m 2 \frac{\pi^{-\frac{1}{2}(1-|N|)}}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)},$$

where Γ is defined in (2.159), (2.172) follows from the transformation $s = \frac{r^2 n}{2t(1-t)}$ and $dr = \sqrt{\frac{t(1-t)}{2ns}} ds$, (2.173) follows from canceling of some terms, and (2.174) follows from Lemma 2.A.20. This concludes the proof. □

Proof of Theorem 2.6.6: Let $R \in \mathcal{R}$. From Proposition 2.6.5, we get for all $n \in \mathbb{N}$ and $\varepsilon > 0$ that

$$\left| K_i^n(R) - \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \phi_m \right| < K\varepsilon + L_\varepsilon n^{-\frac{1}{4}}, \qquad \text{where } K, L_\varepsilon > 0.$$

Pick an $\eta > 0$. Let $\varepsilon = \frac{\eta}{2K}$ and N_η such that $L_\varepsilon N_\eta^{-\frac{1}{4}} = \frac{1}{2}\eta$. Then, we have for all $n > N_\eta$ that

$$\left| K_i^n(R) - \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i]\phi_m \right| < \eta.$$

This concludes the proof. □

Proof of Corollary 2.6.8: For all $R \in \mathcal{R}'$, we have that the risk capital function r is partially differentiable in e_N and, so, $p^* = 1$. Then, we get

$$\begin{aligned} WAS(R) &= (E_{\mathbb{Q}_1}[X_i])_{i\in N} && (2.176)\\ &= AS(R), && (2.177) \end{aligned}$$

where (2.176) follows from Theorem 2.6.6 and (2.41), and (2.177) follows from (2.20). This concludes the proof. □

Proof of Proposition 2.6.9: Let $R \in \mathcal{R}$. Suppose $\phi_m > \frac{1}{2}$ for an $m \in \{1, \ldots, p^*\}$ and $p^* > 1$. Note that $\mu(S_\ell \cap S_{\ell'}) = 0$ for all $\ell, \ell' \in \{1, \ldots, p^*\}$ such that $\ell \neq \ell'$, where μ is the surface area on S. Moreover, we have $z \in S$ if and only if $-z \in S$. From these two results, we obtain that there must exists a $z \in S_m$ such that $f_{\mathbb{Q}_m}(z) > \max_{\ell\in\{1,\ldots,p^*\}\setminus m} f_{\mathbb{Q}_\ell}(z)$ and $-z \in S_m$. So, let $z \in S_m$ be such that $f_{\mathbb{Q}_m}(z) > \max_{\ell\in\{1,\ldots,p^*\}\setminus m} f_{\mathbb{Q}_\ell}(z)$. By linearity of $f_{\mathbb{Q}}$, we get $f_{\mathbb{Q}}(-z) = -f_{\mathbb{Q}}(z)$. If $-z \in S_m$, we get $f_{\mathbb{Q}_m}(-z) \geq \max_{\ell\in\{1,\ldots,p^*\}\setminus m} f_{\mathbb{Q}_\ell}(-z)$ or, equivalently,

$$f_{\mathbb{Q}_m}(z) \leq \min_{\ell\in\{1,\ldots,p^*\}\setminus m} f_{\mathbb{Q}_\ell}(z),$$

which is a contradiction. Hence, we have $\phi_m \leq \frac{1}{2}$ or $p^* = 1$ so that $\phi_m = 1$. This concludes the result. □

Proof of Proposition 2.6.10: Let $R \in \mathcal{R}$. It holds that

$$\begin{aligned} FCore(R) &= \text{conv}\{(E_{\mathbb{Q}}[X_i])_{i\in N} : \mathbb{Q} \in Q^*(\rho)\} && (2.178)\\ &= \text{conv}\{(E_{\mathbb{Q}_m}[X_i])_{i\in N} : m \in \{1, \ldots, p^*\}\}, && (2.179) \end{aligned}$$

where $Q^*(\rho)$ is defined in (2.19). Here, (2.178) follows from Theorem 2.3.5, and (2.179) follows from Definition 2.4.3. Moreover, we showed in Theorem 2.6.6 that $WAS(R)$ is given by a convex combination of allocation $(E_{\mathbb{Q}_m}[X_i])_{i\in N})$ for all $m \in \{1, \ldots, p^*\}$. Hence, we have $WAS(R) \in FCore(R)$. □

Chapter 3

Bargaining for Risk Redistributions: The Case of Longevity Risk

This chapter is based on Boonen, De Waegenaere and Norde (2012a).

3.1 Introduction

There is an increasing need for hedging non-marketed risk. Markets serve as mechanism to reallocate risk among firms. However, if for a class of risks markets are non-existent or if there are obstacles to trade, firms could benefit from risk-sharing by trading with each other. If there are two firms involved in this trade, it is called Over-the-Counter (OTC). In this chapter, we investigate the extent to which involved parties can benefit from such risk redistributions.

Although the model that we introduce allows for any type of risk, our focus in this chapter is on redistribution of longevity risk. Longevity risk is the systematic risk in life-contingent liabilities that arises from the fact that death rates in a population change in an unpredictable way. This risk is a major concern for pension funds, since their liabilities are directly linked with longevity. Exposure to longevity risk can be rather substantial for pension funds, as shown by, e.g., Coughlan et al. (2007) and Hári, De Waegenaere, Melenberg and Nijman (2008). Reinsurance contracts do exist, but the capacity of reinsurance is limited (see, e.g., OECD, 2005). Moreover, there is still considerable uncertainty regarding the price of longevity risk. Despite a very active and growing literature on pricing of longevity risk (see, e.g., Bauer, Börger and Russ, 2010), lack of consensus regarding accurate pricing hampers trade.[15] Equilibrium prices do not exist, and, as a consequence, longevity-linked contracts are mainly traded Over-the-Counter.

[15] Blake, Cairns and Dowd (2006) address the main obstacles for trading longevity linked products in the market.

As pointed out by Dowd, Blake, Cairns and Dawson (2006), the market for Over-the-Counter survivor swaps is expected to grow fast. In this chapter, we analyze opportunities for pension funds and insurers to benefit from mutual redistribution of risk via Over-the-Counter trade. Specifically, parties bargain for a "fair" risk redistribution and a "fair" price. This option is potentially attractive in case redistribution of risk allows parties to benefit from natural hedge potential that arises from combining risks that are not perfectly correlated.

Redistribution of risk in order to benefit from natural hedge potential has been investigated also in Tsai, Wang and Tzeng (2010) and Wang, Huang, Yang and Tsai (2010). Our approach differs in three ways. First, the existing literature focuses on reducing the risk in the present value of the liability payments over complete run-off. We consider instead the case where pension funds and insurers redistribute risk in order to reduce the volatility of their *Net Asset Value* (NAV) at a prespecified future date. The Net Asset Value is defined as the difference between the value of the assets and the value of the liabilities.[16] Second, we extend existing literature by allowing for the case where pension funds and insurance companies may have heterogeneous beliefs regarding the probability distribution of future mortality rates. There exists a relatively large variety of models that can be used to predict future mortality rates. The seminal work of Lee and Carter (1992) gave rise to an active literature on mortality modeling (see, e.g., Brouhns, Denuit and Vermunt, 2002; Cairns et al., 2006, 2007, 2008; Cossette, Delwarde, Denuit, Guillot and Marceau, 2007; Plat, 2009; Dowd et al., 2010). Existing literature also shows that the effect of model risk could be substantial, i.e., predictions for future mortality rates may differ significantly when different models are used, or when model parameters are estimated on different datasets (see, e.g., Dowd, Blake and Cairns, 2008). In our model, we therefore allow the parties who wish to redistribute risk to "agree to disagree" on the appropriate forecast model. Third, our approach differs from existing models in that we model the risk redistribution as the outcome of a bargaining process in which the involved parties bargain for a reallocation of risk that benefits all. They will agree to a particular redistribution only if they benefit weakly in expected utility terms. Moreover, when more than two parties are involved, they will only reach a mutual agreement if no subset of parties can be better off by splitting off and instead redistributing the risk amongst each other.

In the first part of the chapter, we characterize Pareto optimal redistributions of risk in a relatively general setting. There exists a large literature on Pareto optimal redistributions of risk under various assumptions regarding the preference relations, including standard expected utility (see, e.g., Borch 1960, 1962; Wilson, 1968; Raviv, 1979), monetary utility functions (see, e.g., Filipovíc and Kupper 2008, Jouini, Schachermayer and Touzi 2008), and convex or coherent risk measures (see, e.g. Burgert and Rüschendorf, 2008; Filipovic and Svindland, 2008, and Chapter 4 of this dissertation). We consider a setting with expected utility preferences and heterogeneous beliefs regarding the underlying probability distribution. There is relatively little literature that

[16]The focus on Net Asset Value is in line with current regulation. Under Solvency II, for example, the regulator requires that the current level of assets is sufficient to reduce the probability of a negative Net Asset Value on a one year horizon to a sufficiently low level.

focuses explicitly on the effects of heterogeneous beliefs regarding the probability distribution; exceptions are Wilson (1968), Riddell (1981), and Acciaio and Svindland (2009). We show that heterogeneity can make redistribution more attractive, i.e., it is more likely that parties are able to achieve Pareto improvement via redistribution when they disagree on the underlying probability distribution. Moreover, because there are in general infinitely many Pareto optimal risk redistributions, we use the constrained Nash bargaining solution to characterize the risk redistribution that arises from a bargaining process in which the involved parties bargain for the specific Pareto optimal redistribution that will be implemented. In the second part of this chapter, we apply these models to investigate the potential benefits from redistributing longevity risk between pension funds and insurers. We model redistribution of risk via longevity swaps with a predetermined maturity date, and quantify the benefits from a particular swap contract by the maximum premium that the pension fund or insurer would have been willing to pay to obtain the same risk reduction via a reinsurance contract (i.e., the indifference price). It is often argued that major obstacles to trade are lack of agreement regarding the probability distribution of future mortality rates, lack of capacity in the insurance market to effectively hedge longevity risk in annuity portfolios, and reluctance to engage in contracts with a long horizon. Our results suggest that the benefits from redistribution are significant, even when the insurer is small relative to the pension fund, and even when the horizon of the swap contract is relatively short. Moreover, the results also suggest that the benefits are significantly larger when parties disagree regarding the underlying probability distribution. With homogeneous beliefs and a death benefit insurer having a portfolio with a best estimate value equal to 20% of the best estimate value of the annuity portfolio of the pension fund, the benefits of the optimal swap contract are significant. Over a horizon of ten years, the premium the pension fund (insurer) would be willing to pay to obtain the same degree of risk reduction amounts to 9.9% (20.1%) of the best estimate value of the liabilities. When the two parties have heterogeneous beliefs, these benefits increase to 11.0% (22.1%).

The remainder of this chapter is organized as follows. We introduce the model for redistribution of risk in Section 3.2. In Section 3.3, we characterize risk redistributions that satisfy some desirable properties, and model the outcome of the bargaining process in which the involved parties agree on a particular risk redistribution. In Section 3.4, we use this model to numerically illustrate the extent to which pension funds and life insurance companies can benefit from redistributing longevity risk. Section 3.5 concludes.

3.2 Risk redistribution

When there is no liquid market for a certain class of risk, firms can approach each other to redistribute risk with each other. In this case, both the risk redistribution and the corresponding prices are set via a bargaining process. If firms behave myopically, i.e., if each firm bargains for a redistribution that maximizes its own utility, the firms will typically not reach an agreement even though each firm knows that if they behave cooperatively, they can all benefit.[17] In this chapter, we therefore consider a setting in which firms behave cooperatively, and strive to reach a redistribution of risk that benefits all. In Subsection 3.2.1, we first introduce some notation and assumptions. In Subsection 3.2.2, we define some properties that the risk redistribution should satisfy if firms behave rationally.

3.2.1 The model

We consider a setting in which firms aim to redistribute risk with each other. The specific setting is as follows:

- Risk is redistributed between a finite number of firms. The set of firms is denoted by N.
- The risk profile of firm i *prior to redistribution* is a random variable X_i on a set of states of nature. The realization of X_i in a specific state of the world represents the future wealth of firm i at a prespecified future date, T, in that state of the world. The set of possible states of the world is denoted by Ω ; Ω is finite. Firms can "agree to disagree" on the probability measure on the state space Ω. The subjective probability measure of firm $i \in N$ is denoted by $\mathbb{P}_i : \Omega \to \mathbb{R}_{++}$, where $\sum_{\omega\in\Omega} \mathbb{P}_i(\{\omega\}) = 1$.
- To evaluate risk profiles, firm $i \in N$ uses a von Neumann-Morgenstern utility function

  $$u_i : D_i \to \mathbb{R},$$

  where D_i is the domain of the utility function. $D_i = (a_i, b_i)$ is an open interval, where $a_i < b_i$ and possibly $a_i = -\infty$ or $b_i = +\infty$. Moreover, u_i is twice continuously differentiable, $u_i'(\cdot) > 0$ and $u_i''(\cdot) < 0$ on D_i for all $i \in N$, and $\lim_{x\to a_i} u_i'(x) = \infty$, and $\lim_{x\to b_i} u_i'(x) = 0$, for all $i \in N$. The risk profile of firm i, X_i, has realizations in D_i, i.e., $X_i : \Omega \to D_i$.
- Firms can change their risk profile through redistribution of risk. We do not impose any restrictions on the redistribution, i.e., we allow firms to bargain for any risk redistribution that leads to *posterior risk profiles* $X_i^{\text{post}} : \Omega \to D_i$, for $i \in N$, that satisfy

  $$\sum_{i\in N} X_i^{\text{post}} = \sum_{i\in N} X_i. \tag{3.1}$$

  We refer to posterior risk profiles that satisfy (3.1) as *feasible* posterior risk profiles. For any risk profile $X_i^{\text{post}}$of firm i, the corresponding expected utility gain of firm $i \in N$ is denoted $\Delta U_i(X_i^{\text{post}})$, i.e.,

  $$\Delta U_i(X_i^{\text{post}}) = \mathbb{E}_{\mathbb{P}_i}\left[u_i(X_i^{\text{post}}) - u_i(X_i)\right], \text{ for all } i \in N. \tag{3.2}$$

[17] In game theory, this phenomenon is known as the Prisoner's dilemma (see, e.g., Boonen, 2010).

- There is complete information about the risk profiles, utility functions and probability measures of all firms.

Throughout the remainder of the chapter, we let the set of firms N, the prior risk profiles $(X_i)_{i\in N}$, the utility functions $(u_i)_{i\in N}$, and the subjective probability measures $(\mathbb{P}_i)_{i\in N}$ be given. We investigate the extent to which the firms will be able to agree on a redistribution of risk that leads to posterior risk profiles $(X_i^{\text{post}})_{i\in N}$. For notational convenience, we use the following shorthand notations:

- For any $S \subseteq N$, we denote $D(S)$ for the set of possible risk profiles of the firms in S, i.e., $D(S) = \prod_{i\in S} D_i^{\Omega}$.
- The vector relation $x \gneqq y$ implies $x \geq y$ componentwise and $x \neq y$.

We note that while our focus is on expected utility preferences, Föllmer and Schied (2002) show that maximizing expected utility is equivalent to minimizing a convex risk measure. This interpretation is particularly relevant in Section 3.4, where we apply the model to redistribution of risk in the net asset value of pension funds and insurers. In Chapter 4 of this dissertation, we study the problem how risk should be optimally redistributed if all firms use risk measures to evaluate risk.

3.2.2 Properties of risk redistributions

As argued before, we consider the case where firms can bargain for any risk redistribution that leads to feasible posterior risk profiles. When cooperation between the firms is not mandated, however, the firms will only agree to a particular redistribution if that redistribution satisfies some properties. If firms act rationally, they will only agree to a particular redistribution if the redistribution does not make them worse off in expected utility terms. Moreover, all firms have incentive to not engage in a particular redistribution if there exists another feasible redistribution that makes each firm weakly better off, and at least one firm strictly better off. To formally define these properties, we first introduce some notation.

For any set of firms $S \subseteq N$, we therefore denote $\mathcal{F}(S)$ for the set of vectors of posterior risk profiles that the firms in S can reach if they redistribute their risks amongst each other, i.e.,

$$\mathcal{F}(S) = \left\{ (X_i^{\text{post}})_{i\in S} \in D(S) : \sum_{i\in S} X_i^{\text{post}} = \sum_{i\in S} X_i \right\}. \tag{3.3}$$

Moreover, for any subset $S \subseteq N$, we denote $\mathcal{NI}(S)$ for the set of posterior risk profiles $(X_i^{\text{post}})_{i\in S} \in D(S)$ for the firms in S such that there does not exist another feasible redistribution that yields weakly higher expected utility gains for each firm, and a strictly higher expected utility gain for at least one firm. Formally,

$$\mathcal{NI}(S) = \left\{(X_i^{\text{post}})_{i\in S} \in D(S) : \nexists (\widetilde{X}_i^{\text{post}})_{i\in S} \in \mathcal{F}(S) \text{ s.t. } (\Delta U_i(\widetilde{X}_i^{\text{post}}))_{i\in S} \gneqq (\Delta U_i(X_i^{\text{post}}))_{i\in S}\right\}. \tag{3.4}$$

If firms act rationally, they will only agree to redistributions that lead to posterior risk profiles that satisfy the following two properties.

Definition 3.2.1 A vector of posterior risk profiles $(X_i^{\text{post}})_{i\in N} \in D(N)$ satisfies *Individual Rationality* if each firm weakly benefits in expected utility terms, i.e., $\Delta U_i(X_i^{\text{post}}) \geq 0$ for all $i \in N$.

Equivalently, $(X_i^{\text{post}})_{i\in N}$ satisfies *Individual Rationality* if

$$(X_i^{\text{post}})_{i\in N} \in \mathcal{IR}(N), \tag{3.5}$$

where

$$\mathcal{IR}(N) = \prod_{i\in N} \mathcal{NI}(\{i\}). \tag{3.6}$$

Definition 3.2.2 A vector of posterior risk profiles $(X_i^{\text{post}})_{i\in N} \in D(N)$ satisfies *No Pareto Improvement* if there does not exist a vector of feasible posterior risk profiles that yields weakly higher expected utility gains for all firms, and a strictly higher expected utility gain for at least one firm. Formally, $(X_i^{\text{post}})_{i\in N}$ satisfies *No Pareto Improvement* if

$$(X_i^{\text{post}})_{i\in N} \in \mathcal{NI}(N). \tag{3.7}$$

Clearly, no firm would be willing to accept a redistribution that yields a posterior risk profile that does not satisfy *Individual Rationality*. Moreover, all firms have an incentive to choose a redistribution that satisfies *No Pareto Improvement*. Together, these conditions guarantee that no firm is better off if it does not engage in the redistribution (and, hence, keeps its prior risk profile), and that the firms collectively cannot reach an alternative distribution that makes at least one firm better off without harming the other firms.

No Pareto Improvement does not rule out the possibility that a subset of firms could be better off if they decide to redistribute risk amongst each other, excluding the other firms from the negotiation. Allowing more firms to cooperate in the redistribution has the potential advantage that the set of posterior risk profiles that a firm can reach increases. However, it may have the drawback that each firm negotiates with a larger number of other firms who each want to benefit from the redistribution. It is therefore not a priori clear that firms cannot do better by excluding some firms from the negotiation. Therefore, we consider the following condition.

Definition 3.2.3 A vector of posterior risk profiles $(X_i^{\text{post}})_{i \in N} \in D(N)$ satisfies *Stability* if for any subset $S \subseteq N$ of firms, there does not exist a redistribution of risk that satisfies $\sum_{i \in S} \widetilde{X}_i^{\text{post}} = \sum_{i \in S} X_i$, and that is weakly preferred by all firms in S and strictly preferred by at least one firm in S. Formally, *Stability* is satisfied if

$$(X_i^{\text{post}})_{i \in S} \in \mathcal{NI}(S), \text{ for all } S \subseteq N. \tag{3.8}$$

If a redistribution leads to posterior risk profiles that do not satisfy *Stability* and firms act rationally, then they will not agree to that redistribution because there exists a subset of firms $S \subseteq N$ that can be better off when they exclude the other firms from the negotiation and redistribute their risk amongst each other. In game-theoretic terms, this condition implies that the reallocation is an element of the core of the corresponding game. The following result follows straightforwardly.

Proposition 3.2.4 *It holds that:*

(i) *Stability implies No Pareto Improvement and Individual Rationality.*

(ii) *If risk is redistributed between two firms, i.e., if $|N| = 2$, then Stability is satisfied if and only if Individual Rationality and No Pareto Improvement are satisfied.*

In the next section, we characterize the set of posterior risk profiles that satisfy *No Pareto Improvement*, as well as the set of posterior risk profiles that satisfy the stronger *Stability* condition.

3.3 Pareto optimality and stability

As discussed in the previous section, if firms act rationally, they will only agree to redistributions of risk if no subset of firms could be better off by excluding the other firms from the negotiation and instead redistributing the risk amongst each other. Proposition 3.2.4(i) shows that a necessary condition for this *Stability* condition to be satisfied is that the redistribution satisfies *No Pareto Improvement.* In Subsection 3.3.1 we therefore first characterize the set of Pareto optimal redistributions. In Subsection 3.3.2 we show that, among the set of infinitely many redistributions that are Pareto optimal, there exist feasible risk redistributions that satisfy the stronger *Stability* condition. In general, however, the set of redistributions that satisfy *Stability* is not single-valued, and the issue arises which redistribution is selected. In Subsection 3.3.3 we model the choice of a particular redistribution as the outcome of a bargaining process that weighs the benefits of all involved parties.

3.3.1 Pareto optimal risk redistributions

A vector of posterior risk profiles $(X_i^{\text{post}})_{i \in N}$ is *Pareto optimal* if it is feasible, and satisfies *No Pareto Improvement*, i.e., the set of Pareto optimal posterior risk profiles is given by

$$\begin{aligned}&\mathcal{PO}(N) = \mathcal{F}(N) \cap \mathcal{NI}(N)\\ &= \left\{ (X_i^{\text{post}})_{i \in N} \in \mathcal{F}(N) : \nexists (\widetilde{X}_i^{\text{post}})_{i \in N} \in \mathcal{F}(N) \text{ s.t. } (\Delta U_i(\widetilde{X}_i^{\text{post}}))_{i \in N} \gneqq (\Delta U_i(X_i^{\text{post}}))_{i \in N} \right\}.\end{aligned} \tag{3.9}$$

All firms have an incentive to choose a redistribution in the set of Pareto optimal redistributions. If a redistribution is not Pareto optimal, there exists another vector of feasible posterior risk profiles that yields weakly higher expected utility gains for all firms, and a strictly higher expected utility gain for at least one firm.

Borch (1962) characterizes Pareto optimal risk reallocations for the case where the domain is unrestricted (i.e., $D(N) = \mathbb{R}^N$), and firms have homogeneous beliefs regarding the probability distribution (i.e., $\mathbb{P}_i = \mathbb{P}$ for all $i \in N$). He shows that a vector of feasible posterior risk profiles $(X_i^{\text{post}})_{i \in N}$ is Pareto optimal if and only if there exists a vector $k \in \mathbb{R}^N_{++}$, such that it maximizes the weighted sum of the expected utility gains:[18]

$$\begin{array}{rccl}(X_i^{\text{post}})_{i \in N} & \in & \text{argmax} & \sum_{i \in N} k_i \Delta U_i(\widetilde{X}_i^{\text{post}})\\ & & \text{s.t.} & (\widetilde{X}_i^{\text{post}})_{i \in N} \in \mathcal{F}(N).\end{array} \tag{3.10}$$

Thus, the set of Pareto optimal posterior risk profiles can be determined by solving optimization problem (3.10) for all strictly positive vectors $k \in \mathbb{R}^N_{++}$. Wilson (1968) shows that this result holds more generally in settings where there are domain restrictions, and agents may have heterogeneous beliefs regarding the underlying probability distribution. A straightforward application of Wilson's result in our setting yields the following theorem.

Theorem 3.3.1 *It holds that:*

(i) $(X_i^{post})_{i \in N} \in \mathcal{PO}(N)$ *if and only if there exists a* $k \in \mathbb{R}^N_{++}$ *such that*

$$k_i u_i'(X_i^{post}(\omega))\mathbb{P}_i(\{\omega\}) = k_j u_j'(X_j^{post}(\omega))\mathbb{P}_j(\{\omega\}), \qquad \text{for all } \omega \in \Omega, i, j \in N, \tag{3.11}$$

and

$$\sum_{i \in N} X_i^{post}(\omega) = \sum_{i \in N} X_i(\omega), \qquad \text{for all } \omega \in \Omega, \tag{3.12}$$

where $X_i^{post}(\omega)$ *denotes the realization of* $X_i^{post} \in D(N)$ *in state* $\omega \in \Omega$ *for all* $i \in N$.

[18] Borch (1962) considers utility levels. We consider expected utility gains, i.e., the difference between the expected utility of the posterior risk profile and the prior risk profile. This, however, does not affect the result.

(ii) For every $k \in \mathbb{R}^N_{++}$, there exists a unique solution to the system of equations (3.11) and (3.12).

The above theorem shows that the set of Pareto optimal posterior risk profiles can be found by solving the system of equations (3.11) and (3.12) for every $k \in \mathbb{R}^N_{++}$. Note that without loss of generality, we can impose as normalization that $k_i = 1$ for some $i \in N$.

The proposition shows the effect of heterogeneous beliefs regarding the probability distribution over the states of the world. For a given $k \in \mathbb{R}^N_{++}$, the corresponding Pareto optimal posterior risk profile for firm $i \in N$ in state $\omega \in \Omega$ is increasing in $\mathbb{P}_i(\{\omega\})$ and decreasing in $\mathbb{P}_j(\{\omega\})$ for $j \neq i$. The reason is that if a firm assigns a higher subjective probability to a future state than other firms, this firm overvalues the outcome in this state and, so, it is Pareto optimal to assign a higher pay-off to that firm in this state. If all probability measures coincide, the effects cancel out as in Borch (1962).

Heterogeneity regarding the subjective probability distributions also has non-trivial effects on the structure of Pareto optimal redistributions. Gerber and Pafumi (1998) show that when firms have homogeneous beliefs, and use either exponential utility functions, or a power utility function with the same risk aversion parameter for all firms, the Pareto optimal posterior risk profiles are of the form

$$X_i^{\text{post}}(\omega) = \delta_i \sum_{j \in N} X_j(\omega) + d_i, \text{ for all } \omega \in \Omega \text{ and } i \in N, \tag{3.13}$$

where $\delta_i \in [0,1]$, $i \in N$, are non-negative fractions satisfying $\sum_{i \in N} \delta_i = 1$, and $d_i \in \mathbb{R}$, $i \in N$, are side-payments satisfying $\sum_{i \in N} d_i = 0$. Thus, in these cases, the Pareto optimal risk redistributions consist of a proportional redistribution of risk and deterministic side-payments. In the next example we show the effect of heterogeneous beliefs on these Pareto optimal posterior risk profiles.

Example 3.3.2 We consider Pareto optimal risk redistributions for the case where firms use either exponential utility functions, power utility functions or quadratic utility functions, allowing for heterogeneous beliefs regarding the probability distribution over the states of the world. We consider Pareto optimal redistributions for all firms.

First, consider the case where the firms use exponential utility functions, i.e., the utility function of firm $i \in N$ is given by

$$u_i(x) = -\frac{1}{\lambda_i} \exp(-\lambda_i x), \text{ for all } x \in \mathbb{R}, \tag{3.14}$$

where $\lambda_i > 0$ denotes the degree of risk aversion of firm i. Now let $\lambda = \left(\sum_{i \in N} \frac{1}{\lambda_i}\right)^{-1}$. Solving (3.11) and (3.12) yields that the Pareto optimal posterior risk profile of firm $i \in N$ corresponding to $k \in \mathbb{R}^N_{++}$ is given by

$$X_i^{\text{post}}(\omega) = \frac{\lambda}{\lambda_i}\sum_{j\in N} X_j(\omega) + \frac{\log(\mathbb{P}_i(\{\omega\}))}{\lambda_i} - \frac{\lambda}{\lambda_i}\sum_{j\in N}\frac{\log(\mathbb{P}_j(\{\omega\}))}{\lambda_j} + \frac{\log(k_i)}{\lambda_i} - \frac{\lambda}{\lambda_i}\sum_{j\in N}\frac{\log(k_j)}{\lambda_j}, \quad (3.15)$$

for all $\omega \in \Omega$. When firms have homogeneous beliefs regarding the probability distribution, i.e., when $\mathbb{P}_1 = \mathbb{P}_2 = \cdots = \mathbb{P}_n$, the second and the third term vanish, and the redistribution of risk is proportional, i.e., firm $i \in N$ is allocated a fraction $\delta_i = \frac{\lambda}{\lambda_i}$ of the aggregate risk. The last two terms reflect deterministic side payments d_i satisfying $\sum_{i\in N} d_i = 0$. When firms have heterogeneous beliefs regarding the probability distribution, the redistribution is no longer proportional. In addition to the fraction δ_i of the aggregate risk, firm $i \in N$ is now also assigned the risk Y_i given by $Y_i(\omega) = \frac{\log(\mathbb{P}_i(\{\omega\}))}{\lambda_i} - \frac{\lambda}{\lambda_i}\sum_{j\in N}\frac{\log(\mathbb{P}_j(\{\omega\}))}{\lambda_j}$, for all $\omega \in \Omega$.[19]

Next, we consider the case where each firm $i \in N$ has a risk profile that takes on positive values only, i.e., $D_i = \mathbb{R}_{++}$, and uses the constant relative risk aversion (CRRA or power) utility function given by

$$u_i(x) = \frac{x^{1-\gamma}}{1-\gamma}, \text{ for all } x \in \mathbb{R}_{++}, \quad (3.16)$$

where $\gamma \in \mathbb{R}_{++}\backslash\{1\}$ denotes the parameter of risk aversion. Solving (3.11) and (3.12) yields that the Pareto optimal posterior risk profile of firm $i \in N$ corresponding to $k \in \mathbb{R}^N_{++}$ is given by

$$X_i^{\text{post}}(\omega) = \frac{(k_i\mathbb{P}_i(\{\omega\}))^{\frac{1}{\gamma}}}{\sum_{j\in N}(k_j\mathbb{P}_j(\{\omega\}))^{\frac{1}{\gamma}}}\left(\sum_{j\in N} X_j(\omega)\right), \text{ for all } \omega \in \Omega. \quad (3.17)$$

In this case, the Pareto optimal risk redistribution does not involve side payments. The redistribution of risk, however, is proportional only if $\mathbb{P}_1 = \mathbb{P}_2 = \cdots = \mathbb{P}_n$.

Finally, consider the case where the risk profile of firm $i \in N$ takes values in $D_i = \left(-\infty, \frac{1}{\alpha_i}\right)$ for $\alpha_i > 0$, and firm i uses a quadratic utility function given by

$$u_i(x) = x - \frac{\alpha_i}{2}x^2, \text{ for all } x \in \left(-\infty, \frac{1}{\alpha_i}\right). \quad (3.18)$$

Let $\alpha = \left(\sum_{i\in N}\frac{1}{\alpha_i}\right)^{-1}$, $\lambda_i(\omega) = k_i\mathbb{P}_i(\{\omega\})\alpha_i$ and $\lambda(\omega) = \left(\sum_{i\in N}\frac{1}{\lambda_i(\omega)}\right)^{-1}$ for all $i \in N$ and $\omega \in \Omega$. Solving (3.11) and (3.12) yields that the Pareto optimal posterior risk profile of firm

[19] Because with an exponential utility function, the certainty equivalent of a risky payoff is the negative of a cash invariant risk measure applied to the payoff, this result also follows from Acciaio and Svindland (2009).

$i \in N$ corresponding to $k \in \mathbb{R}^N_{++}$ is given by

$$X_i^{\text{post}}(\omega) = \frac{\lambda(\omega)}{\lambda_i(\omega)} \left(\sum_{j \in N} X_j(\omega) - \frac{1}{\alpha} \right) + \frac{1}{\alpha_i}, \text{ for all } \omega \in \Omega. \tag{3.19}$$

As was the case when firms use exponential utility functions, the optimal redistribution involves deterministic side-payments, and the redistribution of risk is proportional only if $\mathbb{P}_1 = \mathbb{P}_2 = \cdots = \mathbb{P}_n$. $\triangledown$

The above example illustrates the effects of heterogeneity regarding the subjective probability distributions on the Pareto optimal redistributions. Heterogeneity implies that the redistribution is unlikely to be proportional. This has some interesting implications. For example, whereas with homogeneous beliefs (i.e., when $\mathbb{P}_1 = \cdots = \mathbb{P}_n$), all Pareto optimal posterior risk profiles are risk-free for each firm if and only if $\sum_{i \in N} X_i$ is risk-free, this is no longer the case when firms have heterogeneous beliefs. Even when pooling all risk profiles allows to eliminate all risk (i.e., $\sum_{i \in N} X_i$ is risk-free), different beliefs regarding the likelihood of the different states of the world imply that the Pareto optimal posterior risk profiles are not necessarily riskless for all firms.

Moreover, heterogeneous beliefs may also increase the likelihood that firms (believe that they) can benefit from redistributing their risks. Whereas clearly $\sum_{i \in N} \mathbb{E}_{\mathbb{P}}[X_i^{\text{post}}] = \sum_{i \in N} \mathbb{E}_{\mathbb{P}}[X_i]$ for any vector of feasible posterior risk profiles when firms have homogeneous beliefs, it is possible that there exist feasible posterior risk profiles satisfying $\sum_{i \in N} \mathbb{E}_{\mathbb{P}_i}[X_i^{\text{post}}] > \sum_{i \in N} \mathbb{E}_{\mathbb{P}_i}[X_i]$ in case of heterogeneous beliefs. Thus, heterogeneous beliefs may imply that all firms believe that they can simultaneously gain in expectation. This suggests that heterogeneous beliefs might make redistribution of risk even more attractive. In the next corollary of Theorem 3.3.1, we show that it is unlikely that there is no room to benefit for the firms.

Corollary 3.3.3 *There does not exist an $(X_i^{post})_{i \in N} \in \mathcal{PO}(N)$ with $\Delta U_i(X_i^{post}) \geq 0$ for all $i \in N$ and $\Delta U_i(X_i^{post}) > 0$ for at least one $i \in N$ if and only if for all $j \in N$, $\frac{u_1'(X_1(\omega))\mathbb{P}_1(\{\omega\})}{u_j'(X_j(\omega))\mathbb{P}_j(\{\omega\})}$ does not depend on $\omega \in \Omega$.*

The above corollary yields a necessary and sufficient condition for the existence of a Pareto optimal risk redistribution that weakly benefits all firms and strictly benefits at least one firm. The proposition shows that heterogeneous beliefs regarding the underlying probability distribution make it more likely that such redistributions exist. For example, whereas with homogeneous beliefs, improvement cannot be obtained when the firms have the same prior risk profile and the same risk preferences (because this implies that $\frac{u_1'(X_1(\omega))}{u_j'(X_j(\omega))} = 1$ for all $\omega \in \Omega$ and $j \in N$), improvements can be achieved when the firms have heterogeneous beliefs. Hence, only in very special cases, there is no room for improvement.

3.3.2 Stable risk redistributions

For all firms to be willing to engage in a particular redistribution, the corresponding posterior risk profiles need to satisfy *Stability.* The set $\mathcal{S}(N)$ is the set of all feasible posterior risk profiles that satisfy *Stability*, i.e.,

$$\mathcal{S}(N) = \left\{(X_i^{\text{post}})_{i\in N} \in \mathcal{F}(N) : (X_i^{\text{post}})_{i\in S} \in \mathcal{NI}(S) \text{ for all } S \subseteq N\right\}. \tag{3.20}$$

We know from Proposition 3.2.4(i) that a necessary condition for *Stability* to be satisfied is that the redistribution satisfies *No Pareto Improvement* and *Individual Rationality.* Moreover, when risk is redistributed between only two firms, these conditions are also sufficient. Therefore, it holds that

$$\mathcal{S}(N) \subseteq \left\{(X_i^{\text{post}})_{i\in N} \in \mathcal{PO}(N) : (\Delta U_i(X_i^{\text{post}}))_{i\in S} \geq 0\right\}, \tag{3.21}$$

where the inclusion is an equality when risk is redistributed between two firms. When risk is redistributed between more than two firms, *No Pareto Improvement* and *Individual Rationality* in general is not sufficient to guarantee *Stability.*

There exist infinitely many Pareto optimal posterior risk profiles, i.e., the set $\mathcal{PO}(N)$ contains infinitely many posterior risk profiles. The following theorem shows that the subset of posterior risk profiles that in addition satisfy *Stability* is non-empty.

Theorem 3.3.4 *The set $\mathcal{S}(N)$ is non-empty, i.e., there exist feasible posterior risk profiles that satisfy Stability.*

3.3.3 The bargaining problem

The previous subsection shows that there exist risk redistributions that benefit all firms in the sense that each firm weakly gains from the redistribution in expected utility terms, and no subset of firms can be better off when they exclude the other firms from the negotiation and redistribute their risk amongst each other. In general, however, the set of redistributions that satisfy these criteria is not single-valued, and the issue arises which redistribution is selected. In each redistribution, all firms weakly benefit, but the extent to which a particular firm benefits depends on the particular redistribution that is chosen. In general, a redistribution that yields a high expected utility gain for a particular firm does not yield a high expected utility gain of another involved firm. This implies that the firms bargain over the redistribution that they choose. The selection of a particular redistribution reflects a bargaining process that can be modeled via a *bargaining rule* (Nash, 1950).

For bargaining problems, there exists a variety of bargaining rules. There is one bargaining rule that received considerable attention in the literature. This rule is the Nash bargaining solution, and is characterized by, e.g., Nash (1950) and Kalai (1977).[20] For the risk redistribution problem that we consider, the Nash bargaining solution is given by

$$\mathcal{NB} = \underset{(X_i^{\text{post}})_{i \in N} \in \mathcal{F}(N) \cap \mathcal{IR}(N)}{\text{argmax}} \prod_{i \in N} \Delta U_i(X_i^{\text{post}}). \tag{3.22}$$

The objective function in (3.22) weighs the benefits of the involved parties. By construction, it yields redistributions that satisfy *Individual Rationality*, i.e., $\mathcal{NB} \subset \mathcal{IR}(N)$. Moreover, while the constraint set allows any redistribution that satisfies *Feasibility* and *Individual Rationality*, it is readily verified that the Nash bargaining solution in (3.22) yields redistributions that satisfy also *No Pareto Improvement*. So, it holds that $\mathcal{NB} \subset \mathcal{PO}(N)$. *Stability,* however, in general is a stronger condition than *No Pareto Improvement* and *Individual Rationality*. Therefore, in order to reflect the fact that firms only agree to a redistribution if it satisfies *Stability*, we consider instead the constrained Nash bargaining solution, which is given by:

$$\mathcal{CNB} = \underset{(X_i^{\text{post}})_{i \in N} \in \mathcal{S}(N)}{\text{argmax}} \prod_{i \in N} \Delta U_i(X_i^{\text{post}}). \tag{3.23}$$

This rule is called the *coalitional Nash bargaining solution* (Compte and Jehiel, 2010) in case of Transferable Utility games. The following proposition shows that there exists a solution to the constrained Nash bargaining problem in (3.23).

Proposition 3.3.5 *It holds that* $\mathcal{CNB} \neq \emptyset$.

We conclude this section by showing that when risk is distributed between two firms, there is a unique solution to the constrained Nash bargaining problem. Recall from Proposition 3.2.4 that with two firms, *Stability* is equivalent to *No Pareto Improvement* and *Individual Rationality*. Hence, the set of potential risk redistributions that the firms bargain over is the set of redistributions that satisfy *No Pareto Improvement* and *Individual Rationality*, which implies that $\mathcal{CNB} = \mathcal{NB}$, i.e., the constrained Nash bargaining solution coincides with the Nash bargaining solution. Using Nash (1950), we show in the following proposition that the Nash bargaining solution is single-valued when $|N| = 2$.

Proposition 3.3.6 *If* $|N| = 2$*, then* $\mathcal{CNB} = \mathcal{NB}$*. Moreover,* $\mathcal{CNB}$ *is single-valued.*

In contrast to the case where $|N| = 2$, the constrained Nash bargaining solution need not be single-valued when $|N| > 2$.

[20] Nash (1950) and Kalai (1977) provide a cooperative game-theoretic characterization based on four properties. Rubinstein (1982) provides a bilateral non-cooperative game in which if players are perfectly patient, the equilibrium division converges to the Nash bargaining solution. Moreover, Van Damme (1986) shows that the Nash bargaining solution constitutes the unique equilibrium if two firms have different opinions about what is the appropriate solution concept to use.

3.4 Redistributing longevity risk

In this section we use the model developed in the previous section to investigate the extent to which pension funds and life insurance companies can benefit from redistributing their risks Over-The-Counter. Pension funds and life insurance companies face longevity risk, which is the risk due to uncertain changes over time in survival rates of the insured population. There is substantial uncertainty regarding the future development of survival rates (see, e.g., Pitacco, Denuit, Haberman and Olivieri, 2009), and this uncertainty imposes significant risk on pension funds and life insurance companies. Coughlan et al. (2007), for example, show that on average every additional year of life expectancy adds approximately 3% to 4% to the value of UK pension liabilities. In the Netherlands, unexpectedly high deviations between best estimate mortality trends estimated in 2010 and those estimated in 2007 led to increases in the net premium for old-age pension annuities of up to 12%. This illustrates the potentially huge impact of unanticipated increases in survival rates. A potential way to mitigate these adverse effects of longevity risk on the liabilities of pension funds is to exploit the natural hedge potential that arises from combining life annuities and death benefit insurance.

In Subsection 3.4.1, we model the risk profiles of pension funds offering whole life annuities, and life insurance companies offering death benefit insurance. In Subsection 3.4.2, we numerically illustrate the extent to which a pension fund and a death benefit insurer can benefit from redistributing their risks. It is often argued that major obstacles to trade are potential disagreement regarding the true distribution of future mortality rates, reluctance to engage in contracts with long horizons, and insufficient capacity in the insurance market to yield significant risk reduction for pension funds. In order to investigate these issues, we focus our numerical analysis on the effects of the time horizon, the relative size of the pension fund and the death benefit insurer, and potential heterogeneity in beliefs regarding the true probability distribution of future survival rates.

3.4.1 The risk profiles

We consider the case where the risk profile of pension funds and insurers is the net value of their assets and liabilities, referred to as the *net asset value,* at a prespecified future date T. Thus, the prior risk profile of firm $i \in N$ is given by

$$X_i(T) = A_i(T) - L_i(T), \tag{3.24}$$

where $A_i(T)$ denotes the (market) value of the assets at time T and $L_i(T)$ denotes the date-T value of the liabilities for firm i.

In order to focus on longevity risk, we assume a deterministic risk-free rate (see, e.g., Olivieri, 2001; Brouhns et al., 2002; Cossette et al., 2007), which we denote r.[21] The asset value for firm $i \in N$ on date T then follows from

$$A_i(t) = (1+r)A_i(t-1) - \widetilde{L}_{i,t}, \qquad \text{for } t = 1, \ldots, T, \tag{3.25}$$

where $\widetilde{L}_{i,t}$ denotes the (stochastic) liability payment of firm i at date t. Combining (3.24) and (3.25) yields

$$X_i(T) = [A_i(0) - CL_i(T)]\,(1+r)^T, \tag{3.26}$$

where

$$CL_i(T) = \sum_{\tau=1}^{T} \frac{\widetilde{L}_{i,\tau}}{(1+r)^\tau} + \frac{L_i(T)}{(1+r)^T}, \text{ for } i \in N. \tag{3.27}$$

Thus, firm i's risk profile equals the initial asset value $A_i(0)$ increased by the return on assets, and reduced by the random variable $CL_i(T)(1+r)^T$, where $CL_i(T)$ represents the sum of the present value of liability payments up to year T, and the present value of the date-T value of all payments beyond date T.

It now remains to specify how the date-T liability value $L_i(T)$ is determined. Ideally, $L_i(T)$ would represent the market value on date T of the future liabilities, i.e., the value at which the liabilities can be sold to a third party. Because there is (not yet) a liquid market for longevity-linked products, however, there is not yet a market price, and pension funds and insurance companies have to valuate their liabilities using mark-to-model valuation instead. We consider the case where the liabilities are valued at "best estimate value"in our numerical analysis in Subsection 3.4.2.

On date zero, the prior risk profile $X_i(T)$ is uncertain due to uncertainty in the liability payments $\widetilde{L}_{i,\tau}$, for $\tau = 1, \ldots, T$, as well as due to uncertainty in the value of the remaining liabilities on date T, $L_i(T)$. In Subsection 3.4.2, we discuss the characteristics of the liabilities.

Via redistribution of risk, pension funds and insurers want to arrive at feasible posterior risk profiles $(X_i^{\text{post}}(T))_{i\in N}$ such that they all benefit weakly in expected utility terms. For any given posterior risk profiles, however, there exist infinitely many ways in which the firms can achieve the corresponding risk redistribution. One way that we consider in particular is redistribution of risk via a swap contract in which a single payment occurs on maturity date T.[22] Specifically, for any given posterior risk profiles $(X_i^{\text{post}}(T))_{i\in N}$, we let $(\chi_i(T))_{i\in N}$ be given by

$$\chi_i(T) = X_i(T) - X_i^{\text{post}}(T), \text{ for } i \in N, \tag{3.28}$$

[21] Since r is deterministic, we focus only on the uncertainty of longevity in the liability payments of a firm. In this way, the redistribution takes the form of a contract contingent on the liabilities and a side-payment.

[22] It is also possible to transfer longevity-linked products at time T or via swaps with maturity T and periodic pay-offs.

i.e., $\chi_i(T)$ reflects the net payment from firm i to the other firms on date T. Then, the posterior risk profiles can be written as

$$X_i^{\text{post}}(T) = \left[A_i(0) - CL_i^{\text{post}}(T)\right](1+r)^T, \text{ for } i \in N, \tag{3.29}$$

where

$$CL_i^{\text{post}}(T) = \sum_{\tau=1}^{T} \frac{\widetilde{L}_{i,\tau}}{(1+r)^\tau} + \frac{\chi_i(T) + L_i(T)}{(1+r)^T}, \text{ for } i \in N. \tag{3.30}$$

Thus, $\chi_i(T)$ can be seen as the payoff of a swap contract with fixed maturity $t = T$ and a single payment on date $t = T$. Clearly, it follows from (3.1) and (3.28) that

$$\sum_{i \in N} \chi_i(T) = 0. \tag{3.31}$$

A special case of our model that is considered often in the literature is the case where $T = T^{\max}$, with $T^{\max}$ large enough so that the probability that all participants are deceased is 1. Then, $L_i(T^{\max}) = 0$, and so $CL_i(T^{\max})$ is equal to

$$CL_i(T^{\max}) = \sum_{\tau=1}^{T^{\max}} \frac{\widetilde{L}_{i,\tau}}{(1+r)^\tau}, \text{ for } i \in N, \tag{3.32}$$

i.e., it equals the date-zero present value of all future liabilities over complete run-off. This is the case which is typically considered in the literature (see, e.g., Tsai et al., 2010; Wang et al, 2010). A potential drawback of this approach is that firms agree on a redistribution of risk over complete run-off. Considering instead a shorter horizon, as we do, allows for the possibility to renegotiate. For example, firms can re-evaluate their liabilities according to new mortality data, new regulations, attrition, and new participants. So, at the future date, a new contract can be negotiated.

3.4.2 Benefits from risk redistributions

In this subsection we numerically illustrate the potential benefits from redistributions of longevity risk between a pension fund and a death benefit insurer. We first specify the risk profiles, risk preferences, and subjective probability distributions of the pension fund and the death benefit insurer. Then, we numerically illustrate the benefits from a redistribution of risk that reflects the outcome of the bargaining process according to the Nash bargaining solution. We focus on the effects of the time horizon T, the relative size of the pension fund and the death benefit insurer, and potential heterogeneity in beliefs regarding the true probability distribution of future survival rates.

Prior risk profiles, subjective probabilities, and risk preferences

In this subsection, we discuss the extent to which insurers and pension funds are able to benefit from redistributing their risks. To do so, we have that the set of firms participating in the risk redistribution problem is given by $N = \{PF, DB\}$, where PF is the pension fund and DB the death benefit insurer. The potential benefits from risk redistribution depend on the characteristics of their liabilities (i.e., the liability payments $\widetilde{L}_{i,\tau}$), their risk preferences, and their (subjective) beliefs regarding the probability distribution on the underlying state space.

We start by discussing the characteristics of the liabilities. For the pension fund, $\widetilde{L}_{PF,\tau}$ is a random variable that equals the aggregate annuity payment in year τ to all participants who are alive and older than 65 in that year. For the insurer, $\widetilde{L}_{DB,\tau}$ is a random variable that equals the aggregate death benefit payment in year τ to all participants who died in that year and were younger than 65. The level of these payments is affected by two types of mortality risk. First, the payments are subject to *longevity risk,* which is the risk that arises due to the fact that the survival rates in a given population change over time in an unpredictable way. In addition, the liabilities are subject to *individual mortality risk* which arises due to the fact that, conditional on given survival rates in the population, whether a particular individual survives an additional year is uncertain. Individual mortality risk, however, becomes negligible when portfolio size is large (see, for example, Olivieri, 2001). Given the large portfolio sizes that we consider, we assume that individual mortality risk is negligible, and focus on the impact of longevity risk. Formally, for sufficiently large portfolios, the aggregate portfolio payment $\widetilde{L}_{i,\tau}$ of firm i in year τ can be approximated by

$$\widetilde{L}_{i,\tau} = \begin{cases} \sum_{j \in M_i} \delta_{i,j} \cdot {}_{\tau}p_{x_{i,j},0} \cdot \mathbb{1}_{x_{i,j}+\tau \geq 65}, & \text{if } i = PF, \\ \sum_{j \in M_i} \delta_{i,j} \cdot \left({}_{\tau-1}p_{x_{i,j},0} - {}_{\tau}p_{x_{i,j},0} \right) \cdot \mathbb{1}_{x_{i,j}+\tau < 65}, & \text{if } i = DB, \end{cases} \tag{3.33}$$

where

- M_i denotes the set of insureds of firm $i \in N$;
- $\delta_{i,j} \in \mathbb{R}_+$ and $x_{i,j} \in \mathbb{N}$ denote the insured right and the age, respectively, of insured $j \in M_i$ of firm $i \in N$;
- ${}_{\tau}p_{x,0}$ denotes the *future* probability that an individual belonging to the cohort aged x in year $t = 0$ will survive at least τ more years. Following Cairns et al. (2006) we define $p_{x+s,s} = \mathbb{P}(T_x \geq s+1 | T_x \geq s, \mathcal{F}_\infty)$, where $\mathcal{F}_\infty$ denotes the set that contains of all information regarding mortality rates at all future dates, and where T_x denotes the random remaining lifetime of an individual aged x in year $t = 0$. Then,

$$\begin{aligned} {}_{\tau}p_{x,0} &= \mathbb{P}(T_x \geq \tau | \mathcal{F}_\infty) \\ &= \mathbb{P}(T_x \geq \tau | T_x \geq \tau - 1, \mathcal{F}_\infty) \cdot \mathbb{P}(T_x \geq \tau - 1 | \mathcal{F}_\infty) \end{aligned}$$

$$= \prod_{s=1}^{\tau} \mathbb{P}(T_x \geq s | T_x \geq s-1, \mathcal{F}_\infty)$$
$$= \prod_{s=0}^{\tau-1} p_{x+s,s}.$$

Note that $p_{x+s,s}$ is known at time $s+1$.

Details regarding the number of insureds and their insured rights are provided in Appendix 3.B.

Longevity risk arises from the fact that the future one-year probabilities $p_{x+s,s}$ for $s \geq 0$ are unknown on date zero. Longevity risk affects the date-T net asset value of the firms through its effect on the probability distribution of $(CL_i(T))_{i \in N}$ from (3.27). It affects both the liability payments up to time T and the value of the liability payments after time T, $L_i(T)$. We consider the case where the pension fund and the insurer valuate their liabilities at the *best estimate value* with respect to their own (subjective) beliefs regarding the probability distribution of future death rates.[23,24] We discuss the determination of this best estimate value in Appendix 3.C.

We distinguish the case where the pension fund and the insurer have homogeneous beliefs regarding the underlying probability distribution and the case where they have heterogeneous beliefs. Specifically, we assume that both the pension fund and the death benefit insurer use the standard Lee and Carter (1992) model to estimate the probability distribution of future mortality rates (see Appendix 3.D), but they may disagree on the appropriate historical time period that is used to estimate the model parameters. For the case of homogeneous beliefs, they each estimate the model parameters based on data for Dutch males as reported in the Human Mortality Database for the period 1977 to 2009. For the case of heterogeneous beliefs, however, the death benefit insurer uses data from the shorter time period from 1987 until 2009. We refer to these models as LC(1977-2009) and LC(1987-2009), respectively. We display the corresponding parameter estimates in Appendix 3.D.

Figure 3.1 displays the 95% confidence interval of $CL_i(T)$ as percentage deviation from its expected value, i.e., $\left(CL_i(T) - E_{\mathbb{P}_i}[CL_i(T)]\right)/E_{\mathbb{P}_i}[CL_i(T)] \cdot 100\%$, as a function of the horizon T. The left panel corresponds to the case where the probability distribution of future mortality rates is estimated based on LC(1977-2009). The right panel corresponds to the case where the probability distribution of future mortality rates is estimated based on LC(1987-2009).

[23] Alternatively, the value of the liabilities can include a *market value margin*. The market value margin can, for example, reflect the cost of holding a capital reserve. For a discussion of how the *market value margin* is formulated under Solvency II, see for example Stevens, De Waegenaere and Melenberg (2011).

[24] In some cases, regulators may prescribe the use of a specific probability distribution to determine the best estimate value of the liabilities for regulatory purposes. Then, depending on whether the firm's objective is to reduce the volatility of the Net Asset Value with liability value as prescribed by the regulator, or the Net Asset Value with liability value determined according to its own probability distribution, the firm would use either the exogenously given probability distribution $\mathbb{P}$, or its own subjective probability distribution $(\mathbb{P}_i)$ to valuate the liabilities.

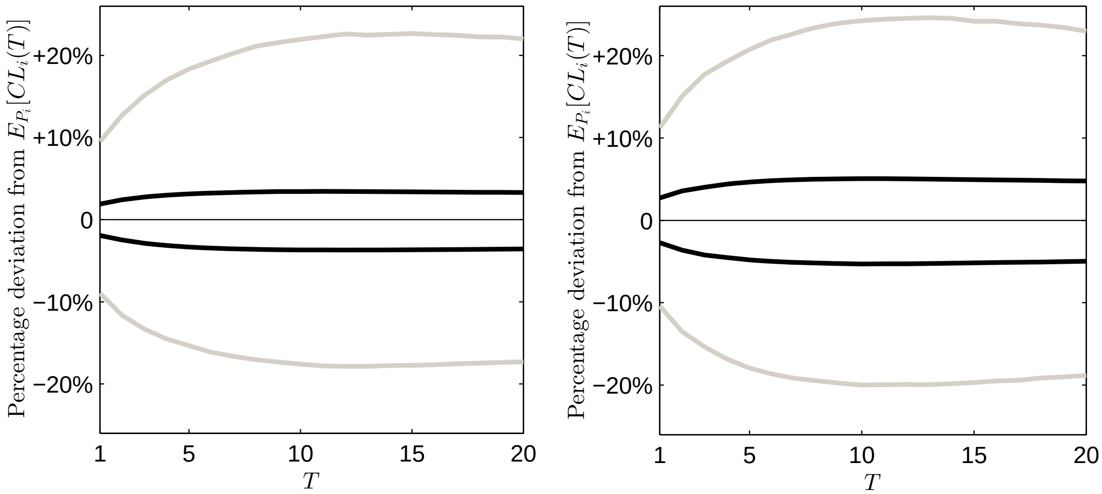


Figure 3.1: The 95% confidence interval of $(CL_i(T) - E_{\mathbb{P}_i}[CL_i(T)])\,/E_{\mathbb{P}_i}[CL_i(T)] \cdot 100\%$ for the pension fund (black) and for the death benefit insurer (grey), as a function of the maturity T. The left panel corresponds to the case where the probability distribution of future mortality rates is estimated based on LC(1977-2009). The right panel corresponds to the case where the probability distribution of future mortality rates is estimated based on LC(1987-2009).

As can be seen from Figure 3.1, the size of the 95% confidence interval is increasing for $T \leq 10$ and, for larger T, it remains approximately constant. The size of the confidence interval for $T = 10$ is not much larger than the size of the confidence interval for $T = 1$ for both the pension fund and the death benefit insurer. A further increase in the length of the horizon has a relatively small effect on the uncertainty. Comparison of the left panel and the right panel shows that the use of a shorter historical period to estimate the parameters of the Lee and Carter model leads to more risky risk profiles for the pension fund and the death benefit insurer, for every time horizon. The same pattern is observed if one considers other risk measures such as, for example, the relative standard deviation of $CL_i(T)$. Detailed summary statistics of $CL_i(T)$ are displayed in Table 3.3 in Appendix 3.E.

An important characteristic of the joint distribution of the net asset values of the pension fund and the death benefit insurer is that, for all horizons T, they exhibit relatively strong negative correlation. Figure 3.2 illustrates the strong negative correlation between $CL_{PF}(T)$ and $CL_{DB}(T)$ for the case where $T = 1$, and the probability distributions of mortality are estimated based on LC(1977-2009) and LC(1987-2009).

Table 3.4 in Appendix 3.C shows that for horizons $T \in \{1, 5, 10, 15, 20, T^{\max}\}$, the correlation between $CL_{PF}(T)$ and $CL_{DB}(T)$ ranges from -0.88 to -0.97 for LC(1977-2009), and from -0.92 to -0.98 for LC(1987-2009). This strong negative correlation suggests that the pension fund and

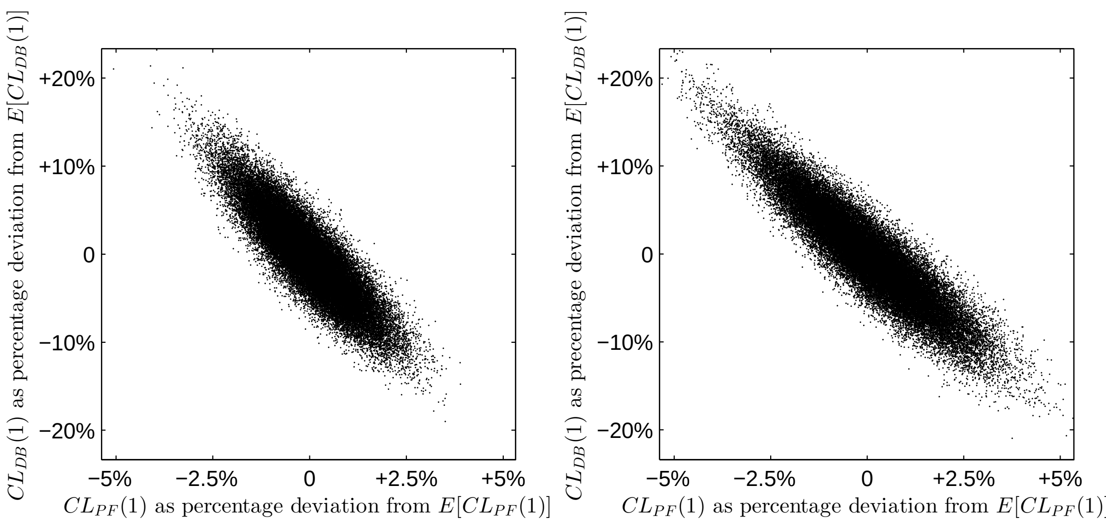


Figure 3.2: The correlation between $CL_{DB}(1)$ for the pension fund and $CL_{DB}(1)$ for the death benefit insurer ($i = DB$), expressed as percentage deviation from their expected values. The left panel corresponds to the case where the probability distribution of future mortality rates is estimated based on LC(1977-2009). The right panel corresponds to the case where the probability distribution of future mortality rates is estimated based on LC(1987-2009).

the death benefit insurer could potentially benefit significantly from natural hedge potential that arises from redistributing the risks.

The extent to which the pension fund and the death benefit insurer can benefit from redistributing their risks also depends on their risk preferences. The pension fund and the death benefit insurer each use an exponential utility function (see (3.14)) with risk aversion parameter

$$\lambda_i = \begin{cases} 10^{-3} & \text{if } i = PF, \\ \frac{5 \cdot 10^{-4}}{\gamma} & \text{if } i = DB. \end{cases} \tag{3.34}$$

These levels of risk aversion imply that the maximum premium that the pension fund would be willing to pay for a full buy-out of its liabilities equals 101.6% of the best estimate value of the liabilities. For the death benefit insurer, the corresponding maximum premium equals 117.3% of the best estimate value.[25]

[25] If the Net Asset Value is divided equally over all participants (insureds), if all participants (insureds) have the same absolute risk aversion, and if the expected utility of the pension fund (insurer) equals the expected utility of its participants (insureds), this implies an individual risk aversion parameter of approximately 50. Note that the monetary unit in our model equals the annual pension right. In case the annual pension right equals 25,000 dollar, this parameter corresponds with a risk aversion parameter of approximately 0.002 when the monetary unit is one dollar.

Benefits from redistribution

The negative correlation between the risk profiles of the pension fund and the death benefit insurer suggests strong potential for hedge benefit for all parties. In this subsection, we numerically illustrate the extent to which the pension fund and the death benefit insurer can benefit from this hedge potential by redistributing their risks amongst each other. Given their prior risk profiles from (3.26), (3.33) and the date-T value of the liabilities, we determine the set of Pareto optimal posterior risk profiles from (3.11) and (3.12). We then use (3.22) to select the particular Pareto optimal posterior risk profiles that reflect the outcome of a bargaining process that equally weighs the benefits of both parties, i.e., the Nash bargaining solution. We note that because the pension fund and the insurer use an exponential utility function, the set of Pareto optimal posterior risk profiles is independent of the initial asset values $A_i(0)$. Hence, posterior distribution of the liabilities in (3.30) corresponding to the Nash bargaining solution, as given in (3.22), is also independent of $A_i(0)$. In Appendix 3.F, we describe the method used to approximate the welfare gains for both the case of homogeneous and heterogeneous beliefs via a partition of the state space.

We first investigate the effect of heterogeneous beliefs on the benefits from redistribution for the case where $T = 1$. In Figure 3.3, we display the probability distributions of the prior date-1 value of the current liabilities $CL_i(1)$ (grey histograms) and the probability distributions of the posterior date-1 value of the current liabilities $CL_i^{\text{post}}(1)$ (black histograms) for the pension fund (left figure) and for the death benefit insurer (right figure), expressed as percentage deviation from their best estimate values. The upper (lower) panels correspond to the case of homogeneous (heterogeneous) beliefs.

The upper panel of Figure 3.3 shows that in case of homogeneous beliefs, the risk redistribution implies that the Net Asset Value of both the pension fund and the insurer becomes significantly less dispersed. Comparing the upper and the lower panel shows that, after redistribution, the probability that the death benefit insurer faces payments lower than the date-zero best estimate value is significantly higher in the case of heterogeneous beliefs. The reason is that the death benefit insurer assigns relatively lower probabilities to states of the world with less extreme outcomes for the aggregate risk, and relatively higher probabilities to states of the world with more extreme outcomes for the aggregate risk. The pension fund assigns almost no probability to these more extreme outcomes. Both parties can therefore benefit by payments in scenarios with more extreme outcomes of the aggregate risk are made by the pension fund to the death benefit insurer.

To quantify the benefits from the redistribution, we consider the following two criteria:

(i) The *percentage decrease in the date-zero expected present value of the liabilities.* We deter-

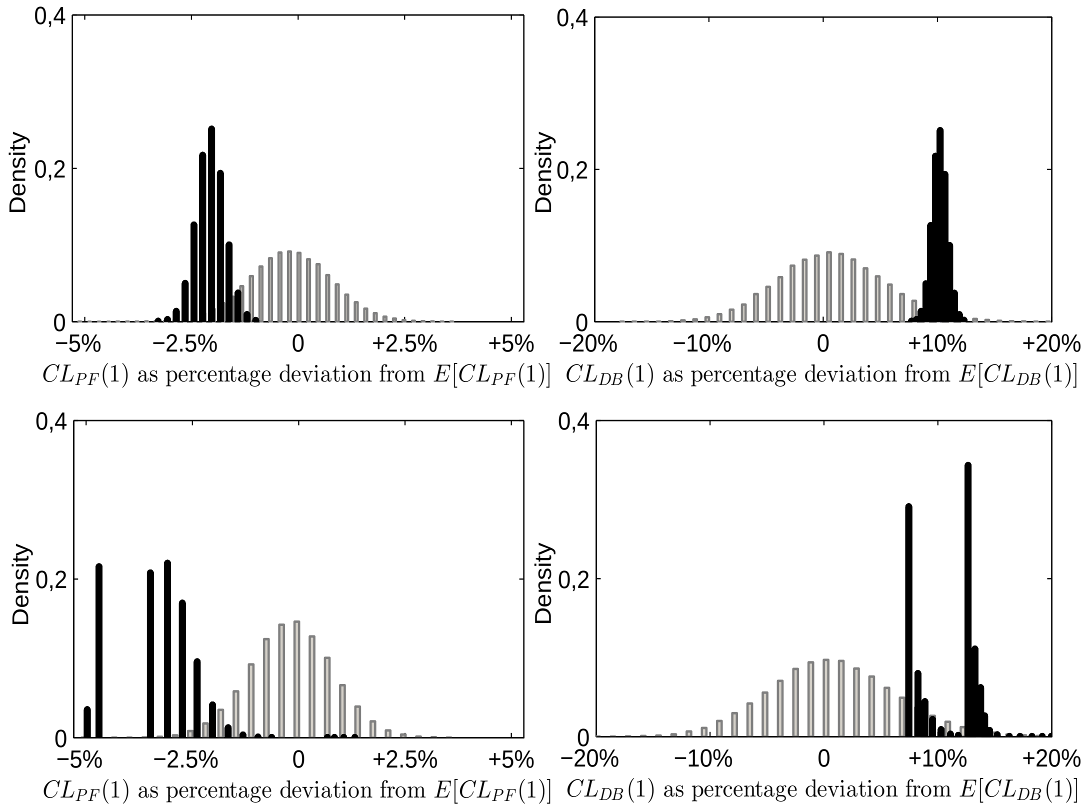


Figure 3.3: The grey (black) histograms represent the probability distributions of the prior (posterior) date-1 value of the current liabilities $CL_i(1)$ of the pension fund (left figure) and the death benefit insurer (right figure), as percentage deviation from the expected value of the date-1 value of the current liabilities $E_{\mathbb{P}_i}[CL_i(1)]$. The posterior distribution is given by (3.22). The upper (lower) panel corresponds to the case of homogeneous (heterogeneous) beliefs.

mine the percentage reduction in the expected value of the liabilities as

$$\%RedEV_i(T) = \frac{E_{\mathbb{P}_i}\left[CL_i(T)\right] - E_{\mathbb{P}_i}\left[CL_i^{\text{post}}(T)\right]}{E_{\mathbb{P}_i}\left[CL_i(T)\right]}, \text{ for } i \in N, \tag{3.35}$$

where $CL_i(T)$ is defined in (3.27) and $CL_i^{\text{post}}(T)$ is defined in (3.30).

(ii) The *relative zero-utility premium.* Recall that the redistribution implies that firm $i \in N$ effectively receives a net payment equal to $X_i^{\text{post}}(T) - X_i(T) = -\chi_i(T)$ on date T. The value to firm i of the risk redistribution can therefore be quantified by determining the maximum premium that firm i would have been willing to pay on date 0 for a contract that yields this net payment on date T. This maximum premium, which we denote $p_i(T) \in \mathbb{R}$, is the premium at which firm i would be indifferent between buying the contract and not buying the contract, and is therefore referred to as the *zero-utility premium.* It is the unique solution of the following equation:

$$\mathbb{E}_{\mathbb{P}_i}\left[u_i(X_i(T) - \chi_i(T) - (1+r)^T p_i(T))\right] = \mathbb{E}_{\mathbb{P}_i}\left[u_i(X_i(T))\right], \text{ for } i \in N. \tag{3.36}$$

We report the value of the zero-utility premium for firm i relative to the date-zero expected value of the firm i's prior liabilities, i.e.,

$$\%ZU_i(T) = \frac{p_i(T)}{E_{\mathbb{P}_i}\left[CL_i(T)\right]}, \text{ for } i \in N. \tag{3.37}$$

Figure 3.4 and Table 3.1 summarize the simulated gains from redistribution, as a function of the time horizon T. The upper (lower) figures correspond to the case where the firms have homogeneous (heterogeneous) beliefs regarding the underlying probability distribution. The left figures display the percentage reduction in the date-zero expected present value of the liabilities, as defined in (3.35). The right figures display the relative zero-utility premium corresponding to the risk distribution, as defined in (3.37). The design of the table is similar.

Panel A: homogeneous beliefs

T	$\%RedEV_i(T)$		$\%ZU_i(T)$	
	Pension fund	Insurer	Pension fund	Insurer
1	1.9%	−9.4%	3.2%	6.9%
5	3.6%	−17.9%	7.4%	15.1%
10	5.7%	−28.3%	9.9%	20.1%
15	5.7%	−28.1%	9.8%	19.8%
20	5.8%	−28.9%	8.8%	17.7%
$T^{\max}$	5.0%	−24.7%	8.3%	16.5%

Panel B: heterogeneous beliefs

T	$\%RedEV_i(T)$		$\%ZU_i(T)$	
	Pension fund	Insurer	Pension fund	Insurer
1	3.2%	−9.9%	3.9%	8.2%
5	6.5%	−26.9%	10.6%	21.5%
10	7.5%	−34.4%	11.0%	22.1%
15	8.1%	−38.4%	11.6%	23.2%
20	7.9%	−37.8%	10.6%	21.1%
$T^{\max}$	6.2%	−32.7%	8.3%	16.4%

Table 3.1: The simulated gains of the risk redistribution according to the Nash bargaining solution from (3.22), as a function of the horizon $T \in \{1, 5, 10, 15, 20, T^{\max}\}$. Here, $\%RedEV_i(T)$ is defined in (3.35) and $\%ZU_i(T)$ is defined in (3.37). The upper panel corresponds to the case where the firms have homogeneous beliefs, namely LC(1977-2009). The lower panel corresponds to the case where the pension fund has again beliefs LC(1977-2009) while the insurer has beliefs LC(1987-2009).

First consider the case with homogeneous beliefs (upper graphs in Figure 3.4 and Panel A in

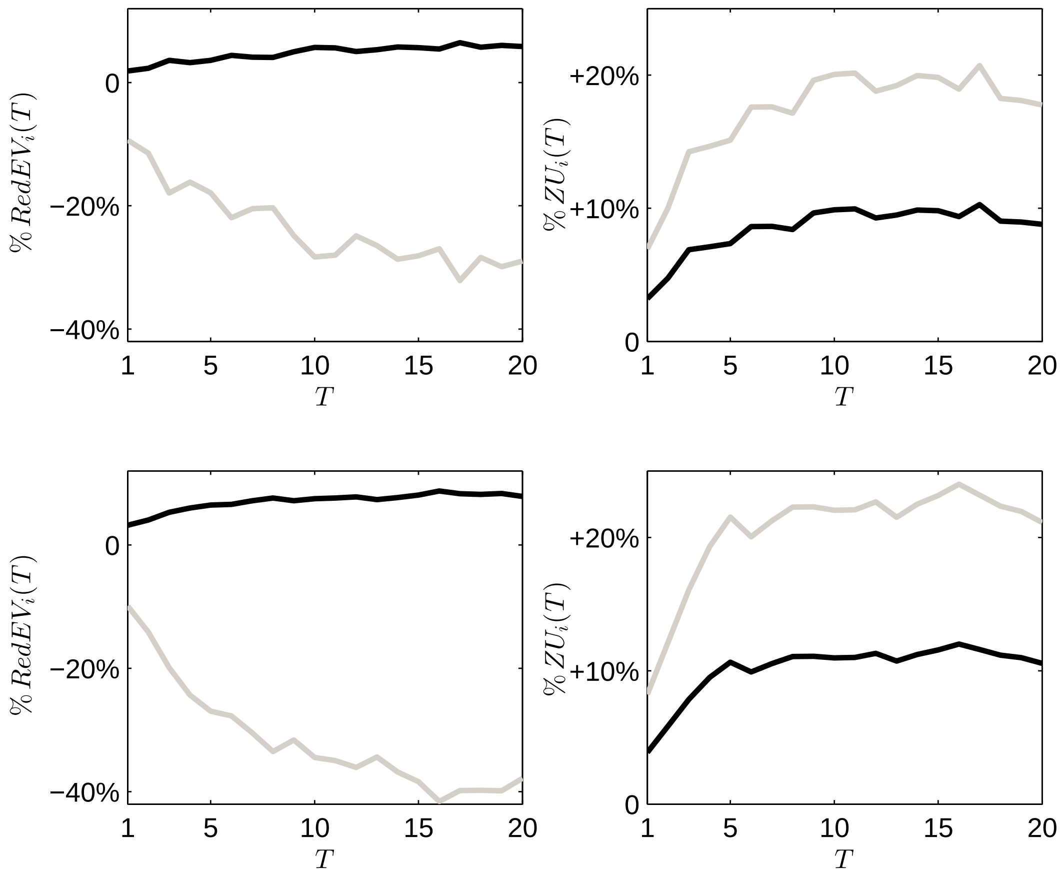


Figure 3.4: The simulated gains of the risk redistribution according to the Nash bargaining solution from (3.22), as a function of the horizon $T \in \{1, 2, \ldots, 20\}$. The black (gray) graphs represent the pension fund (death benefit insurer). Here, $\% RedEV_i(T)$ is defined in (3.35) and $\% ZU_i(T)$ is defined in (3.37). The upper graphs correspond to the case where the firms have homogeneous beliefs, namely LC(1977-2009). The lower graphs correspond to the case where the pension fund has again beliefs LC(1977-2009) while the insurer has beliefs LC(1987-2009).

Table 3.1). Remember that the redistribution occurs according to a swap contract in which, in each state of the world, either the pension fund makes a net payment to the death benefit insurer on date T, or the death benefit insurer makes a net payment to the pension fund on date T. As can be seen from Table 3.1, the risk redistribution is such that in expectation, the death benefit insurer makes a net payment to the pension fund, i.e., the expected value of $CL_{PF}(T)$ decreases, and, hence, the expected value of $CL_{DB}(T)$ increases. However, even though the death benefit insurer loses in expectation and the pension fund gains in expectation, both the pension fund and the insurer benefit from the redistribution in expected utility terms. The relative zero-utility premium of the death benefit insurer is relatively high as compared to the relative zero-utility premium of the pension fund. For example, the relative zero-utility premium according to the risk redistribution with horizon $T = 1$ equals 6.9% of the expected present value of the liabilities for the death benefit insurer, and 3.2% of the expected present value of the liabilities for the pension fund. This difference is due to the fact that the death benefit insurer has a significantly more risky prior risk profile, and hence benefits more from redistribution. As T increases up until $T = 10$, this effect becomes stronger. However, the benefits from choosing a horizon longer than ten years are negligible, i.e., the relative zero-utility premiums for both firms are fairly constant for $T > 10$. Given the potentially important drawbacks of choosing a long horizon, this suggests that parties can benefit most from redistributing their risks over a relatively short horizon.

Next, we compare the results of homogeneous beliefs with heterogeneous beliefs. Consider the case of heterogeneous beliefs (lower graphs in Figure 3.4 and Panel B in Table 3.1). The relative zero-utility premiums of the pension fund and the death benefit insurer are significantly higher than in the case with homogeneous beliefs. This is also what we expected because if a firm believes that a state has a relatively high probability to occur, this firm has a relatively high posterior pay-off in this state (see, e.g., (3.15)). Only if firms use the complete run-off of the liabilities to trade (i.e., the case where $T = T^{\max}$), the difference in relative zero-utility premium between the case of homogeneous beliefs and the case of heterogeneous beliefs is negligible.

Effect of relative size of insurer and pension fund

In this subsection we investigate the effect of the relative size of the pension fund and the death benefit insurer. We quantify the relative size via the ratio of the expected present value of their current liabilities, based on LC(1977-2009), i.e.,

$$\gamma = \frac{\mathbb{E}_{LC(1977-2009)}\left[\sum_{\tau>0} \frac{\widetilde{L}_{DB,\tau}}{(1+r)^{\tau}}\right]}{\mathbb{E}_{LC(1977-2009)}\left[\sum_{\tau>0} \frac{\widetilde{L}_{PF,\tau}}{(1+r)^{\tau}}\right]}. \tag{3.38}$$

We consider various values of $\gamma \in \left\{\frac{k}{100} : k = 1, \ldots, 100\right\}$ by adjusting the number of insureds of the death benefit insurer. All other characteristics of the portfolios are as given in Appendix 3.B.

In Figure 3.5, we display the effect of γ on the benefits from redistribution for the case where $T = 1$. We focus on the benefits for the pension fund, and distinguish the case where the firms have homogeneous beliefs (sold lines) and the case where they have heterogeneous beliefs (dashed lines). The top panel displays the percentage reduction in the expected present value of the liabilities for the pension fund. The bottom panel displays the relative zero-utility premium corresponding to the optimal redistribution.

First consider the case of homogeneous beliefs. Figure 3.5 shows that whereas the percentage increase in the expected Net Asset Value of the pension fund is monotonically increasing in γ,

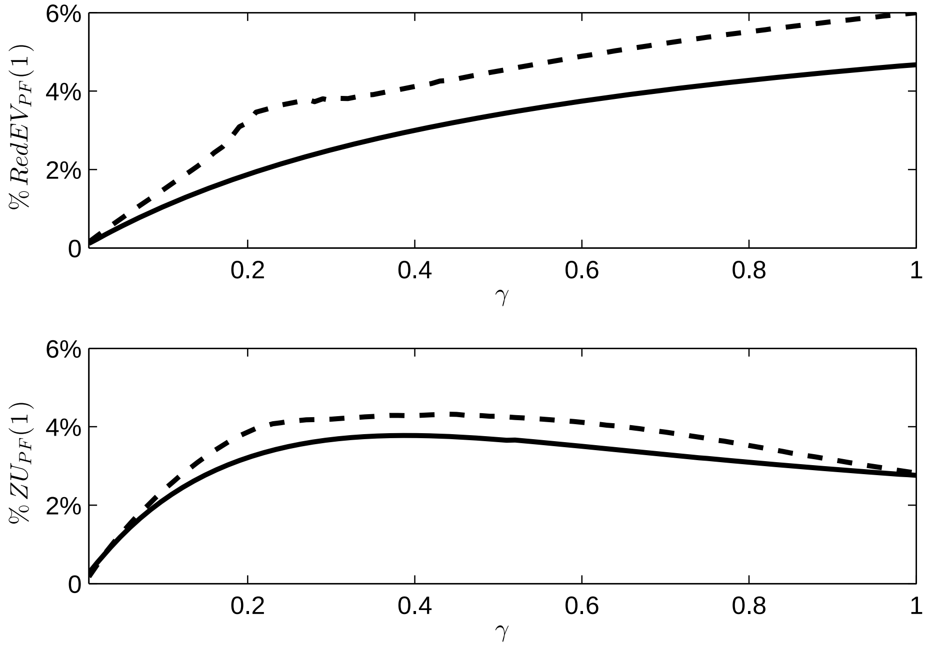


Figure 3.5: The effect of γ on the benefits from redistribution for the pension fund, for the case where $T = 1$. The solid lines correspond to the case of homogeneous beliefs, i.e., both the pension fund and the insurer use LC(1977-2009); the dashed lines correspond the case of heterogeneous beliefs, i.e., the pension fund uses LC(1977-2009), and the insurer uses LC(1987-2009). The upper panel displays the percentage increase in expected value (i.e., $\%RedEV_{PF}(1)$) and the lower panel displays the relative zero-utility premium (i.e., $\%ZU_{PF}(1)$).

the relative zero-utility premium is concave and reaches a maximum approximately at $\gamma = 0.4$. This can be understood as follows. By construction, the Nash bargaining solution that we consider corresponds to a Pareto optimal redistribution of risk. As can be seen from (3.15) with $\mathbb{P}_{PF} = \mathbb{P}_{DB}$, this implies that the aggregate risk is redistributed proportionally with deterministic side-payments. This in turn implies that all Pareto optimal posterior risk profiles for a given γ have the same variance because they only differ in the deterministic side-payments. Moreover, both for the pension fund and for the insurer, the variance of the posterior risk profile is a quadratic convex function of γ that reaches a global minimum.[26] Therefore, there is a value of

[26]Let the prior risk profiles corresponding to $\gamma = 1$ be given by $X_{PF}(T)$ and $X_{DB}(T)$. Then, the prior risk profiles corresponding to $\gamma \in \{\frac{k}{100} : k = 1, \ldots, 100\}$ are given by $X_{PF}(T)$ and $\gamma \cdot X_{DB}(T)$, respectively. Then, it follows from (3.15) that for the posterior risk profiles corresponding to γ, it holds that

$$\begin{aligned} var(X_i^{\text{post}}(T)) &= \frac{\lambda^2}{\lambda_i^2} var(X_{PF}(T) + \gamma \cdot X_{DB}(T)) \\ &= \frac{\lambda^2}{\lambda_i^2} var(X_{PF}(T)) + \frac{\lambda^2}{\lambda_i^2}\gamma^2 var(X_{DB}(T)) - \frac{\lambda^2}{\lambda_i^2}\gamma\, cov(X_{PF}(T), X_{DB}(T)), \end{aligned}$$

for all $i \in N$ and $(X_i^{\text{post}}(T))_{i \in N} \in \mathcal{PO}(N)$.

γ for which the hedge benefits are maximized. In this case, it would be most beneficial for a pension fund to find a death benefit insurer with a value of γ around 0.4.

Now, consider the effect of heterogeneous beliefs on the benefits from redistribution. Comparing the solid and the dashed lines in the bottom panel of Figure 3.5 shows that, for all values of γ, the zero-utility premiums of the risk redistributions are larger in the case of heterogeneous beliefs. This difference is biggest for γ between 0.2 and 0.7. However, whereas the increase in the benefits is rather steep when γ increases from 0 to 0.4, a further increase in γ has a very small effect on the benefits from redistribution. These results suggest that, while the capacity for hedging longevity risk in life annuities via death benefits may indeed be small, most of the benefits already occur with a relatively small death benefit insurer.

3.5 Conclusion

In this chapter we describe a game-theoretic model to determine redistributions of risk that are considered fair by all parties. We then use this model to investigate the potential benefits from redistributions of longevity risk between pension funds and life insurance companies. The redistribution takes the form of a swap contract with a prespecified horizon, i.e., depending on the realized mortality rates at horizon date, a net payment occurs from either the pension fund to the insurer, or vice versa. We allow for the case where the involved parties may have heterogeneous beliefs regarding the underlying probability distribution, arising for example from the use of different models to forecast future mortality rates. A swap contract is acceptable to all parties only if they all believe that they (weakly) benefit from it. Our results suggest that the benefits from redistributing risk can be substantial, even on a one-year horizon. Our results also suggest that heterogeneous beliefs regarding the underlying probability distribution can make redistribution more attractive.

3.A Proofs

Proof of Proposition 3.2.4: (i) The fact that *Stability* implies *No Pareto Improvement* follows immediately from (3.7) and (3.8). To see that *Stability* implies *Individual Rationality,* note that for all $i \in N$, it holds that $\{X_i\} = \mathcal{F}(\{i\})$ and $\Delta U_i(X_i) = 0$. Moreover, it follows from (3.8) that if $(X_i^{\text{post}})_{i\in N}$ satisfies *Stability*, then $X_i^{\text{post}} \in \mathcal{NI}(\{i\})$ for all $i \in N$. Combined with (3.4), this implies that $\Delta U_i(X_i^{\text{post}}) \geq \Delta U_i(X_i)$ for all $i \in N$.

(ii) It suffices to show that *No Pareto Improvement* and *Individual Rationality* implies *Stability.* This follows immediately from (3.7) and (3.8), and the fact that *Individual Rationality* of $(X_i^{\text{post}})_{i\in N}$ implies that $X_i^{\text{post}} \in \mathcal{NI}(\{i\})$ for all $i \in N$. □

Proof of Theorem 3.3.1: (i) This result is shown by Wilson (1968).

(ii) Let $k \in \mathbb{R}^N_{++}$ be given. First note that because the state space is finite, (3.11) and (3.12) is a finite set of constraints. Existence of a solution follows immediately from the fact that $\lim_{x\to a_i} u_i'(x) = \infty$, and $\lim_{x\to b_i} u_i'(x) = 0$, for all $i \in N$. Uniqueness of the solution follows from the fact that, for all $i \in N$, strict concavity of u_i implies strict monotonicity of u_i' for all $i \in N$. □

Proof of Corollary 3.3.3: We start with the "if" part. Suppose that there exists a $k \in \mathbb{R}^N_{++}$, such that $\frac{u_1'(X_1(\omega))\mathbb{P}_1(\omega)}{u_j'(X_j(\omega))\mathbb{P}_j(\{\omega\})} = k_j$ for all $\omega \in \Omega$ and all $j \in N$. Then, the prior risk profiles $(X_i)_{i\in N}$ satisfy (3.11) and (3.12). Hence, according to Theorem 3.3.1, $(X_i)_{i\in N} \in \mathcal{PO}(N)$. It then follows from (3.9) that there do not exist feasible posterior risk profiles $(\widetilde{X}_i^{\text{post}})_{i\in N} \in \mathcal{F}(N)$ such that $(\Delta U_i(\widetilde{X}_i^{\text{post}}))_{i\in N} \gneqq (\Delta U_i(X_i))_{i\in N} = 0$.

Next, we show the "only if" part. Suppose that there does not exist $(X_i^{\text{post}})_{i\in N} \in \mathcal{PO}(N)$ with $\Delta U_i(X_i^{\text{post}}))_{i\in N} \gneqq 0$. Then, it holds that $(X_i)_{i\in N} \in \mathcal{PO}(N)$, and, hence, $(X_i)_{i\in N}$ should satisfy (3.11) and (3.12) for some $k \in \mathbb{R}^N_{++}$ (Theorem 3.3.1). This implies that $k_1u_1'(X_1(\omega))\mathbb{P}_1(\{\omega\}) = k_ju_j'(X_j(\omega))\mathbb{P}_j(\{\omega\})$ for all $\omega \in \Omega$ and for all $j \in N\backslash\{1\}$. This concludes the proof. □

In order to prove Theorem 3.3.4, we introduce the correspondence V that assigns to each set of firms $S \subseteq N$ the set of potential expected utility gains from feasible redistributions of risk, allowing for "free disposal", i.e., for all $S \subseteq N$:

$$V(S) = \left\{a \in \mathbb{R}^S : \exists (X_i^{\text{post}})_{i\in S} \in \mathcal{F}(S) \text{ s.t. } a \leq (\Delta U_i(X_i^{\text{post}}))_{i\in S}\right\}. \tag{3.39}$$

For any $S \subseteq N$, we define the set $\partial V(S)$ as the boundary of $V(S)$. Moreover, for any $S \subseteq N$, we let $\mathcal{PO}(S)$ be the set of Pareto optimal redistributions of risk when the firms in S redistribute their risk, i.e., $\mathcal{PO}(S)$ is given by (3.9) with N replaced by S. Then we have the following proposition.

Proposition 3.A.1 *For every $S \subseteq N$, it holds that:*

(i) $V(S)$ is convex;

(ii) $\partial V(S) = \left\{(\Delta U_i(X_i^{post}))_{i\in S} : (X_i^{post})_{i\in S} \in \mathcal{PO}(S)\right\}$;

(iii) for every $x \in \partial V(S)$, there exists unique $(X_i^{post})_{i\in S} \in \mathcal{PO}(S)$ such that $x = (\Delta U_i(X_i^{post}))_{i\in S}$.

Proof: (i) The proof is a straightforward generalization of the proof of Riddell (1981), who showed this result in case of two firms. Let $S \subseteq N$, $a, b \in V(S)$ and $\gamma \in (0, 1)$. Then, there exist

$(X_i^{\text{post},a})_{i\in S}$ and $(X_i^{\text{post},b})_{i\in S}$ such that $\sum_{i\in S} X_i^{\text{post},a} = \sum_{i\in S} X_i$, $\sum_{i\in S} X_i^{\text{post},b} = \sum_{i\in S} X_i$, $a \leq \left(\Delta U_i(X_i^{\text{post},a})\right)_{i\in S}$ and $b \leq \left(\Delta U_i(X_i^{\text{post},b})\right)_{i\in S}$. Clearly, we have

$$\sum_{i\in S}\left(\gamma X_i^{\text{post},a} + (1-\gamma)X_i^{\text{post},b}\right) = \sum_{i\in S} X_i.$$

Moreover, by concavity of $u_i(\cdot)$, it follows that

$$\Delta U_i(\gamma X_i^{\text{post},a} + (1-\gamma)X_i^{\text{post},b}) \geq \gamma \Delta U_i(X_i^{\text{post},a}) + (1-\gamma)\Delta U_i(X_i^{\text{post},b}) \geq \gamma a_i + (1-\gamma)b_i,$$

for all $i \in S$. Hence, $\gamma a + (1-\gamma)b \in V(S)$ and, therefore, (i) holds true.

(ii) From "free disposal" of $V(S)$ and monotonicity of u_i, it follows that

$$\partial V(S) = \{x \in V(S) : \nexists y \in V(S) \text{ s.t. } y > x\} \tag{3.40}$$

$$= \{x \in V(S) : \nexists y \in V(S) \text{ s.t. } y \gneq x\}. \tag{3.41}$$

Note that $a \in \partial V(S)$ if and only if there does not exist an $(\widetilde{X}_i^{\text{post}})_{i\in S} \in \mathcal{F}(S)$ such that $(\Delta U_i(\widetilde{X}_i^{\text{post}}))_{i\in S} \gneq a$. From this and $\partial V(S) \subset V(S)$, it follows that for every $a \in \partial V(S)$, there exist feasible posterior risk profiles $(X_i^{\text{post}})_{i\in S} \in \mathcal{F}(S)$ such that $a = (\Delta U_i(X_i^{\text{post}}))_{i\in S}$. Moreover, it is verified immediately that $(X_i^{\text{post}})_{i\in S} \in \mathcal{PO}(S)$ implies that $(\Delta U_i(X_i^{\text{post}}))_{i\in S} \in \partial V(S)$. Hence, (ii) holds true.

(iii) Let $x \in \partial V(S)$ be given, and suppose that there exist $(X_i^{\text{post},a})_{i\in S}$, $(X_i^{\text{post},b})_{i\in S} \in \mathcal{PO}(S)$ with $(X_i^{\text{post},a})_{i\in S} \neq (X_i^{\text{post},b})_{i\in S}$ and $(\Delta U_i(X_i^{\text{post},a}))_{i\in S} = (\Delta U_i(X_i^{\text{post},b}))_{i\in S} = x$. Then, by strict concavity of u_i, for $i \in S$, we have that $\Delta U_i(\frac{1}{2}X_i^{\text{post},a} + \frac{1}{2}X_i^{\text{post},b}) \geq \frac{1}{2}\Delta U_i(X_i^{\text{post},a}) + \frac{1}{2}\Delta U_i(X_i^{\text{post},b}) = \Delta U_i(X_i^{\text{post},a}) = \Delta U_i(X_i^{\text{post},b})$ for all $i \in S$ with at least one strict inequality. Because $\frac{1}{2}(X_i^{\text{post},a})_{i\in S} + \frac{1}{2}(X_i^{\text{post},b})_{i\in S} \in \mathcal{F}(S)$, this contradicts the fact that $(X_i^{\text{post},a})_{i\in S}$, $(X_i^{\text{post},b})_{i\in S} \in \mathcal{PO}(S)$. Hence, for every $x \in \partial V(S)$, there exists a unique $(X_i^{\text{post}})_{i\in S} \in \mathcal{PO}(S)$ such that $x = (\Delta U_i(X_i^{\text{post}}))_{i\in S}$. Hence, (iii) holds true. □

Proof of Theorem 3.3.4: Scarf (1967) considers the correspondence $\widehat{V}$ defined as

$$\widehat{V}(S) = \left\{a \in \mathbb{R}^S : \exists (X_i^{\text{post}})_{i\in S} \in \widehat{\mathcal{F}}(S) \text{ s.t. } a \leq (\widehat{u}_i(X_i^{\text{post}}))_{i\in S}\right\}, \tag{3.42}$$

where $\widehat{\mathcal{F}}(S)=\left\{(X_i^{\text{post}})_{i\in S} \in \mathbb{R}^S : \sum_{i\in S} X_i^{\text{post}} = \sum_{i\in S} X_i\right\}$, and for each $i \in N$, $\widehat{u}_i : \mathbb{R} \to \mathbb{R}$ is monotone and concave. He shows that the core of the corresponding NTU-game, i.e., the set

$$C(N, \widehat{V}) = \left\{x \in \widehat{V}(N) : \nexists S \subseteq N \text{ s.t. } (x_i)_{i\in S} \in \widehat{V}(S)\backslash \partial \widehat{V}(S)\right\} \tag{3.43}$$

is non-empty. First, note that the correspondence V defined in (3.39) follows from (3.42) by setting $\widehat{u}_i = \Delta U_i$, for all $i \in N$, and by replacing $\widehat{\mathcal{F}}(S)$ by $\mathcal{F}(S)$ as defined in (3.3), i.e., by allowing the domain D_i, $i \in N$, to be a convex subset of $\mathbb{R}$. Using the fact that, for all $i \in N$, concavity of u_i implies concavity of ΔU_i, that $\lim_{x \to a_i} u_i'(x) = \infty$, $\lim_{x \to b_i} u_i'(x) = 0$, and that $u_i''(\cdot) < 0$, it is verified immediately that the proof in Scarf (1967) extends to the correspondence V. Hence, it follows that the core of the corresponding NTU-game, which is given by

$$C(N,V) = \{x \in V(N) : \nexists S \subseteq N \text{ s.t. } (x_i)_{i \in S} \in V(S) \backslash \partial V(S)\}, \tag{3.44}$$

is non-empty. Next, we show that

$$C(N,V) \subseteq \left\{(\Delta U_i(X_i^{\text{post}}))_{i \in N} : (X_i^{\text{post}})_{i \in N} \in \mathcal{S}(N)\right\}. \tag{3.45}$$

Let $a \in C(N,V)$ be given. This implies that $a \in \partial V(N)$. It therefore follows from Proposition 3.A.1(ii) that there exists an $(X_i^{\text{post}})_{i \in N} \in \mathcal{PO}(N)$ such that $a = (\Delta U_i(X_i^{\text{post}}))_{i \in N}$.

To show that $(X_i^{\text{post}})_{i \in N} \in \mathcal{S}(N)$, we show that if there exist $S \subseteq N$ and $(\widetilde{X}_i)_{i \in S} \in \mathcal{F}(S)$ such that $(\Delta U_i(X_i^{\text{post}}))_{i \in S} \leq (\Delta U_i(\widetilde{X}_i))_{i \in S}$, then $(\Delta U_i(\widetilde{X}_i))_{i \in S} = (\Delta U_i(X_i^{\text{post}}))_{i \in S}$. Suppose there exist $S \subseteq N$ and $(\widetilde{X}_i)_{i \in S} \in \mathcal{F}(S)$ such that

$$(\Delta U_i(X_i^{\text{post}}))_{i \in S} \leq (\Delta U_i(\widetilde{X}_i))_{i \in S}.$$

This implies that $(\Delta U_i(X_i^{\text{post}}))_{i \in S} \in V(S)$. Because $(\Delta U_i(X_i^{\text{post}}))_{i \in N} = a \in C(N,V)$, it follows from (3.44) that $(\Delta U_i(X_i^{\text{post}}))_{i \in S} \in \partial V(S)$. It then follows from Proposition 3.A.1(iii) that there exists a $(\widehat{X}_i)_{i \in S} \in \mathcal{PO}(S)$ such that $(\Delta U_i(X_i^{\text{post}}))_{i \in S} = (\Delta U_i(\widehat{X}_i))_{i \in S}$. Because $(\widehat{X}_i)_{i \in S} \in \mathcal{NI}(S)$, $(\widetilde{X}_i)_{i \in S} \in \mathcal{F}(S)$, and $(\Delta U_i(\widehat{X}_i))_{i \in S} = (\Delta U_i(X_i^{\text{post}}))_{i \in S} \leq (\Delta U_i(\widetilde{X}_i))_{i \in S}$, it follows from (3.4) that $(\Delta U_i(\widetilde{X}_i))_{i \in S} = (\Delta U_i(\widehat{X}_i))_{i \in S}$. Hence, we can conclude that

$$(\Delta U_i(\widetilde{X}_i))_{i \in S} = (\Delta U_i(X_i^{\text{post}}))_{i \in S}.$$

Hence, there do not exist $S \subseteq N$ and $(\widetilde{X}_i)_{i \in S} \in \mathcal{F}(S)$ such that $(\Delta U_i(X_i^{\text{post}}))_{i \in S} \lneqq (\Delta U_i(\widetilde{X}_i))_{i \in S}$. This implies that $(X_i^{\text{post}})_{i \in N} \in \mathcal{S}(N)$. Because $a = (\Delta U_i(X_i^{\text{post}}))_{i \in N}$, we can conclude that the inclusion in (3.45) holds true. Because $C(N,V)$ is non-empty, this concludes the proof. □

Proof of Proposition 3.3.5: First, we show that

$$\{(\Delta U_i(X_i^{\text{post}}))_{i \in N} : (X_i^{\text{post}})_{i \in N} \in \mathcal{S}(N)\} = C(N,V). \tag{3.46}$$

The "$\supseteq$" part is shown in (3.45), so we only need to show the "$\subseteq$" part. Let $x \in \{(\Delta U_i(X_i^{\text{post}}))_{i \in N} : (X_i^{\text{post}})_{i \in N} \in \mathcal{S}(N)\}$. So, there exists an $(\hat{X}_i)_{i \in N} \in \mathcal{S}(N)$ such that $x = (\Delta U_i(\hat{X}_i))_{i \in N}$. Suppose that $x \notin C(N,V)$. Then, there exists an $S \subseteq N$ such that $(x_i)_{i \in S} \in V(S) \backslash \partial V(S)$. Then, there exists an $(\widetilde{x}_i)_{i \in S} \gneqq (x_i)_{i \in S}$ such that $(\widetilde{x}_i)_{i \in S} \in \partial V(S)$. According to Proposition

3.A.1(ii), there exists an $(\widetilde{X}_i)_{i\in S} \in \mathcal{PO}(S)$ such that $(\widetilde{x}_i)_{i\in S} = (\Delta U_i(\widetilde{X}_i))_{i\in S}$. So, it holds that $(\hat{X}_i)_{i\in S} \notin \mathcal{NI}(S)$. This is a contradiction, so $x \in C(N,V)$ and (3.46) holds.

So, it holds that

$$\mathcal{CNB} = \left\{ (X_i^{\text{post}})_{i\in N} \in \mathcal{S}(N) : (\Delta U_i(X_i^{\text{post}}))_{i\in N} \in \operatorname*{argmax}_{x\in C(N,V)} \prod_{i\in N} x_i \right\}. \tag{3.47}$$

The core is given by

$$C(N,V) = V(N) \cap \bigcap_{S\subseteq N} \left\{ x \in \mathbb{R}^N : (x_i)_{i\in S} \in \mathbb{R}^S \backslash (V(S) \backslash \partial V(S)) \right\}. \tag{3.48}$$

Since the set $V(N)$ is closed and the set $\left\{x \in \mathbb{R}^N : (x_i)_{i\in S} \in V(S)\backslash\partial V(S)\right\}$ is open for every $S \subseteq N$, it follows that $C(N,V)$ is closed. Since $C(N,V) \subset V(N)\cap\mathbb{R}^N_+$ and $V(N)\cap\mathbb{R}^N_+$ is bounded, the set $C(N,V)$ is compact. Moreover, in the proof of Theorem 3.3.4, we show that $C(N,V)$ is non-empty. Hence, in (3.47), there exists an $x \in C(N,V)$ such that $x = \operatorname{argmax}_{\widehat{x}\in C(N,V)} \prod_{i\in N} \widehat{x}_i$. Since $C(N,V) \subset \partial V(N)$, it follows that $x \in \partial V(N)$. Hence, it follows from Proposition 3.A.1(ii) that there exists an $(X_i^{\text{post}})_{i\in N} \in \mathcal{PO}(N)$ such that $x = (\Delta U_i(X_i^{\text{post}}))_{i\in N}$. This concludes the proof. □

Proof of Proposition 3.3.6: Consider

$$\widehat{\mathcal{NB}} = \operatorname*{argmax}_{x\in V(N), x\geq 0} \prod_{i\in N} x_i, \tag{3.49}$$

where $V(N)$ is as defined in (3.39). We know from Proposition 3.A.1 that $V(N)$ is convex. Moreover, it is easily verified that $V(N)$ is comprehensive, and that $V(N)\cap\mathbb{R}^N_+$ is non-empty and compact. It therefore follows from Nash (1950) that $\widehat{\mathcal{NB}}$ is non-empty, single-valued, and satisfies $\widehat{\mathcal{NB}} \subset \partial V(N)\cap\mathbb{R}^N_+$. It then follows from Proposition 3.A.1(ii) that for every $x \in \widehat{\mathcal{NB}}$, there exists a $(X_i^{\text{post}})_{i\in N} \in \mathcal{PO}(N)$ such that $x = \Delta U_i(X_i^{\text{post}}))_{i\in N} \geq 0$, i.e.,

$$\widehat{\mathcal{NB}} \subseteq \left\{ (\Delta U_i(X_i^{\text{post}}))_{i\in N} : (X_i^{\text{post}})_{i\in N} \in \mathcal{PO}(N) \cap \mathcal{IR}(N) \right\}.$$

This implies that

$$\begin{aligned} \widehat{\mathcal{NB}} &= \left\{ (\Delta U_i(X_i^{\text{post}}))_{i\in N} : (X_i^{\text{post}})_{i\in N} \in \operatorname*{argmax}_{(\widehat{X}_i)_{i\in N}\in\mathcal{PO}(N)\cap\mathcal{IR}(N)} \prod_{i\in N} \Delta U_i(\widehat{X}_i) \right\} \\ &= \left\{ (\Delta U_i(X_i^{\text{post}}))_{i\in N} : (X_i^{\text{post}})_{i\in N} \in \mathcal{NB} \right\}. \end{aligned}$$

Because $\widehat{\mathcal{NB}}$ is non-empty, it follows that $\mathcal{NB}$ is non-empty. Moreover, because $\widehat{\mathcal{NB}}$ is single-valued, it follows from Proposition 3.A.1(iii), that $\mathcal{NB}$ is single-valued. □

3.B Portfolio characteristics and data

In this appendix, we provide the portfolio characteristics and the data that we use in Subsection 3.4.2. The characteristics of the portfolios are as follows:

1. The pension fund has 50,000 male participants. The participants have accrued rights for a (deferred) single life annuity that yields a nominal yearly payment, with a first payment at the beginning of the year in which the insured reaches age 65, and a last payment at the beginning of the year in which the insured dies. The accrued right depends on age, and is normalized to 1 for a 65 year old.[27] The age composition as well as the accrued rights as a function of age are displayed in Figure 3.6. The average age of the participants is approximately 60.

2. The death benefit insurer has a portfolio of death benefit insurance contracts that pay a lump sum at the end of the year in which the insured dies, in case of decease of the insured before age 65. The age composition is displayed in Figure 3.7. The average age of the death benefit policyholder is approximately 42.[28] Each policyholder has a normalized insured right in case of decease of 10 (10 times the annual annuity payment of a 65 year old).

3. In Subsection 3.4.2, the number of insureds is 19,420. This implies that the relative size γ as defined in (3.38) equals 0.2, i.e., based on LC(1977-2009), the date-0 expected present value of the liabilities of the insurer is 20% of the date-0 expected present value of the liabilities of the pension fund. Moreover, we investigate the effect of the relative size of the pension fund and the insurer by considering the case where the number of insureds is $19,420 \cdot \frac{\gamma}{0.2}$, for $\gamma \in \{\frac{k}{100} : k = 1, \ldots, 100\}$. Then, $\gamma = 1$ corresponds to the case where the date-0 expected present value of the liabilities of the pension fund is equal to the expected present value of the liabilities of the death benefit insurer in case of homogeneous beliefs regarding the underlying probability distribution with LC(1977-2009).

4. The return on assets equals $r = 3\%$.

The age composition of the pension fund and the accrued rights of the participants as a function of their age are displayed in Figure 3.6. These characteristics are based on data from a large Dutch pension fund. The age composition of the death benefit insurer's portfolio is displayed in Figure 3.7.

[27] Note that a normalization is allowed for CARA utility functions as in (3.14), since we adjust the risk aversion parameter λ_{PF} accordingly.

[28] The rationale for young participants of a death benefit insurer is that death benefit insurance is sometimes mandated if individuals buy a mortgage.

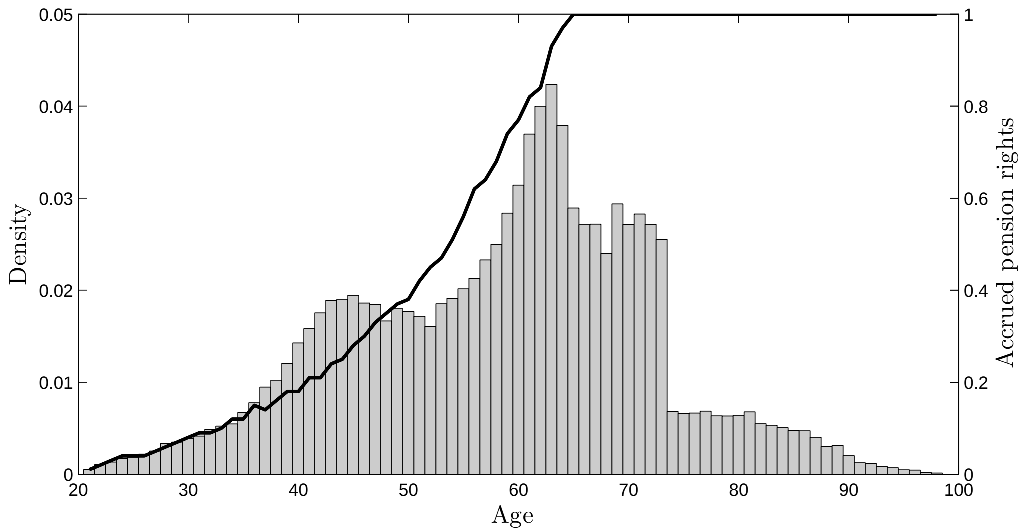


Figure 3.6: The age composition and the accrued pension rights ($\delta_{PF,j}$) of the participants of a pension fund, where we normalize the pension right at retirement to one.

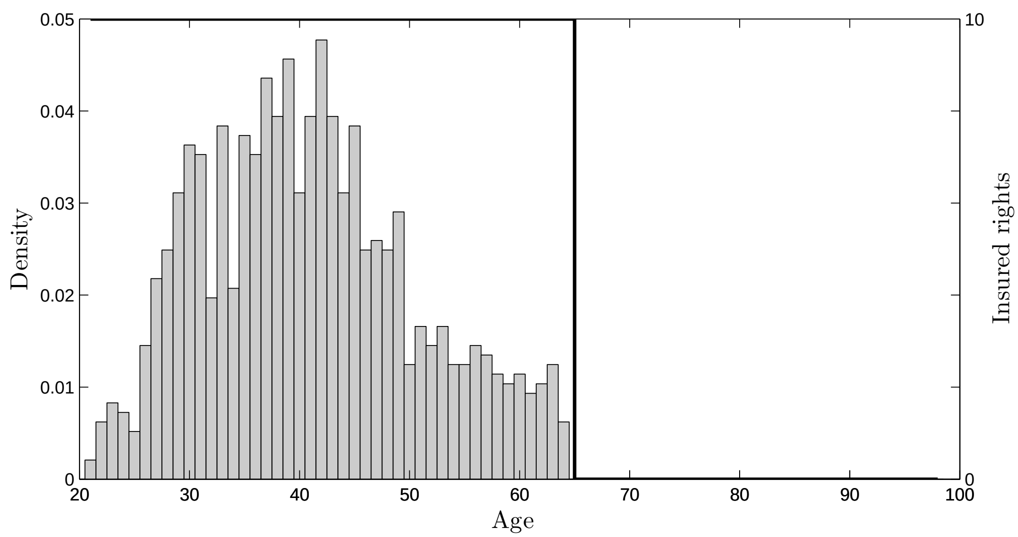


Figure 3.7: The age composition and the insured rights ($\delta_{DB,j}$) of the participants of a death benefit insurer.

3.C Simulation of $CL_i(T)$

In this appendix, we describe the method of simulating that we use to approximate the probability distributions of $CL_i(T)$ for $T \in \{1, 2, \ldots, 20, T^{\max}\}$ and $i \in \{PF, DB\}$. Recall from (3.27) that

$$CL_i(T) = \sum_{\tau=1}^{T} \frac{\widetilde{L}_{i,\tau}}{(1+r)^\tau} + \frac{L_i(T)}{(1+r)^T}.$$

We determine $L_i(T)$ as the best estimate value of the date-T liabilities, i.e.,

$$L_i(T) = \sum_{\tau=1}^{T^{\max}-T} \frac{\widetilde{L}^{(\mathrm{BE})}_{i,T+\tau}}{(1+r)^\tau}, \tag{3.50}$$

where $\widetilde{L}^{(\mathrm{BE})}_{i,T+\tau}$ is the liability payment at future date $T+\tau$ corresponding to the date-T best estimate of the death rates. To simulate the probability distribution of $CL_i(T)$, we use the following procedure:

- We simulate S trajectories for $({}_\tau p_{x,0})_{\tau=1,\ldots,T}$ using (3.51)-(3.54) in Appendix 3.D.
- For every simulated trajectory of $({}_\tau p_{x,0})_{\tau=1,\ldots,T}$,
  - we determine the corresponding value of $\sum_{\tau=1}^{T} \widetilde{L}_{i,\tau}(1+r)^{-\tau}$ using (3.33),
  - we re-estimate the model and estimate best estimates of $({}_\tau p_{x,T})_{\tau\geq 1}$ by setting $\varepsilon_{x,t} = 0$ for all x and all $t > T$ in (3.52), and $e_t = 0$ for all $t > T$ in (3.53). In this way, we construct the "best estimate" projection of future mortality rates, which is used to compute $L_i(T)$ via (3.50).[29]

3.D Lee-Carter model

In this appendix, we describe the Lee-Carter (1992) model. The probability that an individual of age x at time t survives the next year is modeled as

$$p_{x,t} = \exp(-m_{x,t}), \tag{3.51}$$

where $m_{x,t}$ represents the central death rate of a man with age x at time t (see, e.g., Pitacco et al., 2009). The central death rate is given by $m_{x,t} = D_{x,t}/E_{x,t}$, where $D_{x,t}$ is the observed number of deaths in year t in the cohort aged x at the beginning of year t, and $E_{x,t}$ is the corresponding

[29] In some cases, regulators may prescribe the use of a specific probability distribution to determine the best estimate value of the liabilities for regulatory purposes. Then, depending on whether the firm's objective is to reduce the volatility of the Net Asset Value with liability value as prescribed by the regulator, or the Net Asset Value with liability value determined according to its own probability distribution, the firm would use either the exogenously given projection, or its own subjective projection to valuate the liabilities.

number of persons. Lee and Carter (1992) propose the following log-bilinear relationship:

$$\log(m_{x,t}) = a_x + b_x\kappa_t + \varepsilon_{x,t}, \qquad \varepsilon_{x,t} : \mathcal{K}_t \sim^{i.i.d.} N(0, \sigma_x^2), \tag{3.52}$$

for all $t = t_0, \ldots, 0$ and $x = 1, 2, \ldots, 100$, where $\mathcal{K}_t = \{\kappa_{\widetilde{t}} : \widetilde{t} = t_0, \ldots, t\}$ and $t_0 < 0$. Here, t_0 is the first date in the data that is used and $t = 0$ is the last year in the dataset. The following normalizations are imposed: $\sum_{x=1}^{100} b_x = 1$ and $\sum_{t=t_0}^{0} \kappa_t = 0$. Then, the dynamics of the distribution of longevity are captured by κ_t. The estimates of a_x, b_x and κ_t are obtained via singular value decomposition.

Future values of κ_t are forecasted using an ARIMA(0,1,1) model:

$$\kappa_t = \kappa_{t-1} + c + e_t + \theta e_{t-1}, \tag{3.53}$$

for all $t \geq 1$, where we impose the following distribution of the errors:

$$e_t : \mathcal{K}_{t-1} \sim N(0, \sigma^2). \tag{3.54}$$

We include parameter uncertainty in our simulations. So, for every forecast, we take into account that the estimates of a_x and b_x at $t = 0$ are not fixed. Including parameter uncertainty increases the impact of longevity risk. For a discussion, we refer to Hári et al. (2008). In Table 3.2, we show the parameter estimates for Dutch males as reported in the HMD database.

	$\hat{c}$	$\hat{\theta}$	$\hat{\sigma}$
LC(1977-2009)	-2.00	-0.27	2.29
LC(1987-2009)	-2.13	-0.11	2.65

Table 3.2: Estimates of c, θ and σ in the ARIMA(0,1,1) model for male mortality rates, corresponding to (3.53) and (3.54), using HMD data from 1977 to 2009 (first row), and using HMD data from 1987 to 2009 (second row).

3.E Summary statistics of $CL_i(T)$

In this appendix, we present summary statistics of $CL_i(T)$ as a function of the horizon T, where we simulate $S = 50,000$ times. These a shown in Table 3.3. Panels A and B present summary statistics of $CL_i(T)$ for the pension fund and for the death benefit insurer, for the case where the probability distribution of future mortality is estimated based on the Lee-Carter model with parameters estimated based on the historical period 1977-2009. Panels C and D present summary statistics of $CL_i(T)$ for the pension fund and for the death benefit insurer for the case where the probability distribution of future mortality is estimated based on the Lee-Carter model with parameters estimated based on the historical period 1987-2009.

Panel A: pension fund LC(1977-2009)				
	$E[CL_i(T)]$	$\sigma[CL_i(T)]$	$\sigma[CL_i(T)]/E[CL_i(T)]$	buffer_i
$T = 1$	$3.43 \cdot 10^5$	$3.34 \cdot 10^3$	0.97%	1.91%
$T = 5$	$3.42 \cdot 10^5$	$5.62 \cdot 10^3$	1.64%	3.17%
$T = 10$	$3.42 \cdot 10^5$	$6.19 \cdot 10^3$	1.81%	3.48%
$T = 15$	$3.42 \cdot 10^5$	$6.12 \cdot 10^3$	1.79%	3.44%
$T = 20$	$3.42 \cdot 10^5$	$5.97 \cdot 10^3$	1.74%	3.37%
$T = T^{\max}$	$3.42 \cdot 10^5$	$5.80 \cdot 10^3$	1.69%	3.27%
Panel B: death benefit insurer LC(1977-2009)				
	$E[CL_i(T)]$	$\sigma[CL_i(T)]$	$\sigma[CL_i(T)]/E[CL_i(T)]$	buffer_i
$T = 1$	$6.80 \cdot 10^4$	$3.20 \cdot 10^3$	4.70%	9.53%
$T = 5$	$6.83 \cdot 10^4$	$5.84 \cdot 10^3$	8.56%	17.97%
$T = 10$	$6.84 \cdot 10^4$	$6.86 \cdot 10^3$	10.02%	21.29%
$T = 15$	$6.85 \cdot 10^4$	$6.97 \cdot 10^3$	10.18%	21.84%
$T = 20$	$6.85 \cdot 10^4$	$6.79 \cdot 10^3$	9.92%	21.18%
$T = T^{\max}$	$6.85 \cdot 10^4$	$6.43 \cdot 10^3$	9.39%	19.99%
Panel C: pension fund LC(1987-2009)				
	$E[CL_i(T)]$	$\sigma[CL_i(T)]$	$\sigma[CL_i(T)]/E[CL_i(T)]$	buffer_i
$T = 1$	$3.49 \cdot 10^5$	$4.83 \cdot 10^3$	1.38%	2.72%
$T = 5$	$3.49 \cdot 10^5$	$8.38 \cdot 10^3$	2.40%	4.66%
$T = 10$	$3.49 \cdot 10^5$	$9.16 \cdot 10^3$	2.62%	5.07%
$T = 15$	$3.49 \cdot 10^5$	$8.94 \cdot 10^3$	2.56%	4.96%
$T = 20$	$3.49 \cdot 10^5$	$8.59 \cdot 10^3$	2.46%	4.77%
$T = T^{\max}$	$3.489 \cdot 10^5$	$8.09 \cdot 10^3$	2.32%	4.50%
Panel D: death benefit insurer LC(1987-2009)				
	$E[CL_i(T)]$	$\sigma[CL_i(T)]$	$\sigma[CL_i(T)]/E[CL_i(T)]$	buffer_i
$T = 1$	$6.82 \cdot 10^4$	$3.77 \cdot 10^3$	5.53%	11.22%
$T = 5$	$6.87 \cdot 10^4$	$6.77 \cdot 10^3$	9.85%	20.73%
$T = 10$	$6.89 \cdot 10^4$	$7.74 \cdot 10^3$	11.23%	24.23%
$T = 15$	$6.89 \cdot 10^4$	$7.69 \cdot 10^3$	11.15%	24.16%
$T = 20$	$6.89 \cdot 10^4$	$7.31 \cdot 10^3$	10.60%	22.96%
$T = T^{\max}$	$6.88 \cdot 10^4$	$6.70 \cdot 10^3$	9.73%	20.81%

Table 3.3: Summary statistics of $CL_i(T)$ for $T \in \{1, 5, 10, 15, 20, T^{\max}\}$ and $i \in \{PF, DB\}$. Here, σ is the standard deviation and buffer_i is defined by $Q_{0.975,\mathbb{P}_i}(CL_i(T))/E[CL_i(T)] - 1$, where $Q_{0.975,\mathbb{P}_i}(CL_i(T))$ is the 97.5%-quantile with respect to firm i's subjective probability measure.

We note that the expected value of $CL_i(T)$ depends on the horizon T. Fluctuations in $E_{\mathbb{P}_i}[CL_i(T)]$ over time T are therefore not necessarily an indication of simulation error.

Table 3.4 displays the correlation between $CL_{PF}(T)$ and $CL_{DB}(T)$, as a function of the horizon T. We find strong negative correlations. This is also illustrated in the scatter plot in Figure 3.2.

	$T=1$	$T=5$	$T=10$	$T=15$	$T=20$	$T=T^{\max}$
LC(1977-2009)	-0.88	-0.96	-0.97	-0.97	-0.97	-0.97
LC(1987-2009)	-0.92	-0.97	-0.98	-0.98	-0.97	-0.97

Table 3.4: Correlation coefficient of $CL_{PF}(T)$ and $CL_{DB}(T)$ for $T \in \{1, 5, 10, 15, 20, T^{\max}\}$.

3.F Approximation welfare gains

In this appendix, we discuss the method we use to approximate the welfare gains in both the homogeneous and the heterogeneous case. To approximate the welfare gains in the heterogeneous case, we discretize the state space. We let the state space be given by outcomes of the stochastic variable $X(T) = X_{PF}(T) + X_{DB}(T)$ only. We first simulate $X(T)$ $S = 50,000$ times under the probability measures LC(1977-2009) and LC(1987-2009). We impose a partition of 25 equally-sized subintervals of the interval $[\min_{\mathbb{P}_{PF},\mathbb{P}_{DB}}(X(T)), \max_{\mathbb{P}_{PF},\mathbb{P}_{DB}}(X(T))]$. Every sub-interval corresponds with a state $\omega \in \Omega$, so that $|\Omega| = 25$. For every state $\omega \in \Omega$, we compute $\hat{X}(\omega)$ by averaging over all simulated scenarios on the relevant sub-interval. Moreover, for every state $\omega \in \Omega$, we compute $\mathbb{P}_{PF}(\{\omega\})$ and $\mathbb{P}_{DB}(\{\omega\})$ by computing the frequency that the simulated scenarios are in the relevant sub-interval. If $\mathbb{P}_i(\{\omega\}) = 0$ for some $i \in \{PF, DB\}$ and $\omega \in \Omega$, we set this probability at 10^{-10}. The sensitivity of our choice of this probability and the choice of the partition size is negligible for reasonable values. The use of the method to only impose a partition on $X(T)$ instead of $\{X_{PF}(T), X_{DB}(T)\}$ leads to an underestimation of the welfare gains.

To compare the results, we also use a discretization to calculate $E_{\mathbb{P}_i}[X_i(T)]$ for $i \in \{PF, DB\}$. We approximate $X_i(T)$ via a grid with 25 sub-intervals in the same way as we did with $X(T)$. Moreover, the welfare gains in the homogeneous case are also calculated via a discretization. We hereby use approximately the same length of a subinterval.

Chapter 4

Risk Redistribution with Distortion Risk Measures

This chapter is based on Boonen (2013).

4.1 Introduction

In this chapter, we study the question how to redistribute risk if firms use a distortion risk measure in order to evaluate risk. There is a relatively large literature that analyzes optimal redistributions of risk, based on the seminal work of Borch (1962) and Wilson (1968). This chapter mainly differs in terms of the objective of firms. In this chapter, we study optimal risk sharing in the context of distortion risk measures instead of Von Neumann-Morgenstern expected utility. Distortion risk measures are used to define the preference relations of the firms. De Giorgi and Post (2008) show that the equilibrium prices with distortion risk measures are empirically better fitting US stock returns.

Distortion risk measures have applications in both actuarial science and finance, being related both to the dual theory of risk (Yaari, 1987) and coherent risk measures (Artzner, Delbaen, Eber and Heath, 1999). Existing literature has investigated the use of distortion risk measures as pricing mechanisms, e.g., Yaari (1987), Chateauneuf, Kast and Lapied (1996), Wang (1996, 2000), Wang, Young and Panjer (1997), and De Waegenaere, Kast and Lapied (2003). Yaari (1987) characterizes dual utility by a modification of the independence axiom in expected utility theory. Instead of requiring independence with respect to probability mixtures of risks, he requires independence with respect to direct mixing the realizations of the risks. The evaluation of a risk is linear in the pay-offs but non-linear in the probabilities. Distortion risk measures coincide with dual utility if firms are risk averse. Distortion risk measures differ in two fundamental ways from expected utility theory. First, they postulate cash-equivalent preferences, implying that cash payments do not affect risk preferences. For example, in insurance this implies that the price of a risk is independent of the initial wealth of the insurer. Second, distortion risk measures attempt to reflect business practices where Expected Shortfall has been gaining practitioner interest. In line with mean-variance preferences, we can formulate a risk preference based on a distortion risk measure by any trade-off between the expectation and a distortion risk measure. This leads to risk-reward preferences as in De Giorgi and Post (2008).

Ludkovski and Rüschendorf (2008) and Ludkovski and Young (2009) analyze risk redistribution in settings where risk is measured by a distortion risk measure. Their general focus is on Pareto optimality of risk redistributions. The approach that we propose to optimally redistribute risk is twofold. First, we derive all Pareto optimal redistributions, where all transfers such that all risk is redistributed are allowed. Pareto optimality of redistributions based on risk measures is first studied by Jouini, Schachermayer and Touzi (2008) and it is later applied to distortion risk measures by Ludkovski and Young (2009). Jouini et al. (2008) and Ludkovski and Young (2009) characterize all Pareto optimal risk redistributions as a finite sum of stop-loss contracts on the aggregate risk. We introduce two conditions that are jointly sufficient to ensure unique Pareto optimal risk redistributions up to side-payments.

Second, we analyze the problem to determine the size of the side-payments. If the number of firms is large and if there is a well-functioning market where firms act as price-takers, we focus on the *competitive equilibria*. Filipović and Kupper (2008), Dana and Le Van (2010), Dana (2011), and Flåm (2011) analyze existence of equilibria in markets where firms use risk measures. Our focus is on uniqueness of the equilibria. We provide three jointly sufficient conditions to guarantee that there is a unique equilibrium risk redistribution. From this unique equilibrium, we derive a corresponding capital asset pricing model (CAPM) for distortion risk measures. If all firms use the same risk measure, this CAPM model is derived by De Giorgi and Post (2008).

We also characterize optimal redistributions in cases where a well-functioning market does not exist, so that redistributions can only be obtained via multilateral trade. We model this bargaining problem using cooperative game theory. We require Pareto optimal risk redistributions to satisfy four properties. Via those four properties, we characterize a specific risk redistribution. This risk redistribution coincides with the competitive equilibrium. In cooperative game theory, this risk redistribution corresponds with the *Aumann-Shapley value* (Aumann and Shapley, 1974).

If all firms use the same risk measure, the risk redistribution problem can be analyzed using the risk capital allocation game as in, e.g., Denault (2001) and Csóka, Herings and Kóczy (2009). In a risk capital allocation problem, the aim is to find a rule for allocating the aggregate risk capital of a firm to its business units. The Aumann-Shapley value is very prominent in the literature on risk capital allocation problems (see, e.g., Tasche, 1999; Denault, 2001; Myers and Read, 2001; Tsanakas and Barnett, 2003; Kalkbrener, 2005). We relax the assumption that all firms use the same risk measure. If the aim is to redistribute risk among firms as in this chapter, the assumption that all firms use the same risk measure is restrictive. As argued by e.g. Tasche (1999), the risk capital allocation problem is mainly designed for financial performance measurement.

An important application of the problem described in this chapter is insurance. Insurance products are typically traded Over-The-Counter. Particularly in Over-The-Counter trades of insurance products, tranching of the aggregate risk empirically observed. Typically, the idiosyncratic part of insurance risk can be hedged via pooling. This creates incentives for insurance in the first place. Multiple small firms can benefit from this by pooling their risk with other firms. The systemic part of insurance risk cannot be hedged by pooling risk. However, firms can still benefit from trading the systematic part of insurance risk with other firms which face insurance risk in other insurance risk classes. The systematic risk can be shared and, therefore, firms can benefit by redistributing. For instance, the class of longevity risk is an application of risk where there is potential to redistribute risk since death benefit insurers and pension funds have typically naturally hedgeable risk exposures (see, e.g., Tsai, Wang and Tzeng, 2010; Wang, Huang, Yang and Tsai, 2010). This class of risk is typically not marketed and, therefore, traded Over-The-Counter (see, e.g., Chapter 3 of this dissertation).

This chapter is set out as follows. Section 4.2 introduces distortion risk measures and the model. Section 4.3 analyzes Pareto optimality. Section 4.4 derives the competitive equilibrium prices, as well as conditions such that the corresponding equilibrium risk redistribution is unique. Section 4.5 provides a cooperative game-theoretic approach to characterize this competitive equilibrium via, e.g., a fuzzy core criterion. Section 4.6 shows an equivalence result of the competitive equilibria and the fuzzy core. Finally, this chapter concludes in Sections 4.7 and 4.8.

4.2 Distortion risk measures and risk redistribution problems

In this chapter, we study the problem to redistribute risk among firms. The firms use a distortion risk measure in order to evaluate risk. In this section, we briefly introduce distortion risk measures and risk redistribution problems with distortion risk measures.

4.2.1 Distortion risk measures

In this subsection, we discuss distortion risk measures. Distortion risk measures are developed from research on dual utility by Yaari (1987). Moreover, it is developed as premium principle by Wang (1995). Let Ω be a finite state space and $\mathbb{P}$ the physical probability measure on the power set 2^{Ω}. Moreover, denote $\mathbb{R}^{\Omega}$ as the space of all real valued stochastic variables on Ω that are realized at a well-defined reference time. These stochastic variables are referred to as risks. A realization of a risk is interpreted as a future loss.

A risk measure is a function $\rho : \mathbb{R}^\Omega \to \mathbb{R}$, i.e., a risk measure maps risks into real numbers. For every risk $Y \in \mathbb{R}^\Omega$, we refer to $\rho(Y)$ as the risk adjusted value of the liabilities. A distortion risk measure is both related to dual theory (Yaari, 1987) and coherent risk measures (Artzner et al., 1999). In dual theory, Yaari (1987) characterizes distortion risk measures[30] by a modification of the independence axiom in the Von Neumann-Morgenstern expected utility theory. Wang (1995) defines a distortion risk measure by

$$\rho(Y) = \int_0^\infty g^\rho(1 - F_Y(x))\, dx + \int_{-\infty}^0 (g^\rho(1 - F_Y(x)) - 1)\, dx, \text{ for all } Y \in \mathbb{R}^\Omega, \tag{4.1}$$

for a continuous, concave and increasing distortion function $g^\rho : [0,1] \to [0,1]$ with $g^\rho(0) = 0$ and $g^\rho(1) = 1$, where F_Y is the cumulative density function (CDF) of risk Y. Here, convergence of the integrals is guaranteed by boundedness of the risk Y. Several authors, including Yaari (1987), Wang (1996, 2000), Wang et al. (1997), Chateauneuf et al. (1996) and De Waegenaere et al. (2003), suggested to use distortion risk measures to price risk. One of the appealing features of the distortion risk measure is that, unlike under Von Neumann-Morgenstern expected utility, the firm's attitude towards risk is not entangled with its attitude towards wealth. It can be shown that the class of distortion risk measures is equal to the class of spectral risk measures (Acerbi, 2002). There exists a wide literature on spectral risk measures since it is straightforward to apply them empirically (see, e.g., Acerbi, 2002; Adam, Houkari and Laurent, 2008).

We continue with some alternative representations of distortion risk measures that we use throughout this chapter. For a risk $Y \in \mathbb{R}^\Omega$, we order the state space $\Omega = \{\omega_1, \ldots, \omega_p\}$ such that $Y(\omega_1) \geq \cdots \geq Y(\omega_p)$. Then, we get via direct calculations that

$$\rho(Y) = \sum_{k=1}^{p-1} g^\rho(\mathbb{P}(\{\omega_1, \ldots, \omega_k\})) \cdot [Y(\omega_k) - Y(\omega_{k+1})] + Y(\omega_p). \tag{4.2}$$

Moreover, via direct calculations, we get from (4.2) that a distortion risk measure can be written as

$$\rho(Y) = E_{\mathbb{Q}_Y}[Y], \text{ for all } Y \in \mathbb{R}^\Omega, \tag{4.3}$$

where $\mathbb{Q}_Y : 2^\Omega \to (0,1]$ is the additive mapping such that

$$\mathbb{Q}_Y(\{\omega\}) = g^\rho(\mathbb{P}(Y \geq Y(\{\omega\}))) - g^\rho(\mathbb{P}(Y > Y(\{\omega\}))), \text{ for all } \omega \in \Omega. \tag{4.4}$$

Since g^ρ is increasing and such that $g^\rho(0) = 0$ and $g^\rho(1) = 1$, it holds that $\mathbb{Q}_Y$ is a probability measure for a given Y. A distortion risk measure evaluated in risk Y is its expectation under the probability measure $\mathbb{Q}_Y$ that assigns higher probabilities to worst-case realizations of the risk Y.

[30] Yaari (1987) characterizes dual utility. If firms are risk averse, the dual utility preferences can be represented by a distortion risk measure.

Wang et al. (1997) show that every distortion risk measure is coherent. Coherence is later formally introduced by Artzner et al. (1999). As formalized in Chapter 2 of this dissertation for unbounded risks, a risk measure $\rho : \mathbb{R}^\Omega \to \mathbb{R}$ is coherent if and only if it satisfies the following four properties:

- *Sub-additivity*: for all $Y, Z \in \mathbb{R}^\Omega$, we have

$$\rho(Y + Z) \leq \rho(Y) + \rho(Z).$$

- *Monotonicity*: for all $Y, Z \in \mathbb{R}^\Omega$ such that $Y(\omega) \geq Z(\omega)$ for all $\omega \in \Omega$, we have

$$\rho(Y) \geq \rho(Z).$$

- *Positive Homogeneity*: for every $Y \in \mathbb{R}^\Omega$ and every $c \in \mathbb{R}_+$, we have

$$\rho(cY) = c\rho(Y).$$

- *Translation Invariance*: for every $Y \in \mathbb{R}^\Omega$ and every $c \in \mathbb{R}$, we have

$$\rho(Y + c \cdot e_\Omega) = \rho(Y) + c.$$

Similar as in Chapter 2, we denote $e_\Omega \in \mathbb{R}^\Omega$ as the risk with realization one in every state $\omega \in \Omega$. More generally, the risk $e_A \in \mathbb{R}^\Omega$ for $A \subseteq \Omega$ is given by:

$$e_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{otherwise,} \end{cases} \tag{4.5}$$

for all $\omega \in \Omega$.

The relevance of these properties is widely discussed by Artzner et al. (1999). Particularly, *Sub-additivity* of a risk measure implies that aggregate risk adjusted value of the liabilities weakly decreases if risks are pooled. It also implies that there is no incentive for a firm to split its risk into pieces and evaluate them separately. The properties *Translation Invariance* and *Positive Homogeneity* imply that firms valuing risk via a risk measure, would also price risk via its corresponding risk measure, if, in absence of market prices, firms use indifference pricing and the initial wealth is deterministic. This follows from $\rho(w \cdot e_\Omega) = w = \rho(w \cdot e_\Omega + Y - \rho(Y) \cdot e_\Omega)$ for all $Y \in \mathbb{R}^\Omega$, where $w \in \mathbb{R}$ represents the initial deterministic wealth.

Distortion risk measures ρ are characterized as the risk measures that are coherent and satisfying the following two properties (see Wang et al., 1997):

- *Conditional State Independence*: $\rho(Y)$ depends on the risk $Y \in \mathbb{R}^\Omega$ only via its distribution.

- *Comonotonic Additivity*: for all $Y, Z \in \mathbb{R}^\Omega$ that are comonotone, we have

$$\rho(Y + Z) = \rho(Y) + \rho(Z).$$

Risks $Y, Z \in \mathbb{R}^\Omega$ are comonotone if there exists an ordering $(\omega_1, \ldots, \omega_p)$ on the state space Ω such that $\Omega = \{\omega_1, \ldots, \omega_p\}$, $Y(\omega_1) \geq \cdots \geq Y(\omega_p)$ and $Z(\omega_1) \geq \cdots \geq Z(\omega_p)$. According to Acerbi (2002), the property *Conditional State Independence* is in particular crucial to be able to estimate a risk measure from empirical data. Wang et al. (1997) show that, subject to the relatively mild conditions *Conditional State Independence* and *Comonotonic Additivity*, any coherent risk measure can be represented as a distortion risk measure.

Recall from (2.1) in Chapter 2 of this dissertation that a risk measure ρ is coherent if and only if there exists a set of probability measures Q such that

$$\rho(Y) = \sup \{E_{\mathbb{Q}}[Y] : \mathbb{Q} \in Q\}, \text{ for all } Y \in \mathbb{R}^\Omega. \tag{4.6}$$

Denote $\mathcal{P}(\Omega)$ as the set of all probability measures on the state space Ω. As shown in (2.4) in Chapter 2 of this dissertation , it holds that a representation of the set Q in (4.6) for distortion risk measures is given by

$$Q(\rho) = \{\mathbb{Q} \in \mathcal{P}(\Omega) : \mathbb{Q}(A) \leq g^\rho(\mathbb{P}(A)) \text{ for all } A \subset \Omega\}. \tag{4.7}$$

In the sequel, we discuss the problem to redistribute risk where all firms use distortion risk measures to evaluate risk.

4.2.2 The risk redistribution problem

In this subsection, we define the risk redistribution problem with distortion risk measures that we analyze in this chapter. Throughout this chapter, we fix the set of firms and the discrete state space such that:

- the finite collection of firms is given by $N = \{1, \ldots, n\}$;
- the state space is finite and given by Ω. Without loss of generality, it is assumed that $|\Omega| > 1$;[31]
- the physical probability measure is given by $\mathbb{P} : 2^\Omega \to (0, 1]$.[32] This measure is common knowledge.

Next, we define risk redistribution problems with distortion risk measures.

[31] We denote $|\Omega|$ for the cardinality of the state space Ω.
[32] We assume that $\mathbb{P}(\{\omega\}) > 0$ for each state $\omega \in \Omega$, i.e., the probability that a state occurs is strictly positive. As the state space is finite, this is without loss of generality; states with zero probability can be omitted from the state space.

Definition 4.2.1 A *risk redistribution problem with distortion risk measures* is a tuple $(X_i, \rho_i)_{i \in N}$, where

- $X_i \in \mathbb{R}^\Omega$ is the risk held by firm $i \in N$;
- $\rho_i : \mathbb{R}^\Omega \to \mathbb{R}$ is the distortion risk measure that firm $i \in N$ is endowed with. The corresponding distortion function is denoted by g_i.[33]

The class of risk redistribution problems with distortion risk measures is denoted by $\mathcal{RR}$.

In the sequel, we refer to a risk redistribution problem with distortion risk measures as a risk redistribution problem. There is common knowledge about the risks and risk measures of all firms. Moreover, we define the aggregate risk as $X = \sum_{i \in N} X_i$ and, without loss of generality, we order the state space $\Omega = \{\omega_1, \ldots, \omega_p\}$ such that

$$X(\omega_1) \geq \cdots \geq X(\omega_p). \tag{4.8}$$

For a risk redistribution problem, we aim to redistribute the aggregate risk X among firms. The objective of a firm is to minimize its risk adjusted value of the liabilities. We allow for all forms of risk redistributions, as long as the aggregate risk is redistributed. This leads to the following definition of feasible risk redistributions.

Definition 4.2.2 The set of *feasible* risk redistributions of a risk redistribution problem $R \in \mathcal{RR}$ is given by

$$\mathcal{F}(R) = \left\{ (\widetilde{X}_i)_{i \in N} \in (\mathbb{R}^\Omega)^N : \sum_{i \in N} \widetilde{X}_i = X \right\}. \tag{4.9}$$

The set of risk redistributions in (4.9) allows for, e.g., proportional or stop-loss contracts on the aggregate risk X.

The problem is similar to the problem of, e.g., Borch (1962), Wilson (1968) and Aase (1993). The setting in this chapter differs from those papers in that firms use a distortion risk measure ρ_i instead of an expected utility function U_i. An overview of the differences between both criterions on risk preferences is given by, e.g., Wang and Young (1998). A distortion risk measure ρ can be represented as a Von Neumann-Morgenstern expected utility function if and only if the distortion function is given by $g^\rho(x) = x$ for all $x \in [0, 1]$, i.e., if the risk measure is risk-neutral. This follows directly from the fact that the only class of expected utility functions satisfying *Positive Homogeneity* is the class of linear utility functions. Even though a distortion risk measure can generally not be written as an expected utility function, it is a utility function as it satisfies the well-known axioms of rationality of a preference relation.

[33] For notational convenience, we write g_i instead of g^{ρ_i}.

4.3 Pareto optimality

In this section, we first analyze *Pareto optimality* of risk redistributions. We first provide its definition and characterization in Subsection 4.3.1. We analyze Pareto optimal risk redistributions for two prominent special cases of the risk redistribution problem in Subsection 4.3.2. In Subsection 4.3.3, we provide a sufficient condition such that all Pareto optimal risk redistributions are comonotone with the aggregate risk. In Subsection 4.3.4, we provide an additional second condition such that there exists a unique Pareto optimal risk redistribution up to side-payments. Finally, in Subsection 4.3.5 we analyze when the Pareto improvements by redistributing risk can be obtained.

4.3.1 Definition and characterization

In this subsection, we provide the definition and characterization of Pareto optimal risk redistributions. Tsanakas and Christofides (2006), Acciaio (2007), Burgert and Rüschendorf (2008), Filipović and Svindland (2008), Jouini et al. (2008), Kiesel and Rüschendorf (2008), Kaina and Rüschendorf (2009) and Ludkovski and Young (2009) all analyze Pareto optimality of risk redistributions in settings where risk is measured by a risk measure. We summarize results that we need in this chapter.

A risk redistribution is called Pareto optimal if there does not exist another feasible redistribution that is weakly better for all firms, and strictly better for at least one firm. This leads to the following definition.

Definition 4.3.1 The set of *Pareto optimal* risk redistributions of a risk redistribution problem $R \in \mathcal{RR}$ is given by

$$\mathcal{PO}(R) = \left\{(\widetilde{X}_i)_{i\in N} \in \mathcal{F}(R) : \nexists (\hat{X}_i)_{i\in N} \in \mathcal{F}(R) \text{ s.t. } (\rho_i(\hat{X}_i))_{i\in N} \lneqq (\rho_i(\widetilde{X}_i))_{i\in N}\right\}. \quad (4.10)$$

Similar to Borch (1962) where firms use expected utilities, the set $\mathcal{PO}(R)$ only depends on the risks $X_i, i \in N$, via their sum X.

Next, we introduce side-payments. These are used to obtain a characterization of Pareto optimal risk redistributions. A risk $Y \in \mathbb{R}^\Omega$ is a *side-payment* if there exists a constant $c \in \mathbb{R}$ such that $Y = c \cdot e_\Omega$ (i.e., the risk Y is a degenerated stochastic variable). From the following lemma it follows that Pareto optimality of a risk redistribution is unaffected by adding side-payments $(c_i \cdot e_\Omega)_{i\in N}$ such that $\sum_{i\in N} c_i = 0$.

Lemma 4.3.2 *For all $R \in \mathcal{RR}$, it holds that $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$ if and only if $(\widetilde{X}_i + c_i e_\Omega)_{i\in N} \in \mathcal{PO}(R)$ for any $c \in \mathbb{R}^N$ such that $\sum_{i\in N} c_i = 0$.*

Lemma 4.3.2 follows straightforwardly from *Translation Invariance* of ρ_i and is also stated by

Ludkovski and Young (2009).[34] From Lemma 4.3.2 it follows that if we find a Pareto optimal risk redistribution $(\widetilde{X}_i)_{i\in N}$, we can construct a set of Pareto optimal risk redistributions by adding zero-sum side-payments to $(\widetilde{X}_i)_{i\in N}$. If there exists a risk redistribution $(\widetilde{X}_i)_{i\in N}$, such that the set of risk redistributions that arises from adding zero-sum side-payments equals the set of Pareto optimal risk redistributions, we call this risk redistribution $(\widetilde{X}_i)_{i\in N}$ the unique element of $\mathcal{PO}(R)$ up to side-payments.

Definition 4.3.3 For an $R \in \mathcal{RR}$, the risk redistribution $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$ is, *up to side-payments*, the unique element of $\mathcal{PO}(R)$ if

$$\mathcal{PO}(R) = \left\{ (\widetilde{X}_i + c_i e_\Omega)_{i\in N} : c \in \mathbb{R}^N, \sum_{i\in N} c_i = 0 \right\}.$$

The term side-payment is inspired by the procedure that first firms pick a Pareto optimal risk redistribution and, thereafter, add or subtract zero-sum side-payments. Uniqueness up to side-payments is first introduced in the context of risk measures by Jouini et al. (2008) in the case where one firm holds a very specific risk measure.

If firms use expected utilities, one obtains every Pareto optimal risk redistribution by maximizing a weighted sum of expected utilities (see Borch, 1962). Then, under mild regularity conditions, every (normalized) weight-vector yields a unique risk redistribution. For distortion risk measures, we also get the Pareto optimal risk redistributions by minimizing the weighted aggregate risk adjusted value of the liabilities, i.e., for every $k \in \mathbb{R}^N$ such that $k_1 = 1$ we minimize $\sum_{i\in N} k_i \rho_i(\widetilde{X}_i)$ over all $(\widetilde{X}_i)_{i\in N} \in \mathcal{F}(R)$. However, in case of distortion risk measures this minimization problem only has a solution if k equals the ones-vector, i.e., $k_i = 1$ for all $i \in N$. This follows from *Translation Invariance* of the distortion risk measures. If k equals the ones-vector, we get all Pareto optimal risk redistributions. So, as characterized in the following proposition, the set of Pareto optimal risk redistributions is given by the set of all feasible risk redistributions such that the aggregate risk adjusted value of the liabilities is minimal. This is shown by Jouini et al. (2008) and Filipović and Kupper (2008).

Proposition 4.3.4 *For all $R \in \mathcal{RR}$, it holds that $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$ if and only if $(\widetilde{X}_i)_{i\in N} \in \mathcal{F}(R)$ and $\sum_{i\in N} \rho_i(\widetilde{X}_i) = \min \left\{ \sum_{i\in N} \rho_i(\hat{X}_i) : (\hat{X}_i)_{i\in N} \in \mathcal{F}(R) \right\}$.*

4.3.2 Special cases

In this subsection, we analyze Pareto optimality for two special cases of the risk redistribution problem. First, we consider Pareto optimal risk redistributions in case all firms use the same risk measure. Then, there are still risk-sharing benefits due to diversification, which is the effect that pooling risk leads to Pareto improvements. For this special case, all comonotone risk redistributions are Pareto optimal, as shown in the following proposition.

[34] Lemma 4.3.2 also holds true if firms use exponential (CARA) utility functions (see Gerber and Pafumi, 1998).

Proposition 4.3.5 *If $R \in \mathcal{RR}$ is such that $\rho_i = \rho$ for all $i \in N$ and $(\widetilde{X}_i)_{i\in N} \in \mathcal{F}(R)$ such that all $\widetilde{X}_i, i \in N$ are comonotone with each other, then $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$.*

Proposition 4.3.5 implies that if all firms use the same distortion risk measure, any proportional or stop-loss redistribution of the aggregate risk X is Pareto optimal. Hence, there are generally many ways to split the aggregate risk in order to obtain Pareto optimal risk redistributions.

We next provide Pareto optimal risk redistributions for the second special case. The following proposition shows that if there is a firm that is endowed with a smaller distortion function than all other firms, it is Pareto optimal to shift all risk to this firm.

Proposition 4.3.6 *If $R \in \mathcal{RR}$ is such that there exists a firm $i \in N$ such that*

$$g_i(x) \leq g_j(x), \text{ for all } x \in [0,1] \text{ and } j \in N, \tag{4.11}$$

then for $\widetilde{X}_i = X$ and $\widetilde{X}_j = 0 \cdot e_\Omega$ for all $j \in N \backslash \{i\}$, it holds that $(\widetilde{X}_j)_{j\in N} \in \mathcal{PO}(R)$. Moreover, if

$$g_i(x) < g_j(x), \text{ for all } x \in (0,1) \text{ and } j \in N \backslash \{i\}, \tag{4.12}$$

then $(\widetilde{X}_j)_{j\in N}$ is, up to side-payments, the unique element of $\mathcal{PO}(R)$.

Condition (4.11) seems restrictive. However, we next provide examples where this condition holds. These examples focus on the case where all firms use an Expected Shortfall risk measure. This risk measure receives considerable attention in the actuarial and financial literature (see, e.g., Rockafellar and Uryasev, 2000; Acerbi and Tasche, 2002).

Example 4.3.7 In this example, we discuss three cases where condition (4.11) holds. For all $j \in N$, let firm j use the risk measure Expected Shortfall with significance level $\alpha_j \in (0,1)$, which is given by (see, e.g., Acerbi and Tasche 2002 and (2.6) in Chapter 2 of this dissertation):

$$\rho^{ES}_{\alpha_j}(Y) = \alpha_j^{-1}\left(E_{\mathbb{P}}\left[Y \cdot e_{\{\omega\in\Omega: Y(\omega)\geq q_{1-\alpha_j}(Y)\}}\right] - q_{1-\alpha_j}(Y)(\mathbb{P}[Y \geq q_{1-\alpha_j}(Y)] - \alpha_j)\right), \tag{4.13}$$

for all $Y \in \mathbb{R}^\Omega$, where $q_{1-\alpha_j}(Y)$ is the $(1-\alpha_j)$-quantile of risk Y, i.e.,

$$q_{1-\alpha_j}(Y) = \sup\{x \in \mathbb{R} : \mathbb{P}(Y \geq x) > \alpha_j\}, \text{ for all } Y \in \mathbb{R}^\Omega. \tag{4.14}$$

Recall from (2.7) in Chapter 2 of this dissertation that α_j-Expected Shortfall is a distortion risk measure with distortion function $g_j(x) = \min\left\{\frac{x}{\alpha_j}, 1\right\}$, for all $x \in [0,1]$. One can easily verify that condition (4.11) is satisfied, where $i \in \text{argmax}\{\alpha_j : j \in N\}$. Hence, according to Proposition 4.3.6, it is Pareto optimal to redistribute all risk to firm i. Note that it might seem unfair that firm i takes over all risk. Firm i is, however, willing to bear all risk only if the side-payments it receives are sufficiently high. This problem is examined in Sections 4.4 and 4.5.

We proceed with the second example. There exists an $i \in N$ such that firm i is risk-neutral, i.e.,

$$\rho_i(Y) = E_{\mathbb{P}}[Y], \text{ for all } Y \in \mathbb{R}^\Omega. \tag{4.15}$$

This is a distortion risk measure with distortion function $g_i(x) = x$ for all $x \in [0, 1]$. For every firm $j \neq i$, we let the risk measure be given by a distortion risk measure ρ_j. Then, it holds that $g_j(x) \geq x = g_i(x)$ for all $j \in N$. Hence, it is Pareto optimal to redistribute all risk to firm i.

As a third example, we let every firm $j \in N$ use the risk measure Mean-Expected Shortfall with significance levels $\zeta \in [0, 1]$ and $\alpha_j \in (0, 1)$, i.e., firm j uses a weighted average of the mean and the α_j-Expected Shortfall as a risk measure. This risk measure is formally defined as

$$\rho^{MES}_{\zeta,\alpha_j}(Y) = \zeta E_{\mathbb{P}}[Y] + (1 - \zeta)\rho^{ES}_{\alpha_j}(Y), \text{ for all } Y \in \mathbb{R}^\Omega. \tag{4.16}$$

The (ζ, α_j)-Mean-Expected Shortfall is a distortion risk measure with distortion function $g_j(x) = \zeta x + (1 - \zeta) \min\left\{\frac{x}{\alpha_j}, 1\right\}$ for all $x \in [0, 1]$. One can again verify that condition (4.11) is satisfied, where $i \in \text{argmax}\{\alpha_j : j \in N\}$. This result is also shown by Asimit, Basescu and Tsanakas (2013). They show risk redistributions in the case where there are two firms and where the firms have possibly heterogeneous ζ_j and α_j. More generally, if a firm uses a (weighted) average of a distortion risk measure and the expectation as in, e.g., De Giorgi and Post (2008), its preferences can again be formalized via a distortion risk measure. If every firm $j \in N$ uses

$$\rho_j(Y) = \zeta_j E_{\mathbb{P}}[Y] + (1 - \zeta_j)\rho(Y), \text{ for all } Y \in \mathbb{R}^\Omega, \tag{4.17}$$

for some distortion risk measure ρ and $\zeta_j \in [0, 1]$, it is Pareto optimal to shift all risk to a firm $i \in \text{argmax}\{\zeta_j : j \in N\}$. $\triangledown$

4.3.3 Comonotonicity with the aggregate risk

In this subsection, we analyze comonotonicity of Pareto optimal risk redistributions with the aggregate risk. For distortion risk measures, comonotonicity of Pareto optimal risk redistributions is first studied by Ludkovski and Rüschendorf (2008). Their main result, which is based on Landsberger and Meilijson (1994), states that there exists a Pareto optimal risk redistribution such that all individual posterior risks are comonotone with the aggregate risk X. We provide a sufficient condition such that all Pareto optimal risk redistributions are comonotone with the aggregate risk X. We first define a risk measure ρ^*_N that plays a central role in obtaining Pareto optimal risk redistributions.

Definition 4.3.8 The function $g^*_N : [0, 1] \to [0, 1]$ of a risk redistribution problem $R \in \mathcal{RR}$ is given by $g^*_N(x) = \min\{g_i(x) : i \in N\}$ for all $x \in [0, 1]$. Moreover, ρ^*_N is the risk measure as defined in (4.1) with $g^{\rho^*_N} = g^*_N$.

In the following lemma we show that ρ_N^* is a distortion risk measure.

Lemma 4.3.9 *For all $R \in \mathcal{RR}$, the risk measure ρ_N^* is a distortion risk measure.*

Lemma 4.3.9 follows directly from the fact that the function g_N^* is continuous, concave, increasing and such that $g_N^*(0) = 0$ and $g_N^*(1) = 1$. Concavity of the function g_N^* follows from the fact that the minimum of concave functions is concave as well. We show in the sequel of this section that the risk measure ρ_N^* plays a central role in obtaining Pareto optimal risk redistributions.

Next, we provide a closed-form expression of a set of Pareto optimal risk redistributions. To do so, we first define the following set $M(R)$.

Definition 4.3.10 The set of functions $M(R)$ of a risk redistribution problem $R \in \mathcal{RR}$ is given by

$$M(R) = \left\{ m : \{1, \ldots, p-1\} \to N \,\middle|\, m(k) \in \operatorname*{argmin}_{j \in N} \{g_j\left(\mathbb{P}(\{\omega_1, \ldots, \omega_k\})\right)\} \text{ for all } k \in \{1, \ldots, p-1\} \right\}.$$

A function $m \in M(R)$ assigns to every $k \in \{1, \ldots, p-1\}$ a firm i for which the distortion function g_i is minimal at $\mathbb{P}(\{\omega_1, \ldots, \omega_k\})$, i.e., for all $k \in \{1, \ldots, p-1\}$ it holds that

$$g_{m(k)}(\mathbb{P}(\{\omega_1, \ldots, \omega_k\})) = g_N^*(\mathbb{P}(\{\omega_1, \ldots, \omega_k\})). \tag{4.18}$$

On the class of risk redistributions that are comonotone with the aggregate risk X, Ludkovski and Young (2009) characterize the Pareto optimal risk redistributions. The following theorem is based on this result.

Theorem 4.3.11 *For all $R \in \mathcal{RR}$, $m \in M(R)$ and $d \in \mathbb{R}^N$ with $\sum_{i \in N} d_i = X(\omega_p) = \min X$, it holds that $(\widetilde{X}_i)_{i \in N} \in \mathcal{PO}(R)$ where*

$$\widetilde{X}_i = \sum_{k=1}^{p-1} [X(\omega_k) - X(\omega_{k+1})] \cdot \mathbb{1}_{m(k)=i} \cdot e_{\{\omega_1, \ldots, \omega_k\}} + d_i \cdot e_\Omega, \text{ for all } i \in N, \tag{4.19}$$

and $\mathbb{1}_{m(k)=i} = 1$ if $m(k) = i$ and zero otherwise.

Note that (4.19) does not depend on the order on the state space $\Omega = \{\omega_1, \ldots, \omega_p\}$ such that $X(\omega_1) \geq \cdots \geq X(\omega_p)$. The Pareto optimal risk redistributions of the form (4.19) consist of a finite number of long and short positions on various stop-loss contracts (in fact, a portfolio of short positions on binary put options) on the aggregate risk X. Via the risks $d_i \cdot e_\Omega, i \in N$, the functional form (4.19) allows for all side-payments that ensure $(\widetilde{X}_i)_{i \in N} \in \mathcal{F}(R)$. The size of these side-payments is a central topic in the next sections.

In the following proposition, we characterize all Pareto optimal risk redistributions. This result provides the minimum aggregate risk adjusted value of the liabilities in the market after any Pareto optimal risk redistribution.

Proposition 4.3.12 *For all $R \in \mathcal{RR}$, we have $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$ if and only if $(\widetilde{X}_i)_{i\in N} \in \mathcal{F}(R)$ and*

$$\sum_{i\in N} \rho_i(\widetilde{X}_i) = \rho_N^*(X). \tag{4.20}$$

From Proposition 4.3.12 it follows that the Pareto optimal aggregate risk adjusted value of the liabilities depends on the risk measures $\rho_i, i \in N$, via ρ_N^* only. One can interpret this as that there exists a representative agent whose distortion risk measure equals ρ_N^*. The agent is representative in the sense that its preferences are sufficient to calculate the Pareto optimal aggregate risk adjusted value of the liabilities in the market. In contrast to actuarial equilibrium models with utility functions, the representative agent is a hypothetical firm with a risk measure that is least risk-averse in the market instead of an average (see, e.g., Bühlmann, 1980). We can derive that if $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$ then

$$\rho_i(\widetilde{X}_i) = \rho_N^*(\widetilde{X}_i), \text{ for all } i \in N. \tag{4.21}$$

We observe that every risk redistribution of the form (4.19) is comonotone with the aggregate risk X. If the following condition holds, we show that all Pareto optimal redistributions are comonotone with each other.

Condition [SC]: the function g_N^* is strictly concave, i.e.,

$$\lambda g_N^*(x)+(1-\lambda)g_N^*(y) < g_N^*(\lambda x+(1-\lambda)y), \text{ for all } \lambda \in (0,1) \text{ and } x, y \in [0,1] \text{ such that } x \neq y.$$

Lemma 4.3.13 *If $R \in \mathcal{RR}$ is such that condition [SC] holds, it holds for all $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$ that $\widetilde{X}_i(\omega_k) \geq \widetilde{X}_i(\omega_{k+1})$ for all $i \in N$ and all $k \in \{1, \dots, p-1\}$.*

From Lemma 4.3.13 it follows that if the function g_N^* is strictly concave, all Pareto optimal risk redistributions are comonotone with each other (and, so, with the aggregate risk X). Comonotone risks cannot be used as hedges for each other. The individual risks $X_i, i \in N$ are traded such that all posterior risks are weakly increasing with the systematic risk X.

If the function g_N^* is piecewise linear (e.g., if ρ_N^* equals α-Expected Shortfall), comonotonicity with the aggregate risk X is not guaranteed for Pareto optimal risk redistributions. However, if the distortion functions $g_i, i \in N$ are all strictly concave, the function g_N^* is strictly concave as well. Wirch and Hardy (2001) show that strict concavity of a distortion function is a necessary and sufficient condition for risk measures to strongly preserve second order stochastic dominance. From Lemma 4.3.13 we get directly the following result.

Corollary 4.3.14 *If $R \in \mathcal{RR}$ is such that condition [SC] holds and $k \in \{1, \dots, p-1\}$ is such that $X(\omega_k) = X(\omega_{k+1})$, it holds for all $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$ that $\widetilde{X}_i(\omega_k) = \widetilde{X}_i(\omega_{k+1})$ for all $i \in N$.*

Corollary 4.3.14 implies that when considering Pareto optimal risk redistributions, the states $\omega_k, \omega_{k+1} \in \Omega$ such that $X(\omega_k) = X(\omega_{k+1})$ are treated in the same manner.

4.3.4 Uniqueness up to side-payments

In this subsection, we analyze when there exists a unique Pareto optimal risk redistribution up to side-payments. This issue is relevant since if we know that there exists a unique Pareto optimal risk redistribution up to side-payments, the only question left is to determine the size of the side-payments.

From Proposition 4.3.4 and *Sub-additivity* and *Positive Homogeneity* of the risk measures it can be shown that the set of Pareto optimal risk redistributions is convex, i.e., for all $(\widetilde{X}_i)_{i\in N}$, $(\hat{X}_i)_{i\in N} \in \mathcal{PO}(R)$, it holds that $(\lambda\widetilde{X}_i + (1-\lambda)\hat{X}_i)_{i\in N} \in \mathcal{PO}(R)$ for all $\lambda \in [0,1]$. Therefore, the set of Pareto optimal risk redistributions can be large even up to side-payments. In this subsection, we identify two joint conditions under which the Pareto optimal risk redistributions are, up to side-payments, unique.

Uniqueness up to side-payments of Pareto optimal risk redistributions would hold if the risk measures are strictly concave (see Filipović and Svindland, 2008; Kiesel and Rüschendorf, 2009). Distortion risk measures are, however, not strictly concave.[35] To show uniqueness up to side-payments of Pareto optimal risk redistributions, we first introduce the following condition on $\mathcal{RR}$.

Condition [U]: for all $k \in \{1, \ldots, p-1\}$ such that $X(\omega_k) > X(\omega_{k+1})$ there exists a firm $i \in N$ such that for all $m \in M(R)$ it holds that $m(k) = i$.

If [SC] holds, we get from Lemma 4.3.13 that all Pareto optimal risk redistributions are comonotone with each other. From this combined with Theorem 2 of Ludkovski and Young (2009) we get directly the following result. We provide an alternative proof to avoid the introduction of the terminology in Ludkovski and Young (2009).

Theorem 4.3.15 *If $R \in \mathcal{RR}$ is such that condition [SC] holds, there exists a risk redistribution that is, up to side payments, the unique element of $\mathcal{PO}(R)$ if and only if condition [U] holds*

Note that $|M(R)| = 1$ is a sufficient condition for [U] to hold. Hence, if $|M(R)| = 1$ and condition [SC] holds, it follows from Theorem 4.3.15 that all Pareto optimal risk redistributions are uniquely determined up to side-payments. One can determine a Pareto optimal risk redistribution via Theorem 4.3.11. Condition [U] implies that all functions in $M(R)$ differ only on k such that $X(\omega_k) = X(\omega_{k+1})$.

[35] This can be seen from the properties *Comonotonic Additivity* and *Positive Homogeneity* of distortion risk measures as it holds that $\alpha\rho(X) + (1-\alpha)\rho(Y) = \rho(\alpha X + (1-\alpha)Y)$ for all comonotone $X, Y \in \mathbb{R}^\Omega$ such that $X \neq Y$ and $\alpha \in (0,1)$. Hence, distortion risk measures are not strictly concave.

Even if the function g_N^* is not strictly concave, it might be possible to define a strict concave distortion function such that it coincides with the function g_N^* on the relevant subdomain. This subdomain is the finite collection of probabilities $\{\mathbb{P}(\{\omega_1, \dots, \omega_k\}) : k = 1, \dots, p-1\}$. We illustrate this special case in the following example, where we also illustrate the construction of Pareto optimal risk redistributions.

Example 4.3.16 Let $N = \{1, 2\}$, $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $\mathbb{P}(\{\omega\}) = \frac{1}{3}$ for all $\omega \in \Omega$, $g_1(x) = \min\left\{1\frac{1}{2}x, 1\right\}$, $g_2(x) = \sqrt{x}$, $X(\omega_1) = 2$, $X(\omega_2) = 1$, $X(\omega_3) = 0$ and $X_1 = X_2 = \frac{1}{2}X$. So, we consider the case where all benefits from risk redistributions arise from the use of different risk measures. Firm 1 uses $\frac{2}{3}$-Expected Shortfall and Firm 2 uses a so-called proportional hazard distortion risk measure with parameter $\frac{1}{2}$. The function g_N^* is given by:

$$g_N^*(x) = \min\{g_i(x) : i \in N\} = \begin{cases} 1\frac{1}{2}x & \text{if } x \le \frac{4}{9}, \\ \sqrt{x} & \text{otherwise.} \end{cases}$$

The distortion functions g_1, g_2 and g_N^* are displayed in Figure 4.1. From this figure, we see

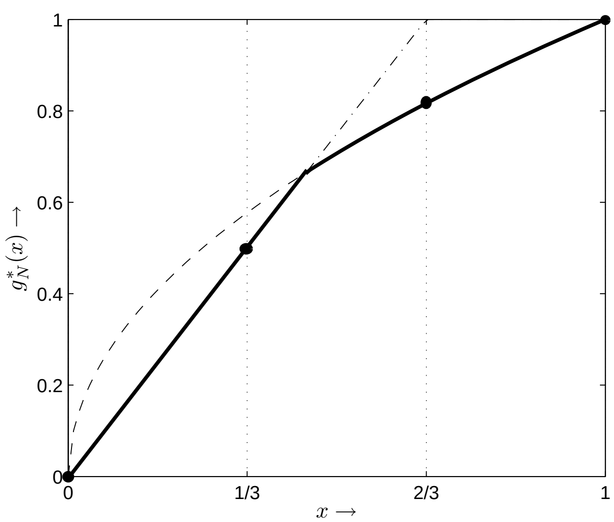


Figure 4.1: Construction of the function g_N^* via the distortion functions g_1 and g_2 corresponding to Example 4.3.16. The function g_1 is the dashed-dotted line, g_2 is the dashed line and g_N^* is the solid line.

that there is a unique $i \in \text{argmax}\{g_j(x) : j \in N\}$ that is minimal at $x = \mathbb{P}(\{\omega_1\}) = \frac{1}{3}$ and at $x = \mathbb{P}(\{\omega_1, \omega_2\}) = \frac{2}{3}$. Hence, it holds that $|M(R)| = 1$ and, therefore, condition [U] holds. Moreover, it holds that $m(1) = 1$ and $m(2) = 2$ for $m \in M(R)$. According to Theorem 4.3.11 with $c_1 = c_2 = 0$, a Pareto optimal risk redistribution is given by $\widetilde{X}_1$ and $\widetilde{X}_2$ such that $\widetilde{X}_1(\omega_1) = 1, \widetilde{X}_1(\omega_2) = 0, \widetilde{X}_1(\omega_3) = 0$, $\widetilde{X}_2(\omega_1) = 1$, $\widetilde{X}_2(\omega_2) = 1$ and $\widetilde{X}_2(\omega_3) = 0$. The construction of this Pareto optimal risk redistribution is shown in Figure 4.2.

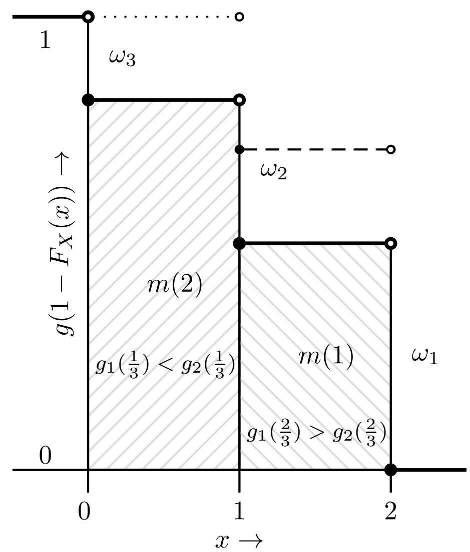


Figure 4.2: Graphical illustration of the construction of Pareto optimal risk redistributions corresponding to Example 4.3.16. The figure displays $g(1 - F_X(x))$ for g equal to g_1 (dotted line), g_2 (dashed line) or g_N^* (solid line). If lines coincide, the solid line is shown. From (4.1) with $\mathbb{P}(X < 0) = 0$ it follows that the area under g_N^* equals $\rho_N^*(X)$. The two shaded blocks are assigned to the firms $m(1) = 1$ and $m(2) = 2$; the risk $\widetilde{X}_1 = e_{\{\omega_1\}}$ is assigned to Firm 1 and the risk $\widetilde{X}_2 = e_{\{\omega_1,\omega_2\}}$ is assigned to Firm 2. The area of the shaded block with $m(i)$ in it represents $\rho_i(\widetilde{X}_i)$, $i = 1, 2$.

One can readily verify that $\rho_2(\widetilde{X}_2) > \rho_2(X_2)$ and, so, Firm 2 needs to receive a side-payment from Firm 1 in order to be willing to trade. We investigate the size of this side-payment in Sections 4.4 and 4.5.

In this example, the function g_N^* is not strictly concave. However, we can define a strictly concave distortion function that coincides with the function g_N^* on the subdomain $\left\{\frac{1}{3}, \frac{2}{3}\right\}$. Since also condition [U] holds, we get from Theorem 4.3.15 that there is a unique Pareto optimal risk redistribution up to side-payments.

The distortion functions g_1 and g_2 cross only once on the domain $(0, 1)$, i.e., $g_1(x) < g_2(x)$ for $x \in (0, x^*)$ and $g_1(x) > g_2(x)$ for $x \in (x^*, 1)$ with $x^* = \frac{4}{9}$. This implies that every Pareto optimal risk redistribution is a stop-loss contract or a deductible on the aggregate risk X and a side-payment; every $(\widetilde{X}_i)_{i \in N} \in \mathcal{PO}(R)$ is such that there exists a $c \in \mathbb{R}$ such that $\widetilde{X}_1 = (X-1)^+ + c \cdot e_\Omega$ and $\widetilde{X}_2 = \min\{X, 1\} - c \cdot e_\Omega$. $\triangledown$

4.3.5 Hedge benefits

In this subsection, we analyze the hedge benefits from risk redistributions. The aggregate hedge benefit of a risk redistribution $(\widetilde{X}_i)_{i\in N}$ is given by

$$\sum_{i\in N}\left[\rho_i(X_i)-\rho_i(\widetilde{X}_i)\right].$$

From Proposition 4.3.4 and Proposition 4.3.12 it follows that

$$\min\left\{\sum_{i\in N}\rho_i(\widetilde{X}_i):(\widetilde{X}_i)_{i\in N}\in\mathcal{F}(R)\right\}=\rho_N^*\left(X\right).$$

From this, it follows that the maximum aggregate hedge benefit over all feasible risk redistributions is given by

$$\max\left\{\sum_{i\in N}[\rho_i(X_i)-\rho_i(\widetilde{X}_i)]:(\widetilde{X}_i)_{i\in N}\in\mathcal{F}(R)\right\}=\sum_{i\in N}\rho_i(X_i)-\rho_N^*\left(X\right).$$

Since $(X_i)_{i\in N}\in\mathcal{F}(R)$, we get from Proposition 4.3.4 and Proposition 4.3.12 that $\sum_{i\in N}\rho_i(X_i)\geq\rho_N^*\left(X\right)$. If $\rho_N^*\left(X\right)=\sum_{i\in N}\rho_i(X_i)$ it follows that no firm can benefit from risk redistribution. Then, a Pareto optimal risk redistribution for firms is to keep their prior risk. The following proposition shows that the aggregate hedge benefits are zero only under two restrictive, joint conditions.

Proposition 4.3.17 *If $R\in\mathcal{RR}$ is such that the function g_N^* is strictly concave, it holds that $\rho_N^*\left(X\right)=\sum_{i\in N}\rho_i(X_i)$ if and only if the following two conditions hold:*

- *all $X_i, i\in N$ are comonotone with each other;*
- *$\rho_i(X_i)=\rho_N^*(X_i)$, for all $i\in N$.*

If the function g_N^* is strictly concave and there is no hedge potential, i.e., if $\rho_N^*\left(X\right)=\sum_{i\in N}\rho_i(X_i)$, all risks are comonotone according to Proposition 4.3.17. Moreover, $\rho_i(X_i)=\rho_N^*(X_i)$ for all $i\in N$. This implies that for all $i\in N$ we have

$$g_i(x)=g_N^*(x),\ \text{for all } x\in\{\mathbb{P}(\{\omega_1,\ldots,\omega_k\}):k\in\{1,\ldots,p-1\}\text{ s.t. }X_i(\omega_k)>X_i(\omega_{k+1})\}.\quad(4.22)$$

Hence, it holds that $\rho_N^*\left(X\right)=\sum_{i\in N}\rho_i(X_i)$ only under two strong conditions. If all firms are Von Neumann-Morgenstern expected utility maximizers, it is a necessary and sufficient condition that $u_i'(X_i)$ is proportional with $u_j'(X_j)$ for all $i,j\in N$ (see Borch, 1962).

We next show the hedge benefits in the risk redistribution problem of Example 4.3.16.

Example 4.3.18 For the risk redistribution problem in Example 4.3.16, it holds that $\rho_1(X_1)=0.75$, $\rho_2(X_2)\approx 0.70$ and $\rho_N^*(X)\approx 1.32$. So, the hedge benefits are equal to $0.32+0.35-0.70\approx$

0.05. Important to note is that even if risks are comonotone, firms can still benefit altogether from redistributing if the risk measures are not identical. This is in contrast to the case where all firms use the same risk measure. $\triangledown$

4.4 Competitive equilibria

In the previous section, we analyzed Pareto optimality of risk redistributions. Under the conditions [SC] and [U] in Theorem 4.3.15, there exists a unique Pareto optimal risk redistribution up to side-payments. In this section, we select risk redistributions as the competitive equilibria. In Subsection 4.4.1, we show that under one additional condition, there is a unique equilibrium risk redistribution. In Subsection 4.4.2, we derive a corresponding capital asset pricing model (CAPM).

4.4.1 Uniqueness of the competitive equilibrium

Competitive equilibria in insurance markets are studied by, e.g., Duffie and Zame (1989) and Aase (1993) for the case where firms use expected utility functions. Under three regularity conditions on the utility functions, Aase (1993) proves existence and uniqueness of the equilibrium. This result is inspired by Borch (1962), who shows this for special cases. For mean-variance investors, the competitive equilibrium corresponds with the classical CAPM equilibrium price of risk as derived by Sharpe (1964). Filipović and Kupper (2008), Dana and Le Van (2010), Dana (2011), and Flåm (2011) analyze existence of competitive equilibria in markets where firms use risk measures. Dana (2011) finds existence of a representative agent in the market where firms use strictly concave risk measures. If firms use distortion risk measures, we show in this subsection under which conditions market prices are comonotone with the aggregate risk X.

Chateauneuf, Dana and Tallon (2000) and Tsanakas and Christofides (2006) analyze equilibria in the case in which firms evaluate risk via $\rho_i(u_i(X_i))$,[36] where u_i is a strictly concave utility function and ρ_i a risk measure. Here, the role of risk measures is to include ambiguity aversion via max-min ambiguity-averse preferences. They assume strict concavity of the utility function as a sufficient condition to have uniqueness of the competitive equilibrium. In this subsection, we relax the assumption of a strictly concave utility function and obtain the competitive equilibria and corresponding capital asset pricing model.

Let there be a complete market. This implies existence of state prices, i.e., prices for the Arrow-Debreu assets $e_{\{\omega\}}$ with $\omega \in \Omega$. The pricing formula is given by $\pi(\hat{p}, Y) = \sum_{k=1}^{p} \hat{p}_k Y(\omega_k)$ for some price vector $\hat{p} \in \mathbb{R}_+^\Omega$. We assume that the risk-free rate is zero, i.e., $\pi(\hat{p}, e_\Omega) = 1$. A competitive equilibrium is a vector of prices $\hat{p} \in \mathbb{R}_+^\Omega$ and a risk redistribution $(\hat{X}_i)_{i\in N} \in (\mathbb{R}^\Omega)^N$ such that given the prices, each firm $i \in N$ individually minimizes $\rho_i(\hat{X}_i)$ under a budget

[36] For a wider class of risk measures ρ_i, this is also called Choquet-, max-min- or rank-dependent expected utility.

constraint, i.e., $\hat{X}_i$ solves

$$\min_{\widetilde{X}_i \in \mathbb{R}^\Omega} \rho_i(\widetilde{X}_i), \tag{4.23}$$

$$\text{s.t.} \quad \pi(\hat{p}, \widetilde{X}_i) \geq \pi(\hat{p}, X_i), \tag{4.24}$$

and the price vector $\hat{p}$ satisfies $\pi(\hat{p}, e_\Omega) = 1$ and induces market-clearing by equating aggregate supply and demand, i.e., $(\hat{X}_i)_{i \in N} \in \mathcal{F}(R)$. Competitive equilibria rely on the assumption that there is a competitive environment, where individual transactions have no influence on the prices. So, the number of firms needs to be large.

The following lemma is shown by Filipović and Kupper (2008). This result is an adjustment of the First Fundamental Welfare Theorem in case firms use distortion risk measures. Filipović and Kupper (2008) show this for the class of monetary utility functions. Monetary utility functions are functions satisfying *Concavity*, *Translation Invariance* and *Monotonicity*. Since it is well-known that risk measures satisfying *Sub-additivity* and *Positive Homogeneity* also satisfy *Concavity*, distortion risk measures are a subclass of monetary utility functions.

Lemma 4.4.1 *For all risk redistribution problems, there exists a competitive equilibrium. Moreover, every equilibrium risk redistribution is Pareto optimal.*

Next, we focus on calculating competitive equilibria. Filipović and Kupper (2008) and Flåm (2011) show for monetary utility functions that a necessary condition for prices to be equilibrium prices is that they are a subgradient[37] of the Pareto optimal aggregate risk adjusted value of the liabilities. We introduce the following condition, which assumes that there are no states in Ω where the realization of the aggregate risk X is the same.

condition [SO]: $X(\omega_1) > \cdots > X(\omega_p)$.

The following theorem shows that the conditions [SC] and [SO] are jointly sufficient to guarantee uniqueness of the equilibrium prices.

Theorem 4.4.2 *If $R \in \mathcal{RR}$ is such that condition [SC] holds, it holds that*

- *the equilibrium price vector $\hat{p}$ is unique if and only if condition [SO] holds;*
- *an equilibrium price vector is given by*

$$\hat{p}_k = g_N^*(\mathbb{P}(\{\omega_1, \ldots, \omega_k\})) - g_N^*(\mathbb{P}(\{\omega_1, \ldots, \omega_{k-1}\})), \text{ for all } k \in \{1, \ldots, p\}. \tag{4.25}$$

[37] A vector $\hat{p}$ is a subgradient of the risk measure ρ_N^* at the point X if for every $\widetilde{X} \in \mathbb{R}^\Omega$ the following inequality holds: $\rho_N^*(X) + \hat{p}(\widetilde{X} - X) \leq \rho_N^*(\widetilde{X})$.

If the function g_N^* is strictly concave and condition [SO] holds, it follows from Theorem 4.4.2 that $\hat{p} \in Q(\rho_N^*)$, where $Q(\rho_N^*)$ is defined in (4.7).[38] If, also, condition [U] holds, we get from Theorem 4.3.15 and Theorem 4.4.2 that the equilibrium risk redistribution is unique.

If [SC] does not hold, we can show that the vector $\hat{p}$ in (4.25) is still the unique equilibrium price vector if [SO] holds. However, the reversed statement does not necessarily hold true. We formally prove this result in Section 4.6.

If we relax the assumption [SO], it follows from Theorem 4.4.2 that the set of equilibrium prices is not single-valued. This does not necessarily imply that the corresponding equilibrium risk redistributions are non-unique. In Sections 4.5 and 4.6, we slightly relax this condition in order to provide a necessary condition for uniqueness of the equilibrium risk redistributions if [SC] holds.

4.4.2 Capital asset pricing model

We show in the following proposition that under the conditions [SC] and [SO], the risk capital of every equilibrium risk redistribution equals the price of the prior risk of the firms.

Proposition 4.4.3 *If $R \in \mathcal{RR}$ is such that conditions [SC] and [SO] hold, the unique equilibrium risk redistribution $(\hat{X}_i)_{i \in N}$ is such that*

$$\rho_i(\hat{X}_i) = E_{\mathbb{Q}_X}[X_i] = \sum_{k=1}^{p} [g_N^*(\mathbb{P}(\{\omega_1, \dots, \omega_k\})) - g_N^*(\mathbb{P}(\{\omega_1, \dots, \omega_{k-1}\}))] X_i(\omega_k), \tag{4.26}$$

for all $i \in N$.

One can interpret $\hat{p} = \mathbb{Q}_X$ as a *risk neutral* probability measure for obtaining the price of all X_i. The Radon-Nikodym derivative is given by

$$\frac{d\mathbb{Q}_X}{d\mathbb{P}}(\{\omega_k\}) = \frac{g_N^*(\mathbb{P}(\{\omega_1, \dots, \omega_k\})) - g_N^*(\mathbb{P}(\{\omega_1, \dots, \omega_{k-1}\}))}{\mathbb{P}(\{\omega_k\})}, \tag{4.27}$$

for all $k \in \{1, \dots, p\}$, which extends the Radon-Nikodym derivative used by Tsanakas (2004) and De Giorgi and Post (2008) for heterogeneous risk measures. So, the equilibrium prices are such that $\pi(\hat{p}, X_i) = E_{\mathbb{P}}\left[X_i \frac{d\mathbb{Q}_X}{d\mathbb{P}}\right]$ for all $i \in N$. This leads to

$$\pi(\hat{p}, X_i) = E_{\mathbb{P}}[X_i] + cov\left(X_i, \frac{d\mathbb{Q}_X}{d\mathbb{P}}\right), \text{ for all } i \in N, \tag{4.28}$$

where $\hat{p}$ is the unique equilibrium price vector. So, if the risk X_i is independent of the aggregate X, firm i only gets a risk where its risk adjusted value of the liabilities equals the expectation of its prior risk X_i. The stochastic variable $\frac{d\mathbb{Q}_X}{d\mathbb{P}}$ is comonotone with the risk X due to concavity

[38] Note that $\mathbb{Q}_X$ is a probability measure whereas $\hat{p}$ is a vector. Here, we mean that $\mathbb{Q}_X(\{\omega_k\}) = \hat{p}_k$ for all $k \in \{1, \dots, p\}$. In the sequel, we interpret $\hat{p}$ as a probability measure.

of the function g_N^*. Therefore, only co-movements with the market risk X are priced. Also, the equilibrium price depends on the aggregate risk X via the ordering on the state space Ω such that $X(\omega_1) \geq \cdots \geq X(\omega_p)$ only.

If the equilibrium prices are unique, it follows from (4.28) that for all $R \in \mathcal{RR}$ such that conditions [SC] and [SO] hold, we have

$$E_{\mathbb{P}}[RR_i] - 1 = \beta_i \left(E_{\mathbb{P}}[RR_m] - 1 \right), \tag{4.29}$$

where $i \in N$, $RR_i = \frac{X_i}{\pi(\hat{p}, X_i)}$ and $RR_m = \frac{X}{\pi(\hat{p}, X)}$ and

$$\beta_i = \frac{cov\left(RR_i, \frac{d\mathbb{Q}_X}{d\mathbb{P}}\right)}{cov\left(RR_m, \frac{d\mathbb{Q}_X}{d\mathbb{P}}\right)}. \tag{4.30}$$

The factor β_i in (4.29) is a market beta in a representation of the CAPM-model with distortion risk measures. Note that the risk-free rate is assumed to be zero. De Giorgi and Post (2008) empirically test the CAPM model for the case where all firms use the same distortion risk measure using US stock returns and find a better fit than the CAPM model with mean-variance investors.

Theorem 4.4.2 states the unique equilibrium prices under the conditions [SC] and [SO]. Moreover, from Theorem 4.3.15 it follows that under condition [U] these unique equilibrium prices correspond with a unique risk redistribution. In Section 4.5, we characterize this risk redistribution via game-theoretic properties. For competitive equilibria one assumes that there is a large number of firms and that there is competition implying that firms cannot influence the prices. If the set of firms is small, one cannot assume prices to be fully competitive and, therefore, we propose a game-theoretic characterization.

4.5 A cooperative game-theoretic approach for the risk redistribution problem

In this section, we discuss the case where the assumption that prices are unaffected by individual transactions is unrealistic. In absence of well-functioning markets, firms can meet each other and trade the risk. The risk redistribution is determined via a cooperative bargaining process. In this section, we propose a risk redistribution rule, the Aumann-Shapley value, that is the result of this cooperative bargaining. We show that this risk redistribution is unique under three conditions. It coincides with the equilibrium risk redistribution if it is unique. Moreover, we characterize the Aumann-Shapley value using four properties.

In Subsection 4.5.1 we introduce how a risk redistribution problem can be formulated as an allocation problem. In Subsection 4.5.2 we define some desirable properties of an allocation rule. Finally, Subsection 4.5.3 characterizes the Aumann-Shapley value as the unique rule satisfying these properties. If the Aumann-Shapley value exists, it corresponds with the competitive equilibrium.

4.5.1 From risk redistribution problems to capital allocation problems

In general, the aim is to find a risk redistribution $(\widetilde{X}_i)_{i\in N} \in \mathcal{F}(R)$ that is perceived as fair by the firms. We require risk redistributions to be Pareto optimal, i.e., $\sum_{i\in N} \rho_i(\widetilde{X}_i) = \rho_N^*(X)$ (see Proposition 4.3.12). In this subsection, we show how we can obtain a risk redistribution via a capital allocation.

Definition 4.5.1 An *allocation* is a vector $a \in \mathbb{R}^N$ such that $\sum_{i\in N} a_i = \rho_N^*(X)$.

To be consistent with the literature of risk capital allocation problems (see e.g. Denault, 2001), we denote in this section the risk adjusted value of the liabilities as risk capital. An allocation assigns to every firm risk capital.

In the next proposition, we show that an allocation corresponds with a Pareto optimal risk redistribution.

Proposition 4.5.2 *For every allocation $a \in \mathbb{R}^N$, there exists a risk redistribution $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$ such that $a_i = \rho_i(\widetilde{X}_i)$ for all $i \in N$. Under the conditions [SC] and [U], this risk redistribution is unique.*

If the conditions [SC] and [U] do not hold, we need to decide which risk redistribution to pick out of a non-empty collection of risk redistributions corresponding to an allocation. For every firm, the risk capital of all the risk redistributions corresponding to an allocation is the same. For every allocation, we pick a specific corresponding Pareto optimal risk redistribution. We denote this injective mapping by ϕ, i.e., ϕ maps allocations in $\{a \in \mathbb{R}^N : \sum_{i\in N} a_i = \rho_N^*(X)\}$ to risk redistributions in $\mathcal{PO}(R)$.

We next introduce allocation rules and risk redistribution rules. In Subsection 4.5.3, we focus on an allocation rule that is not always well-defined on $\mathcal{RR}$.

Definition 4.5.3 An *allocation rule* K maps every risk redistribution problem in $\widetilde{\mathcal{RR}} \subseteq \mathcal{RR}$ into a unique allocation in $\{a \in \mathbb{R}^N : \sum_{i\in N} a_i = \rho_N^*(X)\}$.

Definition 4.5.4 A *risk redistribution rule* ψ maps every risk redistribution problem $R \in \widetilde{\mathcal{RR}} \subseteq \mathcal{RR}$ into a risk redistribution in $\mathcal{PO}(R)$.

An allocation rule K corresponds with a risk redistribution rule ψ via the mapping ϕ, using $\psi = \phi \circ K$. A risk redistribution problem is mapped into an allocation via K. This allocation corresponds with a risk redistribution $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$ using the mapping ϕ. To summarize, we provide an overview in Figure 4.3.

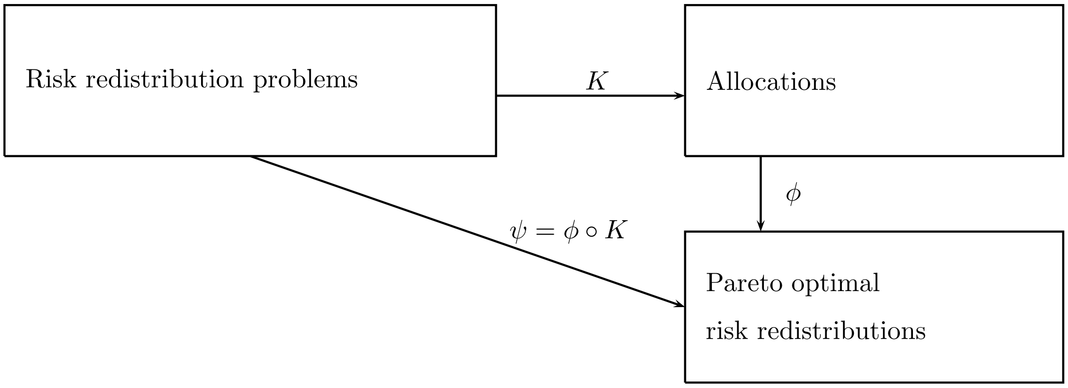


Figure 4.3: An overview of the correspondences between risk redistribution problems, allocations and risk redistributions. Every allocation rule K generates a risk redistribution rule ψ via ϕ.

From now on, we focus on risk capital allocations. There is an impressive amount of literature on allocation problems within the area of game theory. The main sources for this section are Aumann and Shapley (1974), Aubin (1979 and 1981), Billera and Heath (1982), and Mirman and Tauman (1982). These papers use the setting of a production problem with a given production function for multiple goods. The characterizations in these papers are formulated in terms of this production function, but are not directly transferrable in a meaningful way to our setting of allocating risk capital. Specifically for risk capital allocation problems with homogeneous risk measures, Denault (2001) uses a game-theoretic approach as well. The characterization that we provide in this section is similar to his characterization, but we formulate some properties in Subsection 4.5.2 in terms of the risk redistribution problem, while Denault (2001) focuses on a specific structure of cooperation.

4.5.2 Properties of allocation rules

In this subsection, we define five properties of allocation rules. We consider the following properties of an allocation rule $K : \widetilde{\mathcal{RR}} \to \mathbb{R}^N$ on a subclass of risk redistribution problems $\widetilde{\mathcal{RR}} \subseteq \mathcal{RR}$:

1. *Aggregation Invariance*: if for an $a > 0$ and $b \in \mathbb{R}^N$ it holds that $R = (X_j, \rho_j)_{j \in N} \in \widetilde{\mathcal{RR}}$, $\widehat{R} = (\widehat{X}_j, \rho_j)_{j \in N} \in \widetilde{\mathcal{RR}}$ and $\hat{X}_i = a \cdot X_i + b_i \cdot e_\Omega$ for all $i \in N$, then
$$K(R) = a \cdot K(\widehat{R}) + b.$$

2. *Monotonicity*: if $R \in \widetilde{\mathcal{RR}}$ is such that there exist firms $i, j \in N$ such that $g_i(x) \leq g_j(x)$ for all $x \in [0, 1]$ and $X_i(\omega) \leq X_j(\omega)$ for all $\omega \in \Omega$, then
$$K_i(R) \leq K_j(R).$$

3. *No Split-up*: if $R = (X_i, \rho_i)_{i\in N} \in \widetilde{\mathcal{RR}}$, $\widehat{R} = (\widehat{X}_i, \rho_i)_{i\in \widehat{N}} \in \widetilde{\mathcal{RR}}$ and $\ell \in N$ are such that $\widehat{N} = N \cup \{n+1\}$, $\widehat{X}_j = X_j$ for all $j \in N\backslash\{\ell\}$, $\widehat{X}_\ell + \widehat{X}_{n+1} = X_\ell$ and $\rho_\ell = \rho_{n+1}$, then

$$K_j(\widehat{R}) = K_j(R) \text{ for all } j \in N\backslash\{\ell\} \text{ and } K_\ell(\widehat{R}) + K_{n+1}(\widehat{R}) = K_\ell(R).$$

4. *Core Selection*: for all $R \in \widetilde{\mathcal{RR}}$, we have $K(R) \in core(R)$, where $core(R)$ denotes the *core* (Gillies, 1953) of a risk redistribution problem, which is defined as:

$$core(R) = \left\{ a \in \mathbb{R}^N : \sum_{i\in S} a_i \leq \rho_S^* \left(\sum_{i\in S} X_i \right) \text{ for all } S \subset N, \sum_{i\in N} a_i = \rho_N^*(X) \right\}. \quad (4.31)$$

Here, for $S \subseteq N$, ρ_S^* is the distortion risk measure with distortion function $g_S^*(x) = \min\{g_i(x) : i \in S\}$ for all $x \in [0,1]$.

The first three properties are based on, but weaker than, the properties in Billera and Heath (1982) and Mirman and Tauman (1982). Their properties depend only on a specific cost function in the problem to allocate production costs of a firm to the goods, whereas we define properties that are based on the potential to benefit from pooling risks from different firms.

The property *Aggregation Invariance* is necessary to make the allocation rule compatible with the use of risk measures. If, for instance, another currency is used, the relative allocation remains the same.

The property *Monotonicity* is a standard extension of the *Monotonicity* property of risk measures. If a firm holds a portfolio of which its realization is smaller in every state of the world than another firm and is endowed with a smaller distortion function than this other firm, its allocation should be lower than the allocation of this other firm. A smaller distortion function g_i leads to a smaller value of $\rho_i(\widetilde{X}_i)$ for the same risk $\widetilde{X}_i$ (see Proposition 4.3.6). Specifically, from interchanging the firms i and j in the definition of *Monotonicity*, it follows that allocation rules satisfying *Monotonicity* also satisfy the following property:

Symmetry: if $R \in \widetilde{\mathcal{RR}}$ is such that there exist firms $i, j \in N$ where $X_i = X_j$ and $\rho_i = \rho_j$, then

$$K_i(R) = K_j(R).$$

The property *No split-up* implies that firms do not have any incentive to split the firm into two or more parts. Vice versa, firms do not have any incentive to merge with another firm in order to obtain a strict improvement for both firms in the risk redistribution. If there is only one firm, it follows from *Sub-additivity* of the risk measures that this firm is not willing to split-up in multiple firms. In this sense, the set of firms is stable against merging and splitting. Also the allocation to the other firms is independent of whether a firm splitted. This property corresponds directly with the property *Consistency* in Mirman and Tauman (1982).

The property *Core Selection* implies that there does not exist a subgroup of firms that can strictly benefit altogether by splitting off and redistributing risk with only the firms in this subgroup. This property is widely discussed in the game-theoretic literature (see, e.g., Gillies, 1953). The conditions in (4.31) include all individual rationality conditions, i.e., $K_i(R) \leq \rho_i(X_i)$ for all $i \in N$. *Core Selection* implies that an allocation is in the core of the following *cooperative cost game*:[39]

$$c(S) = \rho_S^* \left(\sum_{i \in S} X_i \right) \tag{4.32}$$

$$= \min \left\{ \sum_{i \in S} \rho_i \left(\widetilde{X}_i \right) : (\widetilde{X}_i)_{i \in S} \in \mathcal{F}(R_S) \right\}, \tag{4.33}$$

for all $S \subseteq N$, where R_S is the risk redistribution problem $(X_i, \rho_i)_{i \in S}$. The equality (4.33) can be derived from Proposition 4.3.4 and Proposition 4.3.12, where we replace the set N by S. For a discussion about this cooperative game, we refer to Appendix 4.B.

In (4.32)-(4.33) we show how the capital allocation problem can be seen as a special case of a cooperative game. This suggests that likely candidates for solving the capital allocation problem can be solution concepts proposed in game theory for cooperative games. We focus on solution concepts that assign to the cooperative game (N, c) a vector $a \in \mathbb{R}^N$ such that $\sum_{i \in N} a_i = c(N) = \rho_N^*(X)$. Such a solution concept of cooperative games that received considerable attention is the Shapley value (Shapley, 1953). The corresponding allocation rule does not satisfy all properties defined in this subsection. As discussed by Denault (2001) and Csóka and Pintér (2011), the Shapley value need not be in the core of cooperative cost games where the firms use homogeneous risk measures. Therefore, it does not satisfy *Core Selection*. Also the allocation rules corresponding to other well-known solution concepts such as the Compromise value (Tijs, 1981) and Nucleolus (Schmeidler, 1969) do not satisfy all properties defined in this subsection. The Compromise value does not satisfy *Core Selection* and *No Split-up*, while the Nucleolus does not satisfy *No Split-up*. Later, we show that there exists an allocation rule that satisfies the four properties defined in this subsection. For this reason, we focus on other allocation rules.

The four properties do not necessarily characterize a unique allocation rule. As a solution, we adjust the property *Core Selection*. The more stringent property that we consider is first introduced in the seminal works of Aubin (1979 and 1981) and later imposed by Denault (2001) in the context of risk capital allocation problems with homogeneous risk measures. Let every firm consist of infinitesimally small identical portfolios that can cooperate as a separate party. Before and after risk redistribution, it follows from *Positive Homogeneity* and *Comonotonic Additivity* of ρ_i, $i \in N$, that the total risk capital of a firm is equal to the aggregate risk capital of its

[39] Cooperative cost games consist of the set of firms N and a characteristic function $c : 2^N \to \mathbb{R}$. In the context of the risk redistribution problem considered in this chapter, the characteristic function yields for any subset $S \subset N$ the minimal aggregate risk capital if only the firms in S decide to redistribute.

portfolios. We focus on the following criterion on risk redistributions $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$:

$$\sum_{i\in N} \rho_i(\lambda_i \widetilde{X}_i) \leq \min \left\{ \sum_{i\in N:\lambda_i>0} \rho_i(\hat{X}_i) : \sum_{i\in N:\lambda_i>0} \hat{X}_i = \sum_{i\in N} \lambda_i X_i \right\}, \tag{4.34}$$

for all $\lambda \in [0,1]^N$. The conditions in (4.34) include all individual rationality conditions, because $\lambda = e_i$ yields $\rho_i(\widetilde{X}_i) \leq \rho_i(X_i)$ for all $i \in N$. In general, (4.34) leads to the following property of an allocation rule $K : \widetilde{\mathcal{RR}} \to \mathbb{R}^N$:

5. *Fuzzy Core Selection*: for all $R \in \widetilde{\mathcal{RR}}$, we have $K(R) \in Fcore(R)$, where $Fcore(R)$ denotes the *fuzzy core* (Aubin, 1979) of a risk redistribution problem $R \in \mathcal{RR}$, which is defined as:

$$Fcore(R) = \left\{ a \in \mathbb{R}^N : \sum_{i\in N} \lambda_i a_i \leq r(\lambda) \text{ for all } \lambda \in [0,1]^N, \sum_{i\in N} a_i = r\left(e_N\right) \right\} \tag{4.35}$$

where the fuzzy game $r : \mathbb{R}^N_+ \to \mathbb{R}$ is given by:[40]

$$r(\lambda) = \min \left\{ \sum_{i\in N:\lambda_i>0} \rho_i(\hat{X}_i) : \sum_{i\in N:\lambda_i>0} \hat{X}_i = \sum_{i\in N} \lambda_i X_i \right\}, \text{ for all } \lambda \in \mathbb{R}^N_+. \tag{4.36}$$

We can derive from (4.32) and (4.36) that

$$r(\lambda) = \rho^*_{\{i\in N:\lambda_i>0\}} \left(\sum_{i\in N} \lambda_i X_i \right), \text{ for all } \lambda \in \mathbb{R}^N_+. \tag{4.37}$$

Since $r(e_S) = c(S)$ for all $S \subseteq N$ with c as in (4.32), we get that the fuzzy core is a subset of the core. Hence, every allocation rule satisfying *Fuzzy Core Selection* satisfies *Core Selection*. In the next subsection we characterize a specific allocation rule as the allocation satisfying all properties defined in this subsection.

4.5.3 The Aumann-Shapley value

In this subsection, we introduce the Aumann-Shapley value. We use this allocation rule to obtain a specific allocation, and thus a specific risk redistribution. This risk redistribution coincides with the competitive equilibrium if this equilibrium is unique. Moreover, we characterize this rule via the properties that are defined in the previous subsection.

Let $\mathcal{RR}' \subset \mathcal{RR}$ be the set of all risk redistribution problems for which the fuzzy game

[40] In fuzzy games, as in Aubin (1979 and 1981) and Chapter 2 of this dissertation, firms are able to cooperate partially via fractions between zero and one. Every combination of fractional firms can be seen as a portfolio. One can see the portfolio of a firm as a collection of infinitesimally small identical risks. A fuzzy game is given by a mapping $r : \mathbb{R}^N_+ \to \mathbb{R}$ that is normalized such that full participation of a firm $i \in N$ corresponds with $\lambda_i = 1$. Moreover, it satisfies $r(e_S) = c(S)$ for all $S \subseteq N$, where (N, c) is as defined in (4.32) and e_S is the vector with ones for firms in S and zeros for firms in $N \backslash S$.

$r : \mathbb{R}^N_+ \rightarrow \mathbb{R}$, defined in (4.36), is partially differentiable at $\lambda = e_N$. For fuzzy games, there is one prominent allocation rule, namely the Aumann-Shapley value. This concept is first introduced by Aumann and Shapley (1974) for games with a continuum of players. It is formally defined for risk redistribution problems as follows.

Definition 4.5.5 The *Aumann-Shapley value* for risk redistribution problems, denoted by $AS : \mathcal{RR}' \rightarrow \mathbb{R}^N$, is given by

$$AS_i(R) = \frac{\partial r}{\partial \lambda_i}(e_N), \text{ for all } i \in N, \tag{4.38}$$

where the fuzzy game r is defined in (4.36).

Also for risk capital allocation problems with homogeneous risk measures, the Aumann-Shapley value received considerable attention in the literature.[41]

Next, we state a necessary and sufficient property to guaranty uniqueness of the Aumann-Shapley value if [SC] is satisfied. We first introduce equivalent states in the following definition.

Definition 4.5.6 Two states $\omega, \omega' \in \Omega$ are *equivalent* if $X_i(\omega) = X_i(\omega')$ for all $i \in N$.

If [SC] holds, the following result shows that $R \in \mathcal{RR}'$ if and only if $X(\omega) = X(\omega')$ for equivalent states $\omega, \omega' \in \Omega$ only.

Proposition 4.5.7 *For all $R \in \mathcal{RR}$ such that [SC] holds, it holds that $X(\omega) = X(\omega')$ for equivalent states $\omega, \omega' \in \Omega$ and $X(\omega) \neq X(\omega')$ otherwise if and only if $R \in \mathcal{RR}'$.*

The condition [SO] is sufficient to have no equivalent states. Therefore, it follows from Proposition 4.5.7 that if [SC] holds, [SO] is a sufficient condition for the Aumann-Shapley value to exist. All other instances where the Aumann-Shapley value exists can be neglected since, without loss of generality, we can reformulate the risk redistribution problem such that there are no equivalent states.

Next, we focus on some properties of the Aumann-Shapley value. In the following proposition, which is derived from Mirman and Tauman (1982) in the context of fuzzy games, we show that the Aumann-Shapley value is an allocation rule satisfying the first three properties that are defined in Subsection 4.5.2.

Proposition 4.5.8 *The Aumann-Shapley value satisfies the properties Aggregation Invariance, Monotonicity and No Split-up on $\mathcal{RR}'$.*

Proposition 4.5.8 follows from Mirman and Tauman (1982), since allocation rules satisfying:

[41] For risk capital allocation problems with homogeneous risk measures, there is a wide range of game-theoretic (see, e.g., Denault, 2001; Tsanakas and Barnett, 2003; Kalkbrener, 2005), financial (see, e.g., Tasche, 1999) and economic (see, e.g., Myers and Read, 2001) approaches described in the literature.

- *Additivity* and *Positivity* (see Mirman and Tauman, 1982) satisfy *Aggregation Invariance*;
- *Additivity* (see Mirman and Tauman, 1982) satisfy *Monotonicity*;
- *Consistency* (see Mirman and Tauman, 1982) satisfy *No Split-up*.

Next, we continue with a characterization of the Aumann-Shapley value that is based on Aubin (1979) and focuses on the fuzzy core. To do so, we first show that the fuzzy game r is sub-additive, which implies that r is convex due to positive homogeneity of this fuzzy game.

Lemma 4.5.9 *For all $R \in \mathcal{RR}$, the fuzzy game r is sub-additive, i.e., $r(\lambda) + r(\lambda') \geq r(\lambda + \lambda')$ for all $\lambda, \lambda' \in \mathbb{R}_+^N$.*

Lemma 4.5.9 shows that the fuzzy game r is sub-additive. The following result, which follows from Aubin (1979) and Lemma 4.5.9, shows that the fuzzy core is single-valued on the subclass $\mathcal{RR}'$.

Theorem 4.5.10 *For all $R \in \mathcal{RR}'$, it holds that $AS(R) \in Fcore(R)$ and $Fcore(R)$ is single-valued.*

Aubin (1979) shows this result for sub-additive and positive homogeneous fuzzy games. Theorem 4.5.10 shows that the Aumann-Shapley value, if existent, is an element of the single-valued fuzzy core. In Section 4.6, we extend this result by showing that the fuzzy core is equal to the set of allocations corresponding to the competitive equilibria. Moreover, in Section 4.6, we provide a closed-form expression of the fuzzy core also in case the Aumann-Shapley value does not exist. The following corollary follows directly from Theorem 4.5.10.

Corollary 4.5.11 *The Aumann-Shapley value is the unique allocation rule satisfying Fuzzy Core Selection on $\mathcal{RR}'$.*

Next, we provide a closed-form expression of the Aumann-Shapley value for risk redistribution problems. Tsanakas and Barnett (2003) obtain a closed-form expression of the Aumann-Shapley value in case all firms use the same risk measure and under the assumptions that the probability density function is continuous and the distortion function g is twice differentiable. Then, it holds that $AS_i(R) = E_{\mathbb{P}}[X_i g'(1 - F_X(X))]$ for all $i \in N$. In the following proposition, we show the Aumann-Shapley value for the case of heterogeneous distortion risk measures and discrete risks.

Proposition 4.5.12 *For all $R \in \mathcal{RR}'$, the Aumann-Shapley value is given by:*

$$AS_i(R) = \sum_{k=1}^{p} [g_N^*(\mathbb{P}(\{\omega_1, \ldots, \omega_k\})) - g_N^*(\mathbb{P}(\{\omega_1, \ldots, \omega_{k-1}\}))] X_i(\omega_k), \text{ for all } i \in N. \quad (4.39)$$

For all $R \in \mathcal{RR}'$ such that [SC] holds, we get from Theorem 4.4.2 and Proposition 4.5.12 that

$$AS_i(R) = E_{\mathbb{Q}_X}[X_i] = \pi(\hat{p}, X_i), \text{ for all } i \in N, \tag{4.40}$$

where $\mathbb{Q}_X$ is the probability measure in (4.4) for the risk measure $\rho = \rho_N^*$, and $\hat{p}$ is an equilibrium price vector as defined in (4.25). The Aumann-Shapley value is equal to the allocation corresponding to the competitive equilibrium if the equilibrium is unique, which holds under [SC] and [SO] (see Theorem 4.4.2).

Next, we return to the problem in Example 4.3.16 and compute the Aumann-Shapley value and its corresponding risk redistribution for this case.

Example 4.5.13 In this example, we provide the Aumann-Shapley value for the risk redistribution problem in Example 4.3.16. This is given by

$$AS_1(R) = AS_2(R) = \frac{1}{\sqrt{6}} + \frac{1}{4} \approx 0.66. \tag{4.41}$$

This implies that for the risk redistribution $(\widetilde{X}_i)_{i \in N}$ provided in Example 4.3.16, there is a side-payment of size $\frac{1}{\sqrt{6}} - \frac{1}{4} \approx 0.16$ made by Firm 1 to Firm 2. Then, the corresponding risk redistribution is given by $\hat{X}_1$ and $\hat{X}_2$ such that $\hat{X}_1(\omega_1) \approx 1.16, \hat{X}_1(\omega_2) \approx 0.16, \hat{X}_1(\omega_3) \approx 0.16$, $\hat{X}_2(\omega_1) \approx 0.84$, $\hat{X}_2(\omega_2) \approx 0.84$, and $\hat{X}_2(\omega_3) \approx -0.16$. $\triangledown$

4.6 Competitive equilibria and fuzzy core for risk redistribution problems

In this section, we provide some final comments on the results shown in Section 4.4, where we define conditions for uniqueness of the competitive equilibria. In this section, we discuss some game-theoretic results for the competitive equilibria. Aumann (1964) and Aubin (1981) show that the fuzzy core of risk redistribution problems is equivalent to the set of allocations corresponding to the competitive equilibria. This result is originally shown by Aumann (1964) for games with a continuum of players. Aubin (1981) extends this result in the context of positive homogeneous fuzzy games.

Theorem 4.6.1 (Aumann (1964) and Aubin (1981)) *The fuzzy core of the game r in* (4.36) *equals the set of allocations corresponding to the competitive equilibria, as defined in Section 4.4.*

Theorem 4.6.1 implies that we can obtain all equilibria from elements of the fuzzy core. The corresponding equilibrium prices are given by $\hat{p} \in Q(\rho_N^*)$ such that $\sum_{k=1}^{p} \hat{p}_k X_i(\omega_k) = a_i$ for all $i \in N$. From Theorem 4.6.1, it also follows that the allocation corresponding to the equilibria is unique if and only if the fuzzy core is single-valued. From this, Theorem 4.3.15, Proposition 4.5.7 and Theorem 4.5.10 we directly get the following result.

Theorem 4.6.2 *For all $R \in \mathcal{RR}$ such that [SC] holds, the equilibrium risk redistribution is unique if and only if the following two conditions hold:*

- *for all $k \in \{1, \ldots, p-1\}$ such that the states ω_k and ω_{k+1} are not equivalent, there exists a firm $i \in N$ such that for all $m \in M(R)$ it holds that $m(k) = i$;*[42]
- *$X(\omega) = X(\omega')$ for equivalent states $\omega, \omega' \in \Omega$ only.*

If we formulate a risk redistribution problem without equivalent states, Theorem 4.6.2 strengthens Theorem 4.3.15 and Theorem 4.4.2 by showing that if [SC] holds, the properties [U] and [SO] are necessary for uniqueness of the equilibrium risk redistributions. If firms use expected utilities, Aase (1993) shows that under two regularity conditions on utility functions only, there is existence of the equilibria. Existence of the equilibria in the risk redistribution problem is guaranteed. If firms use expected utilities, there is a third regularity condition necessary to ensure uniqueness of the equilibrium risk redistribution. All three regularity conditions in Aase (1993) are imposed on the utility functions only. Theorem 4.6.2 implies that if firms use distortion risk measures, it is sufficient to impose three conditions to ensure uniqueness of the equilibrium. Two conditions also depend on the aggregate risk.

Theorem 4.6.1 states the equivalence of the fuzzy core and the allocations corresponding to the competitive equilibria. We can find all equilibria from the fuzzy core even when the function g_N^* is not strictly concave or when the fuzzy game r is not partially differentiable at $\lambda = e_N$. The following representation of the fuzzy core follows from Aubin (1979), and is analogue to Theorem 2.3.5 in Chapter 2 of this dissertation.

Theorem 4.6.3 *The fuzzy core is given by*

$$Fcore(R) = \{(E_{\mathbb{Q}}[X_i])_{i \in N} : \mathbb{Q} \in Q^*\}, \text{ for all } R \in \mathcal{RR}, \tag{4.42}$$

where the set Q^* is given by

$$Q^* = \{\mathbb{Q} \in Q(\rho_N^*) : \rho_N^*(X) = E_{\mathbb{Q}}[X]\}. \tag{4.43}$$

Since $Q(\rho_N^*)$ is a convex polytope and, so, compact, it holds that $Q^* \neq \emptyset$. From Theorem 4.6.1 and Theorem 4.6.3 we can derive that the set of all equilibrium prices is given by Q^*.[43] So, all equilibrium prices are comonotone with the aggregate risk X. This leads to the following corollary, which extends Theorem 4.4.2 to the case where the function g_N^* is not strictly concave.

Corollary 4.6.4 *For all $R \in \mathcal{RR}$ such that [SO] holds, the equilibrium price vector $\hat{p}$ is unique. This price vector $\hat{p}$ is defined as in* (4.25).

[42] Alternative formulation: for all $m, m' \in M(R)$ such that $m \neq m'$, it holds that $m(k) = m'(k)$ only if $k \in \{1, \ldots, p-1\}$ is such that the states ω_k and ω_{k+1} are equivalent.

[43] Recall from footnote 38 how we can formulate the equilibrium price vectors from probability measures.

If the fuzzy core is not single-valued, one can select an element based on, e.g., Mertens (1988) and Chapter 2 of this dissertation. They derive the same generalization of the Aumann-Shapley value which is well-defined also in case of non-differentiability of the fuzzy game r in $\lambda = e_N$. Mertens (1988) uses an axiomatic approach whereas we use an asymptotic argument in Chapter 2. This allocation rule is given by a convex combination of "nearby" Aumann-Shapley values $(E_{\mathbb{Q}}[X_i])_{i\in N}$ for $\mathbb{Q} \in Q^*$.

4.7 Future research

In this section, we briefly provide two different directions for future research.

On the individual level, firms might have an incentive to misrepresent their risk measure for the competitive equilibria and, thus, the Aumann-Shapley value. So, both for well-functioning as not well-functioning markets, firms can benefit from misrepresenting their preferences. For instance, if the individual risk X_i of firm $i \in N$ is counter-comonotone with the aggregate risk X and if firm i pretends to have a larger distortion function $\hat{g}_i(x) \geq g_i(x)$ for all $x \in [0,1]$, it holds for any equilibrium risk $\widetilde{X}_i$ that

$$\begin{aligned} \rho_i(\widetilde{X}_i) &= E_{\mathbb{Q}_X}[X_i] && (4.44)\\ &\geq E_{\hat{\mathbb{Q}}_X}[X_i] && (4.45)\\ &= \hat{\rho}_i(\hat{X}_i) && (4.46)\\ &\geq \rho_i(\hat{X}_i), && (4.47) \end{aligned}$$

where ρ_i and $\mathbb{Q}_X$ correspond with the case of no misrepresenting of firm i, and $\hat{X}_i$, $\hat{\rho}_i$ and $\hat{\mathbb{Q}}_X$ correspond with the case of misrepresenting by firm i. Here, (4.44) and (4.46) follow from Proposition 4.5.12, (4.45) follows from the fact that the risks X_i and X are counter-comonotone and (4.47) follows from Proposition 4.3.6. It is not difficult to construct examples where the inequality is strict. Hence, the risk measures $\rho_i, i \in N$ must be observable, whereas the individual risks $X_i, i \in N$ do not. This issue is well-known if firms use expected utilities as in Borch (1962). A direction for future research is to design a risk redistribution that is incentive compatible with respect to representations of the risk measure.

Next, we continue with the second suggestion for future research. If there are firms using an expected utility function (indexed by the set N_1) and there are firms using a distortion risk measure (indexed by the set N_2), we can model the cooperative game as follows. To obtain all Pareto optimal risk redistributions, we need to solve for all $k \in \mathbb{R}^{N_1}_{++}$ the following system:

$$\begin{aligned} &\min -\sum_{i\in N_1} k_i E_{\mathbb{P}}[u_i(-\widetilde{X}_i)] + \sum_{i\in N_2} \rho_i(\widetilde{X}_i) = \min -\sum_{i\in N_1} k_i E_{\mathbb{P}}[u_i(-\widetilde{X}_i)] + \rho^*_{N_2}\left(\sum_{i\in N_2}\widetilde{X}_i\right)\\ &\text{s.t.} \sum_{i\in N_1\cup N_2}\widetilde{X}_i = \sum_{i\in N_1\cup N_2} X_i. \end{aligned}$$

If $g^*_{N_2}$ is strictly concave, it follows from Lemma 4.3.13 that $\widetilde{X}_i, i \in N_2$ are all comonotone with each other. The corresponding cooperative game is a Non-Transferable Utility game. It is still an open question what the distributions are of the competitive equilibria.

4.8 Conclusion

This chapter provides a rule to redistribute risk if the firms use a distortion risk measure to evaluate risk. To do so, it shows how all Pareto optimal risk redistributions are constructed. Throughout this chapter, we introduce the following four conditions for a risk redistribution problem $R \in \mathcal{RR}$:

Condition [SC]: the function g^*_N is strictly concave;

Condition [U]: for all $k \in \{1, \ldots, p-1\}$ such that $X(\omega_k) > X(\omega_{k+1})$ there exists a firm $i \in N$ such that for all $m \in M(R)$ it holds that $m(k) = i$, where $M(R)$ is defined in Definition 4.3.10;

Condition [SO]: $X(\omega_1) > \cdots > X(\omega_p)$;

Condition [SOw]: $X(\omega_k) = X(\omega_{k+1})$ for equivalent states $\omega_k, \omega_{k+1} \in \Omega$ only, where equivalent states are defined in Definition 4.5.6.

The conditions [SC] and [U] are sufficient for Pareto optimal risk redistributions to be unique up to side-payments.

The next problem is to determine the size of the side-payments. We determine the side-payments first via the competitive equilibria and, thereafter, via a game-theoretic characterization. Based on Pareto optimality, we obtain under the Condition [SO] that the equilibrium prices are unique. These prices can be seen as risk-neutral probabilities, where there is a representative agent using the minimal distortion function. As we are mainly interested in risk redistributions and not the corresponding prices, we can relax Condition [SO] to Condition [SOw]. in the sense that the equilibrium risk redistribution is unique under the conditions [SC], [U] and [SOw]. We obtain the corresponding risk redistribution from calculating first the posterior risk adjusted value of the liabilities via the equilibrium prices and, thereafter, determining the corresponding unique risk redistribution.

Thereafter, we characterize this risk redistribution as the one corresponding to the unique allocation satisfying the stability criterion *Fuzzy Core Element*. The basic idea is that firms have the option of participating in a risk redistribution problem by using only a fraction of their initial risk. In cooperative game theory, this risk redistribution rule corresponds with the Aumann-Shapley value, which is the gradient of an appropriately chosen fuzzy game.

We show that even if risks are identical, there can still be benefits from redistributing if the risk measures are heterogeneous.

4.A Proofs

Proof of Proposition 4.3.5: Let $R \in \mathcal{RR}$ be such that $\rho_i = \rho$ for all $i \in N$ and let $(\widetilde{X}_i)_{i\in N} \in \mathcal{F}(R)$ be such that all $\widetilde{X}_i, i \in N$, are comonotone with each other. From *Comonotonic Additivity* of ρ it follows that

$$\sum_{i\in N} \rho\left(\widetilde{X}_i\right) = \rho\left(\sum_{i\in N} \widetilde{X}_i\right) = \rho\left(X\right). \tag{4.48}$$

Generally, from *Sub-additivity* of ρ it follows that

$$\sum_{i\in N} \rho(\hat{X}_i) \geq \rho\left(X\right), \text{ for all } (\hat{X}_i)_{i\in N} \in \mathcal{F}(R). \tag{4.49}$$

Hence, from (4.48) and (4.49) it follows that $(\widetilde{X}_i)_{i\in N} \in \mathcal{F}(R)$ minimizes $\sum_{i\in N} \rho(\hat{X}_i)$ over all $(\hat{X}_i)_{i\in N} \in \mathcal{F}(R)$. Therefore, it follows from Proposition 4.3.4 that $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$. This concludes the proof. □

Proof of Proposition 4.3.6: Let $R \in \mathcal{RR}$ be such that there exists a firm $i \in N$ for which (4.11) holds. One can easily verify that from (4.2) it follows that

$$\rho_i(Y) \leq \rho_j(Y), \text{ for all } Y \in \mathbb{R}^\Omega \text{ and } j \in N. \tag{4.50}$$

From this it follows directly that

$$\begin{aligned}\rho_i(X) &\leq \sum_{j\in N} \rho_i(\hat{X}_j) &(4.51)\\ &\leq \sum_{j\in N} \rho_j(\hat{X}_j), &(4.52)\end{aligned}$$

for all $(\hat{X}_j)_{j\in N} \in \mathcal{F}(R)$, where (4.51) follows from *Sub-additivity* of ρ_i and (4.52) follows from (4.50). Hence, combining (4.51)-(4.52) with $\rho_j(0 \cdot e_\Omega) = 0$ for all $j \in N\backslash\{i\}$ yields that $(\widetilde{X}_j)_{j\in N}$ minimizes $\sum_{i\in N} \rho(\check{X}_i)$ over all $(\check{X}_i)_{i\in N} \in \mathcal{F}(R)$. From this and Proposition 4.3.4 it follows that $(\widetilde{X}_j)_{j\in N} \in \mathcal{PO}(R)$. This concludes the first part of the proof.

Next, let $R \in \mathcal{RR}$ be such that there exists a firm $i \in N$ for which (4.12) holds. Suppose that there exists an $(\hat{X}_j)_{j\in N} \in \mathcal{PO}(R)$ such that for at least one firm $j \neq i$ its risk $\hat{X}_j$ is not a side-payment. For every risk $Y \in \mathbb{R}^\Omega$ that is not a side-payment, it holds that

$$\begin{aligned}\rho_i(Y) &= \sum_{k=1}^{p-1} g_i(\mathbb{P}(\{\omega_1, \dots, \omega_k\}))[Y(\omega_k) - Y(\omega_{k+1})] + Y(\omega_p) && (4.53)\\ &< \sum_{k=1}^{p-1} g_j(\mathbb{P}(\{\omega_1, \dots, \omega_k\}))[Y(\omega_k) - Y(\omega_{k+1})] + Y(\omega_p) && (4.54)\\ &= \rho_j(Y), && (4.55)\end{aligned}$$

for all $j \neq i$, where $Y(\omega_1) \geq \dots \geq Y(\omega_p)$. Here, (4.54) follows from (4.12) and that there exists a $k \in \{1, \dots, p-1\}$ such that $Y(\omega_k) - Y(\omega_{k+1}) > 0$ and $\mathbb{P}(\{\omega_1, \dots, \omega_k\}) \in (0, 1)$. From this it follows directly that

$$\begin{aligned}\rho_i(X) &\leq \sum_{j \in N} \rho_i(\hat{X}_j) && (4.56)\\ &< \sum_{j \in N} \rho_j(\hat{X}_j), && (4.57)\end{aligned}$$

where (4.56) follows from *Sub-additivity* of ρ_i, and (4.57) follows from (4.50) and (4.53)-(4.55). Combining (4.56)-(4.57) with $\rho_j(0 \cdot e_\Omega) = 0$ for all $j \in N \backslash \{i\}$ yields a contradiction with Proposition 4.3.4. Hence, for all $(\hat{X}_j)_{j \in N} \in \mathcal{PO}(R)$ and $j \neq i$ it follows that $\hat{X}_j$ is a side-payment. This concludes the second part of the proof. □

Proof of Proposition 4.3.12: Let $R \in \mathcal{RR}$. From Proposition 4.3.4 we get that it is sufficient to show that $\sum_{i \in N} \rho_i(\widetilde{X}_i) = \rho_N^*(X)$ for an $(\widetilde{X}_i)_{i \in N} \in \mathcal{PO}(R)$. Pick $(\widetilde{X}_i)_{i \in N} \in \mathcal{PO}(R)$ as in (4.19) with $m \in M(R)$ and $d \in \mathbb{R}^N$ such that $\sum_{i \in N} d_i = X(\omega_p)$. The risk $\widetilde{X}_i$ is constructed such that

$$\begin{aligned}\widetilde{X}_i(\omega_k) - \widetilde{X}_i(\omega_{k+1}) &= \sum_{\ell=k}^{p-1} [X(\omega_\ell) - X(\omega_{\ell+1})] \cdot \mathbb{1}_{m(\ell)=i} + d_i && (4.58)\\ &\quad - \left(\sum_{\ell=k+1}^{p-1} [X(\omega_\ell) - X(\omega_{\ell+1})] \cdot \mathbb{1}_{m(\ell)=i} + d_i \right)\\ &= [X(\omega_k) - X(\omega_{k+1})] \cdot \mathbb{1}_{m(k)=i}, && (4.59)\end{aligned}$$

for all $k \in \{1, \dots, p-1\}$ and $i \in N$ and, moreover, $\sum_{i \in N} \widetilde{X}_i(\omega_p) = \sum_{i \in N} d_i = X(\omega_p)$. So, it holds that $\widetilde{X}_i(\omega_1) \geq \dots \geq \widetilde{X}_i(\omega_p)$ for all $i \in N$. From this and (4.58)-(4.59) it follows that

$$\sum_{i \in N} \rho_i(\widetilde{X}_i) = \sum_{i \in N} \left[\sum_{k=1}^{p-1} g_i(\mathbb{P}(\{\omega_1, \dots, \omega_k\}))[\widetilde{X}_i(\omega_k) - \widetilde{X}_i(\omega_{k+1})] + \widetilde{X}_i(\omega_p) \right]$$

$$
\begin{aligned}
&= \sum_{i\in N}\left[\sum_{k=1}^{p-1} g_i(\mathbb{P}(\{\omega_1,\dots,\omega_k\}))[X(\omega_k)-X(\omega_{k+1})]\cdot \mathbb{1}_{m(k)=i} + d_i\right]\\
&= \sum_{k=1}^{p-1}\sum_{i\in N} g_i(\mathbb{P}(\{\omega_1,\dots,\omega_k\}))[X(\omega_k)-X(\omega_{k+1})]\cdot \mathbb{1}_{m(k)=i} + \sum_{i\in N} d_i\\
&= \sum_{k=1}^{p-1}[X(\omega_k)-X(\omega_{k+1})]\sum_{i\in N} g_i(\mathbb{P}(\{\omega_1,\dots,\omega_k\}))\cdot \mathbb{1}_{m(k)=i} + X(\omega_p)\\
&= \sum_{k=1}^{p-1}[X(\omega_k)-X(\omega_{k+1})]\min\{g_i(\mathbb{P}(\{\omega_1,\dots,\omega_k\})) : i\in N\} + X(\omega_p)\\
&= \sum_{k=1}^{p-1}[X(\omega_k)-X(\omega_{k+1})]g_N^*(\mathbb{P}(\{\omega_1,\dots,\omega_k\})) + X(\omega_p)\\
&= \rho_N^*(X).
\end{aligned}
$$

This concludes the proof. □

Proof of Lemma 4.3.13: Let $R \in \mathcal{RR}$ be such that [SC] holds and suppose that $(\widetilde{X}_j)_{j\in N} \in \mathcal{PO}(R)$ is such that there exist $i \in N$ and $k \in \{1,\dots,p-1\}$ where $\widetilde{X}_i(\omega_k) < \widetilde{X}_i(\omega_{k+1})$. Recall from (4.3) and (4.4) that $\rho_N^*(X) = E_{\mathbb{Q}_X}[X]$, where $\mathbb{Q}_X$ is the additive probability measure such that $\mathbb{Q}_X(\{\omega_\ell\}) = g_N^*(\mathbb{P}(\{\omega_1,\dots,\omega_\ell\})) - g_N^*(\mathbb{P}(\{\omega_1,\dots,\omega_{\ell-1}\}))$ for all $\ell \in \{1,\dots,p\}$. One can verify that, due to concavity of the function g_N^*, it holds that $\mathbb{Q}_X \in Q(\rho_N^*) \subseteq Q(\rho_j)$ for all $j \in N$ and, so,

$$\rho_j(\widetilde{X}_j) \geq E_{\mathbb{Q}_X}[\widetilde{X}_j], \text{ for all } j \in N. \tag{4.60}$$

Next, we show that

$$\rho_i(\widetilde{X}_i) > E_{\mathbb{Q}_X}[\widetilde{X}_i]. \tag{4.61}$$

Since $Q(\rho_N^*) \subseteq Q(\rho_i)$, it follows that $\rho_i(\widetilde{X}_i) \geq \rho_N^*(\widetilde{X}_i)$ and, so, it is sufficient to show $\rho_N^*(\widetilde{X}_i) > E_{\mathbb{Q}_X}[\widetilde{X}_i]$. We will show that

$$
\begin{aligned}
&[g_N^*(\mathbb{P}(\{\omega_1,\dots,\omega_{k-1},\omega_{k+1}\})) - g_N^*(\mathbb{P}(\{\omega_1,\dots,\omega_{k-1}\}))]\widetilde{X}_i(\omega_{k+1})\\
&\quad + [g_N^*(\mathbb{P}(\{\omega_1,\dots,\omega_{k+1}\})) - g_N^*(\mathbb{P}(\{\omega_1,\dots,\omega_{k-1},\omega_{k+1}\}))]\widetilde{X}_i(\omega_k)\\
&> [g_N^*(\mathbb{P}(\{\omega_1,\dots,\omega_k\})) - g_N^*(\mathbb{P}(\{\omega_1,\dots,\omega_{k-1}\}))]\widetilde{X}_i(\omega_k)\\
&\quad + [g_N^*(\mathbb{P}(\{\omega_1,\dots,\omega_{k+1}\})) - g_N^*(\mathbb{P}(\{\omega_1,\dots,\omega_k\}))]\widetilde{X}_i(\omega_{k+1}).
\end{aligned} \tag{4.62}
$$

Equivalently, (4.62) can be written for $\upsilon(F) = g_N^*(\mathbb{P}(F))$ for all $F \subseteq \Omega$ and $W = \{\omega_1,\dots,\omega_{k-1}\}$

as

$$\widetilde{X}_i(\omega_{k+1})\left[v(W\cup\{\omega_{k+1}\})-v(W)-v(W\cup\{\omega_k\}\cup\{\omega_{k+1}\})+v(W\cup\{\omega_k\})\right]$$
$$>\widetilde{X}_i(\omega_k)\left[v(W\cup\{\omega_{k+1}\})-v(W)-v(W\cup\{\omega_k\}\cup\{\omega_{k+1}\})+v(W\cup\{\omega_k\})\right]. \quad (4.63)$$

Since the function g_N^* is strictly concave and $\mathbb{P}(\{\omega\})>0$ for all $\omega\in\Omega$, it follows that

$$v(W\cup\{\omega_{k+1}\})-v(W)-v(W\cup\{\omega_k\}\cup\{\omega_{k+1}\})+v(W\cup\{\omega_k\})>0.$$

From this it follows that (4.63) holds and, so, (4.62) holds.

From (4.62), we get that $E_{\mathbb{Q}'}[\widetilde{X}_i]>E_{\mathbb{Q}_X}[\widetilde{X}_i]$ where

$$\mathbb{Q}'(\omega)=\begin{cases} g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_{k+1}\}))-g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_{k-1},\omega_{k+1}\})) & \text{if } \omega=\omega_k,\\ g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_{k-1},\omega_{k+1}\}))-g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_{k-1}\})) & \text{if } \omega=\omega_{k+1},\\ \mathbb{Q}_X(\omega) & \text{otherwise.}\end{cases}$$

From $\mathbb{Q}'\in Q(\rho_N^*)$ it follows that $\rho_N^*(\widetilde{X}_i)\geq E_{\mathbb{Q}'}[\widetilde{X}_i]$. So, we have shown that (4.61) holds. To conclude, it follows that

$$\sum_{j\in N}\rho_j(\widetilde{X}_j)\geq\sum_{j\in N\setminus\{i\}}E_{\mathbb{Q}_X}[\widetilde{X}_j]+\rho_i(\widetilde{X}_i) \quad (4.64)$$
$$>\sum_{j\in N}E_{\mathbb{Q}_X}[\widetilde{X}_j] \quad (4.65)$$
$$=E_{\mathbb{Q}_X}[X] \quad (4.66)$$
$$=\rho_N^*(X), \quad (4.67)$$

where (4.64) follows from (4.60), (4.65) follows from (4.61), and (4.67) follows from (4.3) and (4.4). From Proposition 4.3.12 we get that a risk redistribution $(\hat{X}_j)_{j\in N}\in\mathcal{F}(R)$ is Pareto optimal only if $\sum_{j\in N}\rho_j(\hat{X}_j)=\rho_N^*(X)$. So, it follows that $(\widetilde{X}_j)_{j\in N}$ is not Pareto optimal, which is a contradiction. Hence, we have $\widetilde{X}_j(\omega_1)\geq\cdots\geq\widetilde{X}_j(\omega_p)$ for all $j\in N$ and $(\widetilde{X}_j)_{j\in N}\in\mathcal{PO}(R)$. This concludes the proof. □

In order to show Theorem 4.3.15, we first need the following result.

Lemma 4.A.1 *If $R\in\mathcal{RR}$ is such that [SC] holds and $(\widetilde{X}_i)_{i\in N}\in\mathcal{F}(R)$, it holds that $(\widetilde{X}_i)_{i\in N}\in\mathcal{PO}(R)$ if and only if*

$$\sum_{i\in N}g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_k\}))[\widetilde{X}_i(\omega_k)-\widetilde{X}_i(\omega_{k+1})]=g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_k\}))[X(\omega_k)-X(\omega_{k+1})], \quad (4.68)$$

for all $k\in\{1,\ldots,p-1\}$ and $\sum_{i\in N}\widetilde{X}_i(\omega_p)=X(\omega_p)$.

Proof: The "$\Leftarrow$" ("if") part of the proof follows directly from Theorem 4.3.11. We continue with showing the "$\Rightarrow$" ("only if") part. Let $R \in \mathcal{RR}$ be such that [SC] holds. We first show that $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$ only if

$$\begin{aligned}
&\sum_{i\in N}[g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_k\})) - g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_{k-1}\}))]\widetilde{X}_i(\omega_k) \\
&\qquad = [g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_k\})) - g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_{k-1}\}))]X(\omega_k), \qquad (4.69)
\end{aligned}$$

for all $k \in \{1,\ldots,p\}$. Suppose that $(\widetilde{X})_{i\in N}$ is Pareto optimal and that (4.69) does not hold. From Lemma 4.3.13 it follows that $\widetilde{X}_i(\omega_1) \geq \cdots \geq \widetilde{X}_i(\omega_p)$ for all $i \in N$. Let there exist a $\hat{k} \in \{1,\ldots,p-1\}$ where

$$\begin{aligned}
&\sum_{i\in N}[g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_{\hat{k}}\})) - g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_{\hat{k}-1}\}))]\widetilde{X}_i(\omega_{\hat{k}}) \\
&\qquad < [g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_{\hat{k}}\})) - g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_{\hat{k}-1}\}))]X(\omega_{\hat{k}}).
\end{aligned}$$

Then, for all $i \in N$, we define $\hat{X}_i(\omega_{\hat{k}}) = \widetilde{X}_i(\omega_{\hat{k}})$ and $\hat{X}_i(\omega_{k'}) = \bar{X}_i(\omega_{k'})$ for all $k' \neq \hat{k}$, where $\bar{X}_i$ is as in Theorem 4.3.11 with $d \in \mathbb{R}^N$ such that $\sum_{i\in N} d_i = X(\omega_p)$ and $\hat{X}_i(\omega_1) \geq \cdots \geq \hat{X}_i(\omega_p)$. Then, it holds that

$$\begin{aligned}
\sum_{i\in N}\rho_i(\hat{X}_i) &= \sum_{i\in N}\sum_{k=1}^{p}[g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_k\})) - g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_{k-1}\}))]\hat{X}_i(\omega_k) && (4.70)\\
&= \sum_{k=1}^{p}\sum_{i\in N}[g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_k\})) - g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_{k-1}\}))]\hat{X}_i(\omega_k) && (4.71)\\
&< \sum_{k=1}^{p}[g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_k\})) - g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_{k-1}\}))]X(\omega_k) && (4.72)\\
&= \rho_N^*(X). && (4.73)
\end{aligned}$$

This is a contradiction with Proposition 4.3.12.

Similarly, let there exist a $\hat{k} \in \{1,\ldots,p-1\}$ where

$$\begin{aligned}
&\sum_{i\in N}[g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_{\hat{k}}\})) - g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_{\hat{k}-1}\}))]\widetilde{X}_i(\omega_{\hat{k}}) \\
&\qquad > [g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_{\hat{k}}\})) - g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_{\hat{k}-1}\}))]X(\omega_{\hat{k}}).
\end{aligned}$$

Then, for $(\hat{X}_i)_{i\in N} \in \mathcal{F}(R)$ as in Theorem 4.3.11, it holds due to the contradiction in (4.70)-(4.73) that

$$\sum_{i\in N}\rho_i(\widetilde{X}_i) > \sum_{i\in N}\rho_i(\hat{X}_i),$$

which is due to Proposition 4.3.4 a contradiction with Pareto optimality of $(\widetilde{X}_i)_{i\in N}$.

Finally showing that the expressions (4.68) and (4.69) are equivalent can be shown via re-arranging some terms. □

Proof of Theorem 4.3.15: The "⇒" ("only if") part follows directly from Theorem 4.3.11. We continue with showing the "⇐" ("if") part. Let $R \in \mathcal{RR}$ be such that [SC] holds and for all $k \in \{1, \dots, p-1\}$ such that $X(\omega_k) > X(\omega_{k+1})$ there exists a firm $j \in N$ such that for all $m \in M(R)$ it holds that $m(k) = j$. Then, we have to show that there exists a unique risk redistribution up to side-payments that is of the form (4.19). Suppose that there exists a risk redistribution $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$ that is not of the form (4.19).

From Corollary 4.3.14 it follows that $X(\omega_k) = X(\omega_{k+1})$ implies $\widetilde{X}_i(\omega_k) = \widetilde{X}_i(\omega_{k+1})$ for all $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$ and $i \in N$. Since $(\widetilde{X}_i)_{i\in N} \in \mathcal{PO}(R)$ is not of the form (4.19), there exists a $k \in \{1, \dots, p-1\}$ such that $X(\omega_k) > X(\omega_{k+1})$ and

$$[\widetilde{X}_i(\omega_k) - \widetilde{X}_i(\omega_{k+1})] = [X(\omega_k) - X(\omega_{k+1})] \cdot \mathbb{1}_{m(k)=i} + d_i, \text{ for all } i \in N, \tag{4.74}$$

where $d \in \mathbb{R}^N$ is such that $\sum_{i\in N} d_i = 0$ and $d \neq (0, \dots, 0)$. Note that we assumed that $m(k)$ is the same for all $m \in M(R)$. Let $m \in M(R)$. From Lemma 4.3.13 it follows that $d_j \geq 0$ for all $j \neq m(k)$. Therefore, it holds that $d_{m(k)} < 0$ and, so, we get

$$\begin{aligned}
&\sum_{i\in N} g_i(\mathbb{P}(\{\omega_1, \dots, \omega_k\}))[\widetilde{X}_i(\omega_k) - \widetilde{X}_i(\omega_{k+1})] \\
&\quad = \sum_{i\in N} g_i(\mathbb{P}(\{\omega_1, \dots, \omega_k\}))([X(\omega_k) - X(\omega_{k+1})] \cdot \mathbb{1}_{m(k)=i} + d_i) \qquad (4.75)\\
&\quad = \sum_{i\in N} g_i(\mathbb{P}(\{\omega_1, \dots, \omega_k\}))[X(\omega_k) - X(\omega_{k+1})] \cdot \mathbb{1}_{m(k)=i} + \sum_{i\in N} g_i(\mathbb{P}(\{\omega_1, \dots, \omega_k\}))d_i \\
&\quad > \sum_{i\in N} g_i(\mathbb{P}(\{\omega_1, \dots, \omega_k\}))[X(\omega_k) - X(\omega_{k+1})] \cdot \mathbb{1}_{m(k)=i}, \qquad (4.76)
\end{aligned}$$

where (4.75) follows from (4.74) and (4.76) follows from $\sum_{i\in N} d_i = 0$ and that for all $j \neq m(k)$ it holds that

$$g_{m(k)}(\mathbb{P}(\{\omega_1, \dots, \omega_k\})) < g_j(\mathbb{P}(\{\omega_1, \dots, \omega_k\}))$$

and $d_j \geq 0$. According to Lemma 4.A.1, this is not Pareto optimal. This is a contradiction and, so, $d = (0, \dots, 0)$. This concludes the proof. □

In order to prove Proposition 4.3.17, we first show the following technical result.

Lemma 4.A.2 *For all $R \in \mathcal{RR}$, it holds that*

$$\rho_i(Y) \geq \rho_N^*(Y), \textit{ for all } Y \in \mathbb{R}^\Omega \textit{ and all } i \in N.$$

Proof: This result follows directly from the fact that $g_i(x) \geq \min\{g_j(x) : j \in N\}$ for all $x \in [0,1]$ and, therefore, $Q(\rho_N^*) \subseteq Q(\rho_i)$. □

Proof of Proposition 4.3.17: First, we show the "⇐" ("if") part of the proof. Let $R \in \mathcal{RR}$ be such that $X_i, i \in N$ are all comonotone with each other and let $\rho_i(X_i) = \rho_N^*(X_i)$ for all $i \in N$. Then, $\rho_N^*(X) = \sum_{i \in N} \rho_i(X_i)$ follows directly from:

$$\sum_{i \in N} \rho_i(X_i) = \sum_{i \in N} \rho_N^*(X_i) \tag{4.77}$$

$$= \rho_N^*(X). \tag{4.78}$$

Here, (4.77) follows from $\rho_i(X_i) = \rho_N^*(X_i)$ for all $i \in N$ and (4.78) follows from that all $X_i, i \in N$ are comonotone with each other and *Comonotonic Additivity* of ρ_N^*.

Next, we show the "⇒" ("only if") part of the proof. Let $R \in \mathcal{RR}$ be such that $\rho_N^*(X) = \sum_{i \in N} \rho_i(X_i)$. Generally, it follows from *Sub-additivity* of ρ_N^* and Lemma 4.A.2 that

$$\rho_N^*(X) \leq \sum_{i \in N} \rho_N^*(X_i) \tag{4.79}$$

$$\leq \sum_{i \in N} \rho_i(X_i). \tag{4.80}$$

Since it holds that $\rho_N^*(X) = \sum_{i \in N} \rho_i(X_i)$, the inequalities turn into equalities in (4.79)-(4.80). The equality $\rho_N^*(X) = \sum_{i \in N} \rho_N^*(X_i)$ implies comonotonicity of the risks $X_i, i \in N$ with each other since [SC] holds. This is shown in the proof of Lemma 4.3.13. The equality $\sum_{i \in N} \rho_N^*(X_i) = \sum_{i \in N} \rho_i(X_i)$ implies that $\rho_i(X_i) = \rho_N^*(X_i)$ for every $i \in N$. This follows from Lemma 4.A.2. This concludes the proof. □

Proof of Theorem 4.4.2: Let $R \in \mathcal{RR}$ be such that [SC] holds. First, we show the "⇐" ("if") part of the proof. Let [SO] hold. According to Lemma 4.4.1, there exists an equilibrium. Pick an equilibrium $(\hat{p}, (\hat{X}_i)_{i \in N})$. Lemma 4.4.1 states that every equilibrium risk redistribution is Pareto optimal and, so, we have $(\hat{X}_i)_{i \in N} \in \mathcal{PO}(R)$. From this, strict concavity of the function g_N^* and Lemma 4.3.13 it follows that all risks $\hat{X}_i, i \in N$ are comonotone with each other, i.e., $\hat{X}_i(\omega_1) \geq \cdots \geq \hat{X}_i(\omega_p)$ for all $i \in N$. Hence, the objective function for firm $i \in N$ in (4.23) can be written as

$$\rho_i(\widetilde{X}_i) = \sum_{k=1}^{p} [g_i(\mathbb{P}(\{\omega_1, \ldots, \omega_k\})) - g_i(\mathbb{P}(\{\omega_1, \ldots, \omega_{k-1}\}))]\widetilde{X}_i(\omega_k), \tag{4.81}$$

which is minimized over all $\widetilde{X}_i \in \mathbb{R}^\Omega$ such that $\widetilde{X}_i(\omega_1) \geq \cdots \geq \widetilde{X}_i(\omega_p)$ and $\pi(\hat{p}, \widetilde{X}_i) \geq \pi(\hat{p}, X_i)$. A minimum is obtained in $\widetilde{X}_i = \hat{X}_i$.

Since the constraints and the objective function in (4.81) are all affine, we get that the equilibrium risk redistribution $(\hat{X}_i)_{i\in N}$ satisfies the Kuhn-Tucker conditions. The Kuhn-Tucker conditions are obtained by the first-order conditions of the following function with respect to $\widetilde{X}_i(\omega_k)$:

$$\sum_{k=1}^{p}\Big([g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_k\}))-g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_{k-1}\}))]\widetilde{X}_i(\omega_k)-\lambda_i\hat{p}_k[\widetilde{X}_i(\omega_k)-X_i(\omega_k)]\Big) \\ -\sum_{k=1}^{p-1}\gamma_{i,k}\left[\widetilde{X}_i(\omega_k)-\widetilde{X}_i(\omega_{k+1})\right] \tag{4.82}$$

in $\widetilde{X}_i(\omega_k)=\hat{X}_i(\omega_k)$, for all $k\in\{1,\ldots,p\}$ and $i\in N$, where $\lambda_i\geq 0$ and $\gamma_{i,k}\geq 0$ are the Kuhn-Tucker multipliers of the constraints $\pi(\hat{p},\widetilde{X}_i)\geq\pi(\hat{p},X_i)$ and $\widetilde{X}_i(\omega_k)\geq\widetilde{X}_i(\omega_{k+1})$, respectively. These Kuhn-Tucker conditions are then given by

$$g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_k\}))-g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_{k-1}\}))=\begin{cases}\lambda_i\hat{p}_k+\gamma_{i,k} & \text{if } k=1,\\ \lambda_i\hat{p}_k+\gamma_{i,k}-\gamma_{i,k-1} & \text{if } k=2,\ldots,p-1,\\ \lambda_i\hat{p}_k-\gamma_{i,k-1} & \text{if } k=p,\end{cases} \tag{4.83}$$

for all $k\in\{1,\ldots,p\}$ and $i\in N$, which hold under the constraints $\lambda_i[\pi(\hat{p},\hat{X}_i)-\pi(\hat{p},X_i)]=0,\lambda_i\geq 0,\gamma_{i,k}\left[\hat{X}_i(\omega_k)-\hat{X}_i(\omega_{k+1})\right]=0$ and $\gamma_{i,k}\geq 0$ for all $k\in\{1,\ldots,p-1\}$ and $i\in N$. Since $g_i(0)=0$ and $g_i(1)=1$, it holds that

$$\sum_{k=1}^{p}[g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_k\}))-g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_{k-1}\}))]=1, \tag{4.84}$$

and, moreover, it holds that

$$\sum_{k=1}^{p}\hat{p}_k=1, \tag{4.85}$$

since $\pi(\hat{p},e_\Omega)=1$ and

$$\gamma_{i,1}+\sum_{k=2}^{p-1}(\gamma_{i,k}-\gamma_{i,k-1})-\gamma_{i,p-1}=0. \tag{4.86}$$

From (4.83), (4.84), (4.85) and (4.86) it follows that $\lambda_i=1$ for all $i\in N$ and, so, we can write

(4.83) as:

$$g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_k\})) - g_i(\mathbb{P}(\{\omega_1,\ldots,\omega_{k-1}\})) = \begin{cases} \hat{p}_k + \gamma_{i,k} & \text{if } k=1, \\ \hat{p}_k + \gamma_{i,k} - \gamma_{i,k-1} & \text{if } k=2,\ldots,p-1, \\ \hat{p}_k - \gamma_{i,k-1} & \text{if } k=p, \end{cases} \tag{4.87}$$

for all $k \in \{1,\ldots,p\}$ and $i \in N$. Since $X(\omega_1) > X(\omega_2)$, it holds that there exists at least one firm $i_0 \in N$ such that $\gamma_{i_0,1} = 0$. From this and $\gamma_{j,1} \geq 0$ for all $j \in N$ it follows that

$$\hat{p}_1 = g_{i_0}(\mathbb{P}(\{\omega_1\})) = g_N^*(\mathbb{P}(\{\omega_1\})) \text{ and } \gamma_{i,1} = g_i(\mathbb{P}(\{\omega_1\})) - g_N^*(\mathbb{P}(\{\omega_1\})), \text{ for all } i \in N. \tag{4.88}$$

If $p > 2$, it follows from (4.87) and (4.88) that for $k = 2$ we get

$$g_i(\mathbb{P}(\{\omega_1,\omega_2\})) - g_N^*(\mathbb{P}(\{\omega_1\})) = \hat{p}_2 + \gamma_{i,2}, \text{ for all } i \in N, \tag{4.89}$$

and, so, we get

$$\hat{p}_2 = g_N^*(\mathbb{P}(\{\omega_1,\omega_2\})) - g_N^*(\mathbb{P}(\{\omega_1\})) \text{ and } \gamma_{i,2} = g_i(\mathbb{P}(\{\omega_1,\omega_2\})) - g_N^*(\mathbb{P}(\{\omega_1,\omega_2\})), \tag{4.90}$$

for all $i \in N$. Continuing this procedure for all $k \in \{1,\ldots,p\}$, we obtain by induction the unique equilibrium price vector $\hat{p}$. This concludes the first part of the proof.

Next, we show the "⇒" ("only if") part of the proof. Let there exists a unique equilibrium price vector $\hat{p}$. Suppose that [SO] does not hold. Then, there exists a $k \in \{1,\ldots,p-1\}$ such that $X(\omega_k) = X(\omega_{k+1})$. Construct the price vectors as in (4.25) for every ordering on the state space Ω such that $X(\omega_1) \geq \cdots \geq X(\omega_p)$. These vectors are all equilibrium prices as they satisfy the conditions in (4.83) for some Kuhn-Tucker multipliers. Due to strict concavity of the function g_N^*, these price vectors are not identical. This is a contradiction with that there is a unique equilibrium price vector. This concludes the proof. □

Proof of Theorem 4.4.3: Let $R \in \mathcal{RR}$ be such that conditions [SC] and [SO] hold. From Theorem 4.4.2 it follows that $\hat{p} = \mathbb{Q}_X \in Q(\rho_N^*)$, where $\mathbb{Q}_X(\{\omega_k\}) = g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_k\})) - g_N^*(\mathbb{P}(\{\omega_1,\ldots,\omega_{k-1}\}))$ for all $k \in \{1,\ldots,p\}$. So, from (4.7) it follows that $\hat{p} \in \bigcap_{j\in N} Q(\rho_j) \subseteq Q(\rho_i)$ for all $i \in N$. From this and (4.6) follows that $\pi(\hat{p}, Y) \leq \rho_i(Y)$ for all $i \in N$ and all $Y \in \mathbb{R}^\Omega$. So, any risk $\hat{X}_i \in \mathbb{R}^\Omega$ such that $\rho_i(\hat{X}_i) = \pi(\hat{p}, \hat{X}_i) = \pi(\hat{p}, X_i) = E_{\hat{p}}[X_i]$ minimizes the objective function in (4.23). □

Proof of Theorem 4.5.2: One can determine a Pareto optimal risk redistribution $(\hat{X}_i)_{i\in N} \in \mathcal{PO}(R)$ via Theorem 4.3.11 for any choice of $m \in M(R)$ and $d \in \mathbb{R}^N$ such that $\sum_{i\in N} d_i = X(\omega_p)$.

For every allocation $a \in \mathbb{R}^N$, it holds for the side-payments $c_i \cdot e_\Omega, i \in N$ with $c_i = a_i - \rho_i(\hat{X}_i)$ that $\sum_{i\in N} c_i = 0$, $(\widetilde{X}_i)_{i\in N} = (\hat{X}_i + c_i \cdot e_\Omega)_{i\in N} \in \mathcal{PO}(R)$ and $a = (\rho_i(\widetilde{X}_i))_{i\in N}$.

If the conditions [SC] and [U] hold, it follows from Theorem 4.3.15 that the choice of the Pareto optimal risk redistribution $(\hat{X}_i)_{i\in N}$ is unique up to side-payments. Hence, for every allocation a, there is a unique risk redistribution $(\widetilde{X}_i)_{i\in N}$ such that $a_i = \rho_i(\widetilde{X}_i)$ for all $i \in N$. □

Proof of Proposition 4.5.7: First, we show "⇒" ("only if") part of the proof. Let $R \in \mathcal{RR}$ be such that $X(\omega) = X(\omega')$ for equivalent states $\omega, \omega' \in \Omega$ only. Then, it holds that $\sum_{i\in N} \lambda_i X_i(\omega) = \sum_{i\in N} \lambda_i X_i(\omega')$ for all equivalent states $\omega, \omega' \in \Omega$ and $\lambda \in \mathbb{R}^N_{++}$. For all other states $\omega, \omega' \in \Omega$ such that $X(\omega) > X(\omega')$, it follows from continuity that there exists a neighborhood $\hat{U} \subset \mathbb{R}^N_{++}$ of e_N such that $\sum_{i\in N} \lambda_i X_i(\omega) > \sum_{i\in N} \lambda_i X_i(\omega')$ for all $\lambda \in \hat{U}$. Hence, there exists a neighborhood $U \subset \mathbb{R}^N_{++}$ of e_N such that $\sum_{i\in N} \lambda_i X_i(\omega_1) \geq \cdots \geq \sum_{i\in N} \lambda_i X_i(\omega_p)$ for all $\lambda \in U$. From this, (4.3) and (4.4) it follows that

$$r(\lambda) = E_{\mathbb{Q}_X}\left[\sum_{i\in N} \lambda_i X_i\right], \text{ for all } \lambda \in U. \tag{4.91}$$

This linear function is partially differentiable and, so, the Aumann-Shapley value exists.

We continue by showing the "⇐" ("if") part of the proof. Let $R \in \mathcal{RR}$ be such that [SC] holds and the Aumann-Shapley value exists. Define $\hat{Q} \subset Q(\rho_N^*)$ as all extreme points of the convex polytope $Q(\rho_N^*)$ such that $r(e_N) = E_{\mathbb{Q}}[X]$ for all $\mathbb{Q} \in \hat{Q}$. Note that the set $\hat{Q}$ is non-empty since $Q(\rho_N^*)$ is compact. Since the fuzzy game r is piecewise linear (see in Chapter 2 of this dissertation), there exists a neighborhood $U \subset \mathbb{R}^N_{++}$ of e_N and a $\hat{\mathbb{Q}} \in \hat{Q}$ such that

$$r(\lambda) = E_{\hat{\mathbb{Q}}}\left[\sum_{i\in N} \lambda_i X_i\right], \text{ for all } \lambda \in U. \tag{4.92}$$

For all $\mathbb{Q} \in \hat{Q}$, it holds by definition that

$$E_{\mathbb{Q}}\left[\sum_{i\in N} \lambda_i X_i\right] \leq r(\lambda), \text{ for all } \lambda \in U, \tag{4.93}$$

and, by local linearity of the fuzzy game r on U, it holds for all $\mathbb{Q} \in \hat{Q}$ that

$$E_{\mathbb{Q}}\left[\sum_{i\in N} \lambda_i X_i\right] = r(\lambda), \text{ for all } \lambda \in U. \tag{4.94}$$

Since the partial derivatives of (4.94) exist, it follows that $(E_{\mathbb{Q}}[X_i])_{i\in N}$ is constant for all $\mathbb{Q} \in \hat{Q}$.

For every $\mathbb{Q} \in \hat{Q}$, there exists an ordering on the state space $\Omega = \{\omega_1, \ldots, \omega_p\}$ such that $\mathbb{Q}(\omega_k) = g_N^*(\mathbb{P}(\{\omega_1, \ldots, \omega_k\})) - g_N^*(\mathbb{P}(\{\omega_1, \ldots, \omega_{k-1}\}))$ for all $k \in \{1, \ldots, p\}$. Since [SC] holds, we get from (4.61) in the proof of Lemma 4.3.13 that for all $\mathbb{Q}_1, \mathbb{Q}_2 \in \hat{Q}$ we have $\mathbb{Q}_1(\omega_k) \neq \mathbb{Q}_2(\omega_k)$ only if $X(\omega_{k-1}) = X(\omega_k)$ or $X(\omega_k) = X(\omega_{k+1})$. Let $X(\omega_k) = X(\omega_{k+1})$ and let $\mathbb{Q}_1, \mathbb{Q}_2 \in \hat{Q}$ both be generated by a different ordering on the state space $\Omega = \{\omega_1, \dots, \omega_p\}$ such that $X(\omega_1) \geq \dots \geq X(\omega_p)$ only via interchanging the states ω_k and ω_{k+1}. So, it holds that $\mathbb{Q}_1(\omega) = \mathbb{Q}_2(\omega)$ for all $\omega \in \Omega \backslash \{\omega_k, \omega_{k+1}\}$. From strict concavity of the function g_N^* we get $\mathbb{Q}_1(\omega_k) \neq \mathbb{Q}_2(\omega_k)$. Hence, $(E_{\mathbb{Q}}[X_i])_{i \in N}$ is constant for $\mathbb{Q} \in \{\mathbb{Q}_1, \mathbb{Q}_2\}$ only if $X_i(\omega_k) = X_i(\omega_{k+1})$ for all $i \in N$. Continuing this for all states $\omega_k, \omega_{k+1} \in \Omega$ such that $X(\omega_k) = X(\omega_{k+1})$ yields that the Aumann-Shapley value exists if for all $\omega_k, \omega_{k+1} \in \Omega$ such that $X(\omega_k) = X(\omega_{k+1})$ the states ω_k and ω_{k+1} are equivalent. This concludes the proof. $\square$

Proof of Proposition 4.5.12: Let $R \in \mathcal{RR}'$. Then, the fuzzy game r on a neighborhood of e_N is given in (4.91). Partial differentiating the fuzzy game r in $\lambda = e_N$ yields

$$\frac{\partial r}{\partial \lambda_i}(e_N) = E_{\mathbb{Q}_X}\left[X_i\right], \text{ for all } i \in N.$$

Hence, the Aumann-Shapley value is given by

$$AS_i(R) = \sum_{k=1}^{p} [g_N^*(\mathbb{P}(\{\omega_1, \dots, \omega_k\})) - g_N^*(\mathbb{P}(\{\omega_1, \dots, \omega_{k-1}\}))] X_i(\omega_k), \text{ for all } i \in N.$$

This concludes the proof. $\square$

Proof of Lemma 4.5.9: Let $R \in \mathcal{RR}$. Sub-additivity of the fuzzy game r follows directly from

$$r(\lambda) + r(\lambda') = \rho^*_{\{i \in N: \lambda_i > 0\}}\left(\sum_{i \in N} \lambda_i X_i\right) + \rho^*_{\{i \in N: \lambda'_i > 0\}}\left(\sum_{i \in N} \lambda'_i X_i\right) \tag{4.95}$$

$$\geq \rho^*_{\{i \in N: \lambda_i + \lambda'_i > 0\}}\left(\sum_{i \in N} \lambda_i X_i\right) + \rho^*_{\{i \in N: \lambda_i + \lambda'_i > 0\}}\left(\sum_{i \in N} \lambda'_i X_i\right) \tag{4.96}$$

$$\geq \rho^*_{\{i \in N: \lambda_i + \lambda'_i > 0\}}\left(\sum_{i \in N} (\lambda_i + \lambda'_i) X_i\right) \tag{4.97}$$

$$= r(\lambda + \lambda'), \tag{4.98}$$

for all $\lambda, \lambda' \in \mathbb{R}_+^N$. Here, (4.95) follows from (4.37), (4.96) follows from Lemma 4.A.2, (4.97) follows from *Sub-additivity* of $\rho^*_{\{i \in N: \lambda_i + \lambda'_i > 0\}}$, and (4.98) follows from (4.37). This concludes the proof. $\square$

4.B Discussion of the cooperative cost game

In this appendix, we explain the construction of the cooperative cost game (N, c) as in (4.32). If firms use an expected utility function, the corresponding cooperative game is a Non-Transferable Utility game. If firms use a distortion risk measure, the Non-Transferable Utility game is equivalent to a reduced-form game, namely a Transferable Utility game. This follows from, e.g., Bergstrom and Varian (1985) and the fact that the Pareto optimal risk redistributions are characterized as the ones minimizing the aggregate risk adjusted value of the liabilities (see Proposition 4.3.12). We define this Transferable Utility game for the risk redistribution problem as a cooperative cost game (N, c) as in (4.32). The game (N, c) is a special case of market games that are introduced by Shapley and Shubik (1969) for a wide class of utility functions.

If all firms use the same risk measure, the game (N, c) corresponds with the cooperative cost game in Denault (2001).[44] This follows from the fact that $\rho_i = \rho$ for all $i \in N$ implies $\rho_S^* = \rho$ for all $S \subseteq N$ and, hence, $c(S) = \rho\left(\sum_{i\in S} X_i\right)$. A possible way for determining Pareto optimal risk redistributions corresponding to an allocation is shown in Proposition 4.3.5.

Generally, one can disentangle two marginal effects of a firm $i \in N$ as follows:

$$\begin{aligned} c(S \cup \{i\}) - c(S) = &\underbrace{\rho^*_{S\cup\{i\}}\left(\sum_{j\in S\cup\{i\}} X_j\right) - \rho^*_{S\cup\{i\}}\left(\sum_{j\in S} X_j\right)}_{\text{diversification effect}} \\ &+ \underbrace{\rho^*_{S\cup\{i\}}\left(\sum_{j\in S} X_j\right) - \rho^*_{S}\left(\sum_{j\in S} X_j\right)}_{\text{risk measure effect}}, \end{aligned} \tag{4.99}$$

for every $S \subseteq N\backslash\{i\}$. The diversification effect is due to the hedge benefit from pooling risks. Diversification arises due to different orderings of the risks X_i and $\sum_{j\in S} X_j$. This effect is discussed by Denault (2001). The risk measure effect is the effect that a firm is endowed with a possibly less stringent risk measure. Particularly note that the risk measure effect is independent of risk X_i.

[44] Literally, this is not true as Denault (2001) defines this game for all coherent risk measures and allows for a continuous state space. We neglect these issues.

Bibliography

Aase, K. (1993). Equilibrium in a reinsurance syndicate; existence, uniqueness and characterization. *ASTIN Bulletin 23*, 185–211.

Acciaio, B. (2007). Optimal risk sharing with non-monotone monetary functionals. *Finance and Stochastics 11*, 267–289.

Acciaio, B. and G. Svindland (2009). Optimal risk sharing with different reference probabilities. *Insurance Mathematics and Economics 44*, 426–433.

Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance 26*, 1505–1518.

Acerbi, C. and D. Tasche (2002). On the coherence of expected shortfall. *Journal of Banking and Finance 26*, 1487–1503.

Adam, A., M. Houkari, and J.-P. Laurent (2008). Spectral risk measures and portfolio selection. *Journal of Banking and Finance 32*, 1870–1882.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9*, 203–228.

Asimit, A., A. Badescu, and A. Tsanakas (2013). Optimal risk transfers in insurance groups. *European Actuarial Journal 3*, 159–190.

Aubin, J. (1979). *Mathematical Methods of Game and Economic Theory.* Amsterdam: North-Holland.

Aubin, J. (1981). Cooperative fuzzy games. *Mathematics of Operations Research 6*, 1–13.

Aumann, R. and L. Shapley (1974). *Values of Non-Atomatic Games.* Princeton: Princeton University Press.

Basel Committee on Banking Supervision (2013). Longevity risk transfer markets: market structure, growth drivers and impediments, and potential risks. Joint Forum.

Bauer, D., M. Börger, and J. Russ (2010). On the pricing of longevity-linked securities. *Insurance: Mathematics and Economics 46*, 139–149.

Bergstrom, T. C. and H. R. Varian (1985). When do market games have transferable utility? *Journal of Economic Theory 35*, 222–233.

Billera, L. J. and D. Heath (1982). Allocation of shared costs: A set of axioms yielding a unique procedure. *Mathematics of Operations Research 7*, 32–39.

Blake, D., A. J. G. Cairns, and K. Dowd (2006). Living with mortality: longevity bonds and other mortality-linked securities. *British Actuarial Journal 12*, 153–228.

Boonen, T. (2010). Bargaining for over-the-counter risk redistributions. Master's Thesis, Department of Econometrics and Operations Research, Tilburg University, Tilburg, available at http://arno.uvt.nl/show.cgi?fid=107606.

Boonen, T. (2013). Risk redistribution with distortion risk measures. Manuscript. Available at www.timboonen.com.

Boonen, T., A. De Waegenaere, and H. Norde (2012a). Bargaining for Over-The-Counter risk redistributions: The case of longevity risk. CentER Discussion Paper Series No. 2012-090.

Boonen, T., A. De Waegenaere, and H. Norde (2012b). A game-theoretic generalization of the Aumann-Shapley value for risk capital allocation problems. CentER Discussion Paper Series, No. 2012-091.

Borch, K. (1960). The safety loading of reinsurance premiums. *Scandinavian Actuarial Journal 43*, 163–184.

Borch, K. (1962). Equilibrium in a reinsurance market. *Econometrica 30*, 424–444.

Brouhns, N., M. Denuit, and J. Vermunt (2002). Measuring the longevity risk in mortality projections. *Bulletin of the Swiss Association of Actuaries 2*, 105–130.

Bühlmann, H. (1980). An economic premium principle. *ASTIN Bulletin 11*, 52–60.

Burgert, C. and L. Rüschendorf (2008). Allocations of risks and equilibrium in markets with finitely many traders. *Insurance: Mathematics and Economics 42*, 177–188.

Cairns, A. J. G., D. Blake, and K. Dowd (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance 73*, 687–718.

Cairns, A. J. G., D. Blake, K. Dowd, G. D. Coughlan, and D. Epstein (2007). A quantitative comparison of stochastic mortality models using data from England & Wales and the United States. Pensions Institute Discussion Paper PI-0701.

Cairns, A. J. G., D. Blake, K. Dowd, G. D. Coughlan, D. Epstein, and M. Khalaf-Allah (2008). Mortality density forecasts: An analysis of six stochastic mortality models. Pensions Institute Discussion Paper PI-0801.

Chateauneuf, A., R.-A. Dana, and J.-M. Tallon (2000). Optimal risk-sharing rules and equilibria with Choquet-expected-utility. *Journal of Mathematical Economics 34*, 191–214.

Chateauneuf, A., R. Kast, and A. Lapied (1996). Choquet pricing for financial markets with frictions. *Mathematical Finance 6*, 323–330.

Compte, O. and P. Jehiel (2010). The coalitional nash bargaining solution. *Econometrica 78*, 1593–1623.

Cossette, H., A. Delwarde, M. Denuit, F. Guillot, and E. Marceau (2007). Pension plan valuation and dynamic mortality tables. *North American Actuarial Journal 11*, 279–298.

Coughlan, G. D., D. Epstein, A. Ong, A. Sinha, J. Hevia-Portocarrero, E. Gingrich, M. Khalaf-Allah, and P. Joseph (2007). Lifemetrics: A toolkit for measuring and managing longevity and mortality risks. Technical report, JPMorgan.

Csóka, P., P. Herings, and L. Kóczy (2009). Stable allocations of risk. *Games and Economic Behavior 67*, 266–276.

Csóka, P. and M. Pintér (2011). On the impossibility of fair risk allocation. IEHAS Discussion Papers 1117, Institute of Economics, Hungarian Academy of Sciences.

Damme, E. van (1986). The Nash bargaining solution is optimal. *Journal of Economic Theory 38*, 78–100.

Dana, R.-A. (2011). Comonotonicity, efficient risk-sharing and equilibria in markets with short-selling for concave law-invariant utilities. *Journal of Mathematical Economics 47*, 328–335.

Dana, R.-A. and C. Le Van (2010). Overlapping sets of priors and the existence of efficient allocations and equilibria for risk measures. *Mathematical Finance 20*, 327–339.

De Giorgi, E. and T. Post (2008). Second-order stochastic dominance, reward-risk portfolio selection, and the CAPM. *Journal of Financial and Quantitative Analysis 43*, 525–546.

De Waegenaere, A., R. Kast, and A. Lapied (2003). Choquet pricing and equilibrium. *Insurance: Mathematics and Economics 32*, 359–370.

Delbaen, F. (2000). Coherent risk measures on general probability spaces. In *Avances in Finance and Stochastics, Essays in Honour of Dieter Sondermann*. New York: Springer.

Denault, M. (2001). Coherent allocation of risk capital. *Journal of Risk 4*, 1–34.

Denneberg, D. (1994). *Non-additive measure and integral.* Dordrecht: Kluwer Academic Publishers.

Dhaene, J., M. Goovaerts, and R. Kaas (2003). Economic capital allocation derived from risk measures. *North American Actuarial Journal 7*, 44–59.

Dhaene, J., S. Vanduffel, M. Goovaerts, R. Kaas, Q. Tang, and D. Vyncke (2006). Risk measures and comonotonicity: A review. *Stochastic Models 22*, 573–606.

Dowd, K., D. Blake, and A. J. G. Cairns (2008). Facing up to uncertain life expectancy: The longevity fan charts. Discussion Paper PI-0703, The Pensions Institute, Cass Business School.

Dowd, K., D. Blake, A. J. G. Cairns, and P. Dawson (2006). Survivor swaps. *Journal of Risk and Insurance 73*, 1–17.

Dowd, K., A. J. G. Cairns, D. Blake, G. D. Coughlan, D. Epstein, and M. Khalaf-Allah (2010). Evaluating the goodness of fit of stochastic mortality models. *Insurance: Mathematics and Economics 47*, 255–265.

Duffie, D. and W. Zame (1989). The consumption-based capital asset pricing model. *Econometrica 57*, 1279–1297.

Filipović, D. and M. Kupper (2008). Equilibrium prices for monetary utility functions. *International Journal of Theorerical and Applied Finance 11*, 325–343.

Filipović, D. and G. Svindland (2008). Optimal capital and risk allocations for law- and cash-invariant convex functions. *Finance and Stochastics 12*, 423–439.

Flåm, S. (2011). Exchanges and measures of risks. *Mathematics and Financial Economics 5*, 249–267.

Föllmer, H. and A. Schied (2002). Convex measures of risk and trading constraints. *Finance and Stochastics 6*, 429–447.

Gerber, H. and G. Pafumi (1998). Utility functions: from risk theory to finance. *North American Actuarial Journal 2*, 74–91.

Gillies, D. (1953). *Some Theorems on n-person Games.* Ph. D. thesis, Princeton University.

Gulick, G. van, A. De Waegenaere, and H. Norde (2012). Excess based allocation of risk capital. *Insurance: Mathematics and Economics 50*, 26–42.

Haimanko, O. (2001). Cost sharing: the nondifferentiable case. *Journal of Mathematical Economics 35*, 445–462.

Hári, N., A. De Waegenaere, B. Melenberg, and T. E. Nijman (2008). Longevity risk in portfolios of pension annuities. *Insurance: Mathematics and Economics 42*, 505–519.

Jones, G. (2013). Longevity risk and reinsurance. In Longevity News: Society of Actuaries.

Jouini, E., W. Schachermayer, and N. Touzi (2008). Optimal risk sharing for law invariant monetary utility functions. *Mathematical Finance 18*, 269–292.

Kaina, M. and L. Rüschendorf (2009). On convex risk measures on L^p-spaces. *Mathematical Methods of Operations Research 69*, 475–495.

Kalai, E. (1977). Nonsymmetric Nash solutions and replications of 2-person bargaining. *International Journal of Game Theory 6*, 129–133.

Kalkbrener, M. (2005). An axiomatic approach to capital allocation. *Mathematical Finance 15*, 425–437.

Kiesel, S. and L. Rüschendorf (2009). Characterization of optimal risk allocations for convex risk functionals. *Statistics and Decisions 26*, 303–319.

Laeven, R. and M. Goovaerts (2004). An optimization approach to the dynamic allocation of economic capital. *Insurance: Mathematics and Economics 35*, 299–319.

Landsberger, M. and I. I. Meilijson (1994). Comonotone allocations, Bickel-Lehmann dispersion and the Arrow-Pratt measure of risk aversion. *Annals of Operations Research 52*, 97–106.

Lee, R. D. and L. Carter (1992). Modelling and forecasting the time series of U.S. mortality. *Journal of the American Statistical Association 87*, 659–671.

Ludkovski, M. and L. Rüschendorf (2008). On comonotonicity of Pareto optimal risk sharing. *Statistics & Probability Letters 78*, 1181–1188.

Ludkovski, M. and V. R. Young (2009). Optimal risk sharing under distorted probabilities. *Mathematics and Financial Economics 2*, 87–105.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance 7*, 77–91.

Mertens, J. (1980). Values and derivatives. *Mathematics of Operations Research 5*, 523–552.

Mertens, J. (1988). The Shapley value in the non differentiable case. *International Journal of Game Theory 17*, 1–65.

Mirman, L. and I. Tauman (1982). Demand compatible equitable cost sharing prices. *Mathematics of Operations Research 7*, 40–56.

Moulin, H. (1995). On additive methods to share joint costs. *Japanese Economic Review 46*, 303–332.

Myers, S. C. and J. A. Read (2001). Capital allocation for insurance companies. *Journal of Risk and Insurance 68*, 545–580.

Neyman, A. (2002). Values of games with infinitely many players. Volume 3 of *Handbook of Game Theory with Economic Applications*, pp. 2121 – 2167. Elsevier.

Neyman, A. and R. Smordinsky (2004). Asymptotic values of vector measure games. *Mathematics of Operations Research 29*, 739–775.

OECD (2005). Ageing and pension system reform: Implications for financial markets and economic policies. *Financial Market Trends 1*.

Olivieri, A. (2001). Uncertainty in mortality projections: an actuarial perspective. *Insurance: Mathematics and Economics 29*, 231–245.

Pitacco, E., M. Denuit, S. Haberman, and A. Olivieri (2009). *Modelling Longevity Dynamics for Pensions and Annuity Business*. Oxford: Oxford University Press.

Plat, R. (2009). On stochastic mortality modeling. *Insurance: Mathematics and Economics 45*, 393–404.

Raviv, A. (1979). The design of an optimal insurance policy. *American Economic Review 69*, 84–96.

Riddell, W. C. (1981). Bargaining under uncertainty. *American Economic Review 71*, 579–590.

Rockafellar, R. and S. Uryasev (2000). Optimization of conditional Value-at-Risk. *Journal of Risk 2*, 21–41.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica 50*, 97–109.

Scarf, H. E. (1967). The core of an n person game. *Econometrica 35*, 50–69.

Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal on Applied Mathematics 17*, 1163–1170.

Schmeidler, D. (1986). Integral representation without additivity. *Proceedings of the American Mathematical Society 97*, 255–261.

Shapley, L. (1953). A value for n-person games. In *Contributions to the Theory of Games*, pp. 307–317. Princeton: Princeton University Press.

Shapley, L. (1971). Cores of convex games. *International Journal of Game Theory 1*, 11–26.

Shapley, L. and M. Shubik (1969). On market games. *Journal of Economic Theory 1*, 9–25.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance 19*, 425–442.

Sherris, M. (2006). Solvency, capital allocation, and fair rate of return in insurance. *Journal of Risk and Insurance 73*, 71–96.

Sprumont, Y. (2005). On the discrete version of the Aumann-Shapley cost-sharing method. *Econometrica 73*, 1693–1712.

Stevens, R., A. De Waegenaere, and B. Melenberg (2011). Calculating capital requirements for longevity risk in life insurance products: using an internal model in line with Solvency II. Working paper, Tilburg University.

Tasche, D. (1999). Risk contributions and performance measurement. Preprint, TU Munich.

Tasche, D. (2002). Expected shortfall and beyond. *Journal of Banking and Finance 26*, 1519–1533.

Tijs, S. (1981). Bounds for the core and the τ-value. In *Game Theory and Mathematical Economics*, pp. 123–132. Amsterdam: North-Holland.

Tsai, J., J. Wang, and L. Tzeng (2010). On the optimal product mix in life insurance companies using conditional value at risk. *Insurance: Mathematics and Economics 46*, 235–241.

Tsanakas, A. (2004). Dynamic capital allocation with distortion risk measures. *Insurance: Mathematics and Economics 35*, 223–243.

Tsanakas, A. (2009). To split or not to split: Capital allocation with convex risk measures. *Insurance: Mathematics and Economics 44*, 268–277.

Tsanakas, A. and C. Barnett (2003). Risk capital allocation and cooperative pricing of insurance liabilities. *Insurance: Mathematics and Economics 33*, 239–254.

Tsanakas, A. and N. Christofides (2006). Risk exchange with distorted probabilities. *ASTIN Bulletin 36*, 219–243.

Wang, J., H. Huang, S. Yang, and J. Tsai (2010). An optimal prodict mix for hedging longevity risk in life insurance companies: The immunization theory approach. *Journal of Risk and Insurance 77*, 473–497.

Wang, S. S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics 17*, 43–54.

Wang, S. S. (1996). Premium calculation by transforming the layer premium density. *ASTIN Bulletin 26*, 71–92.

Wang, S. S. (2000). A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance 67*, 15–36.

Wang, S. S. and V. R. Young (1998). Ordering risks: Expected utility theory versus Yaari's dual theory of risk. *Insurance: Mathematics and Economics 22*, 145–161.

Wang, S. S., V. R. Young, and H. H. Panjer (1997). Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics 21*, 173–183.

Wang, Y. (1999). The additivity and dummy axioms in the discrete cost sharing model. *Economics Letters 64*, 187–192.

Wilson, R. (1968). The theory of syndicates. *Econometrica 36*, 119–132.

Wirch, J. and M. Hardy (2001). Distortion risk measures: Coherence and stochastic dominance. Working Paper, available in http://www.gloriamundi.com.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica 55*, 95–115.