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

Tim J. Boonen* Anja De Waegenaere† Henk Norde‡

September 12, 2019

Abstract

The paper proposes a new method to allocate risk capital to divisions or lines of business within a firm. Existing literature advocates an allocation rule that, in game-theoretic terms, is equivalent to using the Aumann-Shapley value as allocation mechanism. The Aumann-Shapley value, however, is only well-defined if a specific differentiability condition is satisfied. The rule that we propose is characterized as the limit of an average of path-based allocation rules with grid size converging to zero. The corresponding allocation rule is equal to the Aumann-Shapley value if it exists. If the Aumann-Shapley value does not exist, the allocation rule is equal to the weighted average of the Aumann-Shapley values of "nearby" capital allocation problems.

Keywords: risk management, capital allocation, risk measure, Aumann-Shapley value, non-differentiability.

JEL-Classification: C71, G32.

1 Introduction

This paper 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 or portfolios within the firm.[1] This problem is called the *risk capital allocation problem*.

*Amsterdam School of Economics, University of Amsterdam.

†Department of Accountancy and Department of Econometrics and OR, Tilburg University, CentER for Economic Research and Netspar.

‡Department of Econometrics and OR, Tilburg University, CentER for Economic Research.

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

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 risky activity (e.g., a specific financial investment) is typically evaluated by means of a risk-return trade-off. Evaluating the performance of a division's activity 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 et al., 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 (e.g., Tasche, 1999; Myers and Read, 2001; Major, 2018), optimization (e.g., Dhaene et al., 2003) and game theory (e.g., Denault, 2001; Tsanakas and Barnett, 2003; Tsanakas, 2004; Powers, 2007; Csóka et al., 2009; Boonen et al., 2017). Our focus in this paper is on game-theoretic approaches to allocating risk capital. Homburg and Scherpereel (2008) and Balog et al. (2017) provide excellent simulation-based game-theoretic comparisons of different methods to allocate capital, with a focus on core-compatibility. This means that allocations are considered as stable, which in game-theoretic terms

[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. Gulick et al. (2012) show the impact of this hedge potential on the allocation of risk capital.

means that it is an element of the core. 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 considers risk capital allocations that are "stable" in the sense that no (set of) division(s) has incentives to withdraw fully or in part from the collective. 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), if it exists, is the only allocation rule that satisfies these additional properties.[3] Aubin (1979) shows that if the Aumann-Shapley value exists, the fuzzy core is single-valued and the Aumann-Shapley value is its unique element. 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 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 an alternative 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

[3]For general production functions, the Aumann-Shapley value is characterized by Aubin (1981), Billera and Heath (1982), and Mirman and Tauman (1982). Moreover, its empirical computation is studied by Bogetoft et al. (2016) using a mathematical programming approach. They propose a lexicographic goal programming technique to overcome non-differentiabilities.

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. When the grid is binary, i.e., when the divisions either participate fully or not at all, this average coincides with the Shapley value. We show that when the grid size converges to zero, this average converges. The allocation rule that we propose in this paper 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.

There is some literature on risk capital allocation in cases where the risk capital function is not differentiable. Gulick et al. (2012) introduce a capital allocation rule in which the excess risks (as measured by the expected loss in excess of the allocated risk capital) are minimized in lexicographical order. This risk capital allocation rule is well-defined in case of non-differentiabilities. Grechuk (2015) uses the characterization of Kalkbrener (2005) to reduce the problem of finding a linear diversifying capital allocation to the problem of selecting the unique center of a convex weakly compact set in a Banach space. The latter problem has a natural solution in the finite-dimensional case, and generalizes the Aumann-Shapley value to the case of non-differentiabilities. Cherny and Orlov (2011) deal with non-differentiability by replacing the derivative of the risk capital function by a directional derivative, which is always well-defined. They refer to this directional derivative as the marginal risk contribution. However, as shown in Balog at al. (2017), the sum of the marginal risk contributions of divisions within a firm would typically not be equal to the firm's aggregate risk, and so the directional derivative approach cannot be used to determine an allocation of risk capital. Centrone and Rosazza Giannin (2018) analyze a family of risk capital allocations for quasi-convex risk measures with non-differentiabilities. Our goal is to characterize a generalization of the Aumann-Shapley value for a class of coherent risk measures.

Aumann and Shapley (1974) show that under very strong assumptions their 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 paper), do not satisfy this strong assumption (also pointed out by Neyman and Smorodinsky, 2004). The allocation rule that we propose in this paper follows a much weaker asymptotic approach than the one used by Aumann and Shapley (1974).

As a consequence, our approach is convergent for all fuzzy games related to risk capital allocation problems. 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 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 risk capital allocations. This specific formula has a geometric interpretation.

The paper is organized as follows. In Section 2, we recall the definition of coherent risk measures and risk capital allocation problems and we introduce the class of risk measures that we will use in this paper. We also briefly summarize existing literature regarding the use of the Aumann-Shapley value as risk capital allocation rule. In Section 3, we introduce a generalization of the Aumann-Shapley value that is well-defined even if the risk capital function is not differentiable. We also provide a closed form expression with a geometric interpretation for the corresponding risk capital allocations. In Section 4 we derive some properties of the new allocation rule. In Section 5, we conclude. All proofs are delegated to the Online Appendix.

2 Risk measures and risk capital allocation problems

In this section we introduce the class of risk measures that we will use in this paper, and we briefly summarize relevant literature regarding game-theoretic approaches to risk capital allocation problems.

2.1 Finitely generated risk measures

Regulators require financial institutions (e.g., pension funds, banks, or insurance companies) to withhold so-called "reserve capital" that needs to be invested safely. The purpose of this reserve capital is to limit the probability of insolvency. A risk is represented by a random variable X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e., Ω is the state space, $\mathcal{F}$ is a σ-algebra on Ω and $\mathbb{P}$ is the physical probability measure on $(\Omega, \mathcal{F})$. The class of probability measures on $(\Omega, \mathcal{F})$ is denoted by $\mathcal{P}(\Omega, \mathcal{F})$. The realization of X can be interpreted as the (net) loss faced at a pre-specified future time. We consider the case where Ω is finite, and we let the σ-algebra $\mathcal{F}$ be its power set, i.e., $\mathcal{F} = 2^{\Omega}$. A risk can then be represented by a vector in $\mathbb{R}^{\Omega}$.

The amount of risk capital to be withheld for a given risk is typically determined using a risk measure, i.e., a function $\rho : \mathbb{R}^\Omega \to \mathbb{R}$ that maps risks into real numbers. Throughout our analyses, we will focus on the case where the risk measure is *coherent* (see Artzner et al. 1999).[4] Artzner et al. (1999) 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

$$\rho(X) = \sup \{E_{\mathbb{Q}}[X] : \mathbb{Q} \in Q\}, \qquad \text{for all } X \in \mathbb{R}^\Omega. \tag{1}$$

We will use the following terminology:

- If for a given risk measure ρ, (1) is satisfied for a set $Q \subset \mathcal{P}(\Omega, \mathcal{F})$, then the set Q is referred to as a *generating probability measure set* for the risk measure ρ.
- Let Q be a *generating probability measure set* for the risk measure ρ, and let $X \in \mathbb{R}^\Omega$ be a risk. A probability measure $\mathbb{Q} \in Q$ such that $\rho(X) = E_{\mathbb{Q}}[X]$ is referred to as a *worst case probability measure* for X.

The generating probability measure set for a given risk measure ρ is typically not unique. We will consider coherent risk measures for which there exists a finite generating probability measure set Q. We will refer to such risk measures as being *finitely generated*.

Definition 2.1 *A coherent risk measure $\rho : \mathbb{R}^\Omega \to \mathbb{R}$ is finitely generated if there exists a finite generating probability measure set, i.e., there exists a finite set $Q \subset \mathcal{P}(\Omega, \mathcal{F})$ such that*

$$\rho(X) = \max \{E_{\mathbb{Q}}[X] : \mathbb{Q} \in Q\}, \qquad \textit{for all } X \in \mathbb{R}^\Omega. \tag{2}$$

The condition in Definition 2.1 may seem restrictive. However, we show in Proposition 2.2 that all coherent risk measures that satisfy *Comonotonic Additivity* (see, e.g., Wang et al., 1997) are finitely generated.[5]

---

[4] A risk measure ρ is called *coherent* if it satisfies the following four properties (Artzner et al., 1999):

- *Sub-additivity*: For all $X, Y \in \mathbb{R}^\Omega$, we have $\rho(X + Y) \leq \rho(X) + \rho(Y)$.
- *Monotonicity*: For all $X, Y \in \mathbb{R}^\Omega$ such that $X \geq Y$, we have $\rho(X) \geq \rho(Y)$.
- *Positive Homogeneity*: For every $X \in \mathbb{R}^\Omega$ and every $c > 0$, we have $\rho(cX) = c\rho(X)$.
- *Translation Invariance*: For every $X \in \mathbb{R}^\Omega$ and every $c \in \mathbb{R}$, we have $\rho(X + c \cdot e_\Omega) = \rho(X) + c$, where $e_\Omega \in \mathbb{R}^\Omega$ is such that $e_\Omega(\omega) = 1$ for all $\omega \in \Omega$.

Note that Artzner et al. (1999) defines risk measures on random variables that are interpreted as gains, whereas we define risk measures on random variables that are interpreted as losses. This affects the definition of the *Monotonicity* property.

[5] A risk measure ρ satisfies *Comonotonic Additivity* if for all $X, Y \in \mathbb{R}^\Omega$ such that X and Y are comonotone, it

Proposition 2.2 *If the risk measure ρ is coherent and satisfies Comonotonic Additivity, then ρ is finitely generated.*

Proposition 2.2 shows that the full class of coherent risk measures that satisfy *Comonotonic Additivity* is finitely generated. We note that while *Comonotonic Additivity* is a sufficient condition for a coherent risk measure to be finitely generated, it is not a necessary condition.

We now show how a finite generating probability measure set can be determined. Let the risk measure ρ be given and suppose it is coherent and satisfies *Comonotonic Additivity*. It then follows from Delbaen (2000) that there exists a supermodular function $\upsilon : \mathcal{F} \to \mathbb{R}_+$ with $\upsilon(\emptyset) = 0$ and $\upsilon(\Omega) = 1$ such that the set

$$Q^{\upsilon} := \{\mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F}) : \mathbb{Q}(A) \geq \upsilon(A) \text{ for all } A \in \mathcal{F}\}, \tag{3}$$

is a generating probability measure set for ρ.[6][7] Now let $\Omega = \{\omega_1, \ldots, \omega_{|\Omega|}\}$, $\sigma : \{1, \ldots, |\Omega|\} \to \{1, \ldots, |\Omega|\}$ be a permutation, and let $\Pi(\Omega)$ be the set of all permutations of $\{1, \ldots, |\Omega|\}$. For every permutation $\sigma \in \Pi(\Omega)$, we define the probability measure:

$$\mathbb{Q}^{\sigma,\upsilon}(\{\omega_{\sigma(j)}\}) := \upsilon\left(\bigcup_{k=1}^{\sigma(j)} \{\omega_{\sigma(k)}\}\right) - \upsilon\left(\bigcup_{k=1}^{\sigma(j)-1} \{\omega_{\sigma(k)}\}\right), \qquad \text{for all } j \in \{1, \ldots, |\Omega|\}. \tag{4}$$

The additive probability measure $\mathbb{Q}^{\sigma,\upsilon}$ on a finite probability space Ω is determined by specifying the values $\mathbb{Q}^{\sigma,\upsilon}(\{\omega\}), \omega \in \Omega$. Supermodularity and non-negativity of υ combined with the fact that $\upsilon(\emptyset) = 0$ and $\upsilon(\Omega) = 1$ imply that $\mathbb{Q}^{\sigma,\upsilon}$ is indeed a probability measure on $(\Omega, \mathcal{F})$.

Proposition 2.3 *Let ρ be a coherent risk measure that satisfies Comonotonic Additivity. Let υ be a supermodular function such that Q^{υ} as defined in (3) is a generating probability measure set for ρ. Then, the set $Q^{FG,\upsilon} := \{\mathbb{Q}^{\sigma,\upsilon} : \sigma \in \Pi(\Omega)\}$ with $\mathbb{Q}^{\sigma,\upsilon}$ as defined in (4) is a finite generating probability measure set for ρ.*

Throughout the paper, we will illustrate our results for the case where *Expected Shortfall* is used as risk measure. In the following example we determine the finite generating probability measure set from Proposition 2.3 for this risk measure.

---

holds that $\rho(X + Y) = \rho(X) + \rho(Y)$ (see, e.g., Wang et al., 1997). Random variables $X, Y \in \mathbb{R}^{\Omega}$ are comonotone if $[X(\omega_1) - X(\omega_2)] \cdot [Y(\omega_1) - Y(\omega_2)] \geq 0$ for all $(\omega_1, \omega_2) \in \Omega \times \Omega$ (e.g., Denneberg, 1994). If a risk measure satisfies *Comonotonic Additivity*, there is no diversification benefit from pooling risks that are comonotone.

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

[7] The supermodular function υ in (3) depends on the risk measure ρ, and is therefore usually denoted as υ^{ρ}. For notational convenience, we do not explicitly denote this dependence.

Example 2.4 *Expected Shortfall with significance level* $\alpha \in (0,1)$*, denoted* $\rho_{\alpha}^{ES}: \mathbb{R}^{\Omega} \to \mathbb{R}$*, is defined as follows (Acerbi and Tasche, 2002, Tasche, 2002):*

$$\rho_{\alpha}^{ES}(Y) = (1-\alpha)^{-1}\left[E_{\mathbb{P}}\left(Y \cdot 1_{\{Y>q_{\alpha}(Y)\}}\right) + q_{\alpha}(Y) \cdot (\mathbb{P}(Y \leq q_{\alpha}(Y)) - \alpha)\right],$$

where $q_{\alpha}(Y)$ *is the smallest* α*-quantile, i.e.,* $q_{\alpha}(Y) := \inf\{x \in \mathbb{R} | \mathbb{P}(Y \leq x) \geq \alpha\}$*. For continuous random variables, the second term would be equal to zero, and so* $\rho_{\alpha}^{ES}(Y)$ *then equals the expected loss conditional on the loss being among the* $(1-\alpha)100\%$ *highest losses (the right tail of the distribution). For discrete distributions, however, the second term is non-zero and corrects for the fact that the probability of exceeding the smallest* α*-quantile need not be equal to* $1-\alpha$*.*

It is well-known that Expected Shortfall is coherent and satisfies Comonotonic Additivity *(see, e.g., Tasche, 2002). Let the supermodular function* υ *be given by* $\upsilon(A) = 1 - g(1-\mathbb{P}(A))$*, for all* $A \in \mathcal{F}$*, where the so-called "distortion function"* g *is given by* $g(x) = \min\left\{\frac{x}{1-\alpha}, 1\right\}$ *for all* $x \in [0,1]$*. Then, the set* Q^{υ} *from (3) is a generating probability measure set for* ρ_{α}^{ES} *(Kusuoka, 2001).*[8] *Because* $g(x)$ *is linear for* $x \leq \alpha$ *and* $\mathbb{Q}(A) \leq 1$ *for all* $\mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F})$*, we can replace the set* $\mathcal{F} = 2^{\Omega}$ *in (3) by the set of singletons* $\{\{\omega\} : \omega \in \Omega\}$*. Hence, the generating probability measure set of* ρ_{α}^{ES} *from (3) is given by:*

$$Q^{\upsilon} = \left\{\mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F}) : \mathbb{Q}(\{\omega\}) \leq \frac{\mathbb{P}(\{\omega\})}{1-\alpha} \text{ for all } \omega \in \Omega\right\}, \tag{5}$$

and the probability measures $\mathbb{Q}^{\sigma,\upsilon}$ *for* $\sigma \in \Pi(\Omega)$*, as defined in (4), are given by:*

$$\mathbb{Q}^{\sigma,\upsilon}(\{\omega_{\sigma(j)}\}) = \min\left\{\frac{\sum_{k=j}^{|\Omega|}\mathbb{P}(\{\omega_{\sigma(k)}\})}{1-\alpha}, 1\right\} - \min\left\{\frac{\sum_{k=j+1}^{|\Omega|}\mathbb{P}(\{\omega_{\sigma(k)}\})}{1-\alpha}, 1\right\}, \text{ for all } j \in \{1, \ldots, |\Omega|\}. \tag{6}$$

Now suppose further that $\Omega = \{\omega_1, \omega_2, \omega_3\}$ *with physical probability measure* $\mathbb{P} = \left(\frac{1}{20}, \frac{9}{20}, \frac{1}{2}\right)$*, and suppose that the confidence level* α *in Expected Shortfall is* 90%*, i.e.,* $\rho = \rho_{0.9}^{ES}$*. Then, using (6) yields the outcomes of* $\mathbb{Q}^{\sigma,\upsilon}(\omega_1), \mathbb{Q}^{\sigma,\upsilon}(\omega_2)$*, and* $\mathbb{Q}^{\sigma,\upsilon}(\omega_3)$ *for all* $\sigma \in \Pi(\Omega)$ *as displayed in Table 1.*

Combined with Proposition 2.3, this implies that a finite generating probability measure set for

[8] Expected Shortfall is a member of the broader class of *distortion risk measures* (Wang, 1996). A distortion risk measures is given by $\rho(X) = \int_0^{\infty} g(\mathbb{P}(X > x))dx + \int_{-\infty}^{0}(g(\mathbb{P}(X > x)) - 1)dx$, for some "distortion function" g that satisfies $g(0) = 0$ and $g(1) = 1$ and is continuous, concave and increasing. For any given distortion function g, the set Q^{υ} from (3) with $\upsilon(A) = 1 - g(1-\mathbb{P}(A))$, for all $A \in \mathcal{F}$, is a generating probability measure set for the corresponding distortion risk measure (Tsanakas, 2004).

$(\sigma(1),\, \sigma(2),\, \sigma(3))$	$\mathbb{Q}^{\sigma}(\{\omega_1\})$	$\mathbb{Q}^{\sigma}(\{\omega_2\})$	$\mathbb{Q}^{\sigma}(\{\omega_3\})$
(1, 2, 3)	0	0	1
(1, 3, 2)	0	1	0
(2, 1, 3)	0	0	1
(2, 3, 1)	$\frac{1}{2}$	0	$\frac{1}{2}$
(3, 2, 1)	$\frac{1}{2}$	$\frac{1}{2}$	0
(3, 1, 2)	0	1	0

Table 1: The construction of the finite generating probability measure set $\mathbb{Q}^{\sigma,\upsilon}$ corresponding to Example 2.4.

$\rho^{ES}_{0.9}$ *is given by* $Q^{FG,\upsilon} = \{\mathbb{Q}_1, \mathbb{Q}_2, \mathbb{Q}_3, \mathbb{Q}_4\}$*, with* $\mathbb{Q}_1 = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$*,* $\mathbb{Q}_2 = \left(\frac{1}{2}, 0, \frac{1}{2}\right)$*,* $\mathbb{Q}_3 = (0, 0, 1)$*, and* $\mathbb{Q}_4 = (0, 1, 0)$*. Note that the elements of* $Q^{FG,\upsilon}$ *are the extreme points of the set in (5).* $\triangledown$

Throughout the remainder of the paper, we fix for a given finitely generated risk measure ρ a finite generating probability measure set, which we denote $Q(\rho)$. None of our results depend on the choice of $Q(\rho)$.

2.2 Risk capital allocation problems and the Aumann-Shapley value

In this section, we recall some definitions and key results regarding game-theoretic approaches to risk capital allocation. Readers familiar with this literature can skip this section.

We consider financial institutions that consist of multiple divisions that each face risk. The divisions are indexed $i \in N$. The risk of division i is represented by a random variable X_i on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Once the amount of risk capital that needs to be withheld for the aggregate risk (i.e., $\rho\left(\sum_{i\in N} X_i\right)$) is determined, the firm needs to decide how to allocate this risk capital to the divisions; it needs to determine amounts of risk capital $(a_i)_{i\in N} \in \mathbb{R}^N$ that satisfy

$$\sum_{i\in N} a_i = \rho\left(\sum_{i\in N} X_i\right). \tag{7}$$

Unless risks are perfectly correlated, diversification benefits typically imply that the risk capital that needs to be withheld at aggregate level is (weakly) lower than the sum of the risk capitals that would need to be held for each division if that division was on its own. Indeed, the *Subadditivity* property of coherent risk measures implies that $\rho\left(\sum_{i\in N} X_i\right) \leq \sum_{i\in N} \rho(X_i)$. Whenever the inequality is strict, there is a diversification benefit from pooling the risks. The goal is then to find an allocation $(a_i)_{i\in N} \in \mathbb{R}^N$ that satisfies (7) and allocates the diversification benefit in a "fair" way to the divisions. We now formally define risk capital allocation problems and risk capital allocations.

Definition 2.5 *Risk capital allocation problems, risk capital allocations, and risk capital allocation rules are defined as follows:*

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

(ii) *A vector* $(a_i)_{i \in N} \in \mathbb{R}^N$ *is a risk capital allocation for* $R = ((X_i)_{i \in N}, \rho) \in \mathcal{R}$ *if and only if* $(a_i)_{i \in N}$ *satisfies (7).*

(iii) *Let* $\widetilde{\mathcal{R}} \subseteq \mathcal{R}$ *be a (sub)domain of risk capital allocation problems. A risk capital allocation rule on* $\widetilde{\mathcal{R}}$ *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* $K(R) \in \mathbb{R}^N$*.*

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, for any given risk capital allocation problem $R = ((X_i)_{i \in N}, \rho)$, he defines the corresponding *risk capital function* $r : \mathbb{R}^N_+ \to \mathbb{R}$ as follows

$$r(\lambda) = \rho\left(\sum_{i \in N} \lambda_i \cdot X_i\right), \qquad \text{for all } \lambda \in \mathbb{R}^N_+. \tag{8}$$

If $\lambda \in [0,1]^N$, $r(\lambda)$ is the amount of risk capital that would need to be withheld for a subportfolio consisting of a fraction λ_i of the risk of division i, for all $i \in N$, if that subportfolio were separated from the rest. The fuzzy core (Aubin, 1979) of the corresponding risk capital function r is given by:

$$FCore(R) = \left\{(a_i)_{i \in N} \in \mathbb{R}^N : \sum_{i \in N} \lambda_i a_i \leq r(\lambda) \text{ for all } \lambda \in [0,1]^N, \sum_{i \in N} a_i = r(e_N)\right\}, \tag{9}$$

where e_N is the unit vector in $\mathbb{R}^N$, and so $r(e_N) = \rho\left(\sum_{i \in N} X_i\right)$. For an allocation in the fuzzy core, the risk capital allocated to a subportfolio consisting of a fraction λ_i of the risk of division i, for all $i \in N$ (which is $\sum_{i \in N} \lambda_i a_i$) is weakly lower than the risk capital that would need to be withheld for this subportfolio if it was on its own (which is $r(\lambda)$). Hence, no subportfolio of fractional risks has an incentive to split off. Therefore, the fuzzy core condition can be seen as a stability condition (see, e.g., also Denault, 2001; Tsanakas and Barnett, 2003).

Denault (2001) focuses on the case where: (i) the risk measure ρ is coherent and satisfies *Comonotonic Additivity*, and, (ii) the corresponding risk capital allocation function r in (8) is partially differentiable at $\lambda = e_N$. He then shows that the fuzzy core in (9) is single-valued and that the unique element is the *Aumann-Shapley value*. Let $\mathcal{R}' \subset \mathcal{R}$ be the set of risk capital allocation problems for which r is partially differentiable at $\lambda = e_N$. Then, the *Aumann-Shapley* allocation rule, denoted $K^{AS} : \mathcal{R}' \to \mathbb{R}^N$, is given by (see Denault, 2001):[9]

$$K_i^{AS}(R) = \frac{\partial r}{\partial \lambda_i}(e_N), \qquad \text{for all } i \in N. \tag{10}$$

Hence, the Aumann-Shapley value is the gradient of the risk capital function r evaluated in the vector e_N of full participation.

In addition to the desirable stability property, financial and economic arguments in favor of the Aumann-Shapley value are provided by, e.g., Tasche (1999) and Myers and Read (2001). A main drawback of the Aumann-Shapley value, however, is that it requires partial differentiability of the risk capital function at the level of full participation, i.e., at $\lambda = e_N$. More generally, for risk capital allocation problems $R = ((X_i)_{i \in N}, \rho) \in \mathcal{R}$ for which the risk capital function is not necessarily partially differentiable at $\lambda = e_N$, it follows immediately from Proposition 4 in Aubin (1979) that the fuzzy core is given by:

$$FCore(R) = \text{conv}\{(E_{\mathbb{Q}}[X_i])_{i \in N} : \mathbb{Q} \in Q^*(\rho)\}, \tag{11}$$

where *conv* denotes the convex hull operator and where $Q^*(\rho)$ is the set of *worst case probability measures* for the aggregate risk $\sum_{i \in N} X_i$, i.e.,[10]

$$Q^*(\rho) = \left\{ \mathbb{Q} \in Q(\rho) : r(e_N) = E_{\mathbb{Q}} \left[\sum_{i \in N} X_i \right] \right\}. \tag{12}$$

This fuzzy core in (11) is non-empty, convex and compact. Now recall that the Aumann-Shapley value, if it exists, is the unique element of the fuzzy core (Aubin, 1981). Combined with (11), this implies that if the Aumann-Shapley value exists, it holds that $(E_{\mathbb{Q}}[X_i])_{i \in N} = (E_{\tilde{\mathbb{Q}}}[X_i])_{i \in N}$ for all

[9] More generally, the Aumann-Shapley value of a fuzzy game $r : [0,1]^N \to \mathbb{R}$ is given by $(a_i)_{i \in N}$ such that $a_i = \int_0^1 \frac{\partial r}{\partial \lambda_i}(\gamma \cdot e_N) d\gamma$, for all $i \in N$, provided that these integrals exist. Due to *Positive Homogeneity* of the risk measure ρ, this integral expression simplifies to (10) in the case of risk allocation problems.

[10] Recall that $r(e_N) = \rho\left(\sum_{i \in N} X_i\right)$. Hence, if $r(e_N) = E_{\mathbb{Q}}\left[\sum_{i \in N} X_i\right]$ for some $\mathbb{Q} \in Q(\rho)$, then the maximum in (2) is attained in $\mathbb{Q}$. Therefore, $\mathbb{Q}$ is referred to as a *worst case probability measure* for the risk $\sum_{i \in N} X_i$.

$\mathbb{Q}, \widetilde{\mathbb{Q}} \in Q^*(\rho)$, and the Aumann-Shapley value is given by:

$$K_i^{AS}(R) = E_{\mathbb{Q}}[X_i], \qquad \text{for all } i \in N, \text{ for any } \mathbb{Q} \in Q^*(\rho). \tag{13}$$

If there exist two worst case probability measures $\mathbb{Q}, \widetilde{\mathbb{Q}} \in Q^*(\rho)$ for the aggregate risk such that $(E_{\mathbb{Q}}[X_i])_{i \in N} \neq (E_{\widetilde{\mathbb{Q}}}[X_i])_{i \in N}$, the Aumann-Shapley value does not exist.

In the next section we propose an allocation rule that can be seen as a generalization of the Aumann-Shapley value, and that is well-defined also for allocation problems $R \in \mathcal{R}$ for which the function r is not partially differentiable at $\lambda = e_N$. The new allocation rule selects a fuzzy core element.

3 The Weighted Aumann-Shapley value

In this section we introduce a generalization of the Aumann-Shapley value. The rule that we propose is inspired by the idea underlying the *Shapley value* (Shapley, 1953) for non-fuzzy cooperative games. In Subsection 3.1, we first discuss some properties of the risk capital function r. Then, in Subsection 3.2 we discretize the participation levels of divisions by considering a finite grid of participation levels, and determine the corresponding path-based allocations. The average of the corresponding risk capital allocations over all possible paths is also a risk capital allocation.[11]

In Subsection 3.3, we show that when the grid size converges to zero, this average converges as well. The allocation rule that we propose in this paper equals this asymptotic value. We will refer to this value as the *Weighted Aumann-Shapley value*.

3.1 Properties of the risk capital allocation function

In this subsection we show that the risk capital function r as defined in (8) is piecewise linear and almost everywhere partially differentiable.

Definition 3.1 *Let $R = ((X_i)_{i \in N}, \rho) \in \mathcal{R}$. For all $\mathbb{Q} \in Q(\rho)$:*

(i) *The function $f_{\mathbb{Q}} : \mathbb{R}^N \to \mathbb{R}$ is defined as follows:*

$$f_{\mathbb{Q}}(\lambda) := \sum_{i \in N} \lambda_i \cdot E_{\mathbb{Q}}[X_i], \qquad \textit{for all } \lambda \in \mathbb{R}^N. \tag{14}$$

[11]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. They consider the units of production goods as fixed, i.e., they do not consider convergence by taking infinitely small fractions of such goods.

(ii) *The set* $A_{\mathbb{Q}} \subseteq [0,1]^N$ *is defined as follows:*

$$A_{\mathbb{Q}} := \left\{ \lambda \in [0,1]^N : r(\lambda) = f_{\mathbb{Q}}(\lambda) \right\}. \tag{15}$$

For all $\mathbb{Q} \in Q(\rho)$, the set $A_{\mathbb{Q}}$ consists of participation profiles λ for which $\mathbb{Q}$ is a *worst case probability measure* for the risk $\sum_{i \in N} \lambda_i \cdot X_i$.[12] It is straightforward to show that $A_{\mathbb{Q}}$ is a closed and convex polytope (in fact, a pointed cone). The participation profile $\lambda = e_{\emptyset}$, where $e_{\emptyset}$ is the zero vector in $\mathbb{R}^N$ is an element of $A_{\mathbb{Q}}$ for all $\mathbb{Q} \in Q(\rho)$. The following result follows immediately from (2).

Proposition 3.2 *For all* $R = ((X_i)_{i \in N}, \rho) \in \mathcal{R}$*, the corresponding risk capital function* r *from (8) is piecewise linear on* $[0,1]^N$*. Specifically, there exist* $\mathbb{Q}_1, \ldots, \mathbb{Q}_p \in Q(\rho)$ *with* $p \leq |Q(\rho)|$ *such that:*

$$\bigcup_{m=1}^{p} A_{\mathbb{Q}_m} = [0,1]^N. \tag{16}$$

In the remainder of the paper we will without loss of generality assume that $\mathbb{Q}_1, \ldots, \mathbb{Q}_p \in Q(\rho)$ are chosen and ranked such that:

- $A_{\mathbb{Q}_i} \neq A_{\mathbb{Q}_j}$ for all $i \neq j$;
- $e_N \in A_{\mathbb{Q}_m}$ for $m \in \{1, \ldots, p^*\}$ and $e_N \notin A_{\mathbb{Q}_m}$ for $m \in \{p^*+1, \ldots, p\}$, for some $p^* \leq p$.

In the next example we illustrate Proposition 3.2.

Example 3.3 *Consider again the setting from Example 2.4 with* $\rho = \rho_{0.9}^{ES}$*, and recall the corresponding finite generating probability measure set* $Q^{FG,v} = \{\mathbb{Q}_1, \ldots, \mathbb{Q}_4\}$ *derived in that example. Now let* $N = \{1, 2\}$ *and suppose the two risks are given by:*

$$X_1 = \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 6 \\ 2 \\ 0 \end{bmatrix}.$$

It follows from (2) with $Q = Q^{FG,v}$ *combined with (14) and (15) that:*

$$A_{\mathbb{Q}_1} = \left\{ \lambda \in [0,1]^N : r(\lambda) = f_{\mathbb{Q}_1}(\lambda) \right\}$$

[12] If $r(\lambda) = f_{\mathbb{Q}}(\lambda)$ for some $\mathbb{Q} \in Q(\rho)$, then $\rho\left(\sum_{i \in N} \lambda_i \cdot X_i\right) = E_{\mathbb{Q}}\left(\sum_{i \in N} \lambda_i \cdot X_i\right)$, and so the maximum in (2) is attained in $\mathbb{Q}$.

$$= \left\{\lambda \in [0,1]^N : f_{\mathbb{Q}_1}(\lambda) \geq f_{\mathbb{Q}_j}(\lambda) \text{ for all } j \in \{2,3,4\}\right\}$$
$$= \left\{\lambda \in [0,1]^N : \lambda_1 + 4\lambda_2 \geq \max\{2\lambda_1 + 3\lambda_2, 4\lambda_1, 2\lambda_1 + 2\lambda_2\}\right\}$$
$$= \left\{\lambda \in [0,1]^N : \lambda_1 \leq \lambda_2\right\}.$$

Likewise, we find $A_{\mathbb{Q}_2} = \left\{\lambda \in [0,1]^N : 3\lambda_2 \geq 2\lambda_1 \geq 2\lambda_2, \right\}$, $A_{\mathbb{Q}_3} = \left\{\lambda \in [0,1]^N : 2\lambda_1 \leq 3\lambda_2\right\}$, *and* $A_{\mathbb{Q}_4} = \{e_\emptyset\}$. *Because* $A_{\mathbb{Q}_4} \subset A_{\mathbb{Q}_m}$ *for* $m \in \{1,2,3\}$, *we can without loss of generality drop* $\mathbb{Q}_4$, *and so (16) is satisfied with* $p = 3$. *Moreover, because* $e_N = (1,1) \in A_{\mathbb{Q}_m}$ *for* $m \in \{1,2\}$ *and* $e_N \notin A_{\mathbb{Q}_m}$ *for* $m = 3$, *it holds that* $p^* = 2$. *Note that non-differentiability of* r *in* e_N *is caused by the firm having the same loss in states* ω_2 *and* ω_3. ▽

We conclude this subsection by showing that the risk capital function r is almost everywhere partially differentiable. If for some given λ, there exists a unique $m \in \{1, \dots, p\}$ such that $\lambda \in A_{\mathbb{Q}_m}$, then 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

$$\frac{\partial r}{\partial \lambda_i}(\lambda) = E_{\mathbb{Q}_m}[X_i], \qquad \text{for all } i \in N. \tag{17}$$

Hence, uniqueness of $m \in \{1, \dots, p\}$ such that $\lambda \in A_{\mathbb{Q}_m}$ is a sufficient condition for the risk capital function r to be partially differentiable in λ. Formally, for all $R \in \mathcal{R}$, we let the set $L(R)$ be given by

$$L(R) = \left\{\lambda \in [0,1]^N : \text{there exists a unique } m \in \{1, \dots, p\} \text{ such that } \lambda \in A_{\mathbb{Q}_m}\right\}. \tag{18}$$

The following result will be relevant when we propose a new risk capital allocation rule later in this section.

Proposition 3.4 *For all* $R \in \mathcal{R}$, *it holds that:*

(i) *The risk capital function* r *is partially differentiable in* λ *if* $\lambda \in L(R)$.

(ii) *The risk capital function* r *is almost everywhere partially differentiable. The set of participation profiles* λ *where the risk capital function* r *is not partially differentiable is a subset of the union of a finite number of hyperplanes passing through* $\lambda = e_\emptyset$.

3.2 Path based allocation rules

In this subsection, we discuss *path based allocation rules*, as introduced in Wang (1999). It extends the idea of the marginal vectors of the Shapley value (Shapley, 1953) by allowing for divisions to participate in the risk capital allocation for only a fraction of their risk. The fractions are represented by the participation levels $\lambda \in [0,1]^N$. We first describe a path based 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

$$G^n = \left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\right\}^N. \tag{19}$$

The starting point on grid G^n is the participation profile $\lambda = e_\emptyset$ 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, where e_i is the i-th unit vector in $\mathbb{R}^N$. 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 total risk capital (which is $\rho\left(\sum_{i\in N} X_i\right) = r(e_N)$) has been allocated to the divisions by then.

Formally, a path is defined as follows.

Definition 3.5 *Let $n \in \mathbb{N}$ be given. A path on the grid G^n is a map $P : \{0,1,2,\ldots,|N|n\} \to G^n$ satisfying:*

(i) $P(0) = e_\emptyset$ and $P(|N|n) = e_N$;

(ii) for every $k \in \{0,\ldots,|N|n-1\}$, there exists a unique $i \in N$ such that

$$P(k+1) - P(k) = \frac{1}{n} \cdot e_i. \tag{20}$$

This unique division i will be denoted $i(P,k)$.

An example of a path P on the grid G^n is given in Figure 1. We denote the collection of all paths on the grid G^n by $\mathcal{P}^n$.

Proposition 3.6 *Let $P \in \mathcal{P}^n$ be a path on the grid G^n. Then, the map $K^{path,P} : \mathcal{R} \to \mathbb{R}^N$ defined*

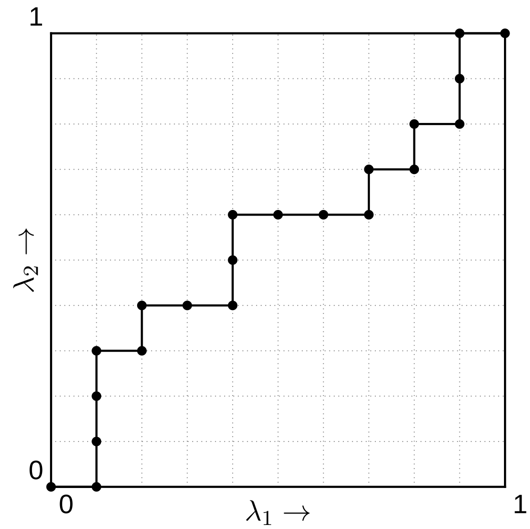

Figure 1: Example of a path $P \in \mathcal{P}^n$ for $|N| = 2$ with $n = 10$. We connected succeeding elements of the path as illustration.

by

$$K^{path,P}(R) = \sum_{k=0}^{|N|n-1} [r(P(k+1)) - r(P(k))] \cdot e_{i(P,k)}, \qquad \text{for all } R \in \mathcal{R}, \tag{21}$$

with risk capital function r as defined in (8), is an allocation rule on $\mathcal{R}$.

We will refer to $K^{path,P}$ as a *path based allocation rule.*

Following Moulin (1995) and Sprumont (2005),we then consider the allocation rule that is given by the average over all paths of the corresponding path based risk capital allocations. Formally, this allocation rule is defined as follows.

Definition 3.7 *Let $n \in \mathbb{N}$ be given. Then, $K^{avg,n} : \mathcal{R} \rightarrow \mathbb{R}^N$ is defined by*

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

where $K^{path,P} : \mathcal{R} \rightarrow \mathbb{R}^N$ for a given path $P \in \mathcal{P}^n$ is the allocation rule defined in (21).

If $n = 1$, the corresponding allocation rule $K^{avg,1}(R)$ equals the Shapley value (Shapley, 1953). In the following subsection, we study its asymptotic behavior when we let n go to infinity.

3.3 The Weighted Aumann-Shapley value

If the Aumann-Shapley value in (10) exists, i.e., if the risk capital function r is partially differentiable along the diagonal, it can be approximated by a path based allocation by using a very small grid and a path close to the diagonal (see, e.g., Aumann and Shapley, 1974, for functions that are globally partially differentiable). However, as mentioned before, the Aumann-Shapley value is not well-defined if the risk capital function r is not partially differentiable along the diagonal. We therefore now propose a generalization that is well-defined even if r is not partially differentiable along the diagonal. The rule that we propose arises from letting the grid size become infinitely small and taking the limit of the corresponding allocations $K^{avg,n}(R)$. In this subsection, we show that $\lim_{n\to\infty} K^{avg,n}(R)$ exists for all $R \in \mathcal{R}$. Our focus is on the intuition behind the result. Detailed proofs are available in the Online Appendix.

Proposition 3.8 *Let $R \in \mathcal{R}$ and $n \in \mathbb{N}$. Let $\overline{P}$ be a path that is randomly selected from $\mathcal{P}^n$ according to the discrete uniform distribution on $\mathcal{P}^n$. Then, it holds that*

$$K_i^{avg,n}(R) = \sum_{\lambda\in G^n:\lambda_i<1} \mathbb{P}\left(\lambda \in \overline{P}\right) \mathbb{P}\left(\lambda + (1/n)\cdot e_i \in \overline{P} | \lambda \in \overline{P}\right) \left[r\left(\lambda + (1/n)\cdot e_i\right) - r(\lambda)\right], \quad (22)$$

for all $i \in N$, where the risk capital function r is as defined in (8).

Hence, the amount of risk capital allocated to division i, $K_i^{avg,n}(R)$, is equal to the sum over all participation profiles $\lambda \in G^n$ with $\lambda_i < 1$ for all i, of the marginal contribution to the risk capital from moving from λ to $\lambda + (1/n)\cdot e_i$ (which is equal to $r(\lambda + (1/n)\cdot e_i) - r(\lambda)$) multiplied by the probability that both λ and $\lambda + (1/n)\cdot e_i$ are on a randomly selected path.

In the proofs in the Online Appendix we show that the following subsets of participation profiles have a negligible contribution to (22) if n becomes sufficiently large:

- **Participation profiles that lie in a neighbourhood of $e_\emptyset$ or e_N.** We show in Lemma C.13 of the Online Appendix that for ε sufficiently small, participation profiles $\lambda \notin G^n_\varepsilon := \left\{\lambda \in G^n : \varepsilon \leq \bar{\lambda} \leq 1-\varepsilon\right\}$ with $\bar{\lambda} = \frac{1}{|N|}\sum_{i\in N}\lambda_i$, have a negligible contribution to $K^{avg,n}(R)$ if n becomes sufficiently large.

- **Participation profiles that are not sufficiently close to the diagonal.** We show in Lemma C.14 of the Online Appendix that participation profiles $\lambda \notin D(n) := \{\lambda \in [0,1]^N :$

$\|\lambda - \bar{\lambda}e_N\| < n^{-\frac{1}{2}+\frac{1}{8|N|}}\}$[13] have a negligible contribution to $K^{avg,n}(R)$ if n becomes sufficiently large.[14]

- **Participation profiles sufficiently close to the set $[0,1]^N \backslash L(R)$.** Let $B(n)$ be the set of participation profiles in an $\frac{1}{n}$-environment of participation profiles in $[0,1]^N \backslash L(R)$, i.e.,

$$B(n) := \left\{ \lambda \in [0,1]^N : \exists \hat{\lambda} \in [0,1]^N \backslash L(R) : \|\lambda - \hat{\lambda}\| < \frac{1}{n} \right\}, \tag{23}$$

where $L(R)$ is as defined in (18). Recall that $L(R)$ contains all participation profiles where r is differentiable (Proposition 3.4(i)). We show in Lemma C.18 of the Online Appendix that all participation profiles in $\lambda \in B(n)$ have a negligible aggregate contribution to $K^{avg,n}(R)$ for sufficiently large n.

We illustrate the sets G_ε^n and $D(n)$ in Figure 2, for the case where there are two divisions ($|N| = 2$).

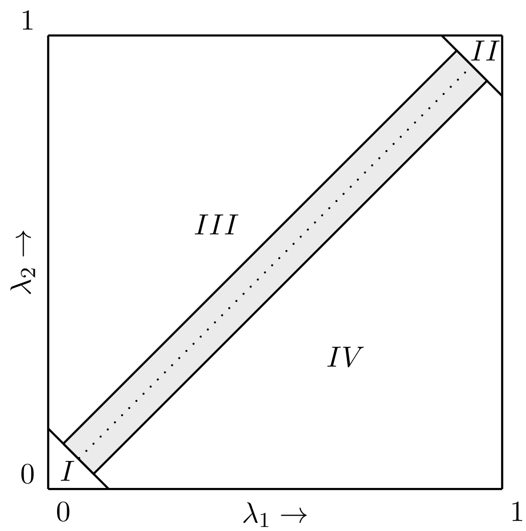

Figure 2: The shaded set is a set of participation profiles with non-negligible aggregate contribution to the Weighted Aumann-Shapley value in case $|N| = 2$. Here, $I \cup II = [0,1]^N \backslash G_\varepsilon$ and $III \cup IV = G_\varepsilon \backslash D(n)$ for an arbitrary choice of $\varepsilon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

Let us provide some intuition here for the case $|N| = 2$. Participation profiles in $I \cup II$ obviously have a small contribution if ε is small. But why is the same true for the profiles in $III \cup IV$ that nearly covers the full participation profile set $[0,1]^N$? The point here is that for a fixed $\varepsilon > 0$ and

[13] Here, $\|\cdot\|$ is the Euclidean norm, i.e., $\|\lambda\| = \sqrt{\sum_{i \in N} \lambda_i^2}$.

[14] Theorem 3.9 can be proven by using a diagonal width $n^{-\frac{1}{2}+\delta}$ for some $\delta \in \left(0, \frac{1}{2(|N|+2)}\right)$. The proofs are based on $\delta = \frac{1}{8|N|}$.

large n the fraction of paths crossing $III \cup IV$ is close to 0, or, stated differently, the fraction of paths that stay within the neighborhood of the diagonal is close to 1. The reason for this is the following. A path from $e_\emptyset$ to e_N can be seen as a sequence of $2n$ steps of which n steps move to the right and n steps move up. So selecting an arbitrary path from $\mathcal{P}^n$ corresponds to a sequence of $2n$ steps $X_1, X_2, \ldots, X_{2n}$ which are random variables taking the value "move to the right" and "move up". These random variables are not independent as a path has to end in e_N: if after k steps more "moves to the right" occurred than "moves up" then the probability of a next "move up" increases. In other words, if the path deviates from the diagonal the probability that the path moves back to the diagonal increases. There is reversion to the diagonal. Now consider as a benchmark a random walk $Y_1, Y_2, \ldots, Y_{2n}$ where this reversion effect to the diagonal is absent: the random variables are independent and identically distributed and take value "move to the right" and "move up" both with probability 1/2. For sure, such a random walk does not need to end up in e_N. The number of moves to the right after the first k steps follows the binomial distribution $Bin(k, 1/2)$ with mean $k/2$ (which corresponds to a position on the diagonal) and standard deviation $\sqrt{k}/2$. Such a standard deviation corresponds to a position at a distance of $\sqrt{2}(\sqrt{k}/2) \cdot (1/n)$ from the diagonal. As $k \leq 2n$ this position is definitely at most $1/\sqrt{n}$ away from the diagonal. The boundaries of the set $D(n)$ are at a distance of $n^{-1/2+1/16}$ from the diagonal. Crossing this boundary corresponds to a random walk that moves at least $n^{1/16}$ standard deviations away from the mean. According to the inequality of Chebychev this happens with probability converging to 0 if n becomes large.

Combined, our results imply that only participation profiles in the set $\Lambda^{NN}(\varepsilon, n) := G^n_\varepsilon \cap D(n) \backslash B(n)$ have a non-negligible contribution to $K^{avg,n}(R)$ if n becomes sufficiently large. In Lemma C.17 of the Online Appendix, we show that for all $\varepsilon > 0$:

$$\Lambda^{NN}(\varepsilon, n) \subset \bigcup_{m \in \{1,\ldots,p^*\}} A_{\mathbb{Q}_m}, \qquad \text{for large } n. \tag{24}$$

Hence, for participation profiles λ that have non-negligible contribution to $K^{avg,n}(R)$, there exists an $m \in \{1, \ldots, p^*\}$ such that $\lambda \in A_{\mathbb{Q}_m}$. This allows us to determine the marginal contribution $r(\lambda + (1/n) \cdot e_i) - r(\lambda)$ in (22) for all these participation profiles. Indeed, because r is partially differentiable in an $\frac{1}{n}$-environment of any $\lambda \notin B(n)$ for n sufficiently large, it follows from (17) that for all $\lambda \in A_{\mathbb{Q}_m} \backslash B(n)$, the marginal contribution is given by:

$$r(\lambda + (1/n) \cdot e_i) - r(\lambda) = \frac{1}{n} \frac{\partial}{\partial \lambda_i} r(\lambda) = \frac{1}{n} E_{\mathbb{Q}_m}[X_i], \text{ for all } i \in N, \tag{25}$$

for n sufficiently large. Combining the fact that only participation profiles in $\Lambda^{NN}(\varepsilon, n)$ have a non-negligible contribution to (22) if n becomes large with (24) and (25) suggests that $\lim_{n\to\infty} K_i^{avg,n}(R)$ exists and is of the form $\sum_{m=1}^{p^*} \phi_m E_{\mathbb{Q}_m}[X_i]$, for some weights ϕ_m. The following theorem provides the formal result and gives a closed form expression for the weights ϕ_m.

Theorem 3.9 *For all $R \in \mathcal{R}$, it holds that $\lim_{n\to\infty} K^{avg,n}(R)$ exists and is given by*

$$\lim_{n\to\infty} K_i^{avg,n}(R) = \sum_{m=1}^{p^*} \phi_m E_{\mathbb{Q}_m}[X_i], \qquad \text{for all } i \in N, \tag{26}$$

with

$$\phi_m = \frac{\mu(S_m)}{\mu(S)}, \tag{27}$$

where μ is the hypersurface measure[15] *and*

$$S = \left\{ z \in \mathbb{R}^N : \sum_{i\in N} z_i = 0, \|z\| = 1 \right\} \tag{28}$$

$$S_m = \left\{ z \in S : f_{\mathbb{Q}_m}(z) = \max_{\ell\in\{1,\ldots,p^*\}} f_{\mathbb{Q}_\ell}(z) \right\}, \text{ for all } m \in \{1,\ldots,p^*\}. \tag{29}$$

The set S is a set of normalized directions perpendicular to the diagonal. The set S_m consists of all directions $z \in S$ that will bring us in the set $A_{\mathbb{Q}_m}$ if we start in e_N and move an infinitesimal amount in the direction z.

Note that

$$\sum_{m=1}^{p^*} \phi_m = 1. \tag{30}$$

If $p^* > 1$ it follows from (13) that for all $m \in \{1,\ldots p^*\}$, the allocation $(E_{\mathbb{Q}_m}[X_i])_{i\in N}$ is the Aumann-Shapley value of the risk capital allocation problem that would arise if the risk capital function r was modified marginally such that $r(\lambda) = \sum_{i\in N} \lambda_i E_{\mathbb{Q}_m}[X_i]$ in a small neighbourhood around $\lambda = e_N$ (so that $p^* = 1$ for the modified problem). Hence, the allocation rule $\lim_{n\to\infty} K^{avg,n}(R)$ is a weighted average of Aumann-Shapley values of "nearby" differentiable allocation problems. Therefore, we refer to $\lim_{n\to\infty} K^{avg,n}(R)$ as the *Weighted Aumann-Shapley value.*

[15] The set S is an $(|N|-2)$-dimensional sphere in $\mathbb{R}^N$. In case $|N| = 3$ the set S is a circle and μ represents arc length, in case $|N| = 4$ the set S is a sphere and μ represents surface area, etcetera.

Definition 3.10 *The allocation rule* $K^{WAS} : \mathcal{R} \to \mathbb{R}^N$ *is given by:*

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

In the next example, we determine $K^{WAS}(R)$ for a given risk capital allocation problem and we illustrate the geometric interpretation of the corresponding weights ϕ_m from (27).

Example 3.11 *We consider a firm with three divisions,* $N = \{1, 2, 3\}$*, and five possible states of the world, i.e.,* $\Omega = \{\omega_1, \ldots, \omega_5\}$ *with equal probabilities* $\mathbb{P}(\{\omega\}) = \frac{1}{5}$ *for all* $\omega \in \Omega$*. We consider the case where the regulator does not allow insolvency, i.e., the amount of risk capital needs to be sufficient to cover the highest possible loss. Hence, the risk measure is defined as follows,*

$$\rho\,(X) = \max_{j\in\{1,2,3,4,5\}} \{X(\omega_j)\}\,,$$

for any $X \in \mathbb{R}^\Omega$*. This risk measure satisfies (2) with* $Q = \{\mathbb{Q}_1, \ldots, \mathbb{Q}_5\}$*, where* $\mathbb{Q}_m(\{\omega\}) = 1$ *if* $\omega = \omega_m$ *and* $\mathbb{Q}_m(\{\omega\}) = 0$ *otherwise, for* $m \in \{1, 2, 3, 4, 5\}$*. Now suppose the three risks are given by*

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

Here, the j*-the element of vector* X_i *represents the loss in state* ω_j*, for* $j \in \{1, 2, 3, 4, 5\}$*. The corresponding fuzzy risk capital game* r *as defined in (8) is given by*

$$r(\lambda) = \rho\left(\sum_{i\in N} \lambda_i \cdot X_i\right) = \max_{j\in\{1,2,3,4,5\}} \left\{\sum_{i\in N} \lambda_i \cdot X_i(\omega_j)\right\}, \text{ for all } \lambda \in [0,1]^N. \tag{31}$$

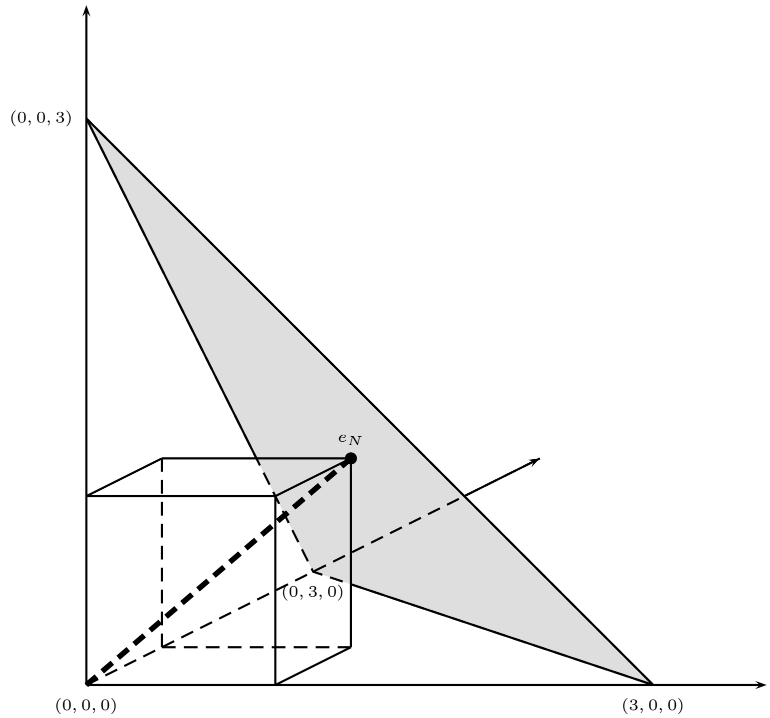

Figure 3: The cube $[0,1]^N$ and its diagonal. The shaded area is the simplex $T := \left\{\lambda \in \mathbb{R}_+^N : \bar{\lambda} = 1\right\}$.

This yields

$$r(\lambda) = \begin{cases} \lambda_1 & \textit{if } \lambda_1 \geq \lambda_2 \geq \lambda_3, \\ \lambda_3 & \textit{if } \lambda_3 \geq \lambda_2 \geq \lambda_1, \\ \lambda_2 & \textit{if } \lambda_2 \geq \lambda_1, \lambda_2 \geq \lambda_3, \lambda_2 \leq \lambda_1 + \lambda_3, \\ \lambda_1 - \lambda_2 + \lambda_3 & \textit{if } \lambda_1 \geq \lambda_2, \lambda_3 \geq \lambda_2, \\ -\lambda_1 + 2\lambda_2 - \lambda_3 & \textit{if } \lambda_2 \geq \lambda_1 + \lambda_3. \end{cases} \tag{32}$$

It follows immediately from (32) that $r(\lambda)$ is not differentiable at $\lambda = (1,1,1)$, and so the Aumann-Shapley value does not exist. Because $e_N \in A_{\mathbb{Q}_m}$ for $m \in \{1,2,3,4\}$ and $e_N \notin A_{\mathbb{Q}_5}$, it holds that $p^* = 4$. The Aumann-Shapley values of the four differentiable fuzzy games that are "nearby" r at $\lambda = (1,1,1)$ are given by $a^m = (E_{\mathbb{Q}_m}[X_i])_{i \in N}$, for $m \in \{1,2,3,4\}$, i.e.,

$$a^1 = (1,0,0),\ a^2 = (0,0,1),\ a^3 = (0,1,0),\ \textit{and } a^4 = (1,-1,1). \tag{33}$$

To determine $K^{WAS}(R)$ from (26), it remains to determine the weights ϕ_m for $m \in \{1,2,3,4\}$, that should be assigned to these four values. We first consider the simplex T defined by $T := \{\lambda \in \mathbb{R}_+^N : \bar{\lambda} = 1\}$, as displayed in Figure 3. This simplex T is also displayed in Figure 4, together

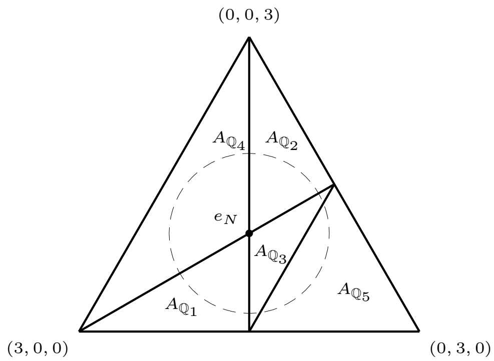

Figure 4: The simplex T and its intersections with the sets $A_{\mathbb{Q}_m}$, $m \in \{1,2,3,4,5\}$, and the set of directions S partitioned in the subsets S_m, $m \in \{1,2,3,4\}$.

with all intersections $T \cap A_{\mathbb{Q}_m}$, $m \in \{1,2,3,4,5\}$. The dashed circle consists of all points $e_N + z$ where directions z are elements of the unit circle S. The subset S_m, $m \in \{1,2,3,4\}$, consists of all directions $z \in S$ that will bring us in the set $A_{\mathbb{Q}_m}$ if we start in e_N and move an infinitesimal amount in direction z. The unit circle S partitioned in the sets S_1, S_2, S_3 and S_4 is displayed in Figure 4 as well. With two dimensions, the measure μ represents arc length and the weights ϕ_m are found by computing the normalized arc lengths of the sets S_m, $m \in \{1,2,3,4\}$. This is equivalent to computing the normalized angles α_m, $m \in \{1,2,3,4\}$. This yields weights $\phi_1 = \frac{1}{6}$, $\phi_2 = \frac{1}{6}$, $\phi_3 = \frac{1}{3}$ and $\phi_4 = \frac{1}{3}$. It then follows from (26) and (33) that

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

4 Properties of the Weighted Aumann-Shapley value

It follows immediately from (11), (26) and (27) that the Weighted Aumann-Shapley value is an element of the fuzzy core. Because the Aumann-Shapley value, if it exists, is the unique element of the fuzzy core (Aubin, 1981), this in turn immediately implies that the Weighted Aumann-Shapley value equals the Aumann-Shapley value if the risk capital function r is differentiable in e_N. Hence, we have the following result.

Proposition 4.1 *It holds that:*

(i) *For all $R \in \mathcal{R}$, it holds that $K^{WAS}(R) \in FCore(R)$.*

(ii) *For all $R \in \mathcal{R}'$, it holds that $K^{WAS}(R) = K^{AS}(R)$.*

If the firm consists of two divisions, the weights ϕ_m of the Weighted Aumann-Shapley value can only take values equal to 0, 1/2, or 1. In that case, the following corollary shows an expression of the Weighted Aumann-Shapley value.

Corollary 4.2 *If $|N| = 2$ and $R \in \mathcal{R}$, it holds that:*

$$K_i^{WAS}(R) = \frac{1}{2}\max\{E_{\mathbb{Q}_m}[X_i] : m \in \{1, \ldots, p^*\}\} + \frac{1}{2}\min\{E_{\mathbb{Q}_m}[X_i] : m \in \{1, \ldots, p^*\}\}, \text{ for all } i \in N.$$

If $|N| = 2$, Corollary 4.2 shows that the Weighted Aumann-Shapley value equals the average of a worst-case Aumann-Shapley value and a best-case Aumann-Shapley value, where the Aumann-Shapley values are taken from the set of Aumann-Shapley values corresponding to "nearby" differentiable allocation problems.

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 $\widetilde{R} = ((\widetilde{X}_i)_{i\in N}, \rho) \in \widetilde{\mathcal{R}}$ where $(\widetilde{X}_i)_{i\in N} = (X_j + c \cdot e_\Omega, X_{-j})$ for some $c \in \mathbb{R}$ and $j \in N$, then $K(\widetilde{R}) = K(R) + c \cdot e_j$.
- *Scale Invariance*: For all $R = ((X_i)_{i\in N}, \rho) \in \widetilde{\mathcal{R}}$, it holds that if $\widetilde{R} = ((\widetilde{X}_i)_{i\in N}, \rho) \in \widetilde{\mathcal{R}}$ where $(\widetilde{X}_i)_{i\in N} = (c \cdot X_i)_{i\in N}$ for some $c > 0$, then $K(\widetilde{R}) = c \cdot K(R)$.
- *Monotonicity*: For all $R \in \widetilde{\mathcal{R}}$ 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 $K(R) \geq 0$.

Denault (2001) shows that the regular Aumann-Shapley value satisfies these three properties on $\mathcal{R}'$. We next show that the Weighted Aumann-Shapley value satisfies these three properties on $\mathcal{R}$.

Theorem 4.3 *The Weighted Aumann-Shapley value satisfies Translation Invariance, Scale Invariance and Monotonicity on $\mathcal{R}$.*

We conclude this section with three remarks.

Remark Because all of our proofs rely only on positive homogeneity and piecewise linearity of the function r, all our results extend to this more general setting by replacing $f_{\mathbb{Q}_m}$ by a more general linear function f_m, for $m \in \{1, \ldots, p\}$, and replacing $E_{\mathbb{Q}_m}[X_i]$ by $\frac{\partial r}{\partial \lambda_i}(\lambda)$ for $\lambda \in [0,1]^N$ such that $r(\lambda) = f_m(\lambda)\}$ for $m \in \{1, \ldots, p\}$. This more general case is analyzed in the working paper version Boonen et al. (2018).

Remark The Weighted Aumann-Shapley value, as presented in (26), can also be computed via the following integral formula:

$$K_i^{WAS}(R) = E\left[\int_0^1 \partial r(\gamma e_N, Y, e_i) d\gamma\right] \tag{34}$$

for every $i \in N$, where r is defined in (8). Here, Y is a random vector in S that is uniformly distributed over S, E is the expectation operator with respect to Y, and

$$\partial r(\gamma e_N, Y, e_i) = \lim_{\varepsilon \downarrow 0} \frac{dr(\gamma e_N, Y + \varepsilon e_i) - dr(\gamma e_N, Y)}{\varepsilon},$$

where $dr(x, y)$ is the directional derivative of r in x in the direction y. In short, in order to compute $K_i^{WAS}(R)$, for every realization y of the random vector Y compute and integrate the directional derivatives of r in the direction e_i on the line segment obtained by shifting the diagonal an infinitesimal amount in the direction y, and finally compute the expectation of these outcomes. We note that for coherent risk measures, *Positive Homogeneity* of the measure implies that the risk capital function r is positive homogeneous (i.e., $r(t\lambda) = tr(\lambda)$ for every $\lambda \in \mathbb{R}_+^N$ and $t > 0$), which implies that $\partial r(\gamma e_N, Y, e_i)$ is independent of γ. Hence, (34) simplifies to

$$K_i^{WAS}(R) = E\left[\partial r(e_N, Y, e_i)\right], \text{ for all } i \in N. \tag{35}$$

The equivalence between (26) and (35) follows from the fact that for any realization y of the random vector Y, it holds that $\partial r(e_N, y, e_i) \in \{E_{\mathbb{Q}_m}[X_i] : m \in \{1, \ldots, p^*\}\}$ and the weights ϕ_m in (26) satisfy $\phi_m = \mathbb{P}(\partial r(e_N, Y, e_i) = E_{\mathbb{Q}_m}[X_i])$ for all $m \in \{1, \ldots, p^*\}$.

Formula (34) is similar to the Mertens value, but differs in the fact that for the Mertens value Y is a random vector in $\mathbb{R}^N$ where the coordinates are independent random variables, each one having the standard Cauchy distribution (Haimanko, 2001). An overview of the Mertens value is given by Neyman (2002). We show in Boonen et al. (2018) that the Weighted Aumann-Shapley value is not identical to the Mertens value.

Remark Our paper focused on the case where the risk measure is finitely generated and the state space is finite. However, the asymptotic approach of Section 3 may be suited for a broader class of risk measures on finite and infinite state spaces for which (34) is well-defined. For example, when the risk capital allocation function r is convex (which is the case for all coherent risk measures), the necessary directional derivatives in (34) exist. As a suggestion for further research, we leave open the characterization of the class of convex risk capital allocation functions for which (34) is well-defined and is equal to the limit of the allocation rules $K^{avg,n}(R)$ as n goes to infinity.

5 Conclusion

This paper considers the allocation problem that arises when the total risk capital withheld by a firm needs to be divided over several portfolios or divisions within the firm. We propose a generalization of the Aumann-Shapley value that is also well-defined if the risk capital allocation function is not partially differentiable at the level of full participation. The allocation rule that we propose is inspired by the Shapley value in a fuzzy setting, but is derived using a much weaker asymptotic approach than the one proposed by Aumann and Shapley (1974), which is not valid for fuzzy games corresponding to risk capital allocation problems. For a given grid on a fuzzy participation set, one can define paths on this grid and for each path one can construct a corresponding path-based allocation rule. The average of these path-based allocation rules is an allocation rule itself. We show that the limit of this average exists when grid size converges to zero. The rule that we propose is equal to this limit.

References

Acerbi, C. and D. Tasche (2002). On the coherence of expected shortfall. *Journal of Banking and Finance 26*, 1487–1503.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9*, 203–228.

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.

Balog, D., T. L. Bátyi, P. Csóka, and M. Pintér (2017). Properties and comparison of risk capital allocation methods. *European Journal of Operational Research 259*, 614–625.

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.

Bogetoft, P., J. L. Hougaard, and A. Smilgins (2016). Applied cost allocation: The DEA-Aumann-Shapley approach. *European Journal of Operational Research 254*, 667–678.

Boonen, T. J., A. De Waegenaere, and H. Norde (2018). The family of weighted Aumann-Shapley values with an application in risk capital allocations. On SSRN: http://ssrn.com/abstract=2181938.

Boonen, T. J., A. Tsanakas, and M. V. Wüthrich (2017). Capital allocation for portfolios with non-linear risk aggregation. *Insurance: Mathematics and Economics 72*, 95–106.

Centrone, F. and E. Rosazza Gianin (2018). Capital allocation á la Aumann-Shapley for non-differentiable risk measures. *European Journal of Operational Research 267*, 667–675.

Cherny, A. and D. Orlov (2011). On two approaches to coherent risk contribution. *Mathematical Finance 21*, 557–571.

Csóka, P., P. J. J. Herings, and L. Kóczy (2009). Stable allocations of risk. *Games and Economic Behavior 67*, 266–276.

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.

Grechuk, B. (2015). The center of a convex set and capital allocation. *European Journal of Operational Research 243*, 628–636.
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.
Homburg, C. and P. Scherpereel (2008). How should the cost of joint risk capital be allocated for performance measurement? *European Journal of Operational Research 187*, 208–227.
Kalkbrener, M. (2005). An axiomatic approach to capital allocation. *Mathematical Finance 15*, 425–437.
Kusuoka, S. (2001). On law invariant coherent risk measures. *Advances in Mathematical Economics 3*, 83–95.
Major, J. A. (2018). Distortion measures and homogeneous financial derivatives. *Insurance: Mathematics and Economics 79*, 82–91.
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. Smorodinsky (2004). Asymptotic values of vector measure games. *Mathematics of Operations Research 29*, 739–775.
Powers, M. R. (2007). Using Aumann-Shapley values to allocate insurance risk: the case of inhomogeneous losses. *North American Actuarial Journal 11*, 113–127.
Shapley, L. (1953). A value for n-person games. In *Contributions to the Theory of Games*, pp. 307–317. Princeton University Press.
Shapley, L. (1971). Cores of convex games. *International Journal of Game Theory 1*, 11–26.
Sprumont, Y. (2005). On the discrete version of the Aumann-Shapley cost-sharing method. *Econometrica 73*, 1693–1712.
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.
Tsanakas, A. (2004). Dynamic capital allocation with distortion risk measures. *Insurance: Mathematics and Economics 35*, 223–243.
Tsanakas, A. and C. Barnett (2003). Risk capital allocation and cooperative pricing of insurance liabilities. *Insurance: Mathematics and Economics 33*, 239–254.
Wang, S. S. (1996). Premium calculation by transforming the layer premium density. *ASTIN Bulletin 26*, 71–92.
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. *Economic Letters 64*, 187–192.

A Proofs of Sections 2.1, 3.1 and 3.2

Proof of Proposition 2.2 Let Q be the generating probability measure set of ρ that is defined in (3), i.e.,

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

where $\upsilon : \mathcal{F} \to \mathbb{R}_+$ is supermodular, $\upsilon(\emptyset) = 0$ and $\upsilon(\Omega) = 1$. 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 convex polytope. Let $\widetilde{Q}$ be the finite collection of extreme points of this convex polytope. Because $\mathbb{Q} \to E_\mathbb{Q}[X]$ is a linear map on Q for every $X \in \mathbb{R}^\Omega$, (1) is a linear programming problem and, therefore, we have

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

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

Proof of Proposition 2.3 The set Q^υ defined in (3) is the core of the Transferable Utility game (Ω, υ), where the state space Ω is now interpreted as a "player" set. Supermodularity of the function υ is equivalent to convexity of the game (Ω, υ) (Shapley, 1971). Moreover, Shapley (1971) shows that the core of a convex game is the convex hull of the marginal vectors. The marginal vectors of the game are the vectors $m^{\sigma,\upsilon} \in \mathbb{R}^\Omega$ with $m^{\sigma,\upsilon}_{\sigma(j)} := \mathbb{Q}^{\sigma,\upsilon}(\omega_{\sigma(j)})$, for all $\sigma \in \Pi(\Omega)$.

Proof of Proposition 3.2 For all $R \in \mathcal{R}$, we have

$$\begin{aligned} r(\lambda) &= \max\left\{E_\mathbb{Q}\left[\sum_{i\in N}\lambda_i X_i\right] : \mathbb{Q} \in Q(\rho)\right\} \\ &= \max\left\{\sum_{i\in N}\lambda_i E_\mathbb{Q}[X_i] : \mathbb{Q} \in Q(\rho)\right\} \\ &= \max\{f_\mathbb{Q}(\lambda) : \mathbb{Q} \in Q(\rho)\}, \end{aligned} \tag{36}$$

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

Proof of Proposition 3.4 (i) Follows directly from the fact that r is the maximum of finitely many linear (hence partially differentiable) functions $f_{\mathbb{Q}_m}, m \in \{1,\dots,p\}$.

We continue with the proof of (ii). We obtain for all $\ell, m \in \{1,\dots,p\}$ that

$$\begin{aligned} A_{\mathbb{Q}_\ell} \cap A_{\mathbb{Q}_m} &= \left\{\lambda \in [0,1]^N : r(\lambda) = f_{\mathbb{Q}_\ell}(\lambda) = f_{\mathbb{Q}_m}(\lambda)\right\} \\ &\subseteq \left\{\lambda \in [0,1]^N : f_{\mathbb{Q}_\ell}(\lambda) = f_{\mathbb{Q}_m}(\lambda)\right\} \\ &= \left\{\lambda \in [0,1]^N : \sum_{i \in N} \lambda_i \left(E_{\mathbb{Q}_\ell}[X_i] - E_{\mathbb{Q}_m}[X_i]\right) = 0\right\}. \end{aligned} \tag{37}$$

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

$$[0,1]^N \backslash L(R) = \bigcup_{\ell, m \in \{1,\dots,p\}: \ell \neq m} A_{\mathbb{Q}_\ell} \cap A_{\mathbb{Q}_m}, \qquad \text{for all } R \in \mathcal{R}. \tag{38}$$

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 3.6 Let $n \in \mathbb{N}$ and $P \in \mathcal{P}^n$. Then, the result follows directly from

$$\sum_{i \in N} K_i^{path,P}(R) = \sum_{i \in N} \sum_{k=0}^{|N|n-1} \left[r(P(k+1)) - r(P(k))\right] \mathbb{1}_{i(P,k)=i} \tag{39}$$

$$= \sum_{k=0}^{|N|n-1} \left[r(P(k+1)) - r(P(k))\right] \sum_{i \in N} \mathbb{1}_{i(P,k)=i} \tag{40}$$

$$= \sum_{k=0}^{|N|n-1} \left[r(P(k+1)) - r(P(k))\right] \tag{41}$$

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

$$= r(e_N), \tag{42}$$

where $\mathbb{1}_{i(P,k)=i} = 1$ if $i(P,k) = i$ and $\mathbb{1}_{i(P,k)=i} = 0$ otherwise. Here, (39) follows from Definition 21, (40) follows by interchanging the summations, (41) follows from the fact that there is precisely one $i \in N$ such that $i(P,k) = i$ for all $k \in \{0,\dots,|N|n-1\}$ and (42) follows from Definition 3.5(i). This concludes the proof. □

B Proof of Proposition 3.8

To prove Proposition 3.8, we first prove the following lemma.

Lemma B.1 *Let $R \in \mathcal{R}$ and $n \in \mathbb{N}$. Then, we have for all $i \in N$ that*

$$K_i^{avg,n}(R) = \sum_{\lambda \in G^n : \lambda_i < 1} t^n(\lambda) p_i^n(\lambda) [r(\lambda + (1/n) \cdot e_i) - r(\lambda)], \tag{43}$$

where

$$t^n(\lambda) = \frac{\prod_{j \in N} \binom{n}{n\lambda_j}}{\binom{|N|n}{|N|n\bar{\lambda}}}, \tag{44}$$

and

$$p_i^n(\lambda) = \frac{1 - \lambda_i}{\sum_{j \in N}(1 - \lambda_j)}, \tag{45}$$

for all $\lambda \in G^n \backslash \{e_N\}$, $\bar{\lambda} = \frac{1}{|N|} \sum_{i \in N} \lambda_i$, for all $\lambda \in \mathbb{R}^N$, and where the risk capital function r is defined in (8).

Proof of Lemma B.1 In this proof, we use the following notation. The set $\widetilde{G}_k^n$ is given by

$$\widetilde{G}_k^n = \left\{ \lambda \in G^n : \sum_{i \in N} \lambda_i = \frac{k}{n} \right\}, \qquad \text{for all } n \in \mathbb{N} \text{ and } k \in \{0, \ldots, |N|n\}. \tag{46}$$

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

$$\widetilde{G}_k^n = \{P(k) : P \in \mathcal{P}^n\}, \qquad \text{for all } n \in \mathbb{N} \text{ and } k \in \{0, \ldots, |N|n\}. \tag{47}$$

Next, we show (43). Then, $K^{avg,n}(R)$ can be rewritten as

$$K^{avg,n}(R) = \frac{1}{|\mathcal{P}^n|} \sum_{P \in \mathcal{P}^n} K^{path,P}(R) \tag{48}$$

$$= \frac{1}{|\mathcal{P}^n|} \sum_{P \in \mathcal{P}^n} \sum_{k=0}^{|N|n-1} [r(P(k+1)) - r(P(k))] \cdot e_{i(P,k)} \tag{49}$$

$$= \sum_{k=0}^{|N|n-1} \sum_{P \in \mathcal{P}^n} \frac{1}{|\mathcal{P}^n|} [r(P(k+1)) - r(P(k))] \cdot e_{i(P,k)}, \tag{50}$$

where (48) follows from Definition 3.7 and (49) follows from Definition 21. Let $i \in N$. Then, we obtain

$$K_i^{avg,n}(R) = \sum_{k=0}^{|N|n-1} \sum_{P \in \mathcal{P}^n : i(P,k)=i} \frac{1}{|\mathcal{P}^n|} [r(P(k+1)) - r(P(k))] \tag{51}$$

$$= \sum_{k=0}^{|N|n-1} \sum_{P \in \mathcal{P}^n : i(P,k)=i} \frac{1}{|\mathcal{P}^n|} [r(P(k) + (1/n) \cdot e_i) - r(P(k))] \tag{52}$$

$$= \sum_{k=0}^{|N|n-1} \sum_{\lambda \in \widetilde{G}_k^n} \sum_{\substack{P \in \mathcal{P}^n: \\ i(P,k)=i, P(k)=\lambda}} \frac{1}{|\mathcal{P}^n|} [r(\lambda + (1/n) \cdot e_i) - r(\lambda)] \tag{53}$$

$$= \sum_{k=0}^{|N|n-1} \sum_{\lambda \in \widetilde{G}_k^n : \lambda_i < 1} [r(\lambda + (1/n) \cdot e_i) - r(\lambda)] \sum_{\substack{P \in \mathcal{P}^n: \\ i(P,k)=i, P(k)=\lambda}} \frac{1}{|\mathcal{P}^n|} \tag{54}$$

$$= \sum_{k=0}^{|N|n-1} \sum_{\lambda \in \widetilde{G}_k^n : \lambda_i < 1} [r(\lambda + (1/n) \cdot e_i) - r(\lambda)]\, t^n(\lambda) p_i^n(\lambda) \tag{55}$$

$$= \sum_{\lambda \in G^n : \lambda_i < 1} [r(\lambda + (1/n) \cdot e_i) - r(\lambda)]\, t^n(\lambda) p_i^n(\lambda), \tag{56}$$

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, (51) follows from (50), (52) follows from (20), (53) follows from (47), (54) follows from the fact that if $k \in \{0, \dots, |N|n-1\}$ 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 $P(k) = \lambda$, (55) follows from the fact that if $k \in \{0, \ldots, |N|n-1\}$ and $\lambda \in \widetilde{G}_k^n$ are such that $P(k) = \lambda$ then $k = |N|n\bar{\lambda}$ and (56) follows from the fact that $\bigcup_{k=1}^{|N|n-1} G_k^n = G^n$ and $G_{k_1}^n \cap G_{k_2}^n = \emptyset$ if $k_1 \neq k_2$.

Next, we show (44). 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,

$$|\mathcal{P}^n| = \frac{(|N|n)!}{(n!)^{|N|}}. \tag{57}$$

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

$$\left|\left\{P \in \mathcal{P}^n : P\left(|N|n\bar{\lambda}\right) = \lambda\right\}\right| = \frac{\left(|N|n\bar{\lambda}\right)!\left(|N|n(1-\bar{\lambda})\right)!}{\prod_{j \in N}(n\lambda_j)!(n(1-\lambda_j))!}. \tag{58}$$

Hence, one can verify that dividing (58) by (57) yields (44). 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 (45). 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 $\lambda + (1/n)e_i$) is given by:

$$\left|\left\{P \in \mathcal{P}^n : P\left(|N|n\bar{\lambda}\right) = \lambda, i\left(P, |N|n\bar{\lambda}\right) = i\right\}\right| = \frac{(|N|n\bar{\lambda})!(|N|n(1-\bar{\lambda})-1)!}{\left[\prod_{j \in N}(n\lambda_j)!\right] \cdot \left[\prod_{j \in N\backslash\{i\}}(n(1-\lambda_j))!\right] \cdot (n(1-\lambda_i)-1)!}. \tag{59}$$

Dividing (59) by (58) yields (45) in a straightforward way. □

Proof of Proposition 3.8 It follows immediately from the proof of Lemma B.1 that 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.

C Proof of Theorem 3.9

We use the following notation.

- 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 $|f(n)| \leq K|g(n)|$ 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 $|g(\varepsilon)| \leq K\varepsilon$ 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 $|f(\varepsilon, n)| \leq K_\varepsilon |g(n)|$ 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 $G_\varepsilon = \left\{\lambda \in [0,1]^N : \varepsilon \leq \bar{\lambda} \leq 1 - \varepsilon\right\}$, and $G_\varepsilon^n = G^n \cap G_\varepsilon$.
- 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 $D(n) = D^{d_n}$, where $d_n = n^{-\frac{1}{2}+\frac{1}{8|N|}}$.[16]

To prove Theorem 3.9, we will prove the following three propositions. The proofs of these propositions are in Subsections C.1, C.2, and C.3, respectively.

Proposition C.1 *Let $i \in N$ and define $Dom = \{(\varepsilon, n, \lambda) : \varepsilon > 0, n \in \mathbb{N}, \lambda \in G_\varepsilon^n\}$. Then, we have*

$$
\begin{aligned}
t^n(\lambda) &= \left(e^{-c(\bar{\lambda})n\|\lambda - \bar{\lambda}\cdot e_N\|^2}\right) b(n, \bar{\lambda})[1 + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}})], && \text{if } \lambda \in D(n), && (60)\\
&= \mathcal{O}^{\varepsilon}(n^{-\infty}), && \text{if } \lambda \notin D(n), && (61)
\end{aligned}
$$

and

$$
\begin{aligned}
p_i^n(\lambda) &= \frac{1}{|N|}[1 + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}})], && \text{if } \lambda \in D(n), && (62)\\
&= \mathcal{O}(1), && \text{if } \lambda \notin D(n), && (63)
\end{aligned}
$$

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

$$
c(\bar{\lambda}) = \frac{1}{2\bar{\lambda}(1 - \bar{\lambda})} > 0, \qquad (64)
$$

[16] Theorem 3.9 can be proven by using a diagonal width $d_n = n^{-\frac{1}{2}+\delta}$ for some $\delta \in \left(0, \frac{1}{2(|N|+2)}\right)$. The proofs are based on $\delta = \frac{1}{8|N|}$.

and

$$b(n,\bar{\lambda}) = (2\pi n)^{\frac{1}{2}(1-|N|)}\sqrt{|N|}(\bar{\lambda}(1-\bar{\lambda}))^{\frac{1}{2}(1-|N|)}. \tag{65}$$

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 $K^{avg,n}$ if n converges to infinity.

To proceed with the proof, we define the function $h^n : [0,1]^N\backslash\{e_\emptyset, e_N\} \to \mathbb{R}_{++}$ as follows:

$$h^n(\lambda) = \left(e^{-c(\bar{\lambda})n\|\lambda-\bar{\lambda}\cdot e_N\|^2}\right) b(n,\bar{\lambda})\frac{1}{|N|}, \tag{66}$$

for all $\lambda \in [0,1]^N\backslash\{e_\emptyset, e_N\}$ and $n \in \mathbb{N}$, where $c(\bar{\lambda})$ is defined in (64) and $b(n,\bar{\lambda})$ in (65).

It follows from Proposition C.1 that

$$t^n(\lambda)p_i^n(\lambda) = h^n(\lambda)[1 + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}})], \tag{67}$$

for all $(\varepsilon, n, \lambda) \in Dom$ such that $\lambda \in D(n)$. This leads to the following approximation.

Proposition C.2 *Let $R \in \mathcal{R}$. Then, for all $i \in N$ we have*

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

where

$$\phi_m^{n,\varepsilon} = \frac{1}{n}\sum_{\lambda\in G_\varepsilon^n\cap D(n)\cap A_{\mathbb{Q}_m}} h^n(\lambda), \tag{68}$$

with p^* and $\mathbb{Q}_m$ as defined in Proposition 3.2.

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 (68) by an integral (see Lemma C.20 and Lemma C.21) and, thereafter, solving this integral.

Proposition C.3 *For all $R \in \mathcal{R}$, it holds that*

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

where ϕ_m for $m \in \{1, \dots, p^*\}$ is as defined in (27).

Proof of Theorem 3.9 Let $R \in \mathcal{R}$. From Proposition C.3, we get for all $n \in \mathbb{N}$ and $\varepsilon > 0$ that

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

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

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

This concludes the proof. □

In the remaining three subsections of this Online Appendix, we present the proofs of Propositions C.1, C.2, and C.3, respectively.

C.1 Proof of Proposition C.1

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

$$G(\lambda) = |N| g\left(\bar{\lambda}\right) - \sum_{i \in N} g(\lambda_i) + |N| g\left(1 - \bar{\lambda}\right) - \sum_{i \in N} g(1 - \lambda_i), \qquad \text{for all } \lambda \in [0,1]^N. \quad (69)$$

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

$$N_1^\lambda = \{i \in N : \lambda_i > 0\} \text{ and } N_2^\lambda = \{i \in N : \lambda_i < 1\}. \tag{70}$$

- 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 $f(n) = \mathcal{O}(n^{-\infty})$.

C.1.1 Some preliminary lemmas

Lemma C.4 *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 C.5 *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 $|R| \leq \frac{1}{3}|N| \min\{\lambda_1, \ldots, \lambda_{|N|}, 1-\lambda_1, \ldots, 1-\lambda_{|N|}\}^{-2} \|\lambda - \bar{\lambda} \cdot e_N\|^3$.

Proof 1. This follows from continuity of g (Lemma C.4).

2. This follows from strict convexity of g (Lemma C.4).

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

$$g(\lambda_i) = g(\bar{\lambda}) + g'(\bar{\lambda})(\lambda_i - \bar{\lambda}) + \frac{g''(\bar{\lambda})}{2}(\lambda_i - \bar{\lambda})^2 + \frac{g'''(\xi_{i,1})}{6}(\lambda_i - \bar{\lambda})^3 \tag{71}$$

$$= g(\bar{\lambda}) + (\ln(\bar{\lambda}) + 1)(\lambda_i - \bar{\lambda}) + \frac{1}{2\bar{\lambda}}(\lambda_i - \bar{\lambda})^2 - \frac{1}{6\xi_{i,1}^2}(\lambda_i - \bar{\lambda})^3, \tag{72}$$

where (71) follows from Taylor's theorem. Note that

$$\sum_{i \in N}(\lambda_i - \bar{\lambda}) = 0. \tag{73}$$

Then, summing the expression (72) of $g(\lambda_i)$ for 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

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

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

Lemma C.6 *Let* $d, \varepsilon > 0$. *Then, for all* $\lambda \in G_\varepsilon \cap D^d$, *we have*

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

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 C.7 *For all* (n, λ) *such that* $n \in \mathbb{N}$ *and* $\lambda \in G^n \backslash \{e_\emptyset, e_N\}$*, we have*

$$t^n(\lambda) = \left(e^{G(\lambda)}\right)^n (2\pi n)^{\frac{1}{2}(1+|N|-|N_1^\lambda|-|N_2^\lambda|)} \sqrt{|N|} \frac{(\bar{\lambda}(1-\bar{\lambda}))^{\frac{1}{2}}}{\prod_{i \in N_1^\lambda} \sqrt{\lambda_i} \prod_{i \in N_2^\lambda} \sqrt{1-\lambda_i}} \cdot \left[1 + \mathcal{O}\left(\frac{1}{n \min(\{\lambda_j : j \in N_1^\lambda\} \cup \{1-\lambda_j : j \in N_2^\lambda\})}\right)\right], \tag{74}$$

where N_1^λ *and* N_2^λ *are defined in (70).*

Proof Using (44), 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

$$\begin{aligned} \ln(t^n(\lambda)) = {} & |N| \ln(n!) + \ln((|N|n\bar{\lambda})!) + \ln((|N|n(1-\bar{\lambda}))!) - \ln((|N|n)!) \\ & - \sum_{i \in N_1^\lambda} \ln((n\lambda_i)!) - \sum_{i \in N_2^\lambda} \ln((n(1-\lambda_i))!). \end{aligned} \tag{75}$$

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 (75) can be written as

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

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

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

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

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 (69). We get

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

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

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

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 C.8 *We have for all $(\varepsilon, n, \lambda) \in Dom$ with $d_n < \frac{1}{2}\varepsilon$ and $\lambda \in D(n)$ that*

$$\frac{(\bar{\lambda}(1-\bar{\lambda}))^{\frac{1}{2}|N|}}{\prod_{i\in N}\sqrt{\lambda_i}\prod_{i\in N}\sqrt{1-\lambda_i}} = 1 + \mathcal{O}^\varepsilon(n^{-1+\frac{1}{4|N|}}). \tag{76}$$

Proof According to Lemma C.6 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

$$\ln(\lambda_i) = \ln(\bar{\lambda}) + \frac{1}{\bar{\lambda}}(\lambda_i - \bar{\lambda}) - \frac{1}{2\xi_{i,1}^2}(\lambda_i - \bar{\lambda})^2, \tag{77}$$

for some $\xi_{i,1} \in [\lambda_i; \bar{\lambda}]$ and for all $i \in N$. From (73) and (77) it follows that

$$\frac{1}{2}\sum_{i\in N}\ln(\lambda_i) = \frac{1}{2}|N|\ln(\bar{\lambda}) - \sum_{i\in N}\frac{1}{4\xi_{i,1}^2}(\lambda_i - \bar{\lambda})^2. \tag{78}$$

Similarly, we obtain

$$\frac{1}{2}\sum_{i\in N}\ln(1-\lambda_i) = \frac{1}{2}|N|\ln(1-\bar{\lambda}) - \sum_{i\in N}\frac{1}{4\xi_{i,2}^2}(\bar{\lambda}-\lambda_i)^2, \tag{79}$$

where $\xi_{i,2} \in [1-\lambda_i; 1-\bar{\lambda}]$ 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 (78) and (79) yields the desired result. □

C.1.2 Proof of Proposition C.1

We now use Lemmas C.4 to C.8 to prove Proposition C.1. We do this in several steps: (60) is shown in Lemma C.9, (61) in Lemma C.10, (62) in Lemma C.11 and (63) in Lemma C.12. 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 C.9 *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 C.6, we then get

$$\min\{\lambda_1,\ldots,\lambda_{|N|}, 1-\lambda_1,\ldots,1-\lambda_{|N|}\} > \frac{1}{2}\varepsilon, \tag{80}$$

and, so,

$$N_1^\lambda = N_2^\lambda = N. \tag{81}$$

Using Lemma C.5.3 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,

$$\begin{aligned} e^{nG(\lambda)} &= e^{-c(\bar{\lambda})n\|\lambda-\bar{\lambda}\cdot e_N\|^2} \cdot e^{\mathcal{O}^\varepsilon(n^{-\frac{3}{2}+\frac{3}{8|N|}})} \\ &= e^{-c(\bar{\lambda})n\|\lambda-\bar{\lambda}\cdot e_N\|^2} \left[1 + \mathcal{O}^\varepsilon\left(n^{-\frac{1}{2}+\frac{3}{8|N|}}\right)\right]. \end{aligned} \tag{82}$$

where (82) follows from the fact that $e^x = 1 + \mathcal{O}(x)$ if $x \in [0, K]$ for some constant $K > 0$. Substituting (76), (80), (81) and (82) in (74) yields the desired result. □

Lemma C.10 *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 (69). The set $G_\varepsilon \backslash D^d$ is compact. Moreover, the function G is continuous (Lemma C.5.1). 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 C.5.2 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 C.7 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 C.5.3 that

$$\begin{aligned}G(\lambda) &= -c(\bar{\lambda})\|\lambda - \bar{\lambda}\cdot e_N\|^2 + R \\ &= -c(\bar{\lambda})\|\lambda - \bar{\lambda}\cdot e_N\|^2\left[1 - R(c(\bar{\lambda}))^{-1}\|\lambda - \bar{\lambda}\cdot e_N\|^{-2}\right],\end{aligned}$$

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 C.6, 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

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

So, then, we obtain that

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

which follows from $c(\bar{\lambda}) \geq 2$, and, hence,

$$e^{nG(\lambda)} < e^{-n^{\frac{1}{4|N|}}}. \tag{83}$$

We get

$$\begin{aligned}t^n(\lambda) &= \mathcal{O}^{\varepsilon}(e^{nG(\lambda)}n^{\frac{1}{2}(1-|N|)}) &(84)\\ &= \mathcal{O}^{\varepsilon}((e^{-1})^{n^{\frac{1}{4|N|}}}n^{\frac{1}{2}(1-|N|)}) &(85)\\ &= \mathcal{O}^{\varepsilon}(n^{-\infty}), &(86)\end{aligned}$$

where (84) follows from Lemma C.7, (85) follows from (83) and (86) 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 C.11 *We have for all $i \in N$ and $(\varepsilon, n, \lambda) \in Dom$ that*

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

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

$$\left|\frac{1-\lambda_i}{\sum_{j\in N}(1-\lambda_j)} - \frac{1}{|N|}\right| = \left|\frac{1-\lambda_i}{(1-\bar{\lambda})|N|} - \frac{1-\bar{\lambda}}{(1-\bar{\lambda})|N|}\right|$$
$$= \frac{|\bar{\lambda}-\lambda_i|}{(1-\bar{\lambda})|N|}$$
$$< \frac{n^{-\frac{1}{2}+\frac{1}{8|N|}}}{(1-\bar{\lambda})|N|} \tag{87}$$
$$\leq \frac{n^{-\frac{1}{2}+\frac{1}{8|N|}}}{\varepsilon|N|}, \tag{88}$$

for all $(\varepsilon, n, \lambda) \in Dom$ such that $\lambda \in D(n)$. Here, (87) follows from $|\bar{\lambda} - \lambda_i| \leq \|\lambda - \bar{\lambda}\cdot e_N\| < d_n = n^{-\frac{1}{2}+\frac{1}{8|N|}}$ and (88) follows from $1-\bar{\lambda} \geq \varepsilon$. This concludes the proof. □

Lemma C.12 *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 $0 \leq p_i^n(\lambda) \leq 1$. □

C.2 Proof of Proposition C.2

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

$$C^n(\lambda) = \left\{\lambda + \frac{1}{n}x : x \in [0,1]^N\right\}. \tag{89}$$

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

$$D'(n) = D^{d'_n}, \qquad \text{where } d'_n = d_n + (\sqrt{|N|}/n) = n^{-\frac{1}{2}+\frac{1}{8|N|}} + (\sqrt{|N|}/n). \tag{90}$$

- 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

$$\nu(C^n(\lambda)) = n^{-|N|}, \qquad \text{for all } \lambda \in G^n, \tag{91}$$

and

$$\nu(D'(n)) = \mathcal{O}(d_n'^{|N|-1}) = \mathcal{O}(n^{-\frac{1}{2}+\frac{1}{8|N|}}), \qquad \text{for all } n \in \mathbb{N}. \tag{92}$$

- Let $R \in \mathcal{R}$ and $\varepsilon > 0$. We define the set $B(R,n)$ by

$$B(R,n) = \left\{ \lambda \in [0,1]^N : \exists \hat{\lambda} \in [0,1]^N \backslash L(R) : \|\lambda - \hat{\lambda}\| < \frac{1}{n} \right\}, \tag{93}$$

  for all $R \in \mathcal{R}$ and $n \in \mathbb{N}$, where $L(R)$ is defined in Definition 3.2. This is the set of all participation profiles close to a participation profile that is an element of multiple sets $A_{\mathbb{Q}_m}$. 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^n_ε have a non-negligible aggregate contribution.

Lemma C.13 *For all $i \in N$, we have*

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

Proof Recall (46) for the definition of $\widetilde{G}^n_k$. We obtain

$$\sum_{\lambda \in G^n \backslash G^n_\varepsilon : \lambda_i < 1} t^n(\lambda) p^n_i(\lambda) = \sum_{k=0}^{\lceil \varepsilon |N| n \rceil - 1} \sum_{\lambda \in \widetilde{G}^n_k : \lambda_i < 1} t^n(\lambda) p^n_i(\lambda) + \sum_{k=\lfloor (1-\varepsilon)|N|n \rfloor + 1}^{|N|n-1} \sum_{\lambda \in \widetilde{G}^n_k : \lambda_i < 1} t^n(\lambda) p^n_i(\lambda) \tag{94}$$

$$\leq \sum_{k=0}^{\lceil \varepsilon |N| n \rceil - 1} \sum_{\lambda \in \widetilde{G}^n_k : \lambda_i < 1} t^n(\lambda) + \sum_{k=\lfloor (1-\varepsilon)|N|n \rfloor + 1}^{|N|n-1} \sum_{\lambda \in \widetilde{G}^n_k : \lambda_i < 1} t^n(\lambda) \tag{95}$$

$$\leq \sum_{k=0}^{\lceil \varepsilon |N| n \rceil - 1} 1 + \sum_{k=\lfloor (1-\varepsilon)|N|n \rfloor + 1}^{|N|n-1} 1 \tag{96}$$

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

$$< 2\varepsilon |N| n + 1 \tag{97}$$

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

Here, (94) follows from (46) and (55), (95) follows from $0 \leq p^n_i(\lambda) \leq 1$ for all $\lambda \in G^n \backslash \{e_N\}$, (96) follows from $\sum_{\lambda \in \widetilde{G}^n_k} t^n(\lambda) = 1$ for all $k \in \{0, \ldots, |N|n - 1\}$ and (97) follows from the fact that $\lceil x \rceil < x + 1$ for all $x \in \mathbb{R}$. □

The following result follows almost directly from Proposition C.1.

Lemma C.14 *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

$$\begin{aligned} \sum_{\lambda \in [G_\varepsilon^n \backslash D(n)]:\lambda_i<1} t^n(\lambda) p_i^n(\lambda) &= \sum_{\lambda \in [G_\varepsilon^n \backslash D(n)]:\lambda_i<1} \mathcal{O}^\varepsilon(n^{-\infty}) && (98) \\ &< (n+1)^{|N|}\mathcal{O}^\varepsilon(n^{-\infty}) && (99) \\ &= \mathcal{O}^\varepsilon(n^{-\infty}), \end{aligned}$$

where (98) follows from Proposition C.1 and (99) follows from $\sharp(\{\lambda \in [G_\varepsilon^n \backslash D(n)] : \lambda_i < 1\}) < \sharp(G^n) = (n+1)^{|N|}$. □

Lemma C.15 *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

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

and

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

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

This concludes the proof. □

Lemma C.16 *For all $i \in N$, we have*

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

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 $t^n(\lambda)p^n_i(\lambda) = h^n(\lambda) = 1$ for all $\lambda \in G^n_\varepsilon$. Next, we let $|N| \geq 2$. For all $\lambda \in G^n_\varepsilon \cap D(n)$, we have

$$\begin{aligned} |t^n(\lambda)p^n_i(\lambda) - h^n(\lambda)| &= \left| h^n(\lambda)[1 + \mathcal{O}^\varepsilon(n^{-\frac{1}{2}+\frac{3}{8|N|}})] \cdot [1 + \mathcal{O}^\varepsilon(n^{-\frac{1}{2}+\frac{1}{8|N|}})] - h^n(\lambda) \right| && (100) \\ &= \left| h^n(\lambda)\mathcal{O}^\varepsilon(n^{-\frac{1}{2}+\frac{3}{8|N|}}) \right| \\ &= \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+\frac{3}{8|N|}}), && (101) \end{aligned}$$

where (100) follows from Lemma C.9 and Lemma C.11 and (101) follows from $h^n(\lambda) = \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)})$. If $y \in C^n(\lambda)$ for a $\lambda \in G^n_\varepsilon \cap D(n)$, we have

$$\begin{aligned} \|y - \bar{y} \cdot e_N\| &\leq \|y - \bar{\lambda} \cdot e_N\| && (102) \\ &\leq \|y - \lambda\| + \|\lambda - \bar{\lambda} \cdot e_N\| && (103) \\ &< (\sqrt{|N|}/n) + n^{-\frac{1}{2}+\frac{1}{8|N|}}, && (104) \end{aligned}$$

where (102) and (103) follow from the triangular inequality and (104) follows from the fact that $\|y - \lambda\| \leq (\sqrt{|N|}/n)$ for all $y \in C^n(\lambda)$. So, we get

$$\bigcup_{\lambda \in G^n_\varepsilon \cap D(n)} C^n(\lambda) \subset D'(n). \quad (105)$$

and, so,

$$\begin{aligned} n^{-|N|}\sharp(G^n_\varepsilon \cap D(n)) &\leq \nu\left(D'(n)\right) && (106) \\ &= \mathcal{O}(d'^{|N|-1}_n) && (107) \end{aligned}$$

$$= \mathcal{O}\left(\left(n^{-\frac{1}{2}+\frac{1}{8|N|}} + (\sqrt{|N|}/n)\right)^{(|N|-1)}\right)$$

$$= \mathcal{O}(n^{-\frac{1}{2}|N|+\frac{5}{8}-\frac{1}{8|N|}}), \tag{108}$$

where (106) follows from (91) and (105), and (107) follows from (92). 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 C.17 *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*

$$G_\varepsilon \cap D(n) \cap A_{\mathbb{Q}_m} = \emptyset.$$

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

$$f_{\mathbb{Q}_\ell}(t \cdot e_N) - f_{\mathbb{Q}_m}(t \cdot e_N) = t(f_{\mathbb{Q}_\ell}(e_N) - f_{\mathbb{Q}_m}(e_N)) \geq t\alpha, \qquad \text{for all } t \in [0, 1]. \tag{109}$$

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

$$f_{\mathbb{Q}_\ell}(\lambda) - f_{\mathbb{Q}_m}(\lambda) = f_{\mathbb{Q}_\ell}(\bar{\lambda} \cdot e_N) - f_{\mathbb{Q}_m}(\bar{\lambda} \cdot e_N) + f_{\mathbb{Q}_\ell}(\lambda - \bar{\lambda} \cdot e_N) - f_{\mathbb{Q}_m}(\lambda - \bar{\lambda} \cdot e_N) \tag{110}$$

$$\geq \bar{\lambda}\alpha + f_{\mathbb{Q}_\ell}(\lambda - \bar{\lambda} \cdot e_N) - f_{\mathbb{Q}_m}(\lambda - \bar{\lambda} \cdot e_N), \tag{111}$$

where (110) follows from linearity of $f_{\mathbb{Q}_\ell}$ and $f_{\mathbb{Q}_m}$ and (111) follows from (109). Moreover, we obtain that

$$\begin{aligned}
|f_{\mathbb{Q}_{m'}}(\lambda - \bar{\lambda} \cdot e_N)| &\leq \|(f_{\mathbb{Q}_{m'}}(e_i))_{i\in N}\| \cdot \|\lambda - \bar{\lambda} \cdot e_N\| && (112)\\
&\leq Mn^{-\frac{1}{4}} && (113)\\
&< MN_\varepsilon^{-\frac{1}{4}} && (114)\\
&= \frac{1}{2}\varepsilon\alpha && (115)\\
&\leq \frac{1}{2}\bar{\lambda}\alpha, && (116)
\end{aligned}$$

for all $m' \in \{1, \ldots, p\}$, where (112) follows from the Cauchy-Schwartz inequality applied to $\sum_{i\in N} f_{\mathbb{Q}_m}(e_i)(\lambda_i - \bar{\lambda})$, (113) follows from $m' \in \{1, \ldots, p\}$ and $\lambda \in D(n)$, (165) follows from $n > N_\varepsilon$, (115) follows from substituting the definition of N_ε, follows from and (116) follows from $\lambda \in G_\varepsilon$. Hence, substituting (116) in (111) yields that $f_{\mathbb{Q}_\ell}(\lambda) - f_{\mathbb{Q}_m}(\lambda) > 0$. Therefore, we have $\lambda \notin A_{\mathbb{Q}_m}$ for every $\lambda \in G_\varepsilon \cap D(n)$ and, hence,

$$G_\varepsilon \cap D(n) \cap A_{\mathbb{Q}_m} = \emptyset. \qquad (117)$$

□

Note that from (16) and Lemma C.17 it follows for all $\varepsilon > 0$ that

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

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 (93), 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 $\lambda \in A_{\mathbb{Q}_m} \backslash B(n)$ for some $m \in \{1, \ldots, p\}$, then $\lambda + (1/n) \cdot e_i \in A_{\mathbb{Q}_m}$ 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 C.18 *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 (38), i.e.,

$$[0,1]^N \backslash L(R) = \bigcup_{\ell,m \in \{1,\ldots,p\}: \ell \neq m} A_{\mathbb{Q}_\ell} \cap A_{\mathbb{Q}_m}.$$

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

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

and $D_\varepsilon = \{\lambda \in G_\varepsilon : \lambda = \bar{\lambda} \cdot e_N\}$. According to Lemma C.17 we have for all $m \in \{p^* + 1, \ldots, p\}$ that $D_\varepsilon \cap A_{\mathbb{Q}_m} = \emptyset$. Since D_ε and $A_{\mathbb{Q}_m}$ are both compact we can define $\alpha_{\varepsilon,m} = \text{dist}(D_\varepsilon, A_{\mathbb{Q}_m}) = \min\{\|x - y\| : x \in D_\varepsilon, y \in A_{\mathbb{Q}_m}\}$. 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 (37) from the proof of Proposition 3.4, i.e.,

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

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 upper bound of the cardinality of $H^n(\ell, m)$, we first derive the Lebesgue measure of the following Euclidean set

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

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 $\|\lambda - \lambda_1\| \leq \frac{\sqrt{|N|}}{n} + \frac{1}{n}$. 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 5. Obviously $\|\lambda - \lambda_2\|^2 = \|\lambda - \lambda_1\|^2 + \|\lambda_1 - \lambda_2\|^2$ and hence $\|\lambda_1 - \lambda_2\| \leq \|\lambda - \lambda_2\| = \|\lambda - \bar{\lambda} \cdot e_N\| < d_n'$. 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 $|\alpha_1| = \|\lambda_2\| = \bar{\lambda}\sqrt{|N|} < \sqrt{|N|}$, $|\alpha_k| \leq \sqrt{\alpha_2^2 + \cdots + \alpha_{|N|-1}^2} = \|\lambda_1 - \lambda_2\| < d_n' = \mathcal{O}(n^{-\frac{1}{2} + \frac{1}{8|N|}})$

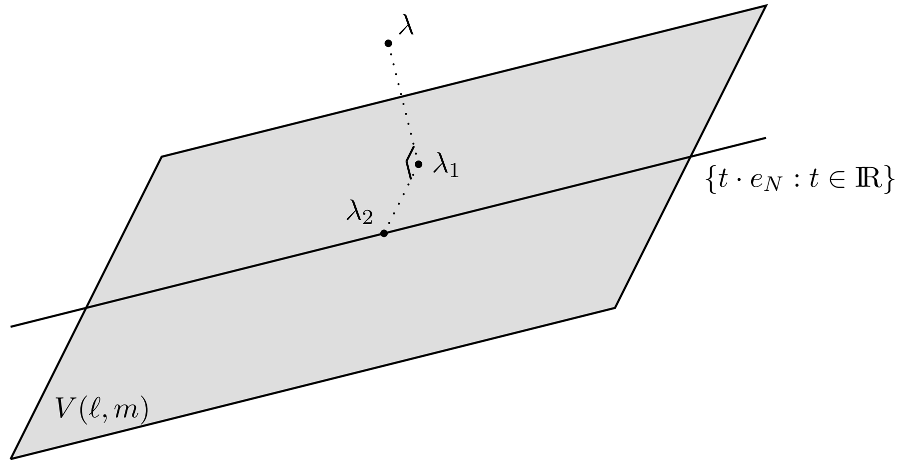

Figure 5: Illustration of λ_1 and λ_2 corresponding to the proof of Lemma C.18.

for all $k \in \{2, \dots, |N| - 1\}$ and $|\alpha_{|N|}| = \|\lambda - \lambda_1\| \le \frac{1}{n} + \frac{\sqrt{|N|}}{n} = \mathcal{O}(n^{-1})$. Hence,

$$\begin{aligned}\nu(\widetilde{H}^n(\ell, m)) &= \mathcal{O}(1)\mathcal{O}(n^{(-\frac{1}{2}+\frac{1}{8|N|})(|N|-2)})\mathcal{O}(n^{-1}) \\ &= \mathcal{O}(n^{-\frac{1}{2}|N|+\frac{1}{8}}).\end{aligned} \tag{118}$$

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

$$\bar{y} = \bar{\lambda} + (\bar{y} - \bar{\lambda}) \begin{cases} \ge \varepsilon - (1/n) > \frac{1}{2}\varepsilon, \\ \le 1 - \varepsilon + (1/n) < 1 - \frac{1}{2}\varepsilon, \end{cases} \tag{119}$$

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

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

From this, (105) and (119), we get

$$\bigcup_{\lambda \in H^n(\ell,m)} C^n(\lambda) \subset \widetilde{H}^n(\ell, m), \qquad \text{for all } n \in \mathbb{N} \text{ such that } n > \frac{2}{\varepsilon}. \tag{120}$$

From (91) and (120) we get

$$n^{-|N|}\sharp(H^n(\ell, m)) \le \nu(\widetilde{H}^n(\ell, m)). \tag{121}$$

Substituting (118) in (121) yields

$$\sharp(H^n(\ell,m)) = \mathcal{O}(n^{\frac{1}{2}|N|+\frac{1}{8}}). \tag{122}$$

Then, we obtain

$$\sum_{\lambda\in G_\varepsilon^n\cap D(n)\cap B(n)} h^n(\lambda) \leq \sum_{\ell,m\in\{1,\dots,p\}:\ell\neq m}\ \sum_{\lambda\in H^n(\ell,m)} h^n(\lambda) \tag{123}$$

$$\leq \begin{pmatrix} p \\ 2 \end{pmatrix} \max_{\ell,m\in\{1,\dots,p\}:\ell\neq m} \sharp(H^n(\ell,m)) \max_{\lambda\in G_\varepsilon} h^n(\lambda) \tag{124}$$

$$= \begin{pmatrix} p \\ 2 \end{pmatrix} \mathcal{O}(n^{\frac{1}{2}|N|+\frac{1}{8}})\mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)}) \tag{125}$$

$$= \mathcal{O}^\varepsilon(n^{\frac{5}{8}}),$$

where (123) follows from (38), (124) follows from $\sharp(\{\ell,m\in\{1,\dots,p\}:\ell\neq m\}) = \begin{pmatrix} p \\ 2 \end{pmatrix}$ and (125) follows from (122) 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 C.2 It is sufficient to show this result for sufficiently large n. We get

$$K_i^{avg,n}(R) = \sum_{\lambda\in G^n:\lambda_i<1} t^n(\lambda)p_i^n(\lambda)\left[r\left(\lambda+(1/n)\cdot e_i\right)-r(\lambda)\right] \tag{126}$$

$$= \sum_{\lambda\in G_\varepsilon^n:\lambda_i<1} t^n(\lambda)p_i^n(\lambda)\left[r\left(\lambda+(1/n)\cdot e_i\right)-r(\lambda)\right] + \mathcal{O}(\varepsilon)+\mathcal{O}(n^{-1}) \tag{127}$$

$$= \sum_{\lambda\in G_\varepsilon^n\cap D(n)} t^n(\lambda)p_i^n(\lambda)\left[r\left(\lambda+(1/n)\cdot e_i\right)-r(\lambda)\right] + \mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon\left(n^{-1}\right) \tag{128}$$

$$= \sum_{\lambda\in G_\varepsilon^n\cap D(n)} h^n(\lambda)\left[r\left(\lambda+(1/n)\cdot e_i\right)-r(\lambda)\right] + \mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{129}$$

$$= \sum_{\lambda\in[G_\varepsilon^n\cap D(n)]\setminus B(n)} h^n(\lambda)\left[r\left(\lambda+(1/n)\cdot e_i\right)-r(\lambda)\right] + \mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{130}$$

$$= \sum_{m=1}^{p}\ \sum_{\lambda\in[G_\varepsilon^n\cap D(n)\cap A_{\mathbb{Q}_m}]\setminus B(n)} h^n(\lambda)\frac{1}{n}E_{\mathbb{Q}_m}[X_i] + \mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{131}$$

$$= \sum_{m=1}^{p^*}\ \sum_{\lambda\in[G_\varepsilon^n\cap D(n)\cap A_{\mathbb{Q}_m}]\setminus B(n)} h^n(\lambda)\frac{1}{n}E_{\mathbb{Q}_m}[X_i] + \mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{132}$$

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

$$= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \frac{1}{n} \sum_{\lambda \in G_\varepsilon^n \cap D(n) \cap A_{\mathbb{Q}_m}} h^n(\lambda) + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}), \tag{133}$$

where (126) follows from Proposition B.1, (127) follows from Lemma C.13 and Lemma C.15, (128) follows from Lemma C.14 and Lemma C.15, (129) follows from Lemma C.15 and Lemma C.16, (130) follows from Lemma C.15 and Lemma C.18, (131) follows from $[0,1]^N \setminus L(R) \subset B(n)$, (132) follows from Lemma C.17 and (133) follows from Lemma C.18. This concludes the proof. □

C.3 Proof of Proposition C.3

Lemma C.19 *The function h^n is differentiable for a 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 (90).

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

$$\frac{\partial h^n}{\partial \lambda_i}(\lambda) = \frac{\partial f^n}{\partial \lambda_i}(\lambda) \cdot h^n(\lambda) + \frac{\partial g}{\partial \lambda_i}(\lambda) \cdot \frac{h^n(\lambda)}{g(\lambda)}, \qquad \text{for all } \lambda \in [0,1]^N \setminus \{e_\emptyset, e_N\}. \tag{134}$$

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

$$\begin{aligned}
\frac{\partial f^n}{\partial \lambda_i}(\lambda) &= -c(\bar{\lambda})n \left[\sum_{k \neq i} 2\left(\lambda_k - \bar{\lambda}\right) \cdot -\frac{1}{|N|} + 2(\lambda_i - \bar{\lambda}) \left(1 - \frac{1}{|N|}\right) \right] \\
&\quad + \frac{1 - 2\bar{\lambda}}{2|N|(\bar{\lambda}(1-\bar{\lambda}))^2} n \|\lambda - \bar{\lambda} \cdot e_N\|^2 \\
&= -c(\bar{\lambda})n2\left(\lambda_i - \bar{\lambda}\right) + \frac{1 - 2\bar{\lambda}}{2|N|(\bar{\lambda}(1-\bar{\lambda}))^2} n \|\lambda - \bar{\lambda} \cdot e_N\|^2 \\
&= \mathcal{O}^\varepsilon(n^{\frac{1}{2}+\frac{1}{8|N|}}) + \mathcal{O}^\varepsilon(n^{\frac{1}{4|N|}}) \qquad (135) \\
&= \mathcal{O}^\varepsilon(n^{\frac{1}{2}+\frac{1}{8|N|}}), \\
\frac{\partial g}{\partial \lambda_i}(\lambda) &= \mathcal{O}^\varepsilon(1),
\end{aligned}$$

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

where (135) follows from $|\lambda_i - \bar{\lambda}| \leq \|\lambda - \bar{\lambda}\cdot e_N\| \leq d'_n = \mathcal{O}(n^{-\frac{1}{2}+\frac{1}{8|N|}})$. Then, the result follows from substituting these equations in (134). □

Lemma C.20 *Let $R \in \mathcal{R}$. Then, we have for all $m \in \{1,\dots,p\}$ that*

$$\sum_{\lambda\in G_\varepsilon^n\cap D(n)\cap A_{\mathbb{Q}_m}} h^n(\lambda) = n^{|N|} \sum_{\lambda\in G_\varepsilon^n\cap D(n)\cap A_{\mathbb{Q}_m}} \int_{C^n(\lambda)} h^n(\lambda^*)d\lambda^* + \mathcal{O}^\varepsilon(n^{\frac{5}{8}}),$$

where $C^n(\lambda)$ is defined in (89)

Proof Let $\varepsilon > 0$. It is sufficient to show this result for all $n \in \mathbb{N}$ such that $n > \frac{2}{\varepsilon}$. Let $\lambda \in G_\varepsilon^n \cap D(n)$. From (105) and (119) it follows that

$$C^n(\lambda) \subset G_{\frac{1}{2}\varepsilon} \cap D'(n). \tag{136}$$

We get from (136) and Lemma C.19 that h^n is differentiable in λ^* for all $\lambda^* \in C^n(\lambda)$. Applying Taylor's theorem yields that

$$h^n(\lambda) - h^n(\lambda^*) = \sum_{i\in N} \frac{\partial h}{\partial \lambda_i}(\chi)(\lambda_i - \lambda_i^*),\ \text{for all } \lambda^* \in C^n(\lambda),\ \text{where } \chi \in \text{conv}\{\lambda, \lambda^*\}. \tag{137}$$

Here, as $\chi \in C^n(\lambda)$, we get from Lemma C.19 that

$$\frac{\partial h}{\partial \lambda_i}(\chi) = \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+1+\frac{1}{8|N|}}), \qquad \text{for all } i \in N. \tag{138}$$

So, as $|\lambda_i - \lambda_i^*| \leq n^{-1}$ for all $\lambda^* \in C^n(\lambda)$ and $i \in N$, we get from (137) and (138) 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

$$h^n(\lambda) - n^{|N|}\int_{C^n(\lambda)} h^n(\lambda^*)d\lambda^* = \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+\frac{1}{8|N|}}), \qquad \text{for all } \lambda \in G_\varepsilon^n \cap D(n). \tag{139}$$

Moreover, from (108) we get

$$\begin{aligned}\sharp(G_\varepsilon^n \cap D(n) \cap A_{\mathbb{Q}_m}) &\leq \sharp(G_\varepsilon^n \cap D(n)) \\ &= \mathcal{O}(n^{\frac{1}{2}|N|+\frac{5}{8}-\frac{1}{8|N|}}).\end{aligned} \tag{140}$$

Hence, from (139) and (140) it follows that

$$\begin{aligned}\left|\sum_{\lambda \in G_\varepsilon^n \cap D(n) \cap A_{\mathbb{Q}_m}} \left(h^n(\lambda) - n^{|N|}\int_{C^n(\lambda)} h^n(\lambda^*)d\lambda^*\right)\right| &\leq \sharp(G_\varepsilon^n \cap D(n) \cap A_{\mathbb{Q}_m})\mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+\frac{1}{8|N|}}) \\ &= \mathcal{O}^\varepsilon(n^{\frac{5}{8}}).\end{aligned}$$

This concludes the result. □

Lemma C.21 *Let $R \in \mathcal{R}$. Then, we have for all $m \in \{1, \ldots, p\}$ that*

$$\sum_{\lambda^* \in G_\varepsilon^n \cap D(n) \cap A_{\mathbb{Q}_m}} \int_{C^n(\lambda^*)} h^n(\lambda)d\lambda = \int_{G_\varepsilon \cap D(n) \cap A_{\mathbb{Q}_m}} h^n(\lambda)d\lambda + \mathcal{O}^\varepsilon(n^{-|N|+\frac{5}{8}}).$$

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 $A = \bigcup_{\lambda^* \in G_\varepsilon^n \cap D(n) \cap A_{\mathbb{Q}_m}} C^n(\lambda^*)$ and $B = G_\varepsilon \cap D(n) \cap A_{\mathbb{Q}_m}$. Moreover, define

$$\begin{aligned}E_1^n &= B(n/(\sqrt{|N|}+1)) \cap G_{\frac{1}{2}\varepsilon} \\ E_2^n &= [G_{\varepsilon-(1/n)} \cap D'(n)] \backslash D''(n) \\ E_3^n &= [D'(n) \cap G_{\varepsilon-(1/n)}] \backslash G_{\varepsilon+(1/n)},\end{aligned}$$

where the set $B(n)$ is defined in (93). We first show

$$(A \backslash B) \cup (B \backslash A) \subset E_1^n \cup E_2^n \cup E_3^n. \tag{141}$$

Let $y_1 \in A \backslash B$, so we have $y_1 \in C^n(\lambda)$ for some $\lambda \in G_\varepsilon^n \cap D(n) \cap A_{\mathbb{Q}_m}$. If $y_1 \notin A_{\mathbb{Q}_m}$, there is a $\lambda' \in [0,1]^N \backslash L(R)$ such that $\lambda' \in \operatorname{conv}\{\lambda, y_1\}$ and, so, $y_1 \in E_1^n$. If $y_1 \notin D(n)$, we have according to (104) that $\|y_1 - \bar{y}_1 \cdot e_N\| < (\sqrt{|N|}/n) + d_n = d'_n$ and, so, $y_1 \in E_2^n$. If $y_1 \notin G_\varepsilon^n$, then $\bar{y}_1 < \varepsilon$ or $\bar{y}_1 > 1 - \varepsilon$ and hence we have according to (119) that $\varepsilon - (1/n) \leq \bar{y}_1 \leq 1 - (\varepsilon - (1/n))$ and, so, $y_1 \in E_3^n$. Now, let $y_2 \in B \backslash A$, so we have $y_2 \in G_\varepsilon \cap D(n) \cap A_{\mathbb{Q}_m}$ and there does not exist a

$\lambda \in G_\varepsilon^n \cap D(n) \cap A_{\mathbb{Q}_m}$ such that $y_2 \in C^n(\lambda)$. Let λ such that $y_2 \in C^n(\lambda)$. If $\lambda \notin A_{\mathbb{Q}_m}$, there exists an $\lambda' \in [0,1]^N \backslash L(R)$ such that $\lambda' \in \operatorname{conv}\{\lambda, y_2\}$ 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 hence $\bar{y}_2 = \bar{\lambda} + (\bar{y}_2 - \bar{\lambda}) < \varepsilon + (1/n)$ or $\bar{y}_2 < 1 - (\varepsilon + (1/n))$ and so, $y_2 \notin G_{\varepsilon+(1/n)}$. So, $y_2 \in E_3^n$. Hence, we have shown (141). Then, we get

$$\left|\int_A h^n(\lambda)d\lambda - \int_B h^n(\lambda)d\lambda\right| \leq \int_{A\backslash B} h^n(\lambda)d\lambda + \int_{B\backslash A} h^n(\lambda)d\lambda \tag{142}$$

$$\leq \int_{E_1^n \cup E_2^n \cup E_3^n} h^n(\lambda)d\lambda \tag{143}$$

$$\leq \sum_{k=1}^{3} \int_{E_k^n} h^n(\lambda)d\lambda \tag{144}$$

$$\leq \sum_{k=1}^{3} \nu(E_k^n)\mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)}) \tag{145}$$

$$= \mathcal{O}^\varepsilon(n^{-|N|+\frac{5}{8}}). \tag{146}$$

Here, (142) follows from $\int_A h^n(\lambda)d\lambda - \int_B h^n(\lambda)d\lambda = \int_{A\backslash B} h^n(\lambda)d\lambda - \int_{B\backslash A} h^n(\lambda)d\lambda$, (143) follows from (141), (144) is a standard rule of integration, (145) follows from $h^n(\lambda) = \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)})$ for all $\lambda \in G_{\frac{1}{2}\varepsilon}$ and (146) follows from $\nu(E_1^n) = \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+\frac{1}{8}})$ (see (118)) and we get in a similar fashion as for (118) 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 C.22 *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:*

$$\Gamma(\kappa) = \int_0^\infty e^{-t} t^{\kappa-1} dt, \qquad \textit{for all } \kappa > 0. \tag{147}$$

Proof We get

$$\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right) - \int_0^{n^{\frac{1}{4|N|}}/2t(1-t)} e^{-s} s^{\frac{1}{2}(|N|-3)} ds = \int_{n^{\frac{1}{4|N|}}/2t(1-t)}^\infty e^{-s} s^{\frac{1}{2}(|N|-3)} ds \tag{148}$$

$$\leq K \int_{n^{\frac{1}{4|N|}}/2t(1-t)}^\infty e^{-\frac{1}{2}s} ds \tag{149}$$

$$= K2e^{-n^{\frac{1}{4|N|}}/4t(1-t)} \tag{150}$$

$$\leq K2e^{-n^{\frac{1}{4|N|}}} \tag{151}$$

$$= \mathcal{O}(n^{-\infty}), \tag{152}$$

where $K > 0$. Here, (148) is a standard integration rule, (149) 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 $s > 1$, (150) follows from $\int_a^b e^{-\frac{1}{2}s} ds = -2(e^{-\frac{1}{2}b} - e^{-\frac{1}{2}a})$ for all $a \leq b$, (151) follows from $4t(1-t) \leq 1$ for all $t \in (0,1)$ and (152) follows from the fact that $(e^{-1})^{n^{\frac{1}{4|N|}}} = \mathcal{O}(n^{-\infty})$. This concludes the proof. □

Proof of Proposition C.3 We get

$$K_i^{avg,n}(R) = \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \frac{1}{n} \sum_{\lambda \in G_\varepsilon^n \cap D(n) \cap A_{\mathbb{Q}_m}} h^n(\lambda) + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{153}$$

$$= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] n^{|N|-1} \sum_{\lambda \in G_\varepsilon^n \cap D(n) \cap A_{\mathbb{Q}_m}} \int_{C^n(\lambda)} h^n(\lambda^*) d\lambda^* + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{154}$$

$$= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] n^{|N|-1} \int_{G_\varepsilon \cap D(n) \cap A_{\mathbb{Q}_m}} h^n(\lambda) d\lambda + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{155}$$

$$= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] n^{\frac{1}{2}(|N|-1)} (2\pi)^{\frac{1}{2}(1-|N|)} |N|^{-\frac{1}{2}} \tag{156}$$

$$\cdot \int_{G_\varepsilon \cap D(n) \cap A_{\mathbb{Q}_m}} \left(e^{-\frac{1}{2\bar{\lambda}(1-\bar{\lambda})} n \|\lambda - \bar{\lambda} \cdot e_N\|^2}\right) (\bar{\lambda}(1-\bar{\lambda}))^{\frac{1}{2}(1-|N|)} d\lambda + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}})$$

$$= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] n^{\frac{1}{2}(|N|-1)} (2\pi)^{\frac{1}{2}(1-|N|)} \tag{157}$$

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

$$
\begin{aligned}
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] n^{\frac{1}{2}(|N|-1)} (2\pi)^{\frac{1}{2}(1-|N|)} \mu(S_m) && (158)\\
&\quad \cdot \int_{\varepsilon}^{1-\varepsilon} \int_0^{d_n} e^{-\frac{1}{2t(1-t)} r^2 n} (t(1-t))^{\frac{1}{2}(1-|N|)} r^{|N|-2} dr dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}})\\
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] n^{\frac{1}{2}(|N|-1)} (2\pi)^{\frac{1}{2}(1-|N|)} \phi_m 2 \frac{\pi^{-\frac{1}{2}(1-|N|)}}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)} && (159)\\
&\quad \cdot \int_{\varepsilon}^{1-\varepsilon} \int_0^{d_n} e^{-\frac{1}{2t(1-t)} r^2 n} (t(1-t))^{\frac{1}{2}(1-|N|)} r^{|N|-2} dr dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}})\\
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \phi_m n^{\frac{1}{2}(|N|-1)} 2^{1\frac{1}{2}-\frac{1}{2}|N|} \frac{1}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)} \left(\frac{2}{n}\right)^{\frac{1}{2}(|N|-2)} && (160)\\
&\quad \cdot \int_{\varepsilon}^{1-\varepsilon} \int_0^{n^{\frac{1}{4|N|}}/2t(1-t)} e^{-s} s^{\frac{1}{2}(|N|-2)} (t(1-t))^{-\frac{1}{2}} \sqrt{\frac{t(1-t)}{2ns}} ds dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}})\\
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \phi_m \frac{1}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)} && (161)\\
&\quad \cdot \int_{\varepsilon}^{1-\varepsilon} \int_0^{n^{\frac{1}{4|N|}}/2t(1-t)} e^{-s} s^{\frac{1}{2}(|N|-3)} ds dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}})\\
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \phi_m \frac{1}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)} \int_{\varepsilon}^{1-\varepsilon} \left(\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right) + \mathcal{O}(n^{-\infty})\right) dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}})\\
& && (162)\\
&= \sum_{m=1}^{p^*} \phi_m E_{\mathbb{Q}_m}[X_i] + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}}) && (163)
\end{aligned}
$$

Here, (153) follows from Proposition C.2, (154) follows from Lemma C.20 and (155) follows from Lemma C.21, (156) follows from substitution of (66), (157) follows from the polar coordinate transformation $\lambda = t \cdot e_N + r\omega$ and $d\lambda = r^{|N|-2}|N|^{\frac{1}{2}} d(t, r, \omega)$, (158) follows from the fact that $\int_{S_m} d\omega = \mu(S_m)$, (159) follows from $\mu(S_m) = \phi_m \mu(S)$ and the well-known result that the hypersurface measure of an $|N|$-dimensional ball is given by

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

where Γ is defined in (147), (160) follows from the transformation $s = \frac{r^2 n}{2t(1-t)}$ and $dr = \sqrt{\frac{t(1-t)}{2ns}} ds$, (161) follows from canceling of some terms and (162) follows from Lemma C.22. This concludes the proof. □

D Proofs of Section 4

Proof of Proposition 4.1 (i) Follows immediately from (11), (26) and (27).

(ii) Follows immediately from (i) and the fact that the Aumann-Shapley value, if it exists, is the unique element of the fuzzy core (Aubin, 1981). □

Proof of Corollary 4.2 If $|N| = 2$, we get

$$S = \{(-\sqrt{0.5}, \sqrt{0.5}), (\sqrt{0.5}, -\sqrt{0.5})\}$$
$$S_m = \{z \in S : z_1 E_{\mathbb{Q}_m}[X_1] + z_2 E_{\mathbb{Q}_m}[X_2] \geq z_1 E_{\mathbb{Q}_\ell}[X_1] + z_2 E_{\mathbb{Q}_\ell}[X_2] \text{ for all } \ell \in \{1, \ldots, p^*\}\}.$$

So, $\mu(S) = |S| = 2$, and $\mu(S_m) = |S_m| = |\{z \in S : z_1 E_{\mathbb{Q}_m}[X_1] + z_2 E_{\mathbb{Q}_m}[X_2] \geq z_1 E_{\mathbb{Q}_\ell}[X_1] + z_2 E_{\mathbb{Q}_\ell}[X_2] \text{ for all } \ell \in \{1, \ldots, p^*\}\}|$. Note that for all $\ell \in \{1, \ldots, p^*\}$, it holds by construction that $E_{\mathbb{Q}_\ell}[X_1] + E_{\mathbb{Q}_\ell}[X_2]$ is the same (and equal to $r(e_N)$). So, if $z = (-\sqrt{0.5}, \sqrt{0.5})$, then $z_1 E_{\mathbb{Q}_m}[X_1] + z_2 E_{\mathbb{Q}_m}[X_2] \geq z_1 E_{\mathbb{Q}_\ell}[X_1] + z_2 E_{\mathbb{Q}_\ell}[X_2]$ for all $\ell \in \{1, \ldots, p^*\}$ holds when $E_{\mathbb{Q}_m}[X_2] = \max\{E_{\mathbb{Q}_\ell}[X_2] : \ell \in \{1, \ldots, p^*\}\}$ or, equivalently, $E_{\mathbb{Q}_m}[X_1] = \min\{E_{\mathbb{Q}_\ell}[X_1] : \ell \in \{1, \ldots, p^*\}\}$. Likewise, if $z = (\sqrt{0.5}, -\sqrt{0.5})$, then $z_1 E_{\mathbb{Q}_m}[X_1] + z_2 E_{\mathbb{Q}_m}[X_2] \geq z_1 E_{\mathbb{Q}_\ell}[X_1] + z_2 E_{\mathbb{Q}_\ell}[X_2]$ for all $\ell \in \{1, \ldots, p^*\}$ holds when $E_{\mathbb{Q}_m}[X_1] = \max\{E_{\mathbb{Q}_\ell}[X_1] : \ell \in \{1, \ldots, p^*\}\}$ or, equivalently, $E_{\mathbb{Q}_m}[X_2] = \min\{E_{\mathbb{Q}_\ell}[X_2] : \ell \in \{1, \ldots, p^*\}\}$. This concludes the proof. □

Proof of Theorem 4.3 It follows immediately from Definition 3.7 and Definition 3.10 that it is sufficient to show that for all $n \in \mathbb{N}$, all $P \in \mathcal{P}^n$, and all $R \in \mathcal{R}$, the properties *Translation Invariance*, *Scale Invariance* and *Monotonicity* are satisfied for the allocation rule $K^{path,P}(R)$ defined in (21).

We start with showing the property *Translation Invariance*. Let $P \in \mathcal{P}^n$, $n \in \mathbb{N}$, $j \in N$, $R = ((X_i)_{i \in N}, \rho) \in \mathcal{R}$ and $\widetilde{R} = ((\widetilde{X}_i)_{i \in N}, \rho) \in \mathcal{R}$ such that $(\widetilde{X}_i)_{i \in N} = (X_j + c \cdot e_\Omega, X_{-j})$ for some $c \in \mathbb{R}$. Let r $(\tilde{r})$ be the fuzzy game corresponding to R $(\tilde{R})$, as defined in (8). Then, we get

$$\begin{aligned}
\widetilde{r}(\lambda) &= \rho\left(\sum_{i \in N} \lambda_i \widetilde{X}_i\right) \\
&= \rho\left(\sum_{i \in N} \lambda_i \cdot X_i + c \cdot \lambda_j \cdot e_\Omega\right) \\
&= \rho\left(\sum_{i \in N} \lambda_i \cdot X_i\right) + c \cdot \lambda_j &\quad (164)\\
&= r(\lambda) + c \cdot \lambda_j, &\quad (165)
\end{aligned}$$

for all $\lambda \in [0,1]^N$, where (164) follows from *Translation Invariance* of ρ. We get

$$
\begin{aligned}
K^{path,P}(\widetilde{R}) &= \sum_{k=0}^{|N|n-1} [\widetilde{r}(P(k+1)) - \widetilde{r}(P(k))] \cdot e_{i(P,k)} && (166)\\
&= \sum_{k=0}^{|N|n-1} [r(P(k+1)) + c \cdot P_j(k+1) - r(P(k)) - c \cdot P_j(k)] \cdot e_{i(P,k)} && (167)\\
&= K^{path,P}(R) + c \cdot \sum_{k=0}^{|N|n-1} [P_j(k+1) - P_j(k)] \cdot e_{i(P,k)} && (168)\\
&= K^{path,P}(R) + c \cdot \sum_{k=0}^{|N|n-1} [P_j(k+1) - P_j(k)] \cdot e_j && (169)\\
&= K^{path,P}(R) + c \cdot [P_j(|N|n) - P_j(0)] \cdot e_j \\
&= K^{path,P}(R) + c \cdot e_j, && (170)
\end{aligned}
$$

where $P_j(k)$ is the j-th element of $P(k)$. Here, (166) follows from (21), (167) follows from (165), (168) follows from (21), (169) follows from $P_j(k+1) - P_j(k) = 0$ if $i(P,k) \neq j$ (see (20)) and (170) follows from Definition 3.5(i). 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 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 $\lambda \leq \lambda^*$. Combined with (8) and (20), this implies that $r(P(k+1)) - r(P(k)) \geq 0$ for all $k \in \{0, \ldots, |N|n-1\}$. It now follows immediately from (21) that $K^{path,P}(R) \geq 0$. This concludes the proof of *Monotonicity*. □