Axiomatizations of the Choquet integral on general decision spaces

A thesis submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy (Ph.D.) in Economics

MIKHAIL TIMONIN

School of Economics and Finance

Queen Mary University of London

21.09.2016

I, Mikhail Timonin, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:

Date: 21 September 2016

Details of collaboration and publications: (Timonin, 2016c,d,a,b, 2015b,a, 2014b,a)

Abstract

We propose an axiomatization of the Choquet integral model for the general case of a heterogeneous product set $X = X_1 \times \cdots \times X_n$. Previous characterizations of the Choquet integral have been given for the particular cases $X = Y^n$ and $X = \mathbb{R}^n$. However, this makes the results inapplicable to problems in many fields of decision theory, such as multicriteria decision analysis (MCDA), state-dependent utility (SD-DUU), and social choice. For example, in multicriteria decision analysis the elements of X are interpreted as alternatives, characterized by criteria taking values from the sets X_i. Obviously, the identicalness or even commensurateness of criteria cannot be assumed a priori. Despite this theoretical gap, the Choquet integral model is quite popular in the MCDA community and is widely used in applied and theoretical works. In fact, the absence of a sufficiently general axiomatic treatment of the Choquet integral has been recognized several times in the decision-theoretic literature. In our work we aim to provide missing results – we construct the axiomatization based on a novel axiomatic system and study its uniqueness properties. Also, we extend our construction to various particular cases of the Choquet integral and analyse the constraints of the earlier characterizations. Finally, we discuss in detail the implications of our results for the applications of the Choquet integral as a model of decision making.

Contents

Contents 4

Introduction 6

1 Non-additive models in decision analysis 9

1.1 Introduction . . . 10
1.2 The Choquet integral . . . 14
1.3 Problem statement and preview of the results . . . 19
1.4 Analysis of previous characterizations . . . 25
1.5 Summary . . . 30

2 Axiomatization of the Choquet integral – two-dimensional case 32

2.1 Introduction . . . 33
2.2 Axioms . . . 36
2.3 Representation theorem . . . 43
2.4 Proof preview . . . 46
2.5 Building additive value functions on **NW** and **SE** . . . 46
2.6 Aligning V^{SE} and V^{NW} . . . 49
2.7 Properties of the intersection of **SE** and **NW** . . . 55
2.8 Extending value functions to extreme points . . . 59
2.9 Constructing a global representation on X . . . 64
2.10 Constructing the capacity and the integral . . . 66
2.11 Uniqueness . . . 66
2.12 Necessity of the axioms . . . 69
2.13 Summary . . . 72

3 Axiomatization of the Choquet integral – n-dimensional case 73
3.1 Introduction . . . 74
3.2 Axioms and definitions . . . 74
3.3 Representation theorem . . . 81
3.4 Proof preview . . . 83
3.5 Additive representations on X^{S_a} . . . 84
3.6 Aligning cardinal representations for different X^{S_a} . . . 89
3.7 Constructing global representation on X . . . 91
3.8 Extending the representation to the extreme points of X . . . 100
3.9 Uniqueness . . . 102
3.10 Necessity of axioms . . . 105
3.11 Equispaced sequences . . . 106
3.12 Technical lemmas . . . 108
3.13 Essentiality and monotonicity . . . 109
3.14 Shape of $\{z_{ij} : i\,\mathbf{E}^z\,j\}$ . . . 110
3.15 Summary . . . 112

4 Extensions and discussion 113
4.1 Introduction . . . 114
4.2 Extensions . . . 115
4.3 Learning the Choquet integral . . . 125
4.4 Interpretations and discussion . . . 134
4.5 Summary . . . 139

Conclusion 141

A Extended literature review 145
A.1 Decisions under uncertainty . . . 146
A.2 MCDA . . . 149
A.3 Characterizations of the Sugeno integral . . . 151

Bibliography 155

Introduction

Axiomatic analysis is important for decision models as it allows for a much deeper understanding of their mechanics. Normally, such analysis pursues one of two approaches: either introducing a novel model, which is based on a set of easily acceptable premises (axioms), or performing an in-depth analysis of a well-known model. The latter case aims to understand the model better, reveal its previously unknown properties, and relate it to other models in the field.

This thesis is dedicated to axiomatic analysis of the Choquet integral. Although it has been widely used since the late 1980s, its theoretical foundations are significantly under-developed. Particularly in the area of multicriteria decision analysis, where the integral enjoys a lot of theoretical and applied interest, it has been long recognized that the theoretical foundations of the model are weak, and the proper decision-theoretic analysis is missing. Several attempts to fill this void have been made in the last 15 years; however, none of them has solved this problem completely. rewrite bette

In this thesis we aim to achieve the following:

A general characterization of the Choquet integral. We will show that earlier axiomatizations of the model catered only for a very special case, which is not applicable in many decision problems, particularly the multicriteria ones. This has been noted several times in the literature – however, no general result has been developed to date.

A novel axiomatic system which generalizes earlier results. The axiomatic systems developed in the previous characterizations are not suitable for the general case that is treated here. Hence, a new set of premises is required. It is interesting to relate these new conditions to the old ones, to see what constraints are being lifted as a result of this generalization.

Analysis of the uniqueness properties. A matter of paramount importance in any decision-theoretic analysis is the question of the model uniqueness. If two instances of the same model (e.g. with two different sets of parameters) can represent the same preference relation, are they related to each other and how?

Again, we expect a significant departure from the earlier result due to the more general case we are considering here.

Extensions of the main characterization result to the ordinal and cardinal special cases of the Choquet integral. The Choquet integral lists a large number of decision models as its special cases. Among those are ordinal models, such as MIN, MAX, or the order statistic (kth smallest element), and the cardinal ones, such as the case when the preferences of the decision maker are in some sense convex or concave. We will extend our results to each of these special cases.

Analysis of the implications of our results for learning of the Choquet integral models. As mentioned previously, the Choquet integral is widely used in applied work. Naturally, one of the most important problems in this context is the process of estimating its parameters, also called "model learning". Here the lack of proper theoretical underpinnings of the model frequently led to using a flawed methodology. We discuss the problems of the current approaches and reveal some new aspects implied by our results.

Discussion of our results. Finally, we discuss the implications our results have for different areas of decision theory and potential connections with other fields, such as psychology. We discuss an interesting concept of implied commensurability, which allows us to construct meaningful correspondences between seemingly incomparable quantities, and its interpretations in various decision problems.

Chapter 1

Non-additive models in decision analysis

1.1 Introduction

Decision theory is a field positioned between economics, psychology, and mathematics. The purpose of decision theory is to introduce and analyse mathematical models of human decision making used to perform forecasting, behavioral and economical analysis, and, in recent years, automated decision making. Hence, the models must be not only accurate, but also highly tractable. The field originated perhaps in economics via the works of Ramsey, De Finetti, von Neumann, and Morgenstern; however, many results were introduced by psychologists, mathematicians, and philosophers such as Kahneman, Tversky, Suppes, Luce, Fishburn, and Krantz. Decision theory includes several sub-fields, such as decision making under uncertainty (DUU), multi-criteria decision analysis (MCDA), and intertemporal decision making.

There is a large number of models used and studied in decision theory, but perhaps the most prominent is the class of the additive models. A well-known example is the expected utility model used in DUU. Its origins are normally attributed to Bernoulli, who proposed the idea to resolve the so-called St. Petersburg paradox. The idea gained prominence after being used in the book of Von Neumann and Morgenstern (1944). They also proposed one of the earliest "axiomatizations" of the expected utility model – that is, a set of conditions characterizing the preferences of the decision maker, under which he would be making decisions *as if* he was maximizing the expected value of some utility function defined on the set of potential outcomes. The authorship of the earliest axiomatization is still a subject of historical research, however the consensus at the present moment attributes it to individually Nash and Marschak (Bleichrodt et al., 2016). Another notable axiomatization of the expected utility is due to Savage (1954), who was the first to provide an axiomatization in a purely "subjective" context, i.e. without referring to any "objective" probabilities in the frequentist sense (e.g. the probability of a horse winning a race as opposed to the probability of getting 6 in a dice roll).

The key elements of the expected utility model in decision under uncertainty are: the *states of the world* – situations which might occur in the future, however the decision maker is not able to predict with certainty which particular one is going to happen; the *outcomes* – consequences that the decision maker would bear as result of his action and the actual state of the world, and finally the *acts*, which are normally understood as mappings between the states of the world and the outcomes – see Fishburn (1970) for alternative formulations.

Let Y be the set of outcomes, S the set of the states of the world (which we hereafter assume to be finite), and the functions $f : S \to Y$ be "acts". Then, the axioms of the expected utility model guarantee the existence of a *probability* p on the set of *events* (2^S), and a *utility* function $U : Y \to \mathbb{R}$, such that f is weakly *preferred* to g (we write $f \succcurlyeq g$), if and only if

$$\sum_{s \in S} p(s)U(f(s)) \geq \sum_{s \in S} p(s)U(g(s)). \tag{1.1}$$

Expected utility transgressed the boundaries of a "descriptive" model and became an extremely influential "normative" paradigm in economics, remaining in this position to the present day. Besides being simple and having a strong intuitive appeal from the descriptive perspective, it also turned out that axioms of the expected utility model have a very strong normative appeal, in other words they seem to prescribe a model of the "rational" behaviour.

A close cousin of the expected utility model is the so-called additive value model. This was introduced by Debreu (1959) and later axiomatized in a number of papers, which culminated in a seminal book by Krantz et al. (1971). Additive value model is different from the expected utility model in that it uses multiple utility (value) functions. The other name of this is multi-attribute utility theory (MAUT) (Keeney and Raiffa, 1976). As the name suggests, this model is mainly used in multicriteria decision analysis. In this case, the preference relation is defined on a set of multi-attribute alternatives $X = X_1 \times \ldots \times X_n$, and for all $x, y \in X$, we have that $x \succcurlyeq y$ if and only if

$$\sum_{i=1}^{n} V_i(x_i) \geq \sum_{i=1}^{n} V_i(y_i). \tag{1.2}$$

The key assumption of classical additive models, whether the expected utility one or the additive value function model, is the *independence* condition. Fishburn and Wakker (1995) and Moscati (2016) give an exciting and enjoyable look into the origins of this condition. Savage (1954) called this axiom the "sure thing principle", whereas in the other literature it is normally called independence or separability. We would be using the term independence from now on.

In MCDA context independence means that if two alternatives share the same values on some coordinates, then the preference between the two should not be affected by the levels of these common values. In decision under uncertainty the condition is known as the sure-thing principle, for which Savage (1954) gives the following informal example. A businessmen considers buying a property, and comes to a conclusion that if a Democratic candidate wins the upcoming election, buying would be preferable. However, if a Republican candidate wins, buying still would be preferable. The sure-thing principle than says that he should buy even without knowing which candidate will win.

Definition 1. *Let* $\succcurlyeq$ *b a binary relation on* $X = X_1 \times \cdots \times X_n$. *We say* $\succcurlyeq$ *satisfies* independence, *if for any* $x_i, y_i \in X_i$, $a_{-i}, b_{-i} \in X_{-i}$, *we have*

$$x_i a_{-i} \succcurlyeq x_i b_{-i} \iff y_i a_{-i} \succcurlyeq y_i b_{-i}. \tag{1.3}$$

The condition has a strong normative and descriptive appeal. However, it turned out to be the most controversial assumption of the additive models, including the expected utility model. In decision under uncertainty Ellsberg (1961) and Allais (1953) used experimental data to show that this condition was systematically violated and yet the experiment subjects considered their behaviour rational. This discovery led to the separation of the notions "uncertainty" and "risk" in decision theory, with the former signifying a situation where the probabilities with which each state occurs are unknown, whereas the latter the opposite situation. Experimenters asked the subjects to choose between gambles involving urns with a known proportion of coloured balls, and those where the proportions were unknown. Participants showed strong violation of independence, always steering away from the urns with an unknown composition. This phenomenon came to be known as "uncertainty aversion".

In multicriteria contexts, violations of independence do not require any complicated experiments and were more or less accepted as given from the very beginning – see Keeney and Raiffa (1976), and Fishburn and Wakker (1995). It became clear that, both in DUU and MCDA, new models were required to deal with the data which did not fit the additive model.

A new generation of decision models emerged in the 1980s. The most important for our work is the class of "rank-dependent" utility models, and in particular the Choquet integral, which is the primary focus of this thesis. The Choquet integral is nowadays widely used in decision analysis, in particular MCDA (Grabisch and Labreuche, 2008), although its use is still somewhat restricted due to both methodological problems and difficulties in practical implementation. Rank-dependent models first appeared in the context of decisions under uncertainty. As discussed above, the Ellsberg (1961) paradox has shown that people can violate Savage's axioms and still consider their behaviour rational. First models accounting for the so-called uncertainty aversion observed in this paradox appeared in the works of Quiggin (1982) and others – see Wakker (1991b) for a review. One particular generalization of the expected utility model (EU) characterized by Schmeidler (1989) is the Choquet expected utility (CEU), where the probability is replaced by a non-additive set function (called *capacity*) and integration is performed using the Choquet integral.

Since Schmeidler's paper, various versions of the same model have been characterized in the literature (e.g. Gilboa, 1987; Wakker, 1991a). CEU has gained some momentum in both theoretical and applied economic literature, being used mainly for analysis of the problems involving Knightian uncertainty. At the same time, rank-dependent models, in particular the Choquet integral, were adopted in the multicriteria decision analysis. Here the integral gained large popularity due to the tractability of non-additive measures in this context (see Grabisch and Labreuche, 2008 for a review). The model allowed for various preferential phenomena, such as criteria interaction, which were impossible to reflect in the traditional additive models. Pioneering works in this area are Sugeno (1974); Mori and Murofushi (1989); Murofushi (1992); Murofushi and Soneda (1993) and Murofushi and Sugeno (1992). The toolbox for interpretation and application, as well as the theoretical understanding of the Choquet integral in MCDA were greatly expanded and popularized in the last 25 years due to the works of Michel Grabisch, Christophe Labreuche, and others (see Grabisch and Labreuche, 2008, 2005; Labreuche and Grabisch, 2003). The concept of a capacity or fuzzy measure became very popular in MCDA because of its tractability in this context. Interaction between criteria is a common and widely recognized phenomenon, and non-additivity of the capacity turned out to be a very nice facility for modelling it. Popularity of the Choquet integral seems to be growing in the recent years. It is also actively used in areas such as machine learning for classification and ranking problems. However, the decision-theoretic foundations of the Choquet integral in MCDA had a large gap, which this work aims to close. Particularly, it quickly became obvious that the existing axiomatizations of the integral by Schmeidler (1989) and others are not general enough to be used in the MCDA context.

1.2 The Choquet integral

As discussed in the previous chapter, the Choquet integral is used to model decision problems in a variety of fields. Below we provide a brief introduction to its use in decision under uncertainty and MCDA. Let us start by giving the definition of a capacity and the Choquet integral.

1.2.1 Definitions

Definition 2. *Let N be a set (of states, criteria, etc) and 2^N its power set.* Capacity (non-additive measure, fuzzy measure) *is a set function $\nu : 2^N \to \mathbb{R}_+$ such that*

1. $\nu(\varnothing) = 0$;
2. $A \subseteq B \Rightarrow \nu(A) \leq \nu(B), \ \forall A, B \in 2^N$.

We also assume that capacities are normalized, i.e. $\nu(N) = 1$.

Definition 3. *The* Choquet integral *of a function $f : N \to \mathbb{R}$ with respect to a capacity ν is defined as*

$$C(\nu, f) = \int_0^{\infty} \nu(\{i \in N \colon f(i) \geq r\})dr + \int_{-\infty}^{0} [\nu(\{i \in N \colon f(i) \geq r\}) - 1]dr. \quad (1.4)$$

If $N = \{1, 2, \ldots, n\}$ is finite, and all values $(f_1, \ldots, f_n)$ of $f : N \to \mathbb{R}$ are non-negative, the definition can be expressed as

$$C(\nu, (f_1, \ldots, f_n)) = \sum_{i=1}^{n} (f_{(i)} - f_{(i-1)}) \nu(\{j \in N \colon f_j \geq f_{(i)}\}), \tag{1.5}$$

where $f_{(1)}, \ldots, f_{(n)}$ is a permutation of $f_1, \ldots, f_n$ such that $f_{(1)} \leq f_{(2)} \leq \cdots \leq f_{(n)}$, and $f_{(0)} = 0$.

1.2.2 The Choquet integral for preference representation

Decision under uncertainty As discussed in the previous section, the Choquet integral was initially introduced in the context of decision making under uncertainty. As before, Y is the set of outcomes, and S is the set of the states of the world, and acts are functions $f : S \to Y$. We say that $\succcurlyeq$ on the set of all acts can be represented by a Choquet integral, if there exists a *capacity* ν and a *utility* function $U : Y \to \mathbb{R}$, such that

$$f \succcurlyeq g \iff C(\nu, (U(f(s_1)), \ldots, U(f(s_n))) \geq C(\nu, (U(g(s_1)), \ldots, U(g(s_n))). \tag{1.6}$$

MCDA Multicriteria decision analysis (MCDA) is a sub-field of decision theory which is focused around making choices between *alternatives* characterized by several *attributes*. For example, consider the example in Table 1.1.

	Top speed	Appearance	Miles per gallon	Comfort	Price
(A) BMW M3	225	Great	21	***	56000
(B) VW Golf	180	Nice	27	**	25000
(C) Volvo	200	Fair	27	***	35000
(D) Toyota Prius	140	Ugly	40	*	20000

Table 1.1: A multicriteria decision problem

We also know that the decision maker prefers A to B, B to C, and C to D ($A \succ B \succ C \succ D$). The decision space is a product set with the dimensions "Top speed", "Appearance", "Miles per gallon", "Comfort", and "Price". Some of the attributes levels are linguistic categories, some are ordinal evaluations, some are numbers. It is important to note that generally the levels of different attributes are not mutually comparable even if they are numerical. Indeed, it is hard to compare "\$25 000" with "Nice appearance" or with "27 mpg".

Formally, the set of the alternatives is a product set $X = X_1 \times \cdots \times X_n$, where sets X_i are the sets of *attribute levels*. We say that $\succcurlyeq$ on $X = X_1 \times \cdots \times X_n$ can be represented by a Choquet integral if there exists a capacity ν and functions $f_i : X_i \to \mathbb{R}$, called value functions, such that

$$x \succcurlyeq y \iff C(\nu, (f_1(x_1), \dots, f_n(x_n)) \geq C(\nu, (f_1(y_1), \dots, f_n(y_n)). \tag{1.7}$$

Mathematically, the connection between MAUT and decision making under uncertainty has been known for a long time. In the case when the number of states is finite, states can be associated with criteria. Accordingly, acts correspond to multicriteria alternatives. Finally, the sets of outcomes at each state can be associated with the sets of criteria values. However, this last transition is not trivial. It is commonly assumed that the set of outcomes is the same in each state of the world (Savage, 1954; Schmeidler, 1989). In multicriteria decision analysis the opposite is true, as has been shown above.

State-dependent utilities The problem can be seen from a different perspective. In decision making under uncertainty, the existing characterizations of the Choquet integral assume that the set of outcomes is exactly the same in all states; moreover, the outcomes in all states are ranked in the same way! This is clearly a very strong assumption which is rarely met in practice. The problem was raised as early as 1971, when Aumann writing to Savage gave an example where the value of an outcome might strongly depend on the state that gets realized (Drèze, 1987). As noted by Karni (1985), the preference relation might display ordinal state dependence, in which case the underlying state may affect the decision makers preferences by altering his ordinal ranking of the consequences; or cardinal state dependence, by altering his risk attitudes; or both.

In this case, we are not able to use the same utility function in every state any more, and the representation of preferences becomes very similar to the multicriteria case. Formally, Y is the set of outcomes, and S is the set of the states of the world, and acts are functions $f : S \to Y$ as previously. However, in the case of the expected utility we now say that $\succcurlyeq$ maximizes the state-dependent expected utility, if

$$\sum_{s\in S} p(s)U(s, f(s)) \geq \sum_{s\in S} p(s)U(s, g(s)). \tag{1.8}$$

We take a further look at the existing axiomatizations of the state-dependent models in Section 4.4.3, at this point it is sufficient to note that they commonly suppose existence of some mapping between elements in various states or a global ordering between these elements. Thus effectively, state-dependence is reduced to state-independence. If such mapping is not assumed, then contrary to the state-independent expected utility, as in equation (1.1), the factorization into p and U in equation (1.8) is not unique. We say that the probabilities and utilities are *confounded*, in the sense that by simple algebraic transformations we can obtain another representation with an arbitrary probability distribution (provided the mass is non-null for all required elements). This mirrors the situation in MCDA where the so-called "weighted average" model with arbitrary weights can be obtained from the additive value (1.2) by simple change of variables. It was shown that in order for the weighted average model to be meaningful, all dimensions must be evaluated on the same scale (Bouyssou et al., 2000, Chapter 6.1), which is clearly usually not the case.

Returning to the Choquet integral, we say that $\succcurlyeq$ on the set of all acts can be represented by a Choquet integral, if there exists a capacity ν and a *state-dependent utility* function $U : Y \times S \to \mathbb{R}$, such that

$$f \succcurlyeq g \iff C(\nu, (U(s_1, f(s_1)), \ldots, U(s_n, f(s_n))) \geq C(\nu, (U(s_1, g(s_1)), \ldots, U(s_n, g(s_n))). \tag{1.9}$$

Again, mathematically state-dependent utility is very similar to multicriteria models. This is so, because every dimension of the decision space now has its own utility function. In the MCDA context, dimensions were associated with various attributes whereas here they are outcomes corresponding to various possible states of the world. Again, these dimensions do not have to be identical – we can even assume a separate set of outcomes Y_s for each state $s \in S$ – and even if they are, the utility functions do not have to be identical, as in the state-independent case.

1.2.3 A unified modelling framework

Before proceeding further, it would be helpful to provide a unified framework for the decision problems we will be considering. Currently, we are considering three separate areas: decision under uncertainty with state-dependent and state-independent utilities and MCDA.

In case of (state-independent) decision making under uncertainty, we have a set of *states of the world* S, which we assume finite in this work, a set of *outcomes* Y and the preference relation $\succcurlyeq$ is defined on the set of *acts*, i.e. functions $f : S \to Y$. In MCDA, the set of *alternatives* is a product set $X = X_1 \times \cdots \times X_n$, where sets X_i are the sets of *attribute levels*. The preference relation $\succcurlyeq$ is on $X = X_1 \times \cdots \times X_n$. Finally, in state-dependent DUU, we have $\succcurlyeq$ defined on the set of all acts $f : S \to Y$ as before, but the evaluation of the outcome now depends on the state as well. In other words, the utility function now depends on the state $U : Y \times S \to \mathbb{R}$.

The alternative way to state this is to say that the preference relation is defined on a product set X. In the DUU case, X is a homogeneous product $X = Y^n$ (note that functions are vectors in Y^n, where $|S|= n$). In the MCDA case, we have a heterogeneous product $X = X_1 \times \cdots \times X_n$ from the start, and in the state-dependent DUU case, we can say that $X = Y_1 \times \cdots \times Y_n$, where Y_i contain "state–prize" elements. This is a more general setup, since outcomes can actually be different in various states, and X is a heterogeneous product set again. This naturally suggests that we use n utility functions $U_i : Y_i \to \mathbb{R}$, instead of a single $U : S \times Y \to \mathbb{R}$.

To conclude, we consider problems where the preference relation $\succcurlyeq$ is defined on a product set $X = X_1 \times \cdots \times X_n$, where the sets X_i can be identical ($X_i = Y$), as in the DUU and state-dependent DUU cases, or distinct (MCDA and state-dependent DUU). If all dimensions are identical, we can have a single utility function or n value (or utility) functions, whereas in the heterogeneous case we have n value (or utility) functions, one for each dimension. We are looking to characterize the Choquet integral representation in this setting.

1.3 Problem statement and preview of the results

The main result of this thesis is an axiomatization of the Choquet integral. Our results are significantly more general than previous ones and allow for a wider range of applications. We also give extensions of our main characterization for the various special cases of the Choquet integral, look at interpretations of our results, and analyse the implications they have for learning of the model.

As we have seen in the previous section, the Choquet integral is effectively a weighted sum; however, the weights actually depend on the ranking of the aggregated elements – see equation (1.5). In decision under uncertainty, ranking would involve comparing utilities of the outcomes attained at various states, which is a perfectly correct operation, as soon as we accept state-independence (or the possibility to compare the outcomes in different states).

However, in MCDA the ranking stage of rank-dependent models would amount to much more challenging comparisons of the levels of different attributes, e.g. comparing appearance to the level of fuel consumption, and maximal speed to comfort. The traditional additive model (Debreu, 1959; Krantz et al., 1971) (see Section 1.1) only implies meaningful comparability of intervals (i.e. differences) on different dimensions, but not of the absolute levels. However, in rank-dependent models such comparability seems to be a necessary condition. This constitutes the main difficulty and novelty in the problem of axiomatic characterization of the Choquet integral in MCDA.

Similarly, in state-dependent utility, ranking would involve comparing outcomes in various states to each other. Some authors assume the possibility of such comparison (Karni and Schmeidler, 2016, and references therein); however, generally it is clear that such approach is not without flaw. In effect, it reduces the model to the state-independent case again.

1.3.1 Brief analysis of the previous results and their restrictions

Before going into the detailed analysis of the previous approaches in Section 1.4 and Appendix A, we will focus on a few key elements of the previous axiomatizations and the difficulties that arise when transferring these results to the more general setting we are considering here.

First is the notion of a *constant act*, which is simply an element of X which has the same level on every dimension, e.g. $(\alpha, \dots, \alpha)$ where $\alpha \in Y$. Constant acts were traditionally used to construct a new relation on the set of outcomes[1] Y, thus $\alpha \succcurlyeq \beta$ iff $(\alpha, \dots, \alpha) \succcurlyeq (\beta, \dots, \beta)$, which is then used to introduce the central notion of **comonotonicity**. The definition of comonotonicity can be slightly different depending on the context. We give a general definition as follows.

Definition 4. *Let A, B be two sets and $f : A \to B$ and $g : A \to B$ two functions. Moreover, $\succcurlyeq$ is a weak order on B and $\succ$ is its asymmetric part. We say that f and g are* comonotonic *if for no $a, b \in A$ we have $f(a) \succ f(b)$ and $g(b) \succ g(a)$.*

An example of constant acts and comonotonic areas is shown in Figure 1.1.

Furthermore, the majority of the previous results are based on the comonotonic variant of the independence condition, i.e. they state that independence holds on comonotonic subsets of X. In here we find it more instructive to look at the comonotonic tradeoff consistency condition used in Wakker (1991a), see Figure 1.2.

Roughly speaking, comonotonic tradeoff consistency states that if in one state, the "interval" between outcomes α and β is the same as the "interval" between outcomes γ and δ, then it must be true also for another state. The equality of "intervals" is attested by the fact that changes from α to β and from γ to δ compensate exactly the same change on another dimension (the magenta segment). The condition holds when all involved points either belong to the same comonotonic subset (as defined by the ordering of the constant acts), or two pairs of points belong to one comonotonic subset and two to another, as in Figure 1.2.

[1] Note that this need not be the "real", i.e. the DM's ordering of outcomes, although it would later be shown to be. More on this in Section 4.4.3

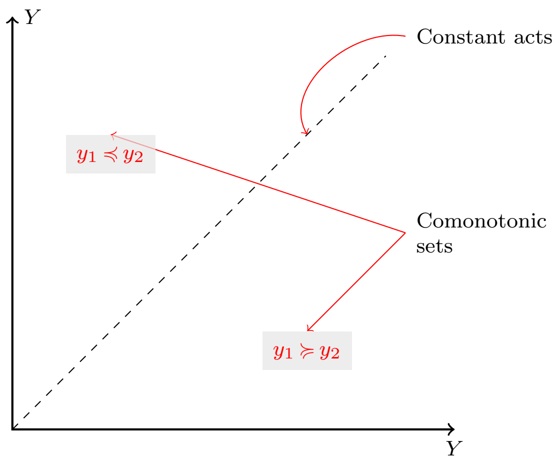


Figure 1.1: Comonotonic subsets and constant acts

Now, if we consider the general case $X = X_1 \times \cdots \times X_n$ which we use in this thesis, the following notions become undefinable:

1. Constant acts, as we cannot have a vector in X with *identical* coordinates, since the dimensions are different sets

2. Comonotonicity, as we are not able to compare elements from different dimension sets a priori or derive such an order from constant acts

3. Using the same "outcomes", i.e. levels on different dimensions, as in the tradeoff consistency condition. Again, this is due to heterogeneity of the decision space.

An illustration of the first two issues is given in Figure 1.3.

1.3.2 Brief introduction of the results

This thesis presents the most general axiomatization of the Choquet integral, relaxing the homogeneity of the decision space. This result should help to fill the theoretic gap in MCDA literature and also can be used in decision under

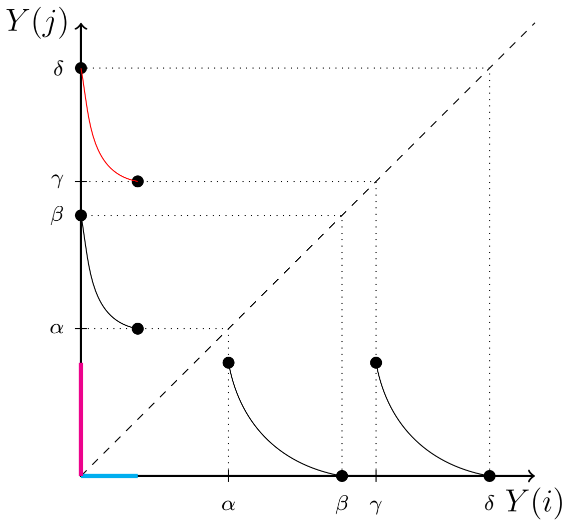


Figure 1.2: Comonotonic tradeoff consistency

uncertainty with state-dependent utilities. Our axioms do not involve any additional constructions or mapping functions, they are given strictly using the preference relation itself. All axioms are testable in the normal sense – a single counterexample is sufficient to refute the condition.

Formally, we propose a representation theorem for the Choquet integral model in the MCDA context. Binary relation $\succcurlyeq$ is defined on a heterogeneous product set $X = X_1 \times \cdots \times X_n$. In multicriteria decision analysis (MCDA), elements of the set X are interpreted as alternatives, characterized by criteria taking values from sets X_i. Previous axiomatizations of the Choquet integral model have been given for the special cases of $X = Y^n$ and $X = \mathbb{R}^n$. As shown above, the results are not transferable to the general case, as they rely on a very specific construction based on "constant acts" and "comonotonicity", artefacts of the decision space homogeneity.

Our approach is to decompose the conditions used in the earlier results, particularly comonotonic tradeoff consistency, into atomic components and generalize each one so that they could be used in our setting. There are two main directions of generalization:

Generalized comonotonicity. Comonotonicity can be viewed as a facility to divide the decision space into subsets within which we can construct additive

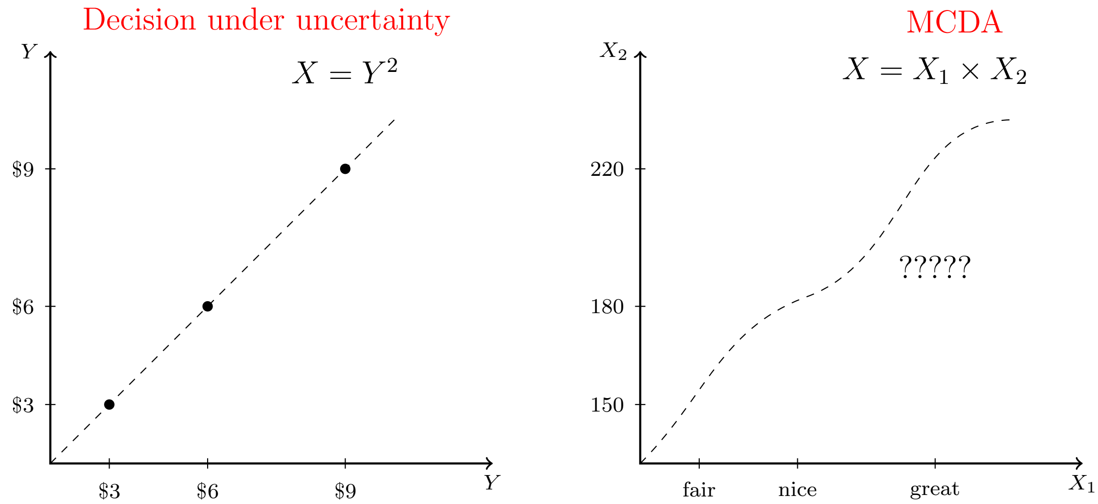
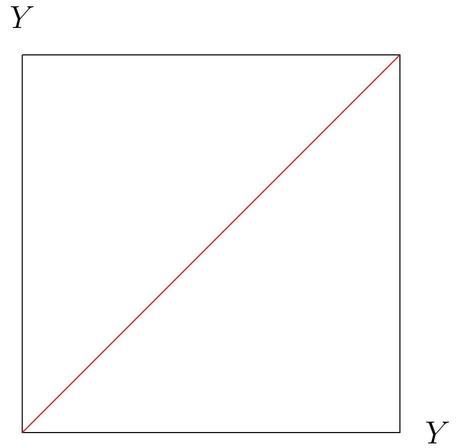


Figure 1.3: $X = Y^n$ vs $X = X_1 \times \ldots \times X_n$

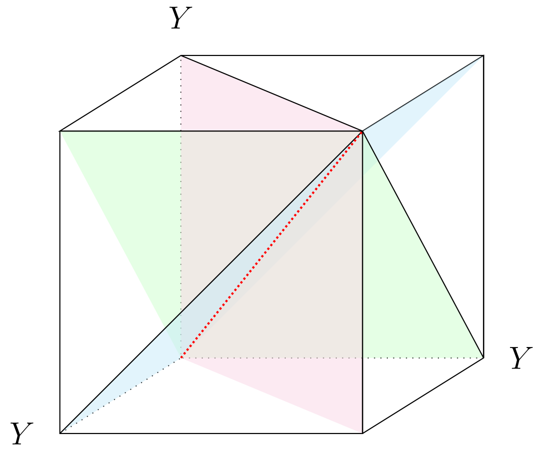
representations. However, it is unnecessary restrictive, by prescribing the certain shape and location of this subsets. It also requires the decision space to be a homogeneous product set. We can part with some of those restrictions by using a weakened scheme of decision space partitioning.

Generalized tradeoff consistency. The main complication with using the tradeoff consistency condition is that it requires the same elements to be present in different dimension sets. However, it is possible to decompose this condition into two parts. First is "intra-coordinate" tradeoff consistency, which Wakker (1991b) calls "generalized triple cancellation", and the second is "inter-coordinate" tradeoff consistency, a condition very similar to the original axiom used in Krantz et al. (1971), on which, in turn, the tradeoff consistency itself is based. None of them require any additional assumptions about the structure of the decision space.

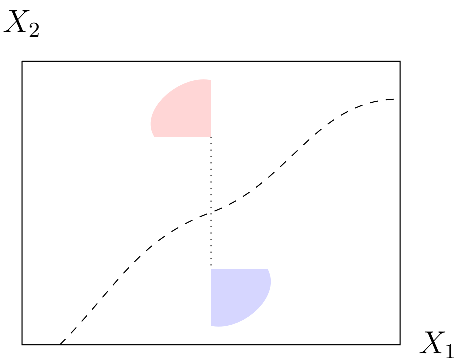
The more interesting of the two is certainly the generalized version of comonotonicity, as it gives way to a much wider range of geometric layouts of the decision space. Previously, comonotonic regions have been dividing the set X into symmetric comonotonic cones, as in Figure 1.4, where cases (a) and (b) depict the comonotonic case, whereas (c) and (d) show more general partitions of the set X.



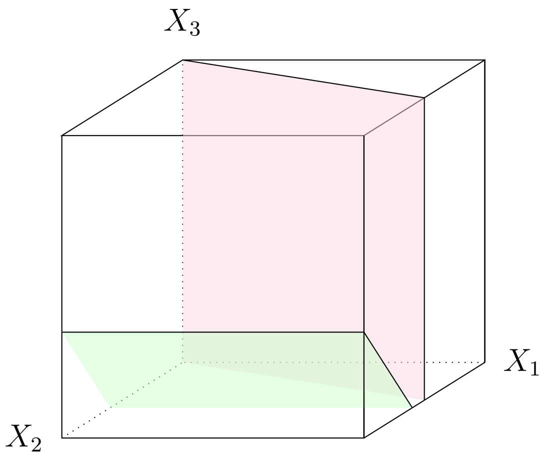
(a) Homogeneous case – two dimensions



(b) Homogeneous case – three dimensions



(c) Heterogeneous case – two dimensions



(d) Heterogeneous case – three dimensions

Figure 1.4: Generalization of comonotonic partitioning

Chapters 2 and 3 of the thesis contain the axiomatization result, first in two and later in n dimensions. There are significant differences between the two which prompt a separated presentation. This is further discussed in Sections 2.1 and 3.1. We present and discuss the axioms and provide a detailed construction of the model. In Chapter 4 we extend our results to the special cases of the Choquet integral. The most interesting ones are ordinal – MIN, MAX, OS_k (order statistic), and $p^{\mathcal{AB}}$ – a lattice polynomial. Cardinal special cases include convex and concave capacities. Finally, we discuss interesting interpretations of our results in various fields of decision analysis.

Apart from the theoretical interest, the characterization also brings to light some interesting problems related to the learning of the Choquet integral models. The common approach in the applied literature currently is to assume the value functions as given and commensurate and to fit the model as a parametric one, using the capacity as the parameter. However, for a proper representation we should also be fitting value functions, thus using a non-parametric representation model (functions are completely non-restricted). Traditionally, the learning theory states that non-uniqueness of the model which is chosen from some parametric family comes from the imbalance between the model complexity and the number of learning data points. Naturally, complicated models overfit small datasets which leads to problems with interpretation of the model. With the Choquet integral, however we face a different problem, somewhat orthogonal to this. Now, if data does not adhere to a certain *structure* we would not get a unique capacity from the learning process no matter how many data points we get. The root cause of this problem is that the lack of required irregularities in the data brings it closer to the additive case, i.e. the "weighted average" model, where the weights, as mentioned previously, can be completely arbitrary. More details on this is provided in Section 4.3.

1.4 Analysis of previous characterizations

This section provides an overview of the known results in the field, focusing on their main shortcomings and useful ideas, upon which our own characterizations are built. A more detailed analysis of the literature can be found in Appendix A.

1.4.1 Decisions under uncertainty

Historically, decision making under uncertainty was the first area in which the Choquet integral was used as a decision model. It remains the central sub-field of decision theory, especially in economics. The first and the majority of later axiomatizations of the Choquet integral were developed for decision making under uncertainty. Following the pioneering works of Quiggin (1982) and Schmeidler (1989), the model gained a lot of popularity in the community and is widely used both in theoretical and applied works. It is therefore important to analyse the existing results and show their weak spots and limitations.

The majority of the results rely heavily on the notions of a "constant act", "comonotonicity", and some weakening of Savage's sure-thing principle (independence). There are also several frameworks: using lotteries or elements of some set Y as outcomes, having an infinite or finite number of states, assuming the existence of a "certainty equivalent" or not, making topological or algebraic assumptions on the set of outcomes. This variation leads to a great variety of techniques being used to demonstrate the results, however, the cornerstone of the results remains constant, and can be summarized using just two words – "comonotonic independence".

One other characteristic feature of the DUU results is that they rely on a very particular decision space, which is a homogeneous product set Y^n. Moreover, the elements of every dimension, i.e. the outcomes in every state, are ordered in exactly the same way, both ordinally and cardinally. It turns out that this state-independence is actually implied by homogeneity and a particular choice of constant acts as the border between "additive" subsets. We return to this question in Section 4.4.3.

Some publications attempted at state-dependence both in additive and rank-dependent models. However, the results are not truly satisfactory when applied to our case. Assuming a mapping between elements of different dimensions or some universal order which encompasses elements from various sets is not easy to justify, especially in the MCDA context, and what is more important, such an order or mapping is not really observable in terms of preferences. We return to this question in Section 4.4.3.

1.4.2 MCDA

As discussed already, the Choquet integral found a lot of success in the MCDA community. This is perhaps mostly due to the fact that the violation of independence, manifested via interaction between criteria, is much more obvious in such setting, so there is no need for complicated experiments (like that of the Ellsberg paradox) to demonstrate it. In fact, as pointed out in Fishburn and Wakker (1995), discussions of the representability in the additive model were taking place as early as 1892.

The popularity of the Choquet integral in MCDA started in the early 1990s and is not showing any signs of abatement. Many interesting results have been achieved in this area, related to interpretation and learning of the capacity (see Grabisch and Labreuche, 2008; Grabisch et al., 2008), which we further touch upon in Section 4.3. Due to the abundance of theoretical tools and its natural fit to multicriteria problems, the integral quickly became very popular in practical work as well. However, one thing that was missing on the theoretical side is a proper conjoint axiomatization of the model. The problem was raised in the literature a few times (e.g. Bouyssou et al., 2009) and this work aims to fill the gap.

Although it was desirable to obtain axiomatization of the Choquet integral in this context, the problem turned out to be not as straightforward as it might seem. It was recognized early on that the fundamental complication and also a fundamental difference from the characterizations used in DUU was the heterogeneity of the decision space. In the context of MCDA, the issue is often referred to as the problem of "criteria incommensurability". The ranking stage of the Choquet integral, and also the notions of the "constant act" and "comonotonicity" in the characterization, all require the dimensions to be level comparable. In terms of state-dependent DUU, we should be able to say something like "outcome x in state s is better than outcome y in state t", which is not always possible. In MCDA terms, it is even more apparent. For example, if the alternatives are characterized by colour and shape, we should be able to make statements like "black is better than a circle".

With this problem in hand, almost all axiomatizations of the Choquet integral and the closely related Sugeno integral to date *assumed* criteria commensurateness. Thus, the problem immediately reduced to the one solved in the DUU context. Several authors in the MCDA community took a somewhat different direction and created axiomatizations of the Choquet integral as an *aggregation function*, i.e. a function from I^n to I, where I is some nontrivial real interval, possibly unbounded, containing 0 (Grabisch et al., 2009). Thus, the individual value functions of the dimensions are considered given and commensurate, and the analysis is concentrated on the aggregation part of the integral. Many of these results contain interesting conditions, which are quite different from those used in DUU (see Appendix A.2). Apart from the obvious translation of the "comonotonic monotonicity" results, a few new and interesting conditions were introduced, such as "horizontal min-additivity". From our perspective, it is notable that many conditions cease to depend on comonotonicity, but still depend strongly on absolute commensurateness of the attributes. This does not allow us to use them directly in conjoint-type axiomatization with heterogeneous sets. One notable idea is "commensurability through interaction" introduced in Labreuche (2012). We will return to it when dealing with the uniqueness part of our characterization, see Section 3.9.

1.4.3 Characterizations of the Sugeno integral

The Sugeno integral was introduced by Sugeno (1974). It could be thought of as an "ordinal" counterpart of the Choquet integral. If we replace summation by maximum and multiplication my minimum in the expression of the Choquet integral, then what we get is exactly the Sugeno integral.

Several characterizations have been given in the literature, which can be divided into three main groups: DUU-type axiomatizations on Y^n, axiomatization of the integral as an "aggregation function", and finally proper conjoint axiomatizations on $X_1 \times \cdots \times X_n$.

Sugeno integral characterizations were the first ones where a step away from comonotonicity and homogeneity was made for characterization of rank-dependent models. The contrast between the conditions relying on comonotonicity, and those that do not is startling (see Appendix, Section A.3). Although neither of the latter axioms can be directly used to characterize the Choquet integral, we see some of their spirit in our own results. One other notable fact is a significant weakening of the uniqueness properties of the representation compared to the homogeneous case. This question was analysed in detail in Bouyssou et al. (2009).

1.4.4 Review summary

Despite the ostensible diversity of various conditions used to characterize the Choquet integral and related models in the literature (see the Appendix for many examples), the majority of them are based on a few fundamental notions:

1. Constant acts – subset of a homogeneous decision space containing points with the same level on all dimensions;
2. Comonotonic sets – sets of points which have the coordinate levels ordered in the same way;
3. Comonotonic independence or its modifications, such as comonotonic trade-off consistency.

None of these can be used in the more general heterogeneous case $X = X_1 \times \cdots \times X_n$. Therefore, we require a new set of axioms which does not rely on comonotonicity or constant acts. In MCDA decision problems both of these notions would require assuming commensurateness between various dimensions, such as colour and shape, or loudness and size. Such a correspondence cannot be assumed a priori, nor is there a simple way to construct it using the preference data. However, as the characterizations of the Sugeno integral show, a solution can be found.

The fundamental question seems to be how to perform sensible comparisons between elements of the different dimension sets. These could be levels of various attributes or outcomes in various states. Classic additive models do not allow us to give a satisfactory answer to this question, since it is impossible to infer any correspondence between such levels purely from the preference data. However, as we will see, in rank-dependent models the situation is dramatically different.

In the following chapters we construct the axioms for a fully general conjoint characterization of the Choquet integral. Our results to some extent amalgamate the conditions used in the DUU context and those used to axiomatize the Sugeno integral. We build upon the following ideas:

1. Tradeoff consistency conditions, including their versions from (Krantz et al., 1971), and the methods of constructing additive representations on irregularly shaped subsets, due to Wakker (1991b).

2. Commensurability through interaction idea due to Labreuche (2012), in our uniqueness analysis.

3. The general "shape" of axioms used in the axiomatizations of the Sugeno integral, especially those by Greco et al. (2004); Bouyssou et al. (2009).

1.5 Summary

In this thesis we present an axiomatization of the Choquet integral decision model. In the introductory chapter we have analysed the limitations of the existing characterizations and highlighted some interesting ideas found in the literature which led to the development of our results. Our primary area of interest is multicriteria decision analysis (MCDA), but the results obtained here also allow for a very nice interpretation in the context of state-dependent utilities.

An informal presentation of our motivation and results was given in Section 1.3. In Section 1.4 we reviewed the key results in the context of decision making under uncertainty, where the Choquet integral was originally introduced, MCDA, where it enjoys a wide theoretical and practical popularity, and finally characterization results of the closely related Sugeno integral. We have outlined the main approaches and analysed the axioms which have been proposed so far. The main characteristic trait of the known results is the homogeneity of the decision space. This naturally allows a whole range of operations which are not available in the more general context, in particular comparing elements of different dimension sets (levels of various attributes, outcomes in different states). This in turn leads to notions such as “constant act” and “comonotonicity”, which are the cornerstones of all existing Choquet integral axiomatizations.

In the following chapter we provide a new characterization which does not rely on these assumptions. To our best knowledge this is the most general result achieved so far. The new axiomatic framework leads to exciting interpretations in MCDA and state-dependent contexts. We also look at the implications of our results for the learning of Choquet integral models, and finally extend our characterization to a number of special cases of the model.

Chapter 2

Axiomatization of the Choquet integral – two-dimensional case

2.1 Introduction

In this thesis we follow the same presentation path as in Krantz et al. (1971), and give the results for the two-dimensional and n-dimensional cases separately. The former case is simpler, requires fewer axioms, and has a more accessible proof. We give the details of the differences in the introduction to the following chapter (Section 3.1), but it is worth mentioning here that the n-dimensional case introduces geometric and combinatoric complexity which is not present when dealing only with two dimensions. This is yet another of many differences of our results from the homogeneous case where the two-dimensional and the n-dimensional cases are not significantly distinct.

Let $X = X_1 \times X_2$ be a (heterogeneous) product set and $\succcurlyeq$ a binary relation defined on this set. In MCDA, elements of the set X are interpreted as alternatives characterized by two criteria taking values from sets X_1 and X_2. In decision making under uncertainty, the factors of the set X usually correspond to outcomes in various states of the world, and an additional assumption $X_1 = X_2 = Y$ is being made. Thus in CEU, the set X is homogeneous, i.e. $X = Y^n$.

As discussed in the previous chapter, earlier axiomatizations of the Choquet integral have been given for this special case of $X = Y^n$ and its particular case $X = \mathbb{R}^n$. The crucial difference between our result and previous axiomatizations is that the notions of "comonotonicity" and "constant act" are no longer available in the heterogeneous case. Recall that two acts are called comonotonic in CEU if their outcomes have the same ordering. A constant act is plainly an act having the same outcome in every state of the world. Apparently, since criteria sets X_1 and X_2 in our model can be completely disjoint, neither of the notions can be used any more due to the fact that there does not exist a meaningful built-in order between elements of sets X_1 and X_2. New axioms and proof techniques must be introduced to deal with this complication.

This chapter is organized as follows. Section 2.1.1 contains the definition of the Choquet integral and looks at its properties. Section 2.2 contains the axioms and their discussion. Section 2.3 gives the representation theorem. Section 4.4 discusses the main result and its economic interpretations. The proof of the theorem is presented in Sections 3.12–2.12.

2.1.1 The Choquet integral

Let $N = \{1, 2, \dots, n\}$ be a set (of criteria) and 2^N its power set.

Definition 5. Capacity (non-additive measure, fuzzy measure) *is a set function* $\nu : 2^N \to \mathbb{R}_+$ *such that:*

1. $\nu(\varnothing) = 0$;
2. $A \subseteq B \Rightarrow \nu(A) \leq \nu(B),\ \forall A, B \in 2^N$.

In this paper, it is also assumed that capacities are normalized, i.e. $\nu(N) = 1$.

Definition 6. *The* Choquet integral *of a function* $f : N \to \mathbb{R}$ *with respect to a capacity* ν *is defined as*

$$C(\nu, f) = \int_0^\infty \nu(\{i \in N \colon f(i) \geq r\})dr + \int_{-\infty}^0 [\nu(\{i \in N \colon f(i) \geq r\}) - 1]dr \tag{2.1}$$

If N is finite, and all values $(f_1, \dots, f_n)$ of $f : N \to \mathbb{R}$ are non-negative, the definition can be expressed as:

$$C(\nu, (f_1, \dots, f_n)) = \sum_{i=1}^n (f_{(i)} - f_{(i-1)})\nu(\{j \in N \colon f_j \geq f_{(i)}\}) \tag{2.2}$$

where $f_{(1)}, \dots, f_{(n)}$ is a permutation of $f_1, \dots, f_n$ such that $f_{(1)} \leq f_{(2)} \leq \dots \leq f_{(n)}$, and $f_{(0)} = 0$.

On of the most useful tools for analysis of the capacity is the so-called Möbius transform. It's a linear transformation of the capacity which is given by:

$$m(A) = \sum_{B \subset A} (-1)^{|A \setminus B|} \nu(B). \tag{2.3}$$

The Choquet integral can be written in a very convenient form using the Möbius transform coefficients:

$$C(\nu, f) = \sum_{A \in N} m(A) \min_{i \in A}(f_i). \tag{2.4}$$

2.1.2 The model

Let $\succcurlyeq$ be a binary relation on the set $X = X_1 \times X_2$. $\succ, \prec, \preccurlyeq, \sim, \not\sim$ are defined in the usual way. We say that $\succcurlyeq$ can be represented by a Choquet integral, if there exists a capacity ν and functions $f_1 : X_1 \to \mathbb{R}$ and $f_2 : X_2 \to \mathbb{R}$, called value functions, such that:

$$x \succcurlyeq y \iff C(\nu, (f_1(x_1), f_2(x_2)) \geq C(\nu, (f_1(y_1), f_2(y_2)). \tag{2.5}$$

As seen in the definition of the Choquet integral, its calculation involves comparison of the f_i's to each other. It is not immediately obvious how this operation can have any meaning in the MAUT context. It is well-known that comparing *levels* of value functions for various attributes is meaningless in the additive model (Krantz et al., 1971) (recall that the origin of each value function can be changed independently). In the homogeneous case $X = Y^n$ this problem is readily solved, as we have a single set of outcomes Y (in the context of decision making under uncertainty). The required order is either assumed as given (Wakker, 1991b) or is readily derived from the ordering of constant acts $(\alpha, \dots, \alpha)$ (Wakker, 1991a). Since there is a single outcome set, we also have a single value (utility) function $U : Y \to \mathbb{R}$, and thus comparing $U(y_1)$ to $U(y_2)$ is perfectly sensible, since U represents the order on the set Y. None of these methods can be readily applied in the heterogeneous case.

2.1.3 Properties of the Choquet integral

Given below are some important properties of the Choquet integral:

remove this section?

1. Functions $f : N \to \mathbb{R}$ and $g : N \to \mathbb{R}$ are comonotonic if for no $i, j \in N$ we have $f(i) > f(j)$ and $g(i) < g(j)$. For all comonotonic f the Choquet integral reduces to the Lebesgue integral with respect to the same additive measure. In the finite case, the integral is accordingly reduced to a weighted sum with the same weights for all comonotonic f.

2. Particular cases of the Choquet integral (e.g. (Grabisch and Labreuche, 2008)).

- If $\nu(\{1\}) = \nu(\{2\}) = 1$, then $C(\nu, (f_1, f_2)) = \max(f_1, f_2)$.
- If $\nu(\{1\}) = \nu(\{2\}) = 0$, then $C(\nu, (f_1, f_2)) = \min(f_1, f_2)$.
- If $\nu(\{1\}) + \nu(\{2\}) = 1$, then $C(\nu, (f_1, f_2)) = \nu(\{1\})f_1 + \nu(\{2\})f_2$

Property 1 states that the set X can be separated into subsets corresponding to particular orderings of the value functions. In the case of two criteria there are only two such sets: $\{x \in X : f_1(x_1) \geq f_2(x_2)\}$ and $\{x \in X : f_2(x_2) \geq f_1(x_1)\}$. Since the integral on each of the sets is reduced to a weighted sum, i.e. an additive representation, we should expect many of the axioms of the additive conjoint model to be valid on this subsets.

2.2 Axioms

Definition 7. *A relation $\succcurlyeq$ on $X_1 \times X_2$ satisfies* triple cancellation*, provided that for every $a, b, c, d \in X_1$ and $p, q, r, s \in X_2$, we have*

$$\left.\begin{aligned} ap &\preccurlyeq bq \\ ar &\succcurlyeq bs \\ cp &\succcurlyeq dq \end{aligned}\right\} \Rightarrow cr \succcurlyeq ds. \tag{2.6}$$

Definition 8. *A relation $\succcurlyeq$ on $X_1 \times X_2$ is* independent*, iff for $a, b \in X_1$, $ap \succcurlyeq bp$ for some $p \in X_2$ implies that $aq \succcurlyeq bq$ for every $q \in X_2$; and, for $p, q \in X_2$, $ap \succcurlyeq aq$ for some $a \in X_1$ implies that $bp \succcurlyeq bq$ for every $b \in X_1$.*

A1. Weak order $\succcurlyeq$ is a weak order.

A2. Weakest separability For any $a_i p_j, b_i p_j \in X$ such that $a_i p_j \succ b_i p_j$, we have $a_i q_j \succcurlyeq b_i q_j$ for all $q_j \in X_j$, for $i, j \in \{1, 2\}$.

The separability condition is weaker than the one normally used. The condition first appeared in (Bliss, 1975), and in this form in (Mak, 1984), the name "weak separability" is used in Bouyssou et al. (2006), however to avoid confusion we call it "weakest separability" here. The condition only rules out a reversal of strict preference. Note, that it implies that for any $a, b \in X_1$ either $ap \succcurlyeq bp$ or $bp \succcurlyeq ap$

for all $p \in X_2$ (symmetrically for the second coordinate). Apparently, transitivity also holds: if $ap \succcurlyeq bp$ for all $p \in X_2$ and $bp \succcurlyeq cp$ for all $p \in X_2$, then $ap \succcurlyeq cp$ for all $p \in X_2$. This allows to introduce the following weak orders:

Definition 9. *For all $a, b \in X_1$ define $\succcurlyeq_1$ as $a \succcurlyeq_1 b \iff ap \succcurlyeq bp$ for all $p \in X_2$. Define $\succcurlyeq_2$ in a analogous way.*

Definition 10. *We call $a \in X_1$* minimal *if $b \succcurlyeq_1 a$ for all $b \in X_1$, and* maximal *if $a \succcurlyeq_1 b$ for all $b \in X_1$. Symmetric definitions hold for X_2.*

Definition 11. *For any $z \in X$ define $\mathbf{SE}^z = \{x : x \in X, x_1 \succcurlyeq_1 z_1$ and $z_2 \succcurlyeq_2 x_2\}$, and $\mathbf{NW}^z = \{x : x \in X, x_2 \succcurlyeq_2 z_2$ and $z_1 \succcurlyeq_1 x_1\}$.*

The "rectangular" cones $\mathbf{SE}^z$ and $\mathbf{NW}^z$ play a significant role in the sequel.

A3. Cone additivity For any $z \in X$, triple cancellation holds either on $\mathbf{SE}^z$ or on $\mathbf{NW}^z$.

The axiom says that the set X can be covered by "rectangular" cones, such that triple cancellation holds within each cone. We will call such cones "3C-cones". The axiom effectively divides X into subsets, defined as follows.

Definition 12. *We say that*

- *$z \in \mathbf{SE}$ if at least one of the following conditions is true:*
  - *Triple cancellation holds on $\mathbf{SE}^z$, z_1 is not maximal and z_2 is not minimal*
  - *z_1 is maximal and for no $x_2, y_2 \in X_2 : z_2 \succcurlyeq_2 x_2 \succcurlyeq_2 y_2$ triple cancellation holds on $\mathbf{NW}^{z_1 x_2}$ but does not hold on $\mathbf{NW}^{z_1 y_2}$*
  - *z_2 is minimal and for no $x_1, y_1 \in X_1 : y_1 \succcurlyeq_1 x_1 \succcurlyeq_1 z_1$ triple cancellation holds on $\mathbf{NW}^{x_1 z_2}$ but does not hold on $\mathbf{NW}^{y_1 z_2}$*
- *$z \in \mathbf{NW}$ if at least one of the following conditions is true:*
  - *Triple cancellation holds on $\mathbf{NW}^z$, z_2 is not maximal and z_1 is not minimal*

– *z_2 is maximal and for no $x_1, y_1 \in X_1 : z_1 \succcurlyeq_1 x_1 \succcurlyeq_1 y_1$ triple cancellation holds on $\mathbf{SE}^{x_1 z_2}$ but does not hold on $\mathbf{SE}^{y_1 z_2}$*

– *z_1 is minimal and for no $x_2, y_2 \in X_2 : y_2 \succcurlyeq_2 x_2 \succcurlyeq_2 z_2$ triple cancellation holds on $\mathbf{SE}^{z_1 x_2}$ but does not hold on $\mathbf{SE}^{z_1 y_2}$*

Define also $\Theta = \mathbf{NW} \cap \mathbf{SE}$.

Presence of maximal and minimal points significantly complicates the definitions of $\mathbf{SE}$ and $\mathbf{NW}$, since at such points some of the sets $\mathbf{SE}^z$ and $\mathbf{NW}^z$ become degenerate and triple cancellation trivially holds. If sets X_1 and X_2 do not contain minimal or maximal points, we can drop the corresponding conditions in each definition and simply state that $\mathbf{SE}$ is the set of all $z \in X$ such that triple cancellation holds on $\mathbf{SE}^z$, whereas $\mathbf{NW}$ is the set of all $z \in X$ such that triple cancellation holds on $\mathbf{NW}^z$

Definition 13. *We say that $i \in N$ is* essential *on $A \subset X$ if there exist $x_i x_j, y_i x_j \in A$, $i, j \in N$, such that $x_i x_j \succ y_i x_j$.*

A4. Intra-coordinate tradeoff consistency For all $a, b, c, d \in X_1$ and $p, q, r, s \in X_2$, we have

$$\left.\begin{array}{r} ap \preccurlyeq bq \\ ar \succcurlyeq bs \\ cp \succcurlyeq dq \end{array}\right\} \Rightarrow cr \succcurlyeq ds, \tag{2.7}$$

provided that either:

a) $ap, bq, ar, bs, cp, dq, cr, ds \in \mathbf{NW}(\mathbf{SE})$, or;

b) $ap, bq, ar, bs \in \mathbf{NW}$ and $i = 2$ is essential on $\mathbf{NW}$ and $cp, dq, cr, ds \in \mathbf{SE}$ or vice versa, or;

c) $ap, bq, cp, dq \in \mathbf{NW}$ and $i = 1$ is essential on $\mathbf{NW}$ and $cp, dq, cr, ds \in \mathbf{SE}$ or vice versa.

Informally, the meaning of the axiom is that ordering between preference differences ("intervals") is preserved irrespective of the "measuring rods" used to measure them. However, contrary to the additive case this does not hold on all X, but only when either points involved in all four relations lie in a single 3C-cone, or points involved in two relations lie in one 3C-cone and those involved in the other two in another.

A5. Inter-coordinate tradeoff consistency For all $i \in N$ we have

$$\left.\begin{array}{r} a_i x_{-i} \preccurlyeq b_i y_{-i} \\ c_i x_{-i} \succcurlyeq d_i y_{-i} \\ a_i y^0_{-i} \sim p_j x^0_{-j} \\ b_i y^0_{-i} \sim q_j x^0_{-j} \\ c_i y^1_{-i} \sim r_j x^1_{-j} \\ d_i y^1_{-i} \sim s_j x^1_{-j} \\ p_j e_{-j} \succcurlyeq q_j f_{-j} \end{array}\right\} \Rightarrow r_j e_{-j} \succcurlyeq s_j f_{-j} \tag{2.8}$$

for all $a_i x_{-i}, b_i y_{-i}, c_i x_{-i}, d_i y_{-i} \in$ **SE**(**NW**) provided i is essential on **SE**(**NW**), $a_i y^0_{-i}, b_i y^0_{-i}, c_i y^1_{-i}, d_i y^1_{-i} \in$ **SE**(**NW**), $p_j x^0_{-j}, q_j x^0_{-j}, r_j x^1_{-j}, s_j x^1_{-j} \in$ **SE**(**NW**) provided j is essential on **SE**(**NW**), $p_j e_{-j}, q_j f_{-j}, r_j e_{-j}, s_j f_{-j} \in$ **SE**(**NW**).

The formal statement of **A5** is rather complicated, but it simply means that the ordering of the intervals is preserved across dimensions. Together with **A4** the conditions are similar to Wakker's tradeoff consistency condition (Wakker, 1991b). The axiom bears even stronger similarity to Axiom 5 (compatibility) from section 8.2.6 of (Krantz et al., 1971). Roughly speaking, it says that if the interval between c_i and d_i is larger than that between a_i and b_i, then projecting these intervals onto another dimension by means of the equivalence relations must leave this order unchanged. We additionally require the comparison of intervals and projection to be consistent – meaning that quadruples of points in each part of the statement lie in the same 3C-cone. Another version of this axiom, which is going to be used frequently in the proofs, is formulated in terms of standard sequences in Lemma 7.

A6. Strong monotonicity. Let $ap, bp, cp, dp \in$ **SE**(**NW**) and $ap \succ bp$. If for some $q \in X_2$ also exist $cq \succ dq$, then $cp \succ dp$. Symmetric condition holds for the second coordinate.

This axiom is similar to "strong monotonicity" in (Wakker, 1991b). We analyze its necessity and the intuition behind it in section 3.13.

A7. Essentiality Both coordinates are essential on X.

A8. Restricted solvability If $x_i a_j \succcurlyeq y \succcurlyeq x_i c_j$, then there exists $b : x_i b_j \sim y$, for $i, j \in \{1, 2\}$.

A9. Archimedean axiom For every $z \in \mathbf{NW}(\mathbf{SE})$ every bounded standard sequence contained in $\mathbf{NW}^z(\mathbf{SE}^z)$ is finite, and if both subsets have only one essential coordinate each, there exists a countable order-dense subset of $X^{\mathbf{S}_i}$.

We also make some non-necessary structural assumptions to simplify the proofs. It seems plausible that all of these can be removed, e.g. by using "cuts" and limits instead of points and values, however, this would inevitably make the proof more technically challenging.

"Collapsed" equivalent points along dimensions. For no $a, b \in X_1$ we have $ap \sim bp$ for all $p \in X_2$. Similarly, for no $p, q \in X_2$ we have $ap \sim aq$ for all $a \in X_1$.

If such points exist, say $ap \sim bp$ for all $p \in X_2$, then we can build the representation for a set $X_1' \times X_2$ where $X_1' = X_1 \setminus a$, and later extend it to X by setting $f_1(a) = f_1(b)$.

Density. Whenever $x \succ y$ there exists z such that $x \succ z \succ y$. From this and restricted solvability immediately follows that $\succcurlyeq_i$ is order dense as well, in other words, whenever $a_i p_j \succ b_i p_j$ there exists $c \in X_i$ such that $a_i p_j \succ c_i p_j \succ b_i p_j$, for $i, j \in N$.

"Closedness" of SE and NW. Whenever exist $ap \notin \mathbf{NW}$ and $bp \notin \mathbf{SE}$, there exist also $cp \in \Theta$. Similarly, whenever exist $ap \notin \mathbf{NW}$ and $aq \notin \mathbf{SE}$, there exist also $ar \in \Theta$.

This assumption says that sets **SE** and **NW** are "closed". In the representation this translates into existence of the inverse for all points where value functions f_1 and f_2 are equal, provided there is a violation of independence.

2.2.1 Discussion of axioms

Roughly speaking, for two dimensional sets the Choquet integral can be characterized by saying that X is divided into two subsets such that $\succcurlyeq$ on each of them has an additive representation, while the intersection of these subsets (in the representation) is the line $\{x : f_1(x_1) = f_2(x_2)\}$. In the previous characterizations locating these subsets was straightforward, as they are nothing else but the comonotonic subsets of X. In this paper we take a different approach. Instead, we state that X can be separated into two subsets without imposing any additional constraints on their location and then use additional axioms to characterize the intersection of these subsets and to show that it is mapped to the line $\{x : f_1(x_1) = f_2(x_2)\}$.

Previous characterizations were usually based on a single condition, such as comonotonic independence Schmeidler (1989) or comonotonic tradeoff consistency Wakker (1991a). It seems quite difficult to translate any of these to our setting, because, as was discussed previously, the general heterogeneous case has several important differences from the homogeneous one (e.g. absence of constant acts and comonotonic subsets). However, it turns out that each of these conditions can be decomposed into several more simple properties which we can try to generalize. In fact, our axiom **A3** can be viewed as a generalization of comonotonicity, whereas **A4** and **A5** constitute a more general form of the trade-off consistency condition.

A3 basically states that the set X can be divided into subsets, such that within every such subset the preference relation can be represented by an additive function. At every point $z \in X$ it is possible to build two "rectangular cones": $\{x : x_1 \succcurlyeq_1 z_1 \text{ and } z_2 \succcurlyeq_2 x_2\}$, and $\{x : x_2 \succcurlyeq_2 z_2 \text{ and } z_1 \succcurlyeq_1 x_1\}$. The axiom states that triple cancellation must then hold on at least one of these cones. An example of such cones is given in Figure 2.1, dashed line shows the boundary between two additive subsets of X. However, contrary to the homogeneous case, this border can be of a more general form. The axiom allows it to be a set of non-null measure, and it could even be empty (however we rule that out with one of the structural assumptions). If the set X was a metric one, we could also say that contrary to the homogeneous case, the border does not have to be a straight line, informally it could have any “quasiconcave” shape.

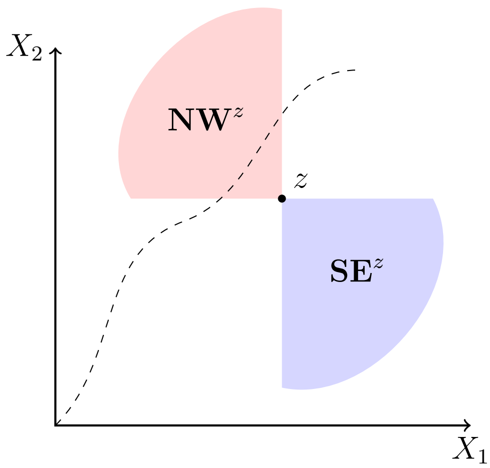


Figure 2.1: Axiom **A3** - weakening of comonotonicity

Axioms **A4** and **A5** are very closely related to Wakker's version, however they are based not on comonotonicity but on **A3**. Also, Wakker's version require homogeneity, as it makes use of the same outcomes but on different dimensions (states). The main idea of the axioms is to say that the additive representations on **SE** and **NW** are interrelated, in particular trade-offs are consistent across subsets both within the same dimension and for different ones. **A4**, similar to the condition used in (Wakker, 1991a) states that triple cancellation holds across cones, while **A5**, similar to (Krantz et al., 1971) (section 8.2), says that the ordering of intervals on any dimension must be preserved when they are projected onto another dimension by means of equivalence relations - see examples of both conditions in Figures 2.2 and 2.3.

These axioms are complemented by a new condition called here monotonicity (**A6**), which basically states, that if a variable is essential “somewhere” within **SE** or **NW**, then it is essential “everywhere” on the same subset.

Weakest separability (**A2**) allows to order elements of the dimension sets,

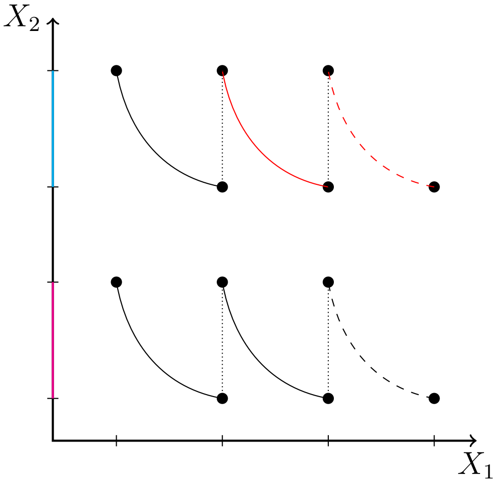


Figure 2.2: Intra-coordinate tradeoff consistency

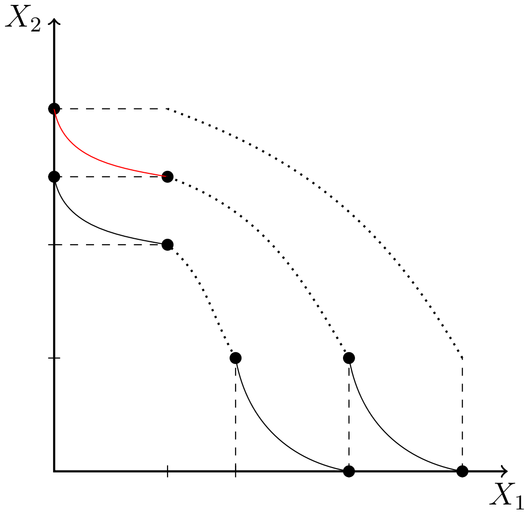


Figure 2.3: Inter-coordinate tradeoff consistency

as was discussed already. Finally, the standard essentiality,"comonotonic" Archimedean axiom and restricted solvability (**A7,A8,A9**) complete the list. $\succcurlyeq$ is supposed to be a weak order (**A1**) and X is dense.

The most important axioms - **A3,A4,A5,A6**, are not only sufficient, but also necessary, the details are given in Section 2.12. Necessity of some of the remaining axioms is well-known Wakker (1991b); Bouyssou and Pirlot (2004).

2.3 Representation theorem

Theorem 1. *Let $\succcurlyeq$ be a binary relation on X and let the structural assumptions hold. Then, if axioms **A1-A9** are satisfied, there exists a capacity ν and value functions $f_1 : X_1 \to \mathbb{R}$, $f_2 : X_2 \to \mathbb{R}$, such that $\succcurlyeq$ can be represented by the Choquet integral:*

$$x \succcurlyeq y \iff C(\nu, (f_1(x_1), f_2(x_2))) \geq C(\nu, (f_1(y_1), f_2(y_2))), \tag{2.9}$$

for all $x, y \in X$.

Moreover, the representation has the following uniqueness properties:

Theorem 2. *Assume $\succcurlyeq$ on X has a representation as in equation* (2.9) *and triple cancellation does not hold throughout X. ν is unique and for any functions $g_1 : X_1 \to \mathbb{R}, g_2 : X_2 \to \mathbb{R}$ such that* (2.9) *holds with f_i substituted by g_i, we have the following relations between f_i and g_i.*

- *Both coordinates are essential on* **SE** *and* **NW**.

$$f_i(x_i) = \alpha g_i(x_i) + \beta, \tag{2.10}$$

  where $\alpha > 0$. Moreover, $\nu(\{1\}) \in (0, 1)$ and $\nu(\{2\}) \in (0, 1)$ and $\nu(\{1\}) + \nu(\{2\}) \neq 1$.

- *One of coordinates is essential on* **SE** *or* **NW**, *and two are essential on the other subset.*

$$f_i(x_i) = \alpha g_i(x_i) + \beta, \tag{2.11}$$

  where $\alpha > 0$, for all x_i such that for some $a_j \in X_j$ the point $x_i a_j$ belongs to the subset with two essential variables, and

$$f_i(x_i) = \psi_i(g_i(x_i)), \tag{2.12}$$

  where ψ_i are increasing functions such that $f_1(x_1) = f_2(x_2) \iff g_1(x_1) = g_2(x_2)$, for all other x_i. Moreover, either $\nu(\{2\}) \in \{0, 1\}$ and $\nu(\{1\}) \in (0, 1)$, or $\nu(\{2\}) \in (0, 1)$ and $\nu(\{1\}) \in \{0, 1\}$.

- *Both* **SE** *and* **NW** *have a single essential variable.*

$$f_i(x_i) = \psi_i(g_i(x_i)), \tag{2.13}$$

  where ψ_i are increasing functions such that $f_1(x_1) = f_2(x_2) \iff g_1(x_1) = g_2(x_2)$. Moreover, $\nu(\{1\}) = \nu(\{2\}) = 0$ or $\nu(\{1\}) = \nu(\{2\}) = 1$.

If triple cancellation holds throughout X, then $\succcurlyeq$ can be represented by an additive model

$$x \succcurlyeq y \iff \phi_1(x_1) + \phi_2(x_2) \geq \phi_1(y_1) + \phi_2(y_2). \tag{2.14}$$

If exist some other functions ϕ'_1 and ϕ'_2 which can be used in place of ϕ_1 and ϕ_2 in this representation, they are related as follows:

$$\phi'_i(x_i) = \alpha\phi_i(x_i) + \beta_i. \tag{2.15}$$

However, the additive expression $\phi_1(x_1) + \phi_2(x_2)$ can also be rewritten in the form (2.9). In particular, each value function can be factorized as $\phi_i(x_i) = \alpha_i f_i(x_i)$, with the condition $\alpha_1 + \alpha_2 = 1$. This "weighted average" model

$$x \succcurlyeq y \iff \alpha f_1(x_1) + (1-\alpha) f_2(x_2) \geq \alpha f_1(y_1) + (1-\alpha) f_2(y_2) \tag{2.16}$$

is a Choquet integral with respect to an arbitrary additive capacity, provided $\nu(\{1\}) \neq 0$ and $\nu(\{2\}) \neq 0$. The weights α and $1-\alpha$ can be further factorized, so that a non-additive capacity can also be used. For this the locations of functions f_i (i.e. the constants β_i) must be chosen in a way that allows the rewriting the integral in the Mobius transform form with a non-additive capacity gives an equivalent additive form. For example, if we set the location of f_1 and f_2 so that $f_1(x_1) \leq f_2(x_2)$ for all $x \in X$, we can rewrite the weighted average form as follows:

$$\alpha f_1(x_1) + (1-\alpha) f_2(x_2) = (\alpha - m_{12}) f_1(x_1) + (1-\alpha) f_2(x_2) + m_{12}(f_1(x_1) \wedge f_2(x_2)), \tag{2.17}$$

which is a Möbius transformation form of the Choquet integral. In this case, the capacity ν does not have to be additive anymore.

2.3.1 Absence of independence implies commensurateness

As the uniqueness part of Theorem 1 states, unless $\succcurlyeq$ can be represented by an additive functional on all of X, the representation implies commensurateness of *levels* of utility functions defined on different factors of the product set. Indeed, we have that if $f_1(x_1) \geq f_2(x_2)$ in one representation, then necessarily $g_1(x_1) \geq g_2(x_2)$ in another one. This is a much stronger uniqueness result in comparison to the traditional additive models. In Section 4.4 we discuss some economic implications of this.

2.4 Proof preview

The proof contains five main stages. First, we construct additive representations on the sets **SE** and **NW** with extreme points removed. Next we show that value functions on the same dimensions are proportional, so we can rewrite the additive representations using the same value functions, but with different weights, for both sets **SE** and **NW**. The third step is to show that we can rescale these value functions so that for all points from Θ, i.e. for points which belong to the border between **SE** and **NW**, levels from both dimensions get the same value. Finally, we show that two additive representations can be unified in a single global Choquet integral representation, including also the extreme points of X. The only thing left after this is to analyse the uniqueness of the constructed model.

Construction of the additive representations within **NW** and **SE** is done in Section 2.5, proportionality is shown in 2.6, equality of value functions for points from Θ is proved in Section 2.7, extension to the extreme points is done in 2.8, whereas construction of the Choquet integral is performed in 2.9 and 2.10. Finally, Section 2.11 contains the analysis of the uniqueness properties.

2.5 Building additive value functions on NW and SE

In this section we assume that $\mathbf{SE}(\mathbf{NW})$ has two essential coordinates.

2.5.1 Covering SE and NW with maximal $\mathbf{SE}^z$ and $\mathbf{NW}^z$

In the sequel we could have covered areas **SE** and **NW** by sets $\mathbf{SE}^z(\mathbf{NW}^z)$ for all $z \in \mathbf{SE}(\mathbf{NW})$, but it is convenient to introduce the following lemma.

Lemma 1. *For every $x \in \mathbf{SE}$ there exists $z \in \Theta$ such that $\mathbf{SE}^x \subset \mathbf{SE}^z$. Accordingly, for every $y \in \mathbf{NW}$ there exists $z \in \Theta$ such that $\mathbf{NW}^y \subset \mathbf{NW}^z$.*

Proof. Take $x \in \mathbf{SE}$ such that $x \notin \mathbf{NW}$. If there exists $y \in \mathbf{NW}^x$ such that $y \notin \mathbf{SE}$, then consider points x_1y_2 and y_1x_2. By **A3** they are either in **SE** or **NW** and by hence closedness assumption there must exist either $ay_2 \in \Theta$ (and thus $x \in \mathbf{SE}_{ay_2}$) or $x_1 p \in \Theta$ (and $x \in \mathbf{SE}^{x_1 p}$). If such y does not exist, $X = \Theta$. Other cases are symmetrical. $\square$

It follows from Lemma 1 that $\mathbf{SE} = \bigcup_{z \in \Theta} \mathbf{SE}^z$, while $\mathbf{NW} = \bigcup_{z \in \Theta} \mathbf{NW}^z$. Comparing this to definitions of **SE** and **NW** we are able to define also the following notions:

Definition 14. *We write $x \in \mathbf{SE}_{ext}$ and say that $x \in X$ is* extreme in **SE** *if $x \in \Theta$ and [x_2 is minimal or x_1 is maximal]. We write $x \in \mathbf{NW}_{ext}$ and say that $x \in X$ is* extreme in **NW** *if $x \in \Theta$ and [x_1 is minimal or x_2 is maximal]. $x \in X$ is extreme if it is extreme in* **SE** *or in* **NW**.

Note that contrary to the homogeneous case $X = Y^n$, extreme points for **SE** and **NW** can be asymmetric, i.e. if a point z is extreme in **SE** it is not necessarily extreme in **NW**.

2.5.2 Representations within $\mathbf{SE}^z$

In the following we will build an additive representation on **SE**. The case of **NW** is symmetric. We proceed by building representations on sets $\mathbf{SE}^z$ for all $z \in \Theta \setminus \mathbf{SE}_{ext}$ (i.e. for all non-extreme points of Θ).

Essential coordinates. For now we assume that both coordinates are essential on **NW** and **SE**.

Theorem 3. *For any $z \in \Theta \setminus \mathbf{SE}_{ext}$ there exists an additive representation of $\succcurlyeq$ on $\mathbf{SE}^z$:*

$$x \succcurlyeq y \Leftrightarrow V_1^z(x_1) + V_2^z(x_2) \geq V_1^z(y_1) + V_2^z(y_2). \tag{2.18}$$

Proof. $\mathbf{SE}^z$ is a Cartesian product, $\succcurlyeq$ is a weak order on $\mathbf{SE}^z$, $\succcurlyeq$ satisfies triple cancellation on $\mathbf{SE}^z$, $\succcurlyeq$ satisfies Archimedean axiom on $\mathbf{SE}^z$, both coordinates are essential. It remains to show that $\succcurlyeq$ satisfies restricted solvability on $\mathbf{SE}^z$.

Assume that for some $xa, y, xc \in \mathbf{SE}^z$, we have $xa \succcurlyeq y \succcurlyeq xc$, hence exists $b \in X_2 : xb \sim y$. We need to show that $xb \in \mathbf{SE}^z$. If $xb \sim xa$ or $xb \sim xc$, then the result is immediate. Hence, assume $xa \succ xb \succ xc$. By definition, $x \in \mathbf{SE}^z$, if

$x \succcurlyeq_1 z_1$, and $z_2 \succcurlyeq_2 b$. For xb we need to check only the latter condition. It holds, since $xa \succ xb \succ xc$, and by weak separability $a \succcurlyeq_2 b$.

Therefore all conditions for the existence of an additive representation are met (Krantz et al., 1971).

□

2.5.3 Joining representations for different $\mathbf{SE}^z$ (or $\mathbf{NW}^z$)

This section closely follows (Wakker, 1991b).

Theorem 4. *There exists an additive interval scale V^{SE} on $\bigcup \mathbf{SE}^z$, with $z \in \Theta \setminus \mathbf{SE}_{ext}$, which represents $\succcurlyeq$ on every $\mathbf{SE}^z$.*

Proof. Choose the "reference" points - pick any $r \in \mathbf{SE}$ and any $r^0, r^1 \in \mathbf{SE}_r$ such that $r_1^1 p \succcurlyeq r_1^0 p$ for every $p \in X_2$. Set $V_1^r(r_1^0) = 0, V_2^r(r_2^0) = 0, V_1^r(r_1^1) = 1$. Now, we align representations on the other sets $\mathbf{SE}^z$ with the reference one. Assume that for some $z \in \Theta$ we have already obtained an additive representation V^z on $\mathbf{SE}^z$. Observe that V^z and V^r are additive value functions for $\succcurlyeq$ on $\mathbf{SE}^z \cap \mathbf{SE}_r$. Morevover $\mathbf{SE}^z \cap \mathbf{SE}_r = \mathbf{SE}_q$, where $q_1 = r_1$ if $r_1 \succcurlyeq_1 z_1$ and $q_1 = z_1$ if the opposite is true. Similarly, $q_2 = r_2$ if $az_2 \succcurlyeq ar_2$ for all $a \in X_1$ and $q_2 = z_2$ in the opposite case. Hence, uniqueness results from Krantz et al. (1971) can be applied. In particular, this means that on $\mathbf{SE}^z \cap \mathbf{SE}_r$ we have $V_i^r = \alpha V_i^z + \beta_i$, so the functions are defined up to a common unit and location.

We choose the unit and location of V_1^z so that $V_1^z(x_1) = V_1^r(x_1)$ for all $x \in \mathbf{SE}^z \cap \mathbf{SE}_r$. Next, we choose the location of V_2^z so that it coincides with V_2^r on $\mathbf{SE}^z \cap \mathbf{SE}_r$.

Finally, we show that $V_i^s(x_i) = V_i^t(x_i)$ for any $s, t \in \Theta$ and $x \in \mathbf{SE}_s \cap \mathbf{SE}_t$. This immediately follows, since V^s and V^t coincide (with V^r) on $\mathbf{SE}_s \cap \mathbf{SE}_t \cap \mathbf{SE}_r$. This defines their unit and locations, hence they also coincide on $\mathbf{SE}_s \cap \mathbf{SE}_t$. Now define V^{SE} as a function which coincides with V^{z_i} on the respective domains $\mathbf{SE}^{z_i}$. By the above argument, this function is well-defined. □

Theorem 5. *Representation V^{SE} obtained in Theorem 4 is globally representing on $\mathbf{SE} \setminus \mathbf{SE}_{ext} = \bigcup_{z \in \Theta \setminus \mathbf{SE}_{ext}} \mathbf{SE}^z$.*

Proof. Let $x \succcurlyeq y$. There can be two cases. First, assume that $x_2 \succcurlyeq_2 y_2$, but $y_1 \succcurlyeq_1 x_1$ (or vice versa). In this case, x and y belong to the same $\mathbf{SE}^z$ (e.g. $\mathbf{SE}^x$) and therefore V^{SE} is a valid representation.

Next, assume that $x_j \succcurlyeq_j y_j$ for all $i, j \in N$. Assume, that $x \in \mathbf{SE}_s, y \in \mathbf{SE}_t$. Observe that $x_1 y_2 \in \mathbf{SE}_s \cap \mathbf{SE}_t$ because by the made assumptions, $x_1 y_2 \in \mathbf{SE}^x, x_1 y_2 \in \mathbf{SE}^y$. By definition of $\succcurlyeq_i$ we have $x_1 x_2 \succcurlyeq x_1 y_2 \succcurlyeq y_1 y_2$, hence $V_1(x_1) + V_2(x_2) \geq V_1(x_1) + V_2(y_2) \geq V_1(y_1) + V_2(y_2)$, with first inequalities lying in $\mathbf{SE}_s$, and second in $\mathbf{SE}_t$. The reverse implication is also true. $\square$

2.6 Aligning V^{SE} and V^{NW}

First we will show that it is not possible for the common domain of V_i^{SE} and V_i^{NW} for some i to contain a single point.

2.6.1 Analysis of the common domain of V^{SE} and V^{NW}

Lemma 2. *Let $a_0 \succcurlyeq_1 b_0$, and for some $p \in X_2$ we have $a_0 p, b_0 p \in \Theta$. Define $X_{a_0 b_0} = \{x_1 : x_1 \in X_1, b_0 \succcurlyeq_1 x_1 \succcurlyeq_1 a_0\}$. Then, triple cancellation holds everywhere on $X_{a_0 b_0} \times X_2$.*

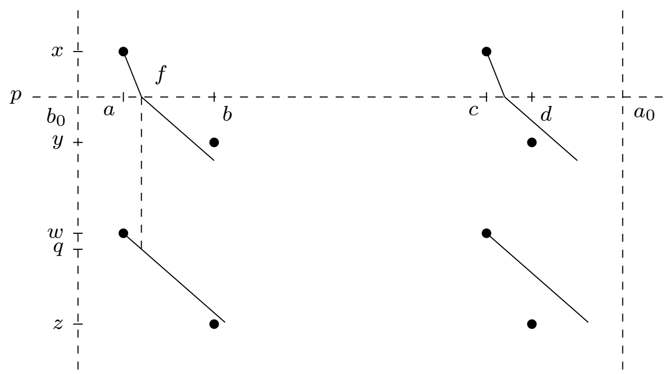


Figure 2.4: Lemma 2

Proof. All points in the below proof are from $X_{a_0b_0} \times X_2$. Let $ax \preccurlyeq by, aw \succcurlyeq bz, cx \succcurlyeq dy$. We will show that together with the assumptions of the Lemma, this implies $cw \succcurlyeq dz$.

The case when all points belong to **SE** or **NW**, or two pairs belong to **SE** and two to **NW** is covered by **A4**. Thus, assume wlog $x \succcurlyeq_2 p$, so that $ax, cx \in$ **NW** and the remaining points are in **SE** (Fig. 2.4). Assume also $dp \succcurlyeq cp$ and $b \succcurlyeq_1 a$. Assume also $ax \succ ap$ (hence by independence also $cx \succ cp$), $bp \succ by$ (hence also $dp \succ dy$), otherwise the result immediately follows by **A4** (e.g. if $ax \sim ap$, we can replace ax by ap and cx by cp in the assumptions of the lemma, which brings all points to **SE**).

1. $ax \preccurlyeq by$.
   $bp \succ by$, hence $ax \prec bp$. $ax \succ ap$, hence $bp \succ by \succcurlyeq ax \succ ap$, $bp \succ ap$, therefore, by restricted solvability exists $fp \sim ax$. Also, $fp \succ ap, bp \succ fp$.

2. $cx \succcurlyeq dy$. There can be two cases:

   a) If $cx \preccurlyeq dp$, then $dp \succcurlyeq cx \succ cp$, hence exists $gp \sim cx$.

   b) $cx \succ dp$.

3. $aw \succcurlyeq bz$.
   Solve for q: $aw \sim fq$. By the results in point 1 and independence we have $fw \succ aw \succcurlyeq bz \succ fz$, therefore by restricted solvability exists $q : fq \sim aw$.

4. Cases correspond to those in point 2 above:

   a) $fp \sim ax, gp \sim cx, aw \sim fq$, hence by **A4** $cw \sim gq$
   $fp \preccurlyeq by, gp \succcurlyeq dy, fq \succcurlyeq bz$, hence by **A4** $gq \succcurlyeq dz$ and $cw \succcurlyeq dz$.

   b) $ax \prec bp, cx \succ dp, aw \succcurlyeq bz$, hence by **A4** $cw \succcurlyeq dz$.

□

From this it follows that it is impossible that for some i the common domain of V_i^{SE} and V_i^{NW} includes a single point. Let (wlog) $i = 1$ and $a \in X_1$ be such a point. Apparently $ap \in \Theta$ for all $a \in X_1$. Then, from Lemma 2 it follows that **SE** = **NW** = X.

2.6.2 Aligning representations on SE and NW

There can be four cases, depending on the number of essential coordinates on **NW** and **SE**:

1. Both areas have two essential coordinates;
2. One area has two essential coordinate, another has one essential coordinate;
3. Both areas have one essential coordinates;
4. An area does not have any essential coordinates.

We start with the case where both coordinates are essential on **NW** and **SE**.

2.6.2.1 Both coordinates are essential

Lemma 3. *Choose $r_1^0 \in X_1$ and $r_1^1 \in X_1$ from the common domain of V_1^{SE} and V_1^{NW} such that $r_1^1 \succcurlyeq_1 r_1^0$, and set $V_1^{SE}(r_1^0) = V_1^{NW}(r_1^0) = 0$ and $V_1^{SE}(r_1^1) = V_1^{NW}(r_1^1) = 1$. Then, $V_1^{SE}(x_1) = V_1^{NW}(x_1)$ on all x_1 from their common domain.*

Proof. This follows directly from **A4**. Assume, we want to show that $V_1^{SE}(y_1) = V_1^{NW}(y_1)$ for some $y_1 \in X_1$. Starting from r_1^0 build any standard sequence on X_1 in **SE**, say $\{\alpha_1^{(i)} : \alpha_1^{(i)} y_1^s \sim \alpha_1^{(i+1)} y_2^s\}$. Then, all $\alpha_1^{(i)} y_1^n, \alpha_1^{(i)} y_2^n$ which are in **NW** also form a sequence in **NW**: if $\alpha_1^{(i)} y_1^s \sim \alpha_1^{(i+1)} y_2^s, \alpha_1^{(i+1)} y_1^s \sim \alpha_1^{(i+2)} y_2^s$ and $\alpha_1^{(i)} y_1^n \sim \alpha_1^{(i+1)} y_2^n$, for some $y_1^n, y_2^n \in X_2$, then by **A4**, necessarily $\alpha_1^{(i+1)} y_1^n \sim \alpha_1^{(i+2)} y_2^n$.

Now let

$$\begin{aligned} 1 = &V_1^{SE}(r_1^1) \approx n[V_2^{SE}(y_2^s) - V_2^{SE}(y_1^s)] \\ &V_1^{SE}(y_1) \approx m[V_2^{SE}(y_2^s) - V_2^{SE}(y_1^s)] \approx \frac{m}{n}. \end{aligned} \tag{2.19}$$

Such n and m exist by the Archimedean axiom. By the argument above we get

$$\begin{aligned} 1 = &V_1^{NW}(r_1^1) \approx n[V_2^{NW}(y_2^n) - V_2^{NW}(y_1^n)] \\ &V_1^{NW}(y_1) \approx m[V_2^{NW}(y_2^n) - V_2^{NW}(y_1^n)] \approx \frac{m}{n} \end{aligned} \tag{2.20}$$

By denserangedness, approximations become exact in the limit, so we obtain $V_1^{SE}(y_1) = V_1^{NW}(y_1)$ on all $y_1 \in X_1$ from their common domain. □

Lemma 4. *Assume, V^{SE} is an additive representation of $\succcurlyeq$ on $\mathbf{SE} \setminus \mathbf{SE}_{ext}$, and V^{NW} is a representation on $\mathbf{NW} \setminus \mathbf{NW}_{ext}$, with V_1^{SE} and V_1^{NW} scaled so that they have a common zero and unit (as in Lemma 3). Then, $V_2^{SE} = \lambda V_2^{NW}$ on the common domain.*

Proof. By Lemma 3, $V_1^{SE} = V_1^{NW}$ on the common domain. Assume $V_2^{SE}(r_2^2) = \lambda, V_2^{NW}(r_2^2) = 1$. We will now show that $V_2^{SE}(x_2) = \lambda V_2^{NW}(x_2)$ for all $x_2 \in X_2$ from the common domain of these functions. Construct a standard sequence within $\mathbf{SE}^z$, this time on X_2. By **A4**, it is also a sequence in **NW**. We obtain

$$\begin{aligned} \lambda =& V_2^{SE}(r_2^2) \approx n[V_1^{SE}(x_1^s) - V_1^{SE}(x_1^s)] \\ & V_2^{SE}(x_2) \approx m[V_1^{SE}(x_1^s) - V_1^{SE}(x_1^s)] \approx \frac{\lambda m}{n} \end{aligned} \tag{2.21}$$

By the argument above we get

$$\begin{aligned} 1 =& V_2^{NW}(r_2^2) \approx n[V_2^{NW}(x_1^n) - V_2^{NW}(x_1^n)] \\ & V_2^{NW}(x_2) \approx m[V_1^{NW}(x_1^n) - V_1^{NW}(x_1^n)] \approx \frac{m}{n} \end{aligned} \tag{2.22}$$

From this in the limit we obtain $V_2^{SE}(x_2) = \lambda V_2^{NW}(x_2)$ on all $x_2 \in X_2$ from the common domain of V_2^{SE} and V_2^{NW}. □

At this point we can drop superscripts and say that we have representations $V_1 + V_2$ on **SE** and $V_1 + \lambda V_2$ on **NW**. Fix two non-extreme points in $\Theta : r^0$ and r^1, such that $r_1^1 \succcurlyeq_1 r_1^0$ and $r_2^1 \succcurlyeq_2 r_2^0$. If such points do not exist, then by Lemma 2 triple cancellation holds everywhere and $\succcurlyeq$ can be represented by an additive function (i.e. $\lambda = 1$). Rescale V_1 and V_2 so that $V_1(r_1^0) = 0, V_2(r_2^0) = 0, V_1(r_1^1) = 1$. Assume that after rescaling we get $V_1(r_2^1) = k$. Define $\phi_2(x_2) = V_2(x_2)/k$, i.e. $\phi_2(r_2^1) = 1$. Define $\phi_1(x_1) = V_1(x_1)$. Thus we get representations $\phi_1 + k\phi_2$ on **SE** and $\phi_1 + \lambda k \phi_2$ on **NW**. Finally rescale in the following way: $\frac{1}{1+k}\phi_1 + \frac{k}{1+k}\phi_2$ on **SE** and $\frac{1}{1+\lambda k}\phi_1 + \frac{\lambda k}{1+\lambda k}\phi_2$ on **NW**. We have thus defined the following representations:

$$\begin{aligned} \phi^{SE}(x) &= \frac{1}{1+k}\phi_1(x_1) + \frac{k}{1+k}\phi_2(x_2) \\ \phi^{NW}(x) &= \frac{1}{1+\lambda k}\phi_1(x_1) + \frac{\lambda k}{1+\lambda k}\phi_2(x_2). \end{aligned} \tag{2.23}$$

Note, that it follows that $\phi^{SE}(r^1) = \phi^{NW}(r^1) = 1$.

2.6.2.2 One area has a single essential coordinate

Assume **SE** has two essential coordinates and **NW** only has $i = 1$ essential. After an additive representation V^{SE} has been built on **SE**, and re-scaled as in (2.23) we have values ϕ_1 and ϕ_2 for all points in **SE**, in particular those in Θ. Let $\phi^{NW}(x) = \phi_1(x_1) + 0\phi_2(x_2)$ (in other words, set $\lambda = 0$ in (2.23)) for those x_i where ϕ_i are defined. By structural assumption, bi-independence and additivity ϕ^{NW} represents $\succcurlyeq$ on those points for which it is defined. For example, let $ap, bp \in \mathbf{NW}$ be such that $ap \succ bp$. Since both coordinates are essential on **SE** by bi-independence we get also $aq \succ bq$ for all $q \in X_2$ such that $aq, bq \in \mathbf{SE}$. Additivity implies $\phi_1(a) > \phi_1(b)$. For the remaining $x_1 \in X_1$, i.e. for $x_1 \in X_1$ such that there are no points in Θ first coordinate of which is x_1, build a simple ordinal representation. Structural assumptions trivially imply that values for all $x_2 \in X_2$ have already been defined at this point. Other cases are similar.

2.6.2.3 Both areas have a single essential coordinate

An interesting result is that **A3** is sufficient for characterization of cases where both **SE** and **NW** have one essential coordinate. There are two cases in total:

1. $i = 1$ is essential on **NW**, $i = 2$ is essential on **SE**;
2. $i = 2$ is essential on **NW**, $i = 1$ is essential on **SE**.

We will need the following lemma.

Lemma 5. *Let $i = 1$ be essential on* **NW** *and $i = 2$ be essential on* **SE** *or $i = 2$ be essential on* **NW** *and $i = 1$ be essential on* **SE**. *Then, either for all $x \in \mathbf{SE}$ exists $z \in \Theta$ such that $z \sim x$, or for all $y \in \mathbf{NW}$ exists $z \in \Theta$ such that $z \sim y$.*

Proof. We only consider one case, others being symmetrical. Assume $i = 1$ is essential on **SE** and $i = 2$ is essential on **NW**. Assume also there exists $x \in \mathbf{SE}$ such that $x \succ z$ for all $z \in \Theta$, in particular some maximal $z^{\max}$. We will show that this implies that there does not exist $y \in \mathbf{NW}$ such that $y \succ z$ or $z \succ y$ for all $z \in \Theta$.

Assume such y exists. Take x_1y_2. By **A3** it belongs either to **SE** or **NW**. If it belongs to **NW**, by closedness assumption exists $x_1t \in \Theta$. We get $x_1t \sim x \succ z^{\max}$ - a contradiction. If $x_1y_2 \in$ **SE**, exists $ay_2 \in \Theta$. We have $y \sim ay_2$ which contradicts both $y \succ z$ and $z \succ y$ for all $z \in \Theta$.

Finally, we need to show that Θ does not have gaps. Assume there exists $y \in$ **NW** and $z^1, z^2 \in \Theta$ such that $z^1 \succ y \succ z^2$ but there is no $z \in \Theta$ such that $z \sim y$. Since only $i = 2$ is essential on **NW**, we get $z_2^1 \succcurlyeq_2 y_2$ and hence $z_1^1 y_2 \in$ **SE**. By closedness assumption exists $x_1 \in X_1$ such that $x_1y_2 \in \Theta$. Since only $i = 2$ is essential we conclude $x_1y_2 \sim y$, which is a contradiction. Therefore, for every $y \in$ **NW** there exists $z \in \Theta$ such that $y \sim z$. □

Defining value functions. Lemma 5 guarantees that for all points in **SE** or all points in **NW** exists an equivalent point in Θ. Assume for example that Θ is such that for all $x \in$ **SE** exists $z \in \Theta$. Assume also that $i = 1$ is essential on **SE** and $i = 2$ is essential on **NW**. Now define value functions $\phi_1 : X_1 \to \mathbb{R}$ and $\phi_2 : X_2 \to \mathbb{R}$ as follows. Choose ϕ_1 to be any real-valued function such that $\phi_1(x_1) > \phi_1(y_2)$ iff $x_1 \succcurlyeq_1 y_1$. Now for all z from Θ set $\phi_2(z_2) = \phi_1(z_1)$. Finally, extend ϕ_2 to the whole X_2 by choosing any function such that $\phi_2(x_2) > \phi_2(y_2)$ iff $x_2 \succcurlyeq_2 y_2$. At this point functions have been defined for all $x_1 \in X_1$ and all $x_2 \in X_2$.

2.6.2.4 Areas without essential coordinates

Lemma 6. *If **A1 - A9** and the structural assumption hold, there can not be* $\mathbf{NW}^z(\mathbf{SE}^z)$ *with no essential coordinates.*

Proof. Assume for some $z \in \Theta$ the set $\mathbf{NW}^z$ has no essential coordinates. By bi-independence and the structural assumption it follows that there are no essential coordinates on any $\mathbf{NW}^z$. This implies (by **A7**) that both coordinates are essential on **SE**. Take $ap, bp \in \mathbf{NW}^z$. Apparently, $ap \sim bp$. By structural assumption there must exist $q \in X_2$ such that $aq \succ bq$. It can't be that $aq, bq \in$ **NW**, hence $aq, bq \in$ **SE**.

By closedness assumption there exist $w, z \in X_2$ such that $aw, bz \in \Theta$. Also, since no coordinate is essential in **NW** we have $aw \sim bz$. Since $aq \succ bq$ it must be $bz \succ bw$, since otherwise ($aw \succ bw \succcurlyeq bz$) it can't be that $aw \sim bz$.

By independence we have $aw \succ bw$. By definition of **NW** and **SE** we have $aw \in$ **SE** (since by weak separability $a \succcurlyeq_1 b$) and $aw \in$ **NW** (since by weak separability $z \succcurlyeq_2 w$). Hence, by independence it must be $az \succ aw$ (since $az \in$ **SE**) but we have $az \sim aw$ (since $az \in$ **NW**). We have arrived at a contradiction. $\square$

2.7 Properties of the intersection of SE and NW

We would need a technical lemma first - in fact a restatement of the axiom **A5.** Figure 2.3 in Section 2.2.1 gives an illustration of this condition.

Lemma 7. *Axiom **A5** implies the following condition. Let* $\{g_1^{(i)} : g_1^{(i)}y_0 \sim g_1^{(i+1)}y_1, g_1^{(i)} \in X_1, i \in N\}$ *and* $\{h_2^{(i)} : x_0h_2^{(i)} \sim x_1h_2^{(i+1)}, h_2^{(i)} \in X_x, i \in N\}$ *be two standard sequences, each entirely contained in* **NW** *or* **SE**. *Assume also, that there exist* $z_1, z_2 \in X$, $p, q \in X_2, a, b \in X_1$ *such that* $g_1^{(i)}p, g_1^{(i)}q \in$ **NW** *or* **SE**, *and* $ah_2^{(i)}, bh_2^{(i)} \in$ **NW** *or* **SE** *for all* i, *and* $g_1^{(i)}p \sim bh_2^{(i)}$ *and* $g_1^{(i+1)}p \sim bh_2^{(i+1)}$, *then for all* $i \in N$, $g_1^{(i)}p \sim bh_2^{(i)}$.

Proof. The proof is very similar to the one from Krantz et al. (1971) (Lemma 5 in section 8.3.1). Assume wlog that $\{g_1^{(i)} : g_1^{(i)}y_0 \sim g_1^{(i+1)}y_1\}$ is an increasing standard sequence on X_1, which is entirely in **SE**, whereas $\{h_1^{(i)} : x_0h_1^{(i)} \sim x_1h_1^{(i+1)}\}$ is an increasing standard sequence on X_2, and lies entirely in **NW**. Assume also for some k it holds $g_1^{(k)}y_0 \sim x_0h_2^{(k)}, g_1^{(k+1)}y_0 \sim x_0h_2^{(k+1)}$. We need to show that $g_1^{(i)}y_0 \sim x_0h_2^{(i)}$ for all i. We will show that $g_1^{(k+2)}y_0 \sim x_0h_2^{(k+2)}$ from which everything holds by induction.

Assume $x_0h_2^{(k+2)} \succ g_1^{(k+2)}y_0$. Since the sequences are increasing, by restricted solvability exists $g \in X_2$ such that $g_1^{(k+2)}y_0 \sim x_0g$. By **A5**, $g_1^{(k)}y_0 \sim g_1^{(k+1)}y_1, g_1^{(k+1)}y_0 \sim g_1^{(k+2)}y_1, x_0h_2^{(k)} \sim x_1h_2^{(k+1)}$ imply $x_0h_2^{(k+1)} \sim x_1g$. By definition of $\{h_2^{(i)}\}$, $x_0h_2^{(k+1)} \sim x_1h_2^{(k+2)}$. Thus, $x_1h_2^{(k+2)} \sim x_1g$ and by independence $x_0h_2^{(k+2)} \sim x_0g$, hence also $g_1^{(k+2)}y_0 \sim x_0h^{(k+2)}$, a contradiction. The case $x_0h_2^{(k+2)} \prec g_1^{(k+2)}y_0$ is symmetrical. Showing that $g_1^{(k-1)}y_0 \sim x_0h_2^{(k-1)}$ can be done in a similar fashion. $\square$

The main result of this section is the following lemma.

Lemma 8. *For any non-extreme $x \in X$ we have:*

$$x \in \Theta \Rightarrow \phi_1(x_1) = \phi_2(x_2), \tag{2.24}$$

unless $\succcurlyeq$ can be represented by an additive function (i.e $\lambda = 1$ in (2.23)).

For the case when both **NW** and **SE** have a single essential coordinate the result holds by definition of ϕ_i, so for the remainder of this section we assume that **SE** or **NW** has two essential coordinates.

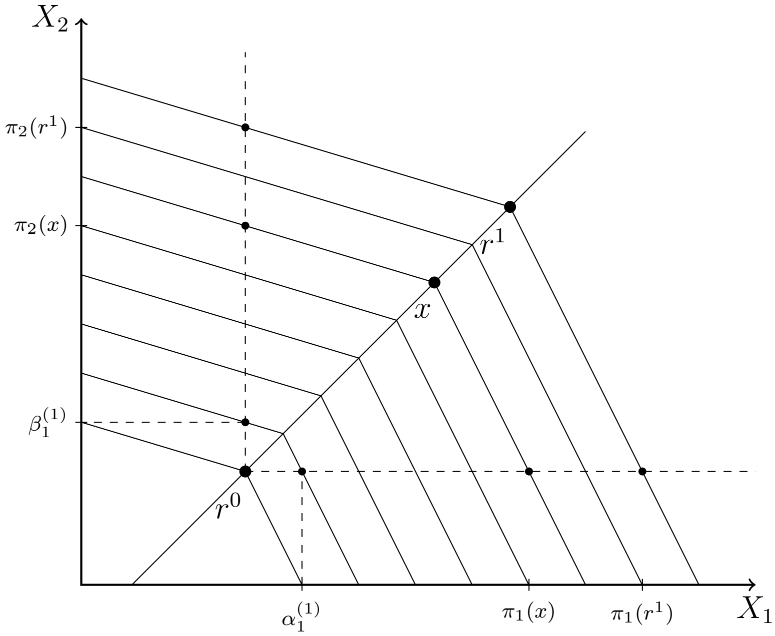


Figure 2.5: Lemma 8

Proof. We start with a case where both coordinates are essential on **SE** and **NW**. Assume also $x \succcurlyeq r^0$ (without loss of generality, other cases are symmetrical and can be proved by the same technique). We are going to show that $x \in \Theta \Rightarrow \phi_1(x_1) = \phi_2(x_2)$ or $\lambda = 1$.

Assume for now that we can find the following solutions:

- Solve for $\pi_1(r^1) : \pi_1(r^1) r_2^0 \sim r^1$.
- Solve for $\pi_2(r^1) : r_1^0 \pi_2(r^1) \sim r^1$.

- Solve for $\pi_1(x) : \pi_1(x)r_2^0 \sim x$.
- Solve for $\pi_2(x) : r_1^0\pi_2(x) \sim x$.

Pick $\alpha_1^{(1)} \in X_1$ such that $r_1^1 \succcurlyeq_1 \alpha_1^{(1)}$, it exists by denserangedness. Solve for $\beta_2^{(1)} : \alpha_1^{(1)} r_2^0 \sim r_1^0 \beta_2^{(1)}$, exists by restricted solvability.

Now build an increasing standard sequence $\alpha_1^{(i)} : \alpha_1^{(0)} = r_1^0$ on X_1 which lies in **SE** and an increasing standard sequence $\beta_2^{(i)} : \beta_2^{(0)} = r_2^0$ on X_2 which lies in **NW** (see Fig. 2.5). Since $\pi_1(r^1)r_2^0 \sim r_1^0\pi_2(r^1)$, by **A5** (Lemma 7) we have for some m:

$$\alpha_1^{(m-1)} r_2^0 \succcurlyeq \pi_1(r^1) r_2^0 \succcurlyeq \alpha_1^{(m)} r_2^0 \iff r_1^0 \beta_2^{(m-1)} \succcurlyeq r_1^0 \pi_2(r^1) \succcurlyeq r_1^0 \beta_2^{(m)}. \tag{2.25}$$

From this (and since $\phi_1(r_1^0) = \phi_2(r_2^0) = 0$) it follows that $\phi_1(\pi_1(r^1)) \approx m\phi_1(\alpha_1^{(1)}), \phi_2(\pi_2(r^1)) \approx m\phi_2(\beta_2^{(1)})$ and, since $\phi^{SE}(\pi_1(r^1)r_2^0) = \phi^{SE}(r^1) = \phi^{NW}(r^1) = \phi^{NW}(r_1^0\pi_2(r^1))$, we obtain:

$$\frac{1}{1+k} m\phi_1(\alpha_1^{(1)}) = \frac{\lambda k}{1+\lambda k} m\phi_2(\beta_2^{(1)}). \tag{2.26}$$

Similarly, $\pi_1(x)r_2^0 \sim r_1^0\pi_2(x)$, so by **A5** (Lemma 7) we have

$$\alpha_1^{(n-1)} r_2^0 \succcurlyeq \pi_1(x) r_2^0 \succcurlyeq \alpha_1^{(n)} r_2^0 \iff r_1^0 \beta_2^{(n-1)} \succcurlyeq r_1^0 \pi_2(x) \succcurlyeq r_1^0 \beta_2^{(n)}. \tag{2.27}$$

From this follows that $\phi_1(\pi_1(x)) \approx n\phi_1(\alpha_1^{(1)}), \phi_2(\pi_2(x)) \approx n\phi_2(\beta_2^{(1)})$ and by (2.26) it follows that $\phi^{SE}(\pi_1(x)r_2^0) = \phi^{NW}(r_1^0\pi_2(x))$. Hence

$$\frac{1}{1+k}\phi_1(x_1) + \frac{k}{1+k}\phi_2(x_2) = \frac{1}{1+\lambda k}\phi_1(x_1) + \frac{\lambda k}{1+\lambda k}\phi_2(x_2), \tag{2.28}$$

and so $\phi_1(x_1) = \phi_2(x_2)$ or $\lambda = 1$ (i.e. the structure is additive).

We need to revisit the case where solutions mentioned in the beginning do not exist. Consider Figure 2.6. Assume this time that there does not exist $\pi_1(r^1)$ such that $r^1 \sim \pi_1(r^1)r_2^0$. If we choose the step in the standard sequence $\alpha_1^{(i)}$ small enough so that there exist $\alpha_1^{(k+1)}, \alpha_1^{(k+2)}$ such that $\alpha_1^{(k+1)} r_2^0 \succcurlyeq r_0^1 r_2^0$ and $\alpha_2^{(k+2)} r_2^0 \succcurlyeq r_0^1 r_2^0$ (which we can do by non-maximality of r^1 and denserangedness of $\succcurlyeq$), then we can "switch" from the standard sequence $\alpha_1^{(i)}$ on X_1 to the standard sequence $\gamma_2^{(i)}$ on X_2 keeping the same increment in value between subsequent members of the sequence. Indeed, $\gamma_2^{(k)} : r_1^1\gamma_2^{(k)} \sim \alpha_1^{(k)}r_2^0$ and $\gamma_2^{(k+1)} : r_1^1\gamma_2^{(k+1)} \sim \alpha_1^{(k+1)}r_2^0$ exist by monotonicity and restricted solvability, so does $x_1 \colon x_1\gamma_2^{(k)} \sim r_1^1\gamma_2^{(k+1)} \sim \alpha_1^{(k+1)}r_2^0$, and by **A5** (Lemma 7) we get $[r_1^1\gamma_2^{(k)} \sim r_1^0\beta_2^{(k)}, r_1^1\gamma_2^{(k+1)} \sim r_1^0\beta_2^{(k+1)}] \Rightarrow \gamma_2^{(i)} \sim \beta_2^{(i)}$ for all i. Note that $\gamma_2^{(i)}r_2^1$ are in **SE** for all i such that $r^1 \succcurlyeq \gamma_2^{(i)}$ since r^0 and r^1 are in **SE**. By monotonicity and restricted solvability there exists i such that $x_1r_2^1 \succcurlyeq x_1\gamma_2^{(i+1)} \succcurlyeq r^1 \succcurlyeq x_1\gamma_2^{(i)}$. Finally, the increment in value is the same between members of α_i and $\gamma_2^{(i)}$ since $\alpha_k \sim \gamma_2^{(k)}$ and $\alpha_{k+1} \sim \gamma_2^{(k+1)}$. The result then follows as above.

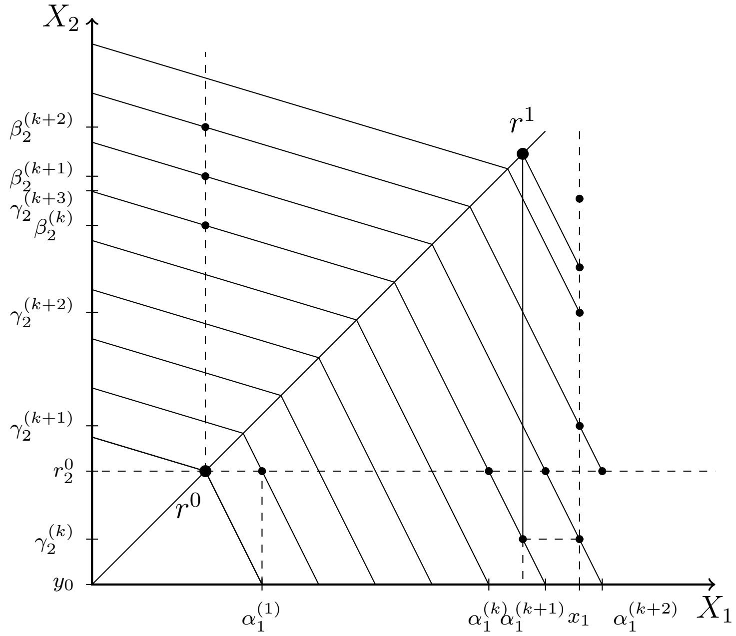


Figure 2.6: Lemma 8 - changing direction

Finally, we look at the case where only one coordinate is essential on either **NW** or **SE**. First assume that $i = 2$ is essential on **NW**. We defined $\phi^{NW}(x) = 0\phi_1(x_1) + \phi_2(x_2)$. Definition implies $\phi^{NW}(r^0) = 0, \phi^{NW}(r^1) = 1$. Build a standard sequence $\{\alpha_1^{(i)}\}$ on X_1 from r^0 to r^1 (in case there exists a solution for $r^1 \sim \pi_1(r^1)r_2^0$, otherwise use the approach detailed in the previous paragraph), setting $\alpha_1^{(0)} = r_1^0$. Take $\alpha_1^{(1)}r_2^0$ and $\alpha_1^{(2)}r_2^0$. By restricted solvability there must exist $\beta_2^{(1)}$ and $\beta_2^{(2)}$, such that $\alpha_1^{(1)}r_2^0 \sim r_1^0\beta_2^{(1)}$ and $\alpha_1^{(2)}r_2^0 \sim r_1^0\beta_2^{(2)}$. By closedness assumption for $\beta_2^{(1)}, \beta_2^{(2)}$ there must exist x_1, x_2 such that $x_1\beta_2^{(1)} \in \Theta$, $x_2\beta_2^{(2)} \in \Theta$. Also, since only $i = 2$ is essential, we get $x_1\beta_2^{(1)} \sim r_1^0\beta_2^{(1)}$, $x_2\beta_2^{(2)} \sim r_1^0\beta_2^{(2)}$. By weak monotonicity and definition of **SE**, $\alpha_1^{(2)}\beta_2^{(1)} \succcurlyeq \alpha_1^{(2)}r_2^0 \sim x_2\beta_2^{(2)} \succcurlyeq x_2\beta_2^{(1)}$, hence by restricted solvability exists $z_1 : x_2\beta_2^{(2)} \sim z_1\beta_2^{(1)}$. By **A5** then $x_2\beta_2^{(1)} \sim z_1r_2^0$. By additivity $x_2\beta_2^{(2)} \sim z_1\beta_2^{(1)}$ and $x_2\beta_2^{(1)} \sim z_1r_2^0$ entail $\phi_2(b_2) - \phi_2(b_1) = \phi_2(b_1) - \phi_2(r_2^0)$. From this the result follows as in the proof above. If now $i = 1$ is essential on **NW** repeat the proof as above this time starting the sequence from r^1 "towards" r^0.

□

Lemma 9. *The following statements hold or $\succcurlyeq$ has an additive representation:*

1. *If $ap \in \Theta$ then for no $b \in X_1$ holds $bp \in \Theta$ and also for no $q \in X_2$ holds $aq \in \Theta$.*

2. *$x \in$ **SE** $\Rightarrow \phi_1(x_1) \geq \phi_2(x_2)$, $y \in$ **NW** $\Rightarrow \phi_2(y_2) \geq \phi_1(y_1)$.*

Proof.

1. Assume $\succcurlyeq$ does not have an additive representation, hence $[x \in \Theta] \Rightarrow \phi_1 = \phi_2$. We have $ap \in \Theta$, hence $\phi_1(a) = \phi_2(p)$. Now take bp, assume $b \succ_2 a$. Hence, $\phi_1(b) > \phi_1(a)$ or otherwise $i = 1$ is inessential in both **SE** and **NW** which contradicts assumptions. Thus, $\phi_1(b) > \phi_1(a)$ and hence bp can't be in Θ.

2. Pick any $bq \in$ **SE**. By Lemma 1 there exists $ap \in \Theta$ such that $bq \in \mathbf{SE}_{ap}$, hence $b \succcurlyeq_1 a$ and $p \succcurlyeq_2 q$. By Lemma 8 $\phi_1(a) = \phi_2(p)$. We also have $\phi_1(b) \geq \phi_1(a), \phi_2(p) \geq \phi_2(q)$. The result follows. **NW** case is symmetric.

□

2.8 Extending value functions to extreme points

Value functions for the case when both **SE** and **NW** have a single essential coordinate were fully defined in Section 2.6.2.3. Thus in what follows we will consider cases where **SE** or **NW** have two essential coordinates.

As indicated in (Wakker, 1991b), value functions might be driven to infinite values at the maximal/minimal points of rank-ordered subsets, nevertheless not implying existence of infinite standard sequences residing entirely within comonotonic cones. Put it another way, it might be not possible to "reach" a maximal/minimal point with a sequence lying entirely in **NW** or **SE**. Yet another way to say it is that for some maximal/minimal point z, the set $\mathbf{NW}^z(\mathbf{SE}^z)$ contains no standard sequences (see also (Wakker, 1991b) Remark 24).

The cornerstone of this section is Lemma 8. It plays the same role as proportionality of value functions plays in (Wakker, 1991b), effectively guaranteeing that both value functions ϕ_1 and ϕ_2 are limited if maximal/minimal elements exist.

Lemma 10. *Assume that* **SE** *has two essential coordinates. The following statements hold:*

- *If there exist a maximal $M_1 \in X_1$, ϕ_1 is bounded from above.*
- *If there exist a minimal $m_2 \in X_2$, ϕ_2 is bounded from below.*

Assume that **NW** *has two essential coordinates. The following statements hold:*

- *If there exist a minimal $M_1 \in X_1$, ϕ_1 is bounded from below.*
- *If there exist a maximal $m_2 \in X_2$, ϕ_2 is bounded from above.*

Proof. We shall only prove the first one. First, notice that there must exist $p \in X_2$ such that $M_1 p \in \mathbf{SE}$. Take $x_1 \in X_1$ and $v_2, w_2 \in X_2$ such that $v_2 \succcurlyeq_2 w_2$, and $x_1 v_2 \in \mathbf{SE}$. If such points cannot be found, X has an additive representation (all $x \in \mathbf{NW}$), and the result follows. So we assume such points exist. By definition of $\mathbf{SE}^{x_1 v_2}$ it follows that $M_1 w_2, x_1 v_2, M_1 v_2 \in \mathbf{SE}$. Hence, we can evoke the argument from Wakker (1991b) Lemma 20.

If $M_1 w_2 \preccurlyeq x_1 v_2$ then we have an upper bound: $V_1(M_1) \leq V_1(x_1) + V_2(v_2) - V_2(w_2)$. If $M_1 w_2 \succ x_1 v_2$ then by monotonicity $M_1 v_2 \succcurlyeq M_1 w_2 \succ x_1 v_2$ and hence exists $z_1 \in X_1$ such that $M_1 w_2 \sim z_1 v_2$, hence $z_1 v_2 \succcurlyeq \beta w_2$ for all $\beta \in X_1$. □

Lemma 11. *If $x_1 x_2 \in \Theta$ and $x_1 x_2$ is extreme, then*

$$\lim_{z \in \Theta, z_2 \to x_2} \phi_2(z_2) = \lim_{z \in \Theta, z_1 \to x_1} \phi_1(z_1). \tag{2.29}$$

Proof. For the case when **SE** or **NW** have two essential variables the result follows from Lemma 8, otherwise it is by definition of ϕ_i (see Section 2.6.2.3). □

2.8.1 Extending value functions to extreme elements of Θ

The following lemma characterizes the cases when extreme elements of Θ are the only representatives of maximal/minimal equivalence classes of $\mathbf{SE}(\mathbf{NW})$.

Lemma 12. *Let $i = 1$ be essential on $\mathbf{SE}$. If there exists $z \in \Theta$ such that z_2 is minimal, then $x \succ z$ for all $x \in \mathbf{SE}$. If $i = 2$ is essential on $\mathbf{SE}$ and there exists $z \in \Theta$ such that z_1 is maximal, then $z \succ x$ for all $x \in \mathbf{SE}$. Similarly, if $i = 1$ is essential on $\mathbf{NW}$ and there exists $z \in \Theta$ such that z_2 is maximal, then $z \succ x$ for all $x \in \mathbf{NW}$. If $i = 2$ is essential on $\mathbf{NW}$ and there exists $z \in \Theta$ such that z_1 is minimal, then $x \succ z$ for all $x \in \mathbf{NW}$.*

Proof. We provide the proof just for one of the cases. Let $\mathbf{NW}$ have two essential variables. Assume z_2 is maximal. Since $z \in \Theta$, for all $x \in \mathbf{NW}$ holds $z_1 \succcurlyeq_1 x_1$ and by maximality $z_2 \succcurlyeq_2 x_2$. Hence, by essentiality and **A6**, $z \succ x$ for all $x \in \mathbf{NW}$. The case with the minimal z_1 is symmetric. □

The following results look at the uniqueness of definition of ϕ_i at the extreme elements of Θ.

Lemma 13. *If both coordinates are essential on $\mathbf{SE}$ and $\mathbf{NW}$ the values of ϕ_i for extreme $x \in \Theta$ are uniquely defined. Moreover, $\phi_1(x_1) = \phi_2(x_2)$.*

Proof. Assume, for example $x_1x_2 \in \Theta$ and x_1 is minimal. Then any z_1x_2 such that $z_1 \succcurlyeq_1 x_1$, belongs to $\mathbf{SE}$, and any equivalence relation within $\mathbf{SE}$ involving z_1x_2 uniquely defines $\phi_2(x_2)$ (see (2.23)). Similarly, any x_1z_2 such that $z_2 \succcurlyeq_2 x_2$, belongs to $\mathbf{NW}$, and any equivalence relation within $\mathbf{NW}$ involving x_1z_2 uniquely defines $\phi_1(x_1)$. By Lemma 11 these values are equal. □

Lemma 14. *If both coordinates are essential on $\mathbf{SE}$ but only one on $\mathbf{NW}$ (or vice versa) the values of ϕ_i for extreme $x \in \Theta$ can be set as follows:*

- *If $x_1x_2 \in \Theta$, $\mathbf{NW}$ has two essential coordinates and x_2 is maximal, then $\phi_2(x_2)$ is uniquely defined, $\phi_1(x_1)$ can be set to any value greater or equal to $\phi_2(x_2)$.*

- *If $x_1x_2 \in \Theta$, **NW** has two essential coordinates and x_1 is minimal, then $\phi_1(x_1)$ is uniquely defined, $\phi_2(x_2)$ can be set to any value less or equal to $\phi_1(x_1)$.*

- *If $x_1x_2 \in \Theta$, **SE** has two essential coordinates and x_2 is minimal, then $\phi_2(x_2)$ is uniquely defined, $\phi_1(x_1)$ can be set to any value less or equal to $\phi_2(x_2)$.*

- *If $x_1x_2 \in \Theta$, **SE** has two essential coordinates and x_1 is maximal, then $\phi_1(x_1)$ is uniquely defined, $\phi_2(x_2)$ can be set to any value greater or equal to $\phi_1(x_1)$.*

Proof. Consider the first case. $\phi_2(x_2)$ is defined uniquely as in the proof of Lemma 13. However, this is not possible for $\phi_1(x_1)$. This is because x is the only point in **NW** having x_1 as the first coordinate, and, by Lemma 12 there is no equivalence relation within **NW** which involves x. If $i = 1$ is essential on **SE** then all points from the equivalence class which includes x also have x_1 as their first coordinate, which does not allow to elicit $\phi_1(x_1)$. If only $i = 2$ is essential on **SE**, then the representations of equivalences involving x_1 do not include $\phi_1(x_1)$. $\square$

In the case where only one coordinate is essential on both **SE** and **NW** no special treatment is required for the extreme elements of Θ.

Lemma 15. *If $x \in \mathbf{SE}_{ext}$ then for any $y \in NW$ such that $x \sim y$, we have:*

$$\phi^{NW}(x) = \phi^{NW}(y). \tag{2.30}$$

If further, x_1 is maximal, then $\phi^{SE}(x) > \phi^{SE}(y)$ for all $y \in \mathbf{SE}$. If x_2 is minimal, then $\phi^{SE}(x) < \phi^{SE}(y)$ for all $y \in \mathbf{SE}$.

If $x \in \mathbf{NW}_{ext}$ then for any $y \in SE$ such that $x \sim y$, we have:

$$\phi^{SE}(x) = \phi^{SE}(y). \tag{2.31}$$

If further, x_1 is minimal, then $\phi^{NW}(x) < \phi^{NW}(y)$ for all $y \in \mathbf{NW}$. If x_2 is maximal, then $\phi^{NW}(x) > \phi^{NW}(y)$ for all $y \in \mathbf{NW}$.

Proof. In case when **SE** and **NW** have the same number of essential variables, value functions at the extreme points are defined uniquely (Lemma 13 for case when both variables are essential and by definition otherwise) and the second parts of each statement follow immediately by Lemma 12. If only **SE** or **NW** have two essential variables, the result follows by Lemma 14, Lemma 12, and definition of ϕ. □

Lemma 16. *For any $x \in X$ we have:*

$$\phi_1(x_1) = \phi_2(x_2) \Rightarrow x \in \Theta. \tag{2.32}$$

Proof. Assume $\phi_1(x_1) = \phi_2(x_2)$ and $x \notin \mathbf{NW}$. By Lemma 1 exists $z \in \Theta$ such that $x \in \mathbf{SE}^z$. By structural assumption we have $\phi_2(z_2) \geq \phi_2(x_2) = \phi_1(x_1) \geq \phi_1(z_1)$, with at least one inequality being strict (otherwise $x = z$).

If z is non-extreme then by Lemma 8 we have $\phi_1(z_1) = \phi_2(z_2)$ - a contradiction. If z is extreme, the only cases when $\phi_2(z_2) > \phi_1(z_1)$ can hold is when either z_2 is minimal or z_1 is maximal. But it is easy to see that in this case the only points for which it is not possible to find a non-extreme z, are the extreme points themselves. □

Lemma 17. *For all $x \in X$ such that $\phi_1(x_1) \geq \phi_2(x_2)$ we have $x \in \mathbf{SE}$. If $x \in X$ is such that $\phi_2(x_2) \geq \phi_1(x_1)$ then $x \in \mathbf{NW}$.*

Proof. For non-extreme points this follows from Lemma 9 and Lemma 16. Assume $\phi_1(x_1) \geq \phi_2(x_2)$. If $\phi_1(x_1) = \phi_2(x_2)$ then by Lemma 16 $x \in \Theta$, so we are done. Therefore, assume $\phi_1(x_1) > \phi_2(x_2)$. If $x \in \mathbf{NW}$, then by Lemma 9 it must be $\phi_2(x_2) \geq \phi_1(x_1)$, a contradiction. Therefore, $x \in \mathbf{SE}$. For extreme points the result follows from Lemma 14. □

Finally, we can formulate:

Theorem 6. *The following statements hold:*

- *If both **NW** and **SE** have two essential variables, then for all $x \in X$:*

$$x \in \Theta \iff \phi_1(x_1) = \phi_2(x_2), \tag{2.33}$$

unless $\succcurlyeq$ can be represented by an additive function (i.e $\lambda = 1$ in (2.23)).

- *If only* **NW** *or only* **SE** *have two essential variables, the for all non-extreme $x \in X$:*

$$x \in \Theta \Rightarrow \phi_1(x_1) = \phi_2(x_2), \tag{2.34}$$

while at extreme $x \in X$, $\phi_1(x_1)$ and $\phi_2(x_2)$ are related as in Lemma 14. Finally, for all $x \in X$:

$$\phi_1(x_1) = \phi_2(x_2) \Rightarrow x \in \Theta, \tag{2.35}$$

- *If both* **NW** *and* **SE** *have only one essential variable, then for all $x \in X$:*

$$x \in \Theta \iff \phi_1(x_1) = \phi_2(x_2). \tag{2.36}$$

Proof. Follows from Lemmas 8, 16, 13, 14.

□

2.9 Constructing a global representation on X

For all $x \in X$ let $\phi^x(x)$ be equal to $\phi^{SE}(x)$ if $x \in$ **SE** or $\phi^{NW}(x)$ if $x \in$ **NW**. For points in Θ values of two latter functions coincide, so $\phi^x(x)$ is well-defined.

Lemma 18. *Let $\phi^x(x) > \phi^y(y)$. Then, $x \succ y$.*

Proof. If x and y belong to **SE** or **NW** the conclusion is immediate, so we only need to look at the remaining case. Assume $x \in$ **SE**, $y \in$ **NW**.

First we will show that it can't hold that $x \succ z, y \succ z$ or $z \succ x, z \succ y$ for all $z \in \Theta$. Assume $x \succ z, y \succ z$ for all $z \in \Theta$. Let $x_1 \succcurlyeq_1 y_1$. If $x_1y_2 \in$ **SE**, then exists $z_1y_2 \in \Theta$ such that $x_1y_2 \succcurlyeq z_1y_2 \succcurlyeq y_1y_2$, a contradiction. If $x_1y_2 \in$ **NW**, then exists $x_1z_2 \in \Theta$, such that $x_1y_2 \succcurlyeq x_1z_2 \succcurlyeq x_1x_2$, again a contradiction. Other cases are symmetrical.

Hence, assume there exists $z^1 \in \Theta$, such that $x \succcurlyeq z^1, y \succcurlyeq z^1$ and $z^2 \in \Theta$ such that $z^2 \succcurlyeq x$ or $z^2 \succcurlyeq y$. The only non-trivial case is $z^2 \succ x \succ z^1, z^2 \succ y \succ z^1$ (in other cases one of the points z^1 or z^2 immediately leads to the conclusion). We have

$$\phi(z^2) > \phi(x) > \phi(y) > \phi(z^1), \tag{2.37}$$

hence also $\phi_1(z_1^2) = 0.5\phi(z^2) > 0.5\phi(x) > 0.5\phi(y) > 0.5\phi(z^1) = \phi_1(z_1^1)$. By denserangedness of ϕ_1 (see (Wakker, 1991b) equation (16)), we can find a point c_1 such that $0.5\phi(x) > \phi(c_1) > 0.5\phi(y)$. We have $c_1 z_2^1 \in \mathbf{SE}, c_1 z_2^2 \in \mathbf{NW}$, hence there exists c_2 such that $c_1 c_2 \in \Theta$. Since $c_1 c_2$ is not extreme, we have $\phi(c_1 c_2) = 2\phi_1(c_1)$, and hence $\phi(x) > \phi(c) > \phi(y)$. The first inequality is in $\mathbf{SE}$, while the second is in $\mathbf{NW}$, hence we conclude that $x \succ y$. □

Lemma 19. *Let $x \succ y$. Then, $\phi(x) > \phi(y)$.*

Proof. By Lemma 18 we have $x \succ y \Rightarrow \phi(x) \geq \phi(y)$. Hence, we need to show that $\phi(x) \neq \phi(y)$. Assume, $x \in \mathbf{SE}, y \in \mathbf{NW}$. If $\mathbf{SE}$ or $\mathbf{NW}$ have only one essential coordinate, then by structural assumptions exists $z \in \Theta$, equivalent either to x or y, from which the conclusion is immediate. Hence, assume both areas have two essential coordinates.

x_1 and x_2 can't be both minimal, because otherwise $x \succ y$ cannot hold by pointwise monotonicity, so assume x_1 is not minimal. We will find a point z in $\mathbf{SE}$ such that $x \succ z \succcurlyeq y$. Take some z_1 such that $x_1 \succcurlyeq_1 z_1$ and $z_1 x_2 \in \mathbf{SE}$ (it can be found by closedness and order density). If $z_1 x_2 \succcurlyeq y$, we have $\phi(x) > \phi(z_1 x_2) \geq \phi(y)$, otherwise by restricted solvability we can find w_1 such that $w_1 x_2 \sim y$, $w_1 \in \mathbf{SE}$, and hence $\phi(x) > \phi(w_1 x_2) = \phi(y)$ (equality follows from Lemma 18). The case when x_2 is not maximal is identical. □

Theorem 7. *For any $x, y \in X$ we have*

$$x \succcurlyeq y \iff \phi(x) \geq \phi(y). \tag{2.38}$$

Proof. Immediate by Lemmas 18 and 19. □

2.10 Constructing the capacity and the integral

Representations ϕ^{SE} and ϕ^{NW} uniquely define a capacity ν. For the case when **SE** or **NW** has two essential coordinates, set (using (2.23)):

$$\begin{aligned} \nu(\{1\}) &= \frac{1}{1+k} \qquad (\text{from } \phi^{SE}) \\ \nu(\{2\}) &= \frac{\lambda}{1+\lambda k} \quad (\text{from } \phi^{NW}) \\ \nu(\{1,2\}) &= 1. \end{aligned} \tag{2.39}$$

Thus, we obtain

$$\begin{aligned} C(\nu,\phi(x)) = \phi^{SE}(x) &= \frac{1}{1+k}\phi_1(x_1) + \frac{k}{1+k}\phi_2(x_2), && \text{for all } x \in \mathbf{SE}, \\ C(\nu,\phi(x)) = \phi^{NW}(x) &= \frac{1}{1+\lambda k}\phi_1(x_1) + \frac{\lambda k}{1+\lambda k}\phi_2(x_2), && \text{for all } x \in \mathbf{NW}. \end{aligned} \tag{2.40}$$

Assume now, that **SE** and **NW** has only one essential coordinate. If $i = 1$ is essential on **SE** set $\nu(1) = 1$, otherwise zero. If $i = 2$ is essential on **NW** set $\nu(2) = 1$, otherwise zero. As above, set $\nu(\{1,2\}) = 1$ We obtain:

$$\begin{aligned} C(\nu,\phi(x)) &= \phi_1(x_1), && \text{if } i = 1 \text{ is essential on the area containing } x, \\ C(\nu,\phi(x)) &= \phi_2(x_2), && \text{if } i = 2 \text{ is essential on the area containing } x, \end{aligned} \tag{2.41}$$

in particular $C(\nu,\phi(x)) = \max(\phi_1(x_1),\phi_2(x_2))$ if $i = 1$ is essential on **SE**, $i = 2$ is essential on **NW**, $C(\nu,\phi(x)) = \min(\phi_1(x_1),\phi_2(x_2))$ if $i = 2$ is essential on **SE**, $i = 1$ is essential on **NW**.

2.11 Uniqueness

Uniqueness properties are similar to those obtained in the homogeneous case $X = Y^n$, but are modified to accommodate for the heterogeneous structure of the set X in this paper. We have shown that axioms **A1-A9** are sufficient for constructing the Choquet integral representation of $\succsim$:

$$x \succsim y \iff C(\nu, (f_1(x_1), f_2(x_2))) \geq C(\nu, (f_1(y_1), f_2(y_2))), \qquad \text{(RepCh)}$$

for all $x, y \in X$.

2.11.1 Non-additive case

Lemma 20. *Assume that triple cancellation does not hold on all of X (in other words* $\mathbf{NW} \neq \mathbf{SE}$*) Representation* (2.9) *has the following uniqueness properties:*

1. *If both coordinates are essential on* $\mathbf{NW}$ *and* $\mathbf{SE}$*, then for any functions $g_1 : X_1 \to \mathbb{R}, g_2 : X_2 \to \mathbb{R}$ such that (2.9) holds with f_i substituted by g_i, we have $f_i(x_i) = \alpha g_i(x_i) + \beta$.*

2. *If both coordinates are essential on* $\mathbf{NW}$*, but only one coordinate is essential on* $\mathbf{SE}$*, then for any functions $g_1 : X_1 \to \mathbb{R}, g_2 : X_2 \to \mathbb{R}$ such that (2.9) holds with f_i substituted by g_i, we have :*

   - *$f_i(x_i) = \alpha g_i(x_i) + \beta$, for all x such that $f_1(x_1) < \max f_2(x_2)$ and $f_2(x_2) > \min f_1(x_1)$;*
   - *$f_i(x_i) = \phi_i(g_i(x_i))$, where ϕ_i is an increasing function, otherwise.*

3. *If both coordinates are essential on* $\mathbf{SE}$*, but only one coordinate is essential on* $\mathbf{NW}$*, then for any functions $g_1 : X_1 \to \mathbb{R}, g_2 : X_2 \to \mathbb{R}$ such that (2.9) holds with f_i substituted by g_i, we have :*

   - *$f_i(x_i) = \alpha g_i(x_i) + \beta$, for all x such that $f_2(x_2) < \max f_1(x_1)$ and $f_1(x_1) > \min f_2(x_2)$;*
   - *$f_i(x_i) = \psi_i(g_i(x_i))$, where ψ_i is an increasing function, otherwise.*

4. *If one coordinate is essential on* $\mathbf{NW}$ *and another one on* $\mathbf{SE}$*, then for any functions $g_1 : X_1 \to \mathbb{R}, g_2 : X_2 \to \mathbb{R}$ such that (2.9) holds with f_i substituted by g_i, we have : $f_i(x_i) = \psi_i(g_i(x_i))$ where ψ_i are increasing functions such that $f_1(x_1) = f_2(x_2) \iff g_1(x_1) = g_2(x_2)$.*

Proof. 1. Direct by uniqueness properties of additive representations.

2. Direct by uniqueness properties of additive representations, Lemma 13, Theorem 7.

3. Assume **NW** has two essential coordinates. If there exists an element in Θ such that x_1 is minimal or x_2 is maximal, then, by Lemma 10, there exist respectively a minimal $\phi_1(x_1)$ and maximal $\phi_2(x_2)$. Points x such that $f_1(x_1) < \max f_2(x_2)$ and $f_2(x_2) > \min f_1(x_1)$ are precisely the elements for which there exist $a \in X_1$ or $p \in X_2$ such that either $ax_2 \in \mathbf{NW}$ or $x_1p \in \mathbf{NW}$. From this follows that uniqueness of ϕ_i for these points is defined by the uniqueness properties of **NW** and definition of ϕ_i on **SE**, i.e. $f_i(x_i) = \alpha g_i(x_i) + \beta$. For the remaining points (including extreme elements of Θ), uniqueness is derived from the uniqueness of ordinal representations.

4. The proof is identical to the one in the previous point.

5. Uniqueness properties are derived from the uniqueness of ordinal representations and definition of ϕ_i (Section 2.6.2.3).

□

Theorem 2 directly follows from Lemma 20.

2.11.2 Additive case

If triple cancellation holds throughout X, in other words if $\mathbf{NW} = \mathbf{SE} = X$, then $\succcurlyeq$ has an additive value model representation - exist value functions $f_1 : X_1 \to \mathbb{R}$ and $f_2 : X_2 \to \mathbb{R}$, such that for any $x, y \in X$ we have

$$x \succcurlyeq y \iff f_1(x_1) + f_2(x_2) \geq f_1(y_1) + f_2(y_2). \tag{2.42}$$

Uniqueness of this representation is well-known - a different pair of functions g_i satisfying (2.42) is related to f_i as follows:

$$g_i(x_i) = \alpha f_i(x_i) + \beta_i. \tag{2.43}$$

The representation (2.42) can be recast as a Choquet integral representation (2.9). In order to do this, we can substitute f_1 by $\alpha f_1'$ and f_2 by $(1-\alpha)f_2'$, where

$\alpha \in (0,1)$. Apparently α can be chosen arbitrarily. Next, we can construct a capacity that "embeds" the weights α and $(1-\alpha)$. If β_i are chosen so that ranges of f_1' and f_2' intersect, the capacity must be additive (so that the weights are equal on both comonotonic cones $f_1' \leq f_2'$ and $f_2' \leq f_1'$), however if $f_1'(x_1) \geq f_2'(x_2)$ for all $x_1x_2 \in X$, we can choose a non-additive capacity as well. It's easy to see this using the Möbius form of the integral. Assume $f_1' \leq f_2'$ for all points from X. Then

$$m_1 f_1'(x_1) + m_2 f_2'(x_2) + m_{12} f_1'(x_2) \wedge f_2'(x_2) = (m_1 + m_{12}) f_1'(x_1) + m_2 f_2'(x_2), \quad (2.44)$$

and the mass $(m_1 + m_{12})$ can be arbitrarily distributed between m_1 and m_{12}.

2.12 Necessity of the axioms

A3. Necessity of **A3** is immediate (in the representation one of the regions $\mathbf{NW}^z$ and $\mathbf{SE}^z$ is necessarily contained in a comonotonic subset of $\mathbb{R}^2$).

A4. Let $ap \preccurlyeq bq, ar \succcurlyeq bs, cp \succcurlyeq dq$ and assume $cr \prec ds$. Let also $ap, bq, cp, dq \in \mathbf{NW}$, $ar, bs, cr, ds \in \mathbf{SE}$ and X_1 to be symmetric in $\mathbf{NW}$ (the other cases are symmetric). We obtain:

$$\begin{aligned}
\alpha_1 f_1(a) + \alpha_2 f_2(p) &\leq \alpha_1 f_1(b) + \alpha_2 f_2(q) \\
\alpha_1 f_1(c) + \alpha_2 f_2(p) &\geq \alpha_1 f_1(d) + \alpha_2 f_2(q) \\
\beta_1 f_1(a) + \beta_2 f_2(r) &\geq \beta_1 f_1(b) + \beta_2 f_2(s) \\
\beta_1 f_1(c) + \beta_2 f_2(r) &< \beta_1 f_1(d) + \beta_2 f_2(s)
\end{aligned} \quad (2.45)$$

From the first two inequalities and essentiality of $i = 1$ ($\alpha_1 \neq 0$) follows $f_1(a) + f_1(d) \leq f_1(b) + f_1(c)$. Last two inequalities imply $f_1(a) + f_1(d) > f_1(b) + f_1(c)$, a contradiction.

We also give the "necessity" proof of the condition in Lemma 7, since comparing it with the necessity proof of **A5** allows to elicit some interesting implications of essentiality.

Lemma 7 Let $\{g_1^{(i)} : g_1^{(i)} y_0 \sim g_1^{(i+1)} y_1, g_1^{(i)} \in X_1, i \in N\}$ and $\{h_2^{(i)} : x_0 h_2^{(i)} \sim x_1 h_2^{(i+1)}, h_2^{(i)} \in X_x, i \in N\}$ be two standard sequences, the first entirely contained in **NW** and the second in **SE**. Assume also, that there exist $z_1, z_2 \in X$, $p, q \in X_2, a, b \in X_1$ such that $g_1^{(i)} p, g_1^{(i)} q \in$ **NW**, and $ah_2^{(i)}, bh_2^{(i)} \in$ **SE** for all i, and $g_1^{(i)} p \sim bh_2^{(i)}$ and $g_1^{(i+1)} p \sim bh_2^{(i+1)}$. Finally, assume, $g_1^{(i+2)} p \succ bh_2^{(i+2)}$. Other cases are symmetric.

$$\alpha_1 f_1(g_1^{(i)}) + \alpha_2 f_2(y_0) = \alpha_1 f_1(g_1^{(i+1)}) + \alpha_2 f_2(y_1)$$
$$\alpha_1 f_1(g_1^{(i)}) + \alpha_2 f_2(y_0) = \alpha_1 f_1(g_1^{(i+1)}) + \alpha_2 f_2(y_1)$$

$$\begin{aligned} \beta_1 f_1(x_0) + \beta_2 f_2(h_1^{(i)}) &= \beta_1 f_1(x_1) + \beta_2 f_2(h_1^{(i+1)}) \\ \beta_1 f_1(x_0) + \beta_2 f_2(h_1^{(i+1)}) &= \beta_1 f_1(x_1) + \beta_2 f_2 f_2 f_2(h_1^{(i+2)}) \end{aligned} \tag{2.46}$$

$$\alpha_1 f_1(g_1^{(i)}) + \alpha_2 f_2(p) = \beta_1 f_1(b) + \beta_2 f_2(h_1^{(i)})$$
$$\alpha_1 f_1(g_1^{(i+1)}) + \alpha_2 f_2(p) = \beta_1 f_1(b) + \beta_2 f_2(h_1^{(i+1)})$$

First two equations imply $\alpha_1(f_1(g_1^{(i)}) - f_1(g_1^{(i+1)})) = \alpha_1(f_1(g_1^{(i+1)}) - f_1(g_1^{(i+2)}))$. The following two imply $\beta_1(f_2(h_1^{(i)}) - f_2(h_1^{(i+1)})) = \beta_1(f_2(h_1^{(i+1)}) - f_2(h_1^{(i+2)}))$. Finally, the last two equations imply $\alpha_1(f_1(g_1^{(i)}) - f_1(g_1^{(i+1)})) = \beta_1(f_2(h_1^{(i)}) - f_2(h_1^{(i+1)}))$. Apparently $\alpha_1(f_1(g_1^{(i+1)}) - f_1(g_1^{(i+2)})) < \beta_1(f_2(h_1^{(i+1)}) - f_2(h_1^{(i+2)}))$ is then a contradiction.

If we were to add an essentiality condition to Lemma 7, the statement can be made stronger as shown below.

A5. Assume $ap \preccurlyeq bq, cp \succcurlyeq dq$ and $ay_0 \sim x_0\pi(a), by_0 \sim x_0\pi(b), cy_1 \sim x_1\pi(c), dy_1 \sim x_1\pi(d)$, and also $e\pi(a) \succcurlyeq g\pi(b)$. Also, $i = 1$ is essential on the set (**NW** or **SE**) which includes ap, bq, cp, dq, and $i = 2$ is essential on the set (**NW** or **SE**), which includes $x_0\pi(a)$ and $x_0\pi(b)$. Finally, assume $e\pi(c) \prec g\pi(d)$.

We get

$$\alpha_1 f_1(a) + \alpha_2 f_2(p) \le \alpha_1 f_1(b) + \alpha_2 f_2(q)$$
$$\alpha_1 f_1(c) + \alpha_2 f_2(p) \ge \alpha_1 f_1(d) + \alpha_2 f_2(q)$$

$$\beta_1 f_1(e) + \beta_2 f_2(\pi(a)) \ge \beta_1 f_1(g) + \beta_2 f_2(\pi(b))$$
$$\beta_1 f_1(e) + \beta_2 f_2(\pi(b)) < \beta_1 f_1(g) + \beta_2 f_2(\pi(d))$$
$$\tag{2.47}$$
$$\gamma_1 f_1(a) + \gamma_2 f_2(y_0) = \delta_1 f_1(x_0) + \delta_2 f_2(\pi(a))$$
$$\gamma_1 f_1(b) + \gamma_2 f_2(y_0) = \delta_1 f_1(x_0) + \delta_2 f_2(\pi(b))$$

$$\gamma_1 f_1(c) + \gamma_2 f_2(y_1) = \delta_1 f_1(x_1) + \delta_2 f_2(\pi(c))$$
$$\gamma_1 f_1(d) + \gamma_2 f_2(y_1) = \delta_1 f_1(x_1) + \delta_2 f_2(\pi(d))$$

First two inequalities and the essentiality of $i = 1$ ($\alpha_1 \neq 0$) imply $f_1(a) - f_1(b) \le f_1(c) - f_1(d)$. Second pair of inequalities yields $f_2(\pi(c)) - f_2(\pi(d)) < f_2(\pi(a)) - f_2(\pi(b))$, while the final pair of equations leads to $\gamma_1(f_1(c) - f_1(d)) = \delta_2(f_2(\pi(c)) - f_2(\pi(d)))$. Combining these results and due to essentiality of $i = 2$ (hence $\delta_2 \neq 0$) we get:

$$\gamma_1(f_1(a)-f_1(b)) \le \gamma_1(f_1(c)-f_1(d)) = \delta_2(f_2(\pi(c))-f_2(\pi(d))) < \delta_2(f_2(\pi(a))-f_2(\pi(b))), \tag{2.48}$$

which contradicts the third pair of inequalities above, which yield $\gamma_1(f_1(a) - f_1(b)) = \delta_2(f_2(\pi(a)) - f_2(\pi(b)))$.

A6. For the Choquet integral representation on a heterogeneous product set $X = X_1 \times X_2$, strong monotonicity is actually a necessary condition because of the following. Assume $ap, bp, cp, dp \in \mathbf{SE}$ and $ap \succ bp, cp \sim dp$. Assume also there exist $cq, dq \in \mathbf{NW}$ such that $cq \succ dq$. Then, provided the representation exists, we get

$$\begin{aligned} \alpha_1 f_1(a) + \alpha_2 f_2(p) &> \alpha_1 f_1(b) + \alpha_2 f_2(p) \\ \alpha_1 f_1(c) + \alpha_2 f_2(p) &= \alpha_1 f_1(d) + \alpha_2 f_2(p) \\ \beta_1 f_1(c) + \beta_2 f_2(q) &> \beta_1 f_1(d) + \beta_2 f_2(q). \end{aligned} \tag{2.49}$$

The first inequality entails $\alpha_1 \neq 0$. From this and the following equality follows $f_1(c) = f_1(d)$, which contradicts with the last inequality. Thus $cq \succ dq$ implies $cp \succ dp$ but only in the presence of $ap \succ bp$ in the same "region" (**SE** or **NW**).

2.13 Summary

We have presented the axiomatization of the Choquet integral for two-dimensional heterogeneous product sets. Our axiom system does not rely on the notions of comonotonicity or constant acts which are not meaningful in the heterogeneous setting. The novel condition is **A3**, which states, roughly speaking, that we should be able to create an additive representation on at least one of the two cones $\mathbf{SE}^z$ and $\mathbf{NW}^z$ built at any point $z \in X$. This axiom is in fact a generalized version of comonotonic additivity, as is explained in the next chapter. The results we have presented include ordinal, cardinal and mixed cases. We constructed the representation and studied its uniqueness properties, which are somewhat weaker than in the homogeneous case. These results prepare us for the n-dimensional case presented in the following chapter.

Chapter 3

Axiomatization of the Choquet integral – n-dimensional case

3.1 Introduction

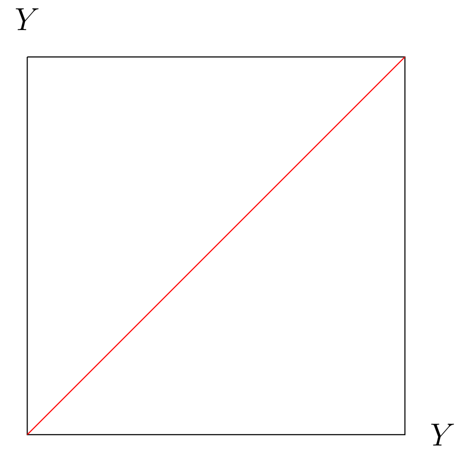
This chapter contains a full n-dimensional characterization of the Choquet integral. There are several major differences from the two-dimensional case. First, we will have to introduce an additional condition, acyclicity, which was not required for two dimensions. Second, the proof flow will have to be changed significantly due to a much larger number of possible "geometric" layouts of the set X (see details in Section 3.7). In fact, the present case introduces one more dimension of generality as compared to the previous results for the homogeneous sets. Not only is the shape and location of the border between additive areas not certain, but so too is the way that these borders intersect each other. The difference between "comonotonicity" and our setup is presented in Figure 3.1. Homogeneous case is typically symmetrical, with $n!$ comonotonic cones, which intersect along the main diagonal of the hypercube. In contrast, in our case the symmetry is gone, and in $n \geq 3$ dimensions, we also can have a much wider range of geometrical layouts. Not necessarily all $n!$ cones are present, even less so must all of them intersect together. This introduces significant complications when constructing the global representation and removes us further from the proof methods used in previous works.

3.2 Axioms and definitions

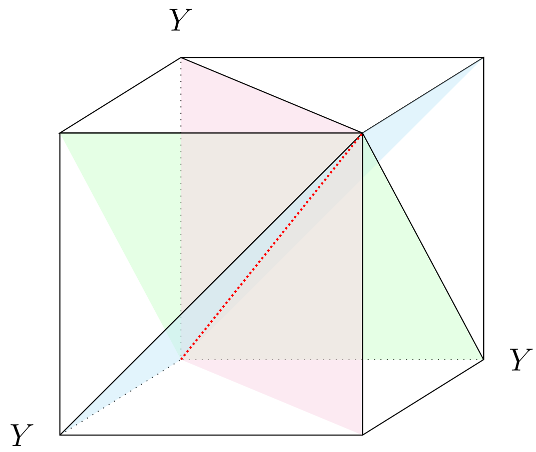
Let $\succcurlyeq$ be a binary relation on the set $X = X_1 \times \ldots \times X_n$. $\succ, \prec, \preccurlyeq, \sim, \not\sim$ are defined in the usual way. In MCDA, elements of set X are interpreted as alternatives characterized by criteria from the set $N = \{1, \ldots, n\}$. Set X_i contains criteria values for criterion i. We say that $\succcurlyeq$ can be represented by a Choquet integral, if there exists a capacity ν and functions $f_i : X_i \to \mathbb{R}$, called value functions, such that:

$$x \succcurlyeq y \iff C(\nu, (f_1(x_1), \ldots, f_n(x_n)) \geq C(\nu, (f_1(y_1), \ldots, f_n(y_n)). \tag{3.1}$$

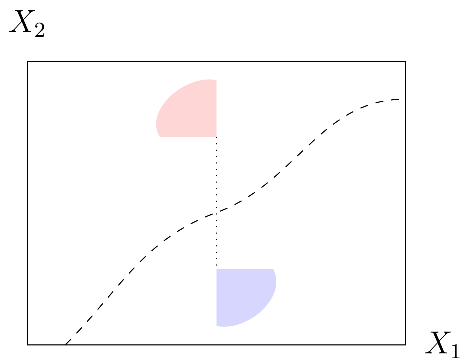
Definition 15. *Given $i, j \in N$, a relation $\succcurlyeq$ on $X_1 \times \ldots \times X_n$ satisfies ij-*triple cancellation (**ij-3C**)*, if for all $a_i, b_i, c_i, d_i \in X_i$, $p_j, q_j, r_j, s_j \in X_j$, and all*



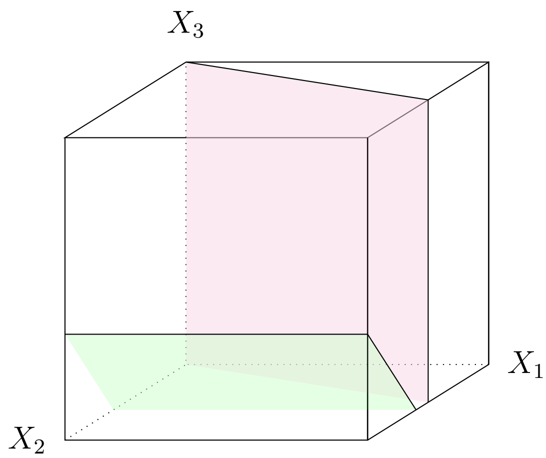
(a) Homogeneous case - 2 dimensions



(b) Homogeneous case - 3 dimensions



(c) Heterogeneous case - 2 dimensions



(d) Heterogeneous case - 3 dimensions

Figure 3.1: Generalization of comonotonic partitioning

$z_{-ij} \in X_{-ij}$ *holds:*

$$\left.\begin{aligned} a_i p_j z_{-ij} &\preccurlyeq b_i q_j z_{-ij} \\ a_i r_j z_{-ij} &\succcurlyeq b_i s_j z_{-ij} \\ c_i p_j z_{-ij} &\succcurlyeq d_i q_j z_{-ij} \end{aligned}\right\} \Rightarrow c_i r_j z_{-ij} \succcurlyeq d_i s_j z_{-ij}. \tag{3.2}$$

A1 - Weak order. $\succcurlyeq$ is a weak order.

A2 - Weakest separability. For all i, if $a_ix_{-i} \succ b_ix_{-i}$ for some $a_i, b_i \in X_i$,$x_{-i} \in X_{-i}$, then $a_iy_{-i} \succcurlyeq b_iy_{-i}$ for all $y_{-i} \in X_{-i}$.

Note, that from this follows, that for any $a_i, b_i \in X_i$ either $a_ix_{-i} \succcurlyeq b_ix_{-i}$ or $b_ix_{-i} \succcurlyeq a_ix_{-i}$ for all $x_{-i} \in X_{-i}$. This allows to introduce the following definition:

Definition 16. *For all $a_i, b_i \in X_i$ define $\succcurlyeq_i$ as $a_i \succcurlyeq_i b_i \iff a_ix_{-i} \succcurlyeq b_ix_{-i}$ for all $x_{-i} \in X_{-i}$.*

Definition 17. *For any $z \in X$ define $\mathbf{SE}^z_{ij} = \{x_ix_jz_{-ij} \in X \colon x_i \succcurlyeq_i z_i, z_j \succcurlyeq_j x_j\}$, and $\mathbf{NW}^z_{ij} = \{x_ix_jz_{-ij} \in X \colon z_i \succcurlyeq_i x_i, x_j \succcurlyeq_j z_j\}$.*

A3 - Coordinate Ordering Completeness. For any $z \in X$, and all $i, j \in N$, ij-triple cancellation holds either on $\mathbf{SE}^z_{ij}$ or on $\mathbf{NW}^z_{ij}$.

This new property would allow us to divide X into subsets without the need to use the notion of comonotonicity. We can introduce the following binary relations:

Definition 18. *We write:*

1. *$i \, \mathbf{R}^z \, j$ if ij-triple cancellation holds on the set $\mathbf{SE}^z_{ij}$.*

2. *$i \, \mathbf{S}^z \, j$ if [NOT $j \, \mathbf{R}^z \, i$].*

3. *$i \, \mathbf{E}^z \, j$ if [$i \, \mathbf{R}^z \, j$ AND $j \, \mathbf{R}^z \, i$].*

Note that $\mathbf{R}^z$ is complete (which is why we have called axiom **A3** "Coordinate Ordering Completeness") and $\mathbf{S}^z$ is partial.[2] Since N is finite, there is only a finite number of various partial orders $\mathbf{S}^z$, so we can index them $(\mathbf{S}_a, \mathbf{S}_b, \ldots)$ and drop the superscripts when not needed. Also, each of the partial orders $\mathbf{S}_k$ uniquely defines the corresponding $\mathbf{R}_k$ - $i \, \mathbf{R}_k \, j$ if [NOT $j \, \mathbf{S}_k \, i$].

In contrast to the case with two variables, this property alone is not sufficient to construct a representation. Comparing value functions for different

[2] If the relation is empty for all z, other axioms entail the existence of an additive representation on X.

attributes suggests some sort of transitivity. For example, $f_i(x_i) > f_j(x_j)$ and $f_j(x_j) > f_k(x_k)$ imply $f_i(x_i) > f_k(x_k)$. The property we introduce is weaker - it is acyclicity. There are two conditions required actually - cardinal acyclicity and ordinal acyclicity. These axioms effectively define how the set X is partitioned.

A3-OA - Ordinal Acyclicity. For all $z \in X$, $\mathbf{S}^z$ is acyclic. In other words,

$$i\,\mathbf{S}^z\,j\,\mathbf{S}^z \ldots \mathbf{S}^z\,k \Rightarrow i\,\mathbf{R}^z\,k. \tag{3.3}$$

The *cardinal acyclicity* condition bears some similarity to the condition necessary for existing of a potential game (Monderer and Shapley, 1996). We call it cardinal acyclicity, as it prohibits an existence of "positive improvement cycles". The condition is a generalization of **A5** from the two-dimensional case, so we just give it as **A5** below.

We also introduce the following notions:

Definition 19. *Define* $\mathbf{SE}_{ij}$ *as a union of the following three sets:*

- *All* $z \in X$ *such that* $i\,\mathbf{R}^z\,j$*, if* z_i *is not maximal and* z_j *is not minimal;*
- *All* $z \in X$ *such that* z_i *is maximal and for no* $x_j, y_j \in X_j : z_j \succsim_j x_j \succsim_j y_j$ *we have* $j\,\mathbf{R}^{x_j z_{-j}}\,i$ *and NOT* $j\,\mathbf{R}^{y_j z_{-j}}\,i$*;*
- *All* $z \in X$ *such that* z_j *is minimal and for no* $x_i, y_i \in X_i : y_i \succsim_i x_i \succsim_i z_i$ *we have* $j\,\mathbf{R}^{x_i z_{-i}}\,i$ *and NOT* $j\,\mathbf{R}^{y_i z_{-i}}\,i$*.*

Define $\mathbf{NW}_{ij}$ *as a union of the following three sets:*

- *All* $z \in X$ *such that* $j\,\mathbf{R}^z\,i$*, if* z_j *is not maximal and* z_i *is not minimal;*
- *All* $z \in X$ *such that* z_i *is minimal and for no* $x_j, y_j \in X_j : y_j \succsim_j x_j \succsim_j z_j$ *we have* $i\,\mathbf{R}^{x_j z_{-j}}\,j$ *and NOT* $i\,\mathbf{R}^{y_j z_{-j}}\,j$*;*
- *All* $z \in X$ *such that* z_j *is maximal and for no* $x_i, y_i \in X_i : z_i \succsim_i x_i \succsim_i y_i$ *we have* $i\,\mathbf{R}^{x_i z_{-i}}\,j$ *and NOT* $i\,\mathbf{R}^{y_i z_{-i}}\,j$*.*

Presence of maximal and minimal points significantly complicates the definitions of $\mathbf{SE}_{ij}$ and $\mathbf{NW}_{ij}$, since at such points some of the sets $\mathbf{SE}_{ij}^z$ and $\mathbf{NW}_{ij}^z$ become degenerate and condition **3C-ij** trivially holds. If sets X_i and X_j do not contain minimal or maximal points, we can drop the corresponding conditions in each definition and simply state that $\mathbf{SE}_{ij} = \{z : i\,\mathbf{R}^z\, j\}$ and $\mathbf{NW}_{ij} = \{z : j\,\mathbf{R}^z\, i\}$.

Partial orders $\mathbf{S}_i$ define subsets of the set X as follows.

Definition 20. *We write* $X^{\mathbf{S}_i} = \bigcap_{(k,j):k\,\mathbf{R}_i\, j} \mathbf{SE}_{kj}$

It is well known that the sufficient property for an additive representation to exist on a Cartesian product is strong independence Krantz et al. (1971). In the $X = Y^n$ case, the Choquet integral was previously axiomatized using comonotonic strong independence (or comonotonic trade-off consistency Wakker (1991a)). In this paper we will be using sets $X^{\mathbf{S}_i}$ to formulate a similar condition.

Definition 21. *We say that* $i \in N$ *is essential on* $A \subset X$ *if there exist* $x_ix_{-i}, y_ix_{-i} \in A$*, such that* $x_ix_{-i} \succ y_ix_{-i}$.

A4 - Intra-coordinate trade-off consistency

$$\left.\begin{aligned} a_ix_{-i} &\preccurlyeq b_iy_{-i} \\ a_iw_{-i} &\succcurlyeq b_iz_{-i} \\ c_ix_{-i} &\succcurlyeq d_iy_{-i} \end{aligned}\right\} \Rightarrow c_iw_{-i} \succcurlyeq d_iz_{-i}, \tag{3.4}$$

provided that either:

a) Exists $X^{\mathbf{S}_j}$ such that $a_ix_{-i}, b_iy_{-i}, a_iw_{-i}, b_iz_{-i}, c_ix_{-i}, d_iy_{-i}, c_iw_{-i}, d_iz_{-i} \in X^{\mathbf{S}_j}$

b) Exist $X^{\mathbf{S}_j}, X^{\mathbf{S}_k}$ such that $a_ix_{-i}, b_iy_{-i}, a_iw_{-i}, b_iz_{-i} \in X^{\mathbf{S}_j}$, i is essential on $X^{\mathbf{S}_j}$, and $c_ix_{-i}, d_iy_{-i}, c_iw_{-i}, d_iz_{-i} \in X^{\mathbf{S}_k}$, or;

c) Exist $X^{\mathbf{S}_j}, X^{\mathbf{S}_k}$ such that $a_ix_{-i}, b_iy_{-i}, c_ix_{-i}, d_iy_{-i} \in X^{\mathbf{S}_j}$, i is essential on $X^{\mathbf{S}_j}$, and $a_iw_{-i}, b_iz_{-i}, c_iw_{-i}, d_iz_{-i} \in X^{\mathbf{S}_k}$.

Informally, the meaning of the axiom is that ordering between preference differences ("intervals") is preserved irrespective of the "measuring rods" used to measure them. However, contrary to the additive case this does not hold on all X, but only when either points involved in all four relations lie in the same "3C-set" X^{S_j}, or points involved in two relations lie in one such set and those involved in the other two in another.

A5 - Inter-coordinate trade-off consistency Let $i, j, k, \ldots, m$ be a sequence made of K coordinates and let

$$\left.\begin{array}{r}
a_i e^0_{-i} \sim b_i f^0_{-i} \\
a_i x^0_{-i} \sim p_j y^0_{-j} \\
b_i x^0_{-i} \sim q_j y^0_{-j} \\
\\
p_j e^1_{-j} \sim q_j f^1_{-j} \\
r_j e^1_{-j} \sim s_j f^1_{-j} \\
r_j x^1_{-j} \sim g_k y^1_{-k} \\
s_j x^1_{-i} \sim h_k y^1_{-k} \\
\\
g_k e^2_{-k} \sim h_k f^2_{-k} \\
\ldots \\
\\
t_m e^K_{-m} \sim u_m f^K_{-m} \\
w_m e^K_{-m} \sim z_m f^K_{-m} \\
w_m x^K_{-m} \sim c_i y^K_{-i} \\
z_m x^K_{-m} \sim d_i y^K_{-i}
\end{array}\right\} \Rightarrow c_i e^0_{-i} \succcurlyeq d_i f^0_{-i} \tag{3.5}$$

for all points in X provided all points containing a_i, b_i, c_i, d_i reside in the same $X^{\mathbf{S}_v}$ such that i is essential on $X^{\mathbf{S}_v}$, all points containing p_j, q_j, r_j, s_j reside in the same $X^{\mathbf{S}_y}$ such that j is essential on $X^{\mathbf{S}_y}$, etc.

The formal statement of the **A5** is rather complicated, but it simply means that the ordering of the "intervals" is preserved across dimensions. This version is an extension of the two-dimensional version and provides also the necessary "cardinal acyclicity" property. Together with **A4** the conditions are similar to Wakker's trade-off consistency condition Wakker (1991b). The axiom bears even stronger similarity to Axiom 5 (compatibility) from section 8.2.6 of Krantz et al. (1971). Roughly speaking, it says that if the "interval" between c_i and d_i is "larger" than that between a_i and b_i, then "projecting" these intervals onto another dimension by means of the equivalence relations must leave this order unchanged. We additionally require the comparison of intervals and "projection" to be consistent - meaning that each quadruple of points in each part of the statement belongs to the same $X^{\mathbf{S}_i}$. Another version of this axiom, which is used frequently in proofs, can be formulated in terms of standard sequences (Lemma 39).

A6 - Strong monotonicity Let $a_i x_{-i}, b_i x_{-i}, c_i x_{-i}, d_i x_{-i} \in X^{\mathbf{S}_i}$ and $a_i x_{-i} \succ b_i x_{-i}$. If for some $y_{-i} \in X_{-i}$ we have $c_i y_{-i} \succ d_i y_{-i}$, then $c_i x_{-i} \succ d_i x_{-i}$ for all $i \in N$.

This axiom is similar to "strong monotonicity" in Wakker (1991b). We analyze its necessity and the intuition behind it in section 3.13.

A7 - Essentiality All coordinates are essential on X.

A8 - Restricted solvability If $a_i x_{-i} \succcurlyeq y \succcurlyeq b_i x_{-i}$, then there exists $c : c_i x_{-i} \sim y$ for $i \in N$.

A9 - Archimedean axiom Every bounded standard sequence contained in some $X^{\mathbf{S}_i}$ is finite, and in the case of only one essential coordinate, there exists a countable order-dense subset of $X^{\mathbf{S}_i}$.

Finally, we can introduce a notion of *interacting* coordinates.

Definition 22. *Coordinates i and j are* interacting *if exists $z \in X$, such that $i\,\mathbf{S}^z\,j$ or $j\,\mathbf{S}^z\,i$. We call a set $A \subset N$ an* interaction group *if for each $i, j \in A$ we can build a chain of coordinates $i, k, \ldots, j$, such that every two subsequent coordinates in the chain are interacting.*

Interaction groups play an important role in the uniqueness properties of the representation. In what follows we will be considering only groups of maximal possible size if not specified otherwise.

3.2.1 Additional assumptions

The following additional assumptions are made. The reasoning behind each one is explained below. As in the two-dimensional case we are strongly convinced that the construction of the representation can be done without these, although with some technical complications.

"Collapsed" equivalent points along dimensions. For no $i \in N$ and no $a_i, b_i \in X_i$ holds $a_i x_{-i} \sim b_i x_{-i}$ for all $x_{-i} \in X_{-i}$.

If this wasn't true, we could have value functions assigning the same value to several points in the same set X_i. To simplify things we exclude such case, however, it can be easily reconstructed once the representation is built.

Density. We assume that for all $i \in N$, whenever $a_i x_{-i} \succ b_i x_{-i}$, there exists $c_i \in X_i$ such that $a_i x_{-i} \succ c_i x_{-i} \succ b_i x_{-i}$ (X is order dense).

"Closedness". For every i and j, if there exist $x_i x_j z_{-ij}$ such that $i\, \mathbf{S}^{x_i x_j z_{-ij}}\, j$ and $y_i x_j z_{-ij}$ such that $j\, \mathbf{S}^{y_i x_j z_{-ij}}\, i$, then exists $z_i \in X_i$ such that $i\, \mathbf{E}^{z_i x_j z_{-ij}}\, j$.

This assumption says that sets $\mathbf{SE}_{ij}$ and $\mathbf{NW}_{ij}$ are "closed". In the representation this translates into existence of the inverse for all points where value functions f_i and f_j are equal, provided i and j are interacting. This is a technical simplifying assumption and the proof can be done without it.

3.3 Representation theorem

As follows from the definition of the Choquet integral (Section 2.1.1), every point $x \in X$ uniquely corresponds to a set of weights $p_i^x : p_i^x \geq 0, \sum_{i \in N} p_i^x = 1$. This notation is used to simplify the statement of the following theorems.

Theorem 8. *Let $\succcurlyeq$ be a binary relation on X and the structural assumptions hold. Then, if axioms **A1-A9** are satisfied, there exists a capacity ν and value functions $f_1 : X_1 \to \mathbb{R}, \ldots, f_n : X_n \to \mathbb{R}$, such that $\succcurlyeq$ can be represented by the Choquet integral:*

$$x \succcurlyeq y \iff C(\nu, (f_1(x_1), \ldots, f_n(x_n))) \geq C(\nu, (f_1(y_1), \ldots, f_n(y_n))), \tag{3.6}$$

for all $x, y \in X$.

Capacity and value functions have the following uniqueness properties.

Theorem 9. *Let* $\mathcal{I} = \{A_1, \ldots, A_k\}$ *be a partition of* N *into interaction groups* A_i. *Then, exist unique capacities* $\nu_{A_1}, \ldots, \nu_{A_k}$ *and value functions* $f_1 : X_1 \to \mathbb{R}, \ldots, f_n : X_n \to \mathbb{R}$, *such that* $\succcurlyeq$ *can be represented by a sum of the Choquet integrals:*

$$x \succcurlyeq y \iff \sum_{i=1}^{k} C(\nu_{A_i}, f_{A_i}(x_{A_i})) \geq \sum_{i=1}^{k} C(\nu_{A_i}, f_{A_i}(y_{A_i})), \tag{3.7}$$

for all $x, y \in X$, *where* $f_{A_i}(x_{A_i})$ *is a shortcut for* $(f_{i_1}(x_{i_1}), \ldots, f_{i_t}(x_{i_t}))$, *where* $\{i_1, \ldots, i_t\} = A_i$, *in other words a vector including value functions of coordinates from the set* A_i.

It is easier to characterize the uniqueness properties of the representation (3.7), as in the case of the model (3.6) value functions in each interaction group can be re-scaled independently, which is "compensated" by re-scaling the corresponding capacity elements and re-normalizing (see example below). On contrary, the representation (3.7) gives us a very clean uniqueness results due to the fact that all capacities ν_{A_i} are normalized. Note that contrary to the representation (3.6), capacities ν_{A_k} are *unique.*

To simplify the uniqueness statement, we separate two cases - the fully cardinal one, where each X^{S_i} has at least two essential coordinates, and the fully ordinal one, where only one coordinate is essential on each X^{S_i}. These kind of assumptions are customary in the literature (e.g. Wakker (1991b) assumes all variables to be essential).

Add analysis of the general case

Theorem 10. *Let each* X^{S_a} *has at least two essential coordinates (cardinal case). Let* $g_1 : X_1 \to \mathbb{R}, \ldots, g_n : X_n \to \mathbb{R}$ *be such that* (3.7) *holds with* f_i *substituted by* g_i. *Let* $i \in A_m$. *Then,*

$$g_i(x_i) = \alpha f_i(x_i) + \beta_{A_m}. \tag{3.8}$$

Theorem 11. *Let each* X^{S_a} *has only one essential coordinate (ordinal case). Let* $g_1 : X_1 \to \mathbb{R}, \ldots, g_n : X_n \to \mathbb{R}$ *be such that* (3.7) *holds with* f_i *substituted by* g_i.

Then, ν in representation (3.6) *is unique, all variables are in the same interaction group, and*

$$g_i(x_i) = \psi_i f_i(x_i), \tag{3.9}$$

where ψ_i is an increasing function, and for all $i, j \in N$ we have $f_i(x_i) \geq f_j(x_j) \iff g_i(x_i) \geq g_j(x_j)$.

The extreme cases of the representation (3.7) are the case where all dimensions are in the same interaction group, in which case we get the representation (3.6) with a unique capacity, and the case where all interaction groups are singletons, i.e. there is no interaction, and we get an additive value model (capacities ν_{A_i} are defined on $\{\emptyset, i\}$).

The uniqueness properties of the original representation (3.6) are much weaker. As shown above, it is strongly conditional on the absence of additivity throughout X. This is so because we can change the origins of the value functions in each interaction group independent of those in other interaction groups. This allows to "move" the representation of X between various comonotonic cones, hence allowing for the use of different capacities. Some examples of non-uniqueness are given in Section 3.9.

.

3.4 Proof preview

The proof contains five main stages. First, we construct additive representations on the sets X^{S_i} with extreme points removed. Next, we show that value functions on the same dimensions are proportional across different X^{S_i}, so we can rewrite the additive representations using the same value functions, but with different weights, for all X^{S_i}. The third step is to show that we can rescale the value functions so that for all points z, where $i \, \mathbf{E}^z \, j$, levels from both dimensions i and j get the same value. Finally, we show that the additive representations can be unified in a single global Choquet integral representation, including also the extreme points of X. The only thing left after this is to analyse the uniqueness of the constructed model.

Construction of the additive representations within X^{S_i} is done in Section

3.5, proportionality is shown in 3.6, equality of value functions for points where $i\,\mathbf{E}\,j$, construction of the global representation, and construction of the Choquet integral are presented in Section 3.7. Extension to the extreme points is done in 3.8. Finally, Section 3.9 contains the analysis of the uniqueness properties. Additional technical results are presented in Sections 3.12 - 3.14.

3.5 Additive representations on X^{S_a}

We start by removing maximal and minimal elements from the sets X_i. The representation will be extended to these points in Section 3.8.

Similar to (Wakker, 1991b) we will be covering the sets X^{S_a} with "rectangular" subsets. Given a point $z \in X^{S_a}$ we construct a "rectangular" set $X^{z(S_a)}$ in the following way:

- If j is minimal in $\mathbf{S}_a$, then $X_j^{z(S_a)} = x_j \in X_j : z_j \succcurlyeq_j x_j$.
- If j is maximal in $\mathbf{S}_a$, then $X_j^{z(S_a)} = x_j \in X_j : x_j \succcurlyeq_j z_j$.
- If j is neither maximal not minimal, then $X_j^{z(S_a)} = [x_j \in X_j : x_j \succcurlyeq_j z_j, x_j z_{-j} \in X^{S_a}]$.
- If for no k we have $j\,\mathbf{S}_a\,k$ or $k\,\mathbf{S}_a\,j$, then $X_j^{z(S_a)} = X_j$.

3.5.1 Constructing additive representation on $X^{z(S_a)}$

We assume that X^{S_a} has at least two essential coordinates. By Lemma 42, all sets $X^{z(S_a)}$ therefore have at least two essential coordinates. Moreover, the essential coordinates are the same across all sets.

Theorem 12. *For any $z \in X^{S_a}$ there exists an additive representation of $\succcurlyeq$ on $X^{z(S_a)}$:*

$$x \succcurlyeq y \Leftrightarrow \sum_{i=1}^{n} V_i^z(x_i) \geq \sum_{i=1}^{n} V_i^z(y_i), \tag{3.10}$$

for all $x, y \in X^{z(S_a)}$.

Proof. $X^{z(S_a)}$ is a Cartesian product, $\succcurlyeq$ is a weak order on $X^{z(S_a)}$, $\succcurlyeq$ satisfies generalized triple cancellation on $X^{z(S_a)}$, $\succcurlyeq$ satisfies Archimedean axiom on $X^{z(S_a)}$, at least two coordinates are essential. It remains to show that $\succcurlyeq$ satisfies restricted solvability on $X^{z(S_a)}$.

Assume that for some $x_i z_{-i}, w, y_i z_{-i} \in X^{z(S_a)}$, we have $x_i z_{-i} \succcurlyeq w \succcurlyeq y_i z_{-i}$, hence exists $z_i \in X_i : z_i z_{-i} \sim w$. We need to show that $z_i z_{-i} \in X^{S_a}$. If $w \sim x_i z_{-i}$ or $w \sim y_i z_{-i}$, then the conclusion is immediate (since either point belongs to $X^{z(S_a)}$). Hence, assume $x_i z_{-i} \succ z_i z_{-i} \succ y_i z_{-i}$. This means that $x_i \succcurlyeq_i z_i \succcurlyeq_i y_i$. Since z_i is "sandwiched" between x_i and y_i we conclude that for any $j \in N \setminus i$, $i \, \mathbf{S}^{x_i z_{-i}} \, j$ and $i \, \mathbf{S}^{y_i z_{-i}} \, j$ imply also $i \, \mathbf{S}^{z_i z_{-i}} \, j$, and symmetrically $j \, \mathbf{S}^{x_i z_{-i}} \, i$ and $j \, \mathbf{S}^{y_i z_{-i}} \, i$ imply $j \, \mathbf{S}^{z_i z_{-i}} \, i$. Hence, it is also in X^{S_a}.

Therefore all conditions for the existence of an additive representation are met (Wakker, 1991a). □

3.5.2 Joint representation V^{S_a} on X^{S_a}

This section is based on (Wakker, 1991b) with some modifications.

Theorem 13. *There exists an additive interval scale $V^{S_a}(z) = \sum_{i=1}^{n} V_i^{S_a}(z_i)$ on X^{S_a}, which represents $\succcurlyeq$ on every $X^{z(S_a)}$ with $z \in X^{S_a}$.*

Proof. Choose the reference set - pick any $r \in X^{S_a}$ such that $X_i^{r(S_a)}$ contains more than one point for any X^{S_a}-essential i. Choose a "zero" point - any $r^0 \in X^{r(S_a)}$, and a "unit mark" - a point $r_k^1 r_{-k}^0 \in X^{r(S_a)}$, such that:

- k is essential on X^{S_a},
- $r_k^1 \succ_k r_k^0$.

Set $V_i^r(r_i^0) = 0$ for all $i \in N$ and $V_k^r(r_k^1) = 1$. This uniquely defines unit and locations of all $V_i^r, i \in N$.

In the following we assume that sets $X_i^{z(S_a)}, X_k^{z(S_a)}$ each contain at least two points, otherwise, alignment is trivial.

Assume $X_i^{r(S_a)} \cap X_i^{z(S_a)} = \emptyset$ and $X_k^{r(S_a)} \cap X_k^{z(S_a)} = \emptyset$ (variations are all covered by the below procedure). We will construct two auxiliary points z' and r' such that $X_i^{z'(S_a)} \subset X_i^{z(S_a)}$, $X_i^{z'(S_a)} \subset X_i^{r(S_a)}$, $X_k^{r'(S_a)} \subset X_k^{z(S_a)}$, $X_k^{r'(S_a)} \subset X_k^{r(S_a)}$. It would allow us to align first V^r and $V^{r'}$, then $V^{r'}$ and $V^{z'}$, and finally $V^{z'}$ and V^z. See Figure 3.2 for an example: green rectangle is for z: $X_i^{z(S_a)} \times X_k^{z(S_a)}$, yellow is for z', red is for r', and blue is for r.

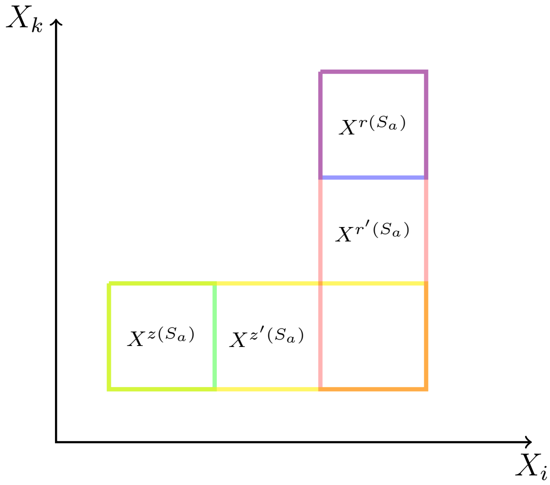


Figure 3.2: Aliging representations V^z and V^r

Construct the point z' by taking coordinate-wise maxima of r and z for coordinates j such that $j\,\mathbf{R}_a\,i$, not including i itself, and coordinate-wise minima of r and z for coordinates j, such that $i\,\mathbf{R}_a\,j$ and i itself. In the short notation the first point is $z' := \max(r_j, z_j)_{j:j\,\mathbf{R}_a\,i} \min(r_j, z_j)_{j=i,j:i\,\mathbf{R}_a\,j}$. The second point r' is constructed by taking coordinate-wise maxima of r and z for coordinates j such that $j\,\mathbf{R}_a\,k$, not including k itself, and coordinate-wise minima of r and z for coordinates j, such that $k\,\mathbf{R}_a\,j$ and k itself. In the short notation the second point looks like $r' := \max(r_j, z_j)_{j:j\,\mathbf{R}_a\,k} \min(r_j, z_j)_{j=k,j:k\,\mathbf{R}_a\,j}$.

Note that both points are in X^{S_a} since relations $j\,\mathbf{R}_a\,l$ remain intact for all pairs j, l. Note also, that $X_i^{z'(S_a)}$ contains both $X_i^{z(S_a)}$ and $X_i^{r(S_a)}$, and $X_k^{r'(S_a)}$ contains both $X_k^{z(S_a)}$ and $X_k^{r(S_a)}$.

Now we have that sets $(X_i^{z(S_a)} \times X_k^{z(S_a)}) \cap (X_i^{z'(S_a)} \times X_k^{z'(S_a)})$, $(X_i^{z'(S_a)} \times X_k^{z'(S_a)}) \cap (X_i^{r'(S_a)} \times X_k^{r'(S_a)})$, $(X_i^{r'(S_a)} \times X_k^{r'(S_a)}) \cap (X_i^{r(S_a)} \times X_k^{r(S_a)})$ are all non-empty, and each dimension contains more than two points. Relation $\succcurlyeq_{i,k}$ on these sets satisfies Archimedean axiom, restricted solvability, and **A4**. Hence we can apply standard uniqueness properties of additive representations. We first align $V_k^{r'}$ with V_k^r and $V_i^{r'}$ with V_i^r, then $V_k^{z'}$ with $V_k^{r'}$ and $V_i^{z'}$ with $V_i^{r'}$, and finally V_k^z with $V_k^{z'}$ and

V_i^z with $V_i^{z'}$ by changing the common unit and locations of corresponding value functions.

Having aligned like this V_i^z and V_k^z with V_i^r and V_k^r for all $z \in X^{S_a}$ we can perform the same alignment operation for all remaining essential coordinates j, using pairs V_j^z and V_k^z. At this stage, functions V_k^z are already aligned, hence have a correct unit and location. As above, uniqueness properties of additive representations of relation $\succsim_{j.k}$ imply that the unit of functions V_j^z is already aligned with that of V_j^r and only location change has to be performed. This can also be done as above.

Once such alignment has been performed for all essential coordinates, we can verify that this is done consistently throughout X^{S_a}. In particular, for any s and t from X^{S_a} we must be able to show that for any essential $j \in N$, we have $V_j^s = V_j^t$ on $X_j^{s(S_a)} \cap X_j^{t(S_a)}$. To show this a following argument can be used. During the initial alignment of V_j^s and V_j^t, auxiliary points t' and s' were used, such that $X_j^{s'(S_a)}$ includes $X_j^{s(S_a)}$ and $X_j^{r(S_a)}$, and $X_j^{t'(S_a)}$ includes $X_j^{t(S_a)}$ and $X_j^{r(S_a)}$. Hence, functions $V_j^{s'}$ and $V_j^{t'}$ coincide with V_j^r on $X_j^{r(S_a)}$. To show that they coincide on all common domain, including $X_j^{s(S_a)} \cap X_j^{t(S_a)}$, we just need to follow the same procedure as before and construct a point that contains $X_k^{s'(X_a)}$ and $X_k^{t'(X_a)}$ for some essential k. Then a uniqueness argument can be evoked once again, and since $V_j^{s'}$ and $V_j^{t'}$ coincide on $X_j^{r(S_a)}$, they would necessarily coincide also on the remaining common domain, which includes $X_j^{s(S_a)} \cap X_j^{t(S_a)}$. Finally, since $V_j^s = V_j^{s'}$ on $X_j^{s(S_a)}$, and $V_j^t = V_j^{t'}$ on on $X_j^{t(S_a)}$, we get that $V_j^s = V_j^t$ on $X_j^{s(S_a)} \cap X_j^{t(S_a)}$.

At this point we can drop the superscripts and define functions $V_i^{S_a}$ which coincide with $V_i^{z(S_a)}$ for all $z \in X^{S_a}$ on the corresponding domains. By the above argument, these functions are well-defined.

□

3.5.3 V^{S_a} is globally representing on X^{S_a}

Lemma 21. *For all X^{S_a}-essential $i \in N$, $V_i^{S_a}$ represents $\succsim_i$ on $X_i^{S_a}$.*

Proof. Let $\alpha_i, \beta_i \in X_i^{S_a}$ be such that $\alpha_i \succsim_i \beta_i$. Similarly to the construction of r' and z' in the proof of theorem 13, we can show that always exists x_i such that

$\alpha_i x_{-i}, \beta_i x_{-i} \in X^{S_a}$. The conclusion follows. $\square$

Theorem 14. *Representation V^{S_a} obtained in Theorem 13 is globally representing on X^{S_a}.*

Proof. We need to show that $x \succcurlyeq y \iff V^{S_a}(x) \geq V^{S_a}(y)$.

- If exists z such that $x, y \in X^{z(S_a)}$ then the result is immediate.
- If the above is not true, we will show that exists $x' \sim x$ such that $V^{S_a}(x) = V^{S_a}(x')$ and $x'_i \succcurlyeq_i y_i$ for all i.

The procedure is identical to Wakker (1991a) with some minor modifications.

1. Find i such that $y_i \succ_i x_i$ and $x_k \succcurlyeq_k y_k$ for all k such that $k\,\mathbf{S}_a\,i$. We have $y_i x_{-i} \in X^{S_a}$ (since for all $k \in N$ such that $k\,\mathbf{S}_a\,i$ we have $x_k \succcurlyeq y_k$, hence $k\,\mathbf{R}^{y_i y_{-i}}\,i$ implies $k\,\mathbf{R}^{y_i x_{-i}}\,i$, whereas for all $t \in N$ such that $i\,\mathbf{S}_a\,t$ we have $i\,\mathbf{R}^{x_i x_{-i}}\,t$, hence $i\,\mathbf{R}^{y_i x_{-i}}\,t$).
2. Similarly, find j such that $x_j \succ_j y_j$ and $y_k \succcurlyeq_k x_k$ for all k such that $j\,\mathbf{S}_a\,k$. By similar reasoning, $y_j x_{-j} \in X^{S_a}$.
3. We are increasing x_i and decreasing x_j and thus move in the direction of y.
4. Note, that $x_{-ij} y_i y_j \in X^{S_a}$.
5. If $x_{-ij} y_i y_j \succcurlyeq x$, then by restricted solvability ($x_{-ij} y_i y_j \succcurlyeq x \succcurlyeq x_{-ij} x_i y_j$) exists $x' := x_{-ij} x'_i y_j \sim x$, where $y_i \succcurlyeq_i x'_i \succcurlyeq_i x_i$. If $x \succ x_{-ij} y_i y_j$, then by restricted solvability ($x_{-ij} y_i x_j \succcurlyeq x \succ x_{-ij} y_i y_j$) exists $x' := x_{-ij} y_i x'_j \sim x$, and $x_j \succcurlyeq_j x'_j \succcurlyeq_J y_j$.
6. In both cases, the resulting point x' is in X^{S_a}, moreover $x', x \in X^{z(S_a)}$ where $z := x_{ij} x_i y_j$, hence x' has the same V^{S_a}-value as x, but one more coordinate becomes identical to that of y.
7. After repeating the procedure unless $x'_i \succcurlyeq_i y_i$ (at most n times), we get the result by Lemma 21.

8. Moreover, if $x \sim y$, we at the end of the procedure we would necessarily arrive to y itself (by strong monotonicity as in Lemma 42, and a structural assumption). Hence we get $x \succ y \Rightarrow V^{S_a}(x) > V^{S_a}(y)$ and $x \sim y \Rightarrow V^{S_a}(x) = V^{S_a}(y)$, which implies that $x \succcurlyeq y \iff V^{S_a}(x) \geq V^{S_a}(y)$

□

3.6 Aligning cardinal representations for different X^{S_a}

There can be several cases depending on what variables are essential on various sets X^{S_i}. We start with the case where exist X^{S_a} and X^{S_b} having at least two essential variables each.

3.6.1 Exist at least two sets X^{S_i} with at least two essential coordinates

Theorem 15. *Assume that at least two coordinates are essential on X^{S_a} and X^{S_b}. For any $i \in N$ that is essential on both areas, it holds $V_i^{S_a}(z_i) = \lambda_i^{ab} V_i^{S_b}(z_i)$ for all z_i from the common domain of $V_i^{S_a}(z_i)$ and $V_i^{S_b}(z_i)$, if a common location is chosen for both functions.*

Proof. If the common domain of $V_i^{S_a}(z_i)$ and $V_i^{S_b}(z_i)$ is empty or contains just one point, the result is trivial. Assume that i, j are essential on X^{S_a} and i, l are essential on X^{S_b}. First, we will establish that a standard sequence on coordinate i in X^{S_a} is also a standard sequence in X^{S_b} (provided all points of the sequence lie within a common domain of $V_i^{S_a}(z_i)$ and $V_i^{S_b}(z_i)$). This follows from **A4**. Build any standard sequence $X_i^{S_a}$, say $\{\alpha_i^k : \alpha_i^k v_j x_{-ij} \sim \alpha_i^{k+1} w_j x_{-ij}\}$. Then, $\{\alpha_i^k : \alpha_i^k t_l x_{-il} \sim \alpha_i^{k+1} u_l x_{-il}\}$ is a standard sequence in S_b, i.e. if exist $t_l, u_l \in X_l$ such that $\alpha_i^k t_l x_{-il} \sim \alpha_i^{k+1} u_l x_{-il}$ for some k, then by **A4**:

$$\left.\begin{aligned} \alpha_i^k v_j x_{-ij} &\sim \alpha_i^{k+1} w_j x_{-ij} \\ \alpha_i^k t_l x_{-il} &\sim \alpha_i^{k+1} u_l x_{-il} \\ \alpha_i^{k+1} v_j x_{-ij} &\sim \alpha_i^{k+2} w_j x_{-ij} \end{aligned}\right\} \Rightarrow \alpha_i^{k+1} t_l x_{-il} \sim \alpha_i^{k+2} u_l x_{-il} \tag{3.11}$$

Pick two points r_i^0 and r_i^1 in the common domain and set $V_i^{S_a}(r_i^0) = V_i^{S_b}(r_i^0) = 0$. Assume we now have $V_i^{S_a}(r_i^1) = v_a$ and $V_i^{S_b}(r_i^1) = v_b$. We need to show that for any point z_i from the common domain of $V_i^{S_a}$ and $V_i^{S_b}$ we have $V_i^{S_a}(z_i) = \lambda_i^{ab} V_i^{S_a}(z_i)$, where $\lambda_i^{ab} = \frac{v_b}{v_a}$.

Build standard sequences from r_i^0 to r_i^1, and from r_i^0 to z_i. We have

$$\begin{aligned} V_i^{S_a}(r_i^1) - V_i^{S_a}(r_i^0) &\approx n[V_j^{S_a}(v_j) - V_j^{S_a}(w_j)] \\ V_i^{S_a}(z_i) - V_i^{S_a}(r_i^0) &\approx m[V_j^{S_a}(v_j) - V_j^{S_a}(w_j)]. \end{aligned} \tag{3.12}$$

$V_i^{S_a}(r_i^0) = 0$, hence

$$V_i^{S_a}(z_i) \approx \frac{m V_i^{S_a}(r_i^1)}{n}. \tag{3.13}$$

Such n and m exist by the Archimedean axiom. By the argument above we get

$$\begin{aligned} V_i^{S_b}(r_i^1) - V_i^{S_b}(r_i^0) &\approx n[V_j^{S_b}(t_l) - V_j^{S_b}(u_l)] \\ V_i^{S_b}(z_i) - V_i^{S_b}(r_i^0) &\approx m[V_j^{S_b}(t_l) - V_j^{S_b}(u_l)]. \end{aligned} \tag{3.14}$$

Similarly,

$$V_i^{S_b}(z_i) \approx \frac{m V_i^{S_b}(r_i^1)}{n}. \tag{3.15}$$

By density, we can pick an arbitrary small step of the standard sequences, so the ratio $\frac{m}{n}$ converges to a limit. Thus, finally

$$V_i^{S_a}(z_i) = \frac{V_i^{S_a}(r_i^1)}{V_i^{S_b}(r_i^1)} V_i^{S_b}(z_i) = \frac{v_a}{v_b} V_i^{S_b}(z_i) = \lambda_i^{ab} V_i^{S_b}(z_i). \tag{3.16}$$

$\square$

We proceed by picking common locations for all value functions. Since r^0 belongs to all X^{S_a}, we can set $V_i^a(r_i^0) = 0$. At this point we can drop superscripts and say that we have representations

$$\lambda_i^a V_i + \ldots + \lambda_n^a V_n \tag{3.17}$$

on each X^{S_a}, defining also $\lambda_i^a := 0$ for variables i that are inessential on the set X^{S_a}.

3.7 Constructing global representation on X

At this stage we need to show that the representations on individual X^{S_i} can be re-aligned to assign the same value to equivalence classes of $\succcurlyeq$ in all X^{S_i}. We start by showing that values of the equivalence classes in subsets X^{S_a}, X^{S_b} that are defined by switching the $\mathbf{S}$-order of two coordinates i, j, are identical.

3.7.1 Aligning the value functions

Pick some X^{S_a} which has at least two essential *interacting* variables i, j, i.e. $i\,\mathbf{S}_a\,j$ and assume moreover that for no k we have $i\,\mathbf{S}_a\,k\,\mathbf{S}_a\,j$. Choose two points r^0 and r^1 from X^{S_a} such that $i\,\mathbf{E}^{r^0}\,j$ and $i\,\mathbf{E}^{r^1}\,j$, moreover, $r^1_k \succ_k r^0_k$ for all $k \in N$. why do these exist?

Set $V_i(r^0_i) = V_j(r^0_j) = 0$ and $V_i(r^1_i) = 1$. We now have $V_j(r^1_j) = k_j$, define $\phi_j := \frac{V_j}{k_j}$, and $\phi_i := \frac{V_i}{k_i}$ ($k_i = 1$). Additive representations on various X^{S_a} now have the form $\lambda^a_1 V_1(x_1) + \ldots + \lambda^a_i k_i \phi_i(x_i) + \lambda^a_j k_j \phi_j(x_j) + \ldots + \lambda^a_n V_n(x_n)$.

Now also, define ϕ_m for $m \neq i, j$ by letting $\phi_m := k_m V_m$, where k_m is such that $V_m(r^1_m) - V_m(r^0_m) = k_m$. Hence, we get $\phi_m(r^1_m)j - \phi_m(r^0_m) = 1$ for all m, including i and j.

The representation now is

$$\phi^a(x) = \lambda^a_1 k_1 \phi_1(x_1) + \ldots + \lambda^a_i k_i \phi_i(x_i) + \lambda^a_j k_j \phi_j(x_j) + \ldots + \lambda^a_n k_n \phi_n(x_n). \quad (3.18)$$

Finally, re-normalize it once again, by dividing by the sum of coefficients:

$$\phi^a(x) = \frac{\lambda^a_1 k_1}{\sum_{i=1}^n \lambda^a_i k_i}\phi_1(x_1) + \ldots + \frac{\lambda^a_n k_n}{\sum_{i=1}^n \lambda^a_i k_i}\phi_n(x_n) \quad (3.19)$$

Define $\alpha^a_m := \frac{\lambda^a_m k_m}{\sum_{i=1}^n \lambda^a_i k_i}$, and get the representation:

$$\phi^a(x) = \alpha^a_1 \phi_1(x_1) + \ldots + \alpha^a_n \phi_n(x_n). \quad (3.20)$$

With this in hand we can introduce the following lemma:

Lemma 22. *Let S_a and S_b be identical apart from $i, j \in N$ such that $i\,\mathbf{S}_a\,j$ and $j\,\mathbf{S}_b\,i$. Then, $\alpha^a_m = \alpha^b_m$ for all $m \neq i, j$.*

Proof. First, note that

$$\phi^a(r^1) - \phi^a(r^0) = \sum_{i=1}^{n} \alpha_i^a(\phi(r_i^1) - \phi(r_i^0)) = \sum_{i=1}^{n} \alpha_i^a = 1 \tag{3.21}$$

and similarly, $\phi^b(r^1) - \phi^b(r^0) = 1$. We can build equispaced sequences (see Section 3.11) from r^0 to r^1 in accordance with the orders $\mathbf{R}_a$ and $\mathbf{R}_b$, e.g. if m is minimal in $\mathbf{R}_a$ and $\mathbf{R}_b$, we first go from r^0 to $r_m^1 r_{-m}^0$, etc. The sequences can be constructed so that one resides entirely in X^{S_a} and another in X^{S_b}. By **A5**, they will have the same number of steps. By the above argument, global value increment per step would be identical in both cases as well. Finally, we can follow these sequences from r^0 or r^1 to $r_A^0 r_{-A}^1$, again, in accordance with orderings $\mathbf{R}_a$ and $\mathbf{R}_b$. Again, by **A5** the number of steps in both sequences would be the same, hence also the utility increment. We get for all A which either do not contain the "switched" pair i, j or contain both i and j:

$$\begin{gathered} \phi^a(r_A^1 r_{-A}^0) - \phi^a(r^0) = \phi^b(r_A^1 r_{-A}^0) - \phi^b(r^0) \\ \sum_{i \in A} \alpha_i^a(\phi_i(r_i^1) - \phi_i(r_i^0)) = \sum_{i \in A} \alpha_i^b(\phi_i(r_i^1) - \phi_i(r_i^0)) \\ \sum_{i \in A} \alpha_i^a = \sum_{i \in A} \alpha_i^b. \end{gathered} \tag{3.22}$$

The result follows. □

Corollary 1. *Let S_a and S_b be identical apart from $i, j \in N$ such that $i\,\mathbf{S}_a\,j$ and $j\,\mathbf{S}_b\,i$. Then $\alpha_i^a + \alpha_j^a = \alpha_i^b + \alpha_j^b$, and $\phi^a(r^0) = \phi^b(r^0)$ and $\phi^a(r^1) = \phi^b(r^1)$ (where r^0 and r^1 is defined as above).*

Note that we can change the location of any $\phi_m, m \neq i, j$ independently, and also change it's scale, by redefining $\phi'_m = k'_m \phi_m$ and then re-normalizing, independently and without violation of anything proved above. Indeed, for example $\phi^a(r^0) = \phi^b(r^0)$ and change the location of ϕ_m by letting $\phi'_m := \phi_m + \gamma_m$. By the uniqueness properties of the additive representations, this should lead to a valid additive representations of X^{S_a} and X^{S_b}. Moreover, apparently, $\alpha_m^a = \alpha_m^b$ still holds, and also $\phi^a(r^0) + \alpha_m^a \gamma_m = \phi^b(r^0) + \alpha_m^b \gamma_m$ is true.

If now we want to change the scale of ϕ_m by letting $\phi'_m = k'_m \phi_m$, we need to re-normalize the representations ϕ^a and ϕ^b by diving by $\alpha_1^a + \ldots + k'_m \alpha_m^a + \ldots + \alpha_n^a$ and $\alpha_1^b + \ldots + k'_m \alpha_m^b + \ldots + \alpha_n^b$ respectively. Since $\alpha_m^a = \alpha_m^b$, these sums are equal, so therefore everything above remains valid, including the equality of coefficients.

3.7.2 Equivalence classes have the same value

To simplify the construction in the main theorem of this section, we introduce the following lemma.

Lemma 23. *For every X^{S_a} and every $z \in X^{S_a}$, such that $r^0 \succ z$, we can find z', such that $z' \sim z$, $z' \in X^{S_a}$, and $r_i^0 \succcurlyeq_i z_i$. Likewise, for every $y \in X^{S_a}$, such that $y \succ r^0$, we can find y', such that $y' \sim y$, $y' \in X^{S_a}$, and $y_i \succcurlyeq_i r_i^0$.*

Proof. We can use the same procedure as was used in the proof of Theorem 14. For the case $y \succ r^0$ the procedure is exactly the same, while for $r^0 \succ z$ it is symmetric, as we are moving z this time, and not r^0. □

Notice that as a result of the rescaling made in Section 3.7.1, points r^0 and r^1 have the same values (0 and 1) in X^{S_a} and X^{S_b} (since $\phi_i = \phi_j$, $\alpha_k^a = \alpha_k^b$ for all $k \neq i, j$ and $\alpha_i^a + \alpha_j^a = \alpha_i^b + \alpha_j^b$).

Theorem 16. *For any $x \in X^{S_a}, y \in X^{S_b}$ we have $x \succcurlyeq y$ iff $\phi^a(x) \succcurlyeq \phi^b(y)$.*

Proof. First take $x \sim y$, such that $x \in X^{S_a}$, and $y \in X^{S_b}$. If $x \sim y \sim r^0$ or $x \sim y \sim r^1$, the conclusion is immediate, so assume otherwise. Let $x \succ r^0$. Using Lemma 23 we construct $x' \in X^{S_a}$ and $y' \in X^{S_b}$, such that $x \sim y \sim x' \sim y'$ and $x'_i \succcurlyeq_i r_i^0$, while $y'_i \succcurlyeq_i r_i^0$.

Next, build equispaced sequences from r^0 to r^1 in X^{S_a} and X^{S_b}, such that first steps of each sequence are equivalent (see details in Section 3.11). By **A5** the number of steps in both sequences is equal.

Finally, build sequences from r^0 to x' and y' (coordinate-wise dominance simplifies construction of the sequences). The number of steps again must be equal, hence the ratios between the number of steps it takes to reach r^1 and x', and between the number of steps it takes to reach r^1 and y' are equal, and hence taking the limit, we get $\phi^a(x) = \phi^a(x') = \phi^b(y') = \phi^b(y)$.

The same approach applies for $x \succ y$. By **A5** the number of steps in the equispaced sequence from r^0 to x must be greater, than in the sequence from r^0 to y. Hence also $\phi^a(x) > \phi^b(x)$.

We now have $x \sim y \Rightarrow \phi^a(x) = \phi^b(y)$ and $x \succ y \Rightarrow \phi^a(x) > \phi^b(y)$. This implies that $x \succcurlyeq y \iff \phi^a(x) \geq \phi^b(y)$. $\square$

3.7.3 Connecting all X^{S_i}

We have shown that equivalence classes have the same values in two areas X^{S_a} and X^{S_b}, where S_a an S_b are equivalent apart from one pair i, j. We need to employ the same technique to align the remaining areas X^{S_i} with each other and with X^{S_a} and X^{S_b}.

As shown above, every value function ϕ_m apart from i, j can be re-scaled and shifted (i.e. its origin changed) independently of each other and without affecting the equality of the value functions between X^{S_a} and X^{S_b}. Assume now, that exist another variable k, such that $j\,\mathbf{S}_a\,k$ and for no m we have $j\,\mathbf{S}_a\,m\,\mathbf{S}_a\,k$. Let S_c be equal to S_b apart from the order of variables j and k as before. We can repeat the construction exactly as in two previous sections and show that equivalence classes take the same values on X^{S_c} and X^{S_b} (hence also X^{S_a}). Note that we can temporarily adjust ϕ_i which was aligned with ϕ_j previously for construction purposes and then bring it back to the original position.

Due to the assumption that all subsets have at least two essential variables, two reference points on each coordinate set the location and the scaling factor k_i of each value function ϕ_i. We cannot have “breaks in cardinality”, in the sense that for no i there can exist x_i such that for no y_{-i}, the point $x_i y_{-i}$ belongs to X^{S_a} where i is essential. This follows from the structural assumption, otherwise by density we would be able to find another z_i with the same property and $x_i y_{-i} \sim z_i y_{-i}$ for all y_{-i}. It is easy to show that various areas provide a sufficient overlap of $X_i^{S_a}$, as we did in Section 3.6. review later

Thus we can move from one subset to another, moving also along a connected component of an interaction graph. When functions for all dimensions within the connected component are aligned, we can repeat the process for another component. One situation that is not covered is when a cycle is present in the connected component. We have aligned the value function for the first coordinate in the cycle with the second, then second with the third, etc. However, when we reach the last value function, we still need to align it with the first, but we cannot change the location and scaling factor of the first value function any more. However, due to the acyclicity axioms (**A3-OA** and **A5**), we can conclude that all value functions are now aligned, and hence the equivalence classes obtain the same value in all X^{S_a}

commented out cases wit a single vari- able on some X^{S_a}

Lemma 24. *Given* $\mathbf{S}_a$ *and* $\mathbf{S}_b$*, such that exists* $A \subset N$*, for which we have* $i\,\mathbf{R}_a\,j$ *iff* $i\,\mathbf{R}_b\,j$ *for all* $i \in A, j \in N \setminus A$*, the following is true*

$$\sum_{i \in A} \alpha_i^a = \sum_{i \in A} \alpha_i^b. \tag{3.23}$$

Proof. Consider $r_A^1 r_{-A}^0$, which belongs both to X^{S_a} and X^{S_b}. By Theorem 16 we have $\phi^a(r_A^1 r_{-A}^0) = \phi^b(r_A^1 r_{-A}^0)$, hence $\sum_{i \in A} \alpha_i^a \phi(r_i^1) = \sum_{i \in A} \alpha_i^b \phi(r_i^1)$, from which the conclusion follows as $\phi_i(r_i^1) = 1$ for all $i \in N$. □

3.7.4 Constructing the Choquet integral

Now we can proceed with construction of a unique *capacity* $\nu : 2^N \to \mathbb{R}$ from coefficients α_i^a which exist on various X^{S_a}. As shown in Wakker (1989), the condition of Lemma 24 is a necessary requirement for this. Capacity ν also has a unique Möbius transform $m : 2^N \to \mathbb{R}$ (see definition in Section 2.1.2).

We can now construct a representation very similar to the Choquet integral. In order to so, let us define the following function first: $\Phi_\wedge(x, A) := \phi_i(x_i)$ for i such that $j\,\mathbf{R}^x\,i$ for all $j \in A \setminus i$, in case when this is true for several i, any can be chosen. We can now construct a global value function (cf. Section 2.1.2):

$$\phi(x) := \sum_{A \in N} m(A)\Phi_\wedge(x, A). \tag{3.24}$$

It is easy to see that for each $x \in X$ and every X^{S_i}, such that $x \in X^{S_i}$, we have $\phi^a(x) = \phi(x)$. Now we can show that $\phi_i(x_i) = \phi_j(x_j)$ whenever $i\,\mathbf{E}^x\,j$, providing i, j interact.

Lemma 25. *For any non-extreme $x \in X$ it holds:*

$$i\,\mathbf{E}^x\,j \Rightarrow \phi_i(x_i) = \phi_j(x_j), \tag{3.25}$$

unless i and j do not interact.

Proof. Assume $x \in X^{S_a}, x \in X^{S_b}$ such that $k\,\mathbf{S}_a\,l$ whenever $k\,\mathbf{S}_b\,l$ for all $k, l \in N$ apart from i, j, for which we have $i\,\mathbf{S}_a\,j$ and $j\,\mathbf{S}_b\,i$. By Theorem 16, $\phi^a(x) = \phi^b(x)$ and by Lemma 24, it is trivial to show that $\alpha_k^a = \alpha_k^b$ for all $k \neq i, j$, and $\alpha_i^a + \alpha_j^a = \alpha_i^b + \alpha_j^b$.

We have $\alpha_i^a \phi_i(x_i) + \alpha_j^a \phi_j(x_j) = \alpha_i^b \phi_i(x_i) + \alpha_j^b \phi_j(x_j)$ (other sum components cancel out). Dividing by $\alpha_i^a + \alpha_j^a = \alpha_i^b + \alpha_j^b$, we get a convex combination of $\phi_i(x_i)$ and $\phi_j(x_j)$ on both sides. From this follows that either $\phi_i(x_i) = \phi_j(x_j)$ or $\alpha_i^a = \alpha_i^b$ and $\alpha_j^a = \alpha_j^b$.

Assume the latter. Repeating this operation for all possible combinations of X^{S_k} and X^{S_l} would lead us to the conclusion that $m(B) = 0$ for all $B \supset \{i, j\}$, as weights $\alpha_i^k, \alpha_i^l, \alpha_j^k, \alpha_j^l$ do not change when we move from X^{S_k} to X^{S_l}, and, accordingly, from ϕ^k to ϕ^l. The conclusion results from equation (3.24).

Finally, we can show that this implies that i and j do not interact. This means that ij-triple cancellation -

$$\left.\begin{aligned} a_i p_j z_{-ij} &\preccurlyeq b_i q_j z_{-ij} \\ a_i r_j z_{-ij} &\succcurlyeq b_i s_j z_{-ij} \\ c_i p_j z_{-ij} &\succcurlyeq d_i q_j z_{-ij} \end{aligned}\right\} \Rightarrow c_i r_j z_{-ij} \succcurlyeq d_i s_j z_{-ij}, \tag{3.26}$$

holds for all $a_i, b_i, c_i, d_i \in X_i$, $p_j, q_j, r_j, s_j \in X_j$, and all $z_{-ij} \in X_{-ij}$. To show this, use equation (3.24) to write the values for all involved points, grouping the sum components as follows. For example, for $a_i p_j z_{-ij}$:

$$\begin{aligned} \phi(a_i p_j z_{-ij}) = &\sum_{\substack{A \supset i \\ A \not\supset j}} m(A)\Phi_\wedge(a_i p_j z_{-ij}, A) + \sum_{\substack{A \supset j \\ A \not\supset i}} m(A)\Phi_\wedge(a_i p_j z_{-ij}, A) + \\ &+ \sum_{A \not\supset i,j} m(A)\Phi_\wedge(a_i p_j z_{-ij}, A) + \sum_{A \supset i,j} m(A)\Phi_\wedge(a_i p_j z_{-ij}, A). \end{aligned} \tag{3.27}$$

Notice, that due to the above argument, we have $\sum_{A \supset i,j} m(A)\Phi_\wedge(a_i p_j z_{-ij}, A) =$

0. Also notice that $\sum_{\substack{A\supset j\\ A\not\supset i}} m(A)\Phi_\wedge(a_i p_j z_{-ij}, A)$ does not depend on a_i, and $\sum_{A\not\supset i,j} m(A)\Phi_\wedge(a_i p_j z_{-ij}, A)$ does not depend on a_i or p_j.

Assume, towards a contradiction, that $d_i s_j z_{-ij} \succ c_i r_j z_{-ij}$. Writing values for all points, and summing the first and the third, and the second and the fourth inequalities gives:

$$\begin{aligned}
&\sum_{\substack{A\supset i\\ A\not\supset j}} m(A)[\Phi_\wedge(a_i p_j z_{-ij}, A) + \Phi_\wedge(d_i p_j z_{-ij}, A)] \leq \sum_{\substack{A\supset i\\ A\not\supset j}} m(A)[\Phi_\wedge(b_i p_j z_{-ij}, A) + \Phi_\wedge(c_i p_j z_{-ij}, A)]\\
&\sum_{\substack{A\supset i\\ A\not\supset j}} m(A)[\Phi_\wedge(a_i p_j z_{-ij}, A) + \Phi_\wedge(d_i p_j z_{-ij}, A)] > \sum_{\substack{A\supset i\\ A\not\supset j}} m(A)[\Phi_\wedge(b_i p_j z_{-ij}, A) + \Phi_\wedge(c_i p_j z_{-ij}, A)],
\end{aligned} \tag{3.28}$$

which is a contradiction. Hence, ij-triple cancellation holds for all $a_i, b_i, c_i, d_i \in X_i$, $p_j, q_j, r_j, s_j \in X_j$, and all $z_{-ij} \in X_{-ij}$, and thus i and j do not interact. $\square$

Lemma 26. *If for some $z \in X$ we have $i\,\mathbf{S}^z\,j$, then $\phi_i(z_i) > \phi_j(z_j)$.*

Proof. It is easy to verify that (due to the "Closedness" assumption), there exists $x_{ij}z_{-ij}$, such that $i\,\mathbf{E}^{x_{ij}z_{-ij}}\,j$ and $z_i \succcurlyeq_i x_i$, whereas $x_j \succcurlyeq_j z_j$. Since ϕ_i represents $\succcurlyeq_i$, by Lemma 25, and the fact that $\succcurlyeq_i$ is asymmetric (due to the structural assumption), we have $\phi_i(z_i) > \phi_i(x_i) = \phi_j(x_j) > \phi_j(z_j)$. $\square$

Now we have that $[i\,\mathbf{E}^x\,j] \Rightarrow [\phi_i(x_i) = \phi_j(x_j)]$ for interacting i,j, and $[i\,\mathbf{S}^x\,j] \Rightarrow [\phi_i(x_i) > \phi_j(x_j)]$, hence we conclude that $[i\,\mathbf{R}^x\,j] \iff [\phi_i(x_i) \succcurlyeq \phi_j(x_j)]$ for all interacting i,j, which allows us to finally rewrite (3.24) as the Choquet integral:

$$\phi(x) = \sum_{A\subset N} m(A) \min_{i\in A} \phi_i(x_i). \tag{3.29}$$

To summarize the results of this section we state the following lemma:

Lemma 27. *Let $X'_i := X \setminus \{$ maximal and minimal elements of $X\}$. Let $X' := X'_1 \times \ldots \times X'_n$. Assume that all sets ${X'}^{S_a}$, defined as previously, have at least two essential variables. Then exists a capacity ν and value functions $\phi_i : X_i \to \mathbb{R}$, such that for every $x, y \in X'$ we have*

$$x \succcurlyeq y \iff \phi(x) \geq \phi(y), \tag{3.30}$$

where $\phi(x)$ is the Choquet integral of $(\phi_1(x_1), \ldots, \phi_n(x_n))$ with respect to the capacity ν.

3.7.5 Case with a single essential variable on every X^{S_a}

For this case we only need **A3** to construct the representation. Since $\succcurlyeq$ is a weak order and each X^{S_i} has a countable order-dense subset, there exists a function $F : X \Rightarrow \mathbb{R}$, such that $x \succcurlyeq y \iff F(x) \geq F(y)$. To perform the construction of the value functions we need the following lemma.

Lemma 28. *Let $a_i w_{-i} \in X^{S_a}$ and $a_i x_{-i} \in X^{S_b}$. Let also i be the only essential coordinate on X^{S_a} and X^{S_b}. Then, $a_i w_{-i} \sim a_i x_{-i}$.*

Proof. The idea of the proof is to "trace a path" from X^{S_a} to X^{S_b} by constructing a sequence of points and subsets X^{S_j}. We will keep a_i unchanged, but will move the remaining coordinates, in order to show that any two subsequent points in the sequence belong to the same subset, moreover i will be the only essential coordinate on all X^{S_j}, so all points in the sequence will be equivalent.

Step 1. Assume there exist some k such that $i \,\mathbf{S}_a\, k$ but $k \,\mathbf{S}_b\, i$. If such points do not exist, move to Step 3. By closedness assumption, we can find y_k between x_k and w_k, such that $i \,\mathbf{E}\, k$ at $a_i y_k w_{-ik}$ and $a_i y_k x_{-ik}$. Symmetrically we can find such points for the reverse case.

Step 2. We can construct $a_i y_A w_{-iA}$ and $a_i y_A x_{-iA}$, where subset A includes all variables k from Step 1, however, generally, such points would not be in the same X^{S_j} as the initial points (i.e. not in X^{S_a} and X^{S_b}). Nevertheless we can easily construct intermediate points which would help to establish the equivalence between $a_i y_A w_{-iA}$ and $a_i w_{-i}$ on the one hand, and $a_i y_A x_{-iA}$ and $a_i x_{-i}$ on the other.

First, note that by **A3**, i stays the only essential variable if X^{S_a}, where i is essential, and X^{S_j} which is constructed by performing a swap ($\mathbf{S}$-wise) of a pair of coordinates, that does not change their relative position with i.

Hence, we can prove the equivalence of $a_i y_A w_{-iA}$ and $a_i w_{-i}$ algorithmically. For example, take the largest ($\mathbf{S}$-wise) $k \in A$, such that $i \,\mathbf{S}_a\, k$. Consider $a_i y_k w_{-ik}$. If this point belongs to some $X^{S_{a'}}$ but not to X^{S_a}, then exist at least one w_j, such that $j\,\mathbf{S}_a\,k$ but $k\,\mathbf{S}_{a'}\,j$. By closedness assumption, exists y'_k such that $j\,\mathbf{E}\,k$ at $a_iy'_kw_{-ik}$. Starting with the smallest ($\mathbf{S}$-wise) such j, we can construct a sequence of points, such that any two subsequent points belong to the same X^{S_j}, moreover by the above arggument i remains the only essential variable on all X^{S_j}. Hence, all points in the sequence are equivalent. Hence, $a_iy_kw_{-ik} \sim a_iw_{-i}$.

Having repeated this process for all $k \in A$, we can conclude $a_iy_Aw_{-iA} \sim a_iw_{-i}$ and $a_iy_Ax_{-iA} \sim a_ix_{-i}$.

Step 3. At this stage all remaining coordinates from $-iA$ are $\mathbf{S}$-ordered in the same way relatively to i in both points. It remains to build a sequence from $a_iy_Aw_{-iA}$ to $a_iy_Aw_{-iA}$. This can be done by substituting w_{-iA} by x_{-iA} point by point and following the same construction principle as in Step 2. We conclude that $a_iy_Aw_{-iA} \sim a_iy_Ax_{-iA}$, and hence $a_iw_{-i} \sim a_ix_{-i}$.

□

Using this lemma, we can now define value functions $\phi_i(x_i) = F(x)$ by picking $x \in X^{S_a}$ where i is the essential coordinate. Lemma shows that the functions are well-defined. It remains to construct a capacity. We can do so, by letting $\alpha_i^a = 1$ for essential i and $\alpha_j^a = 0$ for the remaining coordinates. It is easy to show a result similar to Lemma 24 (see Step 2 in the proof of Lemma 28). This implies that there exist a capacity ν, and the preference relation can be represented by a Choquet integral with respect to this capacity and value functions, defined as above.

Lemma 29. *Let the conditions of Theorem 8 hold and let there be only one essential variable on each X^{S_a}. Then, all variables are in the same interaction group.*

Proof. Assume there are two interaction groups - A and B. By **A7** all variables are essential, hence exists a set X^{S_a} such that some $i \in A$ is essential on X^{S_a}. Assume that $\mathbf{S}_b$ is obtained by switching the ($\mathbf{S}$) order of one pair of variables in $\mathbf{S}_a$, without changing the ordering in other pairs. As shown previously (see proof of Lemma 28), if such switch does not involve i, it will still remain the essential variable one X^{S_b}. If the switch involves i and some other variable j, then either i or j, but not both, are essential on X^{S_b}. Since variables from A do not interact with variables from B, a switch involving a variable from A and a variable from B can never occur, hence on all subsets X^{S_i} we can only have essential variables from A, which contradicts **A7**. □

Lemma 30. *Let the conditions of Theorem 8 hold and let there be only one essential variable on each X^{S_a}. Then, the capacity ν in the representation* (3.6) *is unique.*

Proof. Follows from Lemma 29 and by construction. Further we will show that this capacity only takes values 0 and 1 (see Lemma 46). □

3.8 Extending the representation to the extreme points of X

3.8.1 Definition value functions at maximal and minimal points of X_i

need a better proof

In contrast to the homogeneous case in (Wakker, 1991b) we cannot show with a single statement that all value functions are bounded on X_i. Instead, we need to look at which functions are bounded on $X_i^{S_a}$.

Consider a maximal point M_i. There must exist z_{-i} such that $M_i z_{-i} \in X^{S_a}$ and i is essential on X^{S_a}. By assumption, there exists at least one other variable j essential on X^{S_a}. There can be two cases:

- For some j we can find two points $v_j \succcurlyeq_j w_j$ such that for some y_{-ij} we have $M_i v_j y_{-ij}, M_i w_j y_{-ij} in X^{S_a}$. Then, as in Lemma 20 in (Wakker, 1991b), we can build a bound $t_i v_j y_{ij} \succcurlyeq M_i w_j y_{-ij}$, and so ϕ_i is bounded.

- Otherwise, we are not able to change any of the z_j's. This implies that such z_j's are actually maximal and moreover all of them are such that $NOT i\, \mathbf{S}_a\, j$. In this case we should be able to adjust M_i itself slightly (as above) and show that ϕ_j are bounded. Finally, by a chain of **E** relations and Lemma 25, we have $\lim_{z_i \to M_i} \phi_i(z_i) = \lim_{z_j \to M_j} \phi_i(z_j)$, and hence ϕ_i is bounded from above.

We can now define $\phi(M_i) := \lim_{z_i \to M_i} \phi_i(z_i)$, and $\phi(m_i) := \lim_{z_i \to m_i} \phi_i(z_i)$ for all $i \in N$. Assigning values to minimal elements of X_i can be done in a similar manner. Finally we can prove the final theorem.

3.8.2 Global representation on the whole X

Theorem 17. *For any $x, y \in X$, we have $x \succcurlyeq y \iff \phi(x) \succcurlyeq \phi(y)$.*

Proof. We proceed as in Wakker (1991b) (Lemma 21) with some modifications. For points that do not contain any maximal or minimal coordinates, this has been already proved (Section 3.7). Thus, assume that x or y contain maximal or minimal coordinates. Let $x \succ y$, and let $x \in X^{S_a}, y \in X^{S_b}$. Find S_a-minimal j such that x_j is maximal. We can also assume that for all k, such that $j \, \mathbf{S}_a \, k$, we have $j \, \mathbf{S}^x \, k$. If this is not the case, then x belongs to several X^{S_i} (by definition of these sets), and there must be one where this condition holds. By Lemma 40 we can find $x'_j : x_j \succcurlyeq_j x'_j$ such that $x'_j x_{-j} \in X^{S_a}$ and still $x'_j x_{-j} \succ y$. Proceeding like this we get x' which does not contain any maximal coordinates, and $x' \succ y$. We now need to show that $\phi(x') > \phi(y)$. Similarly, we can replace minimal coordinates of y, and so it is now required to show that $\phi(x') > \phi(y')$. So we can assume that x has no maximal and y has no minimal coordinates. x must have a non-minimal X^{S_a}-essential coordinate, find a S_a-maximal one i. Again, we can assume that for all k, such that $j \, \mathbf{S}_a \, k$, we have $j \, \mathbf{S}^x \, k$. ;By Lemma 40 we can decrease it slightly and find $x'_i : x_i \succ x'_i$ and $x'_i x_{-i} \in X^{S_a}$ and still $x' := x'_i x_{-i} \succ y$. So we need to show now only that $\phi(x') \geq \phi(y)$. If we replace all minimal coordinates of x' by non-minimal ones and all maximal coordinates of y by non-maximal ones, then by monotonicity the preference between them is not affected, and by Lemma 27, we have $\phi(x') > \phi(y)$. Thus any small increase of minimal and small decrease of maximal values leads to a strict inequality. By definition of ϕ_i at extreme elements of X_i, we have that $\phi(x')$ is the infimum of all such ϕ-values, and $\phi(y)$ is the supremum. Hence, $\phi(x') \succcurlyeq \phi(y)$. $x \succ y \Rightarrow \phi(x) > \phi(y)$ also implies $\phi(x) \geq \phi(y) \Rightarrow x \succcurlyeq y$.

Now let $\phi(x) > \phi(y)$. x cannot have all it's essential coordinates minimal, so find S_a-maximal j, such that x_j is not minimal. By denserangedness of ϕ_j, we can find a non-minimal $x'_j : x_j \succ_j x'_j$ and still $\phi(x'_j x_{-j}) > \phi(y)$. By the above argument, we have $x'_j x_{-j} \succcurlyeq y$, and by strict monotonicity we have $x \succ y$. $\square$

3.9 Uniqueness

We have shown, that the preference relation $\succcurlyeq$ can have the following representation:

$$x \succcurlyeq y \iff C(\nu, (f_1(x_1), \ldots, f_n(x_n))) \geq C(\nu, (f_1(y_1), \ldots, f_n(y_n))). \quad \text{(RepCh-n)}$$

In this section we analyse the uniqueness of the value functions and the capacity. It would be convenient to use the Möbius transformation form of the Choquet integral.

$$C(\nu, x) = \sum_{A \subset N} m(A) \min_{i \in A} \phi_i(x_i), \quad \text{(RepCh-Mob)}$$

where m is the Möbius transform of the capacity ν.

Uniqueness properties are based on two fundamental results, namely Lemma 25, which states that for interacting i and j $i\,\mathbf{R}^z\,j$ implies that $\phi_i(z_i) \geq \phi_j(z_j)$, and the uniqueness properties of the additive representations.

In contrast to the traditional additive value model, the Choquet integral representation includes not only the value functions but also the scaling factors, or the "weights", which form the capacity ν. As we will see below, it is not always possible to uniquely disentangle value functions from the weights. In particular, the possibility to do so is strongly conditional on the *absence* of separability between two variables (in other words, on whether ij-triple cancellation holds on all X_{ij} or not). We will show that the decomposition is unique for the variables within an interaction group, but across groups such property does not hold. In fact, the extreme case when *all* variables are pairwise separable is just the (weighted) additive value model, where the weights and the value functions are completely confounded. Indeed, if we have a representation of the form

$$F(x) = w_1 f_1(x_1) + \ldots + w_n f_n(x_n), \quad (3.31)$$

we can change the weights w_i arbitrarily, by multiplying f_i by some factor t_i and making a change of variables $w_i f_i = w'_i f'_i$, where $w'_i = \frac{w_i}{t_i}$ and $f'_i = t_i f_i$.

The following Lemma shows that the capacity is getting "separated" along the lines of interaction groups of N.

Lemma 31. *Let $A_1, A_2, ...$ be the interaction groups of N. Then, for any $B : B \cap A_i \neq \emptyset, B \cap A_j \neq \emptyset$, we have $m(B) = 0$. Also, if for two sets A_1 and A_2 we have $m(B) = 0$ for all $B : B \cap A_1 \neq \emptyset, B \cap A_2 \neq \emptyset$, then coordinates from A_1 do not interact with coordinates from A_2.*

Proof. Let $i \in A_1$ and $j \in A_2$. We need to show that for any $B : i, j \in B$, we have $m(B) = 0$, and vice versa, if $m(B) = 0$ for every such B, then i and j do not interact. Assume that for some such B we have $m(B) \neq 0$. Then, we can find $x_{ij}z_{-ij} \in X^{S_a}$, $y_{ij}z_{-ij} \in X^{S_b}$, such that $\alpha_k^a = \alpha_k^b$ for all $k \neq i, j$, and $\alpha_i^a \neq \alpha_i^b, \alpha_j^a \neq \alpha_j^b$. This implies that ij-trade-off consistency does not hold on all X_{ij}, hence the variables interact. To show the reverse, note that we have $\alpha_i^a = \alpha_i^b, \alpha_j^a = \alpha_j^b$ for all possible points in X, which implies ij-trade-off consistency on all X_{ij}. $\square$

When defining functions ϕ_i in Section 3.7, we choose the origin and scaling factor independently for every interaction group $A \in N$. In the expression (RepCh-Mob) for all i from some interaction group A_k we can redefine value functions by setting $f'_i(z_i) = \frac{1}{t_k} f_i(z_i)$. Accordingly, all Möbius transform coefficients containing variables from A_k must be multiplied by t_k. As the above Lemma shows, this transformation will not impact any of the summands including value functions of coordinates outside A_k. Thus, the newly formed expression remains a valid representation. We can also define $m'(B) = t_k m(B)$ for $B \subset A_k$ and $m'(B) = m(B)$ for all other B. Apparently, $m'(B)$ is not a capacity (unless $t_k = 1$), as $\sum_{B \subset N} m'(B) \neq 1$. However, we can re-normalize the representation by dividing everything by $\sum_{B \subset N} m'(B)$. It is easy to see that the resulting expression is once again a Choquet integral.

To avoid this ambiguity in stating the uniqueness results, it is convenient to introduce the following lemma.

Lemma 32. *Let $A_1, \ldots, A_k$ be the interaction groups of N, and let $\succcurlyeq$ have the* *representation* (3.6)*. Then the following is also a representation of* $\succcurlyeq$*:*

$$x \succcurlyeq y \iff \sum_{i=1}^{k} C(\nu_{A_i}, f_{A_i}(x_{A_i})) \geq \sum_{i=1}^{k} C(\nu_{A_i}, f_{A_i}(y_{A_i})), \tag{Add-Ch}$$

, where ν_{A_i} *are capacities defined on the sets of all subsets of* A_i*. Moreover, the capacities* A_i *are unique.*

Proof. The lemma follows from the above argument. We only need to pick values t_i in a way that $\sum_{B \subset A_i} t_i m(B) = 1$ for each interaction group A_i. The uniqueness comes from the fact that within the interaction group we are not able to re-scale the value functions individually (see further Lemma 33). □

Finally, we can state the uniqueness theorem for the representation (3.7). We prefer this representation, as the capacities ν_{A_i} are unique, contrary to the general Choquet integral representation (3.6), where we can change the capacity almost arbitrarily (see the argument above).

Lemma 33. *Let* $A_1, \ldots, A_k$ *be the interaction groups of* N*. Let* $g_1 : X_1 \to \mathbb{R}, \ldots, g_n : X_n \to \mathbb{R}$ *be such that (3.7) holds with* f_i *substituted by* g_i*. Let* $i \in A_m$*. Then,*

- *In case every* X^{S_i} *has at least two essential variables (cardinal case),*

$$f_i(x_i) = \alpha g_i(x_i) + \beta_{A_m}. \tag{3.32}$$

- *In case every* X^{S_i} *has only one essential variable (ordinal case),*

$$f_i(x_i) = \psi_i(g_i(x_i)), \tag{3.33}$$

  where ψ_i *is an increasing function, and for all* A_m *and* $i, j \in A_m$*, we additionally have*

$$f_i(x_i) \geq f_j(x_j) \iff g_i(x_i) \geq g_j(x_j). \tag{3.34}$$

Proof. Follows from uniqueness properties of additive and ordinal representations and Lemma 25. □

Appendix to Chapter 3

The following sections contain technical and auxiliary results.

3.10 Necessity of axioms

Lemma 34. ***A3*** *is necessary.*

Proof. At any point $z \in X$ and for every i, j, we must have either $f_i(z_i) \geq g_i(z_i)$ or $g_i(z_i) \geq f_i(z_i)$. From this everything follows trivially (write the condition using Mobius representation of the integral). $\square$

Lemma 35. ***A6*** *is necessary.*

Proof. Assume $a_i x_{-i}, b_i x_{-i}, c_i x_{-i}, d_i x_{-i} \in X^{S_a}$ and $a_i x_{-i} \succ b_i x_{-i}, c_i x_{-i} \sim d_i x_{-i}, c_i y_{-i} \succ d_i y_{-i}$. There can be three cases:

1. $c_i y_{-i}, d_i y_{-i} \in X^{S_b}$. We have additive representations on X^{S_a} and X^{S_b}, so

$$\begin{aligned} \alpha_i f_i(a_i) + \sum_{j \in N \setminus i} \alpha_j f_j(x_j) &> \alpha_i f_i(b_i) + \sum_{j \in N \setminus i} \alpha_j f_j(x_j) \\ \alpha_i f_i(c_i) + \sum_{j \in N \setminus i} \alpha_j f_j(x_j) &= \alpha_i f_i(d_i) + \sum_{j \in N \setminus i} \alpha_j f_j(x_j) \\ \beta_i f_i(c_i) + \sum_{j \in N \setminus i} \beta_j f_j(y_j) &> \beta_i f_i(d_i) + \sum_{j \in N \setminus i} \beta_j f_j(y_j). \end{aligned} \tag{3.35}$$

The first inequality entails $\alpha_i \neq 0$. From this and the following equality follows $f_i(c_i) = f_i(d_i)$, which contradicts with the last inequality. Thus $c_i y_{-i} \succ d_i y_{-i}$ implies $c_i x_{-i} \succ d_i x_{-i}$ but only in the presence of $a_i x_{-i} \succ b_i x_{-i}$ in the same X^{S_a} (the case when $c_i y_{-i}$ and $d_i y_{-i}$ are not both in the same X^{S_b} can be reduced to this one. This is also the reason behind the name we gave to this condition - "weak bi-independence".

2. $c_i y_{-i}, d_i y_{-i} \in X^{S_a}$. In this case we get $\alpha_i f_i(c_i) > \alpha_i f_i(d_i)$ and $\alpha_i f_i(c_i) = \alpha_i f_i(d_i)$, a contradiction.

3. $c_i y_{-i} \in X^{S_a}, d_i y_{-i} \in X^{S_b}$. As above $\alpha_i \neq 0$, so it follows that $f_i(a_i) = f_i(b_i)$. But then we must have $d_i y_{-i} \in X^{S_a}$ (value functions are all equal to those for $c_i y_{-i}$), and hence the conclusion follows as in the previous case. $\square$

3.11 Equispaced sequences

A usual standard sequence goes along a single dimension as defined above. In this paper we often require to move along several dimensions, one at a time, maintaining the increment between steps constant in some sense. In order to achieve this we will introduce the concept of *equispaced* sequences [3]. Figure 3.3 illustrates the process.

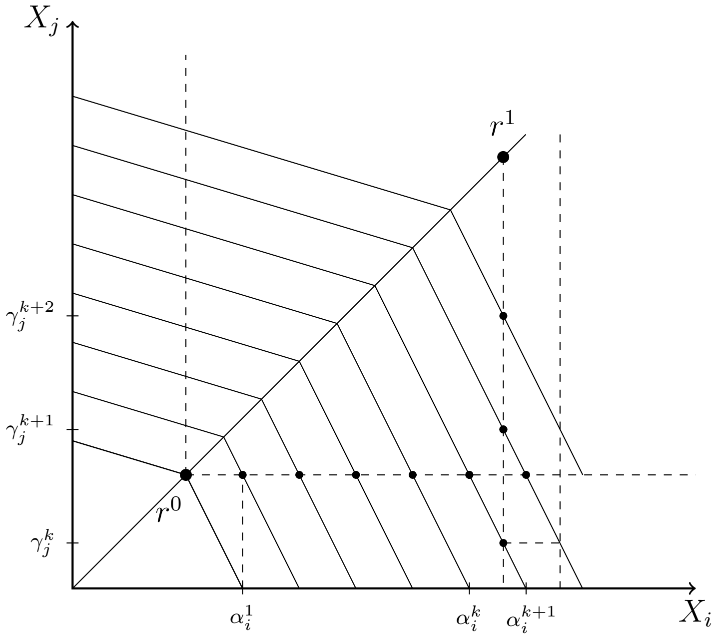


Figure 3.3: Equispaced sequences in two dimensions

Assume that r^0, r^1 are such that $i \, \mathbf{R}^{r^0} \, j$ and $i \, \mathbf{R}^{r^1} \, j$. We would like to build a sequence from r^0 to r^1 staying in the area where $i \, \mathbf{R} \, j$. We can choose the size of the sequence step arbitrarily. However, the problem is that r^1 does not have an equivalent point with the second coordinate equal to r^0_j, so we cannot build a

[3] See also (Bouyssou and Marchant, 2010) for a similar idea

"normal" standard sequence to achieve that. Our aim is to maintain the sequence within set where $i\,\mathbf{R}\,j$. We also assume that X_i and X_j do not have maximal or minimal elements (or they have been removed).

By density and the absence of maximal and minimal elements, we can find $\alpha_i^k, \alpha_i^{k+1}$ such that $\alpha_i^{k+1} \succcurlyeq_i r_i^1 \succcurlyeq_i \alpha_i^k$. We need to change the direction of the sequence from the dimension i to the dimension j at r_i^1. We construct a point equivalent to a_i^k and a point equivalent to a_i^{k+1} such that their i's coordinate is r_i^1 (points γ_j^k and γ_j^{k+1}). Since we can choose the step of the sequence arbitrarily, by density and absence of maximal elements, we can move on and construct a standard sequence on the coordinate j using these two points.

Remarkably the spacing between subsequent members of the equispaced sequence $\alpha^1, \ldots, \alpha^{k-1}, \gamma^k, \gamma^{k+1}, \ldots$ stays in a certain sense the same, no matter along which dimension we are moving. Once an additive interval scale is constructed, the vague notion of the equal spacing will convert into a clear constant difference of values for subsequent members of the sequence.

Extension of A5 to equispaced sequences Construction of equispaced sequences allows us to extend the statement of **A5**, and more precisely that of Lemma 39 to equispaced sequences.

Lemma 36. *If g^k and h^k are two equispaced sequences entirely lying in X^{S_a} and X^{S_b} correspondingly, and for some i we have $g^i \sim h^i$ and $g^{i+1} \sim h^{i+1}$, then for all j such that exist g^j and h^j we have $g^j \sim h^j$.*

Proof. Without loss of generality assume that $g^i := g_k^i g_{-k}$ and $g^{i+1} := g_k^{i+1} g_{-k}$, while $h^i := h_l^i h_{-l}$ and $h^{i+1} := h_l^{i+1} h_{-l}$, i.e. in both cases the points are from subsequences on the same dimensions. Assume further, that $h^{i+2} := h_l^{i+2} h_{-l}$ and $g^{i+2} := g_m^{i+2} g'_{-m}$, i.e. there is a change of dimension in the sequence g^k. We will show that this entails that the following steps of the equispaced sequence g^k will be equivalent to the corresponding steps of the sequence h^k as long as both keep going along dimensions m and l correspondingly. If either of them changes its direction the technique can be applied again. We have by construction (see Figure 3.3) $g_k^{i+2} g_{-k} \sim g_m^{i+2} g'_{-m}$ and $g_k^{i+3} g_{-k} \sim g_m^{i+3} g'_{-m}$. Therefore, $g_m^{i+2} g'_{-m} \sim h_l^{i+2} h_{-l}$ and $g_m^{i+3} g'_{-m} \sim h_l^{i+3} h_{-l}$. The statement follows by Lemma 39 (**A5**). $\square$

3.12 Technical lemmas

Lemma 37. *If $\succcurlyeq$ satisfies triple cancellation then it is independent.*

Proof. $a_i p_{-i} \preccurlyeq a_i p_{-i}, a_i q_{-i} \succcurlyeq a_i q_{-i}, a_i p_{-i} \succcurlyeq b_i p_{-i} \Rightarrow a_i q_{-i} \succcurlyeq b_i q_{-i}$. □

Lemma 38. $X = \bigcup X^{S_i}$.

Proof. Immediate by **A3**. □

Definition 23. *For any set I of consecutive integers, a set $\{g_i^k : g_i^k \in X_i, k \in I\}$ is a* standard sequence *on coordinate i if exist z_{-ij}, y_j^0, y_j^1 such that $y_j^0 \not\sim_j y_j^1$ and for all $i, i+1 \in I$ we have $g_i^k y_j^1 z_{-ij} \sim g_i^{k+1} y_j^0 z_{-ij}$. Further, we say that $\{g_i^k : k \in I\}$* is contained in X^{S_a} *if z_{-ij}, y_j^0, y_j^1 can be chosen in such a way, that all resulting points are in X^{S_a}.*

Lemma 39. *Axiom **A5** implies the following condition. Let $\{g_i^k : k \in I\}$ and $\{h_j^k : k \in I\}$ be two standard sequences, each contained in some X^{S_a}. Assume also, that for some $m \in I$, $g_i^m y_l^0 z_{-il} \sim h_j^m w_n^0 x_{-jn}$ and $g_i^{m+1} y_l^0 z_{-il} \sim h_j^{m+1} w_n^0 x_{-jn}$. Then, for all $k \in I$, $g_i^k y_l^0 z_{-il} \sim h_j^k w_n^0 x_{-jn}$.*

Proof. The proof is very similar to the one from (Krantz et al., 1971) (Lemma 5 in section 8.3.1). Assume wlog that $\{g_i^k : k \in I, g_i^k y_l^1 z_{-il} \sim g_i^{k+1} y_l^0 z_{-il}\}$ is an increasing standard sequence on i, which is contained in X^{S_a}, whereas $\{h_j^k : h_j^k w_n^1 x_{-jn} \sim h^{k+1} w_n^0 x_{-jn}\}$ is an increasing standard sequence on j, and lies entirely in X^{S_b}. We will show that $g_i^{m+2} y_l^0 z_{-il} \sim h_j^{m+2} w_n^0 x_{-jn}$ from which everything follows by induction.

Assume $h_j^{m+2} w_n^0 x_{-jn} \succ g_i^{m+2} y_l^0 z_{-il}$. Then, by restricted solvability exists $\hat{h}_j \in X_j$ such that $\hat{h}_j w_n^0 x_{-jn} \sim g_i^{m+2} y_l^0 z_{-il}$. By **A5**,

$$\left.\begin{aligned}
g_i^m y_l^1 z_{-il} &\sim g_i^{m+1} y_l^0 z_{-il} \\
g_i^{m+1} y_l^1 z_{-il} &\sim g_i^{m+2} y_l^0 z_{-il} \\
g_i^m y_l^0 z_{-il} &\sim h_j^m w_n^0 x_{-jn} \\
g_i^{m+1} y_l^0 z_{-il} &\sim h_j^{m+1} w_n^0 x_{-jn} \\
g_i^{m+1} y_l^0 z_{-il} &\sim h_j^{m+1} w_n^0 x_{-jn} \\
g_i^{m+2} y_l^0 z_{-il} &\sim \hat{h}_j w_n^0 x_{-jn} \\
h_j^m w_n^1 x_{-jn} &\sim h_j^{m+1} w_n^0 x_{-jn}
\end{aligned}\right\} \Rightarrow h_j^{m+1} w_n^1 x_{-jn} \sim \hat{h}_j w_n^0 x_{-jn}. \tag{3.36}$$

By definition of $\{h_j^k\}$, we have $h_j^{m+1} w_n^1 x_{-jn} \sim h_j^{m+2} w_n^0 x_{-jn}$. Thus, $h_j^{m+2} w_n^0 x_{-jn} \sim \hat{h}_j w_n^0 x_{-jn}$, hence also $g_i^{m+2} y_l^0 z_{-il} \sim h_j^{m+2} w_n^0 x_{-jn}$, a contradiction. The other cases are symmetrical. $\square$

Lemma 40. *Let there be x, y, i such that $x \succ y$, $x \in X^{S_a}$, $i \,\mathbf{S}^x\, k$ for all $k : i \,\mathbf{S}_a\, k$, and x_i is non-minimal if i is minimal in $\mathbf{S}_a$. Then exists z_i, such that $z_i \succcurlyeq_i x_i$, $i \,\mathbf{S}^{z_i x_{-i}}\, k$ for all $k : i \,\mathbf{S}_a\, k$, and $z_i x_{-i} \succ y$. Similarly, let there be x, y, i such that $y \succ x$, $x \in X^{S_a}$, $k \,\mathbf{S}^x\, i$ for all $k : k \,\mathbf{S}_a\, i$, and x_i is non-maximal if i is maximal in $\mathbf{S}_a$. Then exists z_i, such that $x_i \succcurlyeq_i z_i$, $k \,\mathbf{S}^{z_i x_{-i}}\, i$ for all $k : k \,\mathbf{S}_a\, i$, and $y \succ z_i x_{-i}$.*

Proof. By restricted solvability and monotonicity. See Wakker (1991b) Lemma 11. $\square$

3.13 Essentiality and monotonicity

The essentiality of coordinates within various X^{S_i} is critical. The central mechanism to guarantee consistency in the number of essential coordinates within various X^{S_i} is the strong monotonicity axiom **A6**, which is closely related to "comonotonic strong monotonicity" of Wakker (1989).

In the Choquet integral representation problem for a heterogeneous product set $X = X_1 \times \ldots \times X_n$, strong monotonicity is actually a necessary condition. It is directly implied by **A6** and the structural assumption.

Lemma 41. *Pointwise monotonicity. If for all $i \in N$ it holds $x_i \succcurlyeq_i y_i$, then $x \succcurlyeq y$.*

Proof. $x \succcurlyeq y_1 x_{-1} \succcurlyeq y_{12} x_{-12} \succcurlyeq \ldots \succcurlyeq y$. $\square$

Lemma 42. *If i is essential on X^{S_a}, then $a_i \succcurlyeq_i b_i$ iff $a_i x_{-i} \succ b_i x_{-i}$ for all $a_i x_{-i}, b_i x_{-i} \in X^{S_a}$.*

Proof. If $a_i \succcurlyeq_i b_i$ then by the structural assumption exists $a_i z_{-i} \succ b_i z_{-i}$. The result follows by (**A6**). $\square$

Conceptually, Lemma 42 implies that if a coordinate i is essential on some subset of X^{S_a}, then it is also essential on the whole X^{S_a}. This allows us to make statements like "coordinate i is essential on X^{S_a}".

3.14 Shape of $\{z_{ij} : i\,\mathbf{E}^z\,j\}$

Shape of the boundary between subsets of X_{ij} where $i\,\mathbf{R}\,j$ and $j\,\mathbf{R}\,i$ is an interesting and important question. Axiom **A3** only guarantees that this boundary is in a certain sense "quasiconvex", i.e. an increase along i cannot be matched by a decrease along j. Strengthening this statement requires invoking other axioms. We only consider the cardinal case in this section.

In section 3.7 we have shown, that in the representation value functions for sets X_i and X_j are equal for points z where $i\,\mathbf{E}^z\,j$. Theorem 18 provides a qualitative version of this statement. The assumption we must make is that i and j are essential on X^{S_a} and X^{S_b}, such that S_a and S_b differ only with respect to order of i and j. In the below proof, we assume that i, j are S_a and S_b-maximal, but this can be easily changed, by starting from some $r^1_A r^0_{-A}$ instead of r^0.

Theorem 18. *Let $r^0 : i\,\mathbf{E}^{r^1}\,j, r^1 : i\,\mathbf{E}^{r^1}\,j$ and a^k_i and b^k_j are two standard sequences such that $a^0_i r^0_{-i} \sim r^0_i r^0_{-i}$ and $b^0_j r^0_{-j} \sim r^0_j r^0_{-j}$ and $r^1_i r^0_{-i} \sim a^m_i r^0_{-i}$ whereas $r^1_j r^0_{-j} \sim b^m_j r^0_{-j}$. Assume r^2 is such that $i\,\mathbf{E}^{r^2}\,j$ and $r^2_i r^0_{-i} \sim a^n_i r^0_{-i}$. Then $r^2_j r^0_{-j} \sim b^n_j r^0_{-j}$.*

Proof. Build two equispaced sequences from r^0 to $r^1_{ij} r^0_{-ij}$:

- e^k starting from r^0_i via $r^1_i r^0_{-i}$, and
- w^k starting from r^0_j via $r^1_j r^0_{-j}$,

such that $e^1_i r^0_{-i} \sim w^1_j r^0_{-j}$. By Lemma 39 (**A5**) it follows then that all corresponding steps of two sequences are equivalent, in other words, $e^k \sim w^k$ for all k. Consequently, there is the same number of steps both sequences make between r^0 and $r^1_{ij} r^0_{-ij}$, say K.

For some $s < K$ we have $r^1_i r^0_{-i}$ lying between e^s and e^{s+1}, i.e. $e^{s+1} \succ r^1_i r^0_{-i} \succcurlyeq e^s$. Similarly, for some $t < K$ we have $w^{t+1} \succ r^1_j r^0_{-j} \succcurlyeq w^t$. We can write:

$$[r^0, r^1_i r^0_{-i}] \approx n a^k_i \approx s e^k, \tag{3.37}$$

which means: $r^1_i r^0_i$ lies between $a^n_i r^0_{-i}$ and $a^{n+1}_i r^0_{-i}$ and also between e^s and e^{s+1}.

Similarly,

$$\begin{aligned} [r^0, r_j^1 r_{-j}^0] &\approx nb_j^k \approx tw^k \\ [r_i^1 r_{-i}^0, r_{ij}^1 r_{-ij}^0] &\approx nb_j^k \approx (K-s)e^k \\ [r_j^1 r_{-j}^0, r_{ij}^1 r_{-ij}^0] &\approx na_i^k \approx (K-t)w^k. \end{aligned} \tag{3.38}$$

Two last statements are possible because by density we can can get arbitrarily close to points $r_j^1 r_{-j}^0$ and $r_i^1 r_{-i}^0$ by choosing finer sequences e^k and w^k.

For point $r_{ij}^2 r_{-ij}^0$ we have:

$$\begin{aligned} [r^0, r_i^2 r_{-i}^0] &\approx ma_i^k \approx \frac{m}{n} se^k \\ [r_j^2 r_{-j}^0, r_{ij}^2 r_{-ij}^0] &\approx ma_i^k \approx \frac{m}{n}(K-t)w^k. \end{aligned} \tag{3.39}$$

Assume that the number of steps on two other segments is different:

$$\begin{aligned} [r^0, r_j^2 r_{-j}^0] &\approx lb_j^k \approx \frac{l}{n} tw^k \\ [r_i^2 r_{-i}^0, r_{ij}^2 r_{-ij}^0] &\approx lb_j^k \approx \frac{l}{n}(K-s)e^k. \end{aligned} \tag{3.40}$$

Summing up parts for both paths to $r_{ij}^2 r_{-ij}^0$ we get $\frac{ms+l(K-s)}{n}e^k$ for $\mathbf{SE}_{ij}$ and $\frac{m(K-t)+lt}{n}w^k$ for $\mathbf{NW}_{ij}$. By Lemma 39 (**A5**) the number of steps must be identical, so:

$$\frac{ms + l(K-s)}{n} = \frac{m(K-t) + lt}{n}, \tag{3.41}$$

or

$$m(s+t-K) = l(s+t-K). \tag{3.42}$$

There are two possible solutions:

- $m = l$, and
- $t = K - s$, which means that trade-offs are consistent throughout X, hence i and j do not interact, a contradiction.

The result follows. □

Corollary 2. *If $a_i p_j : i\,\mathbf{E}^{a_i p_j z_{-ij}}\,j$ and $b_i p_j : i\,\mathbf{E}^{b_i p_j z_{-ij}}\,j$, then $x_i y_j : i\,\mathbf{E}^{x_i y_j z_{-ij}}\,j$ for all $x_i \in X_i$, $y_j \in Y_j$.*

Corollary 3. *If* $a_ip_j : i\,\mathbf{E}^{a_ip_jz_{-ij}}\,j$, *then*

- *for any* b_i, *such that* $b_i \succsim_i a_i$ *we have* $i\,\mathbf{S}^{b_ip_jz_{-ij}}\,j$,
- *for any* b_i, *such that* $a_i \succsim_i b_i$ *we have* $j\,\mathbf{S}^{b_ip_jz_{-ij}}\,i$,
- *for any* q_j, *such that* $q_j \succsim_j p_j$ *we have* $j\,\mathbf{S}^{a_iq_jz_{-ij}}\,i$,
- *for any* q_j, *such that* $p_j \succsim_j q_j$ *we have* $i\,\mathbf{S}^{a_iq_jz_{-ij}}\,j$.

or i, j *do not interact.*

removed alte
native ordina
case treatme

3.15 Summary

In this chapter we have presented an axiomatization of the Choquet integral which is the most general of the available so far. Compared to the results in the previous chapter we have made a significant addition to our axiom system – the acyclicity conditions. Overall, various possible geometric layouts of the decision space make our results more general compared to the two-dimensional heterogeneous and n-dimensional homogeneous cases. The most remarkable implications of our characterization are due to the uniqueness properties of the model. In the absence of homogeneity, the correspondence between elements of various dimension sets is unique and meaningful only in the presence of interaction, i.e. in the absence of separability between the dimensions. In general, the Choquet integral is decomposed into a sum of the Choquet integrals, one per interaction group. The extreme cases are the additive value model, where all interaction groups are of size one, and the single interaction group, which is what we also get in the homogeneous-comonotonic case. Notably, the uniqueness properties of a purely ordinal model are stronger as we always get a single interaction group including all variables. In the following chapter we will revisit the homogeneous model and show that comonotonic independence is not a necessary condition for the existence of a Choquet integral representation on such sets. In fact, comonotonicity implies state-independence which is not always a desirable property.

Chapter 4

Extensions and discussion

4.1 Introduction

This chapter contains extensions of our results to the particular interesting special cases of the Choquet integral, analysis of some aspects of the Choquet integral model learning, and a discussion of the applications of our results in decision theory. The Choquet integral is a powerful aggregation operator which lists many well-known models as its special cases. In this chapter we look at these special cases and provide their axiomatic analysis. In cases where an axiomatization has been previously given in the literature, we connect the existing results with the framework that we have developed.

Next we turn to the question of learning, which is especially important for the practical applications of the model. So far, learning of the Choquet integral has been mostly confined to the learning of the capacity. Such an approach requires making a powerful assumption that all dimensions (e.g. criteria) are evaluated on the same scale, which is rarely justified in practice. Too often categorical data is given arbitrary numerical labels (e.g. AHP), and numerical data is considered cardinally and ordinally commensurate, sometimes after a simple normalization. Such approaches clearly lack scientific rigour, and yet they are commonly seen in all kinds of applications. We discuss the pros and cons of making such an assumption and look at the consequences which our uniqueness results have for the learning problems.

Finally, we revisit some of the applications we discussed in the Introduction. Apart from MCDA, which is the main area of interest for our results, we also discuss how the model can be interpreted in the social choice context. We look in detail at the state-dependent utility, and show how comonotonicity, central to the previous axiomatizations, actually implies state-independency in the Choquet integral model. We also discuss the conditions required to have a meaningful state-dependent utility representation and show the novelty of our results compared to the previous methods of building state-dependent models.

4.2 Extensions

4.2.1 Ordinal models

Notable ordinal special cases of the Choquet integral are:

- Min/Max
- Order statistic (k-smallest element) OS_k
- Lattice polynomial $p^{\mathcal{AB}}$.

Moreover, Min/Max are special cases of OS_k ($k = 1$ and $k = n$ correspondingly), and OS_k is a special case of the lattice polynomial model, as becomes evident from the following definitions.

Definition 24. *$\succcurlyeq$ can be represented by* MIN*, if exist value functions $\phi_i : X_i \to \mathbb{R}$ such that for all $x, y \in X$ we have*

$$x \succcurlyeq y \iff \bigwedge_{i \in N} \phi_i(x_i) \geq \bigwedge_{i \in N} \phi_i(y_i), \tag{4.1}$$

where $\bigwedge$ means minimum.

Definition 25. *$\succcurlyeq$ can be represented by* MAX*, if exist value functions $\phi_i : X_i \to \mathbb{R}$ such that for all $x, y \in X$ we have*

$$x \succcurlyeq y \iff \bigvee_{i \in N} \phi_i(x_i) \geq \bigvee_{i \in N} \phi_i(y_i), \tag{4.2}$$

where $\bigvee$ means maximum.

Definition 26. *$\succcurlyeq$ can be represented by an* order statistic *OS_k, if exist value functions $\phi_i : X_i \to \mathbb{R}$ such that for all $x, y \in X$ we have*

$$x \succcurlyeq y \iff \phi_{(k)}(x_{(k)}) \geq \phi_{(k)}(y_{(k)}), \tag{4.3}$$

where $\phi_{(k)}(z_{(k)})$ stands for kth smallest element of $(\phi_1(z_1), \ldots, \phi_n(z_n))$.

An order statistic can be written in a CNF and DNF-like[4] forms (e.g. Ovchin-

[4]Conjunctive normal form and disjunctive normal form.

nikov, 1996):

$$OS_k = \bigwedge_{\substack{K\subset N\\|K|=k}} \bigvee_{i\in K} \phi_i(x_i) = \bigvee_{\substack{K\subset N\\|K|=n-k+1}} \bigwedge_{i\in K} \phi_i(x_i). \tag{4.4}$$

Obviously, MIN and MAX are particular cases of OS_k with $k = 1$ and $k = n$ correspondingly.

Definition 27. *$\succcurlyeq$ can be represented by a* lattice polynomial $p^{\mathcal{AB}}$*, if exist value functions $\phi_i : X_i \to \mathbb{R}$ such that for all $x, y \in X$ we have*

$$x \succcurlyeq y \iff p^{\mathcal{AB}}(\phi_1(x_1), \dots, \phi_n(x_n)) \geq p^{\mathcal{AB}}(\phi_1(y_1), \dots, \phi_n(y_n)), \tag{4.5}$$

where $p^{\mathcal{AB}}(\phi_1(z_1), \dots, \phi_n(z_n))$ is an expression which includes elements of $(\phi_1(z_1), \dots, \phi_n(z_n))$ and symbols $\vee$ and $\wedge$.

We can write any lattice polynomial in DNF and CNF as well:

$$p^{\mathcal{AB}}(\phi_1(z_1), \dots, \phi_n(z_n)) = \bigwedge_{K\subset \mathcal{A}} \bigvee_{i\in K} \phi_i(x_i) = \bigvee_{M\subset \mathcal{B}} \bigwedge_{i\in M} \phi_i(x_i),, \tag{4.6}$$

where $\mathcal{A} \subset 2^N$ and $\mathcal{B} \subset 2^N$ are some collection of subsets of N. Obviously, order statistic, hence MIN and MAX are special cases of an order polynomial.

The following result states that all aforementioned models are special cases of the Choquet integral.

Theorem 19 (Murofushi and Sugeno, 1993)**.** *The Choquet integral with respect to a capacity ν is a lattice polynomial function if and only if ν is a 0–1 capacity (i.e. only takes values 0 or 1). Moreover, any lattice polynomial function on $\mathbb{R}$ is a Choquet integral with respect to a 0–1 capacity.*

4.2.2 Previous characterizations of the ordinal models

Some known characterizations of the models presented in the previous section are due to Bouyssou et al. (2002), see also Sounderpandian (1991) and Segal and Sobel (2002).

Theorem 20 (Bouyssou et al., 2002)**.** *$\succcurlyeq$ can be represented by MAX if $\succcurlyeq$ is a weak order and the following equivalent conditions hold:*

1. *For all $i \in N, x_i, y_i \in X_i, a_{-i}, b_{-i} \in X_{-i}$ and $w \in X$, we have*

$$[x_i a_{-i} \succsim w] \Rightarrow [y_i a_{-i} \succsim w \text{ } OR \text{ } x_i b_{-i} \succsim w] \tag{4.7}$$

2. *For all $x, y \in X, i \in N$:*

$$[x_i y_{-i} \succsim x] \text{ } OR \text{ } [y_i x_{-i} \succsim x] \tag{4.8}$$

3. *For all $i \in N, y_i \in X_i, z_{-i} \in X_{-i}, x \in X$:*

$$[y_i x_{-i} \succ x] \Rightarrow [y_i z_{-i} \succ x]. \tag{4.9}$$

Theorem 21 (Bouyssou et al., 2002)**.** *$\succsim$ can be represented by MIN if $\succsim$ is a weak order and the following equivalent conditions hold:*

1. *For all $i \in N, x_i, y_i \in X_i, a_{-i}, b_{-i} \in X_{-i}$ and $w \in X$, we have*

$$[w \succsim x_i a_{-i}] \Rightarrow [w \succsim y_i a_{-i} \text{ } OR \text{ } w \succsim x_i b_{-i}] \tag{4.10}$$

2. *For all $x, y \in X, i \in N$:*

$$[x \succsim x_i y_{-i}] \text{ } OR \text{ } [x \succsim y_i x_{-i}] \tag{4.11}$$

3. *For all $i \in N, y_i \in X_i, z_{-i} \in X_{-i}, x \in X$:*

$$[x \succ y_i x_{-i}] \Rightarrow [x \succ y_i z_{-i}]. \tag{4.12}$$

Theorem 22 (Bouyssou et al., 2002)**.** *$\succsim$ can be represented by OS_{n-1} if $\succsim$ is a weak order and the following equivalent conditions hold:*

1. *For all $i, j \in N (i \neq j), x_i, y_i \in X_i, x_j, y_j \in X_j, a_{-i} \in X_{-i}, b_{-j} \in X_{-j}, c_{-ij} \in X_{-ij}$ and $w \in X$, we have*

$$[x_i a_{-i} \succsim w \text{ } AND \text{ } x_j b_{-j} \succsim w] \Rightarrow [y_i a_{-i} \succsim w \text{ } OR \text{ } y_j b_{-j} \succsim w \text{ } OR \text{ } x_{ij} c_{-ij} \succsim w] \tag{4.13}$$

2. *For all $x, y \in X, i, j \in N (i \neq j)$:*

$$[x_i y_{-i} \succcurlyeq x \;\; AND \;\; x_j y_{-j} \succcurlyeq x] \;\; OR \;\; [y_{ij} x_{-ij} \succcurlyeq x] \tag{4.14}$$

3. *For all $x, y \in X$, all $i, j \in N (i \neq j)$, and all $z_{-ij} \in X_{-ij}$:*

$$[y_i x_{-i} \succ x \;\; AND \;\; y_j x_j \succ x] \Rightarrow [y_{ij} z_{-ij} \succ x]. \tag{4.15}$$

4.2.3 Unified characterization of the ordinal models: $p^{\mathcal{A}\mathcal{B}}$ and subcases

Since MIN and MAX are special cases of OS_k, which in turn is a special case of the lattice polynomial models $p^{\mathcal{A}\mathcal{B}}$, it is desirable to build a unified characterization for all of them. In this section we provide some steps towards such result.

Theorem 23. *$\succcurlyeq$ can be represented by a lattice polynomial $p^{\mathcal{A}\mathcal{B}}$ if $\succcurlyeq$ is a weak order, satisfies **A2**, and for any $w, x \in X$ exist $K \in \mathcal{A}, M \in \mathcal{B}$ with $K \cap M \neq \emptyset$, such that for any $a_{-K} \in X_{-K}$ and $b_{-M} \in X_{-M}$ we have:*

$$\begin{cases} w \succcurlyeq x \Rightarrow w \succcurlyeq a_{-K} x_K, K \in \mathcal{A}, \\ x \succcurlyeq w \Rightarrow b_{-M} x_M \succcurlyeq w, M \in \mathcal{B}. \end{cases} \tag{4.16}$$

Note that, because sets $\mathcal{A}$ and $\mathcal{B}$ are finite, the axiom can also be re-written similar to the conditions in the previous section, i.e. using "OR" statements. However, we feel this form is more compact. Particular cases of the above axiom include OS_k and MIN/MAX.

proof needed

Lemma 43. *$\succcurlyeq$ can be represented by OS_k if $\succcurlyeq$ is a weak order, satisfies **A2**, and for any $w, x \in X$ there exist $K : K \subset N, |K| = k$ and $M : M \subset N, |M| = n - k + 1$ with $K \cap M \neq \emptyset$, such that for any $a_{-K} \in X_{-K}$ and $b_{-M} \in X_{-M}$ we have*

$$\begin{cases} w \succcurlyeq x \Rightarrow w \succcurlyeq a_{-K} x_K, \\ x \succcurlyeq w \Rightarrow b_{-M} x_M \succcurlyeq w. \end{cases} \tag{4.17}$$

Lemma 44. *$\succcurlyeq$ can be represented by MIN if $\succcurlyeq$ is a weak order and for any*

$w, x \in X$ *exists* $i \in N$*, such that for any* $a_{-i} \in X_{-i}$ *we have*

$$\begin{cases} w \succcurlyeq x \Rightarrow w \succcurlyeq a_{-i}x_i, \\ x \succcurlyeq w \Rightarrow x \succcurlyeq w. \end{cases} \tag{4.18}$$

Lemma 45. $\succcurlyeq$ *can be represented by MAX if* $\succcurlyeq$ *is a weak order and for any* $w, x \in X$ *exists* $i \in N$*, such that for any* $b_{-i} \in X_{-i}$ *we have*

$$\begin{cases} w \succcurlyeq x \Rightarrow w \succcurlyeq x, \\ x \succcurlyeq w \Rightarrow b_{-i}x_i \succcurlyeq w. \end{cases} \tag{4.19}$$

The second condition in two last lemmas is trivial and is given only to emphasize the similarity of the axiom to the one used above. Note also, that the first conditions in MIN/MAX characterizations are identical to those given in Section 4.2.2.

Although the condition in two last lemmas is sufficient for characterization of MIN and MAX, in general, variations of (4.16) are not powerful enough to characterize $p^{\mathcal{AB}}$ and OS_k. One reason for this is that in the MIN/MAX case the axioms imply our **A2** (the axiom that is called **AC1** in Bouyssou et al. (2009)) – in other words they imply existence of weak orders on individual dimensions. This does not seem to be the case for the $p^{\mathcal{AB}}$ and OS_k conditions that we gave. Hence, we had to add **A2** to the first two results.

4.2.4 Characterization of the ordinal models in our framework

In Section 3.7.5 we gave details of the construction of the Choquet integral for cases when every subset X^{S_i} has only one essential variable. We now provide more details on this result.

Lemma 46. *Let the conditions of Theorem 8 hold and let there be only one essential variable on each* X^{S_a}*. Then,* ν *is a 0–1 capacity.*

Proof. This immediately follows by construction (see Section 3.7.5). As at every $x \in X$ we have $C(\nu, x) = f_i(x_i)$, where i is the variable essential on $X^{S_i} \ni x$, by the definition of the Choquet integral and monotonicity of ν it follows that ν only takes values 0 and 1. $\square$

Lemma 47. *Let the conditions of Theorem 8 hold and let there be only one essential variable on each X^{S_a}.*

- *$\succcurlyeq$ can be represented by $p^{\mathcal{AB}}$;*
- *If the essential variable on every X^{S_i} is the $\mathbf{R}$-minimal one, then $\succcurlyeq$ can be represented by MIN;*
- *If the essential variable on every X^{S_i} is the $\mathbf{R}$-maximal one, then $\succcurlyeq$ can be represented by MAX;*
- *If the essential variable on every X^{S_i} is the $\mathbf{R}$-k-minimal one, then $\succcurlyeq$ can be represented by $OS-k$.*

Proof. The first statement follows from Theorem 19. Other follow by construction and from the uniqueness properties of the representation (3.6) in the ordinal case (see Theorem 10). If **S** ordering is incomplete, then only one **R** ordering can exists which does not contradict **A3,A7** and the condition that only one variable is essential on every X^{S_a}. This follows from the uniqueness of the capacity and the uniqueness properties of the value functions. $\square$

4.2.5 Cardinal models

The particular cases of the Choquet integral in the case of cardinal value functions are related to the convexity of the capacity. We give a characterization of the convex capacity (the concave case is easily obtainable by reversing the preference). Note that in the two-dimensional case, the class of the Choquet integrals with respect to convex capacities coincides with the class of Gilboa–Schmeidler maximin models. In the general case of n dimensions, every Choquet integral with respect to a convex capacity is a Gilboa–Schmeidler model – the integral is a minimum of integrals with respect to probability distributions from the capacity's core (Gilboa and Schmeidler, 1994) – but not other way round.

To our knowledge, this is the first result which characterizes convexity of a capacity using only the primitives of $\succcurlyeq$ and works in ordinal or mixed as well as purely cardinal cases, i.e. it is suitable for situations when standard sequences are not available.

Theorem 24. *Let conditions **A1–A9** and structural assumptions hold. Then, we have*

A10 – Convexity *For all $i, j \in N$ and for all $a_i, b_i, c_i, d_i \in X_i$, $p_j, q_j, r_j, s_j \in X_j$, and all $z_{-ij} \in X_{-ij}$ we have*

$$\left.\begin{aligned} a_i p_j z_{-ij} &\sim b_i q_j z_{-ij} \\ a_i r_j z_{-ij} &\sim b_i s_j z_{-ij} \\ c_i p_j z_{-ij} &\sim d_i q_j z_{-ij} \\ d_i &\succcurlyeq_i c_i \\ r_j &\succcurlyeq_j s_j \end{aligned}\right\} \Rightarrow c_i r_j z_{-ij} \succcurlyeq d_i s_j z_{-ij}, \tag{4.20}$$

provided $j\,\mathbf{R}\,i$ at $a_i p_j z_{-ij}, b_i q_j z_{-ij}, a_i r_j z_{-ij}, \sim b_i s_j z_{-ij}, c_i p_j z_{-ij}, d_i q_j z_{-ij}$ and $i\,\mathbf{R}\,j$ at $c_i r_j z_{-ij}$ and $d_i s_j z_{-ij}$,

if and only if ν is a convex capacity.

Proof. Since conditions **A1–A9** and structural assumptions hold, there exists a Choquet integral representation of $\succcurlyeq$. We can use it to prove the statement of the theorem. A capacity is convex if for all $i, j \in N, A \subset N, i \neq j$ we have (Chateauneuf and Jaffray, 1989):

$$\sum_{i,j \in B \subset A} m(B) \geq 0. \tag{4.21}$$

First, let $c_i r_j z_{-ij} \prec d_i s_j z_{-ij}$. We can write the conditions above using the Möbius form of the Choquet integral. All subsets of N can be separated into four groups:

- $A : A \ni i, A \not\ni j$
- $A : A \ni j, A \not\ni i$
- $A : A \ni i, A \ni j$

- $A : A \not\ni i, A \not\ni j$.

Hence, the value function for each of the points in the axiom can be written as follows. For example, for $a_i p_j z_{-ij}$ (note that we have merged $A : A \ni i, A \not\ni j$ and $A : A \ni i, A \ni j$ groups by virtue of $j \, \mathbf{R} \, i$ at $a_i p_j z_{-ij}$):

$$\sum_{A \ni i} m(A) \min_{k \in A-ij} [f_i(a_i), f_k(z_k)] + \sum_{\substack{A \ni j \\ A \not\ni i}} m(A) \min_{k \in A-ij} [f_i(p_j), f_k(z_k)] + \sum_{A \ni i,j} m(A) \min_{k \in A-ij} [f_k(z_k)]. \tag{4.22}$$

Writing down all four conditions like this and after some trivial algebraic transformations which we omit in the name of readability (sum first two conditions, add to the sum of the last two conditions and simplify), we get

$$\begin{aligned} &\sum_{A \ni i,j} m(A) \left(\min_{k \in A-ij} [f_i(d_i), f_k(z_k)] - \min_{k \in A-ij} [f_i(c_i), f_k(z_k)] \right) \\ &\quad + \sum_{A \ni i,j} m(A) \left(\min_{k \in A-ij} [f_j(r_j), f_k(z_k)] - \min_{k \in A-ij} [f_j(s_j), f_k(z_k)] \right) < 0. \end{aligned} \tag{4.23}$$

We will show that both summands of the above expression are non-negative. Consider

$$\sum_{A \ni i,j} m(A) \left(\min_{k \in A-ij} [f_i(d_i), f_k(z_k)] - \min_{k \in A-ij} [f_i(c_i), f_k(z_k)] \right). \tag{4.24}$$

The difference $f_i(d_i), f_k(z_k)] - \min_{k \in A-ij} [f_i(c_i), f_k(z_k)$ is

- always non-negative, as $d_i \succcurlyeq_i c_i$
- maximal, when $A = \{i, j\}$
- non-increasing as A grows larger.

Note that, by convexity, $m(\{i, j\}) \geq 0$. Hence,

$$m(\{i,j\}) \left(\min_{k \in \emptyset} [f_i(d_i), f_k(z_k)] - \min_{k \in \emptyset} [f_i(c_i), f_k(z_k)] \right) = m(\{i,j\}) \left(f_i(d_i) - f_i(c_i) \right) \geq 0. \tag{4.25}$$

Next, find a maximal $f_{k^1}(z_{k^1}), k^1 \in N \setminus i, j$. Note that in the above expression we will only have one element $\min_{k\in A-ij}[f_i(d_i), f_k(z_k)] - \min_{k\in A-ij}[f_i(c_i), f_k(z_k)]$ where k^1 is not redundant (since it's maximal). We get

$$\begin{aligned}
&m(\{i,j\})\,(f_i(d_i) - f_i(c_i)) + m(\{i,j,k^1\})\,(\min[f_i(d_i), f_{k^1}(z_{k^1})] - \min[f_i(c_i), f_{k^1}(z_{k^1})]) \\
&\geq [m(\{i,j\}) + m(\{i,j,k^1\})]\,(\min[f_i(d_i), f_{k^1}(z_{k^1})] - \min[f_i(c_i), f_{k^1}(z_{k^1})]) \geq 0.
\end{aligned} \tag{4.26}$$

The first inequality is since $m(\{i,j\}) \geq 0$ and the second is since $m(\{i,j\}) + m(\{i,j,k^1\}) \geq 0$, by convexity criterion. Now pick the second largest $f_{k^2}(z_{k^2}), k^2 \in N \setminus i, j, k^1$. Using the same arguments we get

$$\begin{aligned}
&m(\{i,j\})\,(f_i(d_i) - f_i(c_i)) + m(\{i,j,k^1\})\,(\min[f_i(d_i), f_{k^1}(z_{k^1})] - \min[f_i(c_i), f_{k^1}(z_{k^1})]) \\
&\quad + m(\{i,j,k^2\})\,(\min[f_i(d_i), f_{k^2}(z_{k^2})] - \min[f_i(c_i), f_{k^2}(z_{k^2})]) \\
&\quad + m(\{i,j,k^1,k^2\})\,(\min[f_i(d_i), f_{k^2}(z_{k^2})] - \min[f_i(c_i), f_{k^2}(z_{k^2})]) \geq [m(\{i,j\}) \\
&\qquad + m(\{i,j,k^1\}) + m(\{i,j,k^2\}) + m(\{i,j,k^1,k^2\})]\,(\min[f_i(d_i), f_{k^2}(z_{k^2})] \\
&\qquad - \min[f_i(c_i), f_{k^2}(z_{k^2})]) \\
&\quad \geq 0.
\end{aligned} \tag{4.27}$$

Continuing like this we can add more and more elements and eventually conclude that

$$\sum_{A \ni i,j} m(A) \left(\min_{k\in A-ij}[f_i(d_i), f_k(z_k)] - \min_{k\in A-ij}[f_i(c_i), f_k(z_k)] \right) \geq 0. \tag{4.28}$$

Similarly,

$$\sum_{A \ni i,j} m(A) \left(\min_{k\in A-ij}[f_j(r_j), f_k(z_k)] - \min_{k\in A-ij}[f_j(s_j), f_k(z_k)] \right) \geq 0. \tag{4.29}$$

Hence we have shown that the axiom necessarily holds if the capacity is convex. To show the inverse, assume that the axiom holds on X. Writing down conditions of the axiom and simplifying as before, we get that everywhere on X we should have

$$\begin{aligned}
&\sum_{A\ni i,j} m(A)\left(\min_{k\in A-ij}[f_i(d_i),f_k(z_k)]+\min_{k\in A-ij}[f_j(r_j),f_k(z_k)]\right)\\
&\quad\geq \sum_{A\ni i,j} m(A)\left(\min_{k\in A-ij}[f_i(c_i),f_k(z_k)]-\min_{k\in A-ij}[f_j(s_j),f_k(z_k)]\right).
\end{aligned}\tag{4.30}$$

Assume i,j interact. If this is not the case, the convexity criterion is trivially satisfied for i,j as all $m(A)$ in the expression above are 0 (see Lemma 31). Assume also all variables are in the same interaction group. If this is not the case, $m(A)$ for A containing variables not in the same interaction group as i,j are again 0, and can be discarded.

With this assumption made, we can now pick points, such that $f_i(\cdot)$ and $f_j(\cdot)$ are the smallest value functions. Hence, the above expression reduces to

$$[f_i(d_i)+f_j(r_j)]\sum_{A\ni i,j} m(A)\geq [f_i(c_i)+f_j(s_j)]\sum_{A\ni i,j} m(A). \tag{4.31}$$

Since $[f_i(d_i)+f_j(r_j)]\geq[f_i(c_i)+f_j(s_j)]$, we conclude that $\sum_{i,j\in A\subset N} m(A)\geq 0$.

Now pick points such that only $f_{k^1}(z_{k^1})$ is less than $f_i(\cdot)$ and $f_j(\cdot)$. We get

$$\begin{aligned}
&[f_i(d_i)+f_j(r_j)]\sum_{\substack{A\ni i,j\\ A\not\ni k^1}} m(A)+2f_{k_1}(z_{k^1})\sum_{A\ni i,j,k^1} m(A)\\
&\quad\geq [f_i(c_i)+f_j(s_j)]\sum_{\substack{A\ni i,j\\ A\not\ni k^1}} m(A)+2f_{k_1}(z_{k^1})\sum_{A\ni i,j,k^1} m(A),
\end{aligned}\tag{4.32}$$

or

$$[f_i(d_i)+f_j(r_j)]\sum_{\substack{A\ni i,j\\ A\not\ni k^1}} m(A)\geq [f_i(c_i)+f_j(s_j)]\sum_{\substack{A\ni i,j\\ A\not\ni k^1}} m(A). \tag{4.33}$$

From this we conclude that $\sum_{i,j\in A\subset N\setminus k^1} m(A)\geq 0$.

Continuing like this we can check all necessary sums for the convexity condition and for all pairs i,j. So, we have shown that the capacity is convex provided the axiom holds. □

4.3 Learning the Choquet integral

Learning the model means deriving model parameters from data. This step is essential in any practical application, and it is normally performed towards at least one of two goals: analysis of the data, by means of interpreting model parameters, or prediction – in other words, "training" the model on some dataset to use it with some other data.

It is well known that the quality of fit of a model depends on the model complexity and the available data. Learning a very complex model using only a few data points would not achieve satisfactory results, just as using a very simple model might conceal some important properties of a large and complicated dataset.

An important aspect of the learning process is its computational viability. Indeed, from the practical perspective, using a simpler but faster model which is capable of delivering approximate answers in real-time fashion, might be preferable to employing a more precise but also more expensive model which takes hours or days to be built.

In this section we look at various aspects of the Choquet integral learning and emphasize the consequences which our axiomatization results have for this process. We start by an overview of the current learning techniques and then look at difficulties which arise when learning the Choquet integral model in the full generality.

maybe We also suggest new learning techniques using our axiomatic construction methods.

To learn the Choquet integral we need to derive two parts of the model from data:

- value functions $f_i : X_i \to R$, and
- capacity ν.

The following sections provide details on each of these components.

4.3.1 Learning the capacity

The majority of the theoretical and applied literature so far has concentrated on learning ("identification") of the capacity *only*. In this approach, the value functions are assumed as given. Normally, for numerical coordinates $f_i(x_i) = x_i$ are taken (probably after some rescaling). For categorical data, sometimes arbitrary numerical labels are used (see e.g. AHP), although the theoretical problems of this approach are quite apparent.

A good review of the existing methods of capacity construction can be found in Grabisch et al. (2008). In the majority of cases, the learning process is based on minimization of some loss function (MSE, MAE, or similar), or on finding the extremum of some meaningful expression, such as variance or entropy.

Typically, data is used to formulate constraints on the space of possible parameters (i.e. capacities). For example, if $x \succcurlyeq y$, then ν must be such that $C(\nu, f(x)) \geq C(\nu, f(y))$ (remember the value functions are considered known). Since the integral is a linear function of the capacity, we get linear constraints. Eventually the polyhedron of all possible capacities is defined by the following data:

Learning set. Pairwise preferences between elements of the "learning set" X.

Criteria importance. The most intuitive way to describe a multicriteria model qualitatively is, perhaps, to define the relative weights of its components. The process is semantically similar to that for additive models; however, due to non-additivity we can not rely only on values for singletons any more, but must also take into account all other subsets of N.

Criteria interaction. A more complicated type of knowledge about criteria is the character of their combined influence. In particular, criteria can complement each other, which is also known under the name of positive synergy, or else be redundant (resp. negative synergy).

Veto and favour criteria. Sometimes the model also includes criteria of an immense importance, so that the alternatives having low valuations on them will also inevitably receive low overall judgements. This kind of criterion is usually called "veto" in the literature. The opposite situation is having a criterion (or criteria) such that a high value on them automatically justifies a high overall valuation. Such elements are called "favour".

Complexity controls. Often it is deemed that interactions in groups larger than k can be ignored to improve the computational properties of the model. The mechanism which allows us to achieve this is called *k-additivity*. Most frequently, 2-additive capacities are used.

The following indices were originally applied for behavioral analysis of non-additive measures. However, they also allow us to formulate and solve the inverse problem of capacity identification (see Marichal and Roubens, 2000 and references therein).

Definition 28 (Shapley, 1953)**.** *The Shapley value is an additive measure $\phi_\nu : 2^N \to [0,1]$ defined as*

$$\phi_\nu(i) = \sum_{T \subset N \setminus i} \frac{(|N|-|T|-1)!\,|T|!}{|N|!}[\nu(T \cup i) - \nu(T)]. \tag{4.34}$$

It can also be expressed via the Möbius transform coefficients:

$$\phi_m(i) = \sum_{T \subset N \setminus i} \frac{1}{|T|+1} m(T \cup i). \tag{4.35}$$

The semantic interpretation given to the Shapley value of a criterion $i \in N$ in the literature is the relative importance of the said criterion in the decision problem. More formally, it amounts to the average marginal input of that criterion to all subsets of N. Being a probability measure, the Shapley value sums up to 1 over all $i \in N$. Table 4.1 demonstrates how the Shapley value can be used in capacity identification problems (δ_{SH} is some small value – the indifference coefficient). Intuition about the relative importance of a criterion can be expressed

Table 4.1: Criteria importance modelling

The criterion i is more important than j	$\phi_\nu(i) - \phi_\nu(j) \geqslant \delta_{SH}$
Criteria i and j are equally important	$-\delta_{SH} \geqslant \phi_\nu(i) - \phi_\nu(j) \leqslant \delta_{SH}$

as $\phi_\nu(i) = k$ or $\phi_\nu(i) \in [k^l, k^u]$, although, just like in the additive case, doing so is not strictly sensible.

The measure of criteria interaction character and strength was introduced by Murofushi and Soneda (1993) for pairs of elements and later generalized by Grabisch (1997b).

Definition 29. *The interaction index of a subset $T \subset N$ is defined as*

$$I_\nu(T) = \sum_{k=0}^{|N|-|T|} \xi_k^{|T|} \sum_{K \subset Z \setminus T, |K|=k} \sum_{L \subset T} (-1)^{|T|-|L|} \nu(L \cup K), \tag{4.36}$$

where

$$\xi_k^p = \frac{(|N| - k - p)!\, k!}{(|N| - p + 1)!}. \tag{4.37}$$

For practical problems we are particularly interested in the index expression for pairs $\{i, j\}$:

$$I_\nu(ij) = \sum_{T \subset N \setminus ij} \xi_{|T|}^2 \left[\nu(T \cup ij) - \nu(T \cup i) - \nu(T \cup j) + \nu(T)\right], \tag{4.38}$$

or, when expressed with the Möbius transform coefficients:

$$I_m(ij) = \sum_{T \subset N \setminus ij} \frac{1}{|T|+1} m(T \cup ij). \tag{4.39}$$

The interaction index for singletons coincides with the Shapley value. The index can be interpreted as the degree of interaction between elements in the set T. Its values lie in the interval $[-1; 1]$, with 1 corresponding to the maximal positive interaction (complementarity), and -1, accordingly, to the maximal negative interaction (redundancy). Table 4.2 summarizes index usage in identification problems.

To model "veto" and "favour" criteria we can proceed in the following way (Grabisch, 1997a). If some criterion i is a "veto" one, then

$$\nu(A) = 0 \qquad \forall A \not\supseteq i. \tag{4.40}$$

Else, if some criterion i is a "favour" one, then

$$\nu(A) = 1 \qquad \forall A \supseteq i. \tag{4.41}$$

Table 4.2: Modelling criteria interactions

Criteria i and j complement each other	$0 \leqslant I_\nu(i,j) \leqslant 1$
Criteria i and j complement each other stronger than k and l	$I_\nu(i,j) - I_\nu(k,l) \geqslant \delta_I$
Criteria i and j interact in a way similar to k and l	$-\delta_I \geqslant I_\nu(i,j) - I_\nu(k,l) \leqslant \delta_I$

Finally, if the problem allows us to employ a learning set, the DM might be asked to express his preferences with regard to its elements. In an identification problem this corresponds to linear constraints (since the integral is linear in ν) outlined in Table 4.3.

Table 4.3: Preferences over learning set objects

The alternative z_1 is preferred to z_2	$C(\nu, f(z_1)) - C(\nu, f(z_2)) \geqslant \delta_{LS}$
The DM is indifferent between z_1 and z_2	$-\delta_{LS} \geqslant C(\nu, f(z_1)) - C(\nu, f(z_2)) \leqslant \delta_{LS}$

Having the available information expressed as a set of linear constraints we obtain the set $\mathcal{U}$. Summing up the results of the previous section, $\mathcal{U}$ can be written down as shown in equation (4.42).

Notably, all constraints are linear, and thus the set $\mathcal{U}$ is a polyhedron in $\mathbb{R}_+^{2^n}$. Its dimension can be reduced to $2^n - 2$ if we exclude the $\emptyset$ and N coordinates, which have fixed values. It can be reduced even further by using k-additive capacities which, however, is not always possible. By solving the feasibility problem

$$\begin{aligned} &\min_\nu 1 \\ \text{s.t. } &\nu \in \mathcal{U}, \end{aligned} \tag{4.43}$$

we can check if there exists at least one capacity compliant with the given data. If such capacity cannot be found, the following problem can be solved:

$$\begin{aligned} &\min_\nu \mathcal{L}(\mathcal{U}) \\ \text{s.t. } &\nu \text{ is a capacity}, \end{aligned} \tag{4.44}$$

$\mathcal{U}:$

$$
\begin{aligned}
&\textbf{Information from the DM}\\
&\phi_\nu(i) - \phi_\nu(j) \geqslant \delta_{SH}, \quad i, j \in 1, \dots, n\\
&\dots\\
&-\delta_{SH} \geqslant \phi_\nu(i) - \phi_\nu(j) \leqslant \delta_{SH}, \quad i, j \in 1, \dots, n\\
&\dots\\
&I_\nu(i,j) - I_\nu(k,l) \geqslant \delta_I, \quad i, j \in 1, \dots, n\\
&\dots\\
&-\delta_I \geqslant I_\nu(i,j) - I_\nu(k,l) \leqslant \delta_I, \quad i, j \in 1, \dots, n\\
&\dots\\
&C(\nu, f(z_i)) - C(\nu, f(z_j)) \geqslant \delta_{LS}, \quad i, j \in 1, \dots, n\\
&\dots\\
&-\delta_{LS} \geqslant C(\nu, f(z_i)) - C(\nu, f(z_j)) \leqslant \delta_{LS}, \quad i, j \in 1, \dots, n\\
&\dots\\
&\nu(A) = 1, \forall A \supset \text{favour criteria}\\
&\nu(A) = 0, \forall A \not\supset \text{veto criteria}\\
&\textbf{Technical constraints}\\
&\nu(\emptyset) = 0\\
&\nu(N) = 1\\
&\nu(B) \geq \nu(A) \quad \forall B \subset A \subset N\\
&\textbf{Additional constraints}\\
&k - \text{additivity. Not always applicable.}
\end{aligned}
\tag{4.42}
$$

Figure 4.1: Encoding the information as constraints on the set of capacities

where $\mathcal{L}(\mathcal{U})$ is some loss function of the data (e.g. the number of preference reversals). The loss function, whether an error-based one or some other as mentioned above, is typically a convex function, so the optimization problem is quite efficient. If the model is built for forecasting purposes, regularization techniques can also be used (Tehrani and Huellermeier, 2013; Tehrani et al., 2012; Tehrani, Cheng, Dembczyński and Hüllermeier, 2011; Tehrani, Cheng and Hüllermeier, 2011). Additionally, identification problems can have more than one solution, which induces the problem discussed below.

4.3.2 Learning the value functions

Learning the value functions on the other hand is a different matter. Let us consider first how the process is performed in the additive value function model. Recall that the model has the following form:

$$x \succcurlyeq y \iff \sum_{i=1}^{n} f_i(x_i) \geq \sum_{i=1}^{n} f_i(y_i). \tag{4.45}$$

The data in such a learning problem is typically given as pairwise preferences for some points from the set X. The resulting problem is then an LP, because additive value functions are linear with respect to each f_i that we are aiming to learn. A well-known family of learning methods related to learning of the additive value models are called the "UTA methods" (Siskos et al., 2005). The value functions are assumed to be linear interpolations of the learning points (i.e. they are piecewise linear), but sometimes polynomial or spline-based versions are used (Sobrie et al., 2016). Still, the process remains computationally efficient.

Note that the value functions learned in this manner do not provide any "qualitative" information about the data to the analyst. They can be used for forecasting purposes, but due to the restrictions of the additive model, no statements about the "importance" of criteria or similar notions can be made. In contrast, learning value functions and the capacity in the Choquet integral is valuable even if the value functions are learned in a non-parametric manner. Indeed, it is the capacity that is capable of showing the qualitative relations between criteria of the multidimensional problem, as is to some extent attested by the majority of the existing practical applications. However, this process has two complications: the computational complexity and the confounding of the capacity and the value functions.

As mentioned above, the vast majority of the theoretical and practical contributions to the literature assume the existence of value functions, or what is the same, of a common scale on which all attributes of the problem are being measured. This is clearly a very strong assumption, but it also leads to a significant simplification of the learning process. Indeed, in this case we only need to learn the capacity, which is generally a convex minimization problem. In contrast, when learning *both* the capacity and the value functions, we must solve a difficult non-convex optimization problem. Only a few papers have attempted to tackle this issue (Angilella et al., 2004; Goujon and Labreuche, 2013; Angilella et al., 2015), all of them offering some heuristic methods and small-scale examples. This is not surprising. Indeed, consider the data point $x \succcurlyeq y$ for some $x, y \in X$. In the Choquet integral model, it is represented by the following expression: $C(\nu, f(x)) \geq C(\nu, f(y))$. Since the integral is a sum of products of elements of ν and $f(x)$, the constraint is not linear in contrast to the case where only capacity is considered unknown. Moreover, it is generally non-convex. Hence, the process of construction of the capacity and the value functions involves solving a non-convex optimization problem, which is known to be computationally hard.

4.3.3 Confounding of the capacity and the value functions

The second issue in the Choquet integral learning problems is the non-uniqueness of the resulting capacity. Even in cases where only capacity is being learned, the exponential number of the coefficients ($2^n - 2$, excluding $\nu(\emptyset)$ and $\nu(N)$) means that the task of model learning quickly becomes very difficult as the number of dimensions of the model increases. Typically a learning dataset which is not sufficiently large does not allow the capacity to be learned in a precise way. This is a very well-known problem in general learning theory (Hüllermeier and Tehrani, 2012) and it can be addressed by a number of methods. Among these we can mention the general regularization approaches (Tehrani and Huellermeier, 2013; Tehrani et al., 2012; Tehrani, Cheng, Dembczyński and Hüllermeier, 2011; Tehrani, Cheng and Hüllermeier, 2011), but also some specialized methods which can be applied when the model is used in particular applications, such as sorting (Angilella et al., 2015, 2010). Additionally, a number of methods were developed for robust decision making with the Choquet integral. Thus, in Timonin (2013) we proposed an algorithm for regret-minimizing optimization when the capacities are only known to belong to a certain set, whereas Benabbou et al. (2015, 2014) looked at the problem of the robust capacity construction using interactive data.

Axiomatization introduced in this work adds another level of complexity to the uniqueness problem. Indeed, the uniqueness results state that meaningful and unique decomposition of the capacity and the value functions is only possible when the model exhibits sufficient levels of non-separability. In particular, pairwise violation of ij-triple cancellation should be present to a sufficient extent to obtain a unique capacity (in particular, all variables should be in the same interaction group, see Section 3.9). Thus, even an indefinite amount of data, not containing a sufficiently rich structure of preferences, would lead to a strongly non-unique capacity. In fact, it is easy to show that the capacity in such cases can be taken almost arbitrarily. Consider the extreme example, when there is no pairwise interaction in the model. In this case, we have n interaction groups of size 1 or, in other words, an additive value model. In the expression $w_1 f_1(x_1) + \cdots + w_n f_n(x_n)$ we can arbitrarily change the "weights" w_i by compensating their increase or decrease by a proportional change in f_i. The whole model can then be rescaled so that the weights sum up to 1. It is trivial that these modifications do not affect the validity of the representations.

Non-uniqueness of the capacity is not necessarily a problem for prediction applications; however, qualitative conclusions, commonly made based on capacity indices, become meaningless. For example, consider the paper of Li et al. (2012). Here, data from hotel evaluations on the tripadvisor website is analysed with the Choquet integral. Each hotel is reviewed based on several criteria, such as price, location, etc. In addition, every hotel gets an overall mark, which allows the authors to construct the relation between general attractiveness of the hotel and its particular features or their combinations. Reviewers are categorized into several social groups ("American businessmen", "European families", etc). The paper shows which attributes and combinations of attributes are important for every group by finding capacities that provide the best fit of the 2-additive Choquet integral to the corresponding dataset. Shapley values and interaction indices of these capacities provide the required information.

From our perspective, the important point is that the evaluations are *assumed* to be on the same scale. Every criterion is given from one to five stars, and so is the global evaluation. Of course it seems not completely unreasonable to suggest that various incommensurable notions such as "5 minutes from the train station" and "very clean" are somehow mapped onto a global "satisfaction" scale in the mind of the reviewer, indeed there are many examples of such "cross-modality" mappings in the psychological literature (see Section 4.4.2). However, there is no real evidence supporting this claim, and we can also assume that stars on each dimension signify just the ranking within the dimension itself and not across dimensions as the authors conjecture. The other consequence of such assumption is that the scale is equispaced, in the sense that the (cardinal) difference between one and two stars is the same as between two and three and between four and five. Apparently, this does not have to be true and often is not.

The possibility to fit not only the capacity but also the value functions resolves these methodological issues. Apparently, it should also improve the quality of the fit. However, in cases when we assume a common scale, the lack of interaction between certain criteria is not an issue – we still obtain a unique capacity (see also axiomatizations in Wakker, 1989 and Schmeidler, 1989) and corresponding indices, which would show a lack of interaction. In contrast, without the commensurability assumption, having even two interaction groups would mean that we are not able to talk about "criteria importance" globally, but only within these groups. The problem here is not with the tools used for capacity interpretation, in this case the Shapley value, but rather comes from the limitations of the model per se. Unfortunately, it is not easy to see how this problem can be resolved, as it is in fact the same issue as the impossibility of meaningfully using the notion of "criteria weights" in the additive model (Bouyssou et al., 2000) (Chapter 6). It is notable, however, that the value of the interaction index would remain zero for any two elements from different interaction groups, no matter how we transform the capacity![5]

4.4 Interpretations and discussion

Motivation for this thesis came primarily from MCDA applications. However, our results can be also applied in several other subfields of decision theory. In this section we discuss two of them – the state-dependent utility and the social

[5]See Lemma 31 and the definition of the interaction index given earlier in this chapter.

choice problems.

4.4.1 Multicriteria decision analysis

MCDA provides perhaps the most natural context for our results. Indeed, in the multicriteria context the heterogeneity of the decision space dimensions is natural and the insufficiency of the previous results is apparent and has been discussed in the literature multiple times (e.g. Bouyssou et al., 2009). We have covered many aspects of the Choquet integral usage in MCDA in the previous chapters. An introduction and an example of a multicriteria model are given in Section 1.2.2, while questions of the model learning and interpretation are discussed in Section 4.3, together with an example of a practical application.

From the theoretical perspective, in the multicriteria context our results imply that the decision maker constructs a mapping between the elements of the criteria sets (their subsets to be precise). Some authors interpret this by saying that criteria elements sharing the same utility values present the same level of "satisfaction" for the decision maker (Grabisch and Labreuche, 2008). Technically, such statements are meaningful, in the sense that permissible scale transformations do not render them ambiguous or incorrect, unless the representation is additive. However, the substance of the statements such as "x_1 on criterion 1 is at least as good as x_2 on criterion 2" (which would correspond to $f_1(x_1) \geq f_2(x_2)$) is not easy to grasp. Apart from the satisfaction interpretation, perhaps one could think about workers performing various tasks within a single project. From the perspective of a project manager, achievements of various workers, serving as criteria in this example, can be level-comparable despite being physically different, if the project has global milestones (i.e. scale) which are mapped to certain personal milestones for every involved person. The novelty of our characterization is that this scale is not given a priori. Instead, we only observe preferences of the project manager and infer all corresponding mappings from them. It is also worth mentioning that value functions for any interacting pair can be seen to form a so-called Guttman scale (or a biorder) (Guttman, 1944; Doignon et al., 1984).

4.4.2 Psychology

An interesting connection is that in psychology there exists a body of results on the so-called cross-modality matching. A large number of studies have been conducted in this area since 1950s, with experiments related to loudness, colour, size, tone, pain, money, etc. (Stevens and Marks, 1980, 1965; Stevens, 1959; Galanter and Pliner, 1974; Krantz, 1972). Kahneman (2011) gives the following example: "A girl learned to read when she was four. How tall is a man who is as tall as Julie was precocious?" Normally, kids start reading at around 5 or 6, so perhaps the girl is somewhat more precocious than average, although not by too much. Therefore, we could say that the man is somewhat higher than the average 180 cm, perhaps his height is 190 or similar. Apparently our ability to answer this question is based on the existence of some information about the distribution of the age when children start reading, and the distribution of height. The information can come in a number of forms: either just a mean value ("on average kids start reading at 5", "an average man is 180 cm high"), or two absolute reference levels on both dimensions – "children start reading between 3 and 6", "men heights are in the range of 165–205 cm". Finally, we can have complete information about both distributions and pick a match based on that. It is this information that allows us to "map" four years to something like 190 cm. We can perhaps consider the probability of a certain value as the universal scale shared by two distinct elements: "75% of children start reading after 4", "75% of men are lower than 190 cm", etc. However, as discussed above, such information is not always available, and there might be other mechanisms by which such mappings are performed.

4.4.3 State-dependent utility

We will show how the traditional comonotonic-based axiomatization *implies* state-independence and how our approach can be used to construct a truly state-dependent model without making additional assumptions about correspondence between outcomes in different states.

The state-dependent utility concept, as introduced in Chapter 1 and further in Appendix A, is evoked when the nature of the state itself is of significance and it is not assumed that outcomes in different states have the same meaning or value to the decision maker. A popular example is healthcare, where various outcomes can have major effects on the personal value of the insurance premium (Karni, 1985). One way to model this is to use different value functions for every state; moreover, we could also consider the notion of state–prize (Karni, 1985; Karni and Schmeidler, 2016), which actually takes the state-dependent model directly to the heterogeneous product set case (dimensions are sets of "state–prizes").

So far the axiomatizations of the state-dependent utility models have been based on the existence of some correspondence between the outcomes in different states (Karni and Schmeidler, 2016; Karni, 1993, 1985; Fishburn, 1973). In essence, this is not different from assuming the homogeneous product set again, albeit with some technical differences (e.g. the decision space might only be a subset of the full product). Although, in principle, the existence of a preference relation on the set of state–prizes is not unrealistic, it is not clear whether this data is observable (contrary to the preferences on acts which are supposed to be always observable). Without such a relation the additive value model (think SD-EU) does not allow us to disentangle probabilities and utilities at all (see discussion in the previous section and earlier). The other question is whether this gives any real methodological advantage compared to using a union of state–prizes on every dimension and proceeding as normal. A detailed discussion of this question is given in Karni and Schmeidler (2016) and references therein, and we do not pursue it further here. Finally, we would like to mention that the problem of state-dependence in rank-dependent models is not well developed – the only paper known to the author being Chew and Wakker (1996), where the authors comment on the meaninglessness of state-dependency in the normal CEU framework, again due to the confounding issues: "with preferences over acts as the only empirical primitive, the factorization $\nu(A)u_A(\cdot)$ becomes meaningless. Only the product $W(x, A) = \nu(A)u_A(x)$ can be derived from preferences".

However, the general axiomatization of the Choquet integral presented in this thesis, is the first (to the author's best knowledge) result where state-dependence can be derived *exclusively* from the preferences over acts. This constitutes a significant difference with all earlier results. As a side result, it is easy to show that comonotonicity-based conditions actually *imply* state-independence of preferences.

Lemma 48. *Let $X = Y^n$. Let conditions of the Theorem 8 hold. If for all $x \in X$ we have $i\,\mathbf{E}^x\,j$ whenever $x_i = x_j$, the representation is state-independent.*

Proof. Saying that $i\,\mathbf{E}\,j$ whenever $x_i = x_j$ in our framework amounts to saying that additive representations exist on the comonotonic subsets of X. The construction implies that $f_i(x_i) = f_j(x_j)$ whenever $x_i = x_j$. This holds for all $i, j \in X$, hence we can use a single utility function $U : Y \to \mathbb{R}$ for all dimensions. This constitutes state-independency. $\square$

Hence, parting with the assumption that the borders between additive regions actually coincide with the borders between *comonotonic* sets, allows us to introduce state-dependency into the model and to do so solely by observing the preferences between acts. The resulting state-dependent utility functions could be used to derive the relation on the set of state–prizes which is assumed as given in earlier works. Note that, as previously, the meaningfulness of this relation is conditional on the violation of pairwise separability in the model, as explained in Section 3.9. In other words, the relation might not exist between prizes of certain state pairs.

4.4.4 Social choice

If we think of the set N as of a society with n agents, then X is the set of all possible welfare distributions. Moreover, contrary to the classical scenario, agents could be receiving completely different goods, for example X_1 might correspond to healthcare options, whereas X_2 to various educational possibilities. In this case it is not a trivial task to build a correspondence between different options across agents. Our result basically states that provided the preferences of the social planner abide by the axioms given in Section 3.2, the decisions are made as if the social planner has associated cardinal utilities with the outcomes of each agent which are *unit* and *level* comparable (cardinal fully comparable or CFC in terms of Roberts (1980)). Such approach is not conventional in social choice problems, where the global (social) ordering is usually not considered as given (there are, however some papers taking this route, e.g. Ben-Porath et al., 1997). Instead, the conditions are normally given on individual utility functions and the "aggregating" functional that is used to derive the global ordering. However, one of the important questions in social choice literature is that of the interpersonal utility comparability and whether it is justifiable to assume it or not (e.g. Harsanyi, 1980). Our results show that if the global ordering of alternatives made by the society (or the social planner) satisfy certain conditions, it is in principle possible to have individual preferences represented by utility functions that are not only unit but also level comparable with each other.

4.5 Summary

We have presented extensions of our characterization for the ordinal and cardinal special cases of the Choquet integral. The ordinal models are the well-known MIN/MAX and the order statistic, and also their generalization – the lattice polynomial. We have shown how these can be characterized in our framework and also related our results to the previously known axiomatizations. On the cardinal side of things, we have shown how it is possible to characterize the Choquet integral with respect to a convex capacity. The axiom is similar to the tradeoff consistency condition and is the first characterization of convex models which can deal with both cardinal and ordinal cases (or a mixture of the two).

Next, we discussed various aspects of the Choquet integral learning. Traditionally, the learning of the integral was confined to capacity learning only. However, this approach suffers from serious methodological difficulties. Namely, it requires a very strong assumption that all criteria are measured on the same scale. We looked at how various preferential information could be used in the capacity identification problem and analysed why the process of capacity identification is relatively computationally effective. In contrast, learning the capacity and the value functions together seems to be computationally very hard. There have been only a few attempts at solving it in the literature, all of them offering only some heuristic methods. Finally, we look at the problem of confounding of the value functions and the capacity. Our characterization results state that a unique decoupling of the capacity and the value functions is possible only when the dimensions of the decision space exhibit sufficient pairwise interaction. This has a profound impact on the learning properties of the Choquet integral, since it guarantees that it is impossible to obtain a unique capacity if the variables are not interacting enough, no matter how much data we have. This means that the usage of the well-known indices such as the Shapley index is limited. An alternative option is to use the "sum of Choquet" representation (3.7), whereby the indices become meaningful within each interaction group.

Finally, we have looked at various interpretations of our results and their applications in decision theory. We started with MCDA, which was the main inspiration for our research. Our axiomatization is a long-missing result in this area and we hope that it will help promote further theoretical research of the Choquet integral in MCDA. The characterization leads to construction of a unique mapping between elements of various criteria sets (dimensions of the decision space). This has interesting connections to the question of cross-modality mapping, which has been extensively studied in psychology since the 1950s. Finally, we discussed two other areas where our results can be applied – the social choice theory and the state-dependent DUU. The latter is especially interesting, as our characterization is the first to construct a meaningful state-dependent model based solely on the preferences among acts. Previous works introduced additional preference relations into the model, in particular the relation on the set of "state–prizes". Conceptually, this amounts to saying that elements of various dimensions are commensurate which does not always have to be the case. Observability of this preference relation is also not apparent. Our results do not require any additional constructs apart from the preference between acts themselves. Yet, we are able to construct a unique mapping between the outcomes in different states (provided the data exhibits sufficient interaction).

Conclusion

Axiomatic analysis is important for decision models as it allows for a much deeper understanding of their mechanics. Sometimes developing an axiomatic system for a well-known model also helps to elicit its previously unknown properties. In this thesis we have presented a characterization of the Choquet integral for the general case of a heterogeneous decision space. Although the Choquet integral has been widely used since the late 1980s, its theoretical foundations were developed well only in decision under uncertainty. Transferring these results to multicriteria decision analysis proved to be a difficult task, which was quite frustrating, as the integral is widely known and used in theoretical and applied MCDA works. Our results aim to fill this gap.

The key results of this thesis are as follows.

Characterization of the Choquet integral for the general case of heterogeneous products sets. Whereas early characterizations only allowed for a very special case of a homogeneous decision space $X = Y^n$, we here considered a fully general $X = X_1 \times \cdots \times X_n$. This has allowed us not only to deal with cases where the dimensions of the model are inherently different, such as different attributes in MCDA models (e.g. colour, price, speed, etc), but also to work very well in the state-dependent utility and social choice settings.

A novel axiomatic system which generalizes earlier results based on comonotonicity. Our axioms are not based on the notion of a constant act or comonotonicity which are not-definable in the heterogeneous setting. Although seemingly quite different from the previous conditions, our axioms actually generalize them. In the homogeneous setting this allows us to deviate from comonotonic additivity and hence from state-independence (more on this below). Our system also parts with the symmetry of the older characterizations, allowing for a much wider range of decision spaces.

Analysis of the model uniqueness properties which turned out to be weaker than previously thought. The increased generality discussed above caused the uniqueness properties of the model to weaken considerably. We showed how uniqueness "strength" became conditional on the amount of interaction (i.e.

on the lack of separability). Our results allow for a full spectrum of models, from the case with a single unique capacity, to the case of an additive value model, which can be represented by a Choquet integral with respect to an (almost) arbitrary capacity.

Extensions of the main characterization result to the ordinal and cardinal special cases of the Choquet integral. We have considered MIN/MAX, order statistic, lattice polynomial, and the case of a convex capacity which in the two-dimensional case is equivalent to the Gilboa–Schmeidler maximin model.

Analysis of the implications of our results for learning of the Choquet integral models. Traditionally, learning of the Choquet integral meant learning of the capacity only, whereas the value functions were assumed to be given. This requires making a strong assumption of dimension commensurability which is not always desirable (especially in MCDA). There were only a few attempts to learn the capacity and the value functions simultaneously. We discussed the difficulties these approaches faced and also the problems caused by the weaker uniqueness properties of the model. In certain cases, when the data does not exhibit enough interaction between dimensions we cannot expect to get a unique capacity no matter how much data we have access to.

Discussion of our results in relation to psychology, social choice, and state-dependent utility theory. One interesting aspect of our characterization is that it allows us to construct a unique mapping between elements of various dimension sets. In the context of state-dependent utility theory this is the state–prize relation, for the first time obtained based solely on the basis of preferences between acts. We also showed that previous Choquet integral characterizations actually implied state-independence. In the MCDA context, this mapping signifies the commensurability of various seemingly incomparable criteria, such as colour and loudness or money. This has interesting connections to psychology where experiments related to the so-called "cross-modality" matching were performed starting from the 1950s. This provides an interesting interpretation of our results in the multicriteria context.

Appendix

Appendix A

Extended literature review

A.1 Decisions under uncertainty

Schmeidler (1989) was the first to provide an axiomatization of the Choquet integral in the context of the decision making under uncertainty. The axiomatization is done in the so-called "Anscombe–Aumann" framework, where the set of *outcomes* Y is composed of lotteries with known probabilities over some set. The model also includes the set of *states of the world* S, and *acts* are defined as a convex subset L of measurable functions from Y to S, which includes constant functions. The main condition in this axiomatization is comonotonic independence, which is defined as follows.

Comonotonic independence For all pairwise-comonotonic acts f, g and h in L and all $\alpha \in (0, 1)$: $f \succ g$ implies $\alpha f + (1 - \alpha)h \succ \alpha g + (1 - \alpha)h$.

Comonotonicity is defined with respect to an ordering of outcomes which is obtained from ordering of constant acts. Usage of this axiom obviously requires the set of outcomes to be a homogeneous Cartesian product, so that the notion of a constant act can be used. One other restriction of this axiomatization is that it *implies* state-independence (more on this in Section 4.4.3).

While Schmeidler's axiomatization is a comonotonic weakening of Anscombe–Aumann's work, several authors proposed a similar extension in Savage's framework (i.e. without using lotteries). The first one was Gilboa (1987), but here we provide a condition from Chew and Wakker (1996). Their main axiom is a comonotonic version of the sure-thing principle. Let S be the set of acts as previously, and Y is a set of outcomes (now it is not assumed that outcomes are lotteries). Acts again are functions from S to Y. The notation $x_A f$ means that the act obtains value x on A and f on $S \setminus A$.

Comonotonic Sure-Thing Principle For all comonotonic $x_A f, x_A g, y_A f, y_A g$ we have $x_A f \succcurlyeq x_A g \iff y_A f \succcurlyeq y_A g$.

Another notable axiomatization of the Choquet integral for decisions under uncertainty is due to Wakker (1991a, 1989). In these works the set of states S is finite, and the decision space is accordingly a Cartesian product $X = Y^n$. Wakker

(1989) gives the result in the topological framework, whereas in Wakker (1991a) the framework is algebraic. The key condition in both cases is "non-revelation of comonotonic-contradictory tradeoffs", which is defined as follows:

Non-revelation of comonotonic-contradictory tradeoffs For all $\alpha, \beta, \gamma, \delta \in Y$ and $x_{-i}, y_{-i} \in X_{-i}$ and $v_{-j}, w_{-j} \in X_{-j}$, such that $x_{-i}\alpha, y_{-i}\beta, x_{-i}\gamma, y_{-i}\delta$ are contained in a comonotonic set on which state i is essential, and also $v_{-j}\alpha, w_{-j}\beta, v_{-j}\gamma, w_{-j}\delta$ are comonotonic, we have

$$\left.\begin{array}{l} x_{-i}\alpha \preccurlyeq y_{-i}\beta \\ x_{-i}\gamma \succcurlyeq y_{-i}\delta \\ v_{-j}\alpha \succcurlyeq w_{-j}\beta \end{array}\right\} \Rightarrow v_{-j}\gamma \succcurlyeq w_{-j}\delta.$$

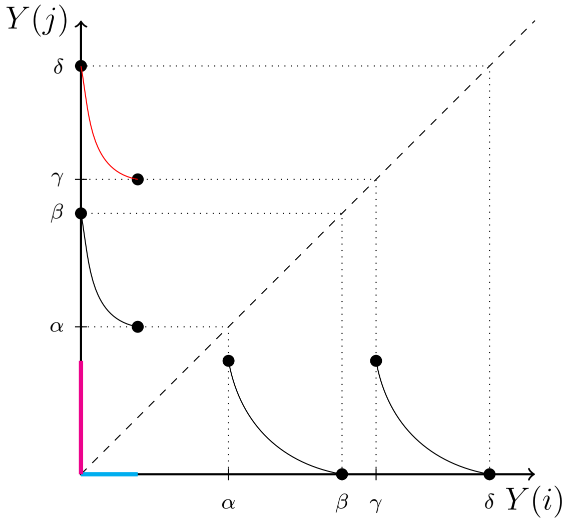


Figure A.1: Non-revelation of comonotonic-contradictory tradeoffs (with equivalences)

The papers rely on another result by Wakker (1991b), which shows how additive representations can be constructed on rank-additive sets. The approach was later generalized in Köbberling and Wakker (2003). These results suffer from the same restrictions:

1. the construction requires comonotonicity, hence homogeneity of the decision space $X = Y^n$;

2. the construction implies that outcomes in all states have the same utility (state-independence).

An excellent review of other axiomatizations of the Choquet integral in decision making under uncertainty can be found in Köbberling and Wakker (2003). Here we briefly mention some of the conditions that are used in these works. The main properties remain unchanged: characterizations rely on comonotonicity, homogeneity of the decision space, and imply state-independence.

Multiple conditions, such as the act-independence of Chew and Karni (1994), are based on the existence if a so-called certainty-equivalent.

Comonotonic act-independence For any n-partition Π, $B \in \Omega$, and $x, x', y \in X^n$ such that $x_i, x'_i \succcurlyeq y_i$ for all i (or $y_i \succcurlyeq x_i, x'_i$ for all i),

$$f_\Pi(x) \succcurlyeq f_\Pi(x') \Rightarrow f_\Pi(m_B(x_1, y_1), \ldots, m_B(x_n, y_n)) \succcurlyeq f_\Pi(m_B(x'_1, y_1), \ldots, m_B(x'_n, y_n)).$$

Visually it is clear that the condition is quite similar to comonotonic independence. Here too, x and x' are "mixed" with a common component y.

One other group of results includes axiomatizations related to state-dependent utility with the Choquet integral. They are of particular importance in our context, as mathematically the setting of a state-dependent model is quite close to the general case that is presented in this thesis. There are not too many works on state-dependency applied to the Choquet integral. The concept proved to be a complicated problem even in the additive setting (Karni, 1985). If we take a look at a state-dependent additive functional, it can be seen that it reduces to the additive value model, and hence from the preference over acts alone it is impossible to define the probability in the unique form. In other words, "beliefs and outcomes are confounded" (Karni, 1993; Chew and Wakker, 1996). In case of just two states:

$$x \succcurlyeq y \iff p_1 u_1(x_1) + p_2 u_2(x_2) \geq p_1 u_1(y_1) + p_2 u_2(y_2).$$

A simple change of variables allows us to rewrite this in the additive value form:

$$x \succcurlyeq y \iff v_1(x_1) + v_2(x_2) \geq v_1(y_1) + v_2(y_2).$$

Yet another change of variables gives a state-dependent form again, but with different probabilities:

$$x \succcurlyeq y \iff p_1' u_1'(x_1) + p_2' u_2'(x_2) \geq p_1' u_1'(y_1) + p_2' u_2'(y_2).$$

Chew and Wakker (1996) axiomatize the so-called cumulative utility (CU) functional, which they suggest can be viewed as "event-dependent" utility, a "natural rank-dependent analogue of state-dependent utility". The condition they use is the comonotonic sure-thing principle which was discussed above. However, they note that the in the case of the Choquet integral the factorization $W(x, A) = \nu(A)u_A(x)$ (where ν is a capacity) is meaningless, as only the product can be derived from the preferences. This is essentially the same problem as in the additive value model example given above.

Two techniques proposed in the literature are to assume an existence of an additional preference relation on the "state–prize" space (Karni, 1985; Karni and Schmeidler, 2016), and also to construct a mapping between outcomes in various states (Karni, 1993). In contrast, our characterization allows for a *meaningful*, i.e. unique, factorization of the state-dependent utility functions into utilities and capacities. More details on this are given in the corresponding chapters.

A.2 MCDA

The earliest characterization in MCDA, very similar to Schmeidler's result but with a much simpler proof due to the atoms of the model, can be found in De Campos and Bolanos (1992). Let $I = [0, 1]$ and $F : I^n \to I$. Then there exists a unique normalized capacity μ such that $F = C_\mu$ if and only if F fulfils the following properties:

1. comonotonic additivity, i.e. $F(x + x') = F(x) + F(x')$,

2. nondecreasing monotonicity,

3. $F(\mathbf{1}_{[n]}) = 1, F(\mathbf{0}) = 0$.

Moreover, μ is defined by $\mu(A) := F(\mathbf{1}_A)$.

Marichal (2002) proposed a characterization based on four properties. Let $I \subset \mathbf{R}, I \ni 0$. A parametric function F_μ is a Choquet integral if:

1. F_μ is linear expression of μ, i.e. exist 2^n functions $f_A : I^n \to I$, such that $F_\mu = \sum_{A \subset N} \mu(A) f_A$.
2. F_μ is non-decreasing.
3. F_μ is interval scale invariant, i.e. $F_\mu(rx_1+s, \ldots, rx_n+s) = rF_\mu(x_1, \ldots, x_n)+s$.
4. F_μ is an extension of μ: $F_\mu(\mathbf{1}_A) = \mu(A)$ for all $A \subset N$.

Couceiro and Marichal (2011) showed that the Choquet integral can be characterized by any of the three conditions below:

Comonotonic additivity A function $f : I^n \to \mathbb{R}$ is *comonotonically additive* if, for every comonotonic $x, x' \in I^n$ such that $x + x' \in I^n$, we have $f(x + x') = f(x) + f(x')$.

Given $x \in I^n$ and $c \in I$, let $[x]_c = x - x \wedge c$ and $[x]^c = x - x \vee c$.

Horizontal min-additivity We say that a function $f : I^n \to \mathbb{R}$ is *horizontally min-additive* if, for every $x \in I^n$ and every $c \in I$ such that $[x]_c \in I^n$, we have $f(x) = f(x \wedge c) + f([x]_c)$.

Horizontal max-additivity We say that a function $f : I^n \to \mathbb{R}$ is *horizontally max-additive* if, for every $x \in I^n$ and every $c \in I$ such that $[x]^c \in I^n$, we have $f(x) = f(x \vee c) + f([x]^c)$.

Labreuche and Grabisch (2003) provide a constructive approach to building value functions and constructing the capacity based on the information from the decision maker. The paper builds on the axiomatization of Marichal (2002), weakening one of the conditions. However, it also supposes the existence of two absolute reference levels on each attribute which are supposed to be commensurate. Mathematically, the approach is based on the MACBETH method (Bana E Costa and Vansnick, 1994), which the authors extend to the Choquet integral.

Note, that the idea of two absolute reference levels on each attribute is connected to cross-modality matching, discussed in Section 4.4.2.

Finally, a very interesting approach to axiomatizing the Choquet integral was proposed by Labreuche (2012). The author operates in the general MCDA framework, so no assumption about commensurability of the dimensions is made. However, the axioms that are given are not conditions on the preference relation itself, but rather on the "global value" function F, and value functions f_i. The main axiom can be explained as follows. Assume that we have values on all dimensions apart from one fixed. Then, we start changing the value of this dimension and look at how this affects the partial derivative of F with respect to *some other* dimension. It turns out that this derivative can take either one value for all possible levels of the criterion that we change, or it can take one value below a certain level and another value above it.

The axiom is called "Commensurability through Interaction" and it is insightful in many aspects, even though it is not a "proper" condition on the preference relation. It tells us that eventual commensurability of different dimensions is the consequence of the lack of independence, moreover the point where the violation of independence happens (i.e. the partial derivative of the global utility changes) can and should be used as an anchor to construct the mapping between levels of two different dimensions. Notably, this condition does not require anything resembling comonotonicity.

A.3 Characterizations of the Sugeno integral

The earliest paper on the subject is probably De Campos and Bolanos (1992), which used the following conditions. Let $I = [0, 1]$ and $F : I^n \to I$.

1. Comonotonic maxitivity – $F(x \vee x') = \max(F(x), F(x'))$.

2. $F(\mathbf{1}_N) = 1$.

3. $\wedge$-homogeneity – for any $\alpha \in (0, 1]$, we have $F(\alpha \wedge x) = \alpha \wedge F(x)$.

The conditions are quite different from the conditions used in the characterizations of the Choquet integral. However, they still use the notion of comonotonicity. Generally speaking, the integral is max- and min-decomposable for comonotonic acts, that is, for all comonotonic f, g we have $S(f \wedge g) = \min(S(f), S(g))$ and $S(f \vee g) = \max(S(f), S(g))$.

The second characterization is due to Marichal (2000). It also relies on three properties:

- nondecreasingness
- $\wedge$-homogeneity
- $\vee$-homogeneity.

In Dubois et al. (1998) the integral was characterized in the DUU context, using the framework of Savage. Two main conditions in the characterization are

RCD – Restrictive Conjunctive Dominance Let f and g be any two acts and y be a constant act. Then, $[g \succ f]$ AND $[y \succ f] \Rightarrow [g \wedge y \succ f]$.

RDD – Restrictive Disjunctive Dominance Restricted max-dominance: Let f and g be any two acts and y be a constant act. Then, $[f \succ g]$ AND $[f \succ y] \Rightarrow [f \succ g \vee y]$.

The conditions are quite different from the ones used in the characterizations given above and those of the Choquet integral. However, they still use the notion of a "constant act".

Conjoint axiomatization of the Sugeno integral is the first result which gives an axiomatization of a rank-dependent model with heterogeneous decision space. To date we are only aware of three papers which have achieved this. We look at the results in detail, as they are helpful in establishing the correspondence between DUU-type axiomatics and the ones of MCDA-type.

The first characterization of the discrete Sugeno integral, given in Greco et al. (2004) and extended and proved in Bouyssou et al. (2009) relies on the following condition. $\succcurlyeq$ is **strongly 2-graded** if for all $i \in N$:

$$\left.\begin{array}{r} x_i a_{-i} \succcurlyeq w \\ y_i b_{-i} \succcurlyeq t \\ w \succcurlyeq t \end{array}\right\} \Rightarrow \left\{\begin{array}{c} z_i a_{-i} \succcurlyeq w \\ or \\ x_i b_{-i} \succcurlyeq t. \end{array}\right.$$

The interpretation of the condition is as follows. Consider, for example, the special case of $w = t$. Suppose also that $NOT[x_i b_{-i} \succsim t]$. We have $y_i b_{-i} \succsim t$ and $NOT[x_i b_{-i} \succsim t]$, which suggests that the level x_i is worse than y_i with respect to t. So we have that $x_i a_{-i} \succsim t$ implies that also $z_i a_{-i} \succsim t$ for all z_i. What this means is that once we have established that a level y_i is better than some other level x_i with respect to an alternative t, we should not be able to find any element in X_i that is worse than x_i, therefore $x_i a_{-i} \succsim t$ implies that for any possible $z_i \in X_i$ we would still have $z_i a_{-i} \succsim t$. Informally, we can say that each $t \in X$ partitions each set X_i into "satisfactory" and "unsatisfactory" levels. The condition then says that these partitions are related for different w and t from X.

In Bouyssou et al. (2009) the axiom is further decomposed into two separate conditions (assuming $\succsim$ is a weak order):

A2 – weakest separability

$$[a_i x_{-i} \succ b_i x_{-i}] \Rightarrow [a_i y_{-i} \succsim b_i y_{-i}]$$

2-graded For all $i \in N$

$$\left.\begin{array}{r} (x_i, a_{-i}) \succsim w \\ (y_i, a_{-i}) \succsim w \\ (y_i, b_{-i}) \succsim t \\ w \succsim t \end{array}\right\} \Rightarrow \left\{\begin{array}{c} (z_i, a_{-i} \succsim w) \\ or \\ (x_i, b_{-i} \succsim t). \end{array}\right.$$

Condition **A2** (the same **A2** that we have in our proofs) basically allows us to construct a weak order $\succsim_i$ on each dimension X_i. Informally, this means that the decision maker is able to rank levels within each attribute, or, in the state-dependent DUU context, outcomes for every state. The interpretation of the "2-graded" condition is similar to that of "strongly 2-graded" given above.

A very recent result is another characterization of the Sugeno integral (Couceiro et al., 2015). In this setup, attribute sets X_i are assumed to be finite chains. Elements 1_i and 0_i denote the maximal and minimal elements of these chains correspondingly. The authors give a single condition which characterizes the Sugeno integral in such setting. For each $k \in N$ we have (condition **PMD**):

$$[a_k x_{-k} \succ 0_k x_{-k}] \text{ AND } [1_k y_{-k} \succ a_k y_{-k}] \Rightarrow [a_k y_{-k} \succsim a_k x_{-k}].$$

This is a very nice and elegant characterization which is based on the fact that the Sugeno integral $S(x)$ is *pseudo-median decomposable*, which means that for each $k \in N$ we can construct a utility function $\phi_k : X_k \to R$ such that $S(x) = \text{med}(S(0_k x_{-k}), \phi_k(x_k), S(1_k x_{-k}))$. Indeed, consider alternatives $a_k x_{-k}$ and $a_k y_{-k}$. Then $a_k x_{-k} \succ 0_k x_{-k}$ means that "downgrading" attribute k makes the corresponding alternative $0_k x_{-k}$ strictly worse. In the same manner $1_k y_{-k} \succ a_k y_{-k}$ means that "upgrading" attribute k makes $1_k y_{-k}$ strictly better than $a_k y_{-k}$. Pseudo-median decomposability implies that since downgrading attribute k made the alternative strictly worse, its overall value was not higher than that of a_k. Similarly, in the second case, the overall value was not lower than value of a_k. From this the conclusion follows.

Bibliography

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine, *Econometrica: Journal of the Econometric Society* **21**(4): 503–546. 12

Angilella, S., Corrente, S. and Greco, S. (2015). Stochastic multiobjective acceptability analysis for the Choquet integral preference model and the scale construction problem, *European Journal of Operational Research* **240**(1): 172–182. 132

Angilella, S., Greco, S., Lamantia, F. and Matarazzo, B. (2004). Assessing non-additive utility for multicriteria decision aid, *European Journal of Operational Research* **158**(3): 734–744. 132

Angilella, S., Greco, S. and Matarazzo, B. (2010). Non-additive robust ordinal regression: A multiple criteria decision model based on the Choquet integral, *European Journal of Operational Research* **201**(1): 277–288. 132

Bana E Costa, C. and Vansnick, J. C. (1994). MACBETH - An interactive path towards the construction of cardinal value functions, *International Transactions in Operational Research* **1**(4): 489–500. 150

Ben-Porath, E., Gilboa, I. and Schmeidler, D. (1997). On the measurement of inequality under uncertainty, *Journal of Economic Theory* **75**(1): 194–204. 138

Benabbou, N., Gonzales, C., Perny, P. and Viappiani, P. (2015). Minimax regret approaches for preference elicitation with rank-dependent aggregators, *EURO Journal on Decision Processes* **3**(1-2): 29–64. 132

Benabbou, N., Perny, P. and Viappiani, P. (2014). Incremental elicitation of Choquet capacities for multicriteria decision making, *ECAI 2014*, pp. 87–92. 132

Bleichrodt, H., Li, C., Moscati, I. and Wakker, P. P. (2016). Nash was a first to axiomatize expected utility, *Theory and Decision* pp. 1–4. 10

Bliss, C. (1975). *Capital theory and the distribution of income*, Advanced textbooks in economics, North-Holland Pub. Co. 36

Bouyssou, D., Greco, S., Matarazzo, B., Pirlot, M. and Slowinski, R. (2002). Characterization of max, min and order statistics multicriteria aggregation functions, *Congrès IFORS 2002*. 116, 117

Bouyssou, D. and Marchant, T. (2010). Additive conjoint measurement with ordered categories, *European Journal of Operational Research* **203**(1): 195–204. 106

Bouyssou, D., Marchant, T. and Pirlot, M. (2009). A conjoint measurement approach to the discrete Sugeno integral, *The Mathematics of Preference, Choice and Order*, Springer, pp. 85–109. 27, 29, 30, 119, 135, 152, 153

Bouyssou, D., Marchant, T., Pirlot, M., Perny, P. and Tsoukiàs, A. (2000). *Evaluation and decision models: a critical perspective*, Kluwer Academic Publishers. 17, 134

Bouyssou, D., Marchant, T., Pirlot, M., Tsoukiās, A. and Vincke, P. (2006). *Evaluation and decision models with multiple criteria: stepping stones for the analyst*, Springer. 36

Bouyssou, D. and Pirlot, M. (2004). Preferences for multi-attributed alternatives: Traces, dominance, and numerical representations, *Journal of Mathematical Psychology* **48**(3): 167–185. 43

Chateauneuf, A. and Jaffray, J. Y. (1989). Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion, *Mathematical Social Sciences* **17**(3): 263–283. 121

Chew, H. S. and Karni, E. (1994). Choquet expected utility with a finite state space: commutativity and act-independence, *Journal of Economic Theory* **62**(2): 469–479. 148

Chew, H. S. and Wakker, P. (1996). The comonotonic sure-thing principle, *Journal of Risk and Uncertainty* **12**(1): 5–27. 137, 146, 148, 149

Couceiro, M., Dubois, D., Prade, H. and Waldhauser, T. (2015). Decision-making with Sugeno integrals, *Order* pp. 1–19. 153

Couceiro, M. and Marichal, J.-L. (2011). Axiomatizations of Lovász extensions of pseudo-Boolean functions, *Fuzzy Sets and Systems* **181**(1): 28–38. 150

De Campos, L. M. and Bolanos, M. J. (1992). Characterization and comparison of Sugeno and Choquet integrals, *Fuzzy Sets and Systems* **52**(1): 61–67. 149, 151

Debreu, G. (1959). Topological methods in cardinal utility theory, *Cowles foundation discussion papers* . 11, 19

Doignon, J.-P., Ducamp, A. and Falmagne, J.-C. (1984). On realizable biorders and the biorder dimension of a relation, *Journal of Mathematical Psychology* **28**(1): 73–109. 135

Drèze, J. (1987). *Essays on economic decisions under uncertainty*, Cambridge University Press. 16

Dubois, D., Prade, H. and Sabbadin, R. (1998). Qualitative decision theory with Sugeno integrals, *Proceedings of the fourteenth conference on uncertainty in artificial intelligence*, Morgan Kaufmann Publishers Inc., pp. 121–128. 152

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms, *The Quarterly Journal of Economics* **75**(4): 643–669. 12, 13

Fishburn, P. C. (1970). *Utility theory for decision making*, Wiley, New York. 11

Fishburn, P. C. (1973). A mixture-set axiomatization of conditional subjective expected utility, *Econometrica: Journal of the Econometric Society* pp. 1–25. 137

Fishburn, P. and Wakker, P. (1995). The invention of the independence condition for preferences, *Management Science* **41**(7): 1130–1144. 11, 13, 27

Galanter, E. and Pliner, P. (1974). Cross-modality matching of money against other continua, *Sensation and Measurement*, Springer, pp. 65–76. 136

Gilboa, I. (1987). Expected utility with purely subjective non-additive probabilities, *Journal of Mathematical Economics* **16**(1): 65–88. 13, 146

Gilboa, I. and Schmeidler, D. (1994). Additive representations of non-additive measures and the Choquet integral, *Annals of Operations Research* **52**(1): 43–65. 120

Goujon, B. and Labreuche, C. (2013). Holistic preference learning with the Choquet integral, *8th conference of the European Society for Fuzzy Logic and Technology (EUSFLAT-13)*, Atlantis Press. 132

Grabisch, M. (1997a). Alternative representations of discrete fuzzy measures for decision making, *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems* **5**(5): 587–607. 128

Grabisch, M. (1997b). k-order additive discrete fuzzy measures and their representation, *Fuzzy Sets and Systems* **92**(2): 167–189. 128

Grabisch, M., Kojadinovic, I. and Meyer, P. (2008). A review of methods for capacity identification in Choquet integral based multi-attribute utility theory: applications of the Kappalab R package, *European Journal of Operational Research* **186**(2): 766–785. 27, 126

Grabisch, M. and Labreuche, C. (2005). Fuzzy measures and integrals in MCDA, *Multiple criteria decision analysis: state of the art surveys* pp. 563–604. 14

Grabisch, M. and Labreuche, C. (2008). A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid, *4OR: A Quarterly Journal of Operations Research* **6**(1): 1–44. 13, 27, 35, 135

Grabisch, M., Marichal, J.-L., Mesiar, R. and Pap, E. (2009). *Aggregation functions (encyclopedia of mathematics and its applications)*, 1st edn, Cambridge University Press, New York, NY, USA. 28

Greco, S., Matarazzo, B. and Słowiński, R. (2004). Axiomatic characterization of a general utility function and its particular cases in terms of conjoint measurement and rough-set decision rules, *European Journal of Operational Research* **158**(2): 271–292. 30, 152

Guttman, L. (1944). A basis for scaling qualitative data, *American Sociological Review* pp. 139–150. 135

Harsanyi, J. C. (1980). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility, *Essays on ethics, social behavior, and scientific explanation*, Springer, pp. 6–23. 139

Hüllermeier, E. and Tehrani, A. F. (2012). On the vc-dimension of the Choquet integral, *Advances on Computational Intelligence*, Springer, pp. 42–50. 132

Kahneman, D. (2011). *Thinking, fast and slow*, Macmillan. 136

Karni, E. (1985). *Decision Making Under Uncertainty: The Case of State-dependent Preferences*, Harvard University Press. 16, 137, 148, 149

Karni, E. (1993). A definition of subjective probabilities with state-dependent preferences, *Econometrica* **61**(1): 187–198. 137, 148, 149

Karni, E. and Schmeidler, D. (2016). An expected utility theory for state-dependent preferences, *Theory and Decision* pp. 1–12. 19, 137, 149

Keeney, R. L. and Raiffa, H. (1976). *Decisions with multiple objectives: preferences and value tradeoffs*, Cambridge University Press. 11, 13

Köbberling, V. and Wakker, P. P. (2003). Preference foundations for nonexpected utility: A generalized and simplified technique, *Mathematics of Operations Research* **28**(3): 395–423. 147, 148

Krantz, D. H. (1972). A theory of magnitude estimation and cross-modality matching, *Journal of mathematical psychology* **9**(2): 168–199. 136

Krantz, D. H., Luce, R. D., Suppes, P. and Tversky, A. (1971). Foundation of measurement. vol. 1: Additive and polynomial representations. 11, 19, 23, 30, 33, 35, 39, 42, 48, 55, 78, 80, 108

Labreuche, C. (2012). An axiomatization of the Choquet integral and its utility functions without any commensurability assumption, *in* S. Greco, B. Bouchon-Meunier, G. Coletti, M. Fedrizzi, B. Matarazzo and R. R. Yager (eds), *Advances in Computational Intelligence*, Vol. 300 of *Communications in Computer and Information Science*, Springer Berlin Heidelberg, pp. 258–267. 10.1007/978-3-642-31724-8_27. 28, 30, 151

Labreuche, C. and Grabisch, M. (2003). The Choquet integral for the aggregation of interval scales in multicriteria decision making, *Fuzzy Sets and Systems* **137**(1): 11–26. 14, 150

Li, G., Law, R., Vu, H. Q. and Rong, J. (2012). Discovering the hotel selection preferences of Hong Kong inbound travelers using the Choquet integral, *Tourism Management* . 133

Mak, K.-T. (1984). Notes on separable preferences, *Journal of Economic Theory* **33**(2): 309–321. 36

Marichal, J.-L. (2000). On Sugeno integral as an aggregation function, *Fuzzy Sets and Systems* **114**(3): 347–365. 152

Marichal, J. L. (2002). An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria, *Fuzzy Systems, IEEE Transactions on* **8**(6): 800–807. 149, 150

Marichal, J. L. and Roubens, M. (2000). Determination of weights of interacting criteria from a reference set, *European Journal of Operational Research* **124**(3): 641–650. 127

Monderer, D. and Shapley, L. S. (1996). Potential games, *Games and Economic Behavior* **14**(1): 124–143. 77

Mori, T. and Murofushi, T. (1989). An analysis of evaluation model using fuzzy measure and the Choquet integral, *5th Fuzzy System Symposium*, pp. 207–212. 13

Moscati, I. (2016). Retrospectives: how economists came to accept expected utility theory: the case of Samuelson and Savage, *The Journal of Economic Perspectives* **30**(2): 219–236. 12

Murofushi, T. (1992). A technique for reading fuzzy measures (i): the Shapley value with respect to a fuzzy measure, *2nd Fuzzy Workshop*, pp. 39–48. 13

Murofushi, T. and Soneda, S. (1993). Techniques for reading fuzzy measures (iii): interaction index, *9th Fuzzy System Symposium*, pp. 693–696. 13, 128

Murofushi, T. and Sugeno, M. (1992). Non-additivity of fuzzy measures representing preferential dependence, 2nd Int, *Conf. on Fuzzy Systems and Neural Networks, Iizuka, Japan.* 13

Murofushi, T. and Sugeno, M. (1993). Some quantities represented by the Choquet integral, *Fuzzy Sets and Systems* **56**(2): 229–235. 116

Ovchinnikov, S. (1996). Means on ordered sets, *Mathematical Social Sciences* **32**(1): 39–56. 115

Quiggin, J. (1982). A theory of anticipated utility, *Journal of Economic Behavior & Organization* **3**(4): 323–343. 13, 26

Roberts, K. W. (1980). Interpersonal comparability and social choice theory, *The Review of Economic Studies* pp. 421–439. 138

Savage, L. (1954). *The foundations of statistics*, Wiley publications in statistics, John Wiley and Sons. 10, 12, 16

Schmeidler, D. (1989). Subjective probability and expected utility without additivity, *Econometrica: Journal of the Econometric Society* **57**(3): 571–587. 13, 14, 16, 26, 41, 134, 146

Segal, U. and Sobel, J. (2002). Min, max, and sum, *Journal of Economic Theory* **106**(1): 126–150. 116

Shapley, L. S. (1953). A value for n-person games, *Contributions to the Theory of Games* **2**: 307–317. 127

Siskos, Y., Grigoroudis, E. and Matsatsinis, N. F. (2005). UTA methods, *Multiple criteria decision analysis: state of the art surveys*, Springer, pp. 297–334. 131

Sobrie, O., Gillis, N., Mousseau, V. and Pirlot, M. (2016). Uta-poly and uta-splines: additive value functions with polynomial marginals, *arXiv preprint arXiv:1603.02626* . 131

Sounderpandian, J. (1991). Value functions when decision criteria are not totally substitutable, *Operations Research* **39**(4): 592–600. 116

Stevens, J. C. and Marks, L. E. (1965). Cross-modality matching of brightness and loudness., *Proceedings of the National Academy of Sciences of the United States of America* **54**(2): 407. 136

Stevens, J. C. and Marks, L. E. (1980). Cross-modality matching functions generated by magnitude estimation, *Perception & Psychophysics* **27**(5): 379–389. 136

Stevens, S. (1959). Cross-modality validation of subjective scales for loudness, vibration, and electric shock., *Journal of Experimental Psychology* **57**(4): 201. 136

Sugeno, M. (1974). *Theory of fuzzy integrals and its applications*, Tokyo Institute of Technology. 13, 28

Tehrani, A., Cheng, W., Dembczyński, K. and Hüllermeier, E. (2011). Learning monotone nonlinear models using the Choquet integral, *in* D. Gunopulos, T. Hofmann, D. Malerba and M. Vazirgiannis (eds), *Machine learning and knowledge discovery in databases*, Vol. 6913 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, pp. 414–429. 130, 132

Tehrani, A. F., Cheng, W., Dembczyński, K. and Hüllermeier, E. (2012). Learning monotone nonlinear models using the Choquet integral, *Machine Learning* **89**(1-2): 183–211. 130, 132

Tehrani, A. F., Cheng, W. and Hüllermeier, E. (2011). Choquistic regression: Generalizing logistic regression using the Choquet integral., *EUSFLAT Conf.*, pp. 868–875. 130, 132

Tehrani, A. F. and Huellermeier, E. (2013). Ordinal choquistic regression., *EUSFLAT Conf.* 130, 132

Timonin, M. (2013). Robust optimization of the Choquet integral, *Fuzzy Sets and Systems* **213**: 27 – 46. 132

Timonin, M. (2014a). Axiomatization of the Choquet integral for heterogeneous 2-dimensional sets, *2014 European Mathematical Psychology Group Meeting Abstracts*, pp. 28–29. 2

Timonin, M. (2014b). Conjoint axiomatization of the Choquet integral for two-dimensional heterogeneous product sets, *Proceedings of DA2PL2014: From Multicriteria Decision Aid to Preference Learning.*, pp. 31–34. 2

Timonin, M. (2015a). Axiomatization of the Choquet integral, *Proceedings of EURO2015: 27th European Conference on Operational Research*, pp. 230–230. 2

Timonin, M. (2015b). Axiomatization of the Choquet integral for 2-dimensional heterogeneous product sets, *arXiv preprint arXiv:1507.04167* . 2

Timonin, M. (2016a). Axiomatization of the Choquet integral and some aspects of its behavioural analysis, *Proceedings of EURO2016: 28th European Conference on Operational Research*, pp. 166–166.
URL: *https://www.euro-online.org/conf/admin/tmp/program-euro28.pdf* 2

Timonin, M. (2016b). Axiomatization of the Choquet integral for heterogeneous product sets, *arXiv preprint arXiv:1603.08142* . 2

Timonin, M. (2016c). Conjoint axiomatization of the Choquet integral for heterogeneous product sets, *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Springer, pp. 46–57. 2

Timonin, M. (2016d). Conjoint axiomatization of the Choquet integral for heterogeneous product sets, *Proceedings of RUD2016*.
URL: *sites.google.com/site/rudconference/2016----paris-dauphine/conference-program* 2

Von Neumann, J. and Morgenstern, O. (1944). *Theory of games and economic behavior.*, Princeton University Press. 10

Wakker, P. (1991a). Additive representations of preferences, a new foundation of decision analysis; the algebraic approach, *Mathematical Psychology*, Springer, pp. 71–87. 13, 20, 35, 41, 42, 78, 85, 88, 146, 147

Wakker, P. (1991b). Additive representations on rank-ordered sets. I. the algebraic approach, *Journal of Mathematical Psychology* **35**(4): 501–531. 13, 23, 30, 35, 39, 40, 43, 48, 59, 60, 65, 80, 82, 84, 85, 100, 101, 109, 147

Wakker, P. P. (1989). *Additive representations of preferences: a new foundation of decision analysis*, Kluwer Academic Publishers Dordrecht. 95, 109, 134, 146