Risk and Welfare: Theory and Experiment

ISBN 978 90 361 0439 5
Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **767** of the Tinbergen Institute Research Series, established through cooperation between Rozenberg Publishers and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Risk and Welfare: Theory and Experiment

Risico en Welvaart: Theorie en Experiment

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus
Prof.dr. R.C.M.E. Engels

and in accordance with the decision of the Doctorate Board.
The public defense shall be held on
Wednesday 2 September 2020, at 9:30 hours

by

JINGNI YANG

born in Shaanxi, China

Erasmus University Rotterdam

Doctoral Committee :

Promotors :	Prof.dr. P.P. Wakker
	Prof.dr. H. Bleichordt
Other Members :	Dr. J.T.R. Stoop
	Prof.dr. F. Maccheroni
	Prof.dr. T.J.S. Offerman

Acknowledgements

I was super lucky to have Peter Wakker, the most intelligent scholar, as my supervisor. Peter started my PhD life with a Cauchy equation, which brought me back to the old days that I was involved with math problems. It always takes me days, weeks or even longer to understand what Peter said. The way he guides students is by asking questions to help me think about a problem deeper and deeper. Peter has the magic of making me do something that I thought I could not. Collaborating with Han has provided me an opportunity to understand the connection between empirical studies and theories, which is very valuable for my future research.

Massimo Marinacci is a model for me, and it was like a dream that I had the chance to know him. I have benefited much from his way of thinking and continuous encouragements. I would like to thank Simon Grant for his trust so that I can work at ANU. Also I thank his family, Michelle, Alexandra and James for their hospitality in Australia.

I would like to thank: Aurelian Baillon, Paul van Bruggen, Jan Heufer, Shaowei Ke, Fabio Maccheroni, Theo Offerman, Fan Wang, Qizhi Wang, and Songfa Zhong, for their encouragements, discussions and very useful comments. Special thanks go to my teachers from my math master program, Sonja Cox, Flora Spieksma, Dorothee Frey, KP Hart, and our super fantastic François Genoud.

Indy, it was quite surprising that we two share so many same ideas about lives and dreams. It is always enjoyable to chat with you even if we are in different continents now. New technologies make it easy to keep updated with each other. We went together all the way of PhD and went though all the happiness and depressions. Finally, we made it! It was a very pleasant experience to drop by with Paul and share possible ideas with him. Also, the experience of coauthoring with him makes me understand the economic reasoning deeper than before.

I would like to thank my friends in Netherlands: Tong, Yan, Xiao, Yuan, Lingwei, Junze, and Doris. Meanwhile, I would also like to thank our ANU big family: Sam, Nik, Tina, Yanan, Yanran, and Haiyan.

Lastly, I thank those close to me in this journey, my parents and Mr. W for their infinite love and support.

Contents

Chapter 1

Introduction

This dissertation contains 5 chapters on decision theory (besides this introduction). The second, third and fifth Chapter aim to derive some results without technical assumptions like differentiability or continuity. The fourth chapter introduces a new testable axiom for the changing of preferences. The last chapter provides a connection between Savage's framework and Luce's framework of modeling uncertainty.

In the second chapter, we show that convexity of preference on the two-dimensional comonotonic cone is equivalent to concave utility and that this result does not depend on differentiability or continuity of utility. Remarkable, we then show that by applying the main result, we can generalize existing results on concave/ convex utility or weighting functions (capacities). In the third chapter, we apply the same idea as in chapter 2 to the choice domain of lotteries, and obtain a surprising connection between Yaari's dual theory and convex preference.

In the fourth chapter, we propose a new axiom, Agreement, to deal with changing preferences. We apply it to reciprocal preferences and then derive the non-parametric implication of Agreement together with GARP. Furthermore, we have conducted an experiment to test the implications. We find that people have a consistent preferences for each context (GARP) and their preferences change from one context to another one always in the same direction (Agreement). The Agreement Axiom can be used in all situations where a decision maker assigns more weight to one attribute than to another.

Proper scoring rules provide an important tool for measuring beliefs. In the fifth chapter, we show that the logarithmic family is the only family of local proper scoring rules for more than two events even when we relax differentiability, continuity, and domain-restrictions.

Luce pointed out a problem of Savage's (1954) framework that many irrelevant events have to be considered when analyzing decision under uncertainty. In the sixth chapter, we use mosaics instead of (σ-)algebras as the collections of events in Savage's framework and prove that in this way, the problem is solved and Luce's framework can be embedded in our generalization.

Chapter 2

A powerful tool for analyzing concave/convex utility and weighting functions[1]

Abstract

This paper shows that convexity of preference has stronger implications for weighted utility models than had been known hitherto, both for utility and for weighting functions. Our main theorem derives concave utility from convexity of preference on the two-dimensional comonotonic cone, without presupposing continuity. We then show that this, seemingly marginal, result provides the strongest tool presently available for obtaining concave/convex utility or weighting functions. We revisit many classical results in the literature and show that we can generalize and improve them.

JEL-classification: D81, C60
Keywords: convex preferences, quasi-concave utility, risk aversion, ambiguity aversion, rank-dependent utility

[1] This chapter is joint with Peter P. Wakker.

2.1 Introduction

Convexity of preference is a standard condition in many fields (De Giorgi and Mahmoud 2016; Debreu 1959; Mas-Colell et al. 1995 p. 44). We examine it for weighted utility models, where its potential has not yet been fully recognized. Our first theorem shows its equivalence to concave utility on the two-dimensional comonotonic cone. This generalizes existing results by not presupposing continuity and by providing flexibility of domain. With this seemingly thin and marginal result we can in one blow generalize virtually all existing theorems on convex or concave utility or weighting functions, and make them more appealing.

The aforementioned theorems concern: (1a) risk aversion for expected utility not only for risk (von Neumann-Morgenstern 1944) but also for (1b) uncertainty (Savage 1954); (1c) Yaari's (1969) comparative risk aversion generalized by allowing for different beliefs; (2) concave/convex utility and weighting functions (2a) for Gilboa's (1987) and Schmeidler's (1989) rank-dependent utility for ambiguity, and (2b) for Tversky and Kahneman's (1992) prospect theory for ambiguity; (3) corresponding results for Ghirardato and Marinacci's (2001) biseparable utility for ambiguity[2]; (4) smooth ambiguity aversion (Klibanoff et al. 2005). Wakker and Yang (2018) show how our main theorem can be applied to decision under risk,[3] providing results on: (1) concave/convex utility and probability weighting for Quiggin's (1982) rank-dependent utility for risk and Tversky and Kahneman's (1992) prospect theory for risk; (2) corresponding results for Miyamoto's (1988) biseparable utility for risk[4]; (3) loss aversion in Köszegi and Rabin's (2006) reference dependent model; (4) inequality aversion for welfare theory (Ebert 2004).

The main contribution of this paper is not to generalize some theorems, which would constitute a marginal contribution, but to provide a general technique to obtain convex/concave utility or weighting functions in a more general and appealing manner than done before. As corollaries, we can generalize and improve virtually all existing results on this topic in the literature. To limit the size of this paper, we focus on uncertainty henceforth. Our theorems can readily be applied, though, not only to risk (Wakker and Yang 2018), but also to discounted utility for intertemporal choice with aversion to variation in outcomes, utilitarian welfare models with aversion to inequality, and other weighted utility models.

[2] This includes many ambiguity theories, such as maxmin expected utility (Alon and Schmeidler 2014; Gilboa and Schmeidler 1989), the α-Hurwicz criterion (Arrow and Hurwicz 1972), multiple priors-multiple weighting (Dean and Ortoleva 2017), contraction expected utility (Gajdos et al. 2008), alpha-maxmin (Ghirardato et al. 2004; Jaffray 1994; Luce and Raiffa 1957 Ch. 13), Hurwicz Expected Utility (Gul and Pesendorfer 2015), binary RDU (Luce 2000) including his rank-sign dependent utility, and binary expected utility (Pfanzagl 1959 pp. 287-288).

[3] We thank an anonymous referee and editor for recommending this organization of our results in two papers.

[4] This is the analog for risk of biseparable utility, and includes many risk theories such as disappointment theory (Bell 1985; Loomes and Sugden 1986) for a disappointment function kinked at 0, RAM and TAX models (Birnbaum 2008), disappointment aversion (Gul 1992), original prospect theory (Kahneman and Tversky 1979) for gains and for losses, Luce's (2000) binary RDU, and prospective reference theory (Viscusi 1989). See Wakker (2010 Observation 7.11.1).

The outline of the paper is as follows. Section 6.2 presents elementary definitions and our main result. To show its usefulness, the following sections apply our main result to a number of well-known classical results in the literature, generalizing and making them more appealing. These applications demonstrate that we have provided a general tool for analyzing concave/convex utility and weighting. Section 6.3 presents implications for uncertainty focusing on classical expected utility. Sections 6.4 and 6.5 turn to ambiguity models, followed by a concluding section and an appendix with proofs. In each proof, we first find a substructure isomorphic to our main theorem, and then extend the desired result to the whole domain considered.

2.2 Definitions and our main theorem

S is the *state space* that can be finite or infinite. $\mathscr{A}$ denotes an algebra of subsets called *events*. The *outcome set* is a nonpoint interval $I \subset \mathbb{R}$, bounded or not. $\mathscr{F}$ denotes a set of functions from S to I called *acts*, which are assumed measurable (inverses of intervals are events). Outcomes are identified with constant acts. We assume that $\mathscr{F}$ contains all *simple*, i.e., finite-valued measurable functions. Other than that, $\mathscr{F}$ can be general, with however the restriction added that all RDU values (defined later) are well-defined. In particular, $\mathscr{F}$ may consist exclusively of simple acts, or contain all bounded acts. By $x = (E_1 : x_1, \ldots, E_n : x_n)$ we denote the simple act assigning outcome x_j to all states in E_j. It is implicitly understood that the E_js are events partitioning S.

A *preference relation*, i.e., a binary relation $\succcurlyeq$ on $\mathscr{F}$, is given; $\succ, \preccurlyeq, \prec, \sim$ are as usual. V *represents* $\succcurlyeq$ on $\mathscr{F}' \subset \mathscr{F}$ if V is real valued with $\mathscr{F}'$ contained in its domain and $x \succcurlyeq y \Leftrightarrow V(x) \geq V(y)$ for all acts $x, y \in \mathscr{F}'$. This implies weak ordering on $\mathscr{F}'$, i.e., $\succcurlyeq$ is transitive and complete there. If we omit "on $\mathscr{F}'$", then $\mathscr{F}' = \mathscr{F}$. Central in this paper are convex combinations $\lambda x + (1-\lambda)y$. Here x and y are acts, $0 \leq \lambda \leq 1$, and the combination concerns the statewise mixing of outcomes. We do not assume that $\mathscr{F}$ is closed under convex combinations. The set of simple acts is, and this provides enough richness for all our theorems.

DEFINITION 1. We call $\succcurlyeq$ *convex* if $x \succcurlyeq y \Rightarrow \lambda x + (1-\lambda)y \succcurlyeq y$ for all $0 \leq \lambda \leq 1$ and acts $x, y, \lambda x + (1-\lambda)y$. □

The condition is only imposed if the mix indeed is an act; i.e., is contained in the domain. Convexity of preference is equivalent to quasi-concavity of representing functions.[5]

An *(event) weighting function* W maps events to $[0,1]$ such that: $W(\emptyset) = 0, W(S) = 1$, and $A \supset B \Rightarrow W(A) \geq W(B)$. *Finitely additive probability measures* P are additive weighting

[5] Unfortunately, terminology in the literature is not uniform, and sometimes terms concave, quasi-convex, or quasi-concave have been used. We use the most common term, convex.

functions. They need not be countably additive. Preference conditions necessary and sufficient for countable additivity are well known (Arrow 1971; Wakker 1993 Proposition 4.4), and can optionally be added in all our theorems.

For a weighting function W and a function $U : I \to \mathbb{R}$, the *rank-dependent utility* (*RDU*) of an act x is

$$\int_{\mathbb{R}^+} W\{s \in S : U(x(s)) > \alpha\} d\alpha - \int_{\mathbb{R}^-} (1 - W\{s \in S : U(x(s)) > \alpha\}) d\alpha. \tag{2.1}$$

An alternative term used in the literature is Choquet expected utility. We impose one more restriction on $\mathscr{F}$: RDU is well defined and finite for all its elements. A necessary and sufficient condition directly in terms of preferences—requiring preference continuity with respect to truncations of acts—is in Wakker (1993). A sufficient condition is that all acts are bounded (with an upper and lower bound contained in I). For a simple act $(E_1 : x_1, \ldots, E_n : x_n)$ with $x_1 \geq \cdots \geq x_n$, the RDU is

$$\sum_{j=1}^{n} (W(E_1 \cup \cdots \cup E_j) - W(E_1 \cup \cdots \cup E_{j-1})) U(x_j). \tag{2.2}$$

Rank-dependent utility (*RDU*) holds on $\mathscr{F}' \subset \mathscr{F}$ if there exist W and strictly increasing U such that *RDU* represents $\succcurlyeq$ on $\mathscr{F}'$. Then U is called the *utility function*. Again, if we omit "on $\mathscr{F}'$," then $\mathscr{F}' = \mathscr{F}$. If $\mathscr{F}'$ contains all constant acts, then strict increasingness of U is equivalent to *monotonicity*: $\gamma > \beta \Rightarrow \gamma \succ \beta$ for all outcomes. We do not require continuity of U. The special case of RDU with W a finitely additive probability measure P is called *expected utility* (*EU*). We sometimes write *subjective EU* if P is subjective. We assume existence of a *nondegenerate event* E, meaning $(E : \gamma, E^c : \gamma) \succ (E : \gamma, E^c : \beta) \succ (E : \beta, E^c : \beta)$ for some outcomes $\gamma > \beta$ with the acts contained in the relevant domain $\mathscr{F}'$. Under sufficient richness, satisfied in all cases considered in this paper, nondegenerateness means $0 < W(E) < 1$.[6] We summarize the assumptions made.

ASSUMPTION 2. [Structural assumption for uncertainty] S is a state space, $\mathscr{A}$ an algebra of subsets (events), and I a nonpoint interval. $\mathscr{F}$, the set of acts, is a set of measurable functions from S to I containing all simple functions, endowed with a binary (preference) relation $\succcurlyeq$. RDU represents $\succcurlyeq$ on a subset $\mathscr{F}'$ of $\mathscr{F}$ (default: $\mathscr{F}' = \mathscr{F}$). There exists a nondegenerate event E. □

To obtain complete preference axiomatizations in the theorems in this paper, we should

[6] If no nondegenerate event exists, then an RDU representation exists with a linear, so surely concave, utility, and in this sense all results below hold true—also regarding convexity of weighting functions as can be demonstrated—without the extra requirement of nondegeneracy. However, then utility is ordinal (Wakker 1989 Observation VI.5.1') so that utility can also be chosen nonconcave, and we should formulate all our results as existence results. We avoid complicating our formulations this way.

state preference conditions for the decision models assumed. Such conditions were surveyed by Köbberling and Wakker (2003) and will not be repeated here.

W is *convex* if $W(A \cup B) + W(A \cap B) \geq W(A) + W(B)$. Elementary manipulations show that this holds if and only if

$$W(A \cup B) - W(B) \leq W(A \cup B') - W(B') \text{ whenever } A \cap B' = \emptyset \text{ and } B \subset B'. \tag{2.3}$$

The latter formulation shows the analogy with increasing derivatives of real-valued convex functions. An interesting implication of convexity of W is that RDU then belongs to the popular maxmin EU model (Wald 1950; Gilboa and Schmeidler 1989) with the set of priors equal to the *Core*, i.e., the set of probability measures that dominate W (Schmeidler 1986 Proposition 3; Shapley 1971).

The following theorem is our main result. Virtually all preceding results in the literature— Debreu and Koopmans (1982) excepted— assumed continuity and often even differentiability, but we do not.

THEOREM 3. *[Main theorem] Assume: (a) Structural Assumption 2; (b)* $S = \{s_1, s_2\}$*; (c)* s_1 *is nondegenerate; (d) EU (= RDU) holds on* $\mathscr{F}' = \{x = (s_1 : x_1, s_2 : x_2) : x_1 \geq x_2\}$. *Then utility is concave if and only if* $\succcurlyeq$ *is convex on* $\mathscr{F}'$. □

In the theorem, nondegeneracy of s_1 is equivalent to nondegeneracy of s_2. The proof of the theorem is more complex than of its analogs on full product spaces that can use hedging, as in the half-half mixture of $(1,0)$ and $(0,1)$ resulting in the sure $(0.5, 0.5)$. Hedging provides a powerful tool for analyzing convex preferences, extensively used in the literature, that we cannot use though because all acts in our domain are maximally correlated. This complicates our proof relative to, for instance, Debreu and Koopmans (1982), its simplification Crouzeix and Lindberg (1986), its generalization Monteiro (1999), and most other predecessors. Therefore, unlike Debreu and Koopmans, we need strictly increasing utility. Example A.3 shows that our theorem does not hold for nondecreasing utility. Because strictly increasing utility is natural in most applications, it does not entail a serious restriction. In return, the flexibility of domain provided by our theorem allows us to apply it to utility functions when expected utility is violated, and to apply it dually so that it speaks to weighting functions. Chateauneuf and Tallon (2002), Ghirardato and Marinacci (2001), and Wakker (1994) did consider comonotonic sets of acts (defined in §6.4) as above. Theorem 3 generalizes their results by showing that continuity/differentiability is redundant. §6.5 gives further details.

2.3 Implications for decision under uncertainty: expected utility

This section considers applications of the main Theorem 3 to classical EU for decision under uncertainty.

COROLLARY 4. *If Structural Assumption 2 and EU hold, then U is concave if and only if $\succcurlyeq$ is convex.* $\square$

Corollary 4 is useful for capturing risk aversion because convexity is directly observable, not involving subjective probabilities. Remarkably, the early Yaari (1965) already pointed out that the traditional definitions of risk aversion, relating to expected value or mean-preserving spreads, cannot be used for subjective EU. He hence tested convexity instead. However, he did not observe that convexity is actually equivalent to the traditional definitions.

Although an early version of Corollary 4 appeared in Debreu and Koopmans (1982 p. 4) and has been used in some works (Section 6.5), the result did not yet receive the attention it deserves and has not been generally known. Alternative, more complex, preference conditions for concave utility under subjective EU are in Baillon et al. (2012), Harvey (1986 Theorem 3), Wakker (1989), Wakker (2010 Eq. 4.8.2), and Wakker and Tversky (1993 §9).

We next turn to comparative results. In what follows, superscripts refer to decision makers. Yaari (1969) provided a well-known characterization of comparative risk aversion under subjective EU, where decision maker $\succcurlyeq^2$ with utility U^2 is more risk averse than decision maker $\succcurlyeq^1$ with utility U^1 if her certainty equivalents are always lower. Then U^2 is a concave transformation of U^1. Unfortunately, Yaari's condition is not necessary and sufficient, but only holds if the two decision makers have the same subjective probabilities. Decision makers with different beliefs cannot be compared because Yaari's condition then is never satisfied. The basic problem is that certainty equivalents depend on probabilities and, thus, involve not only risk attitudes but also beliefs. Yaari's method of comparing certainty equivalents has become a common tool in ambiguity theories, for instance to compare ambiguity aversion across decision makers.[7] Then invariably all other attitude components of the decision makers except the one compared have to be identical. This is implied by the fundamental problem of certainty equivalents of involving all components of decision attitudes. It limits the scope of application. We now show how outcome mixing avoids the aforementioned limitations and works for general beliefs, for Yaari's original EU framework. Generalizations to nonexpected utility theories are left to future work. A preparatory notation: $\alpha_E\beta$ denotes the binary act $(E:\alpha, E^c:\beta)$. Mathematically, we will

[7] See, for instance, Epstein (1999 Definition 2.3), Ghirardato and Marinacci (2001 §4.1), Ghirardato and Marinacci (2002 Definition 4), Izhakian (2017 Definition 4), and Klibanoff et al. (2005 Definition 5).

describe the case where $\succcurlyeq^2$ is risk averse if outcomes are expressed in U^1 units, i.e., units that make $\succcurlyeq^1$ risk neutral. In these outcome units, $\succcurlyeq^2$ should be convex. To capture this idea in a preference condition, we have to avoid the explicit use of theoretical constructs such as U^1. We have to reveal mixtures of acts in U^1 units directly from preferences. Gul (1992) showed a way to do this. Assume, for any event A, with $x_j \geq y_j \geq z_j$ for $j = 1,2$:

$$(x_{1_A} z_1) \sim^1 y_1 \text{ and } (x_{2_A} z_2) \sim^1 y_2. \tag{2.4}$$

This shows that, in U^1 units, y_1 is a mixture of x_1 and z_1, and y_2 is so of x_2 and z_2, with the subjective probabilities $P^1(A)$ and $1 - P^1(A)$ as mixing weights. These weights are not directly observable but this is no problem. All we need for what follows is that these weights are the same in both mixtures. This is enough to infer that, for any event B, the act $(y_{1_B} y_2)$ is a convex mixture of $(x_{1_B} x_2)$ and $(z_{1_B} z_2)$ in U^1 units. Our convexity condition requires that the mixture $(y_{1_B} y_2)$ is preferred to the other two acts if they are indifferent. That is, $\succcurlyeq^2$ is *more outcome-risk averse*[8] than $\succcurlyeq^1$ if, for all events B:

$$(x_{1_B} x_2) \sim^2 (z_{1_B} z_2) \Rightarrow (y_{1_B} y_2) \succcurlyeq^2 (z_{1_B} z_2) \tag{2.5}$$

is implied by Eq. 2.4. Eq. 2.5 is the convexity condition in terms of U^1 units, weakened to the case where the antecedent preference is actually an indifference, and where the mixtures weights are $P^1(A)$ and $1 - P^1(A)))$ for some event A. Under continuity, this weakened version of convexity is strong enough to imply full-force convexity, as Corollary 5 will show. To see intuitively that our condition captures comparative risk aversion, first note that Eq. 2.4 implies, for all events B'

$$(x_{1_{B'}} x_2) \sim^1 (z_{1_{B'}} z_2) \Rightarrow (y_{1_{B'}} y_2) \sim^1 (z_{1_{B'}} z_2) \tag{2.6}$$

because of linearity in probability mixing of the EU^1 functional.[9] The event B' to bring indifference for decision maker 2 may be different than B due to different beliefs and/or state spaces. Comparing Eqs. 2.5 and 2.6 shows that, if decision maker 1 is indifferent to the mix, then decision maker 2 prefers it. This reveals stronger risk aversion, as formalized next.

COROLLARY 5. *Assume that $\succcurlyeq^1$ and $\succcurlyeq^2$ both satisfy Structural Assumption 2 with the same outcome interval I, and both maximize subjective EU with continuous utility functions U^1 and U^2, respectively. Then $\succcurlyeq^2$ is more outcome-risk averse than $\succcurlyeq^1$ if and only if $U^2(.) = \varphi(U^1(.))$ for a concave transformation φ. The two decision makers may have different probabilities and may even face different state spaces.* □

[8] To avoid confusion with Yaari's widely accepted terminology, we add "outcome" to our term.

[9] Gul (1992) used a strengthened version of the implication Eq. 2.4 ⇒ Eq. 2.6 to axiomatize subjective expected utility.

Both the ambiguity aversion of Klibanoff et al.ś (2005) smooth model, and the preference for early resolution of uncertainty of Kreps and Porteus (1978) amount to having one EU utility function more concave than another. Corollary 5 shows a way to obtain these results without involving probabilities as inputs in the preference condition. Note that we have to be able to deal with different state spaces in these applications.

Our outcome-risk aversion condition is more complex, and less appealing, than Yaari's certainty equivalent condition. Its pro is that it delivers a clean comparison of utility and risk attitude, not confounded by beliefs. Both conditions deserve study. Baillon et al. (2012) provided other characterizations of comparative risk aversion that, like our result, do not require same beliefs or given probabilities. They used an endogenous midpoint operation for utilities. Heufer (2014) showed how Yaari's certainty equivalent condition can be elicited from revealed preferences. Our paper propagates the use of preference convexity. Heufer (2012) showed how this convexity can be elicited from revealed preferences.

2.4 Implications for decision under uncertainty: ambiguity

We first discuss applications to RDU, and start with basic definitions. Acts x, y are *comonotonic* if there are no states s, t with $x(s) > x(t)$ and $y(s) < y(t)$. A set of acts is *comonotonic* if each pair of its elements is comonotonic. A *comoncone* is a maximal comonotonic set. It corresponds with an ordering ρ, called *ranking*, of S and contains all acts x with $s\rho t \Rightarrow x(s) \geq x(t)$. Every event E of the form $\{s \in E : s\rho t, s \neq t\}$ is called a *goodnews* event or, more formally, a *rank*. Intuitively, it reflects the good news of receiving all outcomes ranked better than some outcome. A set of acts is comonotonic if and only if it is a subset of a comoncone. A comoncone is *nondegenerate* if it has a nondegenerate goodnews event. On every comoncone with related ordering ρ, RDU agrees with an EU functional with the finitely additive probability measure P_ρ agreeing with W on the goodnews events. The following corollary is a straightforward generalization of the main Theorem 3.

COROLLARY 6. *If Structural Assumption 2 holds and RDU holds on a nondegenerate comoncone $\mathscr{F}'$, then U is concave if and only if $\succcurlyeq$ is convex on $\mathscr{F}'$.* □

A first version of the following result was in Chateauneuf and Tallon (2002), and did not receive the attention it deserves.

COROLLARY 7. *[Main corollary] If Structural Assumption 2 and RDU hold, then $\{U$ is concave and W is convex$\}$ if and only if $\succcurlyeq$ is convex.* □

The result is especially appealing because the two most important properties of RDU jointly follow from one very standard preference condition. The only proof available as yet, by

Chateauneuf and Tallon, assumes differentiability of utility, which is problematic for preference foundations. Differentiability is not problematic, and useful, in most economic applications. For preference foundations the situation is different, though. Preference foundations seek for conditions directly observable from preferences. In this way, preference foundations make theories operational. For general differentiability, there is no clear and elementary preference condition.[10] Hence, differentiability assumptions are better avoided in preference foundations. Strictly speaking, Corollary 7 is then the first joint preference foundation of the two most popular specifications of RDU. Chateauneuf and Tallon (2002) did not present their result very saliently[11]. This, and the use of differentiability, may explain why this appealing result has not yet been as widely known as it deserves to be.

We, finally, give results for a large class of nonexpected utility models. By $\mathscr{F}_E$ we denote the set of binary acts $\gamma_E\beta = (E:\gamma, E^c:\beta)$, and by $\mathscr{F}_E^{\uparrow}$ we denote the subset with $\gamma \geq \beta$. *Biseparable utility* holds if there exist a utility function U, and a weighting function W, such that $RDU(x)$ represents $\succcurlyeq$ on the set of all binary acts x. That is, for all binary acts $\gamma_E\beta$ ($\gamma \geq \beta$) we have an RDU representation $W(E)U(\gamma) + (1 - W(E))U(\beta)$, but for acts with more than two outcomes the representation has not been restricted. Biseparable utility includes many theories (see the Introduction) and the following theorem therefore pertains to all these theories. Statement (ii) characterizes concave utility for all these models. Statement (i) additionally characterizes *subadditivity* of W: $W(A) + W(A^c) \leq 1$. It is equivalent to convexity on binary partitions $\{E, E^c\}$. It is the prevailing empirical finding and has sometimes been taken as definition of ambiguity aversion.

COROLLARY 8. *If Structural Assumption 2 and biseparable utility hold, then:*

(i) U is concave and W is subadditive if and only if $\succcurlyeq$ is convex on every $\mathscr{F}_E$.

(ii) U is concave if and only if $\succcurlyeq$ is convex on every $\mathscr{F}_E^{\uparrow}$. This holds if and only if $\succcurlyeq$ is convex on one set $\mathscr{F}_E^{\uparrow}$ with E nondegenerate. □

[10] Once concavity has been derived, we are close to differentiability (lemma A.1). Then necessary and sufficient conditions for complete differentiability can be stated in terms of vanishing limits of risk premia (Nielsen 1999), a condition which has the same, commonly accepted, observability status as continuity. The task of our paper is, however, to derive concavity.

[11] One has to combine their Proposition 1 with the equivalence of (i) and (iv) in their Theorem 1.

2.5 Further implications for existing results on uncertainty in the literature

Schmeidler (1989), the most famous work in ambiguity theory, assumed an Anscombe-Aumann framework: a set of prizes is given, and the outcome set is the set of simple probability distributions over the prize space. That is, the outcome set is a convex subset of a linear space. Acts map states to outcomes. Utility over outcomes is assumed to be expected utility, i.e., it is linear with respect to probabilistic mixtures of outcomes. This is Structural Assumption 2 but with a multi-dimensional outcome space instead of our one-dimensional I. Schmeidler (1989 the Proposition) showed that W is convex if and only if $\succcurlyeq$ is convex—called uncertainty aversion. This follows from the special case of the main Corollary 7 with utility linear. That the outcome space is multi-dimensional changes nothing in our proofs.[12]

Most studies of multiple priors models, including $\alpha-$maxmin models, used the Anscombe-Aumann framework with linear utility. Exceptions without this restriction include Alon and Schmeidler (2014), Casadesus-Masanell et al. (2000), and Ghirardato et al. (2003). Our results characterize concave utility for the latter studies.

Several studies assumed linear utility as did Schmeidler (1989), and then gave various necessary and sufficient conditions for convex weighting functions alternative to our convexity: Chateauneuf (1991) and Kast and Lapied (2003) for monetary outcomes, and Wakker (1990) for the Anscombe-Aumann framework.

Three results in the literature come close to our main Theorem 3, in deriving concavity not on a full product set but on a comoncone. The first is Wakker (1994 Theorem 24).[13] He applied our main theorem dually to probability weighting instead of utility (similarly as our derivation of convexity of W in Corollary 7). His proof was complex and heavily used continuity in probability weighting and utility. His result will be generalized by Wakker and Yang (2018), who also formalize the aforementioned duality. The second result close to our main Theorem 3 or, more precisely, to our Corollary 6, is Theorem 3 in Chateauneuf and Tallon (2002). They assumed differentiable utility, whence they could skip Steps 2 and 3 of our proof in Appendix A. The third result close to Theorem 3 is Ghirardato and Marinacci (2001 Theorem 17). They assumed continuity, and showed how Debreu and Koopmans (1982) can be used as in Step 1 of our proof in Appendix A. Our proof shows how to add Steps 2 and 3 to their proof.[14]

[12] For necessity of the preference condition, the proof of Lemma B.1 works with the first inequality an equality. For sufficiency, all o terms in the proof of Lemma B.2 are exactly 0. Sufficiency can also be obtained by taking two outcomes $\gamma \succ \beta$ and equating lotteries over them with $I = [0,1]$.

[13] Wakker (1994) used the term quasi-concave preference instead of our term convex preference.

[14] We thus show that the claim $X^0 \subset X^* \cup X_*$ on p. 887 line -13 in their proof holds true by ruling out the existence of β as in Figures A.2 and A.3.

For the special case of linear utility (in an Anscombe-Aumann framework), Cerreia-Vioglio et al. (2011) characterized general preference functionals with convex preferences. Rigotti et al. (2008) examined general convex preferences and specified results for several ambiguity models. Their Remark 1 discussed RDU with convex weighting functions, but did not specify how these are related to convex preferences. Our main Corollary 7 shows that concavity of utility is necessary and sufficient for that relation to be an equivalence.

2.6 Conclusion

We have provided a general technique to obtain convex/concave utility and weighting functions. Fields of application include intertemporal choice, utilitarian welfare aggregations, risk, and, the context chosen in this paper, decision under uncertainty. There this paper generalized and improved virtually all existing theorems, and Wakker and Yang (2018) will do so for risk. Knowledge of Corollaries 4 and 7 will be useful for everyone working in decision theory. Convexity with respect to outcome-mixing is more powerful than had been known before.

Appendix A. Proof of Theorem 3

We first list some well-known properties of concave functions (Van Rooij and Schikhof 1982 §1.2).

LEMMA A.1 *If U is concave and strictly increasing on I, then: (a) U is continuous on I except possibly at $min(I)$ (if it exists). On $int(I)$: (b) U has right derivative U'_r and left derivative U'_ℓ everywhere; (c) $U'_\ell(\alpha) \geq U'_r(\alpha) \geq U'_\ell(\alpha') > 0$ for all $\alpha' > \alpha$; (d) U is differentiable almost everywhere.*

PROOF. As regards positivity in (c), if a left or right derivative were 0 somewhere in $int(U)$ then it would be 0 always after, contradicting strict increasingness of U. The other results are well-known (Van Rooij and Schikhof 1982 §1.2). □

PROOF OF MAIN THEOREM 3. If U is concave then so is the EU functional, so that it is quasi-concave, implying convexity of $\succcurlyeq$. (This also follows from Lemma B.1.)

In the rest of this appendix we assume convexity of $\succcurlyeq$, and derive concavity of U. We write $\pi_1 = W(s_1), \pi_2 = 1 - \pi_1$. By nondegeneracy, $0 < \pi_1 < 1$. We suppress states from acts and write (x_1, x_2) for $(s_1 : x_1, s_2 : x_2)$. As explained in the main text, we cannot use hedging techniques in our proofs. Instead, we will often derive contradictions of convexity of p by constructing a "risky" act $r = (r_1, r_2)$ and a "close-to-certain" act $c = (c_1, c_2)$ with $r_1 > c_1 \geq c_2 > r_2$, such that

$$m := \lambda r + (1-\lambda)c;\ r \succ m; c \succ m \tag{A.1}$$

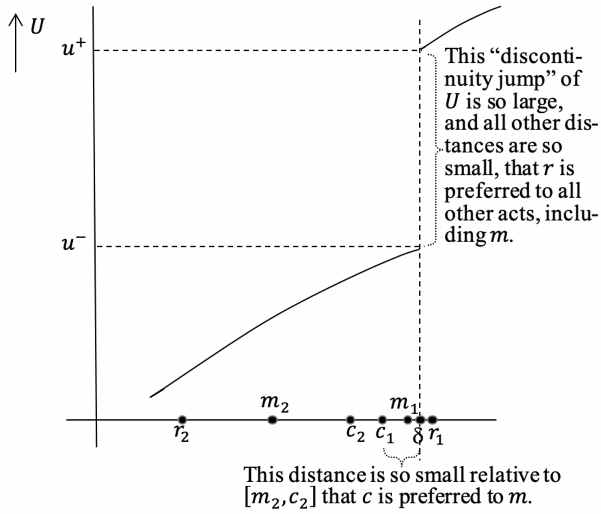
for some $0 < \lambda < 1$ (mostly $\lambda = 0.5$), with m called the "middle" act.

LEMMA A.2 *U is continuous except possibly at $inf\, I$.*

PROOF. See Figure A.1. Assume, for contradiction, that U is not continuous at an outcome $\delta > inf\, I$. We construct r, m, c as in Eq. A.1 with further $r_1 \geq \delta > m_1$.

Define $u^- = sup\{U(\beta) : \beta < \delta\}$. Define $u^+ = U(\delta)$ if $\delta = max(I)$ and $u^+ = inf\{U(\alpha) : \alpha > \delta\}$ otherwise. By discontinuity, $u^+ > u^-$. By taking $r_2 < \delta$ sufficiently close to δ we can get $U(r_2)$ as close to u^- as we want. We take it so close that $\pi_1(u^+) + \pi_2 U(r_2) > u^-$. This will ensure that r, the only act with its first outcome exceeding δ, is strictly preferred to all other acts, in particular, to m. We next choose c_2 strictly between r_2 and δ and define $m_2 = (r_2 + c_2)/2$. We then take c_1 strictly between c_2 and δ so close to δ that $\pi_1(u^-) + \pi_2 U(m_2) < \pi_1 U(c_1)) + \pi_2(U(c_2)$. This will ensure that $c \succ m$ if we ensure that $m_1 < \delta$. For the latter purpose we define

FIGURE A.1. Proof of Continuity of U



A discontinuity of U violates convexity of $\succcurlyeq$.

$r_1 = \delta$ if $\delta = max(I)$, and otherwise $r_1 > \delta$ so close to δ that $m_1 := (r_1 + c_1)/2 < \delta$. In both cases, $U(r_1) \geq u^+$. The acts r, m, c are as in Eq. A.1 with $\lambda = 1/2$. *QED*

Because U is strictly increasing, it suffices to prove concavity outside $inf\, I$. That is, we assume that I has no minimum. Assume for contradiction that U is not concave. Then there exist $0 < \lambda' < 1$ and outcomes $\alpha' < \gamma'$ such that $U(\lambda'\gamma' + (1-\lambda')\alpha') < \lambda' U(\gamma') + (1-\lambda')U(\alpha')$. Define ℓ as the line through $(\alpha', U(\alpha'))$ and $(\gamma', U(\gamma'))$. By continuity of U, we can define α as the maximum outcome between α' and $\lambda'\gamma' + (1-\lambda')\alpha'$ with $(\alpha, U(\alpha))$ on (or above) ℓ, and γ as the minimum outcome between γ' and $\lambda'\gamma' + (1-\lambda')\alpha'$ with $(\gamma, U(\gamma))$ on (or above) ℓ. We have $\gamma > \alpha$ and

$$U(\lambda\gamma + (1-\lambda)\alpha) < \lambda U(\gamma) + (1-\lambda)U(\alpha) \tag{A.2}$$

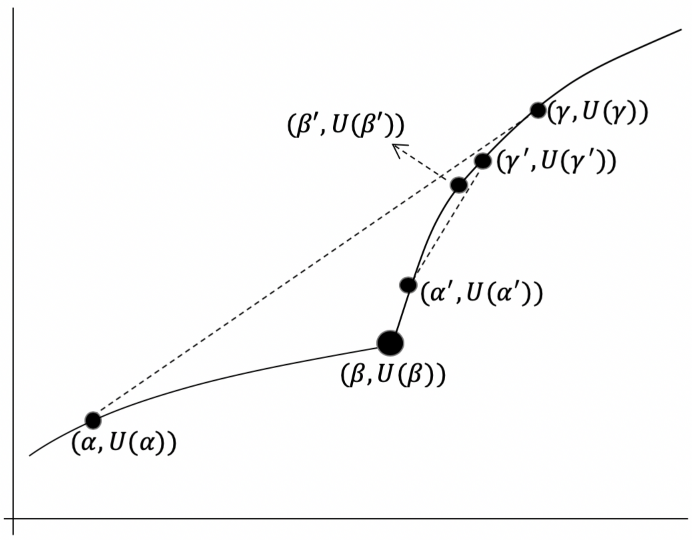
for all $0 < \lambda < 1$ (Figure A.2).

STEP 1 [At most one nonconcavity kink]. Assume that U is not concave on some "middle" interval $M \subset I$, with L and R the intervals in I to the left and right of M, possibly empty. Applying Debreu and Koopmans (1982 Theorem 2) to the additive representation $(1-\pi_1)U(x_1) + \pi_1 U(x_2)$ on $L \times M$ and to the additive representation $(1-\pi_1)U(x_1) + \pi_1 U(x_2)$ on $M \times R$ then implies that U is strictly concave on L and R. Applying this result to smaller and smaller subintervals M of

$[\alpha,\gamma]$ there is one β in $[\alpha,\gamma]$ such that U is concave above and below β.

For the remainder of the proof, we could use Debreu and Koopmans (1982 Theorem 6). They provide a concavity index according to which U would be infinitely *convex* at the nonconcavity kink, then would have to be more, so infinitely, concave at every other point, but being concave there it can be infinitely concave at no more than countably many points, and a contradiction has resulted. This reasoning is advanced and cannot be written formally very easily, because of which we provide an independent proof.

FIGURE A.2. Contradiction from nonconcavity of U: Steps 1 & 2



STEP 1: There exists only one kink β; strict concavity outside β.
STEP 2: $\beta \neq \beta' := \pi_1\gamma + \pi_2\alpha$ cannot be.

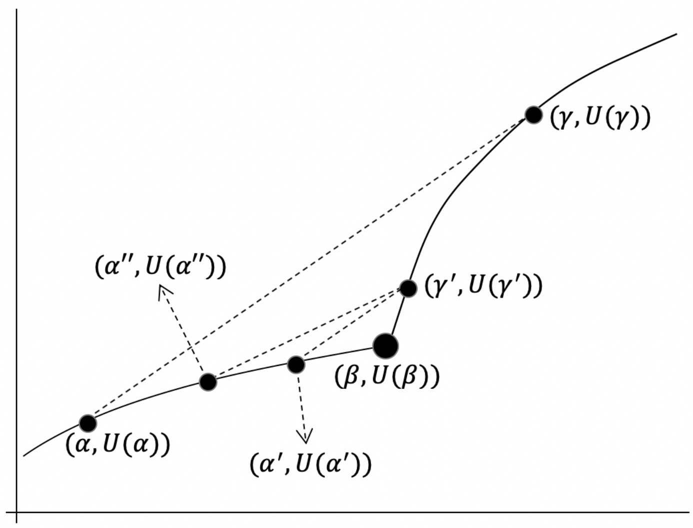
STEP 2 [Nonconcavity kink must be exactly at $\pi_1\gamma+\pi_2\alpha$]. We next show that β can only be $\beta' := \pi_1\gamma+\pi_2\alpha$. Assume, for contradiction, that $\pi_1\gamma+\pi_2\alpha$ were located at a point β' different than β (Figure A.2). Then there would be a small interval around β', not containing α, β, or γ, where U would be strictly concave. We could then find $\gamma' > \beta' > \alpha'$ in the small interval with exactly $\pi_1\gamma'+\pi_2\alpha' = \beta'$ and satisfying the strict concavity inequality

$$\pi_1 U(\gamma') + \pi_2 U(\alpha') < U(\beta'). \tag{A.3}$$

We define some acts, where, on our domain $\mathscr{F}'$ considered, always the best outcome is the first: $c = (\beta',\beta'), r = (\gamma,\alpha), m = (\gamma',\alpha')$, and we show that Eq. A.1 holds. Because β' is the same π_1,π_2 mix of γ and α as of γ' and α', there exists $0 < \mu < 1$ such that $m = \mu r + (1-\mu)c$. We have $r \succ c \succ m$, the first preference by the convexity "risk-seeking" inequality Eq. A.2 (with $\lambda = \pi_1$) and the second by the concavity "risk-averse" inequality Eq. A.3. Eq. A.1 is satisfied.

STEP 3 [Nonconcavity kink cannot be at $\pi_1\gamma + \pi_2\alpha$ either]. We construct, for contradiction, two outcomes γ', α' or γ', α'' that are as γ, α in Eq. A.2 but for whom an analogous β as just constructed does not exist though. See Figure A.3.

FIGURE A.3. Contradiction from nonconcavity of U: Step 3



STEP 3: We replace α by α'' and then Step 2 is violated.

Consider the line through $(\alpha, U(\alpha))$ and $(\gamma, U(\gamma))$. Because of concavity of U on $[\alpha, \beta)$ and $[\beta, \gamma)$, the right derivative of U is decreasing on each of these two intervals. Because the line mentioned lies above the graph of U, its slope exceeds the right derivative of U everywhere on $[\alpha, \beta)$ and is below it everywhere on $[\beta, \gamma)$. Now take any α' strictly between α and β, and take the line parallel to the line through $(\alpha, U(\alpha))$ and $(\gamma, U(\gamma))$. Its first intersection with the graph of U exceeding α' is γ'. Inspection of right derivates shows that $\gamma' > \beta$. If $\pi_1\gamma' + \pi_2\alpha' \neq \beta$ then we have a contradiction with Step 2 and we are done. If we have equality after all, then we take a line through $(\gamma', U(\gamma'))$ that is strictly between the two parallel lines. Its first intersection with the graph of U is $(\alpha'', U(\alpha''))$, and α'' and γ' are as in Eq. A.2 but the outcome β has $\pi_1\gamma' + \pi_2\alpha'' \neq \beta$. Contradiction has resulted and we are done. The proof of Theorem 3 now is complete. $\square$

The following example shows that the main Theorem 3 does not hold if U is only assumed nondecreasing instead of strictly increasing.

EXAMPLE A.3 In this example, all assumptions of Theorem 3 are satisfied except that U is nondecreasing instead of strictly increasing. $\succcurlyeq$ is convex but U is not concave.

$I = \mathbb{R}$, $W(s_1) = \pi_1 = \pi_2 = 0.5$, and U is the sign function. More precisely, $U(\alpha) = 1$ if $\alpha > 0$, $U(0) = 0$, and $U(\alpha) = -1$ if $\alpha < 0$. We show that $\succcurlyeq$ is convex. We consider an exhaustive list of cases of three acts m, r, c, with the cases ordered by the preference value of m. In each case we state some implications from $m \prec r$ and $m \prec c$ that readily preclude m from being a mixture of r and c. We suppress states and denote acts by the utilities of their outcomes. Thus, $(1,-1)$ denotes an act with a positive first outcome and a negative second outcome. We write $d'_j = U(d_j)$ for $d = c, m, r$ and $j = 1, 2$.

(1) If $m'_1 = -1$ then $r'_1 \geq 0$ and $c'_1 \geq 0$. (2) If $(m'_1, m'_2) = (0,-1)$ then $r'_1 \geq 0$ and $c'_1 \geq 0$ and either one of the latter is positive or $(r'_1, r'_2) = (c'_1, c'_2) = (0,0)$. (3) If $(m'_1, m'_2) = (0,0)$ or $(m'_1, m'_2) = (1,-1)$ then both (r'_1, r'_2) and (c'_1, c'_2) are $(1,0)$ or $(1,1)$. (4) If $(m'_1, m'_2) = (1,0)$ then $(r'_1, r'_2) = (c'_1, c'_2) = (1,1)$. (5) $(m'_1, m'_2) = (1,1)$ cannot be. $\square$

Appendix B. Proofs for §6.3 and §6.4

For a weighting function W, we use the following notation: $\pi(E^R) = W(E \cup R) - W(R)$ is the *decision weight* of an outcome in RDU under event E if R is the *rank*, i.e., the event giving outcomes ranked better than the one under E. $\pi^b(E) = W(E)$ and $\pi^w(E) = 1 - W(E^c)$. W is additive (EU) if and only if we have rank independence, i.e., decision weights $\pi(E^R)$ do not dependend on the rank R. W is convex if and only if decision weights $\pi(E^R)$ are nondecreasing in the rank R (Eq. 2.3), which reflects a pessimistic attitude. If W is convex, then its Core consists of all probability measures P_ρ (defined in §6.4) and

$$\text{RDU is the infimum EU with respect to all } P_\rho. \tag{B.1}$$

The following observation provides sufficiency of the preference conditions in all our results.

LEMMA B.1 *If U is concave and W is convex, then RDU is concave and, hence, $\succcurlyeq$ is convex on $\mathscr{F}$ (thus on every $\mathscr{F}' \subset \mathscr{F}$).*

PROOF. Consider three acts $x, y, \lambda x + (1-\lambda)y$, contained in comoncones with orderings $\rho_x, \rho_y, \rho_\lambda$, respectively. Then, with the first inequality due to concave utility and the second due to convexity of W (Eq. B.1): $RDU(\lambda x + (1-\lambda)y) = \int_S (U(\lambda x + (1-\lambda)y)) dP_{\rho_\lambda} \geq \int_S (\lambda U(x) + (1-\lambda)U(y)) dP_{\rho_\lambda} = \lambda \int_S U(x) dP_{\rho_\lambda} + (1-\lambda) \int_S U(y) dP_{\rho_\lambda} = \lambda EU_{P_{\rho_\lambda}}(x) + (1-\lambda) EU_{P_{\rho_\lambda}}(y) \geq$

$\lambda EU_{P_{\rho_x}}(x) + (1-\lambda)EU_{P_{\rho_y}}(y) = \lambda RDU(x) + (1-\lambda)RDU(y)$. $\square$

PROOF OF COROLLARY 4. Necessity of the preference condition follows from Lemma B.1. Sufficiency follows from applying the main Theorem 3 to any two-dimensional subspace $\{\alpha_E\beta \in \mathscr{F} : \alpha \geq \beta\}$ with E nondegenerate. $\square$

PROOF OF COROLLARY 5. Both utilities being strictly increasing and continuous, we define the continuous strictly increasing φ by $U^2(.) = \varphi(U^1(.))$. Deviating from the notation elsewhere in this paper, all outcomes are expressed in U^1 units in this proof. It means that we replace I by $U^1(I)$ which again is a nonpoint interval, that U^1 is linear, and $U^2 = \varphi$. Eq. 2.4 can be rewritten as $y_j = P^1(A)x_j + (1-P^1(A))z_j$. Regarding necessity of the preference condition, if φ is concave then, by Corollary 4, $\succcurlyeq^2$ is convex w.r.t. outcome mixing, which, given the equality just rewritten, implies Eq. 2.5 and, hence, that $\succcurlyeq^2$ is more outcome-risk averse than $\succcurlyeq^1$. The rest of this proof concerns sufficiency. We assume that $\succcurlyeq^2$ is more outcome-risk averse than $\succcurlyeq^1$ and derive concavity of φ.

There exists a nondegenerate event A for $\succcurlyeq^1$, and a nondegenerate event B for $\succcurlyeq^2$. From now on we only consider binary acts $x_{1_A}x_2$ for $\succcurlyeq^1$ and $x_{1_B}x_2$ for $\succcurlyeq^2$, denoting them (x_1, x_2) (no confusion will arise). Define $0 < q = P^1(A) < 1$. Consider any $x = (x_1, x_2) \succcurlyeq^2 (z_1, z_2) = z$. To derive convexity of $\succcurlyeq^2$ on the set of binary acts considered here, we have to show:

$$\forall\, 0 < \lambda < 1 : \lambda x + (1-\lambda)z \succcurlyeq^2 z. \tag{B.2}$$

We first show it for $\lambda = q$. By continuity we can decrease x_1, x_2 into $x_1{}', x_2{}'$ such that $(x_1{}', x_2{}') \sim^2 (z_1, z_2)$. Define $y_j = qx_j{}' + (1-q)z_j$, $j = 1,2$. With these definitions, Eq. 2.4 is satisfied with x' instead of x. By Eq. 2.5, $y \succcurlyeq^2 z$. By monotonicity, $qx + (1-q)z \succcurlyeq^2 y \succcurlyeq^2 z$. Eq. B.2 holds for $\lambda = q$. By repeated application and transitivity, the equation follows for a subset of λs dense in $[0,1]$ and then, by continuity, for all λ. Convexity of $\succcurlyeq^2$ on the two-dimensional set of acts considered here has been proved. Concavity of φ follows from the main Theorem 3. $\square$

PROOF OF COROLLARY 6. RDU on a comoncone coincides with EU on that comoncone w.r.t. a finitely additive probability measure P, which is convex. Hence, concavity of U implies convexity of $\succcurlyeq$ by Lemma B.1. Conversely, assume that $\succcurlyeq$ is convex. Apply the main Theorem 3 to any two-dimensional subspace $\{\alpha_E\beta \in \mathscr{F} : \alpha \geq \beta\}$ of the comoncone SF' with E nondegenerate, and concavity of U follows. $\square$

The following lemma is the main step in deriving implications of the main Theorem 3 for weighting functions. The inequality in the lemma states that the decision weight of s_1 is nondecreasing in rank. It implies the same inequality for s_2 and is equivalent to convexity of W. Showing this for higher dimensions ($n > 2$) goes the same way as for two dimensions, which is why this lemma captures the essence.

LEMMA B.2 *Assume $n = 2$, RDU, and convexity of $\succcurlyeq$. Then* $W(s_1) \leq 1 - W(s_2)$.

PROOF. Take an outcome in $int(I)$, 0 wlog, at which U is differentiable. Wlog, $U(0) = 0$. We consider a small positive α tending to 0, with $o(\alpha)$, or o_α for short, the usual notation for a function with $lim_{\alpha\to 0}\frac{o_\alpha}{\alpha} = 0$. In other words, in first-order approximations o_α can be ignored. We write $\pi_1 = W(s_1), \pi_2{}' = W(s_2)$.

Assume $\pi_1 > 0$ and $\pi_2{}' > 0$; otherwise we are immediately done. Because of continuity of U on $int(I)$ and differentiability at 0, we can obtain, for all α close to 0, the left indifference in

$$(\pi_2{}'\alpha, 0) \sim (0, \pi_1\alpha + o_\alpha) \preccurlyeq (\mu\pi_2{}'\alpha, (1-\mu)(\pi_1\alpha + o_\alpha)). \tag{B.3}$$

The preference is discussed later. We compare two values: the $\mu, 1-\mu$ mixture of the RDU values (which are the same) of the left two acts and the RDU value of their $\mu, 1-\mu$ mixture, which is the right act. We take $\mu > 0$ so small that the left outcome $\mu\pi_2{}'\alpha$ in the mixture is below the right outcome. Informally, by local linearity, in a first-order approximation the only difference between the two values compared is that for the left value the left outcome $\pi_2{}'\alpha$ receives the highest-outcome decision weight π_1 whereas for the right value it receives the lowest-outcome decision weight $1 - \pi_2{}'$. Convexity of $\succcurlyeq$ implies the preference in Eq. B.3, which implies $1 - \pi_2{}' \geq \pi_1$.

Formally, note that different appearances of o_α can designate different functions. Thus we can, for instance, write, for constants k_1 and k_2 independent of α: $k_1 o_\alpha + k_2 o_\alpha = o_\alpha$. The following is most easily first read for linear utility, when all terms o_α are zero. Write $u' = U'(0)$; μ can be chosen independently of α. Here is the comparison of the aforementioned two values: $\mu\pi_1 u'\pi_2{}'\alpha + o_\alpha + (1-\mu)\pi_2{}'u'\pi_1\alpha + o_\alpha \leq (1-\pi_2{}')u'\mu\pi_2{}'\alpha + o_\alpha + \pi_2{}'u'(1-\mu)\pi_1\alpha + o_\alpha$. Dividing by $\mu u'\pi_2{}'\alpha$, we obtain $\pi_1 \leq 1 - \pi_2{}' + \frac{o_\alpha}{\alpha}$. Now $\pi_1 \leq 1 - \pi_2{}'$ follows. □

PROOF OF THE MAIN COROLLARY 7. Necessity of the preference condition follows from Lemma B.1. We, therefore, assume convexity of $\succcurlyeq$. Concavity of U follows from considering any nondegenerate event E and applying the main Theorem 3 to the set of acts $(E : x_1, E^c x_2)$ with $x_1 \geq x_2$. We finally derive convexity of W.

Assume A, B, B' as in Eq. 2.3. Write $C = B' - B, R = S - (A \cup B')$. Take $\gamma > \beta \in int(I)$, and consider $\mathscr{F}^* = \{(B : \gamma, C : x_1, A : x_2, R : \beta) \in \mathscr{F} : \gamma \geq x_1 \geq \beta, \gamma \geq x_2 \geq \beta\}$. This space is isomorphic to the space $\mathscr{F}$ of Lemma B.2 and the convexity inequality needed here follows from the one of that lemma. Details are as follows. Take outcome space $I^* = [\beta, \gamma]$, $s_1^* = C, s_2^* = A$, and weighting function $W^*(E) = \frac{W(E \cup B) - W(B)}{W(A \cup B \cup C) - W(B)}$. (If the denominator is 0, then the convexity inequality is trivially satisfied.) The inequality $W^*(\{s_1, s_2\}) - W^*(s_1) \geq W^*(s_2)$ in Lemma B.2 is the same as the required $W(A \cup B') - W(B') \geq W(A \cup B) - W(B)$. □

PROOF OF COROLLARY 8. Statement (i) follows from the main Corollary 7 because convexity of W on every $\mathscr{F}_E$ is equivalent to subadditivity. Statement (ii) follows immediately from the main Theorem 3. □

REFERENCES

Alon, Shiri and David Schmeidler (2014) "Purely Subjective Maxmin Expected Utility," *Journal of Economic Theory* 152, 382–412.

Arrow, Kenneth J. (1971) "*Essays in the Theory of Risk-Bearing*." North-Holland, Amsterdam.

Arrow, Kenneth J. and Leonid Hurwicz (1972) "An Optimality Criterion for Decision Making under Ignorance." *In* Charles F. Carter and James L. Ford (1972) *Uncertainty and Expectations in Economics*, 1–11, Basil Blackwell and Mott Ltd., Oxford, UK.

Baillon, Aurélien, Bram Driesen, and Peter P. Wakker (2012) "Relative Concave Utility for Risk and Ambiguity, " *Games and Economic Behavior* 75, 481–489.

Bell, David E. (1985) "Disappointment in Decision Making under Uncertainty," *Operations Research* 33, 1–27.

Birnbaum, Michael H. (2008) "New Paradoxes of Risky Decision Making," *Psychological Review* 115, 463–501.

Casadesus-Masanell, Ramon, Peter Klibanoff, and Emre Ozdenoren (2000) "Maxmin Expected Utility over Savage Acts with a Set of Priors," *Journal of Economic Theory* 92, 35–65.

Cerreia-Vioglio, Simone, Fabio Maccheroni, Massimo Marinacci, and Luigi Montrucchio (2011) "Uncertainty Averse Preferences," *Journal of Economic Theory* 146, 1275??1330.

Chateauneuf, Alain (1991) "On the Use of Capacities in Modeling Uncertainty Aversion and Risk Aversion," *Journal of Mathematical Economics* 20, 343–369.

Chateauneuf, Alain and Jean-Marc Tallon (2002) "Diversification, Convex Preferences and Non-Empty Core," *Economic Theory* 19, 509–523.

Crouzeix, Jean-Pierre and P.O. Lindberg (1986) "Additively Decomposed Quasiconvex Functions," *Mathematical Programming* 35, 42–57.

De Giorgi, Enrico G. and Ola Mahmoud (2016) "Diversification Preferences in the Theory of Choice," *Decisions in Economics and Finance* 39, 143–174.

Dean, Mark and Pietro Ortoleva (2017) "Allais, Ellsberg, and Preferences for Hedging" *Theoretical Economics* 12, 377–424.

Debreu, Gérard (1959) "*Theory of Value. An Axiomatic Analysis of Economic Equilibrium.*" Wiley, New York.

Debreu, Gérard and Tjalling C. Koopmans (1982) "Additively Decomposed Quasiconvex Functions," *Mathematical Programming* 24, 1–38.

Ebert, Udo (2004) "Social Welfare, Inequality, and Poverty when Needs Differ," *Social Choice and Welfare* 23, 415–448.

Epstein, Larry G. (1999) "A Definition of Uncertainty Aversion," *Review of Economic Studies* 66, 579–608.

Gajdos, Thibault, Takashi Hayashi, Jean-Marc Tallon, and Jean-Christophe Vergnaud (2008) "Attitude towards Imprecise Information," *Journal of Economic Theory* 140, 27–65.

Ghirardato, Paolo, Fabio Maccheroni, Massimo Marinacci, and Marciano Siniscalchi (2003) "A Subjective Spin on Roulette Wheels," *Econometrica* 71, 1897–1908.

Ghirardato, Paolo and Massimo Marinacci (2001) "Risk, Ambiguity, and the Separation of Utility and Beliefs," *Mathematics of Operations Research* 26, 864–890.

Ghirardato, Paolo and Massimo Marinacci (2002) "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory* 102, 251–289.

Ghirardato, Paolo, Fabio Maccheroni, and Massimo Marinacci (2004) "Differentiating Ambiguity and Ambiguity Attitude," *Journal of Economic Theory* 118, 133–173.

Gilboa, Itzhak (1987) "Expected Utility with Purely Subjective Non-Additive Probabilities," *Journal of Mathematical Economics* 16, 65–88.

Gilboa, Itzhak and David Schmeidler (1989) "Maxmin Expected Utility with a Non-Unique Prior," *Journal of Mathematical Economics* 18, 141–153.

Gul, Faruk (1992) "Savage's Theorem with a Finite Number of States," *Journal of Economic Theory* 57, 99–110. ("Erratum," *Journal of Economic Theory* 61, 1993, 184.)

Gul, Faruk and Wolfgang Pesendorfer (2015) "Hurwicz Expected Utility and Subjective Sources," *Journal of Economic Theory* 159, 465–488.

Harvey, Charles M. (1986) "Value Functions for Infinite-Period Planning," *Management Science* 32, 1123–1139.

Heufer, Jan (2012) "Quasiconcave Preferences on the Probability Simplex: A Nonparametric Analysis," *Mathematical Social Sciences* 65, 21–30.

Heufer, Jan (2014) "Nonparametric Comparative Revealed Risk Aversion," *Journal of Economic Theory* 153, 569–616.

Izhakian, Yehuda (2017) "Expected Utility with Uncertain Probabilities Theory," *Journal of Mathematical Economics* 69, 91–103.

Jaffray, Jean-Yves (1994) “Dynamic Decision Making with Belief Functions.” *In* Ronald R. Yager, Mario Fedrizzi, and Janus Kacprzyk (eds.) *Advances in the Dempster-Shafer Theory of Evidence,* 331–352, Wiley, New York.

Kahneman, Daniel and Amos Tversky (1979) “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica* 47, 263–291.

Kast, Robert and André Lapied (2003) “Comonotonic Book Making and Attitudes to Uncertainty,” *Mathematical Social Sciences* 46, 1–7.

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji (2005) “A Smooth Model of Decision Making under Ambiguity,” *Econometrica* 73, 1849–1892.

Köbberling, Veronika and Peter P. Wakker (2003) “Preference Foundations for Nonexpected Utility: A Generalized and Simplified Technique,” *Mathematics of Operations Research* 28, 395–423.

Köszegi, Botond and Matthew Rabin (2006) “A Model of Reference-Dependent Preferences,” *Quarterly Journal of Economics* 121, 1133–1165.

Kreps, David M. and Evan L. Porteus (1978) “Temporal Resolution of Uncertainty and Dynamic Choice Theory,” *Econometrica* 46, 185–200.

Loomes, Graham and Robert Sugden (1986) “Disappointment and Dynamic Consistency in Choice under Uncertainty,” *Review of Economic Studies* 53, 271–282.

Luce, R. Duncan (2000) “*Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches*.” Lawrence Erlbaum Publishers, London.

Luce, R. Duncan and Howard Raiffa (1957) “*Games and Decisions*.” Wiley, New York.

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green (1995) “*Microeconomic Theory*.” Oxford University Press, New York.

Miyamoto, John M. (1988) “Generic Utility Theory: Measurement Foundations and Applications in Multiattribute Utility Theory,” *Journal of Mathematical Psychology* 32, 357–404.

Monteiro, Paulo Klinger (1999) “Quasiconcavity and the Kernel of a Separable Utility,” *Economic Theory* 13, 221–227.

Nielsen, Lars Tyge (1999) “Differentiable von Neumann-Morgenstern Utility,” *Economic Theory* 14, 285–296.

Pfanzagl, Johann (1959) “A General Theory of Measurement —Applications to Utility,” *Naval Research Logistics Quarterly* 6, 283–294.

Quiggin, John (1982) “A Theory of Anticipated Utility,” *Journal of Economic Behaviour and Organization* 3, 323–343.

Rigotti, Luca, Chris Shannon, and Tomasz Strzalecki (2008) “Subjective Beliefs and ex Ante Trade,” *Econometrica* 76, 1167–1190.

Savage, Leonard J. (1954) “*The Foundations of Statistics*.” Wiley, New York. (Second edition

1972, Dover Publications, New York.)

Schmeidler, David (1986) "Integral Representation without Additivity," *Proceedings of the American Mathematical Society* 97, 255–261.

Schmeidler, David (1989) "Subjective Probability and Expected Utility without Additivity," *Econometrica* 57, 571–587.

Shapley, Lloyd S. (1971) "Cores of Convex Games," *International Journal of Game Theory* 1, 11–26.

Strzalecki, Tomasz (2013) "Temporal Resolution of Uncertainty and Recursive Models of Ambiguity Aversion," *Econometrica* 81, 1039–1074.

Tversky, Amos and Daniel Kahneman (1992) "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5, 297–323.

Van Rooij, Arnoud C.M. and Wilhelmus H. Schikhof (1982) "*A Second Course on Real Functions*." Cambridge University Press, Cambridge, UK.

Viscusi, W. Kip (1989) "Prospective Reference Theory: Toward an Explanantion of the Paradoxes," *Journal of Risk and Uncertainty* 2, 235–264.

von Neumann, John and Oskar Morgenstern (1944, 1947, 1953) "*Theory of Games and Economic Behavior*." Princeton University Press, Princeton NJ.

Wakker, Peter P. (1989) "*Additive Representations of Preferences, A New Foundation of Decision Analysis*." Kluwer Academic Publishers, Dordrecht.

Wakker, Peter P. (1990) "Characterizing Optimism and Pessimism Directly through Comonotonicity," *Journal of Economic Theory* 52, 453–463.

Wakker, Peter P. (1993) "Unbounded Utility for Savage's "Foundations of Statistics," and Other Models," *Mathematics of Operations Research* 18, 446–485.

Wakker, Peter P. (1994) "Separating Marginal Utility and Probabilistic Risk Aversion," *Theory and Decision* 36, 1–44.

Wakker, Peter P. (2010) "*Prospect Theory for Risk and Ambiguity.*" Cambridge University Press, Cambridge, UK.

Wakker, Peter P. and Amos Tversky (1993) "An Axiomatization of Cumulative Prospect Theory," *Journal of Risk and Uncertainty* 7, 147–176.

Wakker, Peter P. and Jingni Yang (2018) "Concave/Convex Weighting and Utility Functions for Risk," in preparation.

Wald, Abraham (1950) "*Statistical Decision Functions*." Wiley, New York.

Yaari, Menahem E. (1965) "Convexity in the Theory of Choice under Risk," *Quarterly Journal of Economics* 79, 278–290.

Yaari, Menahem E. (1969) "Some Remarks on Measures of Risk Aversion and on Their Uses," *Journal of Economic Theory* 1, 315–329.

Chapter 3

Concave/Convex Weighting and Utility Functions for Risk: A New Light on Classical Theorems[1]

Abstract

This paper analyzes concavity and convexity of utility and probability distortion functions for decision under risk (law-invariance). We generalize virtually all existing results and make them more appealing. In particular, we characterize concave utility for virtually all existing models of decision under risk. We also obtain completely general axiomatizations of strictly increasing concave/convex utility and probability distortion functions for rank-dependent utility and prospect theory. Unlike preceding results, we do not need to presuppose any continuity, let be differentiability. An example of a new light shed on classical results: whereas, in general, convexity/concavity with respect to probability mixing is mathematically distinct from convexity/concavity with respect to outcome mixing, in Yaari's dual theory these conditions are not only dual, as was well-known, but also logically equivalent, which had not been known before. It leads to a surprising new way to obtain sub-additivity of law-invariant risk measures.

JEL-classification: D81, C60
Keywords: convex preferences, quasiconcave utility, risk aversion, rank-dependent utility

[1] This chapter is joint with Peter P. Wakker.

3.1 Introduction

This paper generalizes virtually all existing characterizations of concave utility for decision under risk (law-invariance), and makes them more appealing. Surprisingly, using a duality between outcomes and probabilities, we can also apply our results to probability distortions, where we again generalize virtually all existing results on convexity and concavity, and again make them more appealing. Thus, for instance:

(1) We axiomatize concave/convex probability distortion functions for Quiggin's (1982) rank-dependent utility and Tversky & Kahneman's (1992) prospect theory in complete generality, and do so by the simple and standard convexity/concavity of preference (Theorem 15).

In (1), we do not need any prior assumption on continuity, differentiability, or even richness of the outcome set. Our results further include:

(2) Concave utility for Miyamoto's (1988) biseparable utility for risk, which, by Wakker (2010 Observation 7.11.1), includes many risk theories, such as Quiggin's (1982) rank-dependent utility and Tversky & Kahneman's (1992) prospect theory for risk. Further theories included are disappointment theory (Bell 1985; Loomes & Sugden 1986—for a disappointment function kinked at 0), RAM and TAX models (Birnbaum 2008), disappointment aversion (Gul 1992), original prospect theory (Kahneman & Tversky 1979) for gains and for losses, Luce's (2000) binary RDU, and prospective reference theory (Viscusi 1989).

(3) Loss aversion in Köszegi & Rabin's (2006) reference dependent model.

(4) Inequality aversion results for welfare theory (Ebert 2004), and related results of Chew & Mao (1995) for risk.

There is a tight relation between risk attitudes and Artzner et al.'s (1999) risk measures (Belles-Sampera, Guillen, and Santolino 2016; Goovaerts, Kaas, and Laeven 2010a). Thus, our results can be applied to risk measures. §3.5 shows that, whereas properties of risk measures have commonly been derived from mixtures or additions of outcomes, mixtures of probabilities provide an alternative tool.

3.2 Basic definitions

We consider a set $\mathscr{P}$ of probability distributions, called *lotteries* (generic notation P, Q, R), over a set X of *outcomes* (generic notation α, β, γ, or x_j). X can be finite or infinite, and its elements can be monetary or non-monetary. We assume that $\mathscr{P}$ contains all *simple* probability distributions, assigning probability 1 to a finite subset of X, with generic notation $(p_1 : x_1, \ldots, p_n : x_n)$, and possibly more distributions. A *preference relation*, i.e., a binary relation $\succcurlyeq$ on $\mathscr{P}$, is given; $\succ, \preccurlyeq, \prec, \sim$ are as usual. V *represents* $\succcurlyeq$ if $V : \mathscr{P} \to \mathbb{R}$ satisfies $P \succcurlyeq Q \Leftrightarrow V(P) \geq V(Q)$ for all lotteries $P, Q \in \mathscr{P}$. This implies weak ordering on $\mathscr{P}$; i.e., $\succcurlyeq$ is transitive and complete. Outcomes α are identified with degenerate lotteries $(1 : \alpha)$. Thus, $\succcurlyeq$ also denotes preferences over outcomes.

A *(probability) distortion function* w maps $[0,1]$ to $[0,1]$, is strictly increasing, and satisfies $w(0) = 0$ and $w(1) = 1$. We do not assume continuity of w. Discontinuities at $p = 0$ and $p = 1$ are of special empirical interest. For a distortion function w, and a function $U : X \to \mathbb{R}$, the *rank-dependent utility* (*RDU*) of a lottery P is

$$\int_{\mathbb{R}^+} w(P(U(\alpha) > \mu))d\mu - \int_{\mathbb{R}^-} (1 - w(P(U(\alpha) > \mu)))d\mu. \tag{3.1}$$

For a simple lottery $(p_1 : x_1, \ldots, p_n : x_n)$ with $U(x_1) \geq \cdots \geq U(x_n)$, the RDU can be rewritten as

$$\sum_{j=1}^{n} (w(p_1 + \cdots + p_j) - w(p_1 + \cdots + p_{j-1}))U(x_j). \tag{3.2}$$

Rank-dependent utility (*RDU*) holds if there exist w and U such that *RDU* represents $\succcurlyeq$. Then U is the *utility function*, and it represents $\succcurlyeq$ on X. The special case of RDU with w the identity is called *expected utility* (*EU*). We impose one more restriction on $\mathscr{P}$: RDU is well defined and finite for all its elements. A necessary and sufficient condition directly in terms of preferences—requiring preference continuity with respect to truncations of lotteries—is in Wakker (1993). A sufficient condition is that all lotteries are bounded (with an upper and lower bound contained in X).

ASSUMPTION 9. [Structural assumption] $\mathscr{P}$ is a set of lotteries over outcome set X containing all simple probability distributions. RDU holds. X contains at least three nonindifferent outcomes $\gamma \succ \beta \succ \alpha$. □

The main tool in our analysis is Theorem 3 of Wakker and Yang (2019), characterizing concave utility for Choquet expected utility under uncertainty.

3.3 Outcome mixing

This section considers outcome mixing. For this purpose, we reinforce our assumptions.

ASSUMPTION 10. [Structural assumption for monetary outcomes] Assumption 9 holds. Further, $X = I$ is a nonpoint interval and U is strictly increasing. □

We do not presuppose continuity of U. Unlike virtually all axiomatizations in the literature we, similarly do not need to assume continuity of the preference relation (except in Corollary 14). Throughout the past decades, authors have warned against the poblematic empirical status of continuity assumptions in preference axiomatizations (Ghirardato & Marinacci 2001a; Halpern 1999; Khan & Uyan 2018; Krantz et al. 1971 §9.1; Pfanzagl 1968 §6.6 and §9.5; Wakker 1988). The assumption is not merely technical but adds empirical content to the empirical axioms, and the problem is that it is unknown what that added empirical content is. Hence, given the purpose of preference axiomatizations to reveal the empirical content of theories, it is desirable to do without continuity if possible.[2] In our case, continuity of U on $int(I)$ comes free of charge, following from the empirical axioms. At the extremes ($inf(I)$ for concavity and $sup(I)$ for convexity) we have it optional. If continuity is considered to be desirable there, then we can get it by adding the corresponding continuity condition for $\succcurlyeq$.

For simplicity, we restrict the definition of convex preferences to simple lotteries, which will be strong enough to give all the desired implications. Preceding papers (e.g., Yaari 1987) defined the condition for risk by specifying an underlying state space and then extended the condition to nonsimple lotteries.[3] As we will show, imposing the preference conditions only on simple lotteries is enough to give all desired results. We can thus avoid the complications of defining underlying state spaces.

Because it is common in decision under risk to let concavity and convexity refer to probabilistic mixing, considered in the next section, we use a different term for outcome mixing. We call $\succcurlyeq$ *outcome-convex* if for each probability vector $p_1,.,,,,.p_n$ (assumed to add to 1) and $0 < \lambda < 1$ we have

$$\begin{gathered}(p_1 : x_1, \ldots, p_n : x_n) \succcurlyeq (p_1 : y_1, \ldots, p_n : y_n) \Rightarrow \\ (p_1 : \lambda x_1 + (1-\lambda)y_1, \ldots, p_n : \lambda x_n + (1-\lambda)y_n) \succcurlyeq (p_1 : y_1, \ldots, p_n : y_n).\end{gathered} \tag{3.3}$$

We call $\succcurlyeq$ *outcome-convex on* $\mathscr{P}' \subset \mathscr{P}$ if Eq. 3.3 holds whenever all lotteries in it are contained in $\mathscr{P}'$. We next show that this extension of convexity in Euclidean domains to the lottery domain gives convenient axiomatizations of widely used properties. A new result on Yaari's (1987)

[2] We do assume RDU, or biseparable utility, in our results. Köbberling & Wakker (2003) provided preference axiomatizations that do not assume continuity, but a weaker solvability condition. This condition still has observability problems, but to a lesser extent than continuity.

[3] For completeness, we give details. In the following results, the proofs of sufficiency of the preference conditions then remain unaltered because we only need simple lotteries for those. For necessity, concavity of U and convexity of w imply that the representing functional is concave (as in the proof of Wakker & Yang 2019 Lemma B.1) and, hence, surely quasi-concave. Then $\succcurlyeq$ is convex. This implies, in particular, that convexity for all simple lotteries is equivalent to convexity for all lotteries under RDU. For further extensions to nonsimple lotteries, see Mao and Hu (2012). Alternative outcome-operations without a state space can be obtained by taking probabilistically independent combinations of lotteries (Goovaerts, Kaas, and Laeven 2010b).

analog of this extension is given in the next section (Corollary 16).

A *comoncone* is a subset of lotteries $\{(p_1 : x_1, \ldots, p_n : x_n) : x_1 \geq \cdots \geq x_n\}$, with $n \geq 2$ fixed, the probability vector $p_1, \ldots, p_n$ fixed, and $0 < p_1 < 1$. We call $\succcurlyeq$ *comonotonic outcome-convex* if it is outcome-convex on every comoncone; that is, if Eq. 3.3 holds whenever $x_1 \geq \cdots \geq x_n$ and $y_1 \geq \cdots \geq y_n$. The following lemma, used in the proofs of the following theorems, prepares for a duality result discussed later.

LEMMA 11. *Consider a comoncone $\{(p_1 : x_1, \ldots, p_n : x_n) : x_1 \geq \cdots \geq x_n\}$. Under Assumption 10, U is concave if and only if $\succcurlyeq$ is outcome-convex on this comoncone.* □

Ghirardato & Marinacci (2001b) propagated the biseparable utility model for uncertainty. For risk, this was done by Miyamoto (1988), who used the term generic utility. He also emphasized that any result for his theory holds for the many models comprised, referenced in our introduction. We now apply our technique to his model. By $\mathscr{P}_p$ we denote the set of binary lotteries $\gamma_p\beta = (p : \gamma, 1-p : \beta)$, and by $\mathscr{P}_p^{\uparrow}$ we denote the subset with $\gamma \geq \beta$—it is a comoncone if $0 < p < 1$. *Biseparable utility* holds if there exist a utility function U and a distortion function w such that $RDU(\gamma_p\beta) = w(p)U(\gamma) + (1-w(p))U(\beta)$ (for $\gamma \succcurlyeq \beta$) represents $\succcurlyeq$ on the set of all binary lotteries.

THEOREM 12. *If Structural Assumption 10 holds except that biseparable utility holds instead of RDU, then U is concave if and only if $\succcurlyeq$ is outcome-convex on every $\mathscr{P}_p^{\uparrow}$. This holds if and only if $\succcurlyeq$ is outcome-convex on one set $\mathscr{P}_p^{\uparrow}$ with $0 < p < 1$.* □

Thus, we have characterized concave utility for virtually all existing models of risky choice (see introduction). For RDU, the preference conditions are equivalent to the stronger comonotonic outcome-convexity, as is easily verified. For EU, it is equivalent to the even stronger outcome-convexity. For EU, this result amounts to an alternative to the traditional characterizations based on weak risk aversion—preference for expected value—or strong risk aversion—aversion to mean-preserving spreads.

The preceding theorem characterized concavity of utility for RDU (and other theories), and Theorem 15 will characterize convexity of probability distortion for RDU. The following theorem efficiently characterizes the two properties jointly. Such "pessimistic" functionals have been widely used to represent downside risks, rather than overall preference values (Goovaerts, Kaas, and Laeven 2010a).

THEOREM 13. *Under Assumption 10, U is concave and w is convex if and only if $\succcurlyeq$ is outcome-convex.* □

Theorem 13 captures the two most-studied properties of RDU through one basic preference condition. Included are the widely studied law-invariant coherent risk measures that are comonotonically additive (Cornilly, Rüschendorf, and Vanduffel (2018). The theorem provides an interesting alternative to Chew, Karni, & Safra (1987). They showed, assuming differentiability, that concavity of U plus convexity of w is equivalent to aversion to mean-preserving spreads. Quiggin (1993 §6.2) provided an alternative proof, also assuming differentiability. Schmidt & Zank (2008) provided yet another proof, the only one available in the literature that did not assume differentiability; they still did assume continuity. It is desirable to avoid differentiability assumptions in preference axiomatizations because differentiability is even more problematic than continuity: unlike with continuity, for differentiability there is not even a preference condition to axiomatize it. Our derivations, therefore, neither assume differentiability. We obtain the following corollary, where for the definition of continuity and aversion to mean-preserving spreads we refer to Chew, Karni, & Safra (1987). We need to assume continuity because without it there are no results available in the literature on aversion to mean-preserving spreads.

COROLLARY 14. *Under Assumption 10 and continuity, outcome-convexity of $\succcurlyeq$ is equivalent to aversion to mean-preserving spreads.*

The result is remarkable because, at first sight, one condition concerns only outcome mixing whereas the other condition also involves probabilistic mixing. This surprising point was discussed by Quiggin (1993 §9.2) in a somewhat different context. Many papers have used aversion to mean-preserving spreads conditions in various forms. At the end of the appendix, we give details of several results relevant to the preceding analyses. Our paper shows that convexity conditions can serve as appealing alternative. This holds especially if the probabilities involved in mean-preserving spreads are subjective, implying that they are not directly observable, contrary to our preference condition. Thus, our condition can, for instance, serve to make the conditions in §5 of Gul & Pesendorfer (2015) directly observable. Gul & Pesendorfer (2015) used subjective probabilities as inputs in their axioms but subjective probabilities are not directly observable.

3.4 Probabilistic mixing

This section considers probabilistic mixing. For lotteries P, Q, $\lambda P \oplus (1-\lambda)Q$ denotes the probability measure assigning probability $\lambda P(\alpha) + (1-\lambda)Q(\alpha)$ to each outcome α, with a similar probability mix for each subset of outcomes instead of $\{\alpha\}$. Probabilistic mixing can be defined for general outcome sets X.

We call $\succcurlyeq$ *convex* if $P \succcurlyeq Q \Rightarrow \lambda P \oplus (1-\lambda)Q \succcurlyeq Q$. This property, suggesting a deliberate preference for randomization, has been widely studied in the literature (Agranov & Ortoleva

2017; Cerreia-Vioglio et al. 2019; Fudenberg, Lijima, & Strzalecki 2015; Machina 1985; Saito 2015; Sopher & Narramore 2000). The opposite condition is more commonly found empirically: $\succcurlyeq$ is *concave* if $P \succcurlyeq Q \Rightarrow\ P \succcurlyeq \lambda P \oplus (1-\lambda)Q$. It is widely studied for distorted risk measures (Tsanakas 2008). The following theorem offers a characterization in full generality, without using any restrictive assumption, continuity or otherwise.

FIGURE 1: Duality of probabilities versus outcomes, and w versus U.

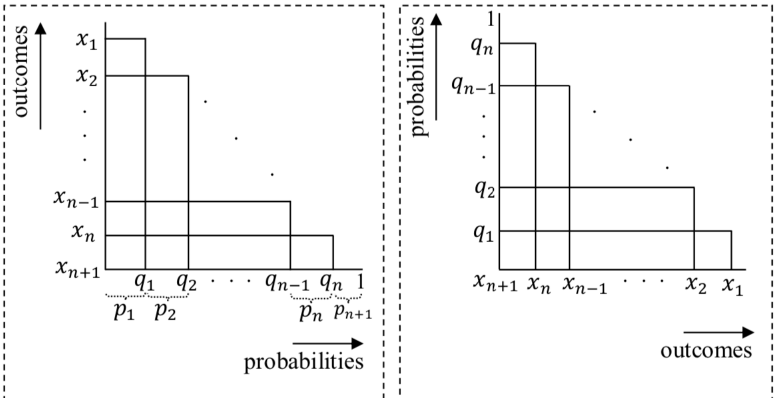


FIG. 1a: U is concave $\Leftrightarrow$ convexity w.r.t. ordinate-mixing (Lemma 3)

FIG. 1b: w is concave $\Leftrightarrow$ convexity w.r.t. ordinate-mixing (Theorem 9)

RDU results from Fig. a by transforming the abcissa by w and the ordinate by U, and then calculating the area of the figure.
RDU results from Fig b by transforming the abcissa by U and the ordinate by w, and then calculating the area of the figure.

THEOREM 15. *Under Assumption 9, convexity of $\succcurlyeq$ is equivalent to concavity of w, and concavity of $\succcurlyeq$ is equivalent to convexity of w.* $\square$

The proof is based on a remarkable duality between outcomes and probabilities—more precisely, goodnews probabilities or ranks q_j(the probability of receiving a better outcome, defined formally in the proof of Theorem 15 in the Appendix). Figure 1 illustrates the duality for a lottery $(p_1 : x_1, \dots, x_n)$ with $x_1 > \cdots > x_{n-1} > x_n = 0$ and $n \geq 3$, with $q_j := p_1 + \cdots + p_j$. It shows that outcomes x_j and ranks q_{n-j} play the same mathematical role in calculating RDU, as do utility and probability distortion. The duality is explained in general in the proof of Theorem 15. Because of it, every theorem on utility, such as Lemma 11, automatically gives a theorem on probability distortion, being Theorem 15 in this case. Similar dualities were exploited by Yaari (1987), Abdellaoui (2002), Abdellaoui & Wakker (2005), and Werner and Zank (2019).[4] The duality in Figure 1 also illustrates that Quiggin's (1982) insight, that one should use differences rather than absolute levels of probability distortion functions, is the dual of the insight of the

[4] Recognizing them in particular situations is nontrivial. Thus, one of us, while well acquainted with Yaari (1987), did not recognize this duality in Wakker (1994 Theorem 24) and Wakker (2010 p. 192 footnote 8).

marginal utility revolution (Jevons 1871; Menger 1871; Walras 1874), being that changes of utility, rather than absolute levels themselves, are often basic.

The axiomatization in Theorem 15 of convexity of w through a widely used preference condition is appealing. If there are only two nonindifferent outcomes, deviating from Assumption 9, then w is only ordinal and can be any strictly increasing function, thus can always be convex but never needs to be. Hence, Theorem 15 has characterized convexity of w as general as can be.

We present yet another surprising equivalence, combining Theorems 13 and 15 for the special case of Yaari's (1987) dual model (RDU with linear utility). Linear utility is also commonly assumed for coherent risk measures (Artzner et al. 1999).

COROLLARY 16. *Under Yaari's (1987) dual model (Assumption 9, with X a nonpoint interval I and U the identity), $\succcurlyeq$ is outcome-convex (concave) if and only if it is concave (convex).* □

Thus the conditions, concerning mixing in two different dimensions—"horizontal" and "vertical"—are not only each others' duals, but they are also logically equivalent here. Röell (1987 §1) discussed these conditions in Yaari's model, but was not aware of their equivalence, nor has anyone else been as yet. The above corollary gives a new way of obtaining sub-additivity of law-invariant (probabilistically sophisticated) risk measures: by mixing probabilities rather than, as is common in the literature, outcomes.

3.5 Further implications for existing results in the literature

Yaari (1987) considered the special case of RDU for risk with linear utility. He characterized convexity of w through aversion to mean-preserving spreads, which is a special case of Chew, Karni, & Safra's (1987) theorem. Quiggin (1993 §9.1) and Röell (1987) similarly derived this result for linear utility. As our Theorem 15 showed, convexity with respect to probabilistic mixing provides an appealing alternative condition. It would have fitted better with the affinity condition for outcome addition that Yaari (1987) used, and the affinity condition for outcome mixing that Röell (1987) used, to axiomatize RDU with linear utility.

A surprising application concerns Köszegi & Rabin's (2006) reference dependent model. Masatlioglu & Raymond (2016) showed how Köszegi & Rabin's choice-acclimating personal equilibrium (CPE) is a special case of RDU. Loss aversion in Köszegi & Rabin's model then holds if and only if the probability distortion function in the equivalent RDU model is convex. Masatlioglu & Raymond's Propositions 3 and 10 used Wakker's (1994) version of our Theorem 15 to characterize loss aversion. They wrote: "we were able to demonstrate a previously unknown relationship between loss aversion/loving behavior and attitudes toward mixing lotteries within the CPE framework" (p. 2792) and "our results allow us to bring 20 years of existing experimental evidence to bear on CPE" (p. 2773). They required monetary outcomes and continuous utility. Our Theorem 15 shows that those restrictions can be dropped, and that the result holds in full generality. Their Proposition 6 uses aversion to mean-preserving spreads to characterize concave utility and loss aversion. Our Theorem 13 shows that their mixture aversion would have provided an appealing alternative characterization.

3.6 Conclusion

We have provided completely general axiomatizations of strictly increasing concave/convex utility and probability distortion functions, using only basic preference conditions. Unlike all preceding results in the literature, we do not need to presuppose any continuity (or differentiability), and the preference conditions used (concavity and convexity) are all basic and appealing. All the richness we need in our analysis is that all simple lotteries are available in the preference domain. We have thus provided the most appealing and most general characterizations of concavity and convexity of utility and probability distortion functions presently available.

Appendix. Proofs and further literature for risk

PROOF OF LEMMA 11. We use Wakker & Yang (2019, Corollary 6). For any fixed probability vector $(p_1,\ldots,p_n)$ and corresponding comoncone we define, in their notation, a state space $S=\{s_1,\ldots,s_n\}$, a probability measure P on S through $P(s_j)=p_j$, and the nonadditive event weighting function $W=w\circ P$. Preferences are then represented by $\sum_{j=1}^{n}(W(s_1\cup\cdots\cup s_j)-W(s_1\cup\cdots\cup s_{j-1}))U(x_j)$. The lemma now follows from their Corollary 6. □

PROOF OF THEOREM 12. We define states and W as in the proof of Lemma 11. The result now follows from Lemma 11. □

PROOF OF THEOREM 13. First assume the properties of U and w. Convexity of w implies convexity of every W constructed in the proof of Lemma 11. Hence, outcome-convexity of $\succcurlyeq$ follows from Wakker & Yang (2019 Corollary 7) applied to such W,U.

Next assume outcome-convexity of $\succcurlyeq$. The definitions in the proof of Lemma 11 and Wakker & Yang (2019 Corollary 7) imply concavity of U and convexity of every W so constructed. Take any probabilities p_1,p_2,p_3 adding to 1 or less. Convexity of the corresponding W implies $W\{s_2,s_1\}-W\{s_1\}\leq W\{s_3,s_2,s_1\}-W\{s_3,s_1\}$, that is, $w(p_2+p_1)-w(p_1)\leq w(p_3+p_2+p_1)-w(p_3+p_1)$. This condition ($w$ differences are increasing), is equivalent to convexity of w. □

PROOF OF THEOREM 15. We use the notation as in Figure 1, using collapsed notation of outcomes (if two outcomes are the same we combine them), but we now allow $x_n\neq 0$, taking x_n general. The *rank* of an outcome is the probability of receiving a better-ranked outcome. Thus, each q_j is the rank of x_{j+1}. We keep the outcomes $(x_1,\ldots,x_n)$ fixed, and consider all probability tuples $(p_1,\ldots,p_n)$. Writing $I:=[0,1]$, we have transformed the set of lotteries into the set $I^{n-1\downarrow}:=\{(q_1,\ldots,q_{n-1})\in I^{n-1}:q_1\geq\cdots\geq q_{n-1}\}$, with every lottery P uniquely corresponding with a $P'\in I^{n-1\downarrow}$, where we denote the induced preference relation as $\succcurlyeq'$. Our correspondence is compatible with convex combination. That is, re-interpreting the q_js as money, i.e., outcomes, a probabilistic mix $\lambda P\oplus(1-\lambda)Q$ corresponds with an outcome-mix $\lambda P'+(1-\lambda)Q'$. Convexity of $\succcurlyeq$ on lotteries is equivalent to outcome-convexity of the corresponding $\succcurlyeq'$ on $I^{n-1\downarrow}$ in the sense of Wakker & Yang (2019, Corollary 6).

We normalize U such that $U(x_n)=0$ and $U(x_1)=1$. Define $\pi'_j:=U(x_j)-U(x_{j+1}), j=1,\ldots,n-1$. Then $RDU(q_1,\ldots,q_{n-1})=\sum_{j=1}^{n-1}\pi'_j w(q_j)$. In other words, we have a representation as in Wakker & Yang (2019, Corollary 6). All conditions of that Corollary are satisfied. Convexity of $\succcurlyeq$ is equivalent to outcome-convexity of $\succcurlyeq'$ and this is, by the aforementioned Corollary, equivalent to concavity of w.

The result for convex w and concave $\succcurlyeq$ follows from the result just obtained by defining $U^* = -U$, $\succeq^* = \preceq$, and $w^* = 1 - w(1-p)$. If outcomes are monetary and monotonicity w.r.t. money is considered desirable, then outcomes can be multiplied by -1. $\square$

FURTHER LITERATURE ON AVERSION TO MEAN-PRESERVING SPEADS. The equivalence of outcome-convexity with aversion to mean-preserving spreads (and its variations discussed below) holds only under RDU. In general, there is no logical relation between these conditions. Before discussing further details, we note that outcome-convexity has been studied in the literature only when an underlying state space was specified, but this is equivalent to our definition for simple lotteries. Under compact continuity, outcome-convexity does imply aversion to mean-preserving spreads (Chateauneuf & Lakhnati 2007 Theorem 4.2; Dekel 1989 Proposition 2). The resulting preferences have been studied for optimal insurances (Ghossoub 2019). If convexity (w.r.t. probabilistic mixing) holds, a condition implied by aversion to mean-preserving spreads under continuity, then by Dekel (1989 Propositions 2 and 3), under weak continuity, aversion to mean-preserving spreads becomes equivalent to convexity.

Bommier, Chassagnon, & Le Grand (2012 Result 3) considered a more-risk-averse than relation weaker than aversion to mean-preserving spreads, with distribution functions crossing once. They should showed for linear w (EU) that their condition is equivalent to concavity of U. They also showed for linear utility (see their proof on pp.1638-1639) that their condition is equivalent to convexity of w. They provided, more generally, comparative results. Chew and Mao (1995, Theorem 2 and Table 2) used a yet weaker elementary risk aversion condition, implied by our outcome-convexity, and showed, under RDU, that it holds if and only if w is convex and U is concave. They assumed differentiability, which under RDU is equivalent to differentiability of w, and continuity. Hence, under the latter two assumptions, they provided an alternative way to obtain our Theorem 13. Ebert (2004, Theorem 2) used a progressive transfer property, equivalent to Chew and Mao (1995) elementary risk aversion, to characterize concavity of U plus convexity of w. Importantly, he did not need differentiability of U, although he did assume continuity. He considered welfare theory where states are reinterpreted as people and probabilities p_j reflect proportions of a population. He used extra structural richness in allowing for any arbitrary replication of any group ("event") in the population. $\square$

References

Abdellaoui, Mohammed (2002) "A Genuine Rank-Dependent Generalization of the von Neumann-Morgenstern Expected Utility Theorem," *Econometrica* 70, 717–736.

Abdellaoui, Mohammed & Peter P. Wakker (2005) "The Likelihood Method for Decision under Uncertainty," *Theory and Decision* 58, 3–76.

Agranov, Marina & Pietro Ortoleva (2017) "Stochastic Choice and Preferences for Randomization," *Journal of Political Economy* 125, 40–68.

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, & David Heath (1999) "Coherent Measures of Risk," *Mathematical Finance* 9, 203–228.

Bell, David E. (1985) "Disappointment in Decision Making under Uncertainty," *Operations Research* 33, 1–27.

Belles-Sampera, Jaume, Montserrat Guillen, & Miguel Santolino (2016) "What Attitudes to Risk Underlie Distortion Risk Measure Choices?," *Insurance: Mathematics and Economics* 68, 101–109.

Birnbaum, Michael H. (2008) "New Paradoxes of Risky Decision Making," *Psychological Review* 115, 463–501.

Bommier, Antoine, Arnold Chassagnon, & François Le Grand (2012) "Comparative Risk Aversion: A Formal Approach with Applications to Saving Behavior," *Journal of Economic Theory* 147, 1614–1641.

Cerreia-Vioglio, Simone, David Dillenberger, Pietro Ortoleva, & Gil Riella (2019), "Deliberately Stochastic," *American Economic Review* 109, 2425–2445

Chateauneuf, Alain & Ghizlane Lakhnati (2007) "From Sure to Strong Diversification," *Economic Theory* 32, 511–522.

Chew, Soo Hong, Edi Karni, & Zvi Safra (1987) "Risk Aversion in the Theory of Expected Utility with Rank Dependent Probabilities," *Journal of Economic Theory* 42, 370–381.

Chew, Soo Hong & Mei-Hui Mao (1995) "A Schur-Concave Characterization of Risk Aversion for Non-Expected Utility Preferences," *Journal of Economic Theory* 67, 402–435.

Cornilly, Dries, Ludger Rüschendorf, Steven Vanduffel (2018) "Upper Bounds for Strictly Concave Distortion Risk Measures on Moment Spaces," *Insurance: Mathematics and Economics* 82, 141–151.

Dekel, Eddie (1989) "Asset Demands without the Independence Axiom," *Econometrica* 57, 163–169.

Ebert, Udo (2004) "Social Welfare, Inequality, and Poverty when Needs Differ," *Social Choice and Welfare* 23, 415–448.

Fudenberg, Drew, Ryota Lijima, & Tomasz Strzalecki (2015) "Stochastic Choice and Revealed Perturbed Utility," *Econometrica* 83, 2371–2409.

Ghirardato, Paolo & Massimo Marinacci (2001a) "Range Convexity and Ambiguity Averse Preferences," *Economic Theory* 17, 599–617.

Ghirardato, Paolo & Massimo Marinacci (2001b) "Risk, Ambiguity, and the Separation of Utility and Beliefs," *Mathematics of Operations Research* 26, 864–890.

Ghossoub, Mario (2019) "Optimal Insurance under Rank-Dependent Expected Utility," *Insurance: Mathematics and Economics* 87, 51–66.

Goovaerts, Marc J., Rob Kaas, & Roger J.A. Laeven (2010a) "Decision Principles Derived from Risk Measures," *Insurance: Mathematics and Economics* 47, 294–302.

Goovaerts, Marc J., Rob Kaas, & Roger J.A. Laeven (2010b) "A Note on Additive Risk Measures in Rank-Dependent Utility," *Insurance: Mathematics and Economics* 47, 187–189.

Gul, Faruk (1992) "Savage's Theorem with a Finite Number of States," *Journal of Economic Theory* 57, 99–110. ("Erratum," *Journal of Economic Theory* 61, 1993, 184.)

Gul, Faruk & Wolfgang Pesendorfer (2015) "Hurwicz Expected Utility and Subjective Sources," *Journal of Economic Theory* 159, 465–488.

Halpern, Joseph Y. (1999) "A Counterexample to Theorems of Cox and Fine," *Journal of Artificial Intelligence Research* 10, 67–85.

Jevons, W. Stanley (1871) "*The Theory of Political Economy*." (5th edn. 1957, Kelley and MacMillan, New York; other edn. Penguin, 1970.)

Khan, M. Ali & Metin Uyan (2018) "Topological Connectedness and Behavioral Assumptions on Preferences: A Two-Way Relationship," working paper.

Kahneman, Daniel & Amos Tversky (1979) "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47, 263–291.

Köbberling, Veronika & Peter P. Wakker (2003) "Preference Foundations for Nonexpected Utility: A Generalized and Simplified Technique," *Mathematics of Operations Research* 28, 395–423.

Köszegi, Botond & Matthew Rabin (2006) "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics* 121, 1133–1165.

Krantz, David H., R. Duncan Luce, Patrick Suppes, & Amos Tversky (1971) "*Foundations of Measurement, Vol. I (Additive and Polynomial Representations)*." Academic Press, New York.

Loomes, Graham & Robert Sugden (1986) "Disappointment and Dynamic Consistency in Choice under Uncertainty," *Review of Economic Studies* 53, 271–282.

Luce, R. Duncan (2000) "*Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches*." Lawrence Erlbaum Publishers, London.

Machina, Mark J. (1985) "Stochastic Choice Functions Generated from Deterministic Preferences over Lotteries," *Economic Journal* 95, 575–594.

Mao, Tiantian & Taizhong Hu (2012) “Characterization of Left-Monotone Risk Aversion in the RDEU Model,” *Insurance: Mathematics and Economics* 50, 413–422.

Masatlioglu, Yusufcan & Collin Raymond (2016) “A Behavioral Analysis of Stochastic Reference Dependence,” *American Economic Review* 106, 2760–2782.

Menger, Karl (1871) “*Principles of Economics*.” Translated into Englisch by James Dingwall & Bert F. Hoselitz, Free Press of Glencoe, New York, 1950.

Miyamoto, John M. (1988) “Generic Utility Theory: Measurement Foundations and Applications in Multiattribute Utility Theory,” *Journal of Mathematical Psychology* 32, 357–404.

Pfanzagl, Johann (1968) “*Theory of Measurement*.” Physica-Verlag, Vienna.

Quiggin, John (1982) “A Theory of Anticipated Utility,” *Journal of Economic Behaviour and Organization* 3, 323–343.

Quiggin, John (1993) “*Generalized Expected Utility Theory — The Rank-Dependent Model*.” Kluwer Academic Press, Dordrecht.

Röell, Ailsa (1987) “Risk Aversion in Quiggin and Yaari’s Rank-Order Model of Choice under Uncertainty,” (Supplement to the) *Economic Journal* 97, 143–160.

Saito, Kota (2015) “ Preference for Flexibility and Randomization under Uncertainty," *American Economic Review* 105, 1246–1271.

Schmidt, Ulrich & Horst Zank (2008) “Risk Aversion in Cumulative Prospect Theory,” *Management Science* 54, 208–216.

Sopher, Barry & J. Mattison Narramore (2000), “Stochastic Choice and Consistency in Decision Making under Risk: An Experimental study,” *Theory and Decision* 48, 323–349.

Tsanakas, Andreas (2008) “Risk Measurement in the Presence of Background Risk,” *Insurance: Mathematics and Economics* 42, 520–528.

Tversky, Amos & Daniel Kahneman (1992) “Advances in Prospect Theory: Cumulative Representation of Uncertainty,” *Journal of Risk and Uncertainty* 5, 297–323.

Viscusi, W. Kip (1989) “Prospective Reference Theory: Toward an Explanantion of the Paradoxes,” *Journal of Risk and Uncertainty* 2, 235–264.

Wakker, Peter P. (1988) “The Algebraic Versus the Topological Approach to Additive Representations,” *Journal of Mathematical Psychology* 32, 421–435.

Wakker, Peter P. (1993) “Unbounded Utility for Savage’s “Foundations of Statistics,” and Other Models,” *Mathematics of Operations Research* 18, 446–485.

Wakker, Peter P. (1994) “Separating Marginal Utility and Probabilistic Risk Aversion,” *Theory and Decision* 36, 1–44.

Wakker, Peter P. (2010) “*Prospect Theory for Risk and Ambiguity.*” Cambridge University Press, Cambridge, UK.

Wakker, Peter P. & Jingni Yang (2019) “A Powerful Tool for Analyzing Concave/Convex Utility and Weighting Functions," *Journal of Economic Theory* 181, 143–159.

Walras, M.E. Léon (1874) "*Elements of Pure Economics*." Translated by William Jaffé, Irwin, Homewood IL, 1954.

Werner, Katarzyna Maria & Horst Zank (2019) " A Revealed Reference Point for Prospect Theory," *Economic Theory* 67, 731–773.

Yaari, Menahem E. (1987) "The Dual Theory of Choice under Risk," *Econometrica* 55, 95–115.

Chapter 4

Giving According to Agreement[1]

Abstract

We propose an axiom we call the 'Agreement axiom' to deal with changing preferences and derive its empirical implications. The resulting revealed preference condition generalizes GARP when preferences are unstable but preferences in one context are still informative about preferences in another context. We apply this idea to a social choice experiment, where a player can respond to another player being kind or relatively unkind (these being the two different contexts). We find that people have a consistent set of preferences for each context, but that preferences are different between the contexts, and that subjects largely act according to the Agreement axiom. We thus provide support for modeling and interpreting responses to the intentions of other players as reciprocity (as opposed to, say, a strategic mistake). Although we apply our method to reciprocity in a social choice setting, it can be used to model changing preferences in many different settings.

[1] This chapter is joint with Jan Heufer and Paul van Bruggen.

4.1 Introduction

Decision makers often act unselfishly: they give to charity, volunteer, and give in dictator games. In experiments, subjects act unselfishly in a manner consistent with utility maximization (Andreoni and Miller 2002) and may thus be considered to have social preferences. Much evidence has been found that these preferences depend on context, not only on final outcomes. Giving is affected by others' intentions. In a typical experiment (e.g. Blount 1995; Falk et al. 2008)), giving to another player is higher in a state where a low endowment is determined by a randomization device compared to when it was chosen by the other player. In other words, they reciprocate.

If people indeed have reciprocal preferences (as opposed to, say, a reciprocal reflex which may be a strategic mistake) this means that while preferences are unstable between different contexts, preferences should still be consistent for a given context. The first question is: do people indeed have well-behaved sets of preferences for a given context? The second question is, if consistent sets of preferences exist for different contexts, can we find some way to integrate these sets of preferences? That is, do preferences in one context allow us to predict preferences in the other contexts?

Existing modeling efforts (e.g. Rabin 1993; Dufwenberg and Kirchsteiger 2004; Falk and Fischbacher 2006; Segal and Sobel 2007) have focused mostly on the game-theoretical implications of reciprocity, rather than on measuring and comparing preferences. As a consequence, they make specific functional form assumptions about important unobservables such as the utility function and notions of fairness. Tests of these models are simultaneously tests of their parametric assumptions. Because these assumptions are made for tractability, not for their realism, they are likely to lead to violations even if the underlying modeling principle is sound.

What is needed when investigating reciprocal preferences is a method that makes a minimum of assumptions about functional forms, yet can establish whether people have consistent social preferences and what these preferences look like. In this paper, we propose an axiom to integrate different sets of preferences and derive its non-parametric implications, which allows predicting choices in one context based on choices in the other context.

The concept underlying our non-parametric method is that the two sets of preferences *agree* on particular allocations. Specifically, if some allocation x is preferred to some allocation y if the other player is being unkind, even though allocation y gives the decision maker more than allocation x, then x should also be preferred to y if the other player is being kind. Choosing x over y if the other is unkind must stem from some consideration of fairness, as the purely selfish motive would suggest picking y. This *Agreement axiom* captures that while different intentions of others may make us more or less selfish, it should not affect how we trade off fairness and efficiency.

We test our proposed method in an experiment. We find that our subjects indeed have different but consistent sets of preferences between the different contexts, and that they largely satisfy the Agreement axiom.

4.2 Agreement

4.2.1 Preferences

Suppose a decision maker has two sets of preferences for allocations of money for herself and someone else (for simplicity, results are derived for this two-dimensional case, but results can be extended to any arbitrary dimension). We assume that the other person is always the same person. Let $\mathbf{X} = \mathbb{R}^2_+$ be the preference domain and for all $a = (a_1, a_2) \in \mathbf{X}$, let a_1 be the money amount she gives to herself and a_2 be the money amount for the other person. Let $\mathbf{P} = \mathbb{R}^2_{++}$ be the set of all prices[2]. The decision maker is characterized by a pair of preferences ($\succsim_A$, $\succsim_S$) on $\mathbf{X}$. Her preference relation $\succsim_A$ reflects an altruistic self and $\succsim_S$ reflects a selfish self. Both preference relations are transitive, complete and continuous. 'Altruistic' and 'selfish' are taken as relative terms here; the selfish preference relation need not be perfectly selfish.[3]

The decision maker faces two contexts to make choices, which we call *context A* and *context S*, where the decision maker is expected to be more altruistic in context A than in context S, for example for reasons of reciprocity. The following axiom gives the relation between these two preference relations.

Axiom 1 (Agreement)**.** *For all $x, y \in \mathbf{X}$ with $x = (x_1, x_2)$ and $y = (y_1, y_2)$,*

$$x \succsim_S y \text{ and } x_1 < y_1 \text{ imply } \quad x \succsim_A y.$$

Intuitively, Agreement states that if the decision maker prefers x to y when choosing according to $\succsim_A$, and she keeps more for herself in choice x than y, then she also prefers x to y according to the more selfish preferences ($\succsim_S$). An equivalent formulation is presented in Proposition 4.1.

Proposition 4.1. *The Agreement axiom is equivalent to the requirement that for all $x, y \in \mathbf{X}$ with $x = (x_1, x_2)$ and $y = (y_1, y_2)$, $x \succsim_A y$ and $x_1 > y_1$ imply $\quad x \succsim_S y$.*

All proofs are in the Appendix.

[2] We follow the convention that $\mathbb{R}^2_+ = \{x \in \mathbb{R}^2 : x_1 \geq 0 \text{ and } x_2 \geq 0\}$ and $\mathbb{R}^2_{++} = \{x \in \mathbb{R}^2 : x_1 > 0 \text{ and } x_2 > 0\}$.

[3] Of course, a decision maker may have more than two preference relations. As the interest of this paper is in comparing preferences, we limit the analysis to two preference relations, but multiple such comparisons can be made between any number of preference relations.

4.2.2 Empirical Implications of Agreement for Revealed Preferences

We start this section with a review of basic concepts of revealed preference.

Definition 4.1. *A set of observations* Ω *is a finite collection of pairs* $(x,p) \in \mathbf{X} \times \mathbf{P}$.

An observation $(x,p) = ((x_1,x_2),(p_1,p_2)) \in \mathbf{X} \times \mathbf{P}$ denotes how much the decision maker chooses to keep, x_1, and to give to the other, x_2, given the choice set $\{(x_1,x_2) \in \mathbb{R}^2_+ : p_1x_1 + p_2x_2 \leq 1\}$. We use superscripts to indicate observations and subscripts to indicate coordinates.

We now define the revealed preference relation R based on Ω. Given $\Omega = \{(x^i, p^i)\}_{i=1}^N$, we say that x^i is *directly revealed preferred* to a, written as x^iR^0a, if $p^ix^i \geq p^ia$; it is *indirectly revealed preferred* to a, written as x^iRa, if there exist $x^{i_1}, x^{i_2}, \cdots, x^{i_m} \in \Omega$ such that $x^iR^0x^{i_1}R^0x^{i_2}R^0 \cdots R^0x^{i_m}R^0a$. We use P^0 (P) to denote the strict preference relation: x^i is *strictly directly revealed preferred* to a, written x^iP^0a, if $p^ix^i > p^ia$. We sayx^i is *strictly revealed preferred* to a, written x^iPa, if there exist $x^{i_1}, x^{i_2}, \cdots, x^{i_m} \in \Omega$ such that $x^iR^0x^{i_1}R^0x^{i_2}R^0 \cdots R^0x^{i_m}R^0a$ and at least one of these R^0 relationship is even a P^0. A *utility function u* is a map from $\mathbf{X}$ to $\mathbb{R}$.

Definition 4.2. *A utility function u* rationalizes *a set of observation* Ω *if* $u(x^i) \geq u(x^j)$ *whenever* x^iRx^j, *and* $u(x^i) > u(x^j)$ *whenever* x^iPx^j.

Afriat (1967) and Varian (1982) have provided an easily testable necessary and sufficient condition for the existence of a utility function that rationalizes a set of observations Ω.

Axiom 2 (GARP)**.** *The set of observations* Ω *satisfies the* General Axiom of Revealed Preference *(*GARP*) if for always* x^iRx^j *implies not* $x^jP^0x^i$.

Theorem 1 (Reny 2015)**.** *The set of observations* Ω *satisfies GARP if and only if there is a semi-strictly increasing and quasiconcave utility function* $u : \mathbf{X} \to \mathbb{R}$ *that rationalizes* Ω.

In our model, we have two sets of observations, one from context A and one from context S, denoted by $\Omega_A = \{x^i, p^i\}_{i=1}^n$ and $\Omega_S = \{y^j, q^j\}_{j=1}^m$. Let R_A and R_S be the revealed preference relations from Ω_A and Ω_S, respectively. Moreover, if Agreement holds then the observations Ω_A (Ω_S) reveal information about $\succsim_S$ ($\succsim_A$) in a way illustrated in Figure 4.1.

In the left graph, given the budget line in context A, the decision maker chooses x^i, that is, x^i is revealed preferred to all the allocations on and below the budget line in context A. Every allocation from the area indicated by the red vertical lines gives the decision maker less than x^i does. Thus, by Agreement, x^i is also revealed preferred to the allocations in the area indicated by the red vertical lines for context S. In the right graph, the decision maker chooses y^j in context S, and any allocation in the area indicated with the blue vertical lines gives the decision maker more than y^j does. Thus, she cannot prefer an allocation from the area indicated by the blue vertical lines over y^j in context A.

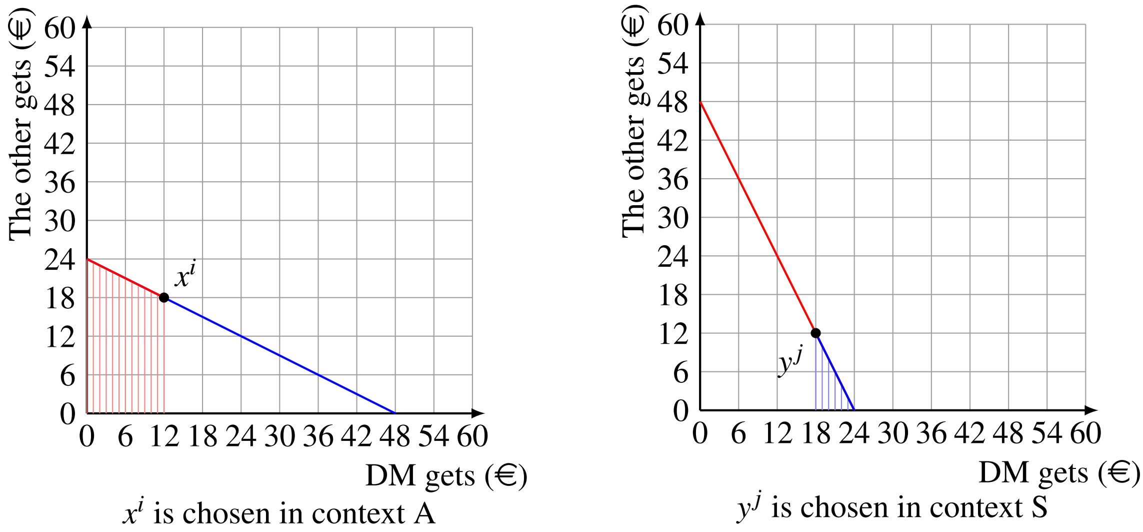


Figure 4.1: Revealed preference implications of Agreement axiom. The observations are revealed preferred to the area below the budget in the same context and to the area indicated with the vertical lines in the other context.

Through Agreement we can thus extend the revealed relation R_A by incorporating the information from Ω_S and extend R_S by incorporating the information from Ω_A. Writing these extensions as $\tilde{R}_A$ and $\tilde{R}_S$, it follows that $x^i\tilde{R}^0_A x^j$, if $x^i R^0_A x^j$ or if $x^i R^0_S x^j$ and $x^i_1 > x^j_1$; and that $x^i\tilde{R}^0_S x^j$, if $x^i R^0_S x^j$ or if $x^i R^0_A x^j$ and $x^i_1 < x^j_1$. $\tilde{P}_A$ and $\tilde{P}_S$ are defined analogously. Let $\Omega = \Omega_A \cup \Omega_S$. We next define rationalization in our model.

Definition 4.3. *(i) An altruistic utility function $u(x)$* Agreement-rationalizes *(*AG-rationalizes*) Ω, if $u(x^i) \geq u(x^j)$ whenever $x^i\tilde{R}_A x^j$.*

(ii) A selfish utility function $v(x)$ AG-rationalizes Ω, if $v(x^i) \geq v(x^j)$ whenever $x^i\tilde{R}_S x^j$.

Both altruistic and selfish utilities represent the extended relations and are derived from the observations Ω. AG-rationalization captures what it means to choose according to one's preferences if these preferences satisfy Agreement.

Axiom 3 (AG-GARP)**.** *The set of observations Ω satisfies* Agreement-GARP *(*AG-GARP*), if $x^i\tilde{R}_A x^j$ implies not $x^j\tilde{P}^0_A x^i$.*

Theorem 2. *Given the observations $\Omega = \Omega_A \cup \Omega_S$ with $\Omega_A = \{x^i, p^i\}^n_{i=1}$ and $\Omega_S = \{y^j, q^j\}^m_{j=1}$, the following conditions are equivalent:*

(a) The set of observations Ω satisfies AG-GARP.

(b) *There exist a non-decreasing and quasiconcave altruistic utility $u(x)$ and selfish utility $v(x)$ that AG-rationalizes Ω. Moreover, for all $x, y \in \mathbf{X}$, $u(x) > u(y)$ with $x_1 > y_1$ implies that $v(x) \geq v(y)$ almost everywhere, and for all $x, y \in \mathbf{X}$, $v(y) > v(x)$ with $y_1 < x_1$ implies that $v(y) \geq v(x)$ almost everywhere.*

Theorem 2 describes that if the set of observations satisfies AG-GARP, the decision maker's choices can be represented by two utility functions, an altruistic utility function and a selfish utility function. Moreover, these two utility functions are connected through the extended relations. Theorem 2 thus captures the empirical implications of Agreement and rationalization.

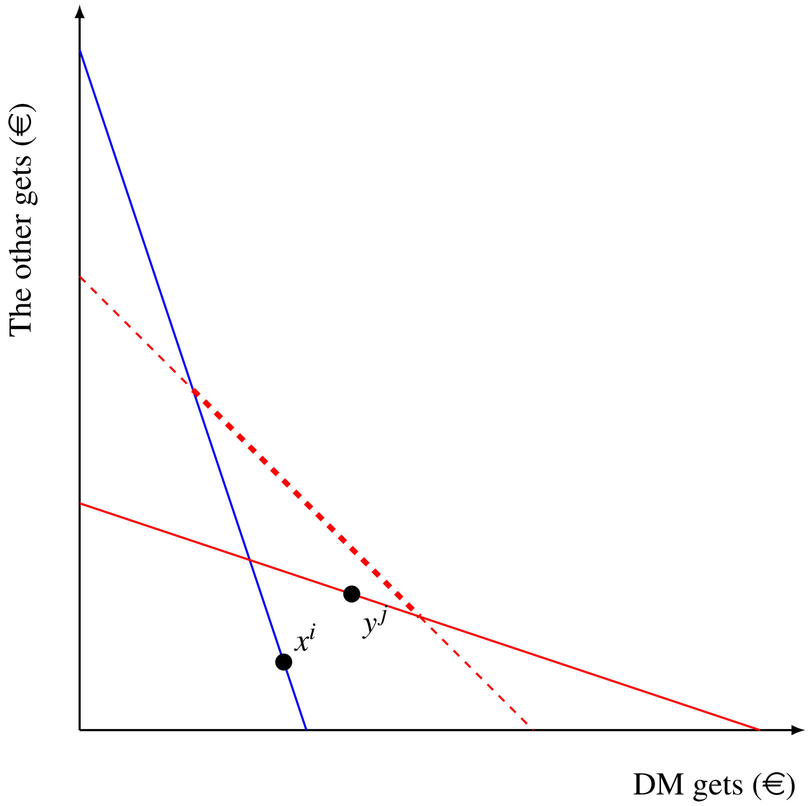


Figure 4.2: AG-GARP predicts that a third choice on the dashed red budget line must be on the thick red part.

We can use AG-GARP to predict choices based on previous observations. This is illustrated in Figure 4.2. x^i is chosen in context A and y^j is chosen in context S. Suppose a person is next presented with the dashed red budget in context S. A person who chooses according to preferences which satisfy AG-GARP must then choose on the thick red part of the budget line. Choosing to keep more violates GARP within context S (and therefore violates AG-GARP); choosing to keep less violates Agreement.[4]

[4] Such predictions can easily be extended to situations where a person does not choose in perfect accordance

Although we have assumed that we have a dataset from a context A and a dataset from a context S, it is not essential that we know for sure that a decision maker is more altruistic in context A than in context S. This is a hypothesis which we can test by testing choices for AG-GARP. If a decision maker is more altruistic in context S this means the Agreement axiom is violated. In some situations we want to model changing preferences without specifying in which context the preferences should be more or less altruistic. In these cases we can let the data speak for itself (if we at least have two clearly distinct sets of data). As we show in the Appendix, satisfying Agreement means violations of GARP can only go in one direction (see Figure 4.5 in the Appendix). We can use this to infer from the data which context is context A and which is context S. If two observations from the two different contexts violate GARP, then the observation which gives the decision maker less than the other observation must come from context A, and hence all observations from that set must belong to Ω_A. AG-GARP is then violated if we find violations of GARP that go in different directions. However, it will be natural in the analysis of our data to specifically assume that one context is context A and that the other is context S.

4.3 Experiment

4.3.1 Experimental design

To measure how social preferences depend on the history of play we ran an interactive experiment with two subjects making choices for each observation. One player made a single Choice (hence we call this player 'player C' or simply 'C') between two allocations: either giving €13 to the other player and keeping €27 or giving €18 and keeping €18. In our experiment, these two possibilities represent the contexts A and S that the second player faced. Using the strategy method, the second player then made 14 choices from 14 Budgets (hence, we call this player 'player B') for each of these two contexts (thus making a total of 28 choices).[5] Crucially, both players were informed that only one choice would be implemented for real: either the choice of the first player (C) would be implemented or one of the choices of player B would be implemented. In the latter case, one budget would be randomly selected from the set of budgets corresponding to the choice made by C. Thus, Player B's choice is only practically relevant as a response to C's intention, but never to a practically implemented choice by C.

For the purposes of this paper, we are only interested in the choices made by player B. Player

with AG-GARP by correcting for efficiency. Efficiency is discussed in Section 4.4.

[5] With 28 budgets one may wonder if we are asking too much of subjects. To test the various aspects of our experiment we performed several pilots. In these pilots it was revealed that our initial choice of budgets was conservative because subjects went through the tasks quickly and accurately. In the interest of power we decided to increase the number of budgets to 28.

C was only part of the experiment to incentivise player B. The budgets player B was faced with were all linear budgets where B could give some money to C at different rates, from giving C €0.33 to giving C €3 for every euro B gave up. The minimum B could keep was always €0, as was the minimum B could give away to C. The maximum amount B could keep or give to C varied by budget and was never higher than €48, respectively €60. Budgets were chosen such that they intersected at many points to get good test power. The average slope was bigger than 1, meaning that on average giving up €1 increased C's payoff by more than €1, to make it attractive for player B to give at least some money to C (if all choices are on the axis, test power is zero).

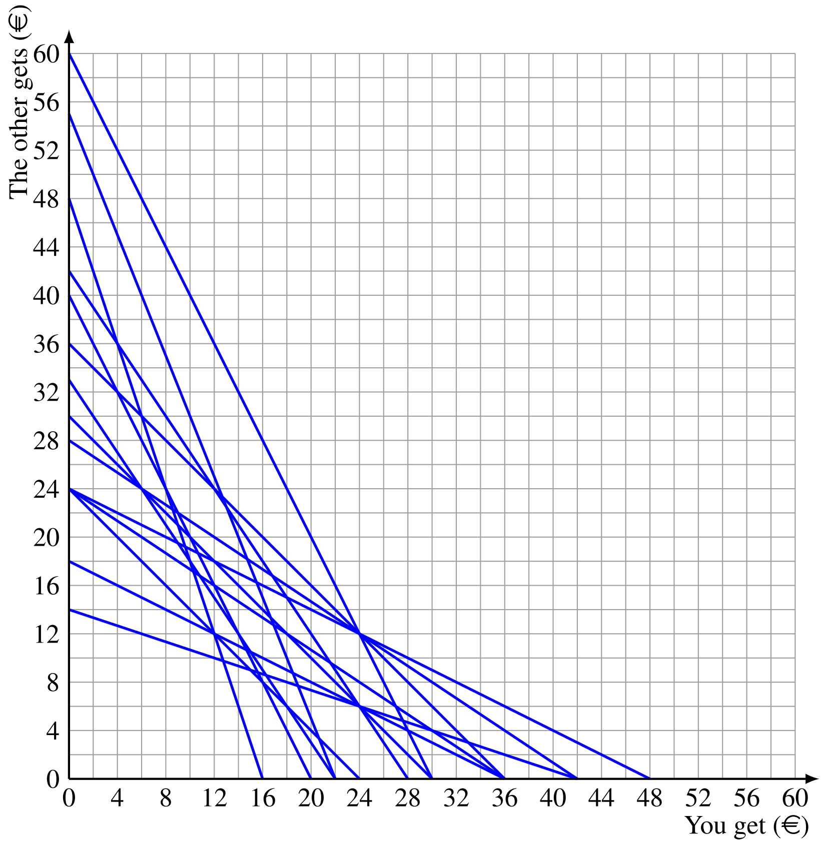


Figure 4.3: Budgets used in the experiment

Player C could indicate their choice by selecting the desired allocation and confirming their choice (they could revise their choice before confirming). The order of the options was randomized at the start of the experiment and then kept the same throughout the experiment in the instructions and in the actual choice. Player B made choices from 14 budgets for the context when C chose the allocation $(13, 27)$ and made choices from the same 14 budgets for the context when C chose $(18, 18)$. Which of these contexts was presented first was randomized at the start of the experiment and then kept the same throughout the experiment in the instructions and in the actual choices. The order of the 14 budgets was randomized for each of the contexts separately, so that the order of the budgets was different between the two contexts.

To make performing the tasks as easy as possible for player B we developed an interface which graphically displayed the available budget. Player B could make choices by clicking on any point on the budget, by typing the amount they wanted to keep, or by typing the amount they wanted to give to player C. When either of these three methods had been used, a dot would appear on the budget line to indicate their choice and the amounts that player C and B would receive were displayed automatically in number fields. Player B could then revise their choice if they were not happy with the resulting allocation by clicking somewhere else on the budget line, by dragging the dot around on the budget line, or by entering in a different number in either of the fields displaying how much they would keep or give away.

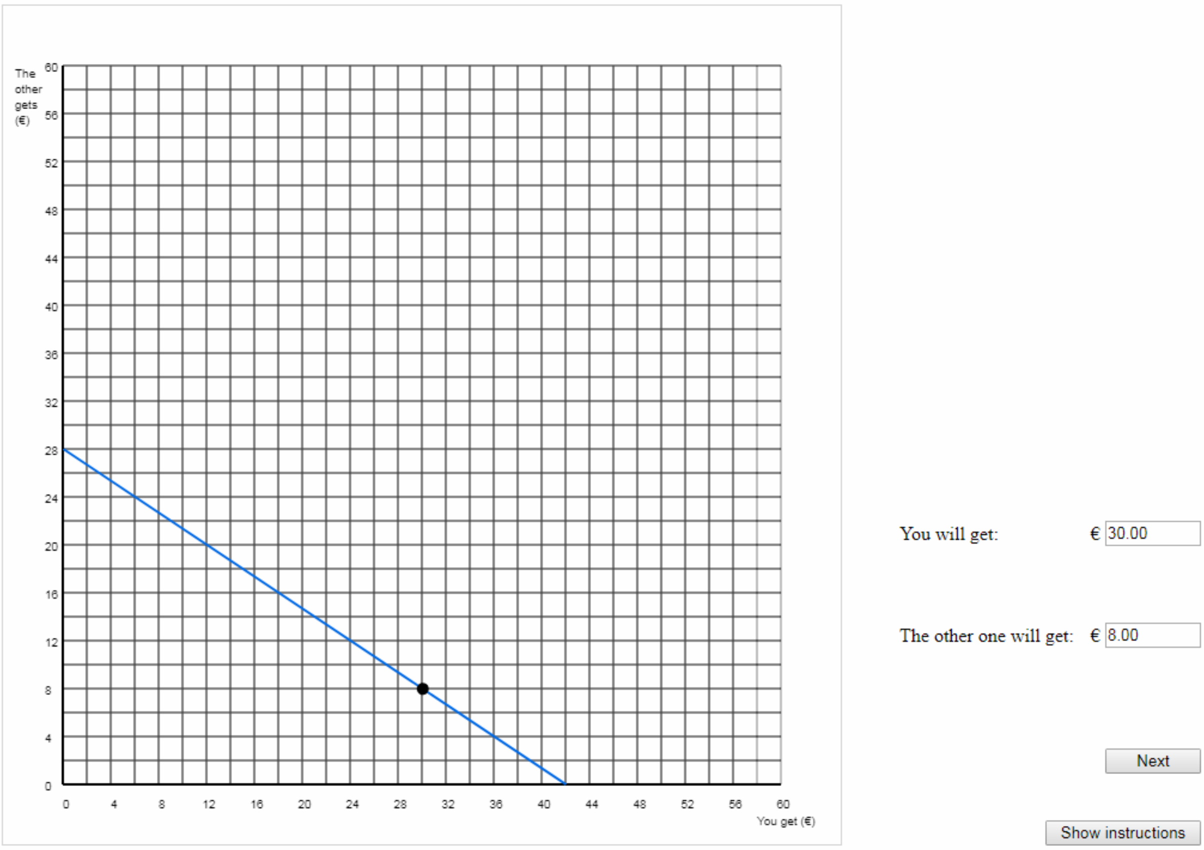


Figure 4.4: A screenshot of the interface for one of the player B's tasks.

The software automatically calculated a minimum step size in multiples of €0.05 and any choice was automatically converted to the closest such step. This ensured that all choices were exactly on the budget line. For example, where player B could give away €3 for every €1 they gave up, the minimum amount B could give up other than 0 was €0.05 and the minimum amount they could give away other than 0 was €0.15. Giving up €0.03 or giving away €0.10 to C was automatically changed to giving up €0.05 and giving away €0.15. Entering an amount not on the budget line (negative amounts or amounts greater than the maximum amount that could be kept or given to C) resulted in a message indicating that the amount entered was invalid. Player B could only continue to the next task after choosing a valid allocation.

Player C was informed that player B could divide money between them at different rates, that the minimum B could give them was €0 and that the maximum differed per budget but was never more than €60. Player C was informed that B made 14 choices for both of C's possible choices. Player B was informed that player C was presented with this information. Both players were informed how they would be matched to each other and were informed about the payment procedure, including that both players would be paid either according to C's choice or to one of B's choices and that this was determined randomly.

After the instructions, player B was given some practice tasks to familiarize them with the interface.[6] To explain the various input methods, in each of the practice tasks only one method of input was enabled (clicking on the budget line; entering the amount to receive; entering the amount for the other player to receive) and the subject was asked to use that method to practice making a choice. Both players were asked to answer three multiple choice comprehension questions. If they answered a question incorrectly they were given immediate feedback as to why their answer was wrong. They could only continue once they had answered the question correctly.

On entering the lab, subjects were assigned a cubicle and asked to wait for the start of the experiment (if there was an odd number of subjects the last subject to arrive was given a show-up fee and did not participate). At the start of the experiment, subjects were handed an envelope containing two slips of paper containing their subject ID. After using their subject ID to log on, we collected their envelopes, leaving one of their IDs at their cubicle, and sorted them into one pile containing player C IDs and one containing player B IDs. Subjects who finished early were then asked to randomly match players by choosing an envelope from each pile (without seeing the ID codes within) and to put the ID code of player B into the envelope containing the ID of player C. Next, subjects were asked to roll a die to determine whether each pair would be paid according to player C or player B's choice (where the probability of C or B being selected was equal). This was marked on the envelope. If the pair was paid according to B's choice, a subject was asked to draw a ball from a bag with 14 balls numbered 1 to 14 inclusive to select according to which budget the pair would be paid out. This too was marked on the envelope. When all subjects had finished subjects were asked one by one to come to the front desk to be paid.

The experiment was run in the ESE-econlab of Erasmus University Rotterdam. There were 9

[6] Player C was not provided with practice tasks because their task was a very simple binary choice which required no more than clicking on the desired option.

sessions of roughly 20 subjects each, with a total of 170 subjects participating. The experiment lasted about an hour and the average payment was €16.47.

4.4 Non-parametric analysis

In our analysis part, we only use the data of player B's choices. We first perform the non-parametric analysis. We take the set of player B's choices responding to player C's possible selfish choice of keeping €27 and giving €13 as context S and the set of player B's choices responding to player C's possible more altruistic choice of keeping €18 and giving €18 as context A. Based on Theorem 1 and Theorem 2, we have the following hypotheses.

Hypothesis 1. *Player B has the same preferences in both contexts:* $\succsim_A = \succsim_S$ *. This implies the following testable conditions:*

$$x^i R_A y^j \Rightarrow \text{ Not } y^j P_S x^i$$

$$y^j R_S x^i \Rightarrow \text{ Not } x^i P_A y^j$$

Hypothesis 2. *Player B has different sets of preferences in each context, which are are connected by Agreement. Player B uses preferences* $\succsim_A$ *in context A and* $\succsim_S$ *in context S. Testable conditions:*

$$x^i R_A y^j \text{ and } x_1^i > y_1^j \Rightarrow \text{ Not } y^j P_S x^i$$

$$y^j R_S x^i \text{ and } x_1^i > y_1^j \Rightarrow \text{ Not } x^i P_A y^j$$

First, we test hypothesis 1 where subjects' preferences are defined over final outcomes only. This is a typical assumption in many models but precludes reciprocity. We compare it to the alternative where subjects have different sets of preferences for each context. Next, we test the hypothesis 2 that people have preferences that satisfy Agreement.

As satisfying GARP is a binary variable, where one either satisfies it or does not and minor and economically unimportant violations of GARP are treated the same as very significant violations, we use the *Afriat efficiency index* (AEI) (Afriat, 1972) instead. This captures how far away someone's choices are from utility maximisation. The index is bounded between 0 and 1. When choices satisfy GARP, the index is 1. When GARP is violated, the index is (weakly) smaller than 1, with bigger violations (relative to the budget) resulting in lower indices. There is no clear benchmark for judging what constitutes 'big' violations of GARP, but an efficiency of 0.95 or higher is typically seen as being close to utility maximisation. We adopt this convention (we also present results for efficiencies of 0.99 and 0.90 for comparison). Where AG-GARP is applicable we calculate the *Agreement Afriat Efficiency Index* (AG-AEI), which has the same interpretation as the AEI but for AG-GARP instead of GARP.

As shown in Table 4.1, for each of player C's two possible choices separately we see that the vast majority of subjects (more than 90%) attains an efficiency of 0.95 or higher. When combining the choices from both contexts, efficiency drops strongly, and only 58% of subjects achieve an efficiency of 0.95 (we will discuss the issue of differences in test power below). Social preferences defined only over final outcomes therefore does not describe choices very well compared to the alternative of separate, consistent sets of preferences.

Separate sets of preferences can never do worse than one set of preferences, and the cost of using separate sets of preferences is generality. The crucial question is therefore whether the separate sets of preferences can be integrated somehow, so that choices in one set can be predicted by the revealed preferences in the other set, such as through our Agreement axiom. Indeed, this is the case: 84% of subjects satisfy Agreement with an (AG-AEI) efficiency of at least 0.95.

Table 4.1: Percentage of subjects with an (AG)-AEI of at least 0.90, 0.95 and 0.99 for various contexts

Index	Context A, AEI	Context S, AEI	Both contexts, AEI	Both contexts, AG-AEI
0.99	77.64%	78.82%	32.94%	62.35%
0.95	92.94%	90.59%	57.64%	83.52%
0.90	98.82%	97.64%	74.12%	94.12%

So far, we have ignored the issue of test power. Combining choices from both contexts means there are more choices and thus that there is a greater likelihood that at least some choices violate (AG)-GARP, and because Agreement is weaker than GARP, efficiency for Agreement will be higher than for GARP. To determine test power for each of the hypotheses, we generated random choices on the budgets used in the experiment.

In Table 4.2 the results of 10000 randomly generated choices are presented in various quantiles. We can see that subjects did considerably better than random choice. Ranking subjects by their AEI, the 5% of subjects with the lowest AEI (the first quantile) achieve an AEI of 0.94 for both contexts, whereas the first quantile of random choices achieve an AEI of only 0.66. Similarly, where fewer than 25% of subjects have an AEI lower than 1.00, this is true for more than 75% of randomly drawn sets of choices.

Table 4.2: Quantiles of AEI for choices from each context and for random choices

	Quantiles				
	0.05	0.10	0.25	0.50	0.75
Context A	0.94	0.95	1.00	1.00	1.00
Context S	0.94	0.96	1.00	1.00	1.00
Random	0.66	0.72	0.80	0.88	0.93

Table 4.3 presents similar calculations for Agreement. Here, both random choices and random choices *that happen to satisfy GARP* for each context were generated. As we can see, subject's choices satisfied Agreement considerably more often than random choices. More than 75% of the subjects had an AG-AEI of at least 0.95, which is true for less than 25% of sets of random choices. Subjects' choices even satisfied Agreement more often than randomly generated choices that satisfy GARP in each of the subsets,[7] with only half such sets having an AG-AEI of at least 0.95. The large number of subjects satisfying Agreement (with an efficiency of at least 0.95) is therefore not due to poor test power.

Table 4.3: Comparison of AG-AEI

	Quantiles				
	0.05	0.10	0.25	0.50	0.75
Subjects	0.87	0.93	0.97	1.00	1.00
Random	0.54	0.58	0.64	0.71	0.78
Random GARP consistent	0.72	0.78	0.88	0.95	1.00

4.5 Parametric analysis

The Agreement axiom can also be used with parametric assumptions. In this section, we therefore perform parametric analysis to complement the non-parametric analysis of Section 4.4. We do so for constant elasticity of substitution (CES) utility functions. The Agreement axiom has an intuitive interpretation for such utility functions, which have the following form:

$$u(a) = (\alpha a_1^{\rho} + (1-\alpha)a_2^{\rho})^{1/\rho} \tag{4.1}$$

Here $a \in \mathbf{X}$ is the allocation. Parameter ρ determines the elasticity of substitution, that is, the curvature of the indifference curves. Thus ρ determines the trade-off between equality and efficiency. Parameter α is the distribution parameter, and captures how the payoff to the decision maker is traded off against the payoff of the other person. If decision makers' utility can be described by (4.1), that is, $u(a) = (\alpha_A a_1^{\rho_A} + (1-\alpha_A)a_2^{\rho_A})^{1/\rho_A}$ for $\succsim_A$ and $v(a) = (\alpha_S a_1^{\rho_S} + (1-\alpha_S)a_2^{\rho_S})^{1/\rho_S}$ for $\succsim_S$, the hypotheses that are in parallel to those of non-parametric analysis are the following (the proof is in the Appendix).

Hypothesis 3. *Player B has the same set of preferences in both contexts: $\succsim_A = \succsim_S$. Parametric implications:*

$$\alpha_A = \alpha_S \text{ and } \rho_A = \rho_S.$$

[7] Note that most subjects violated GARP within a subset at least some of the time, and that such violations are also violations of AG-GARP.

Hypothesis 4. *Player B has different sets of preferences in each context and the different preferences are connected by Agreement. Parametric implications:*

$$\alpha_A \leq \alpha_S \ \ and \ \ \rho_A = \rho_S.$$

Following Andreoni and Muller (2002), we estimate CES functions for each individual for both treatments. With the budget constraint $a_1 + pa_2 = m$ the demand function is

$$a_1(p,m) = \frac{\alpha/(1-\alpha)^{1/(1-\rho)}}{p^{-\rho/(\rho-1)+[\alpha/(1-\alpha)]^{1/(1-\rho)}}} m = \frac{D}{p^r + D} m \tag{4.2}$$

Where $r = -\rho/(1-\rho)$ and $D = [\alpha/(1-\alpha)]^{1/(1-\rho)}$. We first estimate r and D by non-linear least squares, then back out estimates of α and ρ.

For only 59 out of the 85 player B we can fit the CES utility function. Twelve subjects make all choices on the axis (in all cases keeping everything) in either treatment and are therefore excluded, leaving 73 subjects. Additionally, the estimation for 14 player Bs does not converge because they chose on the axis except once or twice in each treatment, or they always chose proportionally in one treatment (the ratio of the money amount player B keeps to the amount he/she gives is constant).

The median $\hat{\alpha}$ and $\hat{\rho}$ for both contexts are reported in Table 4.4: $\hat{\alpha}$ is bigger when player C chooses the less generous allocation than when C chooses the more generous allocation. A Wilcoxon signed-rank test indicates that this difference is significant at the 5% level. By contrast, the difference in $\hat{\rho}$ between the two contexts is not significant.This is in accordance with hypothesis 4.

Table 4.4: Summary of Wilcoxon Signrank Test

	Context A	Context S	P-value
Median $\hat{\alpha}$	0.85	0.98	0.03
Median $\hat{\rho}$	-0.91	-1.00	0.19

4.6 Conclusion

We introduce a revealed preference condition based on a new axiom, which we call Agreement, which generalizes GARP when preferences are context-specific, but where preferences in one context are informative about preferences in the another context. Applying this method to a social choice situation, where the context was generated by the kindness of another player, we find that people have consistent preferences for a given level of kindness of the other player, but that their preferences are different for different levels of kindness. Furthermore, we find that choices are largely consistent with our Agreement axiom, which means choices for one level of kindness of the other player can be used to predict choices given a different level of kindness of the other player. Our findings provide support for the interpretation of reciprocal behavior as being generated by reciprocal preferences. The Agreement axiom and the revealed preference implications we derive can be used to model changing preferences generally, not only in a social choice situation.

4.7 Appendix: Proofs

Proof of Proposition 4.1. Suppose Agreement holds. Assume for contradiction that there are $x, y \in \mathbf{X}$, such that $x \succsim_S y$ with $x_1 < y_1$ and $y \succ_A x$. By continuity of preferences, there exists $\varepsilon > 0$, such that $y \succ_s y* = (y_1 - \varepsilon, y_2)$, $x_1 < y_1 - \varepsilon$, $y^* \succ_A x$. We, therefore, have $x \succ_S y^*$, $x_1 < y_1^*$ and $y^* \succ_A x$. However, by Agreement, $x_1 < y_1^*$ and $y^* \succ_A x$ implies that $y^* \succsim_S x$, so we have a contradiction. The other direction of proposition 4.1 can be shown in a similar way. $\square$

Proof of Theorem 1. To clarify the proof, we first restate Reny's construction here, which is different from Varian (1982).

A data set Ω satisfy GARP if for any finite sequence $(p_1,x_1),(p_2,x_2),\cdots,(p_n,x_n)$ in Ω such that

$$p_1(x_2 - x_1) \leq 0,\ p_2(x_3 - x_2) \leq 0, \cdots, p_{n-1}(x_n - x_{n-1}) \leq 0 \ \Rightarrow \ p_n(x_1 - x_n) \geq 0$$

Choose $\phi : \mathbb{R}^2_+ \to \mathbb{R}$, that is continuous, strictly increasing and quasiconcave on $[0,-1]$. Given $x \in \mathbb{R}^2_+$, we say $(p_1,x_1),(p_2,x_2),\cdots,(p_n,x_n)$ in Ω is *x-feasible* if $p_1(x - x_1), p_2(x_1 - x_2), \cdots p_n(x_{n-1} - x_n)$ are non positive. Now we define the utility follows: Given $x \in \mathbb{R}^2_+$, if there is x' that is a chosen bundle inΩ, then let

$$u(x) := \inf(p_1(x - x_1) + p_2(x_1 - x_2), + \cdots + p_n(x_{n-1} - x_n)),$$

where the infimun is take over all x-feasible sequence in Ω. Otherwise $u(x) = \phi(x)$. Reny (2015) has shown that under this construction, u is semi-strictly increasing, quasiconcave that rationalizes Ω. $\square$

In our model, allocations only contain two dimensions. Then GARP can be simplified by the following Theorem that can help to understand the proofs after.

Theorem 3 (Theorem 1 of Banerjee (2006))**.** *Given $\mathbf{X} \subset \mathbb{R}^2_+$, GARP is equivalent to that $x^i R^0 x^j$ not $x^j P^0 x^i$.*

Let $N_\varepsilon(x) = \{a \in \mathbf{X} : d(a,x) < \varepsilon\}$ be the open neighborhood of x, and let d be the Euclidean distance. The interior of a set S, int S, is the set of all $a \in \mathbf{X}$, such that there exists ε_a such that $N_{\varepsilon_a}(a) \subset S$. The boundary of a set S is d$S = \{a \in S : a \notin \text{int}S\}$. For a finite set $\{s^1,\cdots,s^n\} = S \in \mathbf{X}$, the *convex hull* of S is

$$CH(S) = \{a \in \mathbf{X} : a = \sum_{i=1}^{n} \lambda_i S^i, \sum_{i=1}^{n} \lambda_i = 1 \ \ \lambda_i \in [0,1] \ \text{ for all } i\}.$$

The *convex monotonic hull* of S is,

$$CMH(S) = CH(\{a \in \mathbf{X} : a \geq s^i \ \text{ for some } i = 1, \cdots, n\}).$$

Let $Z = \{z^i\}_{i=1}^k = \{x_i\}_{i=1}^n \cup \{y_j\}_{i=1}^m$, and $P = \{r^i\}_{i=1}^k$ be the set of corresponding prices to Z and $\mathbf{B}(r^i) = \{a \in \mathbf{X} : a \cdot r^i < 1\}$. We define $I(\tilde{R}_S^0, z^m)$ to be the revealed worse set of observation (z^m, r^m). Specifically, if $(z^m, r^m) \in \Omega_S$, $I(\tilde{R}_S^0, z^m) = \{a \in \mathbf{X} : a \cdot r^i < 1\}$ and if $(z^m, r^m) \in \Omega_A$, $I(\tilde{R}_S^0, z^m) = \{a \in \mathbf{X} : a \cdot r^i < 1, \ a_1 \leq z_1^m\}$. Before proving Theorem 1, we first present two Lemmas.

Lemma 4.1. *If Ω satisfies AG-GARP, then for all $z^0 \in Z$, $z^0 \in \mathrm{d}CMH(\{z^i \in Z : z^i \tilde{R}_A z^0\})$ and $z^0 \in \mathrm{d}CMH(\{z^i \in Z : z^i \tilde{R}_S z^0\})$.*

Proof. We prove by contradiction. Assume that $z^0 \notin \mathrm{d}CMH(\{z^i \in Z : z^i \tilde{R}_S z^0\})$. Let $C = CMH(\{z^i \in Z : z^i \tilde{R}_S z^0\})$, $Z^* = \{z^i \in Z : z^i \tilde{R}_S z^0\}$, $z^0 = \{z^i \in Z : z^i \tilde{R}_S^0 z^0\}$ and $\bar{Z} = Z \cap \mathrm{d}C$. WLOG, we set $\bar{Z} = \{z^i\}_{i=1}^l$ with $z_1^{i+1} < z_1^i$ for $i = 1, \cdots, l-1$.

The Lemma holds trivially when $Z = \{z^0\}$. Suppose that $Z \neq \{z^0\}$. If $z^0 \in \mathrm{int}C$, then there exist $z^{i^*} \in \bar{Z}$, and $z^m \in (Z \cap \mathrm{int}C)$ such that $z^{i^*} \tilde{R}_S^0 z^m$. Next we prove the following two claims.

(i) $z^i \tilde{R}_S^0 z^{i+1}$ for $i \in \{1, \cdots, l-1\}$;

(ii) $z^i \tilde{R}_S^0 z^{i-1}$ for $i \in \{2, \cdots, l\}$.

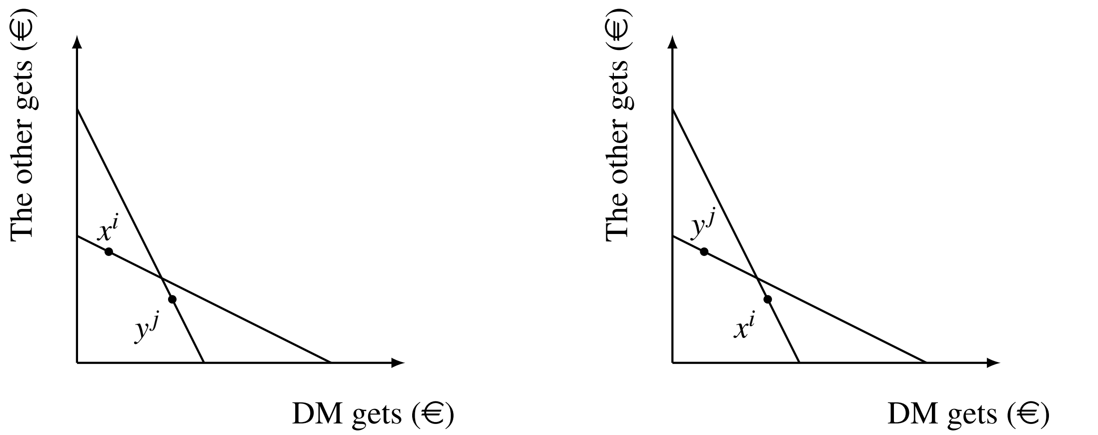
Proof of Claim 1. We start from z^1. Since $z^1 \in \mathrm{d}C$ and $z^1 \tilde{R}_S z^0$, we must have $z^1 \tilde{R}_S^0 z^j$ for some $z^j \in C$. Then $z^j \in \mathbf{B}(r^1)$, which implies $\mathbf{B}(r^1) \cap C \neq \{z^1\}$. Also $z_1^1 > z_1^2$, $z^1 \tilde{R}_S^0 z^2$.

Consider z^2. By AG-GARP, we cannot have $z^2 \tilde{P}_S^0 z^1$. If $z^1 \tilde{R}_S^0 z^2 \tilde{R}_S^0 z^1$, then $\{z^1, z^2\} \in \mathrm{d}\mathbf{B}(r^1)$ and $\mathbf{B}(r^1) = \mathbf{B}(r^2)$. Additionally, $\{z^1, z^1\} \in \mathrm{d}C$, $\mathbf{B}(r^1), \mathbf{B}(r^1)$ are supporting hyperplanes of C. Then both z^1 and z^2 can only be preferred to other choices in C if both are preferred to z^3 and $z^3 \in \mathrm{d}\mathbf{B}(r^2)$. Since $z_1^2 > z_1^3$, then $z^2 \tilde{R}^0 z^3$. If $z^1 \tilde{P}_S^0 z^2$, then $z^2 \tilde{R}_S^0 z^3$ holds the same as we derive $z^1 \tilde{R}_S^0 z^2$. Therefore, by induction, $z^i \tilde{R}_S^0 z^{i+1}$ for $i \in \{1, \cdots, l-1\}$.

Proof of Claim 2. We start from z^l. Since $z^l \in \mathrm{d}C$, we must have $z^l \tilde{R}_S^0 z^j$ for some $z^j \in C$. Because $z_1^l < z_1^i$ for all $z^i \in C$, we cannot have $z^l \tilde{R}_A^0 z^j$. Thus, $z^l \in \{x_i\}_{i=1}^n$ and $z^l R_A^0 z^j$, then $z^l R_A^0 z^{l-1}$. Similarly to the proof of Claim 1, we have $z^i \tilde{R}_S^0 z^{i-1}$ for $i \in \{2, \cdots, l\}$.

(1) The boundary case: if $i^* = 1$, then $z^1 \tilde{R}_S^0 z^m$. Also we have $z^2 \tilde{R}_A^0 z^m$ by definition, then combining $z_1^2 > z_1^1$, we obtain $z^1 \tilde{P}_S^0 z^2$. However by claim (i), $z^1 \tilde{R}_S^0 z^2$, and we have contradiction. A symmetric arguments applies for $i^* = l$.

(2) If $1 < i^* < l$, then if $z_1^m < z_1^l$, the contradiction is the same as for the boundary case $i^* = l$; otherwise it is the same as for $i^* = 1$.



Violation of GARP but not AG-GARP   Violation of both GARP and AG-GARP

Figure 4.5: The relation between GARP and AG-GARP, where x^i is chosen in context A and y^i is chosen in context S.

This finishes the proof of Lemma 4.1. $\square$

Lemma 4.2. *Let $C_A = CMH(\{z^i \in Z : z^i \tilde{R}_A z^0\})$ and $C_S = CMH(\{z^i \in Z : z^i \tilde{R}_S z^0\})$. Let $\{z^i\}_{i=1}^l$ be the set of observed choices on $\mathrm{d}C_A$ or $\mathrm{d}C_S$ such that $z_1^i > z_1^{i+1}$ for all $i = 1, \cdots, l-1$. If Ω satisfies AG-GARP then $z^1 \tilde{R}_A^0 z^2, z^2 \tilde{R}_A^0 z^3, \cdots, z^j \tilde{R}_A^0 z^0$ and $z^l \tilde{R}_A^0 z^{l-1}, z^{l-1} \tilde{R}_A^0 z^{l-2}, \cdots, z^{j+1} \tilde{R}_A^0 z^0$ for some $j \in \{1, \cdots, l\}$; and $z^1 \tilde{R}_A^0 z^2, z^2 \tilde{R}_A^0 z^3, \cdots, z^k \tilde{R}_A^0 z^0$ and $z^l \tilde{R}_S^0 z^{l-1} \tilde{R}_S^0, \cdots, \tilde{R}_S^0 z^{k+1} \tilde{R}_S^0 z^0$ for some $k \in \{1, \cdots, l\}$.*

Proof. By Lemma 4.1, $z^0 \in \mathrm{d}C_A$. There is $z^j \in \{z^i\}_{i=1}^l$, such that (z^j, z^0) and (z^0, z^{j+1}) are on the same supporting hyperplane of C respectively. Following the proof of Lemma 4.1, we have $z^1 \tilde{R}_A^0 z^2, z^2 \tilde{R}_A^0 z^3, \cdots, z^j \tilde{R}_A^0 z^0$ and $z^l \tilde{R}_A^0 z^{l-1}, z^{l-1} \tilde{R}_A^0 z^{l-2}, \cdots, z^{j+1} \tilde{R}_A^0 z^0$ for some $j \in \{1, \cdots, l\}$. $\square$

The Agreement axiom only allows one direction of violations of GARP (see Figure 4.5). Observations in both figures violate GARP. However, if x^i is chosen in context A and y^i is chosen in context S, Agreement allows such choices in the left figure but does not admit those in the right figure. Lemma 4.1 shows that Agreement is still enough to conclude that observations on the boundary of monotonic convex hull are revealed indifferent. Lemma 4.2 explains in detail the relation of observations on the boundary of monotonic convex hull. Next we give the proof of Theorem 1.

Proof of Theorem 2. We only prove the sufficiency of AG-GARP to get the representation in our Theorem. Under Agreement, GARP may be violated when we put Ω_A and Ω_S together. The idea is that we construct virtual budgets for every allocations in Ω_A so that the violations disappear and the information that we learn about the $\succsim_S$ from Ω_A is kept.

Suppose that AG-GARP is satisfied. We show the existence of the more selfish utility function. Start by adding (x^1,p^1) into Ω_S. If $\Omega_S \cup (x^1,p^1)$ satisfies GARP, then we let $\Omega_S^1 = \Omega_S \cup (x^1,p^1)$, Otherwise, let $Z^0 = \{y_j\}_{j=1}^m \cup x^1$, and $C = CMH(\{z^i \in Z^0 : z^i \tilde{R}_S x^1\})$, then $x^1 \in \mathrm{d}C$ by Lemma 4.1. Also by Lemma 4.2, if $z^i \in Z^0$ and x^1 adjacent on dC, then $z^i \tilde{R}_S^0 x^1$. AG-GARP implies that not $x^1 \tilde{R}_S^0 z^i$. Thus $\mathbf{B}(p^1) \cap intC = \emptyset$. Then $I(\tilde{R}_S^0, x^1) \cap C = \emptyset$. Both $I(\tilde{R}_S^0, x^1)$ and C are convex, and by the separating hyperplane theorem, the hyperplane separating $I(\tilde{R}_S^0, x^1)$ and C has the form of $\{a \in \mathbf{X} : (p_1^1 + \theta^1, p_2^1)(a - x^1) = 0\}$ with some $\theta^1 \geq 0$. Denote $p^{1^*} = (p_1^1 + \theta^1, p_2^1)$. For p^{1^*}, $I(\tilde{R}_S^0, x^1) \subset \mathbf{B}(p^{1^*})$ should hold.

There exists $z^i \in Z$ such that the pair of observation (z^i, r^i), (x^1,p^1) violates GARP. The pair satisfies Agreement, so we have $p^1 \cdot x^1 > p^1 \cdot z^i$ and $z^i \tilde{R}_S x^1$ with $z_1^i > x_1^1$. By Lemma 4.2, there exists $z^k \in \mathrm{d}C$ adjacent to x^1 such that $z^k \tilde{R}_S^0 x^1$ with $z_1^k > x_1^1$ and $p^1 \cdot x^1 > p^1 \cdot z^k$. We choose $\hat{\theta}^1 = \frac{p^1 \cdot (x^1 - z^k)}{z_1^k - x_1^1}$, and $\theta^1 = \hat{\theta}^1 + \varepsilon$ with small $\varepsilon > 0$. Thus, $p^{1^*} = (p_1^1 + \theta^1, p_2^1)$ is the budget we construct for allocation x^1. Geometrically, the new budget p^{1^*} is made by rotating the p^1 from the allocation x^1 clockwise, so $I(\tilde{R}_S^0, x^1) \subset \mathbf{B}(p^{1^*})$.

We repeat this for (x^2,p^2). We add (x^2, p^{2^*}) into Ω_S^1. By induction we have $\Omega_S^* = \Omega_S \cup \{x^i, p^{i^*}\}_{i=1}^n$ and Ω_S^* satisfies GARP. In the same way, we have $\Omega_A^* = \Omega_A \cup \{y^j, q^{j^*}\}_{j=1}^n$ and Ω_A^* satisfies GARP.

Let $\{\alpha_i\}_{i=1}^\infty$ be the countable dense of $\mathbb{R}_+^2$, $F_k = \{\alpha_1, \cdots, \alpha_k\}$ and $F_\infty = \{\alpha_1, \cdots, \alpha_k, \cdots\}$.

Lemma 4.3. *Given F_k, there exist two sets of prices $\{l_i\}_{i=1}^k$ and $\{l_i'\}_{i=1}^k$ such that $\Omega_A^* \cup \{\alpha_i, l_i\}_{i=1}^k$ and $\Omega_S^* \cup \{\alpha_i, l_i'\}_{i=1}^k$ satisfy GARP and jointly satisfies Agreement.*

Proof. By Lemma 3 of Reny(2015), there are virtual prices l_i corresponding to α_i for all i, such that $\Omega_A^* \cup \{\alpha_i, l_i\}_{i=1}^k$ satisfies GARP. Then we apply the techniques we did before to rotate the prices so that there are l_i' corresponding to α_i for all i, such that $\Omega_S^* \cup \{\alpha^i, l_i'\}_{i=1}^k$, and the construction guaranteed that agreement is satisfied. ☐

We denote $\Omega_A^k = \Omega_A^* \cup \{\alpha^i, l_i\}_{i=1}^k$ and $\Omega_S^k = \Omega_S^* \cup \{\alpha^i, l_i'\}_{i=1}^k$. And both Ω_A^k and Ω_S^k satisfy GARP, and jointly satisfies Agreement.

Lemma 4.4. *There is monotonic and quasiconcave u_{F_k} rationalizes Ω_A^k such that*

$$u_{F_k}(x) \leq \frac{1}{2^{k'}} \max\left(\frac{x(j)}{\alpha(j)} - 1\right) + u_{F_k}(\alpha^{k'}) \quad \forall k' \leq k \text{ and } x << \alpha^{k'};$$

and there is monotonic and quasiconcave v_{F_k} rationalizes Ω_S^k such that

$$v_{F_k}(x) \leq \frac{1}{2^{k'}} \max\left(\frac{x(j)}{\alpha^{k'}(j)} - 1\right) + v_{F_k}(\alpha_{k'}) \quad \forall k' \leq k \text{ and } x << \alpha^{k'}$$

Moreover, let $O^k = \{x^i\}_{i=1}^n \cup \{y^j\}_{j=1}^m \cup \{\alpha^i\}_{i=1}^k$ *for* $x, y \in O^k$, $u_{F_k}(x) \geq u_{F_k}(y)$ *with* $x_1 > y_1$, *we have that* $v_{F_k}(x) \geq v_{F_k}(y)$.

Proof. Existence of u_{F_k} and u_{F_k}, see Reny(2015). Our construction in Lemma 4.4 guaranteed the relation that for $x, y \in O^k$, $u_{F_k}(x) \geq u_{F_k}(y)$ with $x_1 > y_1$, we have that $v_{F_k}(x) \geq v_{F_k}(y)$. □

Let $k_2 \geq k_1$, then $\Omega_A^{k_1} \subset \Omega_A^{k_2}$, from the construction, $u_{k_2}(x) \leq u(k_1)(x)$, so $u_{F_k}(x)$ is decreasing with k. Similarly, we have $v_{F_k}(x)$ is decreasing with k.

Define altruistic and selfish utilities as

$$u := \inf_k(u_{F_k}) \quad \text{and} \quad v := \inf_k(v_{F_k}).$$

We show the two properties of u and v.

1. u and v rationalize the Ω.
   For all k, u_{F_k} AG-rationalizes Ω, so altruistic utility $\inf_k(u_{F_k})$ AG-rationalizes Ω. Similarly, selfish utility $\inf_k(v_{F_k})$ AG-rationalizes Ω.

2. u and v are monotonic and quasiconcave .
   Given $x \geq y \in \mathbb{R}_+^2$, we have that for all k, $u_{F_k}(x) \geq u_{F_k}(y)$ by the monotonicity of u_{F_k}. Thus, $\inf_k(u_{F_k})(x) \geq \inf_k(u_{F_k})(y)$, that is, $u(x) \geq u(y)$.

   Given $x, y \in \mathbb{R}_+^2$, we have for all k and for all $\lambda \in (0,1)$,

   $$u_{F_k}(\lambda x + (1-\lambda)y) \geq \min\{u_{F_k}(x), u_{F_k}(y)\},$$

   then we have,

   $$\begin{aligned} u(\lambda x + (1-\lambda)y) = \inf_k u_{F_k}(\lambda x + (1-\lambda)y) &\geq \inf_k(\min\{u_{F_k}(x), u_{F_k}(y)\}) \\ &\geq \min\{\inf_k u_{F_k}(x), \inf_k u_{F_k}(y)\}) \\ &= \min(u(x), u(y)) \end{aligned}$$

   Similarly, v is monotonic and quasiconcave.

Next we show the following statement:
For all $x, y \in \mathbf{X}$, $u(x) > u(y)$ with $x_1 > y_1$ implies that $v(x) \geq v(y)$ almost everywhere.

First we show this statement holds for F_∞. Assume for contraction that there are $\alpha^i, \alpha^j \in F_\infty$ with $u(\alpha^i) > u(\alpha^j)$ with $\alpha_1^i > \alpha_1^j$ implies that $v(\alpha^i) < v(\alpha^j)$. By definition of u and v, we have that

$$\inf_k(u_{F_k}(\alpha^i)) > \inf_k(u_{F_k}(\alpha^j)) \text{ with } \alpha_1^i > \alpha_1^j \Rightarrow \inf_k(v_{F_k}(\alpha^i)) < \inf_k(v_{F_k}(\alpha^j))$$

Then there is $m \in \mathbb{N}$, for all $k > m$, $u_{F_k}(\alpha^i) > u_{F_k}(\alpha^j)$ with $\alpha_1^i > \alpha_1^j$ implies that $v_{F_k}(\alpha^i) > v_{F_k}(\alpha^j)$.

Then let k^* be any integer larger than k, we have that for all $k > k^*$ $\alpha_1^i > \alpha_1^j$, $u_{F_k}(\alpha^i) > u_{F_k}(\alpha^j)$ implies that $v_{F_k}(\alpha^i) > v_{F_k}(\alpha^j)$, which contradict to Lemma 4.4.

Until now, we have shown for a dense set of $\mathbb{R}^2_+$, the statement holds. Since u and v are strictly increasing, u and v are continuous almost everywhere. Choose $x, y \in \mathbb{R}^2_+$ such u and v are both continuous at x, y, and $u(x) > u(y)$ with $x_1 > y_1$. We choose two sequences $\{\alpha^i\}_{i=1}^{\infty}, \{\beta^i\}_{i=1}^{\infty} \in F_\infty$ such that $\lim_{i\to\infty}(\alpha^i) \downarrow x$ and $\lim_{i\to\infty}(\alpha^i) \uparrow y$ and $\alpha_1^i > \beta_1^i$. By monotonicity, we have for all i, $u(\alpha^i) > u(\beta^i)$, then $v(\alpha^i) > v(\beta^i)$ for all i, which implies that $v(x) \geq v(y)$. Therefore, the statement holds almost everywhere. This finishes the proof. $\square$

Theorem 4. *Suppose $\succsim_A$ and $\succsim_S$ are represented by CES functions, that is, $u(a) = (\alpha_A a_1^{\rho_A} + (1 - \alpha_A)a_2^{\rho_A})^{1/\rho_A}$ for $\succsim_A$ and $v(a) = (\alpha_S a_1^{\rho_S} + (1 - \alpha_S)a_2^{\rho_S})^{1/\rho_S}$ for $\succsim_S$. Then Agreement is satisfied if and only if $\alpha_A \leq \alpha_S$ and $\rho_A = \rho_S$.*

Proof. $\Leftarrow$ Assume that $\alpha_A \leq \alpha_S$ and $\rho_A = \rho_S = \rho$. If $x \succsim_A y$ and $x_1 > y_1$, then

$$u(x) = (\alpha_A x_1^\rho + (1 - \alpha_A)x_2^\rho)^{1/\rho} \geq u(y) = (\alpha_A y_1^\rho + (1 - \alpha_A)y_2^\rho)^{1/\rho}.$$

Since $x_1 > y_1$ and $\alpha_A \leq \alpha_S$, then

$$v(x) = (\alpha_S x_1^\rho + (1 - \alpha_S)x_2^\rho)^{1/\rho} \geq v(y) = (\alpha_S y_1^\rho + (1 - \alpha_S)y_2^\rho)^{1/\rho},$$

Thus, we have $x \succsim_S y$.

$\Rightarrow$ Assume that Agreement holds. First we show that Agreement implies that marginal rate of substitution (MRS) of the altruistic preference is smaller than that of selfish preference. Given $\varepsilon > 0$, $\lambda(\varepsilon)$ is the value such that the following indifference holds:

$$(x+\varepsilon, y) \sim_A (x, y+\lambda(\varepsilon)) \tag{4.3}$$

By Agreement, 4.3 implies that

$$(x+\varepsilon, y) \succsim_S (x, y+\lambda(\varepsilon)) \tag{4.4}$$

We write 4.3 and 4.4 in utility functions,

$$u(x+\varepsilon, y) = u(x, y+\lambda(\varepsilon)) \Rightarrow v(x+\varepsilon, y) \geq v(x, y+\lambda(\varepsilon)),$$

which is equivalent to,

$$\frac{u(x+\varepsilon,y)-u(x,y)}{\varepsilon}=\frac{u(x,y+\lambda(\varepsilon))-u(x,y)}{\lambda(\varepsilon)}$$

$$\Rightarrow \frac{v(x+\varepsilon,y)-v(x,y)}{\varepsilon}\geq\frac{v(x,y+\lambda(\varepsilon))-v(x,y)}{\lambda(\varepsilon)},$$

That is,

$$\frac{\frac{u(x+\varepsilon,y)-u(x,y)}{\varepsilon}}{\frac{u(x,y+\lambda(\varepsilon))-u(x,y)}{\varepsilon}}=\frac{\varepsilon}{\lambda(\varepsilon)}$$

$$\Rightarrow \frac{\frac{v(x+\varepsilon,y)-v(x,y)}{\varepsilon}}{\frac{v(x,y+\lambda(\varepsilon))-v(x,y)}{\varepsilon}}\geq\frac{\varepsilon}{\lambda(\varepsilon)}.$$

Since u and v are CES functions which are smooth, let $\varepsilon \downarrow 0$ and $K=\lim_{\varepsilon\downarrow 0}\frac{\varepsilon}{\lambda(\varepsilon)}$, then we have

$$\frac{\frac{\partial u}{a_1}}{\frac{\partial u}{a_2}}=MRS_u=K\Rightarrow\frac{\frac{\partial v}{a_1}}{\frac{\partial v}{a_2}}MRS_v\geq K.$$

Therefore, $MRS_{=}v\geq MRS_u$.

Substituting the $MRS_v\geq MRS_u$, with CES functions, we have for all $a=(a_1,a_2)\in\mathbf{X}$,

$$\frac{\alpha_A}{1-\alpha_A}(\frac{a_1}{a_2})^{\rho_A}\geq\frac{\alpha_S}{1-\alpha_S}(\frac{a_1}{a_2})^{\rho_S}$$

Then $a_1\geq a_2$ and $\rho_A=\rho_S$.

□

REFERENCES

AFRIAT, S. N. (1967) "The construction of utility functions from expenditure data," *International economic review* 8, 67–77.

AFRIAT, S. N. (1972) "International Comparisons of Prices and Output" *NBER, chap. The Theory of International Comparisons of Real Income and Prices* , 13–84.

ANDREONI, J. AND J. MILLER (2002) "Giving according to GARP: An experimental test of the consistency of preferences for altruism," *Econometrica* 70, 737–753.

BANERJEE, S. AND J. H. MURPHY (2006) "A simplified test for preference rationality of two-commodity choice," *Experimental Economics* 9, 67–75.

BLOUNT, S. (1995) "When social outcomes aren't fair: the effect of causal attributions on preferences," *Organizational Behavior and Human Decision Processes* 63, 131–144.

DUFWENBERG, M. AND G. KIRCHSTEIGER (2004) "A theory of sequential reciprocity," *Games and economic behavior* 47, 268–298.

FALK, A., E. FEHR, AND U. FISCHBACHER (2008) "Testing theories of fairness - Intentions matter," *Games and economic behavior* 62, 287–303.

FALK, A. AND U. FISCHBACHER (2006) "A theory of reciprocity," *Games and economic behavior* 54, 293–315.

RABIN, M. (1993) "Incorporating fairness into game theory and economics," *American economic review* 54, 1281–1302.

RENY, P. J. (2015) "A characterization of rationalizable consumer behavior," *Econometrica* 83, 175–192.

SEGAL, U. AND J. SOBEL (2007) "Tit for tat: Foundations of preferences for reciprocity in strategic settings," *Journal of Economic Theory* 136, 197–216.

VARIAN, H. R. (1982) "The nonparametric approach to demand analysis," *Econometrica* , 945–973.

Chapter 5

The Uniqueness of Local Proper Scoring Rules: The Logarithmic Family

Abstract

Local proper scoring rules provide convenient tools for measuring subjective probabilities. The only differentiable local proper scoring rule for more than two exclusive events is the logarithmic family. For applications, relaxations of the technical restrictions of differentiability, continuity, and domain should be considered. Our result generalizes many results in the literature and provides new support for the logarithmic family: it is still the only family without those restrictions.

5.1 Introduction

Proper scoring rules play an important role in the measurement of subjective probabilities (Giles 2002; Johnstone 2007). They provide simple questions for measuring subjective probabilities. For comparisons of measurements of subjective probabilities, see Hollard et al (2016). Properness means that only when subjects report their true subjective probabilities, they receive the maximum expected payoff (incentive compatibility). Locality is a desirable property for proper scoring rules. It means that subjects' payoff for a particular event depends only on the subjective probability assigned to that event, and is independent of how the remaining probability mass is distributed over the remaining events. Then, compared to other proper scoring rules, subjects can better understand incentive compatibility and will better satisfy it. This note, therefore, focuses on local proper scoring rules.

In practice, we usually want to know subjects' subjective probabilities for more than two events. Goldstein and Rothschild (2014) showed that subjects can better understand probabilities when presented all events at the same time. Thus, proper scoring rules for more than two events are needed. We can reasonably expect that there will be less probability distortions when proper scoring rules for more than two events are used than when they are used repeatedly. For corrections of such distortions for two events, see Offerman et al 2009).

Locality has traditionally been put forward as a strong argument in favor of the logarithmic family. For more than two exclusive events, Shuford et al (1966), Aczel and Pfanzagl (1967), and Bernardo (1979) proved that the logarithmic family is the only local proper one. Shuford et al (1966), Aczel and Pfanzagl (1967), and Bernardo (1979)'s work on local proper scoring rules were all based on this result. However, all the aforementioned results, and also Parry et al (2012) and Dawid et al (2012) made the assumption of differentiability of the scoring rules.

The logarithmic family becomes too extreme for rare events (it goes to negative infinity). However, in practice, we cannot pay that large amount of money to subjects to elicit their beliefs for rare events. This note examines whether relaxations of differentiability can give more local scoring rules. This research question did not arise from mathematical curiosity, but because it is the question to be asked for practical relevance.[1] Other things equal, differentiability has no practical importance. Every practitioner will, ceteris paribus, immediately give up differentiability if better scoring rules can be regained this way. For applications it is, therefore, important to settle the case without differentiability assumptions. We similarly relax continuity, because, again, ceteris paribus, there is no practical relevance for continuity. We also consider relaxing properness by giving up uniqueness of the optimum of truth telling (weak properness)[2]

[1] I thank Drazen Prelec (personal communication) for having raised this question.

[2] Generally speaking, when requiring only weak properness, we give up the one-to-one relation between types and scores of Myerson (1982)'s and Johnson et al (1990)'s truth revelation. Yet, under this weak condition, people

and by giving up properness at the extremes $p = 0$ and $p = 1$.

Besides all favorable features of the logarithmic family (Diks et al, 2011, Bickel, 2007), this note shows that the logarithmic family is still the only one available after all above relaxations. To the best of my knowledge, only Savage (1971) considered local proper scoring rules without presupposing differentiability or continuity. However, he did not discuss practical relevance, and the literature did not focus on this feature of Savage's work (Carvalho 2016 ; Schlag et al 2015). We generalize Savage (1971) to local weakly proper scoring rules, and provide a simpler proof. Essentially, we replace his advanced geometric arguments by simple first-order optimality conditions, commonly used in economics, and adapted here to the nondifferentiable case. We end with representing Savage's work in section 5.4. Furthermore, his proof is not based on elementary mathematics (section 5.4). Both Savage and we have dealt with the proper scoring rules without technical assumptions (differentiability and continuity), and we assure practitioners that these techincal properties hold automatically.

5.2 Main result

We first discuss our scoring rule when the endpoints $p = 0$ and $p = 1$ are absent, so as to also handle the attempt of giving up properness at the extremes. Let $(0,1)^n$ be the Cartesian product $\underbrace{(0,1) \times \cdots \times (0,1)}_{n}$. We define $\Delta^n = \{(r_1, \ldots, r_n) \in (0,1)^n : \sum_{i=1}^n r_i = 1\}$. The scoring mechanism works as follows. A subject is facing n mutually exclusive and exhaustive events, $(E_1, \ldots, E_n)$. We assume that the subject maximizes subjective expected value, with subjective probabilities $p_1, \ldots, p_n$. Under expected value, proper scoring rules it suffices to identify subjective probabilities. For the moderate payoffs as used in proper scoring rules, linear utility is reasonable. Hence, subjective expected value maximization is a reasonable assumption for prescriptive applications of proper scoring rules. We thus, do not consider ambiguity models, that is, we assume that subjects have no stakes in states, and that ambiguity attitudes do not affect our measurements. Violations of these assumptions are analyzed by Karni and Safra (1995) and Abdellaoui et al (2017). To reveal the subjective probabilities, we ask the subject to report a number, r_i, between 0 to 1 for each event E_i, and those numbers r_i should add to 1. Then a payoff, $f(r_i)$, under every event E_i, is determined by the number r_i, the subject reported for this event. *Locality* means that the payoff under event E_i depends only on r_i.

Definition 5.1. *A family of non-constant functions $\{f_i\}_{i \leq n}$, $f_i : (0,1) \to \mathbb{R}$ is a* local weakly

still have no incentive to misrepresent. Thus, they may still be truth telling.

proper scoring rule *if*

$$\sum_{i=1}^{n} p_i f_i(r_i) \leq \sum_{i=1}^{n} p_i f_i(p_i) \tag{5.1}$$

for all $(r_1,\ldots,r_n),(p_1,\ldots,p_n)\in\Delta^n$.

Under definition 5.1, reporting true subjective probabilities is a (non-unique) dominating strategy.

Theorem 5. *If* $\{f_i\}_{i\leq n}$ *is a local weakly proper scoring rule defined on* $(0,1)^n$ *with* $n\geq 3$*, then there is* $k>0$*, and for all* $i\leq n$ *there is* $a_i\in\mathbb{R}$ *such that*

$$f_i(r) = k\ln r + a_i \quad \textit{for} \quad i\leq n. \tag{5.2}$$

Section 5.3 shows that it is impossible to extend the scoring rules to the extremes $p=0$ or $p=1$ for $k>0$, reinforcing the impossibility result. For $k=0$, local proper scoring rules are constant, which is of limited interest for applications. Moreover, if we extend the range of our scoring rules to $\mathbb{R}\cup\{-\infty\}$, then the domain of the scoring rules in equation (5.2) can be formally extended to $[0,1]^n$, which, however, does not help for applications.

5.3 Proofs

Lemma 5.1. *If* $\{f_i\}_{i\leq n}$ *is a local weakly proper scoring rule with* $n\geq 3$*, then* $\{f_i\}_{i\leq n}$ *is non-decreasing on* $(0,1/2)$.

PROOF. Given $p\in(0,1/2)$, and a small $\varepsilon>0$ such that $\varepsilon<1-p$ and $\varepsilon<p$, probability vectors in definition 5.1 are:

$$(r_1,\ldots,r_n) = (p+\varepsilon, p-\varepsilon, \frac{1-2p}{n-2},\ldots,\frac{1-2p}{n-2}), \quad (p_1,\ldots,p_n) = (p,p,\frac{1-2p}{n-2},\ldots,\frac{1-2p}{n-2}).$$

Then, in a first application of Inequality 5.1 with probability vector $\{p_1,\cdots,p_n\}$ for events, we obtain

$$pf_1(p)+pf_2(p)+\frac{1-2p}{n-2}\sum_{i=3}^{n} f_i\left(\frac{1-2p}{n-2}\right) \geq pf_1(p+\varepsilon)+pf_2(p-\varepsilon)+\frac{1-2p}{n-2}\sum_{i=3}^{n} f_i\left(\frac{1-2p}{n-2}\right).$$

That is,

$$f_2(p)-f_2(p-\varepsilon)\geq f_1(p+\varepsilon)-f_1(p). \tag{5.3}$$

Moreover, a second application of Inequality 5.1 with probability vector $\{r_1,\cdots,r_n\}$, we obtain

$$(p+\varepsilon)f_1(p+\varepsilon)+(p-\varepsilon)f_2(p-\varepsilon)\geq(p+\varepsilon)f_1(p)+(p-\varepsilon)f_2(p)$$

implies

$$(p+\varepsilon)(f_1(p+\varepsilon)-f_1(p)) \geq (p-\varepsilon)(f_2(p)-f_2(p-\varepsilon)). \tag{5.4}$$

Thus, (5.3) multiplied by $p+\varepsilon$ combined with (5.4) and transitivity leads to

$$(p+\varepsilon)(f_2(p)-f_2(p-\varepsilon)) \geq (p-\varepsilon)(f_2(p)-f_2(p-\varepsilon)).$$

That is, $f_2(p)-f_2(p-\varepsilon) \geq 0$, and f_2 is non-decreasing on $(0,1/2)$. By the similar arguments, i.e. we exchange f_2 with any f_i in above reasoning, for all $i \leq n$, f_i is non-decreasing on $(0,1/2)$.

Definition 5.2. *For a function f, we introduce the following notation:*

$$f^{R+}(p) := \limsup_{\varepsilon \downarrow 0} \frac{f(p+\varepsilon)-f(p)}{\varepsilon};$$

$$f^{R-}(p) := \liminf_{\varepsilon \downarrow 0} \frac{f(p+\varepsilon)-f(p)}{\varepsilon};$$

$$f^{L+}(p) := \limsup_{\varepsilon \downarrow 0} \frac{f(p)-f(p-\varepsilon)}{\varepsilon};$$

$$f^{L-}(p) := \liminf_{\varepsilon \downarrow 0} \frac{f(p)-f(p-\varepsilon)}{\varepsilon}.$$

If $f^{R+}(p) = f^{R-}(p)$, we denote them as $f^R(p)$; $f^L(p)$ is defined analogously. If $-\infty < f^R(p) = f^L(p) < \infty$, then f is differentiable at p.

Lemma 5.2. *Let $\{f_i\}_{i \leq n}$ be a local weakly proper scoring rule with $n \geq 3$, and let $p,q \in (0,1)$ with $p+q<1$. If f_1 is differentiable at p, then f_2 is differentiable at q, and $pf_1'(p) = qf_2'(q)$.*

PROOF. Let $\varepsilon > 0$ be small and $\varepsilon < 1-p-q$. Probability vectors in definition5.1 are:

$$(r_1,\ldots,r_n) = (p-\varepsilon, q-\varepsilon, \frac{1-p-q}{n-2},\ldots,\frac{1-2p}{n-2}), \quad (p_1,\ldots,p_n) = (p,q,\frac{1-p-q}{n-2},\ldots,\frac{1-2p}{n-2}).$$

Then, we have

$$\begin{aligned} &pf_1(p)+qf_2(q)+\frac{1-p-q}{n-2}\sum_{i=3}^{n} f_i\left(\frac{1-p-q}{n-2}\right) \\ &\geq pf_1(p-\varepsilon)+qf_2(q+\varepsilon)+\frac{1-p-q}{n-2}\sum_{i=3}^{n} f_i\left(\frac{1-p-q}{n-2}\right), \end{aligned}$$

which implies

$$p\frac{f_1(p)-f_1(p-\varepsilon)}{\varepsilon} \geq q\frac{f_2(q+\varepsilon)-f_2(q)}{\varepsilon}.$$

Because $f_1'(p)$ exists, letting $\varepsilon \downarrow 0$, we have

$$pf_1'(p) \geq qf_2^{R+}(q). \tag{5.5}$$

Similarly,

$$\begin{aligned}&(p-\varepsilon)f_1(p)+(q+\varepsilon)f_2(q)+\frac{1-p-q}{n-2}\sum_{i=3}^{n} f_i\left(\frac{1-p-q}{n-2}\right)\\&\leq (p-\varepsilon)f_1(p-\varepsilon)+(q+\varepsilon)f_2(q+\varepsilon)+\frac{1-p-q}{n-2}\sum_{i=3}^{n} f_i\left(\frac{1-p-q}{n-2}\right)\end{aligned}$$

implies

$$pf_1'(p) \leq qf_2^{R-}(q) \tag{5.6}$$

Thus, (5.5) and (5.6) imply

$$pf_1'(p) = qf_2^{R}(q).$$

By choosing $\varepsilon < 0$, we can similarly show $pf_1'(p) = qf_2^{L}(q)$. We have proved

$$pf_1'(p) = qf_2'(q).$$

Lemma 5.3. *If $\{f_i\}_{i\leq n}$ is a local weakly proper scoring rule with $n \geq 3$, then for all $i \leq n$, f_i is differentiable on $(0,1)$, and moreover, for all $i, j \leq n$ and $p, q \in (0,1)$, $pf_i'(p) = qf_j'(q)$.*

PROOF. Lemma 5.2 implies that for all $i, j \leq n$ with $j \neq i$, and $p, q \in (0,1)$ with $p+q < 1$, if f_i is differentiable at p, f_j is differentiable at q. Statement 4.5 of van Rooij and Schikhof (1982) states that every monotonic function is almost everywhere differentiable. Given $i \leq n$ and $q \in (0,1)$, Lemma 5.1 implies that there is $p \in (0, \min\{1/2, 1-q\})$, such that for some $j \leq n$ with $j \neq i$, f_j is differentiable at p, then f_i is differentiable at q. Thus, for all $i \leq n$, f_i is differentiable on $(0,1)$.

Again by Lemma 5.2, we have for all $i, j \leq n$ and $j \neq i$, $p, q \in (0,1)$ with $p+q < 1$, $pf_i'(p) = qf_j'(q)$. Given $p, q \in (0,1)$, and $i, j \leq n$, there exists $j' \leq n$ with $j' \neq i$ and $j' \neq j$, and exists $q' \in (0,1)$ with $p+q' < 1$ and $q+q' < 1$. Then, we have $pf_i'(p) = q'f_{j'}'(q')$ and $qf_j'(q) = q'f_{j'}'(q')$. Therefore, $pf_i'(p) = qf_j'(q)$ for all $i, j \leq n$ and $p, q \in (0,1)$.

PROOF OF THEOREM 5.1. By Lemma 5.2 and Lemma 5.3, f_i is differentiable on $(0,1)$, and for all $i, j \leq n; p, q \in (0,1)$, $pf_i'(p) = qf_j'(q)$. Thus, we get, $f_i'(p) = \frac{k}{p}$ for all $p \in (0,1), i \leq n$ and some $k > 0$, since f_i is non constant for all i. Therefore, (5.2) is obtained by integrating the latter equations.

PROOF OF THE EXTENSION. We know that if a local scoring rule cannot be defined at 0, then it cannot be defined at 1, because the probabilities must add to 1. Assume, for contradictions, that a local weakly proper scoring rule can be defined at 0, that is, there is $k \in \mathbb{R}$, such that $f_1(0) = k$. f_1 is a logarithmic function on $(0,1)$, so there is $p \in (0,1)$ such that $f_1(p) \leq k$. Thus, a subject can report 0 for an event with subjective probability p, and the probability mass p is distributed over the other events. Because all f_i are strictly increasing, he receives a strictly better payoff. From above analysis, it follows that $f_1(0)$ cannot be defined, and similarly $f_i(0)$ cannot be defined for any $i \leq n$. Therefore, we cannot extended the scoring rule to $[0,1)^n$, or $(0,1]^n$, or $[0,1]^n$.

Now we extend the range of our scoring rules to $\mathbb{R} \cup \{-\infty\}$, and define $f_i(0) = -\infty$ for all $i \leq n$. Assume, for contradiction, that $f_1(1) \leq f_1(r)$ for some $r \in (0,1)$. For a sure event, a subject can report r, and the remaining probability mass $1-r$ goes to the other events. Then, he gets a strictly higher payoff because all f_i are strictly increasing. Therefore,

$$f_i(1) > f_i(r) \text{ for all } r \in (0,1) \text{ and } i \leq n \tag{5.7}$$

Moreover, given $\varepsilon \in (0,1)$, we have

$$\varepsilon f_1(\varepsilon) + (1-\varepsilon) f_2(1-\varepsilon) \geq \varepsilon f_1(1) + (1-\varepsilon) f_2(0), \tag{5.8}$$

which implies, $\liminf_{\varepsilon \uparrow 1} f_1(\varepsilon) \geq f_1(1)$, then for all $i \leq n$, $\liminf_{\varepsilon \uparrow 1} f_i(\varepsilon) \geq f_i(1)$. Combining (5.7),

$$\liminf_{\varepsilon \uparrow 1} f_i(\varepsilon) \geq f_i(1) \geq \limsup_{\varepsilon \uparrow 1} f_i(\varepsilon), \text{ that is, } f_i \text{ is left continous at 1.}$$

Therefore, the domain of (5.2) can be extended to $[0,1]^n$.

5.4 Savage (1971)

In Section 9.4 of Savage (1971), Inequality 5.1 of Definition 5.1 is a strict inequality. We present Savage's proof adapted to our notation here. Given $p, q, w \in (0,1)$, and $\bar{p} = 1-p$, $\bar{q} = 1-q$, probability vectors in definition5.1 with strict inequality as follows:

$$(p_1, \ldots, r_n) = (pw, \bar{p}w, \frac{\bar{w}}{n-2}, \ldots, \frac{\bar{w}}{n-2}), \quad (p_1, \ldots, p_n) = (qw, \bar{q}w, \frac{\bar{w}}{n-2}, \ldots, \frac{\bar{w}}{n-2}).$$

Then we have,

$$q f_1(qw) + \bar{q} f_2(\bar{q}) > q f_1(pw) + \bar{q} f_1(\bar{p}w).$$

Define,

$$g_w(q) = qf_1(qw) + \bar{q}f_2(\bar{q}), \tag{5.9}$$

then Savage showed g_w is strictly convex in $(0,1)$, so g_w is continuous in q, and almost everywhere differentiable in q.

Since $f_1(qw) - f_2(\bar{q}w)$ is the slope of g_w at q, it is locally bounded variation[3]. Thus $f_1(qw) - f_2(\bar{q}w)$ is also almost everywhere differentiable. As a function of y, $f_1(qw) = g_w(q) + \bar{q}[f_1(qw) - f_2(\bar{q}w)]$, which is the summation of two almost everywhere differentiable functions. Therefore, $f_1(qw)$ is almost sure differentiable in q. By a similar argument as in our proof above, we can conclude that $f_1(qw)$ is differentiable for all q. Taking the derivative of q at both sides of (5.9),

$$\frac{d}{dq}g_w(q) = \frac{d}{dq}(qf_1(qw) + \bar{q}f_2(\bar{q})) = w[qf_1'(qw) - \bar{q}f_2'(\bar{q}w)] + f_1(qw) - f_2(\bar{q}w).$$

Thus, we have $qwf_1'(qw) = \bar{q}wf_2'(\bar{q}w)$. Then we conclude that $f_i(r) = k\ln r + a_i$ for $i \leq n$, some $k \geq 0$ and $a_i \in \mathbb{R}$.

Our proof also works for the strict local proper scoring rules. In this sense, we have generalized Savage's work. Both Savage's and our proof tried in their first step to show that local scoring rules are almost sure differentiable, then differentiability of local scoring rules follows from the special relationship among the individual events. The main difference in Savage's proof is that we derived that local scoring rules are monotonic on $(0,1/2)$, then they are almost everywhere differentiable. However, Savage wrote the local scoring rules as the summation of a strictly convex function and a function with locally of bounded variation, that is, the summation of two almost sure differentiable functions. The rests of the proofs are similar in spirit.

5.5 Conclusion

Proper scoring rules provide an important tool to elicit people's subjective beliefs and locality is a desirable property in practice. We searched for more families by relaxing differentiability, continuity, properness, and the domain of local scoring rules. Our result not only generalizes a number of theorems, which assume differentiability, continuity or properness, in the literature, but also provides new support for the logarithmic family.

[3] It is unclear how Savage have the property of locally of bounded variation for this slope function.

References

Abdellaoui, Mohammed, Han Bleichrodt, Emmanuel Kemel and Olivier l'Haridon (2017) "Purely Subjective Maxmin Expected Utility," *Working Paper*.

Aczél, János and Johann Pfanzagl (1967) "Remarks on the measurement of subjective probability and information," *Metrika* 11, 91–105.

Bernardo, José M (1979) "Expected information as expected utility," *the Annals of Statistics* , 686–690.

Bickel, J. Eric (2007) "Some comparisons among quadratic, spherical, and logarithmic scoring rules," *Decision Analysis* 4 , 49–65.

Carvalho, Arthur (2016) "An overview of applications of proper scoring rules," *cision Analysis* 13 , 223–242.

Dawid, A. Philip (2007) "The geometry of proper scoring rules," *Annals of the Institute of Statistical Mathematics* 59 , 77–93.

Dawid, A. Philip, Steffen Lauritzen, and Matthew Parry (2012) "Proper local scoring rules on discrete sample spaces," *The Annals of Statistics* 40, 593–608.

Diks, Cees, Valentyn Panchenko, and Dick Van Dijk (2011) "Likelihood-based scoring rules for comparing density forecasts in tails," *Journal of Econometrics* 163, 215–230.

Geweke J, Amisano G (2011) "Optimal prediction pools," *Journal of Econometrics* 164, 130–141.

Giles J (2002) " Scientific wagers: Wanna bet?," *Nature* 420, 354–355.

Giles J (2002) " Lay understanding of probability distributions," *Judgment and Decision Making* 9, 354–355.

Hollard G, Massoni S, Vergnaud JC (2016) " Ln search of good probability assessors: an experimental comparison of elicitation rules for confidence judgments ," *Theory and Decision* 80, 363–387.

Johnstone DJ (2007) " The value of a probability forecast from portfolio theory," *Theory and Decisiona* 62, 153–203.

Karni E, Safra Z (1995) " The impossibility of experimental elicitation of subjective probabilities," *Theory and Decisiona* 38, 313–320.

Myerson RB (1982) " Optimal coordination mechanisms in generalized principal–agent problems," *Journal of Mathematical Economics* 10, 67–81.

Offerman T, Sonnemans J, Van de Kuilen G, Wakker PP (2009) " A truth serum for non-bayesians: Correcting proper scoring rules for risk attitudes," *The Review of Economic Studies* 76, 1461–1489.

Parry M, Dawid AP, Lauritzen S (2012) " Proper local scoring rules," *Annals of Statistics* 40, 561–592.

van Rooij AC, Schikhof WH (1982) " A Second Course on Real Functions," *CUP Archive* .

Savage LJ (1971) " Elicitation of personal probabilities and expectations," *Journal of the American Statistical Association* 66, 738–801.

Schlag K, Tremewan J, van der Weele J (2015) " A penny for your thoughts: a survey of methods for eliciting beliefs," *Experimental Economics* 18, 457–490.

Shuford EH, Albert A, Edward Massengill H (1966) " Admissible probability measurement procedures," *Psychometrikas* 31, 125–145.

Chapter 6

Reconciling Savage's and Luce's Modeling of Uncertainty: The Best of Both Worlds[1]

Abstract

This paper recommends using mosaics, rather than (σ-)algebras, as collections of events in decision under uncertainty. We show how mosaics solve the main problem of Savage's (1954) uncertainty model, a problem pointed out by Duncan Luce. Using mosaics, we can connect Luce's modeling of uncertainty with Savage's. Thus, the results and techniques developed by Luce and his co-authors become available to currently popular theories of decision making under uncertainty and ambiguity.

[1] This chapter is joint with Junyi Chai, Chen Li, Peter P. Wakker and Tong V. Wang.

6.1 Introduction

Savage (1954) introduced the best-known and most-used model for decision under uncertainty, with gambles[2] mapping states to consequences. A decision maker chooses a gamble, nature independently chooses a state, and the corresponding consequence results. Duncan Luce pointed out some serious drawbacks to Savage's model. Throughout his career, Luce used the following example to illustrate these drawbacks. We use it as the lead example in our paper:

> If one is considering a trip from New York to Boston, there are a number of ways that one might go. Probably the primary ones that most of us would consider are, in alphabetical order, airplane[3] bus, car, and train.

When you consider each transportation alternative, you can focus on the uncertainties relevant to that alternative only. However, Savage's model requires you to consider not only the separate uncertainties regarding each alternative, but also all joint uncertainties. Thus, when choosing between airplane and car, you have to consider your degree of belief that both the airplane and the car (had it been taken) would be delayed jointly. This joint event is, however, irrelevant to the decision to be made. Savage's requirement may lead to large and intractable event and gamble spaces. Further, the resolution of joint uncertainties often is not even observable. For instance, if you had chosen to travel by airplane, then you could never fully learn about the delays of the car trip, which did not even take place.

Luce developed various conditional decision models to avoid the aforementioned drawbacks. In the lead example, one then only considers the uncertainties relevant to (conditioned on) each transportation alternative separately, with no need to inspect the irrelevant joint uncertainties. As pointed out by Luce and others, these models, while avoiding some problems, create other problems. As we will argue in further detail later, one drawback of Luce's models is that they do not have Savage's clean separation between chance (in nature) and the human will of the decision maker. In Luce's models, both the decision maker and nature may choose (conditioning) events to happen. Another drawback is that part of the mathematical elegance of Savage's model is lost (pointed out by Luce 2000 p. 7 and discussed below).

This paper reconciles Savage's and Luce's models, with the aforementioned problems solved and the best of both worlds preserved. For this purpose, we propose a generalization of Savage's

[2] This is Luce's term. Savage (1954) used the term act. We use Luce's (2000) terminology as much as possible.

[3] The exact quote is from Luce (2000 §1.1.6.1). During his childhood, Luce was much interested in airplanes (besides painting), and he majored in aeronautical engineering. His parents advised against an art career, and astigmatism ruled out military flying, so that he turned to academic research. This history may have contributed to the adoption of this example. Luce used the example also in Krantz et al. (1971 §8.2.1) and Luce & Krantz (1971 §2).

model, based on Kopylov's (2007)[4] mosaics. Mosaics relax the intersection-closedness requirement of algebras, which is the cause of the aforementioned problems in Savage's model. Using mosaics we can model Luce's lead example without considering irrelevant and inconceivable combinations of uncertainties. At the same time, we maintain Savage's mathematical elegance and his clear separation of nature's influence and the decision maker's influence. We will show that for every Luce (2000) model there exists an isomorphic Savage model, which implies that this isomorphic model can capture all structures and phenomena that Luce's model can, and it can do so in the same tractable manner. In addition, our model satisfies all principles of Savage's model: One state space captures all uncertainties, and the moves of nature and the decision maker are completely separated. In this sense, our model has the best of both worlds.

Our result shows the usefulness of mosaics. The main conclusion of this paper, entailing a blend of Savage's and Luce's ideas, extends beyond the reconciliation obtained. We recommend the use and study of mosaics rather than (σ-)algebras as the event spaces for decision under uncertainty in general. This raises a research question: To what extent can the appealing and useful mathematical results obtained for algebras in the literature be generalized to mosaics? Kopylov (2007) and Abdellaoui & Wakker (2005) provided several positive results.[5]

This paper is organized as follows. Section 6.2 discusses Savage's (1954) model and §6.3 discusses Luce's (2000) model, the most comprehensive account of his views. Our reconciliation of these two models is in §6.4. Section 6.5 overviews some other deviations from Savage's model, including Luce & Krantz (1971), which contributed to and preceded Luce (2000). Unlike Luce (2000), the models considered there are not isomorphic to (a generalization of) Savage's (1954) models, but we show that they can still be embedded (i.e., are isomorphic to substructures) . Thus we show for all models considered how they can be related to the revealed preference paradigm of economics. Bradley (2007) provides a general logical model that can embed all models considered in this paper as substructures. Section 6.6 presents a discussion and §6.7 concludes. The appendix discusses some other generalizations of Savage's model that Luce considered, being compounding, coalescing, and joint receipts, which are tangential to our main topic: connecting Luce's uncertainty model with other uncertainty models popular in the

[4] In 2007, Kopylov worked at the economics department of the University of California at Irvine, within a mile of Luce's office who was at the psychology department there. Yet Kopylov developed his idea independently of Luce's work.

[5] Kopylov (2007), while working from a different motivation (see below), in fact already gave a positive answer for Savage's (1954) foundation of expected utility, by extending it to mosaics. Abdellaoui & Wakker (2005; written after and building on Kopylov's paper) provided further generalizations and preference foundations for a number of popular nonexpected utility theories for risk and uncertainty: the Quiggin (1982)-Gilboa (1987)-Schmeidler (1989) rank-dependent utility (including Choquet expected utility), Tversky & Kahneman's (1992) prospect theory (which applies not only to risk but also to ambiguity), and Machina & Schmeidler's (1992) probabilistic sophistication. Generalizations of other models of uncertainty to mosaics is a topic for future research, as is the extension of measure-theoretic concepts to mosaics.

literature today. Our connection allows the introduction of Luce's techniques, including those in the appendix and follow-up papers[6], into modern decision theories.

6.2 Savage (1954)

This section reviews Savage's (1954) model. Savage models uncertainty through a state space S. One state $s \in S$ is true and the other states are not true, but it is uncertain which state is the true one. S is endowed with an algebra $\mathscr{E}$ of subsets called *events*.[7] An *algebra* contains S and is closed under union and complementation. It follows from elementary manipulations that an algebra also contains $\emptyset$ and is closed under finite unions and intersections. An event is *true* if it contains the true state of nature. $\mathscr{C}$ is a set of *consequences*; it can be finite or infinite. A decision maker has to choose between *gambles* (generic symbol G), which are mappings from $S$to $\mathscr{C}$ with finite image[8] that are measurable with respect to $\mathscr{E}$. *Measurability* of G means that for each consequence x its inverse under G, $G^{-1}(x)$, is an event. It implies that $G^{-1}(x)$ is an event for every subset $D \subset \mathscr{C}$: $G^{-1}(D)$ is a finite union of events $G^{-1}(x)$ of the elements $x \in D$, where only finitely many of these events $G^{-1}(x)$ are nonempty.

The decision maker's comparisons between gambles constitute a preference relation $\succcurlyeq$. Some approaches do not take states and consequences as primitives, with gambles derived, but take gambles and consequences, or gambles and states, as primitives (Fishburn 1981 §8.4; Karni 2006, 2013) . Yet these approaches can be recast in terms of the original Savage model for the purposes of this paper (Schmeidler & Wakker 1987).

Savage gave a preference foundation for expected utility theory:

$$G \to \int_S U(G(s))dp(s). \tag{6.1}$$

Here $U : \mathscr{C} \to \mathbb{R}$ is a *utility function*, and P is a *probability measure* defined on the events. This paper does not discuss which particular decision theory (such as expected utility theory, prospect theory, multiple priors, and so on) is to be used. Its topic concerns the general modelling of uncertainty.

As regards Savage's drawback of involving a complicated event space, we not only have to

[6] References include Liu (2003), Luce (2010), Luce & Marley (2005), Marley & Luce (2005), and Marley, Luce, & Kocsis (2008).

[7] In the main analysis Savage (1954) assumed that $\mathscr{E}$ is the power set, but he pointed out that it suffices that it is a σ-algebra (§3.4, pp. 42-43) . His preference conditions, especially his P6, imply that S is infinite. Technical aspects such as the difference between σ-algebras and algebras are not important in this paper and we keep these aspects as simple as possible.

[8] We We throughout make this assumption, common in decision theory and made throughout Luce (2000; see his p. 3), to simplify the mathematics.

Table 6.1: Decisions impacting states of nature.

	s_1 :lung cancer	s_2 : no lung cancer
G_1 smoke	Pleasant life, then disease	Pleasant life, healthy
G_2 don't smoke	Unpleasant life, then disease	Unpleasant life, healthy

specify all joint uncertainties but also have to posit axioms sufficiently wide-ranging to generate all likelihoods. Then, further, all gambles whose consequences are contingent on the complicated event space have to be considered. This drawback was elaborated by Luce (2000, p. 6): It is certainly not unreasonable to suppose that each mode of travel entails, as a bare minimum, at least 10 distinct [uncertain events][9]. To place this simple decision situation in the Savage framework we must set $S = A \times B \times C \times T$,[10] and so there are at least 10,000 states of nature. Make the problem a bit more complex and it is easy to see that millions or billions of states must be contemplated. I think very few of us are able or willing to structure decisions in this fashion. Rather, we contemplate each of the alternatives as something unitary. The 10,000 states of nature calculated by Luce involve 9,999 subjective probabilities under subjective expected utility theory, and 210,000 ? 2 nonadditive weights under Luce's rank- and sign-dependent theory.[11] Hence Savage's model quickly becomes intractable for empirical applications.

A crucial assumption in Savage's model is that the decision maker does not have any influence on which state is true (Fishburn 1981 §2.2) . A classical example showing what goes wrong if this assumption is violated is as follows. For two states $S = \{s_1, s_2\}$ and two gambles G_1, G_2, Table 7.1 displays the consequences.

By Savage's expected utility, and even just by dominance, the only rational choice seems to be G_1: "smoke," giving the better consequence and more utility in every state. This analysis obviously misses an essential point: The choice of the decision maker impacts the probability of the state of nature, making G_2: "don't smoke" a good choice. Hence, Savage's model, and the dominance principle, cannot be used in Table 7.1.

In applications, it may be hard to fully satisfy Savage's requirement of a strict separation of the decision maker's and nature's moves. A proper definition of states of nature, specifying the relevant uncertainties beyond control of the decision maker is not easy to achieve in the example of Table 7.1. It is difficult in many applications, e.g. when there is uncertainty about whether you pass an exam next week, whether your next working paper will be accepted by the first outlet

[9] Luce instead used the term outcome. This term commonly refers to uncertain events (states of nature) in probability theory, a convention followed by Luce. In decision theory, however, the term outcome commonly refers to consequences rather than events. To avoid confusion, we do not use this term.

[10] A: airplane; B: bus; C: car; T: train. In his book, Luce used Ω instead of S to denote the universal set.

[11] For simplicity we focus on gains. If there are losses and sign-dependence is allowed, then the number of required calculations doubles for all cases considered in this paper.

tried, whether your new company will survive the first year, whether you will be healthy in five years from now, and so on. Studies of moral hazard in economics (Zeckhauser 1970; Mas-Colell et al. 1995), and of causal decision theory in philosophy (Joyce 1999[12]), do allow for influence on states as an add-on to Savage's model, generalizing it. We will discuss (Sections 6.3 and 6.5) more fundamental deviations from Savage's model. They seek to incorporate influences of the decision maker on states and conditioning events so as to simplify Savage' model rather than to generalize it. We start with Luce (2000).

6.3 Luce (2000)

This section presents Luce's (2000) model of uncertainty, the most comprehensive and most up-to-date version of Luce's views. For a book review of Luce (2000), see Bleichrodt (2001) . An historical development of Luce's ideas and some important historical influences are presented in §6.5. Luce (2000) introduced a new primitive: (chance) experiments in a two-stage decision process. First, there is a set $(E^i)_{i \in I}$ of *chance experiments*, with I an index set. Following Luce (2000), we write the indexes as superscripts. $\mathscr{C}$ denotes a set of consequences. The decision maker first chooses an experiment from some available experiments, and then chooses one of the gambles related to the chosen experiment. One experiment may be traveling by airplane, and another experiment may be traveling by car. Each chance experiment E has its own *universal set* Ω_E which, given that chance experiment, plays the same role as Savage's state space, specifying all uncertain events from there on. It is endowed with an algebra $\mathscr{E}_E$ of *events*. The chance experiment E is often identified with its universal set Ω_E. Given one experiment, Luce's model is like an unconditional Savage model. Once the experiment is chosen, there remain several gambles available conditional on the experiment. For example, if one chooses to travel by car, then one may still have to choose which commitments to take at one's destiny, with different pros and cons resulting that depend on travel delays of the car ride.

A gamble conditional on E is of the form $(E_1 : g_1, \ldots, E_n : g_n)$ where $(E_1, \ldots, E_n)$ are events partitioning Ω_E and consequence g_j is assigned to every element of E_j. Under expected utility theory, the gamble is evaluated by

$$\sum_{j=1}^{n} P(E_j)U(g_j) \tag{6.2}$$

with P a probability measure conditional on E and U utility as usual. These expected utility evaluations are also used for comparisons between different chance experiments. The central decision theory in Luce (2000) is a generalization of expected utility: rank- and sign-dependent utility, introduced by Luce & Fishburn (1991) . It is essentially equivalent to Tversky &

[12] Fishburn (1964 Chs. 2 and 3) similarly considered dependence of probabilities on gambles.

Kahneman's (1992) prospect theory. We focus on the general modelling of uncertainty, without commitment to any specific decision theory. Therefore, an explanation of details of decision theories is not needed for this paper.

Luce's model is more tractable for empirical analyses than Savage's. Consider the lead example where each mode of travel involves 10 distinct further uncertain events. Under Luce's model, the number of states of nature to be contemplated is $4 \times 10 = 40$ instead of $10,000$ as under Savage's model. Under subjective expected utility theory, for each mode of travel 9 probabilities are involved, totalling $4 \times 9 = 36$, considerably fewer than Savage's $9,999$. Under rank- and sign- dependent theory, $4 \times (2^{10} - 2)$ non additive weights are involved, considerably fewer that Savage's $2^{1000} - 2$.

One drawback of Luce's model is that it has a set of chance experiments without further structure or relations between them. This way of modelling is not mathematically elegant. To illustrate a concrete drawback, assume that a richness requirement is needed that the event space is a continuum, as in Savage (1954).[13] In Luce's model this then needs to be imposed on every chance experiment separately instead of only on one unifying state space.

Also valuing mathematical elegance, Luce (2000 p. 7) considered a step halfway in the direction of Savage's (1954) model. For this purpose, different chance experiments are interpreted as mutually exclusive events. Their union $\cup_{i \in I} \Omega_{E^i}$ (called *master experiment*) then is a state space similar to Savage's. However, Luce cautioned against this step conceptually. It leads to a conditional decision model, under which the decision maker first chooses a chance experiment, and then chooses a decision conditional on the experiment chosen. This entails choosing an event (subset of the state space) to come true. Luce (2000 p. 7) wrote critical comments about the danger of confusing nature's moves with the decision maker's moves: A choice of when and how to travel differs deeply from the statistical risks entailed by that choice. The potential for confusion in trying to treat them in a unitary fashion is so great that I eschew this perhaps mathematically more elegant approach. Luce's (2000) approach avoids the most serious problem of the earlier approach in Luce & Krantz's (1971), discussed further in Subsection 6.5.2. However, it still does not achieve the clear separation of influences that Savage did. By choosing a chance experiment the decision maker does impact what the true state of nature will be.

The next section presents an alternative to, and a complete reconciliation of Luce's and Savage's models. Both the mathematical elegance of Savage's model and the empirical tractability of Luce's model are obtained, as is Savage's clean separation between nature and human influence. A special case appeared in Abdellaoui & Wakker (2005 Example 5.4.v).

[13] His probability measure is atomless with full range [0,1] (Savage 1954 Theorem 3.3.3, item 7), which implies a continuum of events.

6.4 Reconciling Savage (1954) and Luce (2000)

In our model, as in Savage's (1954), S denotes the state space, $\mathscr{C}$ the consequence space, and $\mathscr{E}$ the collection of subsets of S called events. To resolve the problems of Savage's model, we relax intersection-closedness. The interest of relaxing intersection-closedness was pointed out before by Luce and his co-authors in Krantz et al. (1971 §5.4.1) and was based on findings of quantum mechanics. There a particle's speed and location can be known separately, but they cannot both be known exactly. Hence, Krantz et al. (1971 §5.4.1) recommended using Dynkin systems[14] to model such uncertainties. Dynkin systems generalize algebras by requiring closedness only for disjoint unions. Then intersection-closedness is no more implied.

An interest in relaxing intersection-closedness also arose recently in the theory of decision under ambiguity (uncertainty without objective or subjective probabilities), where the intersection of two unambiguous events need not be unambiguous. For example, consider an urn with 100 balls where each ball has a number 0 or 1 and a color red or green. Even if the proportion of the numbers and the proportion of the colors are known and unambiguous, their intersection may not be, if the correlation between number and color is unknown. Zhang (2002) therefore recommended taking the set of unambiguous events to be a Dynkin-system.

Kopylov (2007) proposed yet a further generalization and required the set of unambiguous events to be a mosaic (defined later) . He still required that the general domain of events $\mathscr{E}$ be an algebra. He used the mosaic structure only to analyze the subcollection of unambiguous events. We use mosaics for a different purpose. We require that § itself, the whole domain of events, is a mosaic rather than an algebra. We do so to solve the problems in Savage's model pointed out by Luce, while avoiding the problems of Luce's model.

DEFINITION 17. The collection $\mathscr{E}$ of subsets of the state space S is a *mosaic* if it satisfies the following conditions:

(i) It is complementation-closed;

(ii) It contains S;

(iii) For every finite partition $(E_1,\ldots,E_n)$ of S consisting of events (elements of $\mathscr{E}$), $\mathscr{E}$ contains all unions of E_j's. □

To see that mosaics are more general than Dynkin systems, consider two different partitions $(E_1,\ldots,E_n)$ and $(F_1,\ldots,F_n)$. A Dynkin system would still consider the union $E_i \cup F_j$ if $E_i \cap F_j = \emptyset$, but mosaics do not and, therefore, are more flexible for applications involving multiple chance

[14] Other terms are d-systems, QM-algebras, or λ-systems. Such collections play a role in the mathematics of probability theory when constructing probability measures on σ-algebras (Billingsley 1968).

experiments. *Gambles* G are again mappings from S to $\mathscr{C}$ with finite image that are measurable (for each consequence x, $G^{-1}(x)$ is an event) . It still implies that $G^{-1}(D)$ is an event for every subset $D \subset \mathscr{C}$, as $G^{-1}(D)$ is a union of events $\{G^{-1}(x) : x \in D \cap G(S)\}$ from a finite partition $G^{-1}(x) : x \in G(S)\}$. We next show how the flexibility of mosaics enables us to construct, for every Luce model, an isomorphic Savage model. We assume a Luce model with ${E^i}_{I \in I}$, Ω_{E^i}, $\mathscr{E}_{E^i}$, and $\mathscr{C}$ as in §6.3. We define the isomorphic Savage model as follows. We keep $\mathscr{C}$ as it is. We define the state space S of Savage not as the union of the E^is (as in Luce's 2000 master experiment), but instead as their product set $\prod_{i \in I} \Omega_{E^i} \Omega_{E^i}$. In this product there is a dimension for each chance experiment. A Savage-state $s \in S = \prod_{i \in I} \Omega_{E^i} \Omega_{E^i}$ specifies a Luce-state $s^i \in E^i$ for each chance experiment E^i. It may seem at this stage that we still consider joint resolutions of uncertainty as Savage did, because our states refer to such joint resolutions. However, states are not very relevant because they need not be evaluated. Events are relevant for this purpose, and they will not reflect joint uncertainties, as explained next.

The domain $\mathscr{E}$ of *events* in our isomorphic Savage model is defined as $\cup_{j \in I} (\mathscr{C}_{E^j} \times \prod_{i \neq j} \Omega_{E^i}) = \{A := A^j \times \prod_{i \neq j} \Omega_{E^i}, A^j \in \mathscr{E}_{E^j}\}$. Thus an event specifies what happens for one chance experiment E^j, leaving the events for all other chance experiments unspecified. In measure theory, such sets are called *cylinder sets*.

It is well-known that $\mathscr{E}$ must be extended, with many sets added, to turn it into an algebra or a σ-algebra. This extension brings many artificial extra gambles, requiring many extra considerations that are practically irrelevant, and this lies at the heart of Savage's problem. For mosaics, however, no extension is needed, because the collection of cylinders itself is a mosaic. This follows mainly because for a partition of S consisting of events, all those events $A_j \times \prod_{i \neq j} \Omega_{E^i}, A_j \in \mathscr{E}_{E^j}$ so as to be disjoint, must concern the same chance experiment E^j. Their unions are again cylinder sets concerning that same chance experiment and, hence, are again contained in $\mathscr{E}$.

Measurability of a gamble G implies, informally speaking, that G depends on only one chance experiment E^j. To see this point, assume that $\{x_1, \ldots, x_n)\}$ is the image of G. Then the sets $G^{-1}(x_j)$ constitute a partition of S. Hence, as we saw before, these sets all concern the same chance experiment E^j.[15] We call E^j the chance experiment *relevant* to gamble G. We now obtain a one-to-one correspondence between gambles G in Savage's model and gambles G_L in Luce's model. The chance experiment E^j relevant for the Savage gamble G is the domain of the

[15] To see this point, assume that $G^{-1}(x_1)$ refers to chance experiment E^i, imposing restrictions on the ith coordinate of the true state s. Assume, for contradiction, that $G^{-1}(x_2)$ refers to another chance experiment E^j $(j \neq i)$, imposing restrictions on the jth coordinate of the true state s. Then $G^{-1}(x_1) \cap G^{-1}(x_2) \neq \emptyset$, consisting of the states that satisfy the requirements for both coordinate i and coordinate j. This contradicts the disjointness of $G^{-1}(x_1)$ and $G^{-1}(x_2)$.

Luce gamble G_L, and we have

$$G^{-1}(x) = A^j \times \prod_{i \neq j} \Omega_{E^i} \Leftrightarrow G_L^{-1}(x) = A^j \tag{6.3}$$

Although our state space is big, this does not affect tractability because we only consider measurable acts and the event space is tractable. Single states, indeed, never need to be considered or evaluated. The isomorphism of our model with Luce's model shows that we have the same tractability. Consider again the lead example where each mode of travel entails 10 distinct further uncertain events. Our model involves 10000 status of nature, which is as many as Savage (1954) and way more than Luce's (2000) 36 states. Yet, this does not entail a greater complexity because our mosaic does not involve many events. The number of probabilities or weights involved is the same as in Luce's model: 36 probabilities for 36 cylinder sets under subjective expected utility theory, and $4 \times (2^1 0 - 2)$ nonadditive weights of cylinder sets under Luce's rank- and sign dependent theory. Hence, we have maintained the tractability of Luce's model.

We have re-established Savage's principle that nature's moves and the decision maker's moves are completely separated. The decision maker cannot in any way influence which state in S is true. She can influence which partition (chance experiment) $\{G^{-1}(x) : x \in \mathscr{C}\}$ is relevant to her, but this is always the case with every choice of gamble in Savage's model. For instance, choosing to fly by plane in economy class rather than (a) flying in business class or (b) traveling by train 1st class or (c) traveling by train 2nd class, consists of two stages in Luce's model: First, the airplane chance experiment is chosen. Second, economy class is chosen. In our Savage model it is only the one-stage choice of the act of flying economy class. The choice of this act automatically implies that the relevant chance experiment/partition describes the uncertainties about the airplane. In the same way, every choice of any act G in Savage's model always implies the "choice" of the relevant events $\{G^{-1}(x) : x \in \mathscr{C}\}$. Thus we have incorporated Luce's choice of chance events, not as influence on which subset of the universal event happens in a separate stage, but as a standard choice of act that fully maintains Savage's separation of choices of acts from choices of states of nature. Influencing the relevant partition in no way influences the true state of nature.

Using mosaics, we have solved a problem in Savage's model that Luce discussed throughout his career. We fully meet the flexibility required by Luce but at the same time maintain all principles of Savage's model and its elegance with one unifying state space. Using the described isomorphism, all results in Luce (2000) and many papers building on it can now be transposed to the currently popular Savage model.

6.5 Different ways to model uncertainty that preceded and influenced Luce (2000)

Luce was not alone in his search of an alternative to Savage's model. Different models had been proposed before, making different assumptions about where uncertainty comes from. Luce's (2000) model was the culmination of this development. This section discusses this history. We show that all models can be obtained as submodels of Savage (1954) in a mathematical sense, even if at first the concepts of those models seem to be very different.

Savage (1954) is at one extreme of the spectrum: The separation of the decision maker's control over the gamble chosen and nature's control over the state chosen is complete. At the other extreme is Jeffrey's (1965) model: He treats both gambles and events as propositions and uncertainty concerns the truth or falsity of propositions. Thus he does not differentiate at all between moves by the decision maker and moves by nature. Luce throughout took a middle ground between Savage and Jeffrey. In particular, he never recognized Savage's strict separation between the decision maker's and nature's moves.

We next reference some other discussions of models for decision under uncertainty. Fishburn (1981) is an extensive and impressive review of models up to 1981, providing many insights that are still relevant to the literature today. Bradley (2007) presented a general logical model that comprises the models of Savage (1954), Jeffrey (1965), and their intermediate Luce & Krantz (1971) (when interpreted as logical models) as submodels. Spohn (1977) also discussed the models developed up to that point, arguing for the desirability of maintaining Savage's separation.

6.5.1 Jeffrey (1965): the other extreme

Jeffrey (1965) took an approach fundamentally different from Savage. Bolker (1966, 1967) provided generalizations of Jeffrey's model. Whereas Savage distinguishes states and gambles completely, Jeffrey's model makes no such distinction at all, and even equates them. Both probabilities $P(A)$ and utilities $u(A)$ (called desirability by Jeffrey) are carried by propositions A, which we model as subsets of a space S.[16] To Jeffrey, beliefs and desires are just two sorts of attitude toward the same proposition, say A. We can thus assign a probability, but also a degree of desirability, to a proposition such as passing an exam next week. The interpretation of propositions can be very broad. In terms of Savage's model, propositions can refer to events, gambles, consequences, or their mix. A decision problem can thus be modelled using a single

[16] Jeffrey considered a Boolean algebra of propositions, endowed with logical operations. Such an algebra is isomorphic to an algebra of subsets of a set S (Stone 1936) . For the purposes of this paper, the latter is more convenient.

set of propositions. The characteristic condition in Jeffrey's decision theory, making it a relative to Savage's expected utility, is that, whenever $A \cap B = \emptyset$:

$$u(A \cup B) = \frac{P(A)u(A) + P(B)u(B)}{P(A) + P(B)} \tag{6.4}$$

where P denotes a probability measure on S. This formula suggests that $u(A)$ may play a role as a conditional expected utility. This suggestion will be formalized next.

Though Jeffrey's and Savage's models are two extremes of the philosophical spectrum, one can formally obtain every Jeffrey models as a substructure of a Savage models. To see this point, we first start from a Savage model. Take his state space S and consequence space $\mathscr{C}$. We assume expected utility theory with probabilities P and utilities U. Suppose that we want to measure the desirability of a gamble $G : S \to \mathscr{C}$ conditional on an event $A \subset S$, i.e., $\frac{\int_A U(G(s))dP}{P(A)}$. This can be inferred from the *conditional certainty equivalent* $c_{G_A} \in \mathscr{C}$ of G given A, defined by

$$c_{G_A} \textit{ is such that } G' \sim G \textit{ if} : G'(s) = G(s) \textit{ for } s \in S \setminus A \textit{ and } G'(s) \equiv c_{G_A} \textit{ for } s \in A. \tag{6.5}$$

Based on expected utility theory,
$EU(G) = \int_S U(G(s))dP = \int_{S \backslash A} U(G(s))dP + \int_A U(G(s))dP$
$EU(G') = \int_{S \backslash A} U(G'(s))dP + \int_A U(G'(s))dP = \int_{S \backslash A} U(G(s))dP + U(c_{G_A}) \times P(A)$.
Since $G \sim G'$, we have $EU(G) = EU(G')$, resulting in $U(c_{G_A}) = \frac{\int_A U(G(s))dP}{P(A)}$, which can be defined as the desirability of G conditional on event A. Using the same method and keeping G fixed, the desirability of other events ("propositions") can be calculated, which makes the comparison between propositions possible. For example, we can now compare a decision maker's desirability of "walking outside in the rain tomorrow" versus "walking outside in strong wind tomorrow."

Formally, for each fixed gamble G, ordering events A by their conditional certainty equivalents $u(A) = U(c_{G_A})$ yields a Jeffrey structure. Conversely, under some richness, every Jeffrey structure can be identified with such a Savage substructure.[17] This way, Jeffrey preferences between events A become observable from Savagean preferences. In particular, Jeffrey's desirability can now be related to the decision-based revealed preference paradigm prevailing in economics. And hence, at least in a mathematical sense, a Jeffrey model can be interpreted as a conditional certainty equivalent submodel of a Savage model with one gamble fixed. Under revealed preference interpretations of desirability, Jeffrey's model is almost dual to Savage's in the sense that the fixed gamble G has been decided outside the control of the decision maker but now the decision

[17] See Fishburn (1981 p. 186) . Assume countable additivity. Then the measure $P(A)U(A)$ is absolutely continuous with respect to P. With u its Radon-Nikodym derivative, we take U and G in Savage's model such that $u(s) = U(G(s))$ for all s.

maker seems to be choosing between events.

Jeffrey's model can be reinterpretred as an extreme form of state-dependent expected utility.[18] Now utility not only depends on the event considered, but is entirely determined by it. There have been many debates about the identifiability of probabilities under state-dependent utility (Karni 2003). Drèze (1987 Ch. 2) showed that probabilities then can become observable if the decision maker has some influence on the states of nature. Jeffrey's model is in this spirit.

Ramsey's (1931) famous analysis may be between Jeffrey's and Savage's. It never clearly specifies whether events and consequences can be the same or should be separated. Bradley (2004) provides a detailed analysis of Ramsey (1931) . We now turn to another model in between Jeffrey's and Savage's.

6.5.2 Luce & Krantz's (1971) conditional decision model: the middle ground

Inspired by the lead example (see our introduction), Luce & Krantz (1971) developed a conditional decision model.[19] It consists of a set $\mathscr{C}$ of consequences and a state space S with an algebra $\mathscr{E}$ of events $E \subset S$,[20] as in Savage's model. Different from gambles in the unconditional Savage setting, which map the entire space S to $\mathscr{C}$, here *conditional gambles* G_A are considered, which map subsets A of S to $\mathscr{C}$. That is, they are restrictions of gambles G to events A. For example, A can designate the event, decided upon by you, of you traveling by car and G_A then specifies the further uncertainties and results of the travel, laying down a consequence $G(s)$ for each state $s \in A$. If B designates traveling by airplane, then the preference $G_A \succcurlyeq G'_B$ indicates that traveling by car with the consequences specified by G is preferred to traveling by airplane with the consequences specified by G'.

The conditional decision structure synthesizes Savage and Jeffrey. It allows an event, such as A (traveling by car) to be under the control of the decision maker. A choice between G_A and G'_B determines not only the contingencies with which various consequences arise, but also which conditioning event (A or B) will occur. Hence, while Savage's S can be interpreted as one universe, here events can constitute parallel universes, the realization of which depends on the choice of the decision maker. For an event A different conditional decisions G_A, G'_A are conceivable, and this is one way in which the model is richer than Jeffrey's.

[18] This relation was suggested by a referee.

[19] Krantz et al. (1971 Ch. 8) present it with some modifications and extensions.

[20] Null events cannot play the role of conditioning events in what follows. To avoid technicalities, we do not consider null events.

For disjoint events A, B and conditional decisions G_A and G'_B, their union

$$G_a \cup G'_B \tag{6.6}$$

denotes the function restricted to $A \cup B$ and equal to G on A and G' on B. We postpone discussing the problematic interpretation of such unions until the end of this section, and now proceed with the characteristic condition of Luce & Krantz's (1971) decision theory:

$$u(G_A \cup G'_B) = \frac{P(A)u(G_A) + P(B)u(G'_B)}{P(A) + P(B)}. \tag{6.7}$$

Again, P denotes a probability measure on S. As in Jeffrey's model, the condition suggests that $u(G_A)$ plays a role as conditional expected utility. This indeed follows because, under some natural conditions (Luce & Krantz 1971 §4):

$$u(g_A) = \frac{\int_A U(G(s))dP}{P(A)} \tag{6.8}$$

Fishburn (1981 §8.4) reviews related follow-up studies. For example, Balch & Fishburn (1974) considers a combination of conditional gambles with gamble-dependent probabilities.

We next show that for every Savage model, we can consider a Luce & Krantz substructure: For each G_A we construct the conditional certainty equivalent c_{G_A} as in Eq. 6.5. Through conditional certainty equivalents, Luce & Krantz preferences between conditional gambles G_A become observable from Savagean preferences.

Conversely, for every Luce & Krantz model we can obtain a Savage substructure, simply because all primitives of Savage's model have been provided, with unconditional Savagean preferences equated with Luce & Krantz preferences conditioned on the universal event. This "unconditional" subpart of the Luce & Krantz model indeed completely determines U and P, and thus not only the whole Savagean model but also the whole Luce & Krantz model with all its conditional preferences. The Luce & Krantz model can be interpreted as a conditional certainty equivalent submodel of a Savage model but now not with one fixed gamble G considered, as in Jeffrey's model, but with various gambles $G, G', \ldots$ considered.

We finally turn to interpretations of a union $G_A \cup G'_B$ of conditional acts for disjoint events A and B. Whether A or B happens is determined by chance here, with A happening with its probability determined through P, in $\frac{P(A)}{P(A)+P(B)}$. In a choice betweenG_A and G'_B, however, it is under the control of the decision maker whether A or B happens. We agree with Spohn (1977) that this double interpretation of events is hard to conceive of.

6.5.3 From Luce & Krantz (1971) to Luce (2000)

Possibly Luce (2000) came to agree with Spohn (1977) and later works that the double role of events in Luce & Krantz's (1971) approach (Eq. 6.6), with both nature and the decision maker choosing between them, is hard to conceive of. Although not explicitly referring to such unions, Luce's citation at the end of our §6.3, criticizing confusions of nature's and the decision maker's moves, suggests so. His model (Luce 2000), both with and without a master experiment, does not have events that are both under control of nature and the decision maker. Yet Luce (2000) does not achieve the clear separation of influences that Savage did and does involve conditional decisions.

There was yet another reason that probably led Luce to abandon unions of conditional gambles as in Eq. 6.6. Since the 1980s an interest has arisen in theories deviating from expected utility, the most popular one being prospect theory (Tversky & Kahneman 1992) . It has not been very widely known that Luce & Fishburn (1991) essentially introduced the same theory called rank- and sign-dependent utility. This theory is central in Luce (2000) . Once the realm of expected utility is left, it is no longer clear how to update or weigh unions. There then is no undisputed analogue of Eq. 6.7 (Denneberg 1994; Dominiak, Duersch, & Lefort 2012; Miranda & Montes 2015) . This complication will have added to Luce's decision to abandon unions of conditional gambles.

6.6 Discussion

Although our Savage model in §6.4 is isomorphic to Luce's model and can accommodate the same empirical phenomena, we note some formal differences. In Luce's model, different chance experiments are simply different unrelated entities and there is no formal unifying structure. In the step halfway to Savage, using the master experiment and described in §6.3, different chance events would be disjoint subevents of one encompassing master experiment, but Luce does not commit to such a formalization. His claim "the events of one experiment never appear in the formulation of a different experiment" (p. 6) has no clear formal meaning. It does suggest that the universal sets Ω_{Ej} of all possibilities in a chance experiment are differerent, and may be mutually exclusive, for different chance experiments. In our model they are all the same, being the state space S (the universal statement that is always true) . For us, different chance experiments refer to different partitions of the same state space, with that state space reflecting one universal truth.

The term Savage model as used in this paper is defined in §6.2, with states, acts, consequences, events, and a complete preference relation. A technical difference with Savage's (1954) decision theory is that we did not impose his richness conditions, mainly his P6. They imply that S is infinite for instance, whereas we allow finite S. Another difference is that his set of events is an algebra and even a σ-algebra, which is essentially used in his analyses. Hence his preference foundation cannot be directly used, for instance. Here Kopylov's generalization is useful.

In our Savage models that are isomorphic to Luces models, two events $A^i \in \Omega_{E^i}$ and $B^j \in \Omega_{E^j}$ of different chance experiments have a nonempty intersection $A^i \times B^j \times \prod_{k \neq i,j} \Omega_{E^k}$. The true state of nature may be such that both events are true. However, this intersection is no event: it is not contained in the mosaic. It then is not relevant to any decision to be made. We are not interested in it, and need never evaluate, analyze, or even think about it.

In natural language, different choices of gambles, with different partitions of S relevant, can be called different events. In the formal language of decision theory and probability theory, as in Savage's model, such terminology is not possible. Events only refer to subsets of the state space. That is, they only describe moves by nature. Luce never followed this strict separation of nature's moves from decision maker's moves in Savage's model.

In our Savage models isomorphic to Luce's models, any event (except $\emptyset$ and S) appearing in one chance experiment does not appear in any other, in agreement with Luce's assumptions. But this need not hold for general mosaics. In general it may well happen that an event A occurs in two partitions $\{A, B_2, \ldots, B_n\}$ and $\{A, C_2, \ldots, C_m\}$ that have nothing other in common. Here, intersections $B_i \cap C_j$ of other events of the chance experiments may not be events; i.e., they may not be contained in the mosaic. In the lead example, event A could entail that an earthquake destroys the route to the destination, preventing all modes of transportation.[21] We prefer mosaics to Dynkin systems. That is, in the preceding paragraph we prefer not to speculate on whether $B_i \cap C_j = \emptyset$ or on committing to then having $B_i \cup C_j$ as an event (that should be evaluated according to the decision theory considered) . In this regard mosaics better capture a desirable feature of Luce's chance experiments than Dynkin systems do. We consider this to be a desirable extra flexibility of mosaics.

There is a renewed interest in the foundations of uncertainty in the modern literature on uncertainty. Ellsberg (1961) showed that, surely from a descriptive perspective, we often cannot assign traditional probabilities to uncertain events, calling for generalizations ("ambiguity") of Savage's expected utility formula. After initiating work by Gilboa (1987), Schmeidler (1989), and Gilboa & Schmeidler (1989), many ambiguity theories have been developed, including Luce & Fishburn's (1991) rank- and sign-dependent utility. This may have also contributed to Luce's (2000) decision to modify Luce & Krantz's (1971) uncertainty model.

Although many modern nonexpected utility theories stayed within Savage's general uncertainty model[22], there also have been more fundamental deviations. Those involved uncertainties

[21] Such an event is plausible in Luce's domicile, Irvine in California, both in 1971 and in 2000.

[22] For the purposes of this paper, the Anscombe-Aumann (1963) model, popular in modern decision theory, belongs to Savage's general uncertainty model, with a state space and a clear separation between moves by the decision maker and moves by nature.

about the state space, as in unforeseen contingencies (Ahn & Ergin 2010; Dekel, Lipman, & Rustichini 1998; Karni & Viero 2013) and case-based decisions (Gilboa & Schmeidler 2001) . Ahn (2008) proposed an ambiguity model that used Jeffrey's (1965) uncertainty model, also deviating from Savage's general model. Ahn & Ergin (2010) incorporated Luce's partition dependence (see our appendix), unaware of Luce's precedence. Our analysis has shown how Luce's models and theories can be invoked in uncertainty models commonly used today. The many results and tools provided in Luce's book can now be used in modern studies, and new researchers can better become aware of Luce's precedences.

6.7 Conclusion

We have shown how various models of uncertainty can be embedded into the most well-known model today: Savage's (1954) state space model. Thus we have shown how all these models can be related to the decision-based revealed preference paradigm of economics. In particular, we showed that Luce's most comprehensive model, in Luce (2000), not only can be embedded, but even can be related isomorphically to Savage's model. To this effect, we needed a version of Savage's model more general than used by Savage (1954), taking the collection of events as a mosaic rather than as an algebra. This way we avoided the overly restrictive intersection-closedness requirement of algebras. We can now handle parallel uncertainties without having to consider their joint resolutions. The results, techniques, and empirical flexibility of Luce (2000) and the follow-up works by him and his colleagues now become available to modern decision theories.

The main recommendation of our paper is that the literature on decision under uncertainty use and study mosaics rather than the common algebras. This way, more tractable models that are better suited for applications result. It raises a research question for future studies: How can existing results for algebras be generalized to mosaics?

References

Abdellaoui, Mohammed (2000) "Parameter-Free Elicitation of Utility and Probability Weighting Functions," *Management Science* 46, 1497–1512.

Abdellaoui, Mohammed & Peter P. Wakker (2005) "The Likelihood Method for Decision under Uncertainty," *Theory and Decision* 58, 3–76.

Ahn, David S. (2008) "Ambiguity Without a State Space," *Review of Economic Studies* 75, 3–28.

Ahn, David & Haluk Ergin (2010) "Framing Contingencies," *Econometrica* 78, 655–695.

Anscombe, Francis J. & Robert J. Aumann (1963) "A Definition of Subjective Probability," *Annals of Mathematical Statistics* 34, 199–205.

Balch, Michael & Peter C. Fishburn (1974) "Subjective Expected Utility for Conditional Primitives." *In* Michael S. Balch, Daniel L. McFadden, & Shih-Yen Wu (eds.) *Essays on Economic Behaviour under Uncertainty,* 57–69, North-Holland, Amsterdam.

Billingsley, Patrick (1968) "*Convergence of Probability Measures.*" Wiley, New York.

Bleichrodt, Han (2001) "Utility of Gains and Losses by R. Duncan Luce," *Journal of Economic Literature* 39, 130–131.

Bleichrodt, Han & José Luis Pinto (2000) "A Parameter-Free Elicitation of the Probability Weighting Function in Medical Decision Making," *Management Science* 46, 1485–1496.

Bolker, Ethan D. (1966) "Functions Resembling Quotients of Measures," *Transactions of the American Mathematical Society* 124, 292–312.

Bolker, Ethan D. (1967) "A Simultaneous Axiomatization of Utility and Subjective Probability," *Philosophy of Science* 34, 333–340.

Bradley, Richard (2004) "Ramsey's Representation Theorem," *Dialectica* 58, 483–497.

Bradley, Richard (2007) "A Unified Bayesian Decision Theory," *Theory and Decision* 63, 233–263.

Dekel, Eddie, Barton L. Lipman, & Aldo Rustichini (1998) "Recent Developments in Modeling Unforeseen Contingencies," *European Economic Review* 42, 523–542.

Denneberg, Dieter (1994) "*Non-Additive Measure and Integral.*" Kluwer Academic Publishers, Dordrecht.

Dominiak, Adam, Peter Duersch, & Jean-Philippe Lefort (2012) "A Dynamic Ellsberg Urn Experiment," *Games and Economic Behavior* 75, 625–638.

Drèze, Jacques H. (1987) "*Essays on Economic Decision under Uncertainty.*" Cambridge University Press, London.

Ellsberg, Daniel (1961) "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics* 75, 643–669.

Etner, Johanna, Meglena Jeleva, & Jean-Marc Tallon (2012) "Decision Theory under Ambiguity," *Journal of Economic Surveys* 26, 234–270.

Fishburn, Peter C. (1964) "*Decision and Value Theory.*" Wiley, New York.

Fishburn, Peter C. (1981) "Subjective Expected Utility: A Review of Normative Theories," *Theory and Decision* 13, 139–199.

Gilboa, Itzhak (1987) "Expected Utility with Purely Subjective Non-Additive Probabilities," *Journal of Mathematical Economics* 16, 65–88.

Gilboa, Itzhak & Massimo Marinacci (2013) "Ambiguity and the Bayesian Paradigm." *In* Daron Acemoglu, Manuel Arellano, & Eddie Dekel (eds.) *Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress of the Econometric Society* Vol. 1 Ch. 7, Cambridge University Press, Cambridge UK.

Gilboa, Itzhak & David Schmeidler (1989) "Maxmin Expected Utility with a Non-Unique Prior," *Journal of Mathematical Economics* 18, 141–153.

Gilboa, Itzhak & David Schmeidler (2001) "*A Theory of Case-Based Decisions.*" Cambridge University Press, Cambridge.

Jeffrey, Richard C. (1965) "*The Logic of Decision.*" McGraw-Hill, New York. (Second edition 1983, University of Chicago Press, Chicago.)

Joyce, James M. (1999) "*The Foundations of Causal Decision Theory.*" Cambridge University Press, Cambridge.

Karni, Edi (2003) "On the Representation of Beliefs by Probabilities," *Journal of Risk and Uncertainty* 26, 17–38.

Karni, Edi (2006) "Subjective Expected Utility Theory without States of the World," *Journal of Mathematical Economics* 42, 325–342.

Karni, Edi (2013) "Bayesian Decision Making with Action-Dependent Probabilities and Risk Attitudes," *Economic Theory* 53, 335–356.

Karni, Edi & Marie-Louise Vierø(2013) "Reverse Bayesianism: A Choice-Based Theory of Growing Awareness," *American Economic Review* 103, 2790–2810.

Keeney, Ralph L. & Howard Raiffa (1976) "*Decisions with Multiple Objectives.*" Wiley, New York. (Second edition 1993, Cambridge University Press, Cambridge).

Kopylov, Igor (2007) "Subjective Probabilities on "Small" Domains," *Journal of Economic Theory* 133, 236–265.

Krantz, David H., R. Duncan Luce, Patrick Suppes, & Amos Tversky (1971) "*Foundations of Measurement, Vol. I (Additive and Polynomial Representations).*" Academic Press, New York.

Liu, Liping (2003) "A Note on Luce-Fishburn Axiomatization of Rank-Dependent Utility," *Journal of Risk and Uncertainty* 28, 55–71.

Luce, R. Duncan (2000) "*Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches.*" Lawrence Erlbaum Publishers, London.

Luce, R. Duncan (2010) "Behavioral Assumptions for a Class of Utility Theories: A Program of Experiments," *Journal of Risk and Uncertainty* 40, 19–37.

Luce, R. Duncan & Peter C. Fishburn (1991) "Rank- and Sign-Dependent Linear Utility Models for Finite First-Order Gambles," *Journal of Risk and Uncertainty* 4, 29–59.

Luce, R. Duncan & David H. Krantz (1971) "Conditional Expected Utility," *Econometrica* 39, 253–271.

Luce, R. Duncan & Anthony A.J. Marley (2005) "Ranked Additive Utility Representations of Gambles: Old and New Axiomatizations," *Journal of Risk and Uncertainty* 30, 21–62.

Machina, Mark J. (1989) "Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty," *Journal of Economic Literature* 27, 1622–1688.

Machina, Mark J. (2014) "Ambiguity Aversion with Three or More Outcomes," *American Economic Review* 104, 3814–3840.

Machina, Mark J. & David Schmeidler (1992) "A More Robust Definition of Subjective Probability," *Econometrica* 60, 745–780.

Marley, Anthony A.J. & R. Duncan Luce (2005) "Independence Properties vis-à-vis Several Utility Representations," *Theory and Decision* 58, 77–143.

Marley, Anthony A.J., R. Duncan Luce, & Imre Kocsis (2008) "A Solution to a Problem Raised in Luce & Marley (2005)," *Journal of Mathematical Psychology* 52, 64–68.

Mas-Colell, Andreu, Michael D. Whinston, & Jerry R. Green (1995) "*Microeconomic Theory*." Oxford University Press, New York.

Miranda, Enrique & Ignacio Montes (2015) "Coherent Updating of Non-Additive Measures," *International Journal of Approximate Reasoning* 56, 159–177.

Prelec, Drazen (1998) "The Probability Weighting Function," *Econometrica* 66, 497–527.

Quiggin, John (1982) "A Theory of Anticipated Utility," *Journal of Economic Behaviour and Organization* 3, 323–343.

Ramsey, Frank P. (1931) "Truth and Probability." *In* "*The Foundations of Mathematics and other Logical Essays*," 156–198, Routledge and Kegan Paul, London. Reprinted *in* Henry E. Kyburg & Howard E. Smokler (1964, eds.) *Studies in Subjective Probability*, 61–92, Wiley, New York.

Savage, Leonard J. (1954) "*The Foundations of Statistics*." Wiley, New York. (Second edition 1972, Dover Publications, New York.)

Schmeidler, David (1989) "Subjective Probability and Expected Utility without Additivity," *Econometrica* 57, 571–587.

Schmeidler, David & Peter P. Wakker (1987) "Expected Utility and Mathematical Expectation." *In* John Eatwell, Murray Milgate, & Peter K. Newman (eds.) *The New Palgrave: A Dictionary of Economics*, Vol. 2, 229–232, The MacMillan Press, London.

Spohn, Wolfgang (1977) "Where Luce and Krantz Do Really Generalize Savage's Decision Model," *Erkenntnis* 11, 113–134.

Stone, Marshall H. (1936) "The Theory of Representation for Boolean algebras," *Transactions of the American Mathematical Society* 40, 37–111.

Thaler, Richard H. (1985) "Mental Accounting and Consumer Choice," *Marketing Science* 4, 199–214.

Tversky, Amos (1977) "Features of Similarity," *Psychological Review* 84, 327–352.

Tversky, Amos & Daniel Kahneman (1992) "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5, 297–323.

Wakker, Peter P. (1984) "Cardinal Coordinate Independence for Expected Utility," *Journal of Mathematical Psychology* 28, 110–117.

Wakker, Peter P. (2010) "*Prospect Theory for Risk and Ambiguity.*" Cambridge University Press, Cambridge, UK.

Zeckhauser, Richard J. (1970) "Medical Insurance: A Case Study of the Tradeoff between Risk Spreading and Appropriate Incentives," *Journal of Economic Theory* 2, 10–26.

Zhang, Jiankang (2002) "Subjective Ambiguity, Expected Utility and Choquet Expected Utility," *Economic Theory* 20, 159–181.

Summary

This dissertation has mainly focused on connections between theory and empirics in decision theory. The assumptions of differentiablity and continuity mostly have no empirical meaning, i.e., they are not observable, even if to obtain the common uniqueness results of the preference representations, we have to assume that preferences are continuous.

Chapter 2 has generalized the implications of convex preferences without presupposing differentiability or continuity of utility for decision models under uncertainty, and Chapter 3 has done the same for decision models under risk. Chapter 5 has shown that the logarithmic family is the only family of local proper scoring rule for three or more events, even without presupposing differentiability or continuity.

The axioms of revealed preference are directly based on choice data. Chapter 4 has proposed a testable axiom, Agreement, to characterize context-dependent preferences in a general framework. Moreover, Chapter 4 has provided empirical work (a lab experiment) to test the theory derived before. The last chapter has shown that Savage' framework and Luce's framework for decision making under uncertainty can be reconciled when the domain of events is relaxed to be a mosaic.

Nederlandse samenvatting

Deze dissertatie heeft zich voornamelijk gericht op banden tussen theorie en empirie in de besliskunde. De veronderstellingen van differentieerbaarheid en continuïteit hebben meestal geen empirische betekenis, dat wil zeggen ze zijn niet waarneembaar, zelfs als we moeten aannemen dat preferenties continu zijn om de gebruikelijke uniciteit van preferentie-representaties te krijgen.

Hoofdstuk 2 veralgemeniseert de gevolgen van convexiteit van preferenties zonder differentieerbaarheid of continuïteit van utiliteit te veronderstellen voor beslismodellen bij onzekerheid, en hoofdstuk 3 doet hetzelfde voor beslismodellen bij risico. Hoofdstuk 5 toont dat de logaritmische familie de enige familie is van lokale proper scoring rules voor drie of meer gebeurtenissen zelfs als we geen differentieerbaarheid of continuïteit vooronderstellen.

De axioma's van revealed preference zijn direct gebaseerd op observeerbare keuzen. Hoofdstuk 4 stelt een testbaar axioma voor, Agreement, om context-afhankelijke preferenties te karakteriseren in een algemeen raamwerk. Bovendien presenteert hoofdstuk 4 empirisch werk (een laboratorium experiment) om de eerder afgeleide theorie te testen. Het laatste hoofdstuk toont dat Savage's raamwerk en Luce's raamwerk for beslissen bij onzekerheid in overeenstemming kunnen worden gebracht wanneer het domein van gebeurtenissen wordt verruimd tot een mosaïek.