An axiomatic foundation of entropic preferences under Knightian uncertainty

Stefan Baumgärtner[a,b,*] and John-Oliver Engler[b]

[a] Department of Environment and Natural Resources,
University of Freiburg, Germany

[b] Department of Sustainability Science and Department of Economics,
Leuphana University of Lüneburg, Germany

January 9, 2018

Abstract: Decision-making about economy-environment systems is often characterized by deep uncertainties. We provide an axiomatic foundation of preferences over lotteries with known payoffs over known states of nature and unknown probabilities of these outcomes ("Knightian uncertainty"). We elaborate the fundamental idea that preferences over Knightian lotteries can be represented by an entropy function (*sensu* Lieb and Yngvason 1999) of these lotteries. Based on nine axioms on the preference relation and three assumptions on the set of lotteries, we show that there uniquely (up to linear-affine transformations) exists an additive and extensive real-valued function ("entropy function") that represents uncertainty preferences. It represents non-satiation and (constant) uncertainty aversion. As a concrete functional form, we propose a one-parameter function based on Rényi's (1961) generalized entropy. We show that the parameter captures the degree of uncertainty aversion. We illustrate our preference function with a simple decision problem and relate it to other decision rules under Knightian uncertainty (maximin, maximax, Hurwicz, Laplacian expected utility, minimum regret).

Keywords: axiomatic foundation, entropy, Knightian uncertainty, non-expected utility, preferences, Rényi-function

JEL-Classification: D81, H30

*Corresponding author: Chair of Environmental Economics and Resource Management, University of Freiburg, Tennenbacher Str. 4, D-79106 Freiburg, Germany, phone: +49 761 203-3753, email: stefan.baumgaertner@ere.uni-freiburg.de.

1 Introduction

Decision-making about economy-environment systems is often characterized by deep uncertainties. *Knightian uncertainty* denotes income lotteries with known payoffs over known states of nature, but unknown probabilities of these outcomes (Keynes 1921, Knight 1921).[1] It is a deeper form of not-knowing-the-future than risk (where probabilities of outcomes are known) or ambiguity (where people have some, possibly differing, beliefs about the likelihood of outcomes), but less deep than unawareness of payoffs (in some or all states) or unawareness of potential states of nature.

There are compelling reasons to care about Knightian uncertainty. For, it may be clear what are the potential outcomes of an action, but it may be outright impossible to assign probabilities to these outcomes. For example, the system generating the outcomes may be too complex[2] and the time horizon too long to warrant any reasonable probabilistic assessment. A relevant example is our planet's climate where we do not even fully understand every single part of the system yet, let alone all feedback loops contained (Mehta et al. 2009). As a matter of fact, recent climate predictions have been remarkably off (Fyfe, Gillett and Zwiers 2013). Moreover, the fundamental disagreement of expert groups on a certain issue alone, for whatever reason, might invoke situations of Knightian uncertainty (Feduzi and Runde 2011). In such cases, one might be tempted to attach subjective probabilities ("beliefs") to the scenarios, but there are catches: (1) existence of such probabilities cannot be guaranteed (Ellsberg 1961,[3] Halevy 2007), even when experts are asked for their educated guesses (Millner et al. 2013), and (2) espe-

[1] John Maynard Keynes and Frank Knight simultaneously coined the distinction between situations of *risk* and situations of *uncertainty* in 1921 (Keynes 1921, Knight 1921). The term 'Knightian uncertainty' has prevailed in the literature, though.

[2] Here, 'complex' refers to the system consisting of many interacting parts that are interconnected via multiple nonlinear feedback loops (cf. Sornette 2003).

[3] Ellsberg's anomaly or paradox refers to the following: Assume there is an urn that contains 120 balls in total, 40 of which are blue (b) and the other 80 yellow (y) and red (r), but with unknown color frequency ratio. Experiment participants are offered the following two bets: f_1 = 'win 10$ if the ball drawn from the urn is blue' or f_2 = 'win 10$ if the ball drawn from the urn is red' and f_3 = 'win 10$ if the ball drawn from the urn is either blue or yellow' or f_4 = 'win 10$ if the ball drawn from the urn is either red or yellow'. Ellsberg's finding was that the vast majority of participants preferred f_1 to f_2 and f_4 to f_3 which would imply for a probability measure underlying these choices that $P(b) > P(r)$ and $P(r) + P(y) > P(b) + P(y)$, a direct contradiction.

cially experts tend to be overconfident regarding their results (Alpert and Raiffa 1982), which introduces yet another intricacy. And, aside from all this, it seems also justified to ask – at least from time to time – whether any probability is really better than no probability.

A number of decision criteria have been suggested for situations of uncertainty, non-probabilistic as well as probabilistic, and there are apparent problems with both. First and foremost, the probabilistic ones like the "maxmin expected utility" approach (Gilboa and Schmeidler 1989), the "smooth ambiguity' ' model (Klibanoff, Marinacci and Mukerji 2005) or "variational preferences" model (Maccheroni, Marinacci and Rusticchini 2006) require probabilistic information which – as we have just argued – may be unavailable or unreliable. Second, these papers rationalize Ellsberg choices by incorporating an axiom of "ambiguity aversion" into their framework. Yet, it is doubtable whether Ellsberg choices are a desirable feature of a theory of rational decision making, because they imply things such as aversion to information, updating of information based on taste or sensitivity to sunk costs (Al-Najjar and Weinstein 2009). Moreover, as Halevy (2007) has shown in his experimental re-examination of Ellsberg's findings, whether a test person expresses ambiguity aversion is correlated with that person's incapability to apply basic probability calculus. Hence, while these models are successful from a descriptive point of view, they are normatively unsatisfactory. Non-probabilistic models, which seem more in line with what Keynes and Knight had in mind, naturally tend to be minimalistic in terms of what information they take into account. Clearly, they do not rely on anything like probabilities. But often, they also do not use all available information on all possible states of the world. For example, the maximin criterion (Wald 1949) only focuses on the worst outcome and evaluates actions accordingly, the Hurwicz rule (Arrow and Hurwicz 1977) evaluates actions according to weighted average of worst and best possible outcome, clearly an unsatisfying limitation as Gravel, Marchant and Sen (2012) have pointed out. Other rules like the principles of minimum regret (Niehans 1948, Savage 1954) and insufficient reason (Laplace 1820) take all possible states of the world into account, but they lack a formal concept of "uncertainty aversion" or, more fundamentally, a concept of "the degree of uncertainty".

In this paper, we propose and elaborate the fundamental idea that preferences over Knightian lotteries can be represented by an entropy function (sensu Lieb and Yngvason 1999) of these lotteries. In the spirit of Knight (1921), we do not refer to any concept of probability. Instead, we base our argument on Knightian income lotteries which are distributions of monetary payoff over different possible outcomes with unknown probabilities. Based on seven axioms on the preference relation over Knightian acts, we show that there uniquely (up to linear-affine transformations) exists an additive and extensive function ("entropy function") from the set of Knightian lotteries to the real numbers that represents uncertainty preferences. Unlike most approaches so far, we do not relax the Sure Thing Principle (Savage 1954).[4] Instead we show that a Knightian version of it follows from the our basic axioms. We also show that convex preferences are represented by a concave entropy function, which represents uncertainty aversion.

As an example of an entropic preference function under Knightian uncertainty, we propose a one-parameter function based on Rényi's (1961) generalized entropy. The parameter in Rényi's function can be interpreted as the relative weight at which the two fundamental sources of uncertainty are taken into account in the aggregate measure of uncertainty: (1) the pure number of potential states of nature, and (2) the heterogeneity of the payoff-distribution over the given number of states of nature.

We illustrate our preference function with a simple decision problem and relate it to existing decision rules under uncertainty (maximin, maximax, Hurwicz, risk-neutral and risk-averse Laplacian expected utility, minimum regret).

The paper is organized as follows. In Section 3, we axiomatically characterize the uncertainty preference relation and show that it can be represented by an entropic preference function. In Section 4, we discuss this function in terms of uncertainty aversion. In Section 5, we propose one particular function – based on Rényi's (1961) generalized entropy function – as a preference function, and we illustrate it with a simple choice problem. In Section 6, we discuss our findings and relate them to other approaches Knightian uncertainty in the literature: maximin, maximax, the Hurwicz criterion, Laplace's principle of insufficient reason and the principle of minimum regret.

[4]The Sure Thing Principle is sometimes also referred to as 'independence of irrelevant alternatives'.

Section 7 concludes. All proofs are contained in the Appendix.

2 Conceptual clarifications

We present some critical reflections on key concepts of decision theory and related literature in the following.

Risk, uncertainty and ambiguity. Zweifel and Eisen (2012) state that 'the risk of an activity is represented by the probability density $p(x)$ defined over possible consequences x' where consequences may mean utility levels or monetary payoffs. $p(x)$ may be exogenously specified or scientifically calculable objective probabilities (cf. Machina and Rothschild 2008). If we apply this definition to the Keynes-Knight definition of uncertainty, *Knightian uncertainty* would then just amount to the non-existence of such an objective probability density function (PDF), whereas *ambiguity* would imply that there is at least incomplete knowledge concerning chances of outcomes or more than one PDF, and the decision maker is not sure about the 'true' distribution. In other words, if information concerning probabilities is partly missing or if there are several non-identical PDFs over consequences, possibly even weighted by some subjective weighting factors (2nd order probability distributions), then we face an ambiguous situation (cf. Gravel, Marchant and Sen 2012). It is worth noting that even though ambiguity and Knightian uncertainty are in principle very distinct concepts, they are often used interchangeably.

According to Machina and Rothschild (2008), there are two major theory strands concerning choice under Knightian uncertainty: the state-preference approach (Debreu 1959, Arrow 1964, Hirshleifer 1965, Hirshleifer 1966, Yaari 1969) and the hypothesis of probabilistic sophistication. The state-preference approach starts from a set of states of the world $\mathcal{S} = \{s_a, \dots, s_n\}$ and constructs a theory of choice with state-payoff bundles $(c_1, \dots, c_n)$ as objects of choice. Individuals are assumed to have preferences over state-payoff bundles just like regular commodity bundles. L.J. Savage's 1954 contribution was to define an 'act' as a mapping from states to consequences and that there exists a subjective belief, derived from preferences, which substitutes for objective probabilities. Much later, this framework was fortified by the hypothesis of probabilistic sophistication (Machina and Schmeidler 1992) which clarified the notion of subjective probabilities. It states that individuals entertain subjective probabilities which take the form of additive subjective probability measures $\mu(\cdot)$ over the state space $\mathcal{S}$. Anscombe and Aumann (1963) refined Savage's framework by assuming consequences to be risky lotteries rather than simple outcomes. Within this Anscombe-Aumann framework, Bewley (2002) introduced the assumption that individuals may assert that two alternatives are incomparable and that they may only accept an alternative when it is actually preferred to their current status quo. Bewley thus assumes that Knightian preferences are incomplete.

The literature strand that has spawned from the impact of the Ellsberg experiment has been subsumed under the umbrella term 'ambiguity aversion literature' (Al-Najjar and Weinstein 2009). Gilboa and Schmeidler (1989) and Schmeidler (1989) both expanded the Anscombe-Aumann framework to accommodate Ellsberg-type behavior, i.e. the preference of risk to uncertainty. Gilboa and Schmeidler (1989) provided an axiomatic foundation of 'maximin expected utility' (MEU) with multiple priors over the state space, so that the utility of an act is the minimal expected utility resulting from the priors. Schmeidler (1989) introduced the mathematical concepts of capacities and Choquet integration to model ambiguity aversion. Finally, Klibanoff, Marinacci and Mukerji (2005) and Nau (2006) modeled ambiguity via second order probability distributions and ambiguity aversion over the concavity of some second order utility function, i.e. a utility of expected utilities. It is especially this approach that has been applied frequently in climate change economics and related policy analyses (Millner, Dietz and Heal 2010, Traeger 2011, Heal and Millner 2013).

Descriptive and normative decision theory. There is, it seems, a dichotomy of approaches in decision theory. On the one hand, there are descriptive approaches that try to incorporate behavioral findings into existing theories to 'bring theory closer to reality' (Gilboa 2010: 4). Such approaches will be helpful whenever one is interested in descriptive prediction of behavior under ambiguity or uncertainty. On the other hand, the normative approaches treat behavioral peculiarities in conflict with some of their axioms such as famously reported by Allais (1953) and Ellsberg (1961) as errors of human reasoning. The ultimate aim is thus to 'bring reality closer to theory' (ibid.) by pointing out these errors of reasoning to decision makers to make better decisions possible in the future. Such theories can help determining what ought to be done, given that the decision maker agrees with the theory's premises. Obviously, one can have problems with the inherent paternalism of such theories. On the other hand, proponents usually stress that only normative decision theories can help overcoming human thinking biases that irrefutably exist.

Risk and probability. Closely related to the notions of risk and ambiguity are issues of measurability of risk ('riskiness') and probability. These notions are ubiquitous in economics, and yet it seems that their usage can be problematic. Consider the decision of investing in asset A or asset B. The standard way of arguing here is that investment A is said to be riskier than B if the standard deviation of its market price trajectory is larger. This does not seem entirely convincing. As an illustrative example, assume that the choice is between Microsoft stocks and Greek state bonds. Following the standard argument would lead to the conclusion that Microsoft stocks are riskier than Greek state bonds. From recent history, this statement seems questionable. The underlying issue however is whether and to what extent risk can be quantified, possibly even objectively.

If one worries about risk quantification, it entails thinking about probability quantification as well. Many approaches in economic theory require at some point the existence of probabilities that are objectively 'true'. Philosophical details with the concept of truth aside, such probabilities are unlikely to exist in most practical applications. And even if data is abundant, de Finetti's circularity critique[5], which argued that in order to define probabilities the classical or frequentist way, one needs to know the meaning of 'equally probable' first, seems valid (ibid.). De Finetti made these arguments in favor of Bayesian statistics, which states that in principle 'any uncertainty can and should be quantified' (cf. Gilboa 2010: 6). While this is arguably the predominant paradigm in economics to date, there is a catch here as well: Bayesian reasoning requires priors, which are highly subjective, leading to highly subjective results. As Feduzi and Runde

[5] "Therefore, these two ways of defining probability [...] are airy-fairy, unless one states beforehand what 'equally probable' means" (de Finetti 2008 [1979]: 4).

(2011) have pointed out, the consequence might be to face a decision problem under Knightian uncertainty.

3 Characterization and representation of preferences

Our formulation in this section makes use of what is known as the Lieb-Yngvason axiomatization of the second law of thermodynamics (Lieb and Yngvason 1999). They provide an axiomatic characterization of 'entropy' for thermodynamic settings and derive some properties of entropy. As they use a formal axiomatization and derivation of results that is completely independent of the semantic meaning of the formalism, their framework can be applied to other substantive contexts where the formal assumptions have a plausible interpretation. We therefore employ parts of their formal framework, results and proofs, with substantial adaptation and reinterpretation, to a setting of decision-making under Knightian uncertainty.

First, we introduce our setting and notation, and provide some basic definitions in Section 3.1. In Section 3.2, we state the assumptions on the set of Knightian lotteries and the axioms on the preference relation that describe what exactly we mean by uncertainty preferences. In Section 3.3, these axioms are be shown to uniquely constitute a preference function under Knightian uncertainty. Furthermore, we derive and discuss some properties of the function.

3.1 Setting, notation and basic definitions

A *simple Knightian lottery* maps states of nature to a vector $y = (y_1, \ldots, y_n)$ of payoffs, where $n \in \mathbb{N}$ with $n \geq 2$ is the number of potential states of nature and $y_i \in \mathbb{R}$ is the payoff if state i is realized (with $i = 1, \ldots, n$). The *payoff* in each state is an amount of some good which is the same good in all states. For example, this could be money, some primary good, or the level of well-being. One can also think of a simple Knightian lottery y as a discrete and finite payoff distribution over given states of nature. $Y \subseteq \mathbb{R}^n$ denotes the set of potential simple Knightian lotteries (for short: *simple-lottery set*) from which y is taken. We do not make any assumptions whatsoever about the objective or subjective probabilites, or anything like that, with which states of nature are realized.

The total payoff volume over all states for any lottery $y \in Y$ is $\overline{y} = \sum_{i=1}^{n} y_i$. If payoff in all states is non-negative (as in Section 5), $s_i^y = y_i/\overline{y}$ is the relative payoff share that lottery y yields in state i with respect to the total payoff volume $\overline{y}$. By construction, $0 \leq s_i^y \leq 1$ for all $i = 1 \ldots n$ and $\sum_{i=1}^{n} s_i^y = 1$ for any given $y \in Y$. Furthermore, denote by $S \subseteq [0{,}1]^n$ the set containing all possible such distributions s over n states of the world, with any particular element $s = (s_1, \ldots, s_n) \in S$.

We denote by $\underline{0}$ the n-vector $(0, \ldots, 0)$, that is, the lottery that yields zero payoff in all states; and by $\underline{1}$ the n-vector $(1, \ldots, 1)$, that is, the lottery that yields unit payoff in all states. Then, $c\underline{1}$ with $c \in \mathbb{R}$ is a lottery which yields the same amount c of payoff in each of the n potential states of the world. In other words, it yields a payoff of c that is perfectly certain: while one does not know ex-ante which state of nature will actually be realized, one does know ex-ante that – whatever state it will be – the payoff will be c. For any lottery $y \in Y$, we denote by $y^c = (\overline{y}/n)\underline{1}$ the corresponding lottery which yields the same total payoff volume as y, but distributed perfectly evenly over all n potential states of the world (for short: 'pc-corresponding lottery'). That is, the lottery y^c yields a payoff of $\overline{y}/n$ for certain.

From simple Knightian lotteries one may obtain more complex ones through the operations of scaling and compounding. These are defined as follows.

Definition 1 (scaled Knightian lottery)
For any simple Knightian lottery $y \in Y$ with payoff distribution $y = (y_1, \ldots, y_n)$ and any $\lambda > 0$, the *scaled Knightian lottery* λy is the lottery with payoff distribution $(\lambda y_1, \ldots, \lambda y_n)$. It is an element of the set $Y^{(\lambda)}$ that contains all lotteries that are obtained through scaling each simple lottery $y \in Y$ with the factor λ.

By this definition, the λ-scaled lottery λy is obtained from the simple lottery y, by multiplying the payoff in each state of nature by the scalar factor λ, which may be greater or smaller than one. Hence, total payoff volume of the scaled lottery is the λ-fold of that of the simple lottery, $\overline{\lambda y} = \lambda \overline{y}$, while it features the same relative distribution of payoff shares, $s^{\lambda y} = s^y$. From this definition, it also follows that $\mu(\lambda y) = (\mu\lambda)y$ and

$(Y^{(\lambda)})^{(\mu)} = Y^{(\mu\lambda)}$ for all $\mu,\lambda > 0$. We identify $1y = y$ and $Y^{(1)} = Y$.

One of the work horses of decision theory under uncertainty is the concept of a compound lottery. In standard expected-utility theory, a compound lottery is a lottery that has as its potential outcomes again lotteries. Hence, a compound lottery is a multi-stage lottery. For example, the first-stage lottery is flipping a coin, and the outcome of that lottery (heads or tail) determines which out of two second-stage lotteries to play, say, which one out of two urns to draw a colored ball from. Eventually, it is the second-stage lottery actually played that yields a payoff. As the potential outcomes of the first stage-lottery (heads or tail) are mutually exclusive, only one of them is realized and, consequently, only one of the second-stage lotteries (drawing from different urns) is played. As there is uncertainty in the first stage as well as in the second stage and payoff emerges only in the second stage, which is conditional on the first stage, the outcome space has a conjoint algebraic structure.

Here, we build on a very different notion of compounding, which is more similar to the concept of 'concatenation' introduced by Luce (1972). Nevertheless, we will speak of 'compound lotteries' and 'compounding' here.

Definition 2 (compound Knightian lottery)
For any two simple Knightian lotteries $x,y \in Y$, the *compound Knightian lottery* $x \oplus y \in Y \times Y$ is the situation when both simple lotteries x and y are played for sure and independently of each other, and one receives a payoff from both according to the actually realized state in each of the two. If state i is actually realized in the x-lottery and state j in the y-lottery, then payoff from the compound lottery is $x_i + y_j$ (for all $i,j = 1,...,n$).

The basic idea of compounding two simple lotteries, according to this definition, is that both simple lotteries are played for sure and independently of each other, and one receives a payoff from each of the simple lotteries. Payoff, thus, comes from both of the two simple lotteries. Hence, the compound lottery $x \oplus y$ has nn different outcome states (Table 1). Which one out of these is actually realized is determined by which one of the n states of nature of the underlying simple x-lottery is actually realized and which

Table 1: Potential outcomes of the compound lottery $x \oplus y$

	y-state 1	...	y-state j	...	y-state n
x-state 1	$x_1 + y_1$	...	$x_1 + y_j$	...	$x_1 + y_n$
$\vdots$	$\vdots$		$\vdots$		$\vdots$
x-state i	$x_i + y_1$	...	$x_i + y_j$	...	$x_i + y_n$
$\vdots$	$\vdots$		$\vdots$		$\vdots$
x-state n	$x_n + y_1$	...	$x_n + y_j$	...	$x_n + y_n$

one of the n states of nature of the underlying simple y-lottery is actually realized. As payoffs from the x-lottery and from the y-lottery are in units of the same good, and both are received, the actually realized payoffs from both lotteries can simply be added.[6]

The total payoff volume of the compound lottery $x \oplus y$ over all nn states is $\overline{x \oplus y} = n(\overline{x} + \overline{y})$. This corresponds to an addition of the total payoff volumes of the underlying simple lotteries, $\overline{x}$ and $\overline{y}$, weighted by a factor of n for the multiplication of the number of states due to compounding two lotteries defined over n states of nature. As for the mean payoff per state, i.e. total payoff volume divided by number of states, we have $\overline{x \oplus y}/nn = \overline{x}/n + \overline{y}/n$. This is just the sum of the two simple-lottery mean payoffs per state, which is due to the essentially additive character of compounding.

By Definition 2, compounding is essentially an additive operation:[7] the two simple lotteries are not mutually exclusive (that is, only one of them pays out), but both of them pay out so that two payoffs are received. The presupposition that both lotteries in a compound pay out is a major substantive deviation from how we usually think about compounding lotteries. Technically, this presupposition allows us to work without probabilities, which do not exist in the setting studied here. Mathematically, the essentially additive understanding of compounding induces an additive algebraic structure on the outcome space. This leads – with the axioms on the preference relation that become plausible with Definition 2 of compounding (see Section 3.2) – to the preference function

[6]One may construct a simple lottery that is rationally equivalent to the compound lottery $x \oplus y$: it is defined over nn states and has a payoff vector $(x_1 + y_1, ..., x_1 + y_n, ..., x_n + y_1, ..., x_n + y_n)$. This rationally equivalent lottery is neither an element of Y nor of $Y \times Y$, though.

[7]But note: as both simple lotteries are played independently of each other, it may be that in each of them a different state is actually realized. This makes the additive compounding considered here nevertheless non-trivial and essentially different from ordinary addition of payoff-vectors.

being additive (see Proposition 2).

Formally, a compound lottery $x \oplus y$ with $x,y \in Y$ is an element (x,y) in the Cartesian product $Y \times Y$. In general, the Cartesian product has as its elements ordered pairs and, therefore, is neither commutative nor associative. Here, however, from the substantive definition of compounding – the two simple lotteries are played independently of each other, and one receives a separate payoff from both – it is obvious that the operation of forming a compound lottery is a commutative and associative operation: when forming multiple compounds of simple-lotteries, the order and the grouping is irrelevant. Thus, we identify $x \oplus y = y \oplus x$, as well as $(x \oplus y) \oplus z = x \oplus (y \oplus z) = x \oplus y \oplus z$ for all simple lotteries $x,y,z \in Y$. As for scaling compound lotteries, from the two substantive definitions it is obvious that $\lambda(x \oplus y) = (\lambda x) \oplus (\lambda y)$ and $(Y \times Y)^{(\lambda)} = Y^{(\lambda)} \times Y^{(\lambda)}$ for all simple lotteries $x,y \in Y$ and all scaling factors $\lambda > 0$.

From the substantive definition of compounding it is also obvious that to compound an uncertain lottery with a certain one is essentially the same as simply adding the payoff vectors of the uncertain and the certain lottery: for any $y,y' \in Y$, no decision-maker could distinguish the two lotteries $y^c \oplus y'$ and $y^c + y'$. As y^c yields the payoff $\overline{y}/n$ for certain, that is, in each potential state of nature, it does not add any uncertainty to the compound but simply adds a certain payoff in each potential state. Uncertainty in the compound is entirely due to which state is actually realized in the y'-lottery. Therefore, $y^c \oplus y'$ and $y^c + y'$ are equivalent.

By compounding N scaled copies of a simple lottery $y \in Y$ with potentially different scaling factors $\lambda_1,...,\lambda_N > 0$ one can form a multiple scaled copy of y. It is an element in the multiple scaled copy of Y.

Definition 3 (multiple scaled copy of Y)
A *multiple scaled copy of* Y is formed by compounding N scaled copies of the simple-lottery set Y with $\lambda_1,...,\lambda_N > 0$: $Y^{(\lambda_1)} \times ... \times Y^{(\lambda_N)}$.

One particular, and degenerate, multiple scaled copy of Y is Y itself ($N = 1$, $\lambda = 1$). We have now built up the universe of Knightian lotteries between which the preference relation establishes relations. It is constituted by all multiple scaled copies of the simple-lottery set Y.

The core element of our analysis is a binary relation $\succeq$ on all multiple scaled copies of the simple-lottery set Y. It describes for any two lotteries which one is preferred over the other. We employ the following notation and interpretation for the preference relation: for any two lotteries x and y from the universe of all Knightian lotteries, $x \succeq y$ means 'lottery x is (weakly) preferred over lottery y'. If $x \succeq y$ and $y \succeq x$, we write $x \sim y$ and say that 'lottery x is as goods as lottery y'. If $x \succeq y$ and not $y \succeq x$, we write $x \succ y$ and say that 'lottery x is strictly preferred over lottery y'.[8]

3.2 Axioms and assumptions

The following three assumptions give structure to the set Y of simple Knightian lotteries and to the universe of all Knightian lotteries. Although none of them is needed for the existence of the preference function, they will allow us to derive some useful properties of the preference function.

Assumption 1 (Convexity of Y)
For all $x,y \in Y$ and $0 \leq \lambda \leq 1$, $\lambda x + (1-\lambda)y \in Y$.

Convexity of the simple-lottery set Y means that for any two Knightian lotteries from this set, any linear convex combination of the two is also an element of the set. This assumption is needed to show that the preference function is concave on Y.

Assumption 2 (Symmetry of Y)
For all $y \in Y$ and every permutation matrix P, $Py \in Y$.

Symmetry of the simple-lottery set Y means that for any Knightian lottery from this set, all lotteries that can be formed from it through permutation of the statewise payoffs are also elements of the set. For example, if $y = (y_1,y_2,y_3)$ is a simple Knightian lottery form the set Y, then (y_1,y_3,y_2), (y_2,y_1,y_3), (y_2,y_3,y_1), (y_3,y_1,y_2) and (y_3,y_2,y_1) must also be elements of the set. This assumption is needed to show that the preference function is symmetric on Y.

[8]We will also use $x \preceq y$ as synonymous with $y \succeq x$, and $x \prec y$ as synonymous with $y \succ x$.

Assumption 3 (Disjointness of the universe)
Any two sets of lotteries in the universe of all multiple scaled copies of Y are disjoint sets, that is, every lottery in the universe belongs to exactly one set.

We impose the following nine axioms on the preference relation $\succeq$ on all multiple scaled copies of the simple-lottery set Y. We explain and illustrate axioms where we deviate from what can safely be regarded as standard in economic theory (cf. e.g. Savage 1954, Anscombe and Aumann 1963, Gilboa and Schmeidler 1989).

Axiom 1 (Reflexivity)
For all $y \in Y$, $y \sim y$.

Axiom 2 (Transitivity)
For all x,y,$z \in Y$, $x \succeq y$ and $y \succeq z$ implies $x \succeq z$.

Axiom 3 (Completeness)
For all x,y from any multiple scaled copy of Y, either $x \succeq y$ or $y \succeq x$ or both.

Completeness is, as usually, a strong assumption. Here, it is even more so as comparability of any two lotteries should not only hold among all simple lotteries from the set Y, but is should hold beyond that in each multiple scaled copy of Y.[9]

Axioms 1 to 3 are fairly standard to any binary relation that provides a consistent and complete ranking. They do not yet contain any specific substantive content of preferences under uncertainty. Starting with the next axiom, we get more specifically to the meaning of uncertainty preferences.

Axiom 4 (Compounding consistency)
For all x,x',y,$y' \in Y$, $x \succeq x'$ and $y \succeq y'$ implies $x \oplus y \succeq x' \oplus y'$.

The compounding-consistency property means that the preference ordering between simple lotteries carries over to their respective compounds. That is, if some lottery

[9] Here, we take completeness as a basic axiom. Lieb and Yngvason (1999: Sections 3 and 4) show that completeness can be derived as an implication of eight more elementary axioms. As these are more plausible in the thermodynamic context studied by Lieb and Yngvason (1999) than in the decision-under-uncertainty context studied here, we take completeness as an elementary axiom.

x is preferred over x' and another lottery y is preferred over y', then the compound lottery $x \oplus y$ is preferred over the compound lottery $x' \oplus y'$. Hence, the compound lottery obtained from two simple lotteries each of which is preferred over another lottery will also be preferred over the compound lottery of these two other lotteries. With compounding as an essentially additive operation (according to Definition 2), this seems highly plausible.

Axiom 5 (Scaling invariance)
For all x,$y \in Y$ and all $\lambda > 0$, $x \succeq y$ implies $\lambda x \succeq \lambda y$.

While the compounding-consistency property (Axiom 4) refers to consistency in terms of compound lotteries, the scaling-invariance property refers to invariance of the ranking under scaling. This is to say that a statewise proportional change in payoff levels and, hence, total payoff volumes does not alter the preference ordering. Scaling invariance rules out, for example, that at a low level of overall payoff volume a certain lotteries is preferred over an uncertain one (uncertainty aversion), while at a high level of overall payoff volume the scaled-up version of the uncertain lottery is preferred over the scaled-up version of the certain lottery (uncertainty love). It is a strong property of the preference relation, and it induces – together with the other axioms – a strong property on the preference function: the preference function will be homogenous of degree one ('extensive').

Axiom 6 (Splitting and recombination)
For all $y \in Y$ and all $0 < \lambda < 1$, $\quad y \sim \lambda y \oplus (1-\lambda) y$.

The splitting-and-recombination property means that it does not matter for the preference ranking whether one faces some simple lottery y or the compound lottery consisting of some scaled-down versions of y where the scaling factors add up to one. One should be indifferent between the two. With scaling invariance (Axiom 5) and, again, compounding as an essentially additive operation (according to Definition 2), this seems highly plausible, too. It is an independent axiom, though, in that it is the only axiom (among those which establish unique existence of the preference function) that relates a simple lottery and a compound lottery. It is this axiom that establishes, at bottom, how simple lotteries can be compared to compound lotteries.

Axiom 7 (Continuity)
For all x,y,z_0,$z_1 \in Y$ and a sequence ε_k with $\lim_{k\to\infty} \varepsilon_k = 0$, $x \oplus \varepsilon_k z_0 \succeq y \oplus \varepsilon_k z_1$ for $k \to \infty$ implies $x \succeq y$.

The continuity property means that the compounding of each of two lotteries with small-scale perturbations tending to zero does not affect the preference ordering between these two lotteries. This technical property guarantees that there are no discontinuities in the preference relation, and that the preference function is continuous.

A notable point of our axiomatic framework is that we do *not* assume the Knightian analogue of what is commonly referred to as the 'Independence Axiom' or the 'Sure Thing Principle' (Savage 1954) in expected-utility theory. This property (more exactly: its Knightian analogue) follows from the other axioms in our framework.

Lemma 1 (Independence)
It follows from Axioms 1, 2 and 4 through 7 that for all x,y,$z \in Y$, $x \oplus z \succeq y \oplus z$ implies $x \succeq y$.

Proof. See Appendix A.1. □

Lemma 1 states that all axioms introduced so far – with the exception of completeness – imply independence. In other words, the axioms employed here are stronger than the independence assumption alone. The property in Lemma 1 can be interpreted as follows: events occurring anyway do not affect the preference ordering between lotteries. This property is generally considered an important feature of theories of rational choice.[10] One must bear in mind, though, that in our framework compounding has a different definition than usual (Definition 2) and, therefore, the substantive meaning of

[10]However, it has been pointed out by Al-Najjar and Weinstein (2009) to pose a problem to the ambiguity aversion literature as a normative theory, since ambiguity averse agents can be shown to violate this principle. On the descriptive level, people have been shown to systematically violate this principle under sonme conditions which is known as Allais paradox (Allais 1953).

the independence property differs from the usual one.[11]

The two following axioms on the preference relation are not necessary for the unique existence of a preference function. Yet, they are necessary for the preference function to have additional properties that allow its interpretation in terms of uncertainty aversion.

Axiom 8 (Symmetry)
For all $y \in Y$ where Y is symmetric and every permutation matrix P, $y \sim P\,y$.

The symmetry property states that if in the payoff distribution for a given Knightian lottery the payoffs are permutated over states of nature, then the resulting payoff distribution is as good as the original one. For example, one should be indifferent between the lottery (1\$,3\$,5\$) and the permutated lottery (5\$,3\$,1\$). That means, the sequence in which states of nature are numbered is irrelevant for the preference ranking of Knightian lotteries. This property excludes state-dependent utility, where the utility of a payoff depends on the exact state in which it is paid out.

Axiom 9 (Convexity)
For all x,$y \in Y$ where Y is convex and $0 \leq \lambda \leq 1$, $\lambda x + (1-\lambda)y \succeq \lambda x \oplus (1-\lambda)y$.

Here, the left-hand side is an ordinary convex linear combination of two lotteries from the convex set Y. It is, therefore, again an element of this set Y. The right-hand side is a compound lottery. It is an element of $Y^{(\lambda)} \times Y^{(1-\lambda)}$. Hence, the axiom states that an ordinary convex linear combination of two simple lotteries is preferred over a convex compound of these lotteries. This axiom of convexity of the preference relation is different from the normal convexity-of-preferences axiom, which is: a convex linear combination of two lotteries, among which one is indifferent, is preferred over the underlying extreme lotteries. The axiom used here is stronger: it implies (given the other axioms), but is not implied by, the normal convexity property (proof in Appendix A.2).

[11]The usual von-Neumann-Morgenstern axioms on the preference relation (reflexivity/transitivity, completeness and independence) imply that a preference function exists which has an expected-utility form. Here, although we also have reflexivity/transitivity, completeness and independence, we do *not* have a preference function which has an expected-utility form. The reason is that our concept is in a Knightian setting without probabilities and with an essentially additive compounding operation. Hence, our independence property means and implies something different than the independence property in an expected-utility setting.

But, just like the normal convexity property, it implies (for a convex set Y of lotteries) that a lottery where payoff is distributed more evenly over states of nature is preferred to any other lottery with the same total payoff volume where payoff is distributed more unevenly over states. This property is at the core of our notion of uncertainty aversion (to be elaborated in Section 4).

3.3 Existence and properties of the preference function

We have now gathered all building blocks necessary to demonstrate the existence of a preference function that represents the preference relation $\succeq$ on Y. For this, we proceed in two steps. First, we demonstrate that with Axioms 1–7 on the preference relation, a preference function uniquely exists that represents the preference relation (Proposition 1). We also discuss the basic properties of this function (Proposition 2). Second, employing in addition Assumptions 1– 3 as well as Axioms 8 and 9, the preferences function has two more properties (Proposition 3), which make it represent uncertainty aversion (Section 4).

Proposition 1 (Existence and uniqueness of a preference function)
Let Y be a set of simple Knightian lotteries and $\succeq$ a binary relation on the multiple scaled copies of Y. Then, the following two statements are equivalent:

1. *The relation $\succeq$ satisfies Axioms 1–7.*
2. *There exists a continuous function $H : Y \to \mathbb{R}$ that represents the relation $\succeq$ in the following sense. For all $N \geq 1$, $M \geq 1$, all $x^1, \ldots, x^N, y^1, \ldots, y^M \in Y$ and all $\mu_i \geq 0$ and $\lambda_j \geq 0$ with $\mu_1 + \ldots + \mu_N = \lambda_1 + \ldots + \lambda_M$, it holds that*

$$\mu_1 x^1 \oplus \ldots \oplus \mu_N x^N \succeq \lambda_1 y^1 \oplus \ldots \oplus \lambda_M y^M \tag{1}$$

if and only if

$$\sum_{i=1}^{N} \mu_i H(x^i) \geq \sum_{j=1}^{M} \lambda_j H(y^j) \ . \tag{2}$$

The function H is uniquely defined on Y up to a linear-affine transformation. That is, if H characterizes the relation $\succeq$ on Y, then also any other function $\hat{H} : Y \to \mathbb{R}$ with

$\hat{H}(y) = aH(y) + b$ *where* $a, b \in \mathbb{R}$ *and* $a > 0$.

Proof. See Appendix A.3. □

The imposition of Axioms 1 through 7 on the preference relation $\succeq$ implies the existence of a function that maps from the set Y of simple Knightian lotteries to the real numbers such that a more preferred Knightian lottery is assigned a greater real number than a less preferred lottery. Thus, the function $H : Y \to \mathbb{R}$ is a preference function on the set Y. As this function is unique up to linear-affine transformations, it is a cardinal preference index. Note that neither Assumptions 1 (convexity of Y), 2 (symmetry of Y) or 3 (disjointness of the universe) on the sets of Knightian lotteries nor Axioms 8 (symmetry) or 9 (convexity) on the preference relation $\succeq$ are needed for existence and uniqueness of the preference function.

The representation statement in Proposition 1, which is about ranking compound lotteries, is more general than the usual representation statement, which is about ranking simple lotteries. Simple-lottery representation is, of course, just a particular case of the more general representation statement in Proposition 1.

Corollary 1 (Simple-lottery representation)
For any function H *sensu Proposition 1 and all* x,$y \in Y$ *it holds that* $x \succeq y$ *if and only if* $H(x) \geq H(y)$.

Proof. See Appendix A.4. □

Hence, function H represents the preference relation $\succeq$ also in the normal sense, that is, with respect to simple lotteries. In the following, we state two basic properties of the function H.

Proposition 2 (Properties of preference function)
Suppose a preference function H *exists (sensu Proposition 1). If the universe of all multiple scaled copies of the simple-lottery set* Y *fulfills Assumption 3, the function* H *has the following properties for all* x,$y \in Y$ *and all* $\lambda > 0$*:*[12]

[12]To keep notation simple, we denote with the letter H both functions – the preference function $H : Y \to \mathbb{R}$ on the set Y of simple lotteries and the preference function $H : Y \times Y \to \mathbb{R}$ on the set

1. *Additivity:* $H(x \oplus y) = H(x) + H(y)$.
2. *Extensivity:* $H(\lambda y) = \lambda H(y)$.

Proof. See Appendix A.5. □

These two properties make the function H a special preference function. It is additive and extensive: the preference index of a compound lottery is exactly the sum of the preference indices of the simple lotteries (additivity), and if payoff is increased λ-fold in each potential state of nature, the preference index also is increased λ-fold (extensivity). These two properties of the preference function are due to the essentially additive understanding of compounding (Definition 2) and Axioms 4, 5 and 6 on scaling and compounding. With these, there is no loss or gain in terms of preferences from the process of compounding itself.[13]

In thermodynamics, a function that represents a relation with the properties stated in Axioms 1–7 and which has the properties stated in Propositions 1, is called an *entropy function* (Lieb and Yngvason 1999: 24). If it has in addition the property of additivity (Proposition 2), it is called an *additive entropy function*. The preference function H is, thus, an (additive) entropy function. We therefore speak of 'entropic uncertainty preferences', to distinguish our preference concept from other preference concepts under Knightian uncertainty.

Additivity and extensivity of the preference function imply that the preference function is monotonic in some sense. That is, it represents non-satiation of the preference relation: more payoff is strictly preferred over less payoff. The exact meaning of 'more payoff' is specified in the following corollary to Proposition 2.

Corollary 2 (Monotonicity)

Suppose a preference function H exists (sensu Propositions 1 and 2) which is, in particular, additive and extensive. Then it holds for all $y \in Y$, all $\lambda > 1$ and all $c > 0$

$Y \times Y$ of all compound lotteries. There should not be any confusion, as it is obvious in every instance from the argument of H which of the two functions we mean.

[13] Luce et al. (2008) have used an entropy-based modelling approach to account for the utility drawn from the process of compounding itself (which they refer to as gambling), but their approach takes place within an expected utility framework.

that

1. $H(\lambda y) > H(y)$ *or, equivalently,* $\lambda y \succ y$,

and

2. $H(y + c\underline{1}) > H(y)$ *or, equivalently,* $y + c\underline{1} \succ y$.

Proof. See Appendix A.6. □

Here, monotonicity comes in two kinds. First, if one scales up payoff in each state of nature by the same factor $\lambda > 1$, one obtains a lottery that is strictly preferred over the original one (scaling-monotonicity). Second, if one adds the same strictly positive amount $c > 0$ of payoff in each state of nature, one obtains a lottery that is strictly preferred over the original one (adding-monotonicity). This kind of monotonicity of the preference relation is already implied by Axioms A1–A7, which are necessary and sufficient for the existence of the preference function. It does not need to be assumed separately in our framework. Rather, it is an inherent property of any entropic preference function.

With two more properties, symmetry and concavity, which are routinely assumed in statistical physics and information theory, the entropy function becomes a statistical measure of the homogeneity of a distribution: a more even distribution is characterized by a higher entropy than a more uneven distribution. In the context of payoff distributions over potential states of nature, this makes the entropy/preference function to represent uncertainty aversion. We now introduce these two properties, before then discussing uncertainty aversion in Section 4.

Proposition 3 (Symmetry and concavity of the preference function)
Let Y be a set of simple Knightian lotteries and $\succeq$ a binary relation on the multiple scaled copies of Y, and suppose $\succeq$ satisfies Axioms 1–7 such that a preference function H exists (sensu Proposition 1).

(1) If Y fulfills Assumption 2 (symmetry) and $\succeq$ satisfies Axiom 8 (symmetry), then H is a symmetric function on Y: $H(y) = H(Py)$ for every permutation matrix P. Conversely, if H is a symmetric function, then Axiom 8 holds a fortiori.

(2) If Y fulfills Assumption 1 (convexity) and $\succeq$ satisfies Axiom 9 (convexity), then

H is a concave function on Y: $H(\lambda x + (1-\lambda y)) \geq \lambda H(x) + (1-\lambda)H(y)$ for all $x,y \in Y$ and $0 \leq \lambda \leq 1$. Conversely, if H is a concave function, then Axiom 9 holds a fortiori.

Proof. See Appendix A.7 □

4 Uncertainty aversion

We now clarify the notion of aversion against Knightian uncertainty (Section 4.1), and demonstrate that the entropic preference function H represents uncertainty aversion if it is symmetric and concave (Section 4.2). We then discuss how to measure the degree of uncertainty aversion, and how to characterize a preference relation in terms of its degree of uncertainty aversion (Section 4.3). Throughout this section, we follow the standard program of expected-utility theory under probabilistic risk, and in each step develop concepts for Kinghtian uncertainty in analogy.

4.1 Defining uncertainty aversion

To start with, we introduce the concept of 'uncertainty dominance' as a relation between two Knightian lotteries in terms of which one is more uncertain.

Definition 4 (uncertainty dominance[14])
For any two simple Knightian lotteries $x,y \in Y$, y is said to *uncertainty-dominate* x (or: *x is uncertainty-dominated by y*) if

$$\sum_{i=1}^{k} x_i^{\downarrow} \leq \sum_{i=1}^{k} y_i^{\downarrow} \quad \text{for} \quad k = 1,...,n-1 \tag{3}$$

$$\text{and} \quad \overline{x} = \overline{y} \ , \tag{4}$$

where $x^{\downarrow}$ and $y^{\downarrow}$ are the payoff vectors that are obtained from x and y, respectively, by rearranging their components in descending order, such that $x_1^{\downarrow} \geq x_2^{\downarrow} \geq ... \geq x_n^{\downarrow}$ and $y_1^{\downarrow} \geq y_2^{\downarrow} \geq ... \geq y_n^{\downarrow}$. Uncertainty dominance is said to hold *strictly* if (3) holds with strict

[14]What we call 'uncertainty dominance' here is normally called 'majorization' in mathematics ever since the term and formal definition have been introduced by Hardy et al. (1934/1952).

inequality for at least one $k \in \{1,...,n-1\}$.

The idea behind this definition is that a Knightian lottery y is said to uncertainty-dominate another lottery x if both lotteries have the same total payoff volume (Condition 4) and state-wise payoffs are more unequal under y than under x (Condition 3). In other words, the same total payoff volume $\overline{y} = \overline{x}$ is distributed more unequally over all potential states of nature under lottery y than under lottery x. For example, the lottery $y = (1\$,4\$)$ uncertainty dominates the lottery $x = (2\$,3\$)$. Whenever we speak of 'less uncertain' or 'more uncertain' lotteries, we mean it in the sense of uncertainty dominance (Definition 4): a lottery y is said to be 'more uncertain' than another lottery x and, conversely, x is 'less uncertain' than y, if y uncertainty-dominates x.

Uncertainty dominance does not establish a complete order on the set Y of simple Knightian lotteries. It only allows comparison between two Knightian lotteries with the same payoff volume.[15] Therefore, there may exist Knightian lotteries $x,y \in Y$ both of which are uncertain and which cannot be compared in terms of uncertainty-dominance. For example, with $x = (2\$,3\$)$ and $y' = (1\$,5\$)$ it holds neither that y' uncertainty-dominates x nor that x uncertainty-dominates y' according to Definition 4, simply because $\overline{x} \neq \overline{y'}$ so that Condition (4) cannot be met. While one may intuitively be tempted to say that 'payoff from y' is more uncertain than payoff from x' (because $y = (1\$,4\$)$ is more uncertain than $x = (2\$,3\$)$, and y' has even more unequal state-wise payoffs than y), it is also true that lottery y' pays out more overall than lottery x ($\overline{y'} = 6$ and $\overline{x} = 5$). Hence, if one compares their pc-corresponding lotteries, $x^c = (2.50\$, 2.50\$)$ and $y'^c = (3\$,3\$)$, y'^c would be strictly preferred over x^c because of its higher payoff level (due to monotonicity of the preference relation $\succeq$). To clearly distinguish uncertainty aversion from non-satiation, we keep uncertainty-dominance restricted to subsets of simple Knightian lotteries with the same total payoff volume.

One can give further plausibility to the interpretation of uncertainty-dominance as a criterion of the unevenness of a payoff distribution over states of nature. For, a lottery

[15] Strictly speaking, uncertainty dominance is not even a partial order but a preorder, since 'y uncertainty-dominates x' and 'x uncertainty-dominates y' does not imply $x = y$ (Marshall et al. 2011: 18-19).

y uncertainty-dominates another lottery x if and only if the latter can be obtained from the former by redistributing – in the spirit of Pigou (1912: 24) and Dalton (1920: 351) – payoff from states with relatively high payoff to states with relatively low payoff.

Lemma 2 (uncertainty dominance through Pigou-Dalton-transfers)
For any two simple Knightian lotteries x,$y \in Y$, x is strictly uncertainty-dominated by y if and only if x can be obtained from y through a finite sequence of transfers of the following kind: take two states j and k with j,$k \in \{1,...,n\}$, $j \neq k$ and $y_j < y_k$ and make a payoff transfer δ with $0 < \delta < y_k - y_j$ from the high-payoff state k to the low-payoff state j, so that $x_j = y_j + \delta$, $x_k = y_k - \delta$ and $x_l = y_l$ for all $l \in \{1,...,n\} \setminus \{j,k\}$.

Proof. This lemma is Theorem 2.1 in Arnold (1987: 14) and is proven there. □

A Pigou-Dalton-transfer redistributes some amount of payoff from a state of nature with relatively high payoff to a state with relatively low payoff such that the order of states in terms of higher-or-lower-payoff remains the same. It, thus, preserves the total payoff volume over all states of nature, and distributes payoff more evenly over states of nature. For example, consider x = (2$,3$,4$) and y = (1$,3$,5$). The payoff distribution x is more even than y because the former emerges from the latter through a Pigou-Dalton transfer of 1$ from state 3 to state 1 which leaves the payoff in state 2 unaltered and also preserves the rank-ordering of states in terms of payoff. As a result, the payoff in the lowest-payoff state (state 1) is not as low as before the transfer, and the payoff in the highest-payoff state (state 3) is not as high as before the transfer. Hence, payoff is more evenly distributed over all potential states of nature. By the same token, payoff from lottery x is more certain than payoff from lottery y. Hence, the reverse of a Pigou-Dalton transfer could be termed a 'total-payoff-volume-preserving spread of a Knightian lottery',[16] as it leads to a more uncertain lottery.

Compared to the given Knightian lottery y, any other lottery obtained from y through a sequence of Pigou-Dalton transfers is less uncertain: it features the same total payoff volume but with more even distribution over the states of nature. In the

[16]This is in analogy to using the 'mean-preserving spread of a probability distribution' to identify a 'more risky' probability distribution (Rothschild and Stiglitz 1970).

extreme, one may obtain from any Knightian lottery $y \in Y$ through a finite sequence of Pigou-Dalton transfers the pc-corresponding lottery y^c. It is not only less uncertain than y, but it is perfectly certain: from y^c no less uncertain lottery can be obtained through a Pigou-Dalton transfer, because payoff is distributed perfectly evenly over states, i.e. the payoff vector has the same payoff in each state of nature. At the other end of the uncertainty-domination chain for the given Knightian lottery y is the lottery ($\overline{y}/n$,0,...,0), or any permutation of it. This lottery strictly uncertainty-dominates all other lotteries with the same total payoff volume $\overline{y}$ (and which are not simply permutations of it). In this sense, it is the most uncertain lottery with total payoff volume $\overline{y}$.

With this understanding of more and less uncertain lotteries, we can now define uncertainty aversion. The idea is simply that a decision-maker who prefers all less uncertain lotteries over any given uncertain lottery is said to be uncertainty averse.

Definition 5 (uncertainty aversion)
A decision-maker with preference relation $\succeq$ on the simple-lottery set Y is said to be *uncertainty averse* (*neutral*, *loving*) if and only if for all x,$y \in Y$ where y strictly uncertainty-dominates x, $x \succ (\sim, \prec)\ y$.

One straight implication of this definition of uncertainty aversion is that an uncertainty averse decision-maker, who has a choice between some uncertain lottery y and the pc-corresponding lottery y^c – that is, a lottery with the same total payoff volume but distributed evenly over all potential states of nature, so that payoff $\overline{y}/n$ is certain – prefers the pc-corresponding lottery. Likewise, a risk-loving decision-maker rejects the pc-corresponding lottery, and an uncertainty neutral decision-maker is indifferent between the two.

Lemma 3 (uncertainty aversion implies strict preference for pc-corresponding lottery)
Consider a decision-maker with preference relation $\succeq$ on Y. If she is uncertainty averse (neutral, loving), then $y^c \succ (\sim, \prec)\ y$ for all $y \in Y$ with $y \neq y^c$.

Proof. See Appendix A.8 □

For example, if $n = 2$ and $\overline{y} = 10\$$, an uncertainty averse decision-maker would always prefer the pc-corresponding lottery $y^c = (5\$,5\$)$ to any other lottery y where $\overline{y} = 10\$$ is distributed unevenly over the two potential states of nature, such as (3\$,7\$) or (1\$,9\$).

In risk theory, it is normally taken as an elementary definition of risk aversion that a decision maker for all lotteries prefers the expected payoff from the lottery over the risky lottery. In terms of Knightian uncertainty, this would be analogous to saying that the decision maker for any Knightian lottery y prefers the pc-corresponding lottery y^c over the uncertain lottery y. In our approach, this is not the elementary definition of uncertainty aversion, but it is implied by our Definition 5 of uncertainty aversion (Lemma 3). This shows that our definition of uncertainty aversion (Definition 5) is stronger, by demanding that the decision-maker prefers *all less uncertain* lotteries over a given lottery, and *not just the least uncertain* lottery, namely the pc-corresponding one.

4.2 Uncertainty aversion in entropic preferences

Having a clear idea of uncertainty aversion, we now move on to demonstrating that the function H (introduced by Proposition 1) represents uncertainty aversion if it is symmetric and concave. The basic idea of uncertainty aversion according to Definition 5 is that the decision-maker always strictly prefers all less uncertain (that is: uncertainty-dominated) lotteries over a given uncertain lottery. In mathematics, functions that preserve the ordering of majorization (here: 'uncertainty-dominance') are called *Schur-concave*, or *S-concave*, after Schur (1923).

Definition 6 (Schur-concavity)
A function $f : A \to \mathbb{R}$ with $A \subset \mathbb{R}^n$ is said to be *Schur-concave* (*Schur-convex*) on A if for all $x,y \in A$

$$y \text{ majorizes } x \quad \Rightarrow \quad f(x) \geq (\leq)\ f(y)\ .$$

If equality holds only when x is a permutation of y, then f is said to be *strictly* Schur-concave (*strictly* Schur-convex) on A.

By this definition, if a Knightian lottery $y \in Y$ uncertainty-dominates another lottery $x \in Y$, a Schur-concave function would assign a higher value to x, which is the less uncertain lottery, than to y. With this property, a more fitting term for the function would actually be 'uncertainty-decreasing', or 'decreasing in uncertainty', as the higher the uncertainty, the smaller the function value. Likewise, Schur-convex functions are 'uncertainty-increasing' in that they assign a higher function value to more uncertain lotteries. That the function is not called Schur-increasing/decreasing but Schur concave/convex is due to the fact that Schur-concavity/convexity is closely linked to concavity/convexity, which will become obvious later (in Lemma 5). At this point, it is obvious that Schur-concavity of the preference function is the property that makes it represent uncertainty aversion in the sense of uncertainty-dominance (Definition 5).

Lemma 4 (A Schur-concave preference function represents uncertainty aversion)
Suppose a preference function H exists that represents the preference relation $\succeq$ on Y (sensu Proposition 1). If H is strictly Schur-concave (Schur-convex) on Y, it represents the preferences of a decision-maker who is uncertainty averse (loving): for all x,$y \in Y$ such that y strictly uncertainty-dominates x, $H(x) > (<)\ H(y)$.

Proof. Follows directly from Definitions 5 and 6. □

The crucial question now is: what are sufficient conditions for a function to be Schur-concave? For, it is under these conditions that the preference function H represents uncertainty aversion. The answer is well-known in the theory of convex functions.

Lemma 5 (Symmetry and concavity imply Schur-concavity)
Let $f : A \to \mathbb{R}$ where $A \subset \mathbb{R}^n$ is symmetric. If f is symmetric and (strictly) concave on A, then it is (strictly) Schur-concave on A.

Proof. This lemma is Proposition C.2 (C2c for 'strict') in Marshall et al. (2011: 97) and is proven there. □

Symmetry and concavity are jointly a sufficient condition for Schur-concavity. But they are not necessary in general. That is, a Schur-concave function does not need to be concave. A weaker sufficient, yet also not necessary, condition for Schur-concavity is that the function f is quasi-concave and symmetric (Marshall et al. 2011: 98, Theorem C.3). We can now state under which (sufficient) conditions on the preference relation, the decision-maker is uncertainty averse.

Proposition 4 (Uncertainty aversion in entropic preferences)
Consider a decision-maker with preference relation $\succeq$ on Y, and suppose a preference function H exists that represents the preference relation (sensu Proposition 1). If Y satisfies Assumptions 1 (convexity) and 2 (symmetry), and $\succeq$ satisfies Axioms 8 (symmetry) and 9 (convexity), the decision-maker is uncertainty averse.

Proof. See Appendix A.9 □

By this proposition, for the preference relation to describe uncertainty aversion, we build on two more assumptions on the simple-lottery set Y (Assumption 1: convexity, Assumption 2: symmetry) and two more axioms on the preference relation $\succeq$ (Axiom 8: symmetry, Axiom 9: convexity) than for mere existence and uniqueness of the preference function H (cf. Proposition 1). The decision-maker is uncertainty averse if the conditions of Proposition 4 hold, but not only if these conditions hold. As already noted above, convexity of the preference relation (equivalently: concavity of the preference function) is stronger than needed for uncertainty aversion. Schur-concavity of the preference function suffices (Lemma 4).

4.3 Measuring uncertainty aversion

We now want to measure a decision-maker's degree of uncertainty aversion. To this end, we introduce the concept of the certainty equivalent of a Knightian lottery.

Definition 7 (Certainty equivalent)
Consider a decision-maker with preference relation $\succeq$ on the simple-lottery set Y. For any Knightian lottery $y \in Y$, the *certainty equivalent* $C \in \mathbb{R}$ of y is defined through

$$C(y)\underline{1} \sim y \; . \tag{5}$$

In words, the certainty equivalent of a Knightian lottery y is the amount C of payoff that leaves the decision-maker indifferent between playing the uncertain lottery y and receiving the payoff C for certain. In Definition 7 we again formally identify a certain payoff with a payoff vector where this amount of payoff is obtained in each state of nature.[17]

An important question at this point is: does a certainty equivalent – as defined through Condition (5) – exist, and is it unique, for all preference relations and for all Knightian lotteries? To answer this question, we look at the definition of the certainty equivalent again, but now with a preference function.

Lemma 6

Suppose a preference function H exists that represents the preference relation $\succeq$ (sensu Proposition 1) and that is extensive (sensu Proposition 2). For any Knightian lottery $y \in Y$, the certainty equivalent C *of y uniquely exists and is given by*

$$C(y) = \frac{H(y)}{H(\underline{1})} \ . \tag{6}$$

Proof. See Appendix A.10. □

Equation (6) shows that for an entropic preference function which is extensive, the certainty equivalent of any Knightian lottery can explicitly be calculated from H. It also states that for given extensive preference function H, that is, if Y satisfies Assumptions 1–3 and $\succeq$ satisfies Axioms 1–7 (Propositions 1 and 2), for any Knightian lottery the certainty equivalent uniquely exists.

The denominator on the right-hand-side of Equation (6) is a normalization. Remember (from Proposition 1) that H is unique only up to linear-affine transformations. Therefore, one is free to set $H(\underline{1})$ – by choice of the additive constant in a linear-affine transformation of H – to any arbitrary real value. In fact, the concrete function that we will introduce in Section 5 is such that $H(\underline{1}) = 1$, which seems a natural choice for

[17] Here, the certainty equivalent $C(y)$ of an uncertain lottery is analogous to the *equally distributed equivalent income* introduced by Atkinson (1970: 250) to characterize an income distribution over members of society in terms of its inequality.

the normalization. With this, Lemma 6 reveals a remarkable property of the entropic preference function H: the entropic preference function H – with suitable normalization – yields as a preference index for any Knightian lottery y exactly the certainty equivalent C of this lottery. This means that the certainty equivalent is a sure-payoff metric measure of utility. It also implies that the entropic preference function H ranks all Knightian lotteries $y \in Y$ by their respective certainty equivalent $C(y)$. Hence, in an optimization problem under Knightian uncertainty, to find the lottery $y \in Y$ that maximizes $H(y)$ is equivalent to finding the lottery $y \in Y$ that maximizes $C(y)$ – a result that is in full analogy to expected-utility theory under risk (Chavas 2004: 35).

For given preference relation $\succeq$ and given lottery y, the certainty equivalent $C(y)$ (Definition 7) of lottery y may be greater than, equal to, or smaller than the certain payoff from the pc-corresponding lottery y^c, which is the total payoff-volume of lottery y divided by the number of states, $\overline{y}/n$. The difference between the two is the uncertainty premium of the lottery.

Definition 8 (Uncertainty premium)
Consider a decision-maker with preference relation $\succeq$ on the simple-lottery set Y, and suppose the certainty equivalent uniquely exists for all $y \in Y$. For any Knightian lottery $y \in Y$, the *uncertainty premium* $P \in \mathbb{R}$ of y is

$$P(y) = \frac{\overline{y}}{n} - C(y)\ . \tag{7}$$

From Equations (5) and (7) one obtains:[18]

$$\left(\frac{\overline{y}}{n} - P(y)\right)\underline{1} \sim y\ . \tag{8}$$

This means that the uncertainty premium $P(y)$ of a Knightian lottery y is the sure amount of payoff that the decision-maker is willing to forego at maximum to reach a situation of certainty: it makes the decision-maker indifferent between playing the

[18] Alternatively, Equation (8) could be taken as the definition of the uncertainty premium P of lottery y. Then, Equation (7) would be an implication.

uncertain lottery y or receiving for certain its total payoff per state, $\overline{y}/n$, minus the uncertainty premium P. It can thus be interpreted as the decision-maker's shadow costs of bearing uncertainty.

It is obvious from the defining Equation (7) that the uncertainty premium $P(y)$ uniquely exists whenever the certainty equivalent $C(y)$ uniquely exists. If a preference function H exists, it follows from Equations (6) and (7) that the uncertainty premium of any Knightian lottery $y \in Y$ can be calculated as follows:

$$P(y) = \frac{\overline{y}}{n} - \frac{H(y)}{H(\underline{1})} \ . \tag{9}$$

The two concepts of the certainty equivalent and the uncertainty premium of a Knightian lottery can be used to indicate whether a decision-maker is uncertainty avers, uncertainty neutral, or uncertainty loving.

Proposition 5 (Characterization of uncertainty aversion)
Consider a decision-maker with preference relation $\succeq$ on the simple-lottery set Y, and suppose the certainty equivalent uniquely exists for all $y \in Y$. The following three statements are equivalent:

1. *The decision-maker is uncertainty averse (neutral, loving).*
2. *For all $y \in Y$, $C(y) < (= , >) \ \overline{y}/n$.*
3. *For all $y \in Y$, $P(y) > (= , <) \ 0$.*

Proof. to be completed □

This means, uncertainty aversion of a decision maker is characterized by a strictly positive uncertainty premium for all Knightian lotteries: the decision maker would be willing to forego a positive amount of sure payoff for getting into a position of certainty where she receives for sure the mean payoff per state from the lottery minus the uncertainty premium. Equivalently, the certainty equivalent of the lottery, which makes the decision-maker indifferent between receiving this amount of payoff for certain or playing the uncertain lottery, is strictly smaller than the mean payoff per state. Both indicators express that the decision-maker is willing to forego something in order to gain certainty.

The two concepts of the certainty equivalent and the uncertainty premium of a Knightian lottery can also be used to compare the uncertainty attitudes of two different decision-makers. We want to know, who of the two is more uncertainty averse in the following sense.

Definition 9 (More uncertainty averse)
Consider two individuals A and B with preference relations $\succeq_A$ and $\succeq_B$, respectively, on the set Y of simple Knightian lotteries. A is said to be *at least as uncertainty averse as* B if and only if for all $y \in Y$ and all $c \in \mathbb{R}$ with $y \succeq_A c\underline{1}$, $y \succeq_B c\underline{1}$.

This means, A is said to be at least as uncertainty averse as B, if any time A prefers the uncertain lottery y over the certain payoff c, so does B. With this understanding of which decision-maker is more uncertainty averse, again, the certainty premium and the risk premium characterize who is more uncertainty averse.

Proposition 6 (Characterization of more uncertainty averse)
Consider two individuals A and B with preference relations $\succeq_A$ and $\succeq_B$, respectively, on the set Y of simple Knightian lotteries, and suppose the certainty equivalent uniquely exists for all $y \in Y$ for both preference relations. A is at least as uncertainty averse as B if and only if for all $y \in Y$, $C_A(y) < C_B(y)$ or, equivalently, $P_A(y) > P_B(y)$.

Proof. to be completed □

This means that if and only if A is at last as uncertainty averse than B, then for all Knightian lotteries A's certainty equivalent is smaller than B's, and A's uncertainty premium is greater than B's. With this proposition, one can use the certainty equivalent or the risk premium to also characterize the degree of uncertainty aversion of a single decision-maker. If and only if the certainty equivalent is lower, and risk premium is higher, for all Knightian lotteries, then this indicates a higher degree of uncertainty aversion.

With this, we can, for example, address the question: How does the degree of uncertainty aversion change with the sure wealth level at which uncertain lotteries are played? Suppose the decision-maker disposes of a wealth level of $w \in \mathbb{R}$ for certain and, on top of that, plays a lottery $y \in Y$. The resulting lottery is $w\underline{1} \oplus y$, which is equivalent to $w\underline{1} + y$ because wealth w is certain. If for all Knightian lotteries $y \in Y$, the uncertainty premium of $w\underline{1} + y$ decreases (does not change, increases) with w, the decision-maker is said to show *decreasing* (*constant*, *increasing*) *uncertainty aversion*.

Proposition 7 (Constant uncertainty aversion)
Consider a decision-maker with preference relation $\succeq$ on the simple-lottery set Y, such that the preference relation satisfies Axioms 1–7, and suppose that the uncertainty premium uniquely exists for all $y \in Y$. Then, for any Knightian lottery $y \in Y$ and for all wealth levels $w \in \mathbb{R}$,

$$P(w\underline{1} + y) = P(y) \ . \tag{10}$$

That is, a decision-maker with entropic preferences has constant uncertainty aversion.

Proof. See Appendix A.11 □

Constant uncertainty aversion is another consequence of the preferences being additively and extensively entropic. Proposition 7 is obtained without any assumption of convexity or concavity of the preference relation. It therefore holds for uncertainty aversion (positive uncertainty premium $P > 0$) as well as for uncertainty love (negative uncertainty premium $P < 0$).

5 A one-parameter function

We now propose a one-parameter real-valued preference function satisfying Axioms 1–9, which is based on Rényi's (1961) generalized entropy. The positive, real-valued parameter can be interpreted as the degree of uncertainty aversion. We subsequently illustrate the behavior of the preference function with a stylized decision problem between simple Knightian lotteries.

In this section, we let $Y = \mathbb{R}^n_{0+}$ so that the set of simple Knightian lotteries comprises all n-vectors with non-negative payoffs $y_i \geq 0$. With this setting, the universe of all Knightian lotteries naturally satisfies Assumptions 1–3.

Definition 10 (Rényi-entropic preference function)

The *Rényi-entropic preference function* on the set of all simple Knightian lotteries $H : Y \to \mathbb{R}$ is given by

$$H(y) := \frac{\overline{y}}{n} \frac{h_\alpha(s^y)}{h^{max}} \ , \tag{11}$$

where $h^{max} := \max_{s \in S} h_\alpha(s)$, and $h_\alpha : S \to \mathbb{R}$ for all $\alpha \geq 0$ is the Rényi entropy of order α with

$$h_\alpha(s) := \begin{cases} \frac{1}{1-\alpha} \ln \left(\sum\limits_{i=1}^{n} s_i^\alpha \right) & : \alpha \neq 1 \\ -\sum\limits_{i=1}^{n} s_i \ln s_i & : \alpha = 1 \end{cases} . \tag{12}$$

For compound lotteries, the preference function is extended to $H : Y \times Y \to \mathbb{R}$ with[19]

$$H(x \oplus y) := \frac{1}{nh^{max}} \left(\overline{x} h_\alpha(s^x) + \overline{y} h_\alpha(s^y) \right) \ . \tag{13}$$

Here, h is Rényi's (1961) generalized entropy function, which is well known from statistical physics and information theory. The expression for $h_1(s)$ is the continuous extension of the general expression of h_α for the limit $\alpha \to 1$. It has been proposed independently by Shannon (1948) and Wiener (1948). It is sometimes referred to as Shannon-Weaver-entropy because it has been popularized by Shannon and Weaver (1949).[20] Other notable special cases are $h_0(s) = \ln n$, the Hartley entropy (Hartley 1928), and $h_\infty(s) = \min_i\{-\ln s_i\} = -\ln(\max_i\{s_i\})$, which is also known as min-entropy.

The following lemma connects Rényi's generalized entropy function to the context of Knightian uncertainty.

Proposition 8 (Functional representation of preference relation)

The function H as defined in Definition 10 fulfills Axioms 1–8 for all $\alpha \geq 0$, and Axiom 9

[19] Again, we denote with the letter H both functions – the preference function on the set Y of simple lotteries (Equation 11) and the preference function on the set of all compound lotteries (Equation 13). There should not be any confusion, as it should be obvious in every instance from the argument of H which of the two functions we mean.

[20] The base of the logarithm used to calculate the entropies can be arbitrarily chosen. Rényi (1961) introduced his generalized entropy using the ld function, i.e. $\log_2$. Naturally, the choice of a particular base does not affect any result as long as the same base is used consistently.

for $0 < \alpha < 1$. *Hence, it uniquely (up to linear-affine transfomrations) represents the preference relation* $\succeq$.

Proof. See Appendix A.12. □

Thus, because Rényi's entropy fulfills Axioms 1–9, we can interpret it as one possible representation of the prefrence relation. Being one particular such function, it has all the properties that additive entropy functions have in general: continuity, monotonicity, additivity, extensivity and maximality (Proposition 1 and Corollary 1). The next lemma states some specific properties of Rényi's generalized entropy.

Proposition 9 (Properties of the Rényi-entropic preference function)
The Rényi-entropic preference function H_α *(Definition 10) has the following properties for all* $x,y \in Y$ *and all* $\alpha \geq 0$*:*

1. *Symmetry:* $H_\alpha(y) = H_\alpha(Py)$ *for every permutation matrix* P.
2. *Maximality:* $H_\alpha\left(\frac{\overline{y}}{n}\underline{1}\right) = \frac{\overline{y}}{n} > H_\alpha(y')$ *for all* $y' \in Y \setminus \left\{\frac{\overline{y}}{n}\underline{1}\right\}$.
3. *Minimality:* $H_\alpha(P(\overline{y},0,\ldots,0)) = 0 < H_\alpha(y')$ *for all* $y' \in Y \setminus \{P(\overline{y},0,\ldots,0)\}$ *and every permutation matrix* P.
4. *Additivity:* $H_\alpha(x \oplus y) = H_\alpha(x) + H_\alpha(y)$.
5. *Extensivity:* $H_\alpha(\lambda y) = \lambda H_\alpha(y)$ *for all* $\lambda > 0$.
6. *Concavity:* $H_\alpha(s)$ *is concave over* Y *for* $0 \leq \alpha \leq 1$ *and strictly so for* $0 < \alpha \leq 1$. *For* $\alpha > 1$ *it is neither concave nor convex in general but it is strictly quasi-concave. It is Schur-concave over* Y *for all* $\alpha \geq 0$ *and strictly so for* $\alpha > 0$
7. *Dependence on* α*:* $\frac{\mathrm{d}}{\mathrm{d}\alpha}H_\alpha(y) < 0$.

Proof. See Appendix A.13 □

The formal properties of Rényi's generalized entropy stated in Lemma **??** can be interpreted in terms of the preference relation. The symmetry property states that the sequence of the payoff shares that result from an act does not affect the value of h_α so that it does not matter in what sequence these shares are numbered. With regard to decision theory, this is a central assumption with regard to the decision maker's preferences, which will be discussed in greater detail in Section 6. For now, we just note that it implies that the Rényi individual is probabilistically sophisticated (Machina and Schmeidler 1992) with uniform subjective beliefs, i.e. the decision maker implicitly applies the Laplacian Principle of Insufficient Reason (Laplace 1820). The maximality property tells us that H_α reaches its unique maximum for a completely uniform distribution. This maximum value is equal to $\ln n$ and hence independent of α. As discussed in the context of Proposition 2 in Section 3.2, maximality means that the most preferred lottery is one with a completely uniform distribution of payoff over states. Any non-uniform distribution of payoffs over states will lead to a lower level of preference satisfaction than the uniform distribution. Conversely, lotteries where the total payoff volume is concentrated in just one state are generally least preferred (minimality). Any less extreme distribution gives a higher level of preference satisfaction. The last property, dependence on α, means that, for all any given lottery $y \in Y$, it holds that $h_0(s^y) > h_1(s^y) > \ldots > h_\infty(s^y)$. That is, for given lottery and, the greater the parameter α in the function, the smaller the resulting value of the preference index.

Corollary 3 (The Rényi-entropic preference function represents uncertainty aversion) *The Rényi-entropic preference function H_α (Definition 10) represents uncertainty aversion for $\alpha > 0$ and uncertainty neutrality for $\alpha = 0$.*

Proof. See Appendix A.14 □

The last property, dependence on α, is directly relevant for modelling uncertainty aversion. It means that, for any given act $f \in \mathcal{G}(\overline{y})$ of arbitrary dimension n, it holds that $h_0^n(s^f) > h_1^n(s^f) > \ldots > h_\infty^n(s^f)$. That is, for given act and, thus, uncertainty, the greater the parameter α in the utility function, the smaller the resulting value of the utility function. This leads to the following statement.

With a Rényi-entropic preference function $H(y)$ (Definition 10) and Definition 7 of the certainty equivalent, for any Knightian $y \in Y$ the certainty equivalent is given by

$$H(C(y)\underline{1}) = H(y) \tag{14}$$

$$C(y) = \frac{\overline{y}}{n}\frac{h_\alpha(s^y)}{h^{max}} = \frac{\overline{y}}{n \ln n} h_\alpha(s^y) \leq \frac{\overline{y}}{n}, \tag{15}$$

where Proposition 9 (Maximality) has been used for $H(C(y)\underline{1}) = C(y)$. Result (15) shows that the certainty equivalent of any lottery y is smaller than its total payoff volume divided by the number of states, $\overline{y}/n$, depending on the unevenness at which payoff is distributed over states. The latter is captured by the factor $h_\alpha(s^y)/h^{max}$, which varies between one for $s^y = (1/n,...,1/n)$ (perfect evenness, that is, perfect certainty) and zero for $s^y = P(1,0,...,0)$ (maximum unevenness, that is, maximum uncertainty). With this, the certainty equivalent is maximal (for given y) if payoff is perfectly certain, and it vanishes for maximum uncertainty of payoff.

In passing we note that, obviously, for any Rényi-entropic preference function $H(y)$, for any Knightian lottery $y \in Y$ the certainty equivalent $C(y)$ uniquely exists.

We now turn to the interpretation of the parameter α in the Rényi-entropic preference function $H(y)$ (Definition 10). This parameter captures the decision-maker's degree of uncertainty aversion.

Proposition 10 (α parameterizes the degree of uncertainty aversion)
Consider two individuals A and B whose preference relations $\succeq_A$ and $\succeq_B$ on the set Y of simple Knightian lotteries are represented by preference functions H_A and H_B (given by Equation 11) with parameters α_A and α_B, respectively. A is more (less) uncertainty averse than B (sensu Definition 9) if and only if $\alpha_A > (<)\alpha_B$.

Proof. See Appendix A.15 □

Illustration

As exemplary decision problem, we look at the following textbook example: an individual has to take a decision between three acts, f, g and h. The acts are known to generate the following payoffs (in monetary units):

$$\begin{aligned} x &= (300, 150, 250, 300)\,, \\ y &= (60, 60, 60, 820)\,, \\ z &= (15, 280, 340, 365)\,. \end{aligned}$$

Lottery x is very even but does neither have an especially large maximum payoff nor a particularly low minimum possible payoff. In fact, it guarantees the maximal minimum payoff out of the three alternatives. Hence, the maximin criterion would select lottery x. Lottery y offers the potentially highest possible win out of all three uncertain prospects but it only does so in one out of four possible states of the world whereas in the other three states, we end up having only 60 monetary units. Obviously, the maximax criterion would rate y highest and x lowest. Lottery z features the smallest minimum but otherwise it offers three potentially large payoffs as compared to x and y.

In Table 2, we illustrate our proposed index $H_\alpha^4(s^k)$ for the three Knightian lotteries x, y and z for different parameter values of uncertainty aversion α. In our framework, a comparison of differences in the uncertainty utility index is meaningful for the same individual due to Proposition 1 (uniqueness up to linear-affine transformations). We see that, although the preference over the acts always remains $x \succ z \succ y$, the overall level of well-being attached to a single act drastically depends on the degree of uncertainty aversion α. For example, an individual with a very low level of uncertainty aversion ($\alpha = 0.1$), the respective index values of H provided by the three uncertain prospects are within a range of 0.077 from act f (best) to act g (worst), whereas at high levels of uncertainty aversion (e.g. $\alpha = 50$), the difference is 1.016. Thus, an individual relatively uncaring towards Knightian uncertainty would gain relatively little in terms of preference satisfaction when switching from lottery g to lottery f. On the other hand, the very same switch would mean an over six times higher level of preference satisfaction to a highly uncertainty averse decision maker.

Comparison of decision rules

We now review well-established decision criteria under uncertainty from the literature and compare these to our framework using a stylized and static sample decision problem. The criteria from the literature include the maximin rule (Wald 1949) and its optimistic counterpart, the maximax rule, Laplace's principle of insufficient reason (Laplace 1820), the rule of minimum regret (Niehans 1948, Savage 1951) and the Hurwicz criterion

Table 2: H_α^4 scores of the three Knightian lotteries x, y and z for different degrees of uncertainty aversion α. The resulting preference ordering is $x \succ z \succ y$.

uncertainty aversion	uncertainty utility		
α	$H_\alpha^4(s^x)$	$H_\alpha^4(s^y)$	$H_\alpha^4(s^z)$
0.1	1.383	1.306	1.338
0.5	1.369	0.983	1.215
1	1.354	0.660	1.151
3	1.309	0.291	1.103
5	1.283	0.243	1.090
10	1.252	0.216	1.070
20	1.230	0.204	1.048
50	1.214	0.198	1.025

(Arrow and Hurwicz 1977) which is a linear combination of maximin and maximax which weighs possible maximum and minimum payoffs in each state according to the decision maker's optimism. The latter rule is sometimes also called α-maximin. Concise overviews can be found in Luce and Raiffa (1989) and Heal and Millner (2013).

In the following, we take a closer look at these rules. While terms such as maximin ('maximize the minimum over all possible acts') and maximax ('maximize the maximum over all possible acts') are self-explanatory, this is less true for the other three decision rules mentioned. Pierre-Simon Laplace's 1820 principle of insufficient reason[21] states that there is no reason to assume that one specific state of the world is more probable than another one when probabilities are unknown. Hence, they should all be given equal probability weight. Strictly speaking, Laplace's principle is thus a rule for assigning probabilities to outcomes and not a decision rule in itself. However, the wording 'Laplace principle' is often used synonymously with 'Laplacian expected utility', which refers to an expected utility maximizer applying the Laplace principle to calculate expected utility. The rule of minimum regret is based on the idea to minimize the maximum possible 'regret': for each possible state of the world, the act that leads to the highest

[21] The principle was renamed 'the principle of indifference' by Keynes (1921).

payoff is set as reference point relative to which the 'regret' is calculated as possible payoff that would potentially be foregone if the respective state of the world materialized. So, in the optimal case regret is zero. The alternative that minimizes possible regret over all states of the world is considered the best choice in this decision framework. Thus, 'regret' is quite similar to the concept of opportunity cost, but unlike its well-known sibling, it can attain a value of zero. Moreover, focusing on minimizing a negative quantity rather than maximizing a positive one, the rule expresses a very cautious, if not pessimistic attitude of the decision maker towards uncertainty. Mathematically, the rule of minimum regret – sometimes also referred to as 'Savage-Niehans rule' – is to choose the act k from $\mathcal{G}(\overline{y})$ which minimizes possible 'regret', i.e. which minimizes the expression

$$\sum_{i=1}^{n} \left[(\max_k u(y_i^k)) - u(y_i^k) \right] \ . \tag{16}$$

Eventually, the Hurwicz rule generalizes the maximin and maximax criteria: for each alternative k, the function

$$\Phi(y^k) = \lambda \max_i \{u(y_i^k)\} + (1-\lambda) \min_i \{u(y_i^k)\}, \quad 0 \leq \lambda \leq 1 \tag{17}$$

is evaluated and compared to the function values of the alternatives. The associated decision rule is $\max_k \Phi(y^k)$. λ thus reflects the individual's optimism as a greater λ gives more weight to the maximum payoff of one particular act k and hence less weight to the minimum. Similarly, Heal and Millner (2013) give an interpretation of $1 - \lambda$ as representing 'aversion to a lack of knowledge'. Hence, with $\lambda = 1$, we recover the maximax rule while $\lambda = 0$ leads again to the maximin criterion.

As exemplary decision problem, we look at the following textbook example: an individual has to take a decision between three acts, f, g and h. The acts are known to generate the following payoffs (in monetary units):

$$
\begin{aligned}
x &= (300, 150, 250, 300)\,, \\
y &= (60, 60, 60, 820)\,, \\
z &= (15, 280, 340, 365)\,.
\end{aligned}
$$

Lottery x is very even but does neither have an especially large maximum payoff nor a particularly low minimum possible payoff. In fact, it guarantees the maximal minimum payoff out of the three alternatives. Hence, the maximin criterion would select lottery x. Lottery y offers the potentially highest possible win out of all three uncertain prospects but it only does so in one out of four possible states of the world whereas in the other three states, we end up having only 60 monetary units. Obviously, the maximax criterion would rate y highest and x lowest. Lottery z features the smallest minimum but otherwise it offers three potentially large payoffs as compared to x and y. A risk-neutral Laplace individual would be indifferent between the three lotteries, while a risk-averse one would prefer x. The rule of minimum regret would lead to the choice of x while z and y would be tied. The advice that the Hurwicz criterion gives us critically depends on the choice of λ. A rather pessimistic individual ($\lambda = 0.1$) would choose lottery x while for any $\lambda \geq \frac{9}{61}$, lottery y would be preferred. This choice is made by any sufficiently optimistic individual – i.e. $\lambda \geq \frac{9}{61}$ – and as λ is further increased, we can observe a change of the second most preferred lottery from y to z. The complete rankings of acts are given in Table 3.

In summary, combining the results from Tables 3 and 2, we find that the overall ranking of lotteries from our method is different from the other criteria. However, the most preferred option is the same as with the maximin rule and a pessimistic Hurwicz individual. That being said, a risk-averse Laplacian individual would arrive at the very same ranking as our Rényi individual, but this comes with the above disclaimer.

Table 3: Orderings over the Knightian lotteries x, y and z that result from different decision rules from the literature.

decision criterion	choice ordering
maximin	$x \succ y \succ z$
maximax	$y \succ z \succ x$
risk-neutral Laplace EU	$x \sim y \sim z$
risk-averse Laplace EU	$x \succ z \succ y$
minimum regret	$x \succ y \sim z$
Hurwicz, $\lambda = 0.1$	$x \succ y \succ z$
$\lambda = 0.2$	$y \succ x \succ z$
$\lambda = 0.8$	$y \succ z \succ x$

6 Discussion

We now discuss some key features of our approach, and relate it to other approaches in the literature.

Summary. Based on a set of seven axioms on the preference relation over Knightian acts, we have provided a proof that there exists a real-valued, unique (up to linear-affine transformations), additive and extensive representation of this preference relation on the set of all Knightian acts. Moreover, we have shown that this function may represent uncertainty aversion.

We have illustrated our approach with a suitable one-parameter function from information theory fulfilling all axioms of certainty measurement – Rényi's (1961) generalized entropy. The parameter can be interpreted as the relative weight at which the two fundamental sources of uncertainty are taken into account in the aggregate measure of uncertainty: (1) the pure number of potential states of nature, and (2) the heterogeneity of the payoff-distribution over the given number of states of nature.

Finally, we have compared our approach in a simple decision problem to other methods from the literature. We have found that our certainty measure produces a ranking different from the other decision rules under Knightian uncertainty. However, the most certain act coincides with the one preferred by an individual with maximin preferences and with the choice a very pessimistic Hurwicz individual would make.

Comparison to the vNM-approach to expected utility. From seven axioms we obtain a preference function. In vNM, with three axioms one has an expected-utility form of the preference function. But to have a preference function fully specified, one needs – in addition – a Bernoulli utility function. Some of the properties of the preference function derive from the properties of the Bernoulli utility fcn (e.g. non-satiation or risk-aversion). In our approach, we obtain the full preference function from a set of seven to nine basic axioms.

Our axioms, which establish unique existence of preference function, imply independence (which does not need to be assumed)

They also imply two properties, additivity and extensivity, as well as monotonicity, which in the vNM-approach needs to be assumed separately.

(In-)Completeness of preferences. The arguably strongest assumption required to bring our framework to life is Axiom 3. It assumes completeness only on the subset $\mathcal{G}(\overline{y})$ of all simple acts Y. Thus, we do not assume that any two arbitrary Knightian acts from Y can be ranked in terms of certainty, i.e. they may be incomparable. Rather, we require the certainty relation to be complete only on the subset of Knightian acts with the same total payoff volume $\overline{y}$, $\mathcal{G}(\overline{y})$ over all states. Indeed, we think that this kind of completeness on a subset is in fact normatively much more compelling than assuming completeness on the full set of all acts Y, even more so in the case of Knightian uncertainty. In fact, the implications of the completeness assumption for economic theory have been vividly discussed from the outset. Von Neumann and Morgenstern (1944) themselves considered it 'very dubious whether the idealization which treats this postulate as a valid one, is appropriate or even convenient' (ibid.: p. 630). Others like Luce and Raiffa (1957) criticized the possibility of intransitivities if individuals exacted decisions between alternatives that might be 'inherently incomparable'. In the same vein, R.J. Aumann (1962, 1964) doubted the normative appeal of an an a-priori exclusion of the possibility of an individual to be unwilling or unable to arrive at preference statements for at least some acts. In our view, Aumann's point was that the inability to state one's preferences regarding a decision might be the result of rational thinking and judgment, so there is no reason to make completeness of preferences a standard of any rational choice theory, an argument ultimately very similar to Putnam (1986). Despite all this, the only other contribution we know of that discusses and uses incompleteness of preferences in the context of Knightian uncertainty is Bewley (2002). He replaces completeness with an inertia assumption, which states that an alternative is only accepted if preferred to the status quo. Moreover, an individual may assert that two alternatives are incomparable. The main difference to our approach however is that Bewley works within the Anscombe-Aumann framework, which relies on subjective and objective probabilities, a concept we have deliberately avoided here for reasons laid out earlier in this paper.

The role of $\overline{y}$ and its intuitive appeal. The sum of payoffs $\overline{y}$ over all possible states of the world plays a central role in our theory. One might wonder about desirability and intuitive appeal of this feature. First and foremost, the issue is closely related to the above question of whether or not to assume (in)completeness of the certainty relation. We argue in favor of incompleteness here, but it is clear that it naturally comes at a price. From a technical point of view, the problem with an incomplete realtion is that one cannot have a complete representation either, so one might run into issues with dominance. On the other hand, for each $\overline{y} > 0$, $\mathcal{G}(\overline{y})$ is the largest possible set of acts which cannot possibly dominate each other, so it seems natural to build a theory around this. As to intuition, consider the possibility that you are promised a slice of your favorite pie tomorrow, but the size of the slice will depend, for some reason, on (1) what you do today, and (2) what state of the world materializes tomorrow. Even before you worry about your options to act and single outcomes, the most natural question seems 'How large is that pie anyway?'. Only as a second step will you probably think about what is best to do given the possible – fundamentally uncertain – states of the world.

Probabilistic sophistication. A central feature of the Rényi preference representation from Section 5 is the symmetry of $H(s)$ (cf. Lemma **??**), which implies that the sequence in which the given states of nature are numbered does not matter for measuring (un)certainty. The only thing that matters is the distribution of payoffs over all states, not the exact states of the world in which particular payoffs are obtained. This property has been termed 'event exchangeability' (Chew and Sagi 2006: 771), and it interestingly implies the individual to be *probabilistically sophisticated* (Machina and Schmeidler 1992) with uniform subjective beliefs. That is, because any two outcomes of any act are exchangeable for a Rényi individual, the concrete choice of the Rényi function implies that the decision maker follows Laplace's principle of insufficient reason in a non-expected utility framework. In that respect, our result might be one step towards a result that parallels the contribution of Gravel, Marchant and Sen (2012), who provided a complete axiomatic foundation of Laplace's principle in the expected utility setting.

Connection to Laplacian expected utility. In the following, we compare the standard expected utility (EU) framework to our approach. Starting from expected utility with concrete utility function $u(y) = \ln y$, and assuming that we have an individual that applies Laplace's principle, we can generally say that the ranking of acts will coincide with a ranking done by our Rényi individual. We make this explicit in Appendix A.16. In that particular case, the connection between Laplacian EU and our framework can be established, because a strictly increasing monotonous transformation from the EU functional to the Rényi functional can always be constructed. The situation is however not clear for general utility functions $u(y)$ or representations of H other than the one presented here, so a general correspondence between these two frameworks cannot be established here, and its existence seems questionable to us. However, the positive message to be learned from this observation is that the theory developed here is not completely detached from other theories, but much rather shares a boundary point with them.

(Non-) Additive entropies in economics. We have seen that our axiomatic foundation of certainty measurement allows for the existence of an additive certainty representation, and that such functions are called additive entropies in physics and information theory. For completeness, we should say that entropies have been used before in economic theory. Luce, Ng and Marley (2008) have proposed to use entropies to model the utility which individuals derive from the process of gambling itself. This approach seems interesting, because it departs from the typical consequentialist setting in economics, in that it matters to the decision maker how an outcome is obtained rather than just caring about what outcome is obtained. However, this is quite a different use of the concept of entropy, and unrelated to our concept presented here. Second, there are also non-additive entropies, such as the Tsallis entropy (Tsallis 1988). Non-additivity is interesting from the point of view of ranking and valuing compound acts. With additivity, any compound act is just as certain as the sum of numerical certainty values obtained from its constituent acts (cf. Proposition 2). Non-additivity would enable us to model situations where compound acts are more or less certain than just the sum of their constituents. While such non-additive certainty relations, represented by non-additive entropy functions, would be an interesting object of study, we do not know of a formal axiomatization of that.

7 Conclusion and Outlook

In a nutshell, we have shown how a parsimonious set of nine axioms on the preference relation over Knightian lotteries establishes the existence and uniqueness of a preference function which is additive and extensive, and we have provided a one-parameter function as an example. From here, several research fields open up.

First, turning to *preferences over lotteries under Knightian uncertainty*, there is the problem field of how to conceptualize and represent such preferences, and how to measure a decision-maker's type and degree *uncertainty aversion*, theoretically as well as experimentally. Theoretically, one could think of a transfer of concepts from general relativity (Einstein, Minkowski), where measuring curvature that is invariant under certain coordinate transformations in multi-dimensional spaces is formalized, so that a stronger "curvature" of the preference function leads to a higher degree of uncertainty aversion. Experimentally, it would be interesting to assess uncertainty attitudes in different settings and contexts.

Second, based on our concept of uncertainty preferences and uncertainty aversion one may think of *insurance* against Knightian uncertainty as an institution that reduces the uncertainty for all acts. Likewise, one may generalize the concepts of self-insurance and self-protection (Ehrlich and Becker 1972) to situations of Knightian uncertainty. Such generalized insurance would be not based on probabilities, but – given preferences under Knightian uncertainty, or at least some measure of uncertainty aversion – could still be valued in monetary terms, so that it can be brought into the market-context.

Third, as our approach to measuring (un)certainty is centrally build on entropy, and entropy generally measures the heterogeneity/homogeneity of a distribution, our axiomatic framework of measuring (un)certainty, as well as the example of the Rényi-function, may be directly transferred to *measuring the heterogeneity/homogeneity of other kinds of distributions of economic relevance* – e.g. income (in)equality, product diversity, technological diversity, or institutional diversity.

Acknowledgments

We thank Dave Abson, Simon Grant, Nikolai Hoberg, Mark Machina, Jack Stecher, Martin Quaas, Thorsten Upman and Ralph Winkler for critical discussion and helpful comments on earlier versions of this manuscript. Financial support under grant no. 01LL0911E from the German Federal Ministry of Education and Research is gratefully acknowledged.

Appendix

A.1 Proof of Lemma 1

Lemma 1 is identical – up to the re-naming of concepts and a change in notation – to Theorem 2.1 in Lieb and Yngvason (1999: 22). Their proof (Lieb and Yngvason 1999: 22) therefore also proves our lemma.

A.2 Axiom 9 implies normal convexity of $\succeq$

Assume $x \sim y$. By Axiom 5, $\lambda x \sim \lambda y$ and $(1-\lambda)x \sim (1-\lambda)y$. Then, by Lemma 1, $\lambda x \oplus (1-\lambda)y \sim \lambda x \oplus (1-\lambda)x$ and $\lambda x \oplus (1-\lambda)y \sim \lambda y \oplus (1-\lambda)y$. By Axiom 6,

$\lambda x \oplus (1-\lambda)x \sim x$ and $\lambda y \oplus (1-\lambda)y \sim y$. Hence, $\lambda x \oplus (1-\lambda)y \sim \lambda x \oplus (1-\lambda)x \sim x$ and $\lambda x \oplus (1-\lambda)y \sim \lambda y \oplus (1-\lambda)y \sim y$. With that, $\lambda x + (1-\lambda)y \succeq \lambda x \oplus (1-\lambda)y$ implies $\lambda x + (1-\lambda)y \succeq x$ and $\lambda x + (1-\lambda)y \succeq y$.

A.3 Proof of Proposition 1

Proposition 1 is identical – up to the re-naming of concepts and a change in notation – to Theorem 2.2 in Lieb and Yngvason (1999: 24). Their proof (Lieb and Yngvason 1999: 24–29) therefore also proves our proposition.

A.4 Proof of Corollary 1

Follows directly from Proposition 1 by setting $N = M = 1$ and $\mu_1 = \lambda_1$.

A.5 Proof of Proposition 2

Our proposition 2 is included – with re-naming of concepts and a change in notation – in Theorem 2.5 in Lieb and Yngvason (1999: 30). Their proposition is much more general, though, as they consider several simple-lottery sets (in our language) and the relation among them. In contrast, we only consider a single simple-lottery set Y. With this simplification, requirements (i) and (ii) in their Theorem 2.5 are captured by our Assumption 3, and their requirement (iii) is included in our Axiom 3 (completeness). Also, their distinction between a simple-lottery-set-specific entropy function and a universal entropy function, which is formed from the former by suitable calibration of the constants a and b, is irrelevant for our proposition. While in the Lieb-Yngvason treatment, additivity and extensivity only hold as a property of the suitably calibrated universal entropy function, in our setting they hold immediately for the entropy function on Y as introduced by Proposition 1. Their proof of their Theorem 2.5 (Lieb and Yngvason 1999: 30–31) directly proves our proposition.

A.6 Proof of Corollary 2

(1) Scaling-monotonicity. With $\lambda > 1$ one has (from Extensivity, Proposition 2) $H(\lambda y) = \lambda H(y) > H(y)$, or $\lambda y \succ y$.

(2) Adding-monotonicity. With the compounding property $y + c\underline{1} \sim y \oplus c\underline{1}$ one has (due to Additivity and Extensivity of H, Proposition 2)

$$H(y + c\underline{1}) = H(y \oplus c\underline{1}) = H(y) + H(c\underline{1}) = H(y) + cH(\underline{1}) \ . \tag{A.1}$$

As H is unique only up to linear-affine transformations (Proposition 1), one is free to set $H(\underline{1})$ – by choice of the additive constant in a linear-affine transformation of H – to any arbitrary real value. Here, we set $H(\underline{1}) > 0$. Then, as $c > 0$ as well, one has $cH(\underline{1}) > 0$ and Equation (A.1) implies $H(y + c\underline{1}) > H(y)$, or $y + c\underline{1} \succ y$.

A.7 Proof of Proposition 3

(1) Symmetry: Both operations, scaling (Definition 1) and compounding (Definition 2, are not affected by permutating the underlying payoff vectors. Also, none of the Axioms 1–7 and 9 is affected by permutating the underlying payoff vectors. Assuming symmetry of the relation $\succeq$ (Axiom 8), therefore, does not affect any other result, such as e.g. Propositions 1 or 2. It comes in addition and its effect can be studied on top of everything else. With that, Axiom 8 (symmetry of $\succeq$) and symmetry of the function H are obviously equivalent.

(2) Concavity: Proposition 3 is identical – up to the re-naming of concepts and a change in notation – to Theorem 2.8 in Lieb and Yngvason (1999: 34). Their proof (Lieb and Yngvason 1999: 35) therefore also proves our proposition.

A.8 Proof of Lemma 3

Any simple Knightian lottery $y \in Y$ uncertainty-dominates its pc-corresponding lottery y^c, and strictly so if and only if $y \neq y^c$ (by definition of y^c). If the decision-maker is uncertainty averse (sensu Definition 5), she strictly prefers to any $y \in Y$ all those lotteries that are strictly uncertainty-dominated by y. As y^c is strictly uncertainty-dominated by y (for $y \neq y^c$), she strictly prefers y^c to y. Likewise, if the decision-maker is uncertainty loving (neutral), the preference direction is reversed (indifferent).

A.9 Proof of Proposition 4

If Y satisfies Assumption 2 and $\succeq$ satisfies Axiom 8, then H is symmetric on Y (Proposition 3(1)). Further, if Y satisfies Assumption 1 and $\succeq$ satisfies Axiom 9, then H is symmetric on Y (Proposition 3(2)). Any function that is symmetric and concave is Schur-concave (Lemma 5). Hence, H is Schur-concave if Y satisfies Assumptions 1 and 2, and $\succeq$ satisfies Axioms 8 and 9. By Lemma 4, a preference function that is Schur-concave represents uncertainty aversion. Hence, H represents uncertainty aversion under these conditions.

A.10 Proof of Lemma 6

Condition (5), which defines the certainty equivalent of y, can be stated equivalently in terms of H as $H(C(y)\underline{1}) = H(y)$. By extensivity of H (Proposition 2), this can be restated as $C(y)H(\underline{1}) = H(y)$, which can be rearranged into Equation (6). Obviously, for given preference function H this yields a result for any $y \in Y$ (existence), and exactly one result (uniqueness).

A.11 Proof of Proposition 7

The uncertainty premium P of lottery $w\underline{1} + y$ is defined through Relation (8):

$$\left(\frac{\overline{w\underline{1}+y}}{n} - P(w\underline{1}+y)\right)\underline{1} \sim w\underline{1}+y \ . \tag{A.2}$$

Here, $\overline{w\underline{1}+y} = nw + \overline{y}$ such that

$$\left(\frac{\overline{w\underline{1}+y}}{n} - P(w\underline{1}+y)\right) = \left(w + \frac{\overline{y}}{n} - P(w\underline{1}+y)\right) \ . \tag{A.3}$$

With this, Relation (A.2) becomes

$$\left(w + \frac{\overline{y}}{n} - P(w\underline{1} + y)\right)\underline{1} \;\sim\; w\underline{1} + y \tag{A.4}$$

$$w\underline{1} + \left(\frac{\overline{y}}{n} - P(w\underline{1} + y)\right)\underline{1} \;\sim\; w\underline{1} + y \tag{A.5}$$

$$w\underline{1} \oplus \left(\frac{\overline{y}}{n} - P(w\underline{1} + y)\right)\underline{1} \;\sim\; w\underline{1} \oplus y \tag{A.6}$$

$$\left(\frac{\overline{y}}{n} - P(w\underline{1} + y)\right)\underline{1} \;\sim\; y\ , \tag{A.7}$$

where Relations (A.4) and (A.5) are equivalent because of normal vector algebra, Relations (A.5) and (A.6) are equivalent because of the compounding property $y \oplus x^c \sim y + x^c$, and Relations (A.6) and (A.7) are equivalent because the preference relation $\succeq$ satisfies Independence (Lemma 1) with respect to the 'sure thing' $w\underline{1}$. From Relation (A.6) and Relation (8), one has (because of Transitivity of $\succeq$, Axiom 2)

$$\left(\frac{\overline{y}}{n} - P(w\underline{1} + y)\right)\underline{1} \sim \left(\frac{\overline{y}}{n} - P(y)\right)\underline{1}\ . \tag{A.8}$$

This is equivalent, because of scaling-monotonicity of $\succeq$ (Corollary 2, Statement 1), to $P(w\underline{1} + y) = P(y)$.

A.12 Proof of Proposition 8

We take the Rényi-entropic preference function from Definition 10 and suppose it represents the preference relation $\succeq$, such that $H(x) \geq H(y)$ is equivalent to $x \succeq y$ for all $x,y \in Y$. We then show that each of the Axioms 1–9, as expressed through H, makes a true statement. Knowing that the function H from Definition 10 fulfills all axioms, and knowing (from Proposition 1) that the representing preference function is unique up to linear-affine transformations, we then also know that the function H from Definition 10 is – up to linear-affine transformations – the only function that represents $\succeq$.

Axiom 1

to be completed

A.13 Proof of Proposition 9

Rényi's generalized entropy function $h_\alpha : S \to \mathbb{R}$ (Equation 12) is known, or can easily be shown, to have the following properties for all $s \in S$ and all $\alpha \geq 0$:

1. Symmetry: $h_\alpha(s) = h_\alpha(Ps)$ for every permutation matrix P (follows directly from symmetry of the summation).
2. Intensivity: $h_\alpha(s^y) = h_\alpha(s^{\lambda y})$ for all $\lambda > 0$ (follows directly from $s^{\lambda y} = \lambda^y$).
3. Maximality: $h_\alpha\left(\frac{1}{n}\underline{1}\right) = \ln n > h_\alpha(s)$ for all $s \in S \setminus \left\{\frac{1}{n}\underline{1}\right\}$ (Rao 1984: 70).
4. Minimality: $h_\alpha(P(1,0,\ldots,0)) = 0 < h_\alpha(s)$ for all $s \in S \setminus \{P(1,0,\ldots,0)\}$ and every permutation matrix P (Rao 1984: 70).
5. Concavity: $h_\alpha(s)$ is concave over S for $0 \leq \alpha \leq 1$ and strictly so for $0 < \alpha \leq 1$ (Rao 1984: 70, He et al. 2003). For $\alpha > 1$ it is neither concave nor convex in general but it is strictly quasi-concave. Hence, $h_\alpha(s)$ is Schur-concave over S for all $\alpha \geq 0$ and strictly so for $\alpha > 0$ (Marshall et al. 2011, Pliam 2013).
6. Dependence on α: $\frac{\mathrm{d}}{\mathrm{d}\alpha} h_\alpha(s) < 0$ (Beck and Schlögl 1993).

With these properties of h_α, the properties of H are demonstrated as follows.

A.14 Proof of Corrolary 3

For all $\alpha > 0$, $H(y)$ is strictly Schur concave on Y (Proposition 9, Statement 6). With Proposition ???, this means that H represents uncertainty aversion.

In contrast, for $\alpha = 0$ one has $h_\alpha(y) \to \ln n = h^{max}$ and, hence, $H(y) \to \overline{y}/n$. In this case, $H(y)$ is not strictly concave on Y, but it is both concave and convex. For any $y \in Y$ only the total payoff volume $\overline{y}$ matters for the preference ranking, and the relative (un)evenness of the payoff distribution s^y is irrelevant. This represents uncertainty neutrality.

A.15 Proof of Proposition 10

With a Rényi-entropic preference function $H(y)$ (Definition 10), for any $y \in Y$ the certainty equivalent of lottery y is given by Equation (15). In the expression on the right-hand side, the factor $\xi := \overline{y}/(n \ln n)$ is greater than zero (as $n \geq 2$) and independent of

α. Hence,

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}C(y) = \xi\frac{\mathrm{d}}{\mathrm{d}\alpha}h_\alpha(s^y) < 0 \quad \text{for all } y \in Y \ ,$$

because of $dh_\alpha/d\alpha < 0$ (Appendix A.13, Property 6 of Rényi's entropy). This is equivalent to

$$C_A(y) < (>) C_B(y) \quad \text{for all } y \in Y \quad \text{if and only if} \quad \alpha_A > (<) \alpha_B \ .$$

With Proposition 6 (Statement 1), this establishes the proposition.

A.16 Connection to expected utility

If $u(y) = \ln y$, then there is a direct correspondence between a Laplace-EU individual and a Rényi individual. To see this, consider a two-state act with one good outcome, y_H and one bad outcome y_L. The Rényi functional then reads

$$H = \frac{1}{1-\alpha}\ln\left[\left(\frac{y_L}{y_L+y_H}\right)^\alpha + \left(\frac{y_H}{y_L+y_H}\right)^\alpha\right] \tag{A.9}$$

while the Laplacian expected utility reads

$$\mathbb{E}U = \frac{1}{n}(\ln y_L + \ln y_H) = \frac{1}{n}\ln(y_L \cdot y_H) \tag{A.10}$$

For the relative ranking of any two acts, the ln functions matter. The question is thus, whether there exists a strictly monotonous, i.e. order preserving, transformation

$$T : y_L \cdot y_H \mapsto \left(\frac{y_L}{y_L+y_H}\right)^\alpha + \left(\frac{y_H}{y_L+y_H}\right)^\alpha \tag{A.11}$$

At least one such transformation exists:

$$T : x \mapsto \frac{(\frac{x}{y_H})^\alpha + (\frac{x}{y_L})^\alpha}{(\frac{x}{y_L} + \frac{x}{y_H})^\alpha} \tag{A.12}$$

It holds that $T'(x) > 0$ for all $\alpha > 0$ and such a transformation can be constructed for all $n > 2$ as can be seen by complete induction. It is however unclear whether such a transformation can be found for any $u(y)$ and any H.

References

Al-Najjar, N.I. and J. Weinstein (2009), The ambiguity aversion literature: a critical assessment, *Economics and Philosophy* 25, 249–284

Allais, M. (1953), Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine, *Econometrica* 21(4), 503–546

Anscombe, F.J. and R.J. Aumann (1963), A definition of subjective probability, *The Annals of Mathematical Statistics* 34, 199–205

Arnold, B.C. (1987), *Majorization and the Lorenz Order: A Brief Introduction*, Lecture Notes in Statistics, Vol. 43, Springer, Berlin.

Atkinson, A.B. (1970), On the measurement of inequality, *Journal of Economic Theory* 2, 244–263.

Beck, C. and T. Schlögl (1993), *Thermodynamics of Chaotic Systems: An Introduction*, Cambridge University Press, Cambridge, UK.

Chavas, J.-P. (2004), *Risk Analysis in Theory and Practice*, Elsevier Academic Press, Amsterdam.

Dalton, H. (1920), The measurement of the inequality of incomes, *The Economic Journal* 30(119): 348–361.

Gilboa, I. and D. Schmeidler (1989), Maxmin expected utility with non-unique prior, *Journal of Mathematical Economics* 18, 141–153

Hardy, G.H., J.E. Littlewood and G. Pólya (1934/1952), *Inequalities*, 1st ed. 1934, 2nd ed. 1952, Cambridge University Press, London and New York.

He, Y., A.B. Hamza and H. Krim (2003), A generalized divergence measure for robust image registration, *IEEE Transactions on Signal Processing* 51(5), 1211–1220

Keynes, J.M. (1921), *A Treatise on Probability*, Macmillan, London

Knight, F. (1921), *Risk, Uncertainty, and Profit*, Houghton Mifflin Company, The Riverside Press, Cambridge, MA

Laplace, P.S. (1820), *Théorie Analytique des Probabilités*, Mme Ve Courcier, Paris

Lieb, E. and J. Yngvason (1999), The physics and mathematics of the second law of thermodynamics, *Physics Reports* 310, 1–96

Luce, R.D. (1972), Conditional expected, extensive utility, *Theory and Decision* 3(2), 101–106

Luce, R.D., C.T. Ng and A.A.J. Marley (2008), Utility of gambling I: entropy modified linear weighted utility, *Economic Theory* 36, 1–33

Machina, M. and D. Schmeidler (1992), A more robust definition of subjective probability, *Econometrica* 60, 745–780

Marshall, A.W., I. Olkin and B.C. Arnold (2011), *Inequalities: Theory of Majorization and Its Applications*, second edition, Springer, Berlin.

Pigou, A.C. (1912), *Wealth and Welfare*, Macmillan, London.

Pliam, J.O. (2013), Alternating product ciphers: A case for provable security comparisons, in G. Paul and S Vaudenay (eds), *Progress in Cryptology - INDOCRYPT 2013*, Springer, New York.

Rao, C.R. (1984), Convexity properties of entropy functions and analysis of diversity, *Inequalities in Statistics and Probability* 5, 68-77.

Rényi, A. (1961), On measures of entropy and information, in J. Neyman (ed.), *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, University of California Press, Berkeley, CA, pp. 547–561

Rothschild, M. and J. Stiglitz (1970), Increasing risk I: A definition, *Journal of Economic Theory* 2, 225–243.

Savage, L.J. (1954), *The Foundations of Statistics*, Wiley and Sons, New York, NY, 2nd revised and enlarged edition (1972), Dover Publications, New York, NY

Schur, I. (1923), Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinanten-Theorie, *Sitzungsberichte der Berliner Mathematischen Gesellschaft* 22, 9–20