A CONCEPTUAL FOUNDATION FOR THE THEORY OF RISK AVERSION

YONATAN AUMANN

Bar Ilan University

Ramat Gan, Israel

Abstract. Classically, risk aversion is equated with concavity of the utility function. In this paper we explore the conceptual foundations of this definition. In accordance with neo-classical economics, we seek an ordinal definition of risk aversion, based on the decisions maker's preference *order* alone, independent of numerical values. We explore two such definitions. We then show that when cast in quantitative form these ordinal definitions coincide with the classical Arrow-Pratt definition once the latter is defined with respect to the appropriate scale (which, in general is not money), thus providing a conceptual foundation for the classical definition. The implications of the theory are discussed, including, in particular, to defining risk aversion for non-monetary goods, and to disentangling risk aversion from diminishing marginal utility. The entire study is within the expected utility framework.

Keywords: Risk aversion, Utility theory, Ordinal preferences, Multiple objectives decision making.

1. Introduction

1.1. **Risk Aversion - The Classic Approach.** The concept of *risk aversion* is fundamental in economic theory. Classically, it is defined as an attitude under which the certainty equivalent of a gamble is less than the gamble's expected value; e.g., if a decision maker prefers one unit with certainty over a fair gamble between three units and none, then she is deemed risk averse.

Examining this core definition, two fundamental questions arise, the combination thereof drives this work.

First, there is the matter of scale. Consider a decision maker having to choose between lotteries on the temperature-level in her office room. If she prefers 40° F with certainty over a fair gamble between $30^\circ$and 60°– should this be considered risk aversion? The Fahrenheit scale seems rather arbitrary in this case, but it is not clear what other scale should or can be used. In the seminal works of Arrow [2] and Pratt [24], risk aversion was defined with respect to money and the market value of the goods. This, however, limits the notion to monetary (or liquid) goods. A core question is thus if and how risk aversion can be defined for non-monetary goods - temperature, health, love, pain, and the like.

E-mail address: aumann@cs.biu.ac.il.

Date: November 14, 2016.

The second question is a conceptual one. The classic definition of risk aversion seems to be based on the presumption that the natural certainty equivalent of a gamble is its expectation, and risk aversion is defined with respect to this natural certainty equivalent. The question, however, is why this presumption? Clearly, it cannot rest on empirical evidence, as most people are assumed to be risk averse. The justification must be conceptual. But on a conceptual level, it is not clear what reasoning dictates that a fair gamble between \$100 and \$200 "should" be worth \$150; From a decision theoretic perspective, where preference orders and independence curves are the core elements of interest, what is the significance of the arithmetic mean? Providing a conceptual, decision theoretic justification for basing the definition of risk aversion on the arithmetic mean is the second core goal of this paper.

1.2. **An Ordinal Foundation.** In order to address the above questions, we start by seeking a fully ordinal definition of risk aversion, independent of *any* units, and making no use of arithmetic notions such as mean or expectation. We consider two such definitions, as outlined shortly. Both definitions are fully ordinal, based solely on the internal structure of the decision maker's preferences. Having defined risk aversion in purely ordinal terms, we then *derive* a quantitative/numeric form of these definitions. This quantitative form, we show, coincides with the classic Arrow-Pratt definition, once the latter is defined with respect to an appropriate, natural scale. This scale, which in general is *not* money, applies to *any* goods - monetary or non-monetary. Thus, we provide the missing conceptual justification for the use of the expectation as the baseline for defining risk aversion, and determine the "appropriate" scale.

Ordinal Definition I: Repeated Gambles. Consider a lottery L with certainty equivalent c. Arguably, the most extreme form of risk aversion would be exhibited if, with probability 1, the certainty equivalent is inferior to the outcome of the lottery. If that is the case then the decision maker is willing to pay a premium, with certainty, merely to avoid being in an uncertain situation. Such a preference, however, is ruled out by the von Neumann-Morgenstern (NM) axioms; the utility of a lottery must lie between the utilities of its possible outcomes. Interestingly, while such a preference is indeed not possible for any single lottery, it *is* possible once we consider *sequences* of lotteries, and risk aversion as a *policy* - consistently adhered to over multiple gambles. We show that for some preference orders (agreeing with the NM axioms), repeatedly choosing the certainty equivalent of a lottery over the lottery itself can result in an outcome that is inferior to what would have been the outcome of the lotteries, with probability 1. This is thus our ordinal definition of risk aversion: a preference order is deemed risk-averse if adhering to this preference order over repeated lotteries ultimately results in an inferior outcome, with probability 1. Importantly, here "inferior" is according to the decision maker's own preference order, over sequences, not any external market-based criterion. The details of the definition are provided in Section 3.

The above definition requires the consideration of repeated lotteries. The next definition considers the "one-shot" case.

Ordinal Definition II: Hedging. The second definition we consider is that of Richard [25]. Consider two commodities/time-periods[1] and assume that the *certainty* preferences on the commodities are independent; that is, the certainty preferences on each commodity separately are well defined. Then, Richard's definition of (multivariate) risk aversion is the following:

> Let a, A, be two states of one commodity and b, B, two states of the other commodity, with $a \prec A$ and $b \prec B$. Then, a risk averse decision maker always prefers the fair gamble between (A, b) and (a, B) - the gamble wherein the loss in one commodity is (somewhat) hedged against a win in the other - over the fair gamble between the two extreme outcomes (a, b) and (A, B).[2]

Here, risk aversion is equated with a preference for hedging bets, whenever and to the extent possible.

The above definition is stated in terms of two specific commodities, but can also be applied to a partition of the entire space into *independent factors* (collections of commodities for which the certainty preference are well defined), and hedging takes place between the factors. In this case it may seem that the definition of risk aversion can depend on how the commodities are grouped: a person may, say, prefer hedging between today and tomorrow, but dislike hedging between work and pleasure. We show that this is not possible; regardless of how one chooses to partition the commodities into independent factors, a decision maker is risk averse according to one partition if and only if she is risk averse according to any and all other partitions. Thus, this definition of risk aversion reflects an underlying attitude of the decision maker, not a particularity of the specific partition.

Also, the definition can be extended to the case where the partition is into more than two factors, and hedging takes place between two of the factors. In this case, it may again seem plausible that the definition depend on which two factors are chosen. Again, we show that this is not the case; a decision maker is risk averse when considering one pair of factors, if and only if she is risk averse according to any and all other pairs (provided that the certainty preferences on any such pair are well defined). Thus, again, the definition reflects an underlying attitude of the decision maker, not a particularity of the specific pair in consideration.

A Quantitative Form. Having established ordinal definitions of risk-aversion, we show that these ordinal notions can *also* be cast in quantitative form, using an appropriate scale. Such a scale, we show, is provided by the *multi-attribute (additive) value function*, pioneered by Debreu [7, 8],

[1]Here and throughout, the term "commodities" may refer to different types of goods (e.g. apples and oranges), or to the same good at different times (e.g. oranges today and oranges tomorrow), or to any combination thereof (apples and oranges today and tomorrow). However, "commodities" does not refer to *contingent commodities*, as our use of the term specifically refers only to sure outcomes. Preferences over contingent commodities are determined by the lottery preferences.

[2]Following Richard's initial definition, Epstein and Tanny [12] coined the term *correlation aversion* for such a preference. Here, we use Richard's original term.

and commonly used in the theory of multi-attribute decision theory (see [19]). Debreu proves that (under appropriate conditions) the preferences on commodity bundles can be represented by the sum of appropriately defined functions of the individual commodities. Importantly, these Debreu functions are defined solely on the basis of the internal preferences amongst the commodity bundles. Thus, unlike market value - which is determined by external market forces - the Debreu functions represent the decision maker's own preferences. Also, the functions are defined using the preferences on sure outcomes alone, with no reference to gambles. Thus, they provide a natural, intrinsic yardstick with which risk-aversion can be measured.

We show that the two above mentioned ordinal definitions of risk-aversion coincide with the Arrow-Pratt numerical definition, once the latter is defined with respect to the Debreu value function. Essentially, we show that the NM utility function is concave with respect to the associated Debreu function if and only if the given preference order is risk averse, under either of the two ordinal definitions.

It is interesting to note that Richard's ordinal definition, while well known in the literature, has been viewed as separate from the classic one; "a new type of risk aversion unique to multivariate utility functions" - in the words of Richard [25]. We show that the two definitions are one and the same, once the appropriate scale is used.

1.3. **Implications.**

The approach offered in this paper has several implications for the understanding of risk aversion, both conceptual and technical. Here, we mention two. Additional implications are mentioned in the discussion.

Non-monetary Goods. First and foremost, the approach offers a way to define risk aversion for non-monetary goods and goods with no natural scale, such as temperature, pain, and pleasure. Indeed, in the definitions of this paper, externally defined scales (such as market value) do not play any role. Rather, the only scale of interest is the intrinsic Debreu value, which reflects the decision maker's own certainty preferences.

Disentangling Risk Aversion from Diminishing Marginal Utility. On a conceptual level, the approach offered in this paper provides a natural way for disentangling risk aversion from diminishing marginal utility. In this scheme, the curvature of the NM utility function with respect to money is decomposed into two components: the curvature of the Debreu value function with respect to money, and the curvature of the NM utility function with respect to the Debrue value function. With this decomposition, the former may naturally be associated with diminishing marginal utility, while the latter - we argue - represents the risk aversion component.

1.4. **Assumptions.**

Independence. Independence is a key notion and assumption throughout this work. Simply put, a commodity, or set of commodities, is *independent* if the preference order over bundles of this set of commodities is independent of the state in other commodities.[3] Arguably, independence is a strong assumption; having eaten Chinese food today may affect one's gastronomical preferences tomorrow. Nonetheless, independence is a common assumption in economic literature, and in particular with respect to time preferences; indeed, the standard (exponential) discounted-utility model implies independence of any time period (indeed, any subset of the time periods). We use the independence assumption not because we believe it is a 100% accurate representation of reality, but rather because we believe it is a good enough approximation, which allows us to concentrate on and formalize other key notions.

Expected Utility. This work is presented entirely within the expected-utility (EU) framework. The key reason is that the classical Arrow-Pratt definitions were provided within this framework, and we seek to explore the conceptual foundations of these definitions. Additionally, while EU is perhaps not the only possible model, it nonetheless is a possible model; and one that is frequently used in real-world economic and financial applications. So, understanding the notion of risk aversion within this framework is of interest. Extending these ideas to non-EU models is an interesting future research direction.

1.5. **Plan of the Paper.** The remainder of the paper is structured as follows. Immediately following, in Section 2, we present the model, terminology and notation used throughout. The first ordinal definition is presented in Section 3, and its quantitative form in Section 4. Section 5 presents the second definition, with its equivalent quantitative form in Section 6. The relationship between the two definitions in discussed in Section 7. The applications to multi-commodity risk aversion are discussed in Section 8. We conclude the main body of the paper with a discussion in Section 9. All proofs are deferred to an appendix.

2. Model, Terminology and Notation

The Commodity Spaces. Preferences are defined over a product space $\mathcal{S} = \mathscr{C}_1 \times \cdots \times \mathscr{C}_m$, where each $\mathscr{C}_i$ is a real interval representing the consumption space of commodity i.

Lotteries. We consider finite support lotteries over $\mathcal{S}$, and denote by $\Delta(\mathcal{S})$ the space of all such lotteries. The fair lottery between s_1 and s_2 is denoted $\langle s_1, s_2 \rangle$.

Preference Orders. For a space $\mathcal{S}$, two preferences orders are considered:

- *the certainty preferences*: a preference order $\precsim$ on $\mathcal{S}$,[4]
- *the lottery preferences*: a continuous preference order $\overset{\Delta}{\precsim}$ on $\Delta(\mathcal{S})$, which agrees with $\precsim$ on the sure outcomes.

[3] A formal definition is provided in the next section.

[4] A *preference order* is a complete, transitive and reflexive binary relation.

As customary, $\prec$ denotes the strict preference order induced by $\precsim$, and $\sim$ the induced indifference relation; similarly $\overset{\vartriangle}{\prec}$ and $\overset{\vartriangle}{\sim}$ denote the relations induced by $\overset{\vartriangle}{\precsim}$. Continuity of $\overset{\vartriangle}{\precsim}$ means that for any lottery L, the sets $\{s : s\overset{\vartriangle}{\prec}L\}$ and $\{s : s\overset{\vartriangle}{\succ}L\}$ are open (in $\mathcal{S}$). Since $\overset{\vartriangle}{\precsim}$ and $\precsim$ agree on $\mathcal{S}$, this implies that $\precsim$ is also continuous (that is, the sets $\{s : s\overset{\vartriangle}{\prec}s'\}$ and $\{s : s\overset{\vartriangle}{\succ}s'\}$ are open for all $s' \in \mathcal{S}$)

All commodity spaces $\mathscr{C}_i$ are assumed to be *strictly essential* [15]; that is, for each i and $s_{-i} \in \mathscr{C}_{-i}$ (the remaining commodities), there exist $s_i, s'_i \in \mathscr{C}_i$ with $(s_i, s_{-i}) \not\sim (s'_i, s_{-i})$.

We assume throughout that the von Neumann-Morgenstern (NM) axioms hold for all preference orders on lotteries.

Factors and Partitions. The term *factor* refers to a single $\mathscr{C}_i$ or a product of several $\mathscr{C}_i$'s; i.e., a factor is the product of one or more commodity spaces. A *partition* of $\mathcal{S}$ is a representation of $\mathcal{S}$ as a product of factors $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$. An element of $\mathcal{S}$ (or of any factor) is called a *bundle.*

Throughout, a_i, b_i, c_i represent elements of $\mathcal{T}_i$. For i, j, we denote $\mathcal{S}_{-\{i,j\}} = \prod_{t\neq i,j} \mathcal{T}_t$. For $\boldsymbol{c} \in \mathcal{S}_{-\{i,j\}}$, by a slight abuse of notation we denote

$$(a_i, a_j, \boldsymbol{c}) = (c_1, \ldots, c_{i-1}, a_i, c_{i+1}, \ldots, c_{j-1}, a_j, c_{j+1}, \ldots, c_n). \tag{1}$$

Bundle Intervals. For $\underline{s} \precsim \overline{s}$, we denote

$$[\underline{s}, \overline{s}] = \{s : \underline{s} \precsim s \precsim \overline{s}\}$$

That is, $[\underline{s}, \overline{s}]$ is the closed *interval of bundles* between $\underline{s}$ and $\overline{s}$. Hence, we call such an $[\underline{s}, \overline{s}]$ a *bundle interval*, or simply *interval.*

Utility Representations. A function $f : \mathcal{S} \to \mathbb{R}$ *represents* $\precsim$ if for any $s, s' \in \mathcal{S}$,

$$s \precsim s' \iff f(s) \leq f(s').$$

The function $f : \mathcal{S} \to \mathbb{R}$ is an *NM utility* of $\overset{\vartriangle}{\precsim}$ if for any $L_1, L_2 \in \Delta(\mathcal{S})$,

$$L_1\overset{\vartriangle}{\precsim}L_2 \iff E_{L_1}[f(s)] \leq E_{L_2}[f(s)],$$

where $E_{L_j}[f(s)]$ is the expectation of $f(s)$ when s is distributed according to L_j. In that case we also say that f *represents* $\overset{\vartriangle}{\precsim}$.

Independence. Independence is a key notion in our analysis. Simply put, a factor is independent if the preferences on the factor are well defined; i.e., the preferences within the factor are independent of the state in other factors. Formally, for a partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$, we say that factor $\mathcal{T}_i$ is *independent* if there exists a preference order $\precsim^{\mathcal{T}_i}$ on $\mathcal{T}_i$ such that for any $a_i, b_i \in \mathcal{T}_i$ and any $\boldsymbol{c} \in \mathcal{S}_{-i}$ (the remaining factors),

$$a_i \precsim^{\mathcal{T}_i} b_i \iff (a_i, \boldsymbol{c}) \precsim (b_i, \boldsymbol{c}).$$

It is important to stress that independence only refers to the certainty preferences; it does not state or imply that the preferences on lotteries in one factor are independent of the state in other factors. That would be a much stronger assumption, which we do not make.

When no confusion can result, we may write $\precsim$ instead of $\precsim^{\mathcal{T}}$; thus, when $a, a' \in \mathcal{T}$, we may write $a \precsim a'$ instead of $a \precsim^{\mathcal{T}} a'$. It is worth noting that the product of independent factors need not be independent.[5]

A partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ is an *independent partition* if the product of any subset of factors is independent. By Gorman [15], for $n \geq 3$, it suffices to assume that $\mathcal{T}_i \times \mathcal{T}_{i+1}$ is independent for all i, and the independence of all other products then follows.

Relative Convexity/Concavity. Let $f, g : S \to \mathbb{R}$, for some space S, with $g(x) = g(y) \Rightarrow f(x) = f(y)$, for all $x, y \in S$. We say that f *is concave with respect to* g if there is a concave function h with $f = h \circ g$. Similarly for convexity, strict concavity, and strict convexity.

3. Ordinal Definition I: Repeated Lotteries

Our first ordinal definition of risk aversion is set in the context of repeated lotteries. Conceptually, this definition says that risk aversion is a preference that when adhered to repeatedly, ultimately leads to an inferior outcome. More specifically, with a risk averse preference, repeatedly choosing the certainty equivalent of a lottery over the lottery itself ultimately leads to an inferior outcome, with probability 1. To make this definition concrete, we must first define the associated notions, including: *repeated lotteries*, *certainty equivalent of a repeated lottery*, and *ultimately inferior outcome.*

The Space. We consider an infinite sequence of factors $\mathcal{T}_1, \mathcal{T}_2, \ldots$, where $\mathcal{T}_i$ represents the consumption space at time i.[6] We denote $\mathcal{H}^n = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ - the finite history space up to time n. In the following, a_i, b_i, c_i, will always be taken to be in $\mathcal{T}_i$, and lottery L_i will be over $\mathcal{T}_i$.

Preference Orders. While the number of factors is infinite, we only need to consider the preferences on the finite history spaces $\mathcal{H}^n$. We denote by $\precsim^n$ the preference order on $\mathcal{H}^n$, and by $\overset{\vartriangle}{\precsim}^n$ the preference order on $\Delta(\mathcal{H}^n)$. The superscript n is frequently omitted when clear from the context. Each $\mathcal{T}_i$ is assumed to be independent (in the certainty preference orders $\precsim^n$), but not necessarily utility independent (in preference orders $\overset{\vartriangle}{\precsim}^n$).

We call the sequence of preference orders $\overset{\vartriangle}{\precsim} = (\overset{\vartriangle}{\precsim}^1, \overset{\vartriangle}{\precsim}^2, \ldots)$ the *preference policy.*

Lottery Sequences. Let $L_1, L_2, \ldots,$ be a sequence of lotteries (with L_i over $\mathcal{T}_i$). We denote by $(L_1, \ldots, L_n)$ the lottery over $\mathcal{H}^n$ obtained by the independent application of each L_i on its associated factor.

[5]A simple example is the preference on $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z} = (\mathbb{R}^+)^3$ represented by the function $v(x, y, z) = xy + z$. Here, each commodity space is independent, but $\mathcal{Y} \times \mathcal{Z}$ is not independent.

[6]We do not assume that $\mathcal{T}_i = \mathcal{T}_j$, i.e. the state spaces need not be the same at different time periods. In particular, we do not assume any form of stationarity (though it is possible). Similarly, discounting may or may not be applied between consecutive factors. Our discussion here is independent of any such nominal matters.

Certainty Equivalents. Suppose that at time $t = 1$ the decision maker is offered the choice between lottery L_1 and its certainty equivalent c_1. Then, consistent with her preference policy, she may choose c_1, which suppose she indeed does. Now, at time t_2, she is offered the choice between lottery L_2 and its certainty equivalent c_2. Again, consistent with her preference policy, she chooses c_2. Suppose that she is thus offered, in each time period, the choice between a lottery L_i and its certainty equivalent c_i. Then the decision maker can consistently choose c_i, ending up with $(c_1, c_2, \ldots)$.

Accordingly, we say that $\boldsymbol{c} = (c_1, c_2, \ldots)$ is the *repeated certainty equivalent* of $\boldsymbol{L} = (L_1, L_2, \ldots)$ if $(c_1, \ldots, c_{n-1}, c_n) \precsim\!\!\succsim^n (c_1, \ldots, c_{n-1}, L_n)$ for all n.

Ultimate Inferiority. Consider a sequence $\boldsymbol{c} = (c_1, c_2, \ldots)$ of sure states, and a sequence $\boldsymbol{L} = (L_1, L_2, \ldots)$ of lotteries. Let ℓ_i be the realization of L_i. We say that $\boldsymbol{c}$ is *ultimately inferior* to $\boldsymbol{L}$ if

$$\Pr[(c_1, \ldots, c_n) \prec^n (\ell_1, \ldots, \ell_n) \text{ from some } n \text{ on}] = 1.^7$$

Notably, here $\prec^n$ denotes the preference over the *sure* states. Thus, if $\boldsymbol{c}$ is ultimately inferior to $\boldsymbol{L}$, then consistently choosing the sure state c_i over the lottery L_i, will, with probability 1, eventually result in an inferior outcome, and continue being so indefinitely.

Similarly, $\boldsymbol{c}$ is *ultimately superior* to $\boldsymbol{L}$ if

$$\Pr[(c_1, \ldots, c_n) \succ^n (\ell_1, \ldots, \ell_n) \text{ from some } n \text{ on}] = 1.$$

Bounded and Non-Vanishing Lottery Sequences. We now want to define risk aversion as a policy for which the repeated certainty equivalent of a lottery sequence is always ultimately inferior to the lottery sequence itself. However, as such, this definition cannot be a good one since in the case that the "magnitude" of the lotteries rapidly diminishes the overall outcome will be dominated by that of the first lotteries, and we could never obtain an inferior outcome with probability 1. Similarly, if the "magnitude" of the lotteries can grow indefinitely, then for almost any preference policy one can construct a lottery sequence that is ultimately inferior to its repeated certainty equivalent.[8] Hence, we now define the notions of a *bounded* lottery sequence and a *non-vanishing* lottery sequence.

For bundle intervals $[a_1, b_1]$ and $[a_j, b_j]$, we denote $[a_j, b_j] \sqsubseteq [a_1, b_1]$ if $(a_1, \boldsymbol{c}, b_j) \precsim (b_1, \boldsymbol{c}, a_j)$ for all $\boldsymbol{c} \in \mathcal{S}_{-\{1,j\}}$ (see Figure 1). Similarly, $[a_1, b_1] \sqsubseteq [a_j, b_j]$ if $(a_1, \boldsymbol{c}, b_j) \succsim (b_1, \boldsymbol{c}, a_j)$ for all $\boldsymbol{c} \in \mathcal{S}_{-\{1,j\}}$.

A sequence of intervals $[a_1, b_1], [a_2, b_2], \ldots$, is *bounded* if $[a_i, b_i] \sqsubseteq [a_1, b_1]$, for all i. The sequence is *vanishing* if for any $[\tilde{a}_1, \tilde{b}_1]$, there exists a j_0 such that $[a_j, b_j] \sqsubseteq [\tilde{a}_1, \tilde{b}_1]$ for all $j > j_0$. That is, the intervals in the tail of the sequence become infinitely small.

A lottery sequence $\boldsymbol{L} = (L_1, L_2, \ldots)$ is *bounded* if its support is entirely within some bounded interval sequence (that is, there exists a bounded sequence of intervals $[a_1, b_1], [a_2, b_2], \ldots$, with

[7]differently put: $\Pr[\exists N, \forall n \geq N : (c_1, \ldots, c_n) \prec^n (\ell_1, \ldots, \ell_n)] = 1$.

[8]See Appendix B.

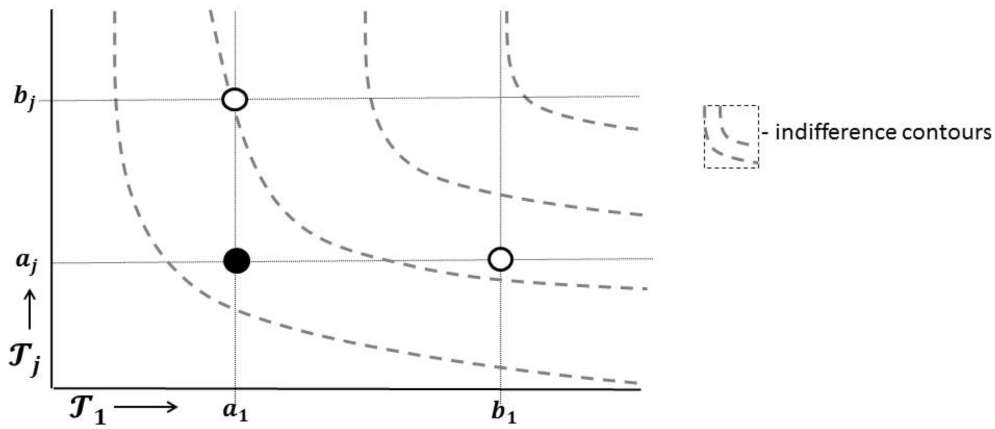

FIGURE 1. Illustration of $[a_j, b_j] \sqsubseteq [a_1, b_1]$ (the factors of $\mathcal{S}_{-\{1,j\}}$ are not depicted).

$L_i \in \Delta([a_i, b_i])$ for all i). The sequence is *non-vanishing* if it includes an infinite sub-sequence of fair lotteries, the support thereof is not entirely within any vanishing interval sequence.

Risk Averse Policies. Equipped with these definitions, we can now define risk aversion:

Definition 1. *We say that preference policy* $\stackrel{\Delta}{\succsim}$ *is:*

- Risk averse *if for any bounded non-vanishing lottery sequence, the repeated certainty equivalent of the sequence is ultimately inferior to the lottery sequence itself.*
- Weakly risk averse *if the repeated certainty equivalent of any bounded lottery sequence is not ultimately superior to the lottery sequence itself.*

Thus, the bias of the risk averse for certainty can never result in an ultimately superior outcome, and on non-vanishing lotteries necessarily leads to an inferior outcome.

Note that the above definition is fully ordinal; it makes no reference to any numerical scale, and indeed, no such scale need exist.

3.1. **Risk Loving and Risk Neutrality.** For readability, we deferred the definitions of risk loving and risk neutrality. We now complete the picture by providing these definitions.

Definition 2. *We say that preference policy* $\stackrel{\Delta}{\succsim}$ *is:*

- Risk loving *if for any bounded non-vanishing lottery sequence, the repeated certainty equivalent of the sequence is ultimately superior to the lottery sequence itself.*
- Weakly risk loving *if the repeated certainty equivalent of any bounded lottery sequence is not ultimately inferior to the lottery sequence itself.*
- Risk neutral *if it is both weakly risk loving and weakly risk averse.*

Thus, the risk loving require an ultimately superior certainty equivalent to forgo their love of risk.

4. Repeated Lotteries: The Quantitative Perspective

The previous section provided a fully ordinal definition of risk aversion. We now show how this ordinal definition can be cast in quantitative form. Specifically, we show that (under some assumptions) this ordinal definition of risk-aversion coincides with the Arrow-Pratt cardinal definition, once the latter is defined with respect to the appropriate scale. This scale, we show, is provided by the Debreu value function, which we review next.

4.1. **Debreu Value Functions.** The theory of multi-attribute decision making considers certainty preferences over a multi-factor space, and establishes that under certain independence assumptions such preferences can be represented by quantitative functions, as follows. Consider the space $\mathcal{H}^n = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ $(n \geq 2)$, with preference order $\precsim^n$. Debreu [7] proves that, if the partition is independent[9] then $\precsim^n$ is *additively separable*;[10] that is, there exist functions $v^{\mathcal{T}_i} : \mathcal{T}_i \to \mathbb{R}$, such that for any $(a_1, \ldots, a_n), (a'_1, \ldots, a'_n)$

$$(a_1, \ldots, a_n) \precsim^n (a'_1, \ldots, a'_n) \iff \sum_{i=1}^{n} v^{\mathcal{T}_i}(a_i) \leq \sum_{i=1}^{n} v^{\mathcal{T}_i}(a'_i).$$

It is important to note that the functions are defined solely on the basis of the certainty preferences.

Debreu's theorem also establishes that the functions are unique up to similar positive affine transformations (that is, multiplication by identical positive constants and addition of possibly different constants).

We call the function $v^{\mathcal{T}_i}$ a *(Debreu) value function* for $\mathcal{T}_i$, and the aggregate function $v_n = \sum_{i=1}^{n} v^{\mathcal{T}_i}$ a *(Debreu) value function* for $\mathcal{H}^n$.[11] We note that Debreu [7] called these functions *utility* functions; but following Keeney and Raiffa [19], we use the term *value* functions, to distinguish them from the NM utility function.

4.2. **Risk Aversion and Concavity.** We now show that our ordinal definition of risk aversion, Definition 1, corresponds to concavity of the NM utility functions with respect to the associated Debreu value functions, provided these value functions exist, and that some consistency properties hold among the preference orders on the $\mathcal{H}^n$'s. The exact conditions are now specified.

Certainty Preference. Consider the case where each consecutive pair of factors $\mathcal{T}_i \times \mathcal{T}_{i+1}$ is independent. Also, assume that the preference orders $\precsim^n$ are *consistent* in the sense that for $n' > n$, the preference order induced by $\precsim^{n'}$ on $\mathcal{H}^n$ is identical to $\precsim^n$. These assumptions yield the existence of value functions, as follows:

---

[9] see page 7.

[10] In the case of two factors $(n = 2)$, the following *Thomsen condition* is also required: for all $a_1, B_1, c_1 \in \mathcal{T}_1$, and $a_2, b_2, c_3 \in \mathcal{T}_2$, if $(a_1, b_2) \sim (b_1, a_2)$ and $(b_1, c_3) \sim (c_1, b_2)$ then $(a_1, c_2) \sim (c_1, a_2)$. For $n > 2$ the Thomsen condition is implied by the independence of the pairs.

[11] This is a slight abuse of notation. More precisely, v is the function on $\mathcal{H}^n$ given by $v(a_1, \ldots, a_n) = \sum_{i=1}^{n} v^{\mathcal{T}_i}(a_i)$.

Proposition 4.1. *There exist Debreu value functions $v^{\mathcal{T}_i} : \mathcal{T}_i \to \mathbb{R}$, $i = 1, 2, \ldots$, such that for all n, $v_n = \sum_{i=1}^{n} v^{\mathcal{T}_i}$ represents $\precsim^n$.*

Lottery Preferences. Whereas the factors are assumed independent, the lottery preferences thereupon need not be independent. That is, the preference order on $\Delta(\mathcal{H}^n)$ induced by $\overset{\vartriangle}{\precsim}^{n+1}$ may depend on the state a_{n+1} in $\mathcal{T}_{n+1}$. We do assume, however, a form of *weak consistency*, whereby there *exists* some $\phi_{n+1} \in \mathcal{T}_{n+1}$ with

$$L \overset{\vartriangle}{\precsim}^n L' \iff (L, \phi_{n+1}) \overset{\vartriangle}{\precsim}^{n+1} (L', \phi_{n+1});$$

that is, the preferences on $\Delta(\mathcal{H}^n)$ are consistent with *some* possible future. We call the sequence $(\phi_2, \phi_3, \ldots)$ a *presumed future*, and assume that it is *internal*,[12] in the following sense. The sequence $(\phi_2, \phi_3, \ldots)$ is *internal* if there exists an $s > 0$ with $v^{\mathcal{T}_i}(\phi_i) \pm s \in v^{\mathcal{T}_i}(\mathcal{T}_i)$ for all i; that is, the presumed future is bounded away from the boundaries of the $\mathcal{T}_i$'s.

4.2.1. *Weak Risk Aversion and (Weak) Concavity.* For each n, let u_n be the NM utility function representing $\overset{\vartriangle}{\precsim}^n$. The next theorem establishes the connection between weak risk aversion and concavity of the u_n's.

Theorem 1. *$\overset{\vartriangle}{\precsim}$ is weakly risk averse if and only if u_n is concave with respect to v_n for all n.*

Thus, Theorem 1 provides the missing conceptual justification for defining risk aversion by concavity of the utility function. It also establishes the appropriate scale - the Debreu value function.

Interestingly, the theorem provides that *all* NM utility functions must be concave, not only from some n on.

4.2.2. *(Strict) Risk Aversion and Strict Concavity.* We would have now wanted to claim that (strict) risk aversion corresponds to strict concavity of the NM utility functions (with respect to the value function). However, strict concavity alone is not enough, as we are considering repeated lotteries, and we cannot expect ultimate inferiority if the "level of concavity" rapidly diminishes. So, we need a condition that ensures that the functions are also "uniformly" strictly concave in some sense. As it turns out, the condition of interest is that the coefficient of absolute risk aversion of the NM utility functions is bounded away from zero (when measured with respect to the value function). The exact definitions follow.

For each n, let $\hat{u}_n$ be the function such that $\hat{u}_n(v_n(a_1, \ldots, a_n)) = u_n(a_1, \ldots, a_n)$. This is well defined, as $\overset{\vartriangle}{\precsim}^n$ and $\precsim^n$ agree on the certainty preferences. Conceptually, $\hat{u}_n$ is the function u_n once the underlying scale is converted to the value function v_n. Denote $\hat{\boldsymbol{u}} = (\hat{u}_1, \hat{u}_2, \ldots)$.

For a twice differentiable function f the *coefficient of absolute risk aversion* of f at x is:

$$A_f(x) = -\frac{f''(x)}{f'(x)}.$$

[12] More precisely, we assume that there exists a presumed future that is internal.

Theorem 2. *If $A_{\hat{u}_n}(x)$ is bounded away from 0, uniformly for all n and x,*[13] *then $\stackrel{\Delta}{\succsim}$ is risk averse (assuming $\hat{u}_n$ is twice differentiable for all n).*

Theorem 2 establishes a sufficient condition for risk aversion. We now proceed to establish a necessary condition, which is "close" to being tight. To do so we need to consider the behavior of the functions $\hat{u}_i$, and the definition of $A_{\hat{u}_i}(\cdot)$, in a little more detail.

Let *risk-prem*$_{\hat{u}_n}(x, \pm\epsilon)$ be the risk premium according to $\hat{u}_n$ of the of the lottery $\langle x+\epsilon, x-\epsilon\rangle$; that is

$$\textit{risk-prem}_{\hat{u}_n}(x, \pm\epsilon) = x - (\hat{u}_n)^{-1}\left(\frac{\hat{u}_n(x+\epsilon) + \hat{u}_n(x-\epsilon)}{2}\right).$$

Now for any ϵ (sufficiently small) define

$$RP_{\hat{\mathbf{u}}}(\epsilon) = \inf_{n,x}\{\textit{risk-prem}_{\hat{u}_n}(x, \pm\epsilon)\}.$$

So, $RP_{\hat{\mathbf{u}}}(\cdot)$ is a function. We will be interested in the rate at which $RP_{\hat{\mathbf{u}}}(\epsilon)$ declines as $\epsilon \to 0$. The condition of interest, we show, is that $RP_{\hat{\mathbf{u}}}(\epsilon)$ declines no faster than ϵ^2.

Theorem 3.

(a) If $RP_{\hat{\mathbf{u}}}(\epsilon) = \Omega(\epsilon^2)$ as $\epsilon \to 0$ then $\stackrel{\Delta}{\succsim}$ is risk averse.[14]
(b) If $RP_{\hat{\mathbf{u}}}(\epsilon) = O(\epsilon^{2+\beta})$ as $\epsilon \to 0$, for some $\beta > 0$, then $\stackrel{\Delta}{\succsim}$ is not risk averse.

The sufficient condition of (a) and the necessary one of (b) are not identical, but are close.

Finally, we establish that the sufficient condition of Theorem 3-(a) and that of Theorem 2 are the same.

Proposition 4.2. *$RP_{\hat{\mathbf{u}}}(\epsilon) = \Omega(\epsilon^2)$ as $\epsilon \to 0$, if and only if $A_{\hat{u}_n}(x)$ is bounded away from 0, uniformly for all n and x (assuming $\hat{u}_n$ is twice differentiable for all n).*

4.3. **Risk Loving and Risk Neutrality.** In analogy to Theorems 1 and 3 we have:

Theorem 4. *For v_n, u_n, and $\hat{\mathbf{u}}$ as in Theorems 1 and 3*

(a) Weak risk loving: $\stackrel{\Delta}{\succsim}$ is weakly risk loving if and only if u_n is convex with respect to v_n for all n.
(b) Risk loving
- *If $(-RP_{\hat{\mathbf{u}}}(\epsilon)) = \Omega(\epsilon^2)$ as $\epsilon \to 0$ then $\stackrel{\Delta}{\succsim}$ is risk loving.*
- *If $(-RP_{\hat{\mathbf{u}}}(\epsilon)) = O(\epsilon^{2+\beta})$ as $\epsilon \to 0$ (for some $\beta > 0$) then $\stackrel{\Delta}{\succsim}$ is not risk loving.*

(c) Risk Neutral: $\stackrel{\Delta}{\succsim}$ is risk neutral if and only if u_n is a linear transformation of v_n for all n.

[13] that is, there exists an constant $\alpha > 0$ such that $A_{\hat{u}_n}(x) \geq \alpha$ for all n and x.

[14] recall that $g(y) = \Omega(h(y))$ as $y \to 0$ if there exists a constant M and y_0 such that $g(y) > M \cdot h(y)$ for all $y < y_0$.

5. Ordinal Definition II: Hedging

5.1. **The Definition.** Consider a space $\mathcal{S}$ and an independent partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$.[15] Recall the notation $\mathcal{S}_{-\{i,j\}} = \prod_{t \neq i,j} \mathcal{T}_t$, and for a_i, a_j, and $\boldsymbol{c} \in \mathcal{S}_{-\{i,j\}}$, the slight abuse of notation $(a_i, a_j, \boldsymbol{c})$ for $(c_1, \ldots, c_{i-1}, a_i, c_{i+1}, \ldots, c_{j-1}, a_j, c_{j+1}, \ldots, c_n)$. The following definition is that of Richard [25].

Definition 3. *Preference order* $\stackrel{\Delta}{\precsim}$ *is* R risk-averse *(Richard risk averse) with respect to the independent partition* $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$*, and the pair of factors* $\mathcal{T}_i, \mathcal{T}_j$*,* $i \neq j$*, if for any* $a_i \prec b_i, a_j \prec b_j$*, and* $\boldsymbol{c} \in \mathcal{S}_{-\{i,j\}}$

$$\langle (a_i, a_j, \boldsymbol{c}), (b_i, b_j, \boldsymbol{c}) \rangle \stackrel{\Delta}{\prec} \langle (a_i, b_j, \boldsymbol{c}), (b_i, a_j, \boldsymbol{c}) \rangle . \tag{2}$$

and weakly R risk averse *if* (2) *holds with weak preference.*

Note that the left-hand side lottery, $\langle (a_i, a_j, \boldsymbol{c}), (b_i, b_j, \boldsymbol{c}) \rangle$, is between two extreme outcomes: 50% probability for getting the better outcome in both commodities, and 50 % probability for getting the lesser in both. In the right-hand side lottery, $\langle (a_i, b_j, \boldsymbol{c}), (b_i, a_j, \boldsymbol{c}) \rangle$, the loss in one commodity is (partially) hedged by the gain in the other. Thus, a decision maker is R risk-averse if she prefers to hedge her bets to the extent possible.

5.2. **Properties.** In definition 3, risk aversion is defined with respect to a specific partition, and a specific pair of factors within the partition. The following proposition establishes that if the definition holds for some partition and some pair, then it holds for any partition and all pairs.

Theorem 5. *If* $\stackrel{\Delta}{\precsim}$ *is R risk averse with respect to some independent partition* $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$*, and some pair of factors* $\mathcal{T}_i, \mathcal{T}_j$*, then it is also R risk averse with respect to any independent partition, and any pair therein. Similarly for weak R risk aversion.*

By Theorem 5, we may drop reference to the specific partition and pair when considering R risk aversion.

5.3. **Perfectly Hedged Lotteries.** In the definition of R risk aversion the decision maker must prefer the hedged version of a lottery over the non-hedged version, for any quadruplet a_i, b_i, a_j, b_j, even if the resultant hedged lottery still entails some risk. A possible alternative definition would require that the decision maker only prefer hedges that completely eliminate any risk, as in the following definition:

Definition 4. *For* $a_i \prec b_i$*,* $a_j \prec b_j$*, say that* (a_i, b_j)*,* (b_i, a_j) *are* perfectly hedged *if* $(a_i, b_j) \sim (b_i, a_j)$ *(see Figure 2).*[16]

[15] Here we use the notation $\mathcal{S}$ rather than $\mathcal{H}^n$ since we will be considering one fixed space $\mathcal{S}$.

[16] The equivalence relation $(a_i, b_j) \sim (b_i, a_j)$ is well-defined as $\mathcal{T}_i \times \mathcal{T}_j$ is independent.

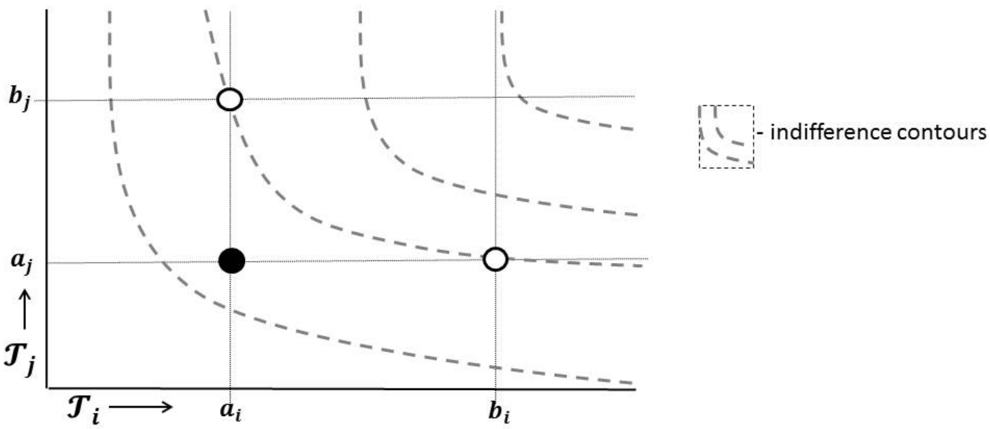

FIGURE 2. Illustration of a perfectly hedged pair.

We say that $\stackrel{\vartriangle}{\succsim}$ *is* perfect-R risk-averse *if for any partition* $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$*, and any* $a_i, b_i \in \mathcal{T}_i, a_j, b_j \in \mathcal{T}_j$*, and* $\boldsymbol{c} \in \mathcal{S}_{-\{i,j\}}$*, if* $(a_i, b_j), (b_i, a_j)$ *are perfectly hedged then,*

$$\langle (a_i, a_j, \boldsymbol{c}), (b_i, b_j, \boldsymbol{c}) \rangle \stackrel{\vartriangle}{\prec} \langle (a_i, b_j, \boldsymbol{c}), (b_i, a_j, \boldsymbol{c}) \rangle \,. \tag{3}$$

The preference order is weakly perfect-R risk averse *if* (3) *holds with a weak preference* ($\stackrel{\vartriangle}{\precsim}$).

The following theorem establishes that R risk aversion and perfect-R risk aversion are in fact equivalent.

Theorem 6. $\stackrel{\vartriangle}{\succsim}$ *is R risk-averse if and only if it is perfect-R risk averse.*

We note that the theorem holds even if there are only two factors, in which case $\precsim$ may fail to be *additively separable.*

5.4. **Risk Loving and Risk Neutrality.**

Definition 5. *We say that* $\stackrel{\vartriangle}{\succsim}$ *is* R risk-loving *if for any* $i \neq j$*, and any perfectly hedged* $(a_i, b_j), (b_i, a_j)$*, and any* $\boldsymbol{c} \in \mathcal{S}_{-\{i,j\}}$

$$\langle (a_i, a_j, \boldsymbol{c}), (b_i, b_j, \boldsymbol{c}) \rangle \stackrel{\vartriangle}{\succ} \langle (a_i, b_j, \boldsymbol{c}), (b_i, a_j, \boldsymbol{c}) \rangle \,. \tag{4}$$

and weakly R risk loving *if the preference in* (4) *is a weak one.*

Similarly, $\stackrel{\vartriangle}{\succsim}$ *is* R risk-neutral *if for any* $i \neq j$*, and any perfectly hedged* $(a_i, b_j), (b_i, a_j)$*, and any* $\boldsymbol{c} \in \mathcal{S}_{-\{i,j\}}$

$$\langle (a_i, a_j, \boldsymbol{c}), (b_i, b_j, \boldsymbol{c}) \rangle \stackrel{\vartriangle}{\sim} \langle (a_i, b_j, \boldsymbol{c}), (b_i, a_j, \boldsymbol{c}) \rangle \,.$$

Theorem 5 holds analogously for risk loving and risk neutrality.

6. Hedging Definition: The Quantitative Perspective

Again, the previous section provided a fully ordinal definition of risk aversion. We now show how this ordinal definition, too, equates with concavity of the utility function with respect to the value function, if and when the latter exists.

6.1. Uniqueness of the Aggregate Debreu Value Function.

We will shortly establish the relation between ordinal risk-aversion as in Definition 3, on the one hand, and the aggregate Debreu value function, on the other. Before we can do so, however, we need to guarantee that the notion of "*the*" aggregate Debreu value function is well defined. Debreu's theorem relates to a specific partition of the space, and asserts that the value functions are unique (up to similar positive affine transformations) for the given partition. It does not assert that a different function may not arise from a different partition. Thus, the notion of a single, unique value function for $\mathcal{S}$ may not be well defined. The following simple theorem, which may be of independent interest, shows that this is not the case; all disparate value functions that may arise from different partitions are identical.

Theorem 7. *For any $\mathcal{S}$, all (aggregate) Debreu value functions for $\mathcal{S}$ are identical up to positive affine transformations.*

6.2. Risk Aversion and the Debreu Value Functions.

Assume that the conditions guaranteeing the existence of a Debreu value function for $\mathcal{S}$ hold.[17] We now show that in this case, the ordinal Definition 3 coincides with concavity of the utility function with respect to the value function.

Theorem 8. *For NM utility u and Debreu value function v,*

- *Risk aversion:*
    - *u is strictly concave with respect to v if and only if $\overset{\vartriangle}{\succsim}$ is R risk averse.*
    - *u is concave with respect to v if and only if $\overset{\vartriangle}{\succsim}$ is weakly R risk averse.*
- *Risk loving:*
    - *u is strictly convex with respect to v if and only if $\overset{\vartriangle}{\succsim}$ is R risk loving.*
    - *u is convex with respect to v if and only if $\overset{\vartriangle}{\succsim}$ is weakly R risk loving.*
- *Risk neutrality: u is linear with respect to v if and only if $\overset{\vartriangle}{\succsim}$ is R risk-neutral.*

In all, we obtain that R risk aversion coincides with Arrow-Pratt risk aversion once concavity is defined with respect to the Debreu value function.

7. Relating the Two Ordinal Definitions

We considered two separate ordinal definitions of risk aversion: Definition 1, based on repeated lotteries, and Definition 3, based on hedging. Technically, the two definitions relate to different

[17]If there are three or more factors in the partition, then the existence of a value function is provided by the independence of the partition. If there are only two factors, the additional Thomsen condition is required (see Footnote 10).

mathematical objects: the first relates to a *preference policy*, which is a sequence of preference orders, while the latter relates to a single preference order. However, the two definitions are closely related, as established by Theorems 3 and 8: both definitions correspond to concavity of the NM utility function with respect to the Debreu value function (when it exists). For weak risk aversion the concavity requirements in both theorem are identical – (weak) concavity. So, when a Debreu value exist, a preference policy is weakly risk averse, according to Definition 1, if and only if each of the preferences orders therein is weakly risk averse, according to the Definition 3. For (strict) risk aversion, the requirement in Theorem 3 is a coefficient of absolute risk aversion bounded away from zero, whereas Theorem 8 requires only strict concavity. So, if the preference policy is (strictly) risk averse then so are all of the preference orders therein, but the opposite does not always hold. The reason is that since we are considering the behavior on recurring gambles we need a "recurring" bound on the strict concavity in all the gambles.

8. Multi-Commodity Risk Aversion

The seminal works of Arrow [2] and Pratt [24] defined risk aversion with respect to a single commodity – money. Ever since, researchers have attempted to extend the definition, and associated measures, to the multi-commodity setting (see [20, 28, 23, 9, 17, 25, 18, 21] for some references in the expected utility model). It is out of the scope of this paper to review this extensive body of research, but a key problem in the multi-commodity setting is that each commodity has its own scale so the question is which scale should be used when measuring the concavity of the utility function. Indeed, some papers (e.g. [9]) keep the multiple scales - in which case the measures of risk aversion become vectors and matrices.

Our approach here takes a different direction, which, in a way is the reverse. We do not start from the single commodity definition and extend it to multi-commodities, but rather start from the multi-commodity setting, and then *derive* the uni-scale case as a quantitative representation of the former. So, the "native" scales of the different commodities are immaterial in our approach. Rather, the only scale of interest is the intrinsically defined Debreu value function, which is shared across all commodities.

We note, again, that the second definition we consider, that of Richard, was presented as "a new type of risk aversion unique to multivariate utility functions" [25]. Scarsini, in a paper based on Richard's definition, writes "[Richard's definition] has nothing to do with what is generally known as risk aversion" [27]. Theorem 8 establishes that the two definitions are one and the same, once the appropriate scale is used.

We now show how, with the definition considered in this paper, the Arrow-Pratt framework carries over to the multi-commodity setting.

CARA Preferences. A (uni-scale) preference order is CARA (constant absolute risk aversion) if the coefficient of absolute risk aversion of its associated NM utility is constant. Arrow [2] showed that a preference is CARA if and only if the preferences on lotteries are independent of the wealth level. Specifically, for wealth level x and lottery L denote by (L, x) the lottery that gives the random outcome L *in addition* to the sure outcome x. Then, Arrow shows that preference order $\stackrel{\Delta}{\precsim}$ is CARA if and only if for all lotteries L, L' and wealth levels x, y

$$(L, x) \stackrel{\Delta}{\precsim} (L', x) \iff (L, y) \stackrel{\Delta}{\precsim} (L', y).$$

Now, in the multi-commodity setting, a natural interpretation of the phrase "the preferences on lotteries are independent of the wealth level" is that the preferences on lotteries in one commodity are independent of the wealth level in other commodities. Using our definition of multi-commodity risk aversion, we get the same correspondence as in the uni-scale case:

Theorem 9. *In the multi-commodity setting (with $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ an independent partition), the NM utility function u has constant coefficient of absolute risk aversion when measured with respect to the Debreu value function v if and only if for any i, lotteries L, L' over $\mathcal{T}_i$, and $\boldsymbol{x}, \boldsymbol{y} \in \Omega_{-\{i\}}$*

$$(L, \boldsymbol{x}) \stackrel{\Delta}{\precsim} (L', \boldsymbol{x}) \iff (L, \boldsymbol{y}) \stackrel{\Delta}{\precsim} (L', \boldsymbol{y}).$$

Furthermore, the following proposition establishes that our definition, in a way, is the only definition that preserves this correspondence.

Proposition 8.1. *Let $\precsim$ be an (additively separable) preference order on $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$, and g a real valued function on $\mathcal{S}$. Suppose that the following holds for any NM utility function u:*

- *u has constant coefficient of absolute risk aversion when measured with respect to g if and only if*

  $$(L, \boldsymbol{x}) \stackrel{\Delta}{\precsim} (L', \boldsymbol{x}) \iff (L, \boldsymbol{y}) \stackrel{\Delta}{\precsim} (L', \boldsymbol{y}).$$

  for any $L, L' \in \Delta(\mathcal{T}_i)$, and $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}_{-\{i\}}$.

Then g is a Debreu value function.

Comparative Multi-Commodity Risk Aversion. As observed by Kihlstrom and Mirman [20], in the multi-commodity setting it is natural to limit comparisons of risk aversion to decision makers agreeing on the certainty preferences. This also holds in our framework, as our definition of risk aversion is always *with respect to* the certainty preferences. For individuals agreeing on the certainty preferences, using our approach the entire Arrow-Pratt framework carries over as is, once the underlying scale is converted to the associated (joint) Debreu value function. In particular, we have the following. Let v be the joint Debreu value function and $\hat{u}_1$ and $\hat{u}_2$ be the NM utility functions of players 1 and 2 - when measured with respect to v. For a lottery L let $ce_j(L)$ be the certainty equivalent of L by $\hat{u}_j$ $(j = 1, 2)$. Then

$$ce_1(L) \precsim ce_2(L)$$

for all lotteries L if and only if

$$A_{\hat{u}_1}(x) \geq A_{\hat{u}_2}(x)$$

for all x (where $A_{\hat{u}_i}(x)$ is the coefficient of absolute risk aversion of $\hat{u}_i$ at x). This follows directly from Arrow-Pratt as their theorems do not specify the scale, and thus also apply when using the value function scale.

9. Discussion

We presented fully ordinal definitions of risk aversion, based entirely on the internal structure of preferences of the decision maker; independent of money or any other units. The first definition equates risk aversion with a policy that, in the long run, necessarily leads to an inferior outcome. The second definition, which is that of Richard [25], equates risk aversion with a preference for hedging bets. We show that when cast in numerical terms, both these ordinal definitions coincide with the Arrow-Pratt definition, once the latter is defined with respect to the Debreu value function associated with the decision maker's preferences over the sure outcomes. This, we suggest, provides the missing conceptual justification for the use of the arithmetic mean as the basis for defining risk aversion, and, at the same time, establishes the appropriate scale to use.

Inter-Commodity and Intra-Commodity Risk Aversion. We should stress that risk-aversion, as considered in this paper, does not relate only to gambles involving multiple commodities or times, but also to gambles *within* a single commodity/time. It may be seen that, given the multi-commodity *certainty* preferences, inter-commodity lottery preferences determine intra-commodity lottery preferences, and vice versa. Thus, inter-commodity and intra-commodity risk attitudes are one and the same. We use the inter-commodity setting as it provides an Archimedean vantage point from which the risk-attitude can be disentangled from the risk-free preferences. Once defined, however, it applies to all manifestations of risk. This is highlighted by the quantitative form using the Debreu function. The multi-commodity setting merely provides us with the appropriate *scale* with which to measure risk aversion, both inter and intra-commodity.

9.1. **Disentangling Risk Aversion from Diminishing Marginal Utility.** It is well known that under the classical definition, risk aversion and diminishing marginal utility are forever entangled. On a conceptual level, however, the two notions are distinct. Indeed, disentangling diminishing marginal utility from risk aversion is one of the earliest motivations for the non-expected utility literature, as Yaari [29] writes: "Two reasons have prompted me to look for an alternative to expected utility theory. The first reason is methodological: In expected utility theory, the agent's attitude towards risk and the agent's attitude towards wealth are forever bonded together. At the level of fundamental principles, risk aversion and diminishing marginal utility of wealth, which are synonyms under expected utility theory, are horses of different colors." Using the concepts of this work, it is possible disentangle the two *within* the expected utility framework. In our scheme, the curvature of the NM utility function with respect to money is decomposed into two components: the curvature of the Debreu value function with respect to money, and the curvature of the NM utility function with respect to the Debrue value function. With this decomposition, the former may naturally be associated with diminishing marginal utility, while the latter - we argue - represents the risk aversion component.

This disentangling may have implications for the language we use to describe (and hence comprehend) key economic behavior. Consider, for example, an aging, retired individual, comfortably living off her savings, who is offered a 50-50 gamble between tripling her savings and losing them all. Common sense has it that rejecting the gamble is a perfectly rational choice for all but the most risk loving individuals. Classical economic language, however, would have to deem such a rejection "risk aversion". The framework of this paper provides us with a more refined language, that allows us to give a more convincing interpretation of the behavior. When measured in terms of the Debreu value function, which reflects the relative benefits provided by each of the possible outcomes, the 50-50 gamble may well be actuarially inferior to the existing state. So, by our definition, the gamble should be rejected by risk neutral (or even some risk loving) individuals.

9.2. **CARA and CRRA.** Arrow and Pratt defined two concrete *measures* of risk aversion: the *coefficient of absolute risk aversion at x*, and the *coefficient of relative risk aversion at x* (defined as $-\frac{x \cdot u''(x)}{u'(x)}$). The measure of absolute risk aversion can naturally be converted to our definition of risk aversion, by simply considering the utility function *with respect to the Debreu value function*, as discussed in Section 8. The notion of relative-risk-aversion w.r.t. the value function, however, is not well defined, as the definition of relative risk aversion requires a well-defined zero point, and the value function is only defined up to an additive constant.[18]

In Section 8 we proved that once considered w.r.t. the value function, constant-absolute-risk-aversion (CARA) has a simple and intuitive meaning. A preference order is CARA w.r.t. the value function if and only if the preferences over lotteries in each individual factor are well defined and independent of the state in the other factors; preferences over apple lotteries are independent of the available amount of oranges and preferences over orange lotteries are independent of the available amount of apples (this is termed *utility independence* in [25, 4, 19]).

In the economic literature, CRRA (constant relative risk aversion) rather than CARA, is the more prevalent model. CRRA, however, is assumed w.r.t. money. Once considered in terms of the value function, the observed CRRA w.r.t. money may actually reflect a combination of an underlying CARA ordinal risk attitude superimposed on a value function that is logarithmic w.r.t. money. This combination yields exactly the known CRRA family of functions:

- ordinal risk aversion: $u(x) = -e^{-\gamma \ln(x)} = -x^{-\gamma}$ ($\gamma > 0$),
- ordinal risk neutrality: $u(x) = \ln(x)$,
- ordinal risk loving: $u(x) = e^{\gamma \ln(x)} = x^{\gamma}$ ($\gamma > 0$).

[18]Indeed, we would argue that determining the zero point is a big problem, mostly overlooked, also when defining relative risk aversion w.r.t. money. What is the right zero point? no money in the bank? no material possessions (no house, no clothes, no food)? no money left after selling a kidney? Choosing any of these zero points results in very different relative risk aversion coefficients.

Interestingly, this means that the utility functions $\ln(x)$ and x^γ actually correspond to risk neutrality and risk loving under our definitions, not risk aversion.

9.3. **Additive Utility.** Under the definitions of this paper the standard intertemporal model, wherein NM utility is additive across time periods, corresponds to risk neutrality. To obtain risk aversion we must venture out of the additive model. This may seem problematic at first, but we note that, irrespective of how one defines risk aversion, restricting the NM utility to an additive form necessarily means that the preferences on lotteries ($\overset{\vartriangle}{\precsim}$) are uniquely determined by the certainty preferences ($\precsim$); that is, all agents that agree on the preferences on the sure outcomes, must also agree on the preferences over risky one (see Appendix C for a formal statement and proof). This, we find, strips the notion of "risk aversion" from any semantic meaning, as the attitude towards risk plays no role in the determining the preferences - it is all determined by the preferences over the risk free outcomes. So, if we want to allow for varying risk attitudes, we must venture beyond additive utility. In a fascinating work, Bommier [5] shows how it is possible to allow for such non-additive models, while retaining both stationarity and consistency, *and* remaining within the expected utility framework.

9.4. **Repeated Games.** The theory of (infinitely) repeated games assumes that the utility in the repeated game is additive, in one way or another, in the utilities of the individual stage games [3, 26, 14]. By our definition, this corresponds to an assumption of risk neutrality. Accordingly, in a sequel work [**?**], we consider a theory of repeated games without this additivity assumption. We show that when players are risk averse - according to our ordinal definitions - new equilibria emerge, unaccounted for by the classical theory. Also, in two person matching pennies games, if one player is risk averse and the other risk loving, then the resulting pure strategy equilibria are biased in favor of the risk loving player. Such biased equilibria are not possible in the classic theory.

9.5. **Strength of Preference and Relative Risk Aversion.** Dyer and Sarin [11] and Bell and Raiffa [4] have suggested measuring risk aversion with respect to the *strength of preference function*, rather than money. It is out of the scope of this paper to review the strength-of-preference theory, but generally speaking this theory assumes that not only do decision makers have a well defined preference order over sure states and lotteries, but also that they have a preference order over *differences* between states; that is, the decision maker can state that she prefers the transition $x_1 \longmapsto x_2$ over the the transition $y_1 \longmapsto y_2$ (where x_1, x_2, y_1, y_2 are states). Assuming such preferences exist (and some additional technical conditions), the theory establishes that there exists a function f (termed *measurable value function* [10]) that *represents* these preferences, in the sense that $f(x_2) - f(x_1) > f(y_2) - f(y_1)$ if and only if the transition $x_1 \longmapsto x_2$ is preferred over the transition $y_1 \longmapsto y_2$. Given such a function, Dyer and Sarin [11] define the notion of *relative risk* *aversion*[19] as the concavity of the NM utility function u with respect to the measurable value function f. Bell and Raiffa [4] similarly define the notion of *intrinsic risk aversion*.

Bell and Raiffa [4] also show how the strength-of-preference function (assuming it exists) can be deduced and identified with a multi-attribute (Debreu) value function (see also [11, Theorem 1]). Thus, Theorem 8 establishes that technically the ordinal of R risk aversion coincides with the Dyer and Sarin notion of *relative risk aversion*, if a Debreu value function exists and relative risk aversion is computed with respect to this function. Conceptually, however, our approach is totally different from that of [11] and [4]. First, we do not suppose, technically or conceptually, any form of preferences over differences. Rather, we only use the standard preferences on bundles and lotteries thereof. Second, conceptually [11] and [4] follow the Arrow-Pratt framework, taking it as given that the "natural value" of a gamble "should be" its expectation. They differ from Arrow-Pratt only in using a different scale. Our approach is the opposite. Our starting point, and all core definitions, are fully ordinal. The numerical representation is then mathematically *derived* from this ordinal theory.

References

[1] N. Alon and J.H. Spencer. *The Probabilistic Method*. Wiley Series in Discrete Mathematics and Optimization. Wiley, 2004.

[2] Kenneth J. Arrow. *Essays in the theory of risk-bearing*. Markham Publishing Company, 1971.

[3] Robert J. Aumann. Survey of Repeated Games. *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, 1981.

[4] David E Bell and Howard Raiffa. Marginal value and intrinsic risk aversion. In David E. Bell, Howard Raiffa, and Amos Tversky, editors, *Decision Making Descriptive, Normative and Prescriptive Interactions*, pages 384–397. Cambridge University Press, 1988.

[5] Antoine Bommier. Life-cycle preferences revisited. *Journal of the European Economic Association*, 11(6):1290–1319, 2013.

[6] G Debreu. Representation of a preference ordering by a numerical function. In R.M. Thrall, R. L. Davis, and C.H. Coombs, editors, *Decision Process*, pages 159–165. John Wiley, New York, 1954.

[7] Gerard Debreu. Topological Methods in Cardinal Utility Theory. Technical Report 76, Cowles Foundation for Research in Economics, Yale University, 1959.

[8] Gerard Debreu. Topological methods in cardinal utility theory. In *Mathematical Economics: Twenty Papers of Gerard Debreu*, chapter 9. Cambridge University Press, 1986.

[9] George T. Duncan. A matrix measure of multivariate local risk aversion. *Econometrica*, 45(4):895–903, 1977.

[10] James S. Dyer and Rakesh K. Sarin. Measurable multiattribute value functions. *Operations Research*, 27(4):810–822, 1979.

[11] James S. Dyer and Rakesh K. Sarin. Relative risk aversion. *Management Science*, 28(8):875–886, 1982.

[12] Larry G. Epstein and Stephen M. Tanny. Increasing Generalized Correlation: A Definition and Some Economic Consequences. *Canadian Journal of Economics*, 13(1):16–34, February 1980.

[13] Peter C. Fishburn. *Utility Theory for Decision Making*. Wiley, New York, 1970.

[19]not to be confused with the Arrow-Pratt coefficient of relative risk aversion

[14] Drew Fudenberg and Eric Maskin. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3):533–554, 1986.

[15] William M. Gorman. The structure of utility functions. *The Review of Economic Studies*, 35(4):367 – 390, October 1968.

[16] Charles M. Grinstead and Laurie J. Snell. *Introduction to Probability.* American Mathematical Society, 2006.

[17] Edi Karni. On multivariate risk aversion. *Econometrica*, 47(6):1391–1401, 1979.

[18] Edi Karni. On the correspondence between multivariate risk aversion and risk aversion with state-dependent preferences. *Journal of Economic Theory*, 30(2):230 – 242, 1983.

[19] R.L. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Trade-Offs.* Wiley series in probability and mathematical statistics. Applied probability and statistics. Cambridge University Press, 1993.

[20] Richard E. Kihlstrom and Leonard J. Mirman. Risk aversion with many commodities. *Journal of Economic Theory*, 8(3):361 – 388, 1974.

[21] Dilip B. Madan. Measures of risk aversion with many commodities. *Economics Letters*, 11(12):93 – 100, 1983.

[22] Richard F. Meyer. Some Notes on Discrete Multivariate Utility. Technical report, Mimeographed manuscript, Havard Business School, 1972.

[23] Jacob Paroush. Risk premium with many commodities. *Journal of Economic Theory*, 11(2):283 – 286, 1975.

[24] John W. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32(1/2):122 – 136, 1964.

[25] Scott F. Richard. Multivariate risk aversion, utility independence and separable utility functions. *Management Science*, 22(1):12–21, 1975.

[26] Ariel Rubinstein. Equilibrium in supergames with the overtaking criterion. *Journal of Economic Theory*, 21(1):1 – 9, 1979.

[27] Marco Scarsini. Dominance conditions for multivariate utility functions. *Management Science*, 34(4):454–460, 1988.

[28] Joseph E Stiglitz. Behavior towards risk with many commodities. *Econometrica*, pages 660–667, 1969.

[29] Menahem E Yaari. The dual theory of choice under risk. *Econometrica*, 55(1):95–115, 1987.

Appendix A. Proofs

For readability, all theorems and propositions are restated in this appendix.

Proofs for Section 4. The proofs in this section follow certain conventions that simplify the presentation:

- x, y, are real number, α, β, δ - with or without indices or primes - are positive reals.
- a_i, b_i, and c_i are points in $\mathcal{T}_i$.
- L_i is a lottery over $\mathcal{T}_i$ and ℓ_i is the realization of L_i.
- Variables not explicitly quantified are taken to be universally quantified, it being understood that the expressions in which they appear are defined.

Proposition 4.1. *There exist Debreu value functions $v^{\mathcal{T}_i} : \mathcal{T}_i \to \mathbb{R}$, $i = 1, 2, \ldots$, such that for all n, $v_n = \sum_{i=1}^n v^{\mathcal{T}_i}$ represents $\precsim^n$.*

Proof. Consider $\mathcal{H}^n$ for $n \geq 3$. By assumption, any product of the $\mathcal{T}_i$'s is independent. Hence, there exist value functions $v_n^{\mathcal{T}_1}, \ldots, v_n^{\mathcal{T}_n}$, with $\sum_{i=1}^n v_n^{\mathcal{T}_i}$ representing $\precsim^n$. We now show that there is actually a *single* function $v^{\mathcal{T}_i}$, for each i, that works for all the $\mathcal{H}^n$'s.

For $i = 1, 2, 3$, set $v^{\mathcal{T}_i} := v_3^{\mathcal{T}_i}$. Suppose $v^{\mathcal{T}_i}$ has been defined for all $i < n$; we inductively define $v^{\mathcal{T}_n}$. By the induction hypothesis, $\sum_{i=1}^{n-1} v^{\mathcal{T}_i}$ represents $\precsim^{n-1}$. By independence of $\mathcal{H}^{n-1}$ in $\precsim^n$, the function $\sum_{i=1}^{n-1} v_n^{\mathcal{T}_i}$ also represents $\precsim^{n-1}$. So, by uniqueness of the value functions, there exist constants $\beta > 0, \xi_i$, such that $v^{\mathcal{T}_i} = \beta v_n^{\mathcal{T}_i} + \xi_i$, for $i = 1, \ldots, n-1$. So, setting $v^{\mathcal{T}_n} = \beta v_n^{\mathcal{T}_n}$, we have that

$$\sum_{i=1}^{n} v^{\mathcal{T}_i} = \sum_{i=1}^{n-1} (\beta v_n^{\mathcal{T}_i} + \xi_i) + \beta v_n^{\mathcal{T}_n} = \beta \sum_{i=1}^{n} v_n^i + \textit{constant},$$

which represents $\precsim^n$, as required. $\square$

From now on we assume w.l.o.g. that the factors are already represented in units of the respective value functions; that is, $v^{\mathcal{T}_i}(a_i) = a_i$ for all i and $a_i \in \mathcal{T}_i$. Then u_n, the NM utility function representing $\overset{\vartriangle}{\precsim}^n$, is actually only a function of the sum of its arguments; i.e. $u_n(a_1, \ldots, a_n) = u_n(b_1, \ldots, b_n)$ whenever $a_1 + \cdots + a_n = b_1 + \cdots + b_n$. Recall that $\hat{u}_n$ is the function such that $u_n(a_1, \ldots, a_n) = \hat{u}_n(a_1 + \cdots + a_n)$. Note that $\hat{u}_n = u_n \circ (v_n)^{-1}$. Thus, u_n is concave with respect to v_n if and only if $\hat{u}_n$ is concave.

Let $(\phi_2, \phi_3, \ldots)$ be the presumed future. By assumption $(\phi_2, \phi_3, \ldots)$ is internal.[20] So, there exists $s > 0$ with $\phi_i \pm s \in \mathcal{T}_i$, for all i.

[20]More precisely, $(\phi_2, \phi_3, \ldots)$ is *a* presumed future that is internal, if there are several presumed futures.

<u>*Proofs for Section 4.2.1.*</u>

Lemma A.1. *Let $X_1, X_2, \dots$ be an infinite sequence of independent uniformly bounded random variables,*[21] *with $E(X_i) = 0$ for all i. Set $S_n = \sum_{i=1}^{n} X_i$. Then*

$$\Pr[S_n \geq 0 \textit{ infinitely often}] > 0. \tag{5}$$

Proof. Denote $v_i = \mathrm{Var}(X_i)$, and $V_n = \sum_{i=1}^{n} v_i$. The X_i's are independent, so $V_n = \mathrm{Var}(S_n)$. Now, either $V_n \to \infty$ or not. We consider each case separately.

If $V_n \to \infty$, applying the central limit theorem for uniformly bounded random variables (e.g. [16], Theorem 9.5) we obtain that

$$\lim_{n\to\infty} \Pr[\frac{S_n}{\sqrt{V_n}} \geq 0] = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-x^2/2} dx = \frac{1}{2}.$$

In particular, $\Pr[S_n \geq 0$ infinitely often$] > 0$.

Next, suppose that V_n does not go to infinity. Each v_i is non-negative. Hence, the V_i's form a monotonically non-decreasing and bounded sequence, and hence converge. Thus, for any $\delta > 0$ there exists an N_δ with $\sum_{i=N_\delta}^{\infty} v_i < \delta$. If all the X_i are identically 0 there is nothing to prove. Otherwise, w.l.o.g. X_1 is not identically 0. Thus there exists an $x > 0$ with $\Pr(X_1 \geq x) = q_x > 0$. Choose $\delta < x^2$. Then by the Chebyshev inequality, for all $n > N_\delta$,

$$\Pr[\sum_{i=N_\delta}^{n} X_i < -x] < \frac{\mathrm{Var}(\sum_{i=N_\delta}^{n} X_i)}{x^2} \leq \frac{\delta}{x^2} < 1.$$

Clearly, there is some probability p^+ for which $\Pr[\max_{n=2,\dots,N_\delta}\{S_n - X_1\} \geq 0] \geq p^+$. So for all n,

$$\Pr[S_n \geq 0] \geq \Pr[X_1 \geq x] \cdot \Pr[\max_{n=2,\dots,N_\delta}(S_n - X_1) \geq 0] \cdot \Pr[\sum_{i=N_\delta}^{n} X_i \geq -x] \geq$$
$$q_x \cdot p^+ \cdot (1 - \frac{\delta}{x^2}) > 0.$$

So, again, in particular, $\Pr[S_n \geq 0$ infinitely often$] > 0$. $\square$

Theorem 1. *$\stackrel{\star}{\succsim}$ is weakly risk averse if and only if u_n is concave with respect to v_n for all n.*

Proof. <u>$\stackrel{\star}{\succsim}$ is weakly risk averse $\Rightarrow$ all $\hat{u}_n$ are concave:</u> Contrariwise, suppose that $\hat{u}_k$ is not concave, for some k. So, $\hat{u}_k$ is not concave on some interval of size $\leq s$. So, there exist $x, \epsilon \leq s$ and $0 < \beta < \epsilon$ with

$$\hat{u}_k(x + \beta) = \frac{1}{2}\left(\hat{u}_k(x - \epsilon) + \hat{u}_k(x + \epsilon)\right).$$

So, by definition of the presumed future also for any $m > k$,

$$\begin{aligned} &\hat{u}_m(x + \phi_{k+1} + \cdots + \phi_m + \beta) = \\ &\quad = \frac{1}{2}\left(\hat{u}_m(x + \phi_{k+1} + \cdots + \phi_m - \epsilon) + \hat{u}_m(x + \phi_{k+1} + \cdots + \phi_m + \epsilon)\right). \end{aligned} \tag{6}$$

[21] that is, the support of all the random variables is included in a real interval $[\underline{b}, \overline{b}]$, with $\underline{b}, \overline{b}$ finite.

We construct a recurring lottery sequence $\boldsymbol{L}$ that is ultimately inferior to its repeated certainty equivalent. By definition, $x = b_1 + \cdots + b_k$, for some $(b_1, \ldots, b_k) \in \mathcal{H}^k$. The sequence $\boldsymbol{L} = (L_1, L_2, \ldots)$ is defined as follows:

- for $i = 1, \ldots, k$: $L_i = b_i$;
- for j odd: $L_{k+j} = \langle(\phi_{k+j} - \epsilon), (\phi_{k+j} + \epsilon)\rangle$;
- for j even: $L_{k+j} = \phi_{k+j} - \beta$.

We now inductively determine the repeated certainty equivalent of $\boldsymbol{L} = (L_1, L_2, \ldots)$, which we denote $(c_1, c_2, \ldots)$. For $i = 1, \ldots, k$, $c_i = b_i$. Consider the lottery at time $k+1$. The (degenerate) lotteries in the previous times have brought us to the point $x = b_1 + \cdots + b_k$, and the lottery at time $k+1$ is $L_{k+1} = \langle(\phi_{k+1} - \epsilon), (\phi_{k+1} + \epsilon)\rangle$. So, by (6), its certainty equivalent is β above the average; that is, $c_{k+1} = \phi_{k+1} + \beta$. The next lottery, at time $k+2$, is the degenerate lottery $L_{k+2} = \phi_{k+2} - \beta$, with certainty equivalent $c_{k+2} = \phi_{k+2} - \beta$. Hence, having chosen the certainty equivalent at all times, after time $k+2$ we are at point $x + c_{k+1} + c_{k+2} = x + \phi_{k+1} + \phi_{k+2}$. So again (6) applies to the lottery at time $k+3$, which is $L_{k+3} = \langle(\phi_{k+1} - \epsilon), (\phi_{k+1} + \epsilon)\rangle$. So $c_{k+3} = \phi_{k+3} + \beta$. This process repeats again and again. So, $c_{k+j} = \phi_{k+j} + \beta$ for j odd and $c_{k+j} = \phi_{k+j} - \beta$ for j even.

Now, assume w.l.o.g. that $E(L_i) = 0$ for all i. Then, for j odd, L_{k+j} is a $\pm\epsilon$ lottery and $c_{k+j} = \beta$. For all other i's, ℓ_i is a degenerate lottery and $c_i = 0$. Let ℓ_i be the realization of L_i. Then,

$$\Pr[(c_1, \ldots, c_n) \succ (\ell_1, \ldots, \ell_n) \text{ from some } n \text{ on}] = \Pr[\frac{n-k}{2}\beta > \sum_{i=1}^{n} \ell_i \text{ from some } n \text{ on}] = 1,$$

where the last equality is by the law of large numbers. So, $(c_1, c_2, \ldots)$ is ultimately superior to $(L_1, L_2, \ldots)$.

All $\hat{u}_n$ are concave $\Rightarrow \stackrel{s}{\precsim}$ is weakly risk averse: Consider a lottery sequence $\boldsymbol{L} = (L_1, L_2, \ldots)$. W.l.o.g. $E(L_i) = 0$ for all i. Denote by $\boldsymbol{c} = (c_1, c_2, \ldots)$ the repeated certainty equivalent of $\boldsymbol{L}$. Since all $\hat{u}_n$'s are concave, also all the functions u_n are concave in each of their arguments. So, $c_i \leq 0$ for all i. So, for any n,

$$\Pr[(\ell_1, \ldots, \ell_n) \prec^n (c_1, \ldots, c_n)] \leq \Pr[\sum_{i=1}^{n} \ell_i < 0].$$

So,

$$\Pr[(\ell_1, \ldots, \ell_n) \prec^n (c_1, \ldots, c_n) \text{ from some } n \text{ on}] \leq (1 - \Pr[\sum_{i=1}^{n} \ell_i \geq 0 \text{ infinitely often}]) < 1.$$

where the last inequality is by Lemma A.1. So, $(c_1, c_2, \ldots)$ is not ultimately superior to $(L_1, L_2, \ldots)$.

$\square$

Proofs for Section 4.2.2. Theorem 2 follows directly from Theorem 3 (a) and Proposition 4.2. So, proceed to prove this theorem and proposition.

For $\alpha > 0$ let $cara_\alpha$ be the function $cara_\alpha(x) = -e^{-\alpha x}$. It is well known that $A_{cara_\alpha}(x) = \alpha$ for all x. For a real-valued lottery L and NM utility function f let $\textit{risk-prem}_f(x, L)$ be the risk-premium according to f of the lottery L applied at wealth x.

Lemma A.2. *$RP_{\hat{\mathbf{u}}}(\epsilon) = \Omega(\epsilon^2)$ as $\epsilon \to 0$ if and only if there exists an α such that*

$$\textit{risk-prem}_{\hat{u}_n}(x, L) \geq \textit{risk-prem}_{cara_\alpha}(x, L) \tag{7}$$

for all n, x and L.

Proof. Suppose that $RP_{\hat{\mathbf{u}}}(\epsilon) = \Omega(\epsilon^2)$. Then there exists ϵ_0 and $\alpha > 0$ with

$$\textit{risk-prem}_{\hat{u}_n}(x, \pm\epsilon) \geq \alpha\epsilon^2 \tag{8}$$

for all n, x and $\epsilon \leq \epsilon_0$.

For the function $cara_\alpha$, using the Taylor expansion of e^ϵ around 0,

$$\begin{aligned}
\frac{cara_\alpha(\epsilon) + cara_\alpha(-\epsilon)}{2} &= \frac{-e^{-\alpha\cdot\epsilon} - e^{\alpha\cdot\epsilon}}{2} \\
&= -\frac{1}{2}(1 - \alpha\epsilon + \frac{\alpha^2\epsilon^2}{2} + 1 + \alpha\epsilon + \frac{\alpha^2\epsilon^2}{2} + O(\epsilon^3)) \\
&= -(1 + \frac{\alpha^2\epsilon^2}{2} + O(\epsilon^3))
\end{aligned} \tag{9}$$

So, for ϵ sufficiently small

$$\frac{cara_\alpha(\epsilon) + cara_\alpha(-\epsilon)}{2} > -(1 + \frac{2\alpha^2\epsilon^2}{3}) > -e^{-\alpha(-2\alpha\epsilon^2/3)} = cara_\alpha(-2\alpha\epsilon^2/3).$$

So,

$$\textit{risk-prem}_{cara_\alpha}(0, \pm\epsilon) < \frac{2}{3}\alpha\epsilon^2.$$

For the function $cara_\alpha$ the risk premium is independent of x, and hence,

$$\textit{risk-prem}_{cara_\alpha}(x, \pm\epsilon) < \frac{2}{3}\alpha\epsilon^2, \tag{10}$$

for all x.

So, combining (8) and (10)

$$\textit{risk-prem}_{\hat{u}_n}(x, \pm\epsilon) > \textit{risk-prem}_{cara_\alpha}(x, \pm\epsilon), \tag{11}$$

for ϵ sufficiently small. But then, by Pratt [24], (11) holds for any lottery L.

Conversely, if $\textit{risk-prem}_{\hat{u}_n}(x, \pm\epsilon) \geq \textit{risk-prem}_{cara_\alpha}(x, \pm\epsilon)$ then by (9)

$$\textit{risk-prem}_{\hat{u}_n}(x, \pm\epsilon) \geq \frac{\alpha\epsilon^2}{2} + O(\epsilon^3),$$

so $RP_{\hat{\mathbf{u}}}(\epsilon) = \Omega(\epsilon^2)$. $\square$

The following simple lemma establishes that any risk premium exhibited by $\hat{u}_k$, for some k, is (re)exhibited by all subsequent $\hat{u}^m$, for $m > k$.

Lemma A.3. *For any $m > k$,*

$$\textit{risk-prem}_{\hat{u}_m}(x + \phi_{k+1} + \ldots, \phi_m, \pm\epsilon) = \textit{risk-prem}_{\hat{u}_k}(x, \pm\epsilon).$$

Proof. Set $\beta = \textit{risk-prem}_{\hat{u}_k}(x, \pm\epsilon)$. By definition

$$\hat{u}_k(x - \beta) = \frac{1}{2}(\hat{u}_k(x - \epsilon) + \hat{u}_k(x + \epsilon)).$$

Let $\boldsymbol{a}_{+\epsilon}, \boldsymbol{a}_{-\epsilon}, \boldsymbol{a}_{-\beta} \in \mathcal{H}^k$ be such that $v_k(\boldsymbol{a}_{+\epsilon}) = x + \epsilon, v_k(\boldsymbol{a}_{-\epsilon}) = x - \epsilon$, and $v_k(\boldsymbol{a}_{-\beta}) = x - \beta$. So,

$$(\boldsymbol{a}_{-\beta}) \stackrel{\vartriangle}{\sim}^k \langle \boldsymbol{a}_{-\epsilon}, \boldsymbol{a}_{+\epsilon} \rangle .$$

By assumption, $\stackrel{\vartriangle}{\precsim}^k$ and $\stackrel{\vartriangle}{\precsim}^m$ agree on the preferences over $\Delta(\mathcal{H}^k)$ when fixing the state in $\mathcal{T}_{k+1} \times \cdots \times \mathcal{T}_m$ to the presumed future $(\phi_{k+1}, \ldots, \phi_m)$. So,

$$(\boldsymbol{a}_{-\beta}, \phi_{k+1}, \ldots, \phi_m) \stackrel{\vartriangle}{\sim}^m \langle (\boldsymbol{a}_{-\epsilon}, \phi_{k+1}, \ldots, \phi_m), (\boldsymbol{a}_{+\epsilon}, \phi_{k+1}, \ldots, \phi_m) \rangle .$$

Hence,

$$\begin{aligned}&\hat{u}_m(x - \beta + \phi_{k+1} + \cdots + \phi_m) = \\ &\frac{1}{2}(\hat{u}_m(x - \epsilon + \phi_{k+1} + \cdots + \phi_m) + \hat{u}_m(x + \epsilon + \phi_{k+1} + \cdots + \phi_m)).\end{aligned}$$

□

The following lemma establishes that if $\hat{u}_k$ exhibits some risk premium, at some point x, then not only is this risk premium re-exhibited by all subsequent utility functions $\hat{u}^m$, but also that it is "reachable" from any state y, of any period K.

Lemma A.4. *For any k, K, x, y, with x in the domain of $\hat{u}_k$ and y in the domain of $\hat{u}_K$, there exist $m \geq \max\{k, K\}$ and $b_{K+1}, \ldots, b_m$, $b_i \in \mathcal{T}_i$, with*

$$\textit{risk-prem}_{\hat{u}_m}(y + b_{K+1} + \cdots + b_m, \pm\epsilon) = \textit{risk-prem}_{\hat{u}_k}(x, \pm\epsilon).$$

Proof. Set $K' = \max\{k, K\}$. If $K < k$ then for $i = K + 1, \ldots, k$, let b_i be any point in $\mathcal{T}_i$ and set $y' = y + b_{K+1} + \cdots + b_k$. Otherwise ($K \geq k$) set $y' = y$.

Let $\delta = y' - x, j = \lceil \delta/s \rceil$, and $m = K' + j$. For $i = K' + 1, \ldots, m$, set $b_i = \phi_i + \delta/j$. Then, $m > \max\{k, K\}$, and $x + \phi_{k+1} + \cdots + \phi_m = y + b_{K+1} + \cdots + b_m$. The result then follows from Lemma A.3 . □

The following Theorem is from Alon and Spencer [1].

Theorem A.5 ([1], Theorem A.1.19)**.** *For every $C > 0$ and $\gamma > 0$ there exists a $\delta > 0$ so that the following holds: Let X_i, $1 \leq i \leq n$, n arbitrary, be independent random variables with $E[X_i] = 0$, $|X_i| \leq C$, and $Var(X_i) = \sigma_i^2$. Set $S_n = \sum_{i=1}^n X_i$ and $\Sigma_n^2 = \sum_{i=1}^n \sigma_i^2$, so that $Var(S_n) = \Sigma_n^2$. Then, for $0 < a \leq \delta \cdot \Sigma_n$*

$$\Pr[S_n > a\Sigma_n] < e^{-\frac{a^2}{2}(1-\gamma)}. \tag{12}$$

Lemma A.6. *Let $X_1, X_2, \ldots,$ be independent random variables with $E[X_i] = 0$, $|X_i| \leq C$, and $Var(X_i) = \sigma_i^2$. Set S_n, σ_i^2 and Σ_n^2 as above. If $\Sigma_n \to \infty$, then for any $\alpha > 0$*

$$\Pr[S_n > \alpha\Sigma_n^2 \text{ infinitely often}] = 0.$$

Proof. Denote by $n(i)$ the least n such that $\Sigma_n^2 \geq i$. Since $\Sigma_n \to \infty$, for any i there exists an $n(i)$. Since $|X_i| \leq C$, $i \leq \Sigma_{n(i)}^2 \leq i + C^2$.

Denote by A_k the event that there exists i, $n(k) < i \leq n(k+1)$, for which $S_i > \alpha\Sigma_i^2$. We bound $\Pr[A_k]$.

Set $\gamma = 0.5$, and let δ be that provided by Theorem A.5. Set $\beta = \min\{\delta, \alpha/2\}$. Then, considering $n(k)$, by Theorem A.5, setting $a = \beta\Sigma_{n(k)}$

$$\Pr[S_{n(k)} > \beta\Sigma_{n(k)} \cdot \Sigma_{n(k)}] < e^{-\frac{\beta^2\Sigma_{n(k)}^2}{2}(1-\gamma)} \leq e^{-\frac{\beta^2 k}{4}} \tag{13}$$

Now consider the random variables X_i for $i = n(k)+1, \ldots, n(k+1)$. Set $D_j = \sum_{i=n(k)+1}^{j} X_i$. Then,

$$Var((D_{n(k+1)}) = \Sigma_{n(k+1)}^2 - \Sigma_{n(k)}^2 \leq (k+1+C^2) - k = 1 + C^2.$$

So, by the Kolmogorov inequality

$$\Pr[\max_{n(k)<j\leq n(k+1)}\{D_j\} \geq \beta\Sigma_{n(k)}^2] \leq \frac{Var(D_{n(k+1)})}{(\beta\Sigma_{n(k)}^2)^2} \leq \frac{1+C^2}{\beta^2 k^2}. \tag{14}$$

Combining (13)-(14), for any k

$$\begin{aligned}
\Pr[A_k] &= \Pr[\exists i, n(k) < i \leq n(k+1), S_i > \alpha\Sigma_i^2] \\
&\leq \Pr[S_{n(k)} \geq \beta\Sigma_{n(k)}^2] + Pr[\max_{n(k)<j\leq n(k+1)}\{D_j\} \geq \beta\Sigma_{n(k)}^2] \\
&\leq e^{-\frac{\beta^2 k}{4}} + \frac{1+C^2}{\beta^2 k^2}.
\end{aligned}$$

So, $\sum_{k=1}^{\infty} \Pr[A_k] < \infty$. So, by the Borel Cantelli lemma

$$\Pr[A_k \text{ occurs infinitely often}] = 0.$$

For any k there is only a finite number of i's with $n(k) < i \leq n(k+1)$. So, $S_i > \alpha\Sigma_i^2$ infinitely often only if A_k occurs infinitely often, and the result follows. ∎

Theorem 3.

(a) If $RP_{\hat{\mathbf{u}}}(\epsilon) = \Omega(\epsilon^2)$ *as* $\epsilon \to 0$ *then* $\overset{\Delta}{\precsim}$ *is risk averse.*[22]

(b) If $RP_{\hat{\mathbf{u}}}(\epsilon) = O(\epsilon^{2+\beta})$ *as* $\epsilon \to 0$*, for some* $\beta > 0$*, then* $\overset{\Delta}{\precsim}$ *is not risk averse.*

Proof. (a): Suppose that $RP_{\hat{\mathbf{u}}}(\epsilon) = \Omega(\epsilon^2)$ as $\epsilon \to 0$.

Let $\boldsymbol{L} = (L_1, L_2, \ldots)$ be a bounded, non-vanishing lottery sequence. W.l.o.g. $E(L_i) = 0$ for all i. Set $\sigma_i^2 = Var(L_i)$, $S_n = \sum_{i=1}^n L_i$ and $\Sigma_n^2 = Var(S_n) = \sum_{i=1}^n \sigma_i^2$. Since $\boldsymbol{L}$ is non-vanishing $\Sigma_n \to \infty$. Since $\boldsymbol{L}$ is bounded, there exists a C such that $|L_i| \le C$ for all i.

By the Taylor expansion,

$$cara_\alpha(\epsilon) = -e^{-\alpha\epsilon} = -1 + \alpha\epsilon - \frac{\alpha^2\epsilon^2}{2} + O(\alpha^3\epsilon^3). \tag{15}$$

Let α_1 be such that the $O(\alpha^3\epsilon^3)$ term in (15) is small for $|\epsilon| \le C$; that is,

$$cara_{\alpha_1}(\epsilon) \approx -1 + \alpha_1\epsilon - \frac{\alpha_1^2\epsilon^2}{2}, \tag{16}$$

for $|\epsilon| \le C$.

Let $(c_1, c_2, \ldots)$ be the repeated certainty equivalent of $\boldsymbol{L}$. Let α_0 be that provided by Lemma A.2. Then, for any $\alpha < \alpha_0$

$$c_i < -\textit{risk-prem}_{cara_\alpha}(0, L_i).$$

Set $\alpha = \min\{\alpha_0, \alpha_1\}$. Suppose that L_i gets values $x_1^i, \ldots, x_m^i$ with probabilities $p_1, \ldots, p_m$, respectively. Then,

$$\begin{aligned}
c_i < -\textit{risk-prem}_{cara_\alpha}(0, L_i) =& cara_\alpha^{-1}\left(\sum_{j=1}^m cara_\alpha(x_j^i)p_j\right) \\
\approx& cara_\alpha^{-1}\left(\sum_{j=1}^m (-1 + \alpha x_j^i - \frac{\alpha^2(x_j^i)^2}{2})p_j\right) \\
=& cara_\alpha^{-1}\left(\sum_{j=1}^m (-1)p_j + \alpha\sum_{j=1}^m x_j^i p_j - \sum_{j=1}^m \frac{\alpha^2(x_j^i)^2}{2}p_j\right) \\
=& cara_\alpha^{-1}\left(-1 + 0 - \frac{\alpha^2\sigma_i^2}{2}\right) \\
\approx& cara_\alpha^{-1}\left(-e^{-\alpha(-\alpha\sigma_i^2/2)}\right) < -\alpha\sigma_i^2.
\end{aligned}$$

So,

$$\left[-\alpha\cdot(\Sigma_n)^2 < S_n\right] \Rightarrow \left[\sum_{i=1}^n c_i < S_n\right] \Rightarrow \left[(c_1, \ldots, c_n) \prec (\ell_1, \ldots, \ell_n)\right]. \tag{17}$$

[22] recall that $g(y) = \Omega(h(y))$ as $y \to 0$ if there exists a constant M and y_0 such that $g(y) > M \cdot h(y)$ for all $y < y_0$.

So, it is sufficient to prove that

$$\Pr[S_n > -\alpha(\Sigma_n)^2 \text{ from some } n \text{ on}] = 1.$$

which is equivalent to saying that

$$\Pr[S_n < -\alpha(\Sigma_n)^2 \text{ infinitely often}] = 0, \tag{18}$$

which is provided by Lemma A.6 (by symmetry).

(b): Suppose that $RP_{\hat{\boldsymbol{u}}}(\epsilon) = O(\epsilon^{2+\beta})$ as $\epsilon \to 0$, with $\beta > 0$. So, there exists α and ϵ_0 such that for any $\epsilon < \epsilon_0$, there exists an i and x with

$$\textit{risk-prem}_{\hat{u}_i}(x, \pm\epsilon) \le \alpha \cdot \epsilon^{2+\beta}. \tag{19}$$

Set $\epsilon_1 = \min\{\epsilon_0^2, s^2\}$. For $j = 1, 2, \ldots$, set a_j as follows:

$$a_j = \begin{cases} \sqrt{\epsilon_1} & \text{if } j = 3^{k^2} \text{ for some integral } k \\ \sqrt{\epsilon_1}\frac{1}{\sqrt{j}} & \text{otherwise} \end{cases}$$

So, by (19), for any j there exists i_j and x_j with

$$\textit{risk-prem}_{\hat{u}_{i_j}}(x_j, \pm a_j) \le \alpha \cdot a_j^{2+\beta}. \tag{20}$$

We construct a bounded, non-vanishing lottery sequence $\boldsymbol{L} = (L_1, L_2, \ldots)$ that is not ultimately superior to its repeated certainty equivalent, which we denote by $(c_1, c_2, \ldots)$. The construction of $\boldsymbol{L}$ is inductive, wherein the lotteries are defined in *chunks*. For each j, the j-th chunk consists of a sequence of degenerate lotteries, followed by a single $\pm a_j$ lottery, with which the chunk ends. We denote by $n(j)$ the index of the last lottery in the j-th chunk. The chunks are constructed as follows. Set $n(0) = 0$. Suppose $L_1, \ldots, L_{n(j-1)}$ have been defined, and that their repeated certainty equivalent is $c_1, \ldots, c_{n(j-1)}$. Let i_j, x_j be as in (20). Set $y_{n(j-1)} = c_1 + \cdots + c_{n(j-1)}$. By Lemma A.4 and (20), there exists $m > \max\{n(j-1), i_j\}$ and $b_{n(j-1)+1}, \ldots, b_m$, with

$$\textit{risk-prem}_{\hat{u}_m}(y_{n(j-1)} + b_{n(j-1)+1} + \cdots + b_m, \pm a_j) \le \alpha a_j^{2+\beta}.$$

Hence also (moving to $m+1$)[23],

$$\textit{risk-prem}_{\hat{u}_{m+1}}(y_{n(j-1)} + b_{n(j-1)+1} + \cdots + b_m + \phi_{m+1}, \pm a_j) \le \alpha a_j^{2+\beta},$$

which means that

$$\begin{aligned} &\hat{u}_{m+1}(y_{n(j-1)} + b_{n(j-1)+1} + \cdots + b_m + \phi_{m+1} - (\alpha a_j^{2+\beta})) \le \\ &\quad \le \frac{1}{2}(\hat{u}_{m+1}(y_{n(j-1)} + b_{n(j-1)+1} + \cdots + b_m + \phi_{m+1} - a_j) + \hat{u}_{m+1}(y_{n(j-1)} + b_{n(j-1)+1} + \cdots + b_m + \phi_{m+1} + a_j)). \end{aligned}$$

[23]We move to $m+1$ with ϕ_{m+1} to guarantee sufficient distance from the boundaries to allow a $\pm a_j$ lottery.

Accordingly, set $L_i = b_i$, for $i = n(j-1)+1, \ldots, m$ and $L_{m+1} = \langle(\phi_{m+1} - a_j), (\phi_{m+1} + a_j)\rangle$. By construction, $c_i = b_i$ for $i = n(j-1)+1, \ldots, m$, and

$$c_{m+1} \geq \phi_{m+1} - \alpha a_j^{2+\beta}. \tag{21}$$

Denote $n(j) = m+1$; that is, $n(j)$ is the index of the $\pm a_j$ lottery.

We now show that $(c_1, c_2, \ldots)$, is not ultimately inferior to $(L_1, L_2, \ldots)$. W.l.o.g. $E(L_i) = 0$ for all i. So, we have that $L_i = \langle(-\sigma_i), (\sigma_i)\rangle$ with

$$\sigma_i = \begin{cases} \sqrt{\epsilon_1} & \text{if } i = n(j) \text{ with } j = 3^{k^2} \text{ for some integral } k \\ \sqrt{\epsilon_1}\frac{1}{\sqrt{j}} & \text{if } i = n(j) \text{ for other } j\text{'s} \\ 0 & \text{otherwise} \end{cases}$$

and

$$c_i \geq \begin{cases} -\alpha(\epsilon_1)^{1+\beta/2} & \text{if } i = n(j) \text{ with } j = 3^{k^2} \text{ for some integral } k \\ -\alpha(\epsilon_1)^{1+\beta/2} \cdot \frac{1}{j^{1+\beta/2}} & \text{if } i = n(j) \text{ for other } j\text{'s} \\ 0 & \text{otherwise} \end{cases}$$

Let $S_n = \sum_{i=1}^n L_i$. So, $Var(S_n) = \sum_{i=1}^n \sigma_i^2$. So, for $n = n(3^{k^2})$,

$$Var(S_{n(3^{k^2})}) \geq \sum_{j=1}^{3^{k^2}} \frac{\epsilon_1}{j} > \sum_{j=1}^{e^{k^2}} \frac{\epsilon_1}{j} > \epsilon_1 \cdot k^2.$$

On the other hand,

$$\sum_{i=1}^{n(3^{k^2})} c_i \geq -\alpha(\epsilon_1)^{1+\beta/2} \left(\sum_{j=1}^{3^{k^2}} \frac{1}{j^{1+\beta/2}} + k \right) > -\alpha(\epsilon_1)^{1+\beta/2} (D+k),$$

for $D = \sum_{j=1}^{\infty} \frac{1}{j^{1+\beta/2}} < \infty$.

Set $\gamma = \alpha(\epsilon_1)^{1+\beta/2}$. Then, for k sufficiently large

$$\begin{aligned}
\Pr[(\ell_1, \dots, \ell_{n(3^{k^2})}) \precsim (c_1, \dots, c_{n(3^{k^2})})] &= \Pr\left[S_{n(3^{k^2})} \leq \sum_{i=1}^{n(3^{k^2})} c_i\right] \geq \\
&\geq \Pr\left[S_{n(3^{k^2})} \leq -\gamma\,(D+k)\right] = \\
&= \Pr\left[\frac{S_{n(3^{k^2})}}{Var(S_{n(3^{k^2})})^{1/2}} \leq -\gamma\frac{(D+k)}{Var(S_{n(3^{k^2})})^{1/2}}\right] \geq \\
&\geq \Pr\left[\frac{S_{n(3^{k^2})}}{Var(S_{n(3^{k^2})})^{1/2}} \leq -\gamma\frac{(D+k)}{\sqrt{\epsilon_1}\cdot k}\right] \geq \\
&\geq \Pr\left[\frac{S_{n(3^{k^2})}}{Var(S_{n(e^{k^2})})^{1/2}} \leq -\gamma\cdot\frac{2}{\sqrt{\epsilon_1}}\right] \approx \\
&\approx \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-2\gamma\epsilon_1^{-1/2}} e^{-x^2/2}dx = p > 0,
\end{aligned}$$

for some constant p. In particular, $(\ell_1, \dots, \ell_{n(3^{k^2})}) \precsim (c_1, \dots, c_{n(3^{k^2})})$ for infinitely many k's, with probability 1. □

Proposition 4.2. *$RP_{\hat{\mathbf{u}}}(\epsilon) = \Omega(\epsilon^2)$ as $\epsilon \to 0$, if and only if $A_{\hat{u}_n}(x)$ is bounded away from 0, uniformly for all n and x (assuming $\hat{u}_n$ is twice differentiable for all n).*

Proof. Follows directly from Lemma A.2 and the fact that $A_{cara_\alpha}(x) = \alpha$ for all x. □

Theorem 4. *For v_n, u_n, and $\hat{\mathbf{u}}$ as in Theorems 1 and 3*

(a) Weak risk loving: $\stackrel{\Delta}{\precsim}$ is weakly risk loving if and only if u_n is convex with respect to v_n for all n.

(b) Risk loving

- *If $(-RP_{\hat{\mathbf{u}}}(\epsilon)) = \Omega(\epsilon^2)$ as $\epsilon \to 0$ then $\stackrel{\Delta}{\precsim}$ is risk loving.*
- *If $(-RP_{\hat{\mathbf{u}}}(\epsilon)) = O(\epsilon^{2+\beta})$ as $\epsilon \to 0$ (for some $\beta > 0$) then $\stackrel{\Delta}{\precsim}$ is not risk loving.*

(c) Risk Neutral: $\stackrel{\Delta}{\precsim}$ is risk neutral if and only if u_n is a linear transformation of v_n for all n.

Proof. The proofs of (a) and (b) are analogous to those of Theorems 3 and 1. (c) follows from combining Theorems 3 and 4. □

Proofs for Sections 5 and 6. Many of the claims of Section 5 become easier to analyze and prove with the aid of the value function introduced in Section 6. Hence, we first provide the proofs for Section 6 and then come back and prove those of Section 5.

Throughout, the following notation is used:

- v denotes a Debreu value function on $\mathcal{S}$, and $v^{\mathcal{T}_i}$ a Debreu value function on the factor $\mathcal{T}_i$.

- u denotes an NM utility function on $\mathcal{S}$. An NM utility for $\mathcal{S}$ necessarily exists since the NM axioms are assumed to hold, and we consider only lotteries with finite support (see Fishburn [13, Theorem 8.2]).

<u>Preliminaries</u>

Lemma A.7. *u is continuous.*

Proof. It suffices to prove that the pre-images of the *open rays* $(-\infty, r)$ and (r, ∞) are open, for all r (these open rays constitute a subbase for the standard topology on the line). Consider $(-\infty, r)$ (the other case is analogous). If $u(s) \geq r$ for all $s \in \mathcal{S}$ then $u^{-1}(-\infty, r) = \emptyset$, which is open. Similarly, if $u(s) < r$ for all $s \in \mathcal{S}$ then $u^{-1}(-\infty, r) = \mathcal{S}$, which is open. Otherwise, there exist $s_1 < r \leq s_2$ and $\hat{s}_1, \hat{s}_2 \in \mathcal{S}$, with $u(\hat{s}_1) = s_1, u(\hat{s}_2) = s_2$. Set $\hat{p} = (r - s_1)/(s_2 - s_1)$. Then, $r = \hat{p}s_1 + (1 - \hat{p})s_2$. Since $\stackrel{\vartriangle}{\precsim}$ is continuous the set

$$u^{-1}(-\infty, r) = \{s : u(s) < r\} = \{s : s \stackrel{\vartriangle}{\prec} \langle \hat{s}_1, \hat{s}_2 : \hat{p}, (1 - \hat{p})\rangle\}$$

is open, by definition (where $\langle \hat{s}_1, \hat{s}_2 : \hat{p}, (1 - \hat{p})\rangle$ is the lottery with value $\hat{s}_1$ with probability $\hat{p}$ and $\hat{s}_2$ with probability $1 - \hat{p}$). $\square$

<u>*Proofs for Section 6.*</u>

Each factor $\mathcal{T} = \mathcal{T}_i$ is a product of some set of commodity spaces, that is $\mathcal{T} = \prod_{j \in T} \mathscr{C}_i$, for some index set T. For factors $\mathcal{T} = \prod_{j \in T} \mathscr{C}_j$ and $\mathcal{R} = \prod_{j \in R} \mathscr{C}_j$, by a slight abuse of notation, we write $\mathcal{T} \cap \mathcal{R}$ for $\prod_{j \in T \cap R} \mathscr{C}_j$, $\mathcal{T} - \mathcal{R}$ for $\prod_{j \in T - R} \mathscr{C}_j$, and $\mathcal{T} \subseteq \mathcal{R}$ if $T \subseteq R$. We say that $\mathcal{T}$ and $\mathcal{R}$ *overlap* if $T \cap R \neq \emptyset$ and neither is contained in the other; the factor $\mathcal{T}$ is *non-degenerate* if $T \neq \emptyset$.

Lemma A.8. *If there exist two non-identical independent partitions $\mathcal{S} = \mathcal{A} \times \mathcal{B}$ and $\mathcal{S} = \mathcal{C} \times \mathcal{D}$, then there exist value functions $v^{\mathcal{A}}, v^{\mathcal{B}}, v^{\mathcal{C}}$, and $v^{\mathcal{D}}$ (for $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$), such that*

(1) $v^{\mathcal{A}} + v^{\mathcal{B}}$ and $v^{\mathcal{C}} + v^{\mathcal{D}}$ both represent $\precsim$,
(2) $v^{\mathcal{A}} + v^{\mathcal{B}} = v^{\mathcal{C}} + v^{\mathcal{D}}$,
(3) if $\hat{v}^{\mathcal{A}}, \hat{v}^{\mathcal{B}}$ are value functions for $\mathcal{A}, \mathcal{B}$, and $\hat{v}^{\mathcal{C}}, \hat{v}^{\mathcal{D}}$, are value functions for $\mathcal{C}, \mathcal{D}$, then $\hat{v}^{\mathcal{A}} + \hat{v}^{\mathcal{B}}$ is a positive affine transformation of $\hat{v}^{\mathcal{C}} + \hat{v}^{\mathcal{D}}$.

Proof. Gorman [15, Theorem 1] proves that if two independent factors $\mathcal{E}$ and $\mathcal{F}$ overlap then $\mathcal{E} \cup \mathcal{F}, \mathcal{E} \cap \mathcal{F}, \mathcal{E} - \mathcal{F}, \mathcal{F} - \mathcal{E}$, and $\mathcal{E} \triangle \mathcal{F} = (\mathcal{E} - \mathcal{F}) \cup (\mathcal{F} - \mathcal{E})$ are all independent.

Set $\mathcal{W} = \mathcal{A} \cap \mathcal{C}, \mathcal{X} = \mathcal{A} \cap \mathcal{D}, \mathcal{Y} = \mathcal{B} \cap \mathcal{C}$, and $\mathcal{Z} = \mathcal{B} \cap \mathcal{D}$. Then, by Gorman's theorem, $\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are independent, as is any product thereof. Since the partitions are not identical, at least three out of $\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are non-degenerate. So, $\mathcal{S} = \mathcal{W} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ is an independent partition with at least 3 factors. So, by Debreu [7], there are value functions $v^{\mathcal{W}}, v^{\mathcal{X}}, v^{\mathcal{Y}}$, and $v^{\mathcal{Z}}$, with $v^{\mathcal{W}} + v^{\mathcal{X}} + v^{\mathcal{Y}} + v^{\mathcal{Z}}$ representing $\precsim$. So, the pair of functions $v^{\mathcal{A}} = v^{\mathcal{W}} + v^{\mathcal{X}}$ and $v^{\mathcal{B}} = v^{\mathcal{Y}} + v^{\mathcal{Z}}$ are value functions for the independent partition $\mathcal{S} = \mathcal{A} \times \mathcal{B}$. Similarly, the functions $v^{\mathcal{C}} = v^{\mathcal{W}} + v^{\mathcal{Y}}$, and $v^{\mathcal{D}} = v^{\mathcal{X}} + v^{\mathcal{Z}}$ are value functions for the independent partition $\mathcal{S} = \mathcal{C} \times \mathcal{D}$, proving (1) and (2). Finally, (3) follows from (2) by the uniqueness of value functions. $\square$

Theorem 7. *For any $\mathcal{S}$, all (aggregate) Debreu value functions for $\mathcal{S}$ are identical up to positive affine transformations.*

Proof. Suppose $\mathcal{S}$ has two different independent partitions $\mathcal{S} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$ and $\mathcal{S} = \mathcal{C}_1 \times \cdots \times \mathcal{C}_m$, with value functions $v^{\mathcal{A}_1}, \ldots, v^{\mathcal{A}_n}$ and $v^{\mathcal{C}_1}, \ldots, v^{\mathcal{C}_m}$, respectively. Since the two partitions are different, there must be some $\mathcal{A}_i$ for which there is no j with $\mathcal{C}_j = \mathcal{A}_i$. W.l.o.g. this is $\mathcal{A}_1$. Set $\mathcal{B} = \mathcal{A}_2 \times \cdots \times \mathcal{A}_n$ and $v^{\mathcal{B}} = \sum_{i=2}^{n} v^{\mathcal{A}_i}$. Similarly, set $\mathcal{D} = \mathcal{C}_2 \times \cdots \times \mathcal{C}_j$ and $v^{\mathcal{D}} = \sum_{i=2}^{j} v^{\mathcal{C}_i}$. Then, $v^{\mathcal{A}_1} + v^{\mathcal{B}}$ represents $\precsim$, as does $v^{\mathcal{C}_1} + v^{\mathcal{D}}$. So, by Lemma A.8-(3), $\sum_{i=1}^{n} v^{\mathcal{A}_i} = v^{\mathcal{A}_1} + v^{\mathcal{B}}$ is an affine transformation of $\sum_{i=1}^{m} v^{\mathcal{C}_i} = v^{\mathcal{C}_1} + v^{\mathcal{D}}$. $\square$

Theorem 8 is essentially a direct corollary of Theorem 4(a) of Epstein and Tanny[**?**], which states the following (the notations and wording have been modified to match those of this paper):

Theorem A.9 ([12], Theorem 4(a))**.** *Let $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$, (with $\mathcal{X}_1, \mathcal{X}_2 \subseteq \mathbb{R}$), be an independent partition. Let $\stackrel{\Delta}{\precsim}^{\mathcal{X}}$ be a preference order on $\Delta(\mathcal{X})$, with NM utility $u(x_1, x_2)$. If $u^{\mathcal{X}}(x_1, x_2) = \phi(\alpha_1 x_1 + \alpha_2 x_2)$, for some function ϕ, and constants $\alpha_1, \alpha_2 > 0$, then $\stackrel{\Delta}{\precsim}^{\mathcal{X}}$ is weakly R risk averse, weakly R risk loving, or R risk neutral according to whether ϕ is concave, convex or linear.*

The theorem also holds for (strict) R risk aversion and strict concavity, as well as (strict) R risk loving and strict convexity.

The following Lemma is essentially Theorem 8 only stated with respect to a specific partition and specific pair therein.

Lemma A.10. *Let $\stackrel{\Delta}{\precsim}$ be an preference on $\Delta(\mathcal{S})$, and $\precsim$ the induced preference on $\mathcal{S}$, with u an NM utility representing $\stackrel{\Delta}{\precsim}$, and v a Debreu value function representing $\precsim$. For any independent partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$, and any pair of factors $\mathcal{T}_i, \mathcal{T}_j$, of the partition, the following holds: $\stackrel{\Delta}{\precsim}$ is R risk averse, (R weakly risk averse, R risk loving, R weakly risk loving, risk neutral) with respect to the partition and the pair of factors, if and only if u is strictly concave (respectively, concave, strictly convex, convex, linear) with respect to v.*

Proof. We prove the claim for weak R risk aversion and concavity. The other claims are analogous. Also, it suffices to prove for the pair of factors $\mathcal{T}_1, \mathcal{T}_2$.

Since there exists a value function v representing $\precsim$ (based on some independent partition), by Lemma A.8, there exist value functions $v^{\mathcal{T}_i}$ for the $\mathcal{T}_i$'s, with $v = \sum_{t=1}^{n} v^{\mathcal{T}_t}$.

Let $I_{1,2}$ be the image of $\mathcal{T}_1 \times \mathcal{T}_2$ under $v^{\mathcal{T}_1} + v^{\mathcal{T}_2}$; that is, $I_{1,2} = \{v^{\mathcal{T}_1}(a_1) + v^{\mathcal{T}_2}(a_2) : (a_1, a_2) \in \mathcal{T}_1 \times \mathcal{T}_2\}$.

Let $\hat{u}$ be such that $u(a_1, \ldots, a_n) = \hat{u}(v(a_1, \ldots, a_n)$ (such a $\hat{u}$ exists since $\stackrel{\Delta}{\precsim}$ and $\precsim$ agree on the certainties). So, for any fixed $\boldsymbol{c} = (c_3, \ldots, c_n)$, $u(a_1, a_2, \boldsymbol{c}) = \hat{u}(v^{\mathcal{T}_1}(a_1) + v^{\mathcal{T}_2}(a_2) + x_c)$, for $x_c = \sum_{j=3}^{n} v^{\mathcal{T}_j}(c_j)$. So, Theorem A.9 applies, with $\mathcal{X}_1 = v^{\mathcal{T}_1}(\mathcal{T}_1)$ and $\mathcal{X}_2 = v^{\mathcal{T}_2}(\mathcal{T}_2)$. So, for each fixed $\boldsymbol{c}$, by Theorem A.9, $\hat{u}$ is concave on $I_{1,2} + x_{\boldsymbol{c}}$ if and only if

$$\langle (a_1, a_2, \boldsymbol{c}), (b_1, b_2, \boldsymbol{c}) \rangle \stackrel{\Delta}{\precsim} \langle (a_1, b_2, \boldsymbol{c}), (b_1, a_2, \boldsymbol{c}) \rangle . \tag{22}$$

for any $a_1 \prec b_1, a_2 \prec b_2$.

In particular, if $\hat{u}$ is concave throughout its domain, then (22) holds for any $\boldsymbol{c}$.

Conversely, if (22) holds for any $\boldsymbol{c}$, then $\hat{u}$ is concave on overlapping intervals covering its entire range, so it is concave. $\square$

Theorem 8. *For NM utility u and Debreu value function v,*

- *Risk aversion:*
  - *u is strictly concave with respect to v if and only if $\overset{\Delta}{\succsim}$ is R risk averse.*
  - *u is concave with respect to v if and only if $\overset{\Delta}{\succsim}$ is weakly R risk averse.*
- *Risk loving:*
  - *u is strictly convex with respect to v if and only if $\overset{\Delta}{\succsim}$ is R risk loving.*
  - *u is convex with respect to v if and only if $\overset{\Delta}{\succsim}$ is weakly R risk loving.*
- *Risk neutrality: u is linear with respect to v if and only if $\overset{\Delta}{\succsim}$ is R risk-neutral.*

Proof. This is exactly Lemma A.10. $\square$

Proofs for Sections 5.

Theorem 5. *If $\overset{\Delta}{\succsim}$ is R risk averse with respect to some independent partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$, and some pair of factors $\mathcal{T}_i, \mathcal{T}_j$, then it is also R risk averse with respect to any independent partition, and any pair therein. Similarly for weak R risk aversion.*

Proof. If there is only one independent partition, with only two factors, then there is nothing to prove. Otherwise, either there are two different partitions, or there are more than two factors in the partition (or both). In either case, there exists a Debreu value function v representing $\precsim$.

Suppose that $\overset{\Delta}{\succsim}$ is R risk averse with respect to some partition and some pair. Then, by Lemma A.10, u is concave with respect to v. So, again, by Lemma A.10, $\overset{\Delta}{\succsim}$ is R risk averse with respect to any other partition and any pair therein. Similarly for weak R risk aversion. $\square$

Proof of Theorem 6.

As it turns out, the proof of Theorem 6 is the most involved in the section. The challenge arises in the case that the partition is with only two factors, in which case a Debreu value function need not exist, and the tools of Section 6 do not apply.

When considering a partition into two factors, we adopt the following notation, which is somewhat different from that used in the rest of the paper. The independent partition is denoted $\mathcal{S} = \mathcal{A} \times \mathcal{B}$. We use a, A, with or without subscripts or superscripts, for points in $\mathcal{A}$, and b, B for points in $\mathcal{B}$. By convention, $a \prec A$ and $b \prec B$.

Let $w^{\mathcal{A}} : \mathcal{A} \to \mathbb{R}$ be a continuous real function representing $\precsim^{\mathcal{A}}$, and similarly $w^{\mathcal{B}}$ a continuous real function representing $\precsim^{\mathcal{B}}$ (such function are exist by Debreu [6] since $\precsim^{\mathcal{A}}$ and $\precsim^{\mathcal{B}}$ are continuous). Define $\boldsymbol{w} : \mathcal{A} \times \mathcal{B} \to \mathbb{R}^2$ as $\boldsymbol{w}(a, b) = (w^{\mathcal{A}}(a), w^{\mathcal{B}}(b))$. Let $I_{\mathcal{A}} \times I_{\mathcal{B}} \subseteq \mathbb{R}^2$ be the image of $\mathcal{A} \times \mathcal{B}$ under $\boldsymbol{w}$.

Lemma A.11. *$u \circ \boldsymbol{w}^{-1} : I_{\mathcal{A}} \times I_{\mathcal{B}} \to \mathbb{R}$ is well defined, increasing in each coordinate, and continuous.*

Proof. If $\boldsymbol{w}(a,b) = \boldsymbol{w}(a',b')$ then $(a,b) \sim (a',b')$, and hence $u(a,b) = u(a',b')$. Thus, $u \circ \boldsymbol{w}^{-1}$ is well defined. It is increasing is each coordinate as u and $w^{\mathcal{A}}, w^{\mathcal{B}}$ agree on the certainty preference.

Denote $\hat{u} = u \circ \boldsymbol{w}^{-1}$, and for $x \in I_{\mathcal{A}}$ define $\hat{u}^{\mathcal{B}}_x : I_{\mathcal{B}} :\to \mathbb{R}$, by $\hat{u}^{\mathcal{B}}_x(y) = \hat{u}(x,y)$. Then, the $\hat{u}^{\mathcal{B}}_x$ is monotone. Also, $\hat{u}^{\mathcal{B}}_x(I_{\mathcal{B}}) = u((w^{\mathcal{A}}(x)^{-1}, \mathcal{B})$ is an interval (since $\mathcal{B}$ is a finite product of connected spaces and u continuous). So, $\hat{u}^{\mathcal{B}}_x$ is continuous for any x. Similarly, the function $\hat{u}^{\mathcal{A}}_y : I_{\mathcal{A}} :\to \mathbb{R}$, defined by $\hat{u}^{\mathcal{B}}_y(x) = \hat{u}(x,y)$ is continuous for any y.

To prove continuity of $\hat{u}$, we prove that the pre-images of the *open rays* $(-\infty, r)$ and (r, ∞) are open, for all r. Consider $(-\infty, r)$ (the other case is analogous). Set $E_r = \{(x,y) : \hat{u}(x,y) < r\}$. If $E_r = \emptyset$ or $E_r = I_{\mathcal{A}} \times I_{\mathcal{B}}$ then there is nothing to prove. Otherwise, consider (x^*, y^*) with $\hat{u}(x^*, y^*) < r - \epsilon$, for some $\epsilon > 0$. We show that there is a neighborhood of (x^*, y^*) fully contained in E_r. Suppose that x^* is not maximal in $I_{\mathcal{A}}$ and y^* not maximal in $I_{\mathcal{B}}$ (the proof for the case that one of them is maximal is similar). The function $\hat{u}^{\mathcal{B}}_{x^*}$ is continuous. So, there exists some y' with

$$0 < \hat{u}^{\mathcal{B}}_{x^*}(y') - \hat{u}^{\mathcal{B}}_{x^*}(y^*) < \frac{1}{2}\epsilon. \tag{23}$$

Similarly, the function $\hat{u}^{\mathcal{A}}_{y'}$ is continuous. Thus, there exists x' with

$$0 < \hat{u}^{\mathcal{A}}_{y'}(x') - \hat{u}^{\mathcal{A}}_{y'}(x^*) < \frac{1}{2}\epsilon. \tag{24}$$

Combining (23) and (24), we obtain

$$\hat{u}(x^*, y^*) < \hat{u}(x', y') + \epsilon < r.$$

Set $\delta = \min\{x' - x^*, y' - y^*\}$. Then, for any (x,y) if $\|(x,y) - (x^*, y^*)\| < \delta$ then $x < x'$ and $y < y'$. So, by monotonicity of $\hat{u}$, $\hat{u}(x,y) < \hat{u}(x',y') < r$. So, the entire ball of size δ around (x^*, y^*) is contained in E_r, as required. □

Lemma A.12. *Let $\mathcal{A} \times \mathcal{B}$ be an independent partition and $a \prec A$, $b \prec B$. Set $a^0 = a$, and while $(a^i, B) \precsim (A, b)$ let a^{i+1} be such that $(a^{i+1}, b) \sim (a^i, B)$ (such an a^{i+1} exists by continuity). Then, there exists an $\bar{i}$ such that $(a^{\bar{i}}, B) \succsim (A, b)$ (that is, the sequence $a^0, a^1, \ldots$ is finite).*

Proof. Contrariwise, suppose there is no such $\bar{i}$. Then, for $i = 1, 2, \ldots$, $(a^i, B) \prec (A, b)$, and hence $a^i \prec A$. Clearly, $a^i \precsim a^{i+1}$. Thus, the sequence $a^1, a^2, \ldots$, is an infinite monotone and bounded sequence, and hence converges to a limit $\hat{a}$. By definition, for each i

$$(a^i, B) \sim (a^{i+1}, b).$$

Thus, by continuity,

$$(\hat{a}, B) \sim (\hat{a}, b),$$

which is impossible since $b \prec^{\mathcal{B}} B$ and $\precsim$ is strictly monotone in each factor. □

Theorem 6. *$\overset{s}{\precsim}$ is R risk-averse if and only if it is perfect-R risk averse.*

Proof. R risk aversion ⇒ perfect-R risk aversion: The requirement of perfect-R risk aversion - (3) - is identical to that of R risk aversion - (2) - only limited to perfectly hedges.

Perfect-R risk aversion ⇒ R risk aversion: Suppose that $\overset{\vartriangle}{\precsim}$ is perfect-R risk averse. First, consider the case that the independent partition is with three or more factors. That is, suppose that $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$, $n \geq 3$. Then, there are Debreu value functions for the partition. Let $v^{\mathcal{T}_i}$ be the value function of $\mathcal{T}_i$. By Theorem 8 u is concave with respect to $v = \sum_{i=1}^n v^{\mathcal{T}_i}$. Consider $a_1 \prec b_1, a_2 \prec b_2$, and $\boldsymbol{c} \in \mathcal{S}_{-\{1,2\}}$. Set $x_i = v^{\mathcal{T}_i}(a_i), y_i = v^{\mathcal{T}_i}(b_i)$ and $z = \sum_{i=2}^n v^{\mathcal{T}_i}(c_i)$. W.l.o.g. $x_1 = x_2 = z = 0$. Set $\lambda = \frac{y_1}{y_1+y_2}$. Then, since u is concave with respect to v

$$\lambda \cdot u(a_1, a_2, \boldsymbol{c}) + (1-\lambda)u(b_1, b_2, \boldsymbol{c}) = \lambda \cdot (u \circ v^{-1})(0) + (1-\lambda)(u \circ v^{-1})(y_1 + y_2) < \tag{25}$$

$$(u \circ v^{-1})(\lambda \cdot 0 + (1-\lambda)(y_1+y_2)) = (u \circ v^{-1})(y_2) = u(a_1, b_2, \boldsymbol{c}). \tag{26}$$

Similarly,

$$(1-\lambda)u(a_1, a_2, \boldsymbol{c}) + \lambda \cdot u(b_1, b_2, \boldsymbol{c}) < u(b_1, a_2, \boldsymbol{c}). \tag{27}$$

Combining (25) and (27)

$$u(a_1, a_2, \boldsymbol{c}) + u(b_1, b_2, \boldsymbol{c}) < u(a_1, b_2, \boldsymbol{c}) + u(b_1, a_2, \boldsymbol{c}),$$

and $\overset{\vartriangle}{\precsim}$ is R risk averse.

Next, suppose that the partition has only two factors: $\mathcal{S} = \mathcal{A} \times \mathcal{B}$. Let $a, A \in \mathcal{A}, b, B \in \mathcal{B}$, with $a \prec A$ and $b \prec B$. We need to show that

$$\langle (a,b), (A,B) \rangle \overset{\vartriangle}{\prec} \langle (A,b), (a,B) \rangle . \tag{28}$$

If $(a, B) \sim (A, b)$ then they are perfectly hedged and (28) holds by the definition of perfect-R risk aversion.

Otherwise, let u be an NM utility for $\overset{\vartriangle}{\precsim}$. set

$$\mathrm{diff} = u(a,b) + u(A,B) - u(a,B) - u(A,b).$$

We show that diff < 0, which establishes (28).

Let $w^{\mathcal{A}}$ be a continuous function representing $\precsim^{\mathcal{A}}$ and $w^{\mathcal{B}}$ a continuous function representing $\precsim^{\mathcal{B}}$ (the certainty preferences). In order to prove that diff < 0, we start out by proving that there exists $a_{\frac{1}{2}}, A_{\frac{1}{2}}, b_{\frac{1}{2}}, B_{\frac{1}{2}}$, with

$$a \precsim a_{\frac{1}{2}} \prec A_{\frac{1}{2}} \precsim A, \quad \text{and} \quad b \precsim b_{\frac{1}{2}} \prec B_{\frac{1}{2}} \precsim B,$$

such that

$$\begin{aligned} &w^{\mathcal{A}}(A_{\frac{1}{2}}) - w^{\mathcal{A}}(a_{\frac{1}{2}}) \leq \frac{1}{2}(w^{\mathcal{A}}(A) - w^{\mathcal{A}}(a)) \quad \text{or} \\ &w^{\mathcal{B}}(B_{\frac{1}{2}}) - w^{\mathcal{B}}(b_{\frac{1}{2}}) \leq \frac{1}{2}(w^{\mathcal{B}}(B) - w^{\mathcal{B}}(b)) \end{aligned} \tag{29}$$

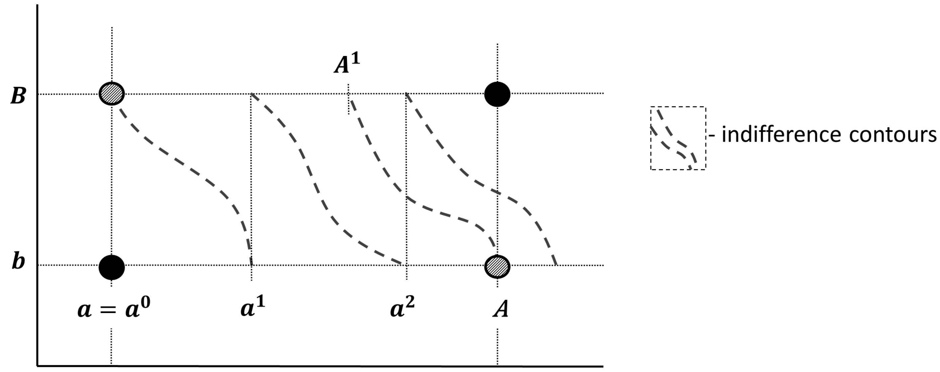

FIGURE 3. Illustration of the proof of Theorem 6. The values a^i are calculated left-to-right, starting at $a = a^0$. Here $\bar{i} = 2$ and the point a^2 is such that $w^{\mathcal{A}}(a^2) \geq \frac{1}{2}(w^{\mathcal{A}}(A) + w^{\mathcal{A}}(a))$ (assuming the picture is scaled according to $w^{\mathcal{A}}$).

and

$$\text{diff} < u(a_{\frac{1}{2}}, b_{\frac{1}{2}}) + u(A_{\frac{1}{2}}, B_{\frac{1}{2}}) - u(a_{\frac{1}{2}}, B_{\frac{1}{2}}) - u(A_{\frac{1}{2}}, b_{\frac{1}{2}}). \tag{30}$$

W.l.o.g. we may assume that $(a, B) \prec (A, b)$; so $(a, b) \prec (a, B) \prec (A, b)$. Thus, since $\precsim^{\mathcal{A}}$ is continuous and $\mathcal{A}$ connected, there exists $a \prec a^1 \prec A$ with

$$(a^1, b) \sim (a, B). \tag{31}$$

Figure 3 illustrates the following argument. Set $a^0 = a$. Given a^i, let a^{i+1} be such that $(a^{i+1}, b) \sim (a^i, B)$. Let $\bar{i}$ be the first index with $(a^{\bar{i}}, B) \succsim (A, b)$; such an $\bar{i}$ exists by Lemma A.12. Then, $(a, B) \prec (A, b) \precsim (a^{\bar{i}}, B)$. Thus, there exists $A^1, a \prec A^1 \precsim a^{\bar{i}}$, such that $(A^1, B) \sim (A, b)$. Clearly, $a^{\bar{i}} \precsim A$. Thus, either

$$w^{\mathcal{A}}(A^1) \leq \frac{1}{2}(w^{\mathcal{A}}(a) + w^{\mathcal{A}}(A)), \tag{32}$$

or

$$w^{\mathcal{A}}(a^{\bar{i}}) \geq \frac{1}{2}(w^{\mathcal{A}}(a) + w^{\mathcal{A}}(A)). \tag{33}$$

We consider each of these cases separately.

First, suppose that (32) holds. Then, by construction $(A^1, B) \sim (A, b)$, and they are perfectly hedged. Hence, by assumption,

$$\left\langle (A^1, b), (A, B) \right\rangle \overset{\vartriangle}{\prec} \left\langle (A^1, B), (A, b) \right\rangle.$$

So,

$$u(A^1, b) + u(A, B) - u(A^1, B) - u(A, b) < 0.$$

Hence,

$$
\begin{aligned}
& u(a,b)+u(A,B)-u(A,b)-u(a,B)= \\
& u(a,b)+u(A^1,B)-u(A^1,b)-u(a,B)+u(A^1,b)+u(A,B)-u(A^1,B)-u(A,b)< \\
& u(a,b)+u(A^1,B)-u(A^1,b)-u(a,B).
\end{aligned}
\tag{34}
$$

Setting $a_{\frac{1}{2}}=a$, $A_{\frac{1}{2}}=A^1$, $b_{\frac{1}{2}}=b$ and $B_{\frac{1}{2}}=B$, by (32) and (34) we get (29) and (30).

Next, suppose that (33) holds. Then, by construction, for $i=1,\ldots,\bar{i}$, $(a^{i-1},B)\sim(a^i,b)$, and each such pair is perfectly hedged. Since $\stackrel{\Delta}{\precsim}$ is ordinally risk averse,

$$\left\langle(a^{i-1},b),(a^i,B)\right\rangle \stackrel{\Delta}{\prec} \left\langle(a^{i-1},B),(a^i,b)\right\rangle,$$

for all i. So,

$$\frac{1}{2\bar{i}}\sum_{i=1}^{\bar{i}}\left(u(a^{i-1},b)+u(a^i,B)\right)<\frac{1}{2\bar{i}}\sum_{i=1}^{\bar{i}}\left(u(a^{i-1},B)+u(a^i,b)\right); \tag{35}$$

and

$$u(a^0,b)+u(a^{\bar{i}},B)<u(a^{\bar{i}},b)+u(a^0,B);$$

so (as $a^0=a$)

$$u(a,b)+u(a^{\bar{i}},B)-u(a^{\bar{i}},b)-u(a,B)<0\,.$$

Hence,

$$
\begin{aligned}
& u(a,b)+u(A,B)-u(A,b)-u(a,B)= \\
& u(a,b)+u(a^{\bar{i}},B)-u(a^{\bar{i}},b)-u(a,B)+u(a^{\bar{i}},b)+u(A,B)-u(a^{\bar{i}},B)-u(A,b)< \\
& u(a^{\bar{i}},b)+u(A,B)-u(a^{\bar{i}},B)-u(A,b).
\end{aligned}
\tag{36}
$$

Setting $a_{\frac{1}{2}}=a^{\bar{i}}$, $A_{\frac{1}{2}}=A$, $b_{\frac{1}{2}}=b$ and $B_{\frac{1}{2}}=B$, by (33) and (36) we get (29) and (30).

Thus, we have established (29) and (30), and we now return to complete the proof that diff <0. Set

$$\mathrm{diff}_{\frac{1}{2}}=u(a_{\frac{1}{2}},b_{\frac{1}{2}})+u(A_{\frac{1}{2}},B_{\frac{1}{2}})-u(a_{\frac{1}{2}},B_{\frac{1}{2}})-u(A_{\frac{1}{2}},b_{\frac{1}{2}}).$$

Then,

$$\mathrm{diff}<\mathrm{diff}_{\frac{1}{2}}.$$

Applying the above halving procedure repeatedly, we obtain that for any $\delta>0$ there exists $(a_\delta,b_\delta),(A_\delta,B_\delta)$, such that

$$w^{\mathcal{A}}(A_\delta)-w^{\mathcal{A}}(a_\delta)\le\delta \quad \text{or} \tag{37}$$

$$w^{\mathcal{B}}(B_\delta)-w^{\mathcal{B}}(b_\delta)\le\delta \tag{38}$$

and

$$\text{diff}_{\frac{1}{2}} < u(a_\delta, b_\delta) + u(A_\delta, B_\delta) - u(a_\delta, B_\delta) - u(A_\delta, b_\delta) =$$

$$(u(A_\delta, B_\delta) - u(a_\delta, B_\delta)) + (u(a_\delta, b_\delta) - u(A_\delta, b_\delta)) = \tag{39}$$

$$(u(A_\delta, B_\delta) - u(A_\delta, b_\delta)) + (u(a_\delta, b_\delta) - u(a_\delta, B_\delta))\,. \tag{40}$$

By Lemma A.11 the function $u \circ (w^{\mathcal{A}}, w^{\mathcal{B}})^{-1}$ is continuous. So it is uniformly continuous on the rectangle $[w^{\mathcal{A}}(a), w^{\mathcal{A}}(A)] \times [w^{\mathcal{B}}(b), w^{\mathcal{B}}(B)]$. That is, for any $\epsilon > 0$, there exists a δ such that if

$$\|(w^{\mathcal{A}}(a'), w^{\mathcal{B}}(b')) - (w^{\mathcal{A}}(a''), w^{\mathcal{B}}(b''))\| < \delta$$

then

$$|u(a', b') - u(a'', b'')| < \epsilon.$$

In particular, if (37) holds then (39) is $\leq 2\epsilon$, and if (38) holds then (40) is $\leq 2\epsilon$. Thus, $\text{diff}_{\frac{1}{2}} \leq 0$, so diff < 0. □

Proofs for Section 8.

Theorem 9. *In the multi-commodity setting (with $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ an independent partition), the NM utility function u has constant coefficient of absolute risk aversion when measured with respect to the Debreu value function v if and only if for any i, lotteries L, L' over $\mathcal{T}_i$, and $\boldsymbol{x}, \boldsymbol{y} \in \Omega_{-\{i\}}$*

$$(L, \boldsymbol{x}) \overset{\vartriangle}{\precsim} (L', \boldsymbol{x}) \iff (L, \boldsymbol{y}) \overset{\vartriangle}{\precsim} (L', \boldsymbol{y}).$$

Proof. Let u be an NM utility representing $\overset{\vartriangle}{\precsim}$. Meyer [22] (quoted in [25]) showed that all $\mathcal{T}_i$'s are utility independent if and only if there exist functions $u^i : \mathcal{T}_i \to \mathbb{R}$, $\beta > 0$ and α, such that one of the following holds:

$$u(a_1, \ldots, a_n) = \sum_{i=1}^{n} u^i(a_i) \tag{41}$$

$$u(a_1, \ldots, a_n) = \alpha + \beta \prod_{i=1}^{n} u^i(a_i)\ , \text{ with } u^i(a_i) > 0 \tag{42}$$

$$u(a_1, \ldots, a_n) = \alpha - \beta \prod_{i=1}^{n} (-u^i(a_i))\ , \text{ with } u^i(a_i) < 0. \tag{43}$$

If (41) holds than the u^i's are Debreu value functions (since $\overset{\vartriangle}{\precsim}$ agrees with $\precsim$). So u is linear with respect to v, and, in particular CARA.

If (42) holds than setting $v^i = \ln(u^i)$ we have that

$$v(a_1, \ldots, a_n) = \sum_{i=1}^{n} v^i(a_i) = \ln(\prod_{i=1}^{n} u^i(a_i)),$$

is a Debreu value function representing $\precsim$. So,

$$u(a_1, \ldots, a_n) = \alpha + \beta e^{v(a_1,\ldots,a_n)},$$

is CARA w.r.t. v.

If (43) holds than setting $v^i = -\ln(-u^i)$ we have that

$$v(a_1, \ldots, a_n) = -\sum_{i=1}^{n} v^i(a_i) = -\ln(\prod_{i=1}^{n} -u^i(a_i))$$

is a Debreu value function, and

$$u(a_1, \ldots, a_n) = \alpha - \beta e^{-v(a_1,\ldots,a_n)},$$

is CARA w.r.t. v. $\square$

Proposition 8.1. *Let $\precsim$ be an (additively separable) preference order on $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$, and g a real valued function on $\mathcal{S}$. Suppose that the following holds for any NM utility function u:*

- *u has constant coefficient of absolute risk aversion when measured with respect to g if and only if*

$$(L, \boldsymbol{x}) \overset{\vartriangle}{\precsim} (L', \boldsymbol{x}) \iff (L, \boldsymbol{y}) \overset{\vartriangle}{\precsim} (L', \boldsymbol{y}).$$

  for any $L, L' \in \Delta(\mathcal{T}_i)$, and $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}_{-\{i\}}$.

Then g is a Debreu value function.

Proof. In the literature, the condition that

$$(L, \boldsymbol{x}) \overset{\vartriangle}{\precsim} (L', \boldsymbol{x}) \iff (L, \boldsymbol{y}) \overset{\vartriangle}{\precsim} (L', \boldsymbol{y}).$$

for any $L, L' \in \Delta(\mathcal{T}_i)$, and $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}_{-\{i\}}$ is termed *utility independence* of $\mathcal{T}_i$ [19].

By assumption there exists a Debreu value function v for $\mathcal{S}$ with $v = \sum_{i=1}^{n} v^{\mathcal{T}_i}$. So, for the NM utility function $u = v$, it holds that each $\mathcal{T}_i$ is utility independent. So, by assumption this u is CARA in g. So, v is CARA in g. If it is linear there is nothing to prove. Otherwise,

$$v = \alpha + \beta e^{\gamma g}, \tag{44}$$

for some α, β, γ.

Now consider another NM utility $u = e^v$. By Theorem 9 under this utility function each $\mathcal{T}_i$ is utility independent. Hence, by assumption, this u must also be CARA in g. But, by (44), $u = e^v = e^{\alpha + \beta e^{\gamma g}}$, which is not CARA in g. $\square$

Appendix B. Unbounded Lottery Sequences

Here we show why in Definition 1 one needs to require that the lottery sequence be bounded. Suppose that the conditions of Section 4 hold. We show that if we allow for unbounded lottery sequences, then for *any* preference policy $\stackrel{\vartriangle}{\precsim} = (\stackrel{\vartriangle}{\precsim}^1, \stackrel{\vartriangle}{\precsim}^2, \ldots)$, there exists a lottery sequence that is ultimately inferior to its repeated certainty equivalent.

Let $v^{\mathcal{T}_i}$ be the value function of $\mathcal{T}_i$. W.l.o.g. suppose that $\mathcal{T}_i$ is already represented in terms of $v^{\mathcal{T}_i}$, that is $v^{\mathcal{T}_i}(a_i) = a_i$ for all $a_i \in \mathcal{T}_i$. Then, the certainty preferences $\precsim^n$ are simply determined by the sum of the coordinates.

Let u_n be a NM utility representing $\stackrel{\vartriangle}{\precsim}^n$. For each n, let b_n be such that

$$2^{-n} \cdot u_n(0, \ldots, 0, b_n) + \left(1 - 2^{-n}\right) u_n(0, \ldots, 0, -1) = u_n(0, \ldots, 0).$$

Let L_n be the lottery obtaining the value b_n with probability 2^{-n} and the value -1 with probability $1 - 2^{-n}$. Then, $c_1, c_2, \ldots$, the repeated certainty equivalent of the lottery sequence $L_1, L_2, \ldots$, has $c_n = 0$ for all n. However,

$$\sum_{n=1}^{\infty} \Pr[\ell_n > -1] = \sum_{n=1}^{\infty} 2^{-n} < \infty.$$

So, by the Borel Cantelli lemma

$$\Pr[\ell_n > -1 \text{ infinitely often}] = 0.$$

So,

$$\Pr[\sum_{i=1}^{n} \ell_i < 0 \text{ from some } n \text{ on}] = 1,$$

and hence

$$\Pr[\sum_{i=1}^{n} \ell_i < 0 = \sum_{i=1}^{n} c_i \text{ from some } n \text{ on}] = 1.$$

So, $L_1, L_2, \ldots$ is ultimately inferior to $c_1, c_2, \ldots$.

Appendix C. Additive Utility Functions

In Section 9.3 we noted that if the NM utility must be additive, then the certainty preferences uniquely determine the lottery preferences. Here, we make this statement precise, and prove it.

For a utility function u over $\Delta(\mathcal{S})$, we say that u is *additively separable* if there exists a partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$, $n \geq 2$, and non-constant functions $u^{\mathcal{T}_i}$, $i = 1, \ldots, n$, such that for $\omega = (a_1, \ldots, a_n)$, $u(\omega) = \sum_{i=1}^{n} u^{\mathcal{T}_i}(a_i)$. Note that the definition of additively separable does not specify the partition. So, two additively separable function may be so with respect to different partitions.

Theorem 10. *Let $\precsim$ be a preference order on $\mathcal{S}$. There exists at most one preference order $\stackrel{\vartriangle}{\precsim}$ on $\Delta(\mathcal{S})$ that agrees with $\precsim$, and for which the corresponding utility function u is additively separable.*

Proof. Let u be an additively separable utility function. Suppose that u agrees with $\precsim$ on the certainties; that is, u represents $\precsim$ (if no such u exists then there is nothing to prove). By definition $u(\omega) = \sum_{i=1}^{n} u^{\mathcal{T}_i}(a_i)$, for some partition. So, u is (also) a Debreu value function. So, by Theorem 7, u is unique up to positive affine transformations. $\square$