AN AXIOMATIC FOUNDATION FOR THE THEORY OF RISK AVERSION WITH APPLICATIONS TO MULTI-COMMODITY RISK AVERSION (WORKING PAPER†)

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. To this end, we introduce two axiomatic definitions of the notion risk aversion, based on the decisions maker's preference order alone, independent of any numerical scale. We then show that when cast in quantitative form these axiomatic 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). The implications of the theory are discussed, including in particular, to defining risk aversion for non-monetary goods, to disentangling risk aversion from diminishing marginal utility, to multi-commodity/multi-period risk aversion, to two-period consumption-saving choices, and to temporal discounting. The entire study is within the expected utility framework.

Keywords: Risk aversion, Decision theory, Utility theory, Multi-commodity risk aversion, Multiple objectives decision making, Consumption-saving choice, Temporal Discounting.

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.

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

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

Date: June 1, 2017.

† For the most updated version see https://sites.google.com/site/aumannbiu/home/publication-by-topic-1.

works of Arrow [2] and Pratt [28], 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.

The matter of scale brings about the second issue, which is a more conceptual one. The classic definition of risk aversion is inherently scale-dependent; the notion of *expectation*, or concavity is only defined with respect to a given scale. In particular, the definition of risk aversion is not preserved under monotone, non-linear transformation of the scale; that is, a decision maker deemed risk averse by one scale, may, for the same set of preferences, be deemed risk loving by merely changing the scale with which the outcomes are described.[1][2] Thus, unlike the predominant trend in neoclassical economics, risk aversion is not defined based on fundamental notions such as preferences and indifference curves alone, but is inherently dependent on the scale one uses to describe the outcomes. Providing a scale-free, preference based definition of risk aversion is the second driving motivation for the current work.

1.2. **An Axiomatic Foundation.** In order to address the above questions, we start by seeking an axiomatic definition of risk aversion, independent of any units, and making no use of arithmetic notions such as mean or expectation. We provide two, related, such definitions, as outlined shortly. Both definitions are based solely on the internal structure of the decision maker's preferences. Having defined risk aversion in purely axiomatic 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* type of goods - monetary or non-monetary. In particular, we obtain a general framework for defining risk aversion for any type of goods, and to determine the scale.

Lottery Sequences. 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 realization 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

[1]As an example, consider preferences over noise intensities. Suppose that a decision maker is indifferent between being exposed to jack hammer noise (100 dB), and a fair gamble between falling leafs (10 dB) and jet-engine (150 dB) noises. Using the decibel measure for these noises, this decision would be deemed risk averse, as the certainty equivalent is larger than the expectation (noise is a "bad", so the negative amounts should be considered when assessing risk aversion). Decibel measures, however, are computed using the base-10 logarithm of the *sound pressure*. The sound pressure (measured is Pa - Pascal - units) of the three possible outcomes is 2 Pa (jack hammer), 6×10^{-5} Pa (falling leafs), and 632 Pa (jet engine). So, the same decision maker would be deemed risk loving under the sound pressure scale.

[2]Indeed, for any preference order, there exists a monotone transformation of the scale that deems the preference risk averse, and another monotone transformation of the scale the deems the same preference order risk loving.

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 first axiomatic 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.

Finite Sequences. The above definition is set in the context of an infinite sequence of time periods. The second definition applies to any number (two or more) of time periods (or alternatively, any other partition of the space into two or more independent factors[3]). When the number of lotteries is finite, there can be no (non-trivial) behavior "with probability 1". So, we cannot require that the realization of the lottery sequence be preferred to the certainty equivalent with probability 1. Rather, we define risk aversion as a preference wherein, for fair lottery sequences, the probability that the realization is (weakly) preferred to the certainty equivalent is greater than the probability of the reverse (that is, the probability that the certainty equivalent is (weakly) preferred to the realization). The exact definition is provided in Section 5.

A Quantitative Form. Having established axiomatic foundation for defining risk-aversion, we show that these definitions 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], and commonly used in the theory of multi-attribute decision theory (see [22]). 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 monetary 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 axiomatic 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 our definitions.

[3] formally defined in Section 2

1.3. **Implications.** The approach offered in this paper has several implications for the understanding of risk aversion, both conceptual and technical, including:

Non-monetary Goods. First, 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 to disentangle 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. This direction is explored in more detail in Section 10.

Multi-commodity Risk Aversion. Ever since the initial works of Arrow [2] and Pratt [28], researchers have considered how to extend the definition and associated measures to the multi-commodity setting, and various approaches have been suggested (see [23, 32, 27, 10, 20, 29, 21, 26] for some references in the expected utility model). The core problem of the multi-commodity setting is the lack of a unified scale; each commodity may have a different (natural) scale. This problem is cricumvented in our approach here, which is inherently multi-commodity definition, by using the *joint* Debreu value scale for all commodities. We show (Section 6) that this allows to naturally extend the Arrow-Pratt framework to the multi-commodity setting, including, in particular, for defining the coefficient of absolute risk aversion, and retaining the characteristics of notions such as DARA (decreasing absolute risk aversion) and CARA (constant absolute risk aversion).

Consumption-Saving Choice. In the *portfolio selection* setting a decision maker needs to split her budget between a risky asset and a non-risky one. Arrow and Pratt have shown how the notions of risk aversion and DARA (decreasing absolute risk aversion) determine the decision makers' behavior in such settings. Their analysis, however, only holds for purely monetary investments, wherein, under certainty, a dollar of one asset is a perfect substitute for a dollar of the other asset. Extending such an analysis to the two-period case - wherein the decision maker needs to decide how to split the funds between (sure) consumption in the first period, and savings with a random return, for consumption in the second period - has proven challenging. Kihlstrom and Mirman [23, 24] provided some extensions, but core problem remained open. In particular, determining conditions under which savings increase with wealth have remained unclear. We show how our approach allows to tackle this and related problems, extending the Arrow-Pratt analysis to the two period case.

Time Discounting. Time discounting is widely considered in the economic literature. One, but not necessarily the only, source of such discounting is uncertainty associated with the future; the decision maker may not live to enjoy later time periods. We show that our framework allows to isolate the contribution of this source to the overall discounting, obtaining a measure of the *survival uncertainty discounting rate*. In addition, we show that the level of such discounting is directly related to the decision maker's risk attitude - risk aversion, risk loving, DARA and CARA, as defined in this paper.

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.[4] 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 classic 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 one; 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 definition based on infinite lottery sequences is presented in Section 3, and its quantitative form in Section 4. Section 5 presents the second definition, together with its equivalent quantitative form. Implications to multi-commodity risk aversion are discussed in Section 6, to the consumption-saving choice in Section 7, and to inter-temporal discounting in Section 8. Some related work is discussed in Section 9. We conclude the main body of the paper with a discussion in Section 10. All proofs are deferred to an appendix.

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

2. Model, Terminology and Notation

The Spaces. Preferences are defined over consumption spaces of the form $\mathcal{S} = \mathscr{C}_1 \times \cdots \times \mathscr{C}_m$, where each $\mathscr{C}_i$ is a separable and arc-connected topological space, representing the consumption space of a specific commodity (at a specific time period). We call each $\mathscr{C}_i$ *a commodity space.*

Lotteries. All lotteries considered are finite support lotteries. The set of all such lotteries over a space $\mathcal{S}$ is denote by $\Delta(\mathcal{S})$. The notation $\langle s_1, \ldots, s_n : p_1, \ldots, p_n \rangle$, denotes the lottery wherein outcome s_k occurs with probability p_k.

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

- *the certainty preferences*: a preference order $\precsim$ on $\mathcal{S}$,[5]
- *the lottery preferences*: a continuous preference order $\overset{\vartriangle}{\precsim}$ on $\Delta(\mathcal{S})$.

It is assumed throughout that $\precsim$ and $\overset{\vartriangle}{\precsim}$ agree on the preferences over the certainties - $\mathcal{S}$ (which for $\overset{\vartriangle}{\precsim}$ are the degenerate lotteries).

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* [18]; 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. Most commonly in this paper, factors represent time periods; that is, the factor $\mathcal{T}_i$ is the product of all commodity spaces associated with consumption at time i. A *partition* of $\mathcal{S}$ is a representation of $\mathcal{S}$ as a product of factors $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$.[6] 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.*

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

[6] For our purposes, we only care about the commodities of the space, not their order. So, it may be that the order of the commodities in the partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ does not correspond to the original order.

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

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

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

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

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* $\overset{\Delta}{\precsim}$.

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.[7]

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 [18], for $n \ge 3$, 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. Lottery Sequences

Our definition of risk aversion is set in the context of lottery sequences. 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: *lottery sequences*, *certainty equivalent of a lottery sequence*, and *ultimately inferior outcome*.

[7] 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.

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.[8] 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.

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 $(c_1, \ldots, c_{n-1}, c_n) \overset{\vartriangle}{\sim}^n (c_1, \ldots, c_{n-1}, L_n)$ 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

$$\Pr[(c_1, \ldots, c_n) \prec^n (\ell_1, \ldots, \ell_n) \text{ from some } n \text{ on}] = 1.$$ [9]

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

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

[8] 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.

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

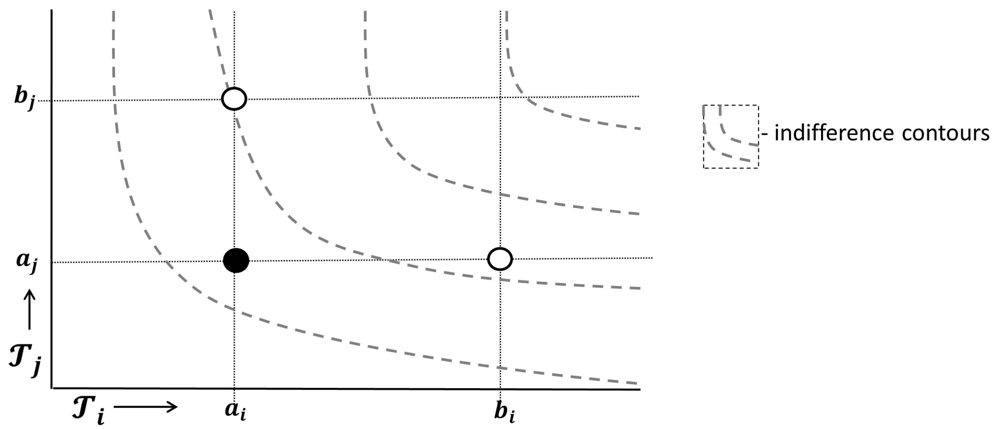

FIGURE 1. Illustration of $[a_j, b_j] \sqsubseteq [a_i, b_i]$ (the factors of $\mathcal{S}_{-\{i,j\}}$ are not depicted).

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.[10] Hence, we now define the notions of a *bounded* lottery sequence and a *non-vanishing* lottery sequence.

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

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 $i > 1$. 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 $L_i \in \Delta([a_i, b_i])$ 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 provide a scale-free definition of risk aversion:

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

- Scale-free (SF) 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.*

[10]See Appendix B.

- Weakly scale-free (SF) 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:*

- Scale-free (SF) 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.*
- Scale-free (SF) weakly risk loving *if the repeated certainty equivalent of any bounded lottery sequence is not ultimately inferior to the lottery sequence itself.*
- Scale-free (SF) 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. Lottery Sequences: The Quantitative Perspective

The previous section provided a scale-free, axiomatic definition of risk aversion. We now show how this axiomatic definition can also be cast in quantitative form. Specifically, we show that (under some assumptions) this scale-free definition of risk-aversion coincides with the Arrow-Pratt scale-dependent definition for a specific choice of scale, which arises intrinsically from the decision maker's preference structure. Specifically, the appropriate 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[11] then $\precsim^n$ is *additively separable*;[12] that is, there exist functions $v^{\mathcal{T}_i} : \mathcal{T}_i \to \mathbb{R}$, such that for any $(a_1, \ldots, a_n), (a'_1, \ldots, a'_n)$

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

[11]see page 7.

[12]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.

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 $\mathcal{H}^n$.[13] We note that Debreu [7] called these functions *utility* functions; but following Keeney and Raiffa [22], 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. For the certainty preferences, assume:

A1 - Independence: $\mathcal{H}^n = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ is an independent partition[14] (with respect to $\precsim^n$), for all n.

A2 - Consistency: for any $n, n', n' > n$, the preference order induced by $\precsim^{n'}$ on $\mathcal{H}^n$ is identical to $\precsim^n$

Note that this is only assumed with respect to the certainty preferences, not the lottery preferences. These assumptions yield the existence of value functions, as follows:

Proposition 4.1. *Assuming A1-A2, 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, the following *weak consistency.*

A3 - Weak consistency: for any n, there *exists* a $\phi_{n+1} \in \mathcal{T}_{n+1}$ 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 $(\phi_2, \phi_3, \ldots)$ the *presumed future*, and assume that it is bounded away from the boundaries of the $\mathcal{T}_i$'s; that is, 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.

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

[14]That is, each consecutive pair of factors $\mathcal{T}_i \times \mathcal{T}_{i+1}$ is independent.

4.2.1. *Weak SF Risk Aversion and (Weak) Concavity.* For each n, let u_n be the NM utility function representing $\overset{\vartriangle}{\succsim}^n$. Since $\overset{\vartriangle}{\succsim}^n$ and $\precsim^n$ agree, u_n is a monotone transformation of v_n. Set $\hat{u}_n = u_n \circ (v_n)^{-1}$; so, $\hat{u}_n(v_n(\boldsymbol{a})) = u_n(\boldsymbol{a})$, for any $\boldsymbol{a} \in \mathcal{H}^n$. Conceptually, $\hat{u}_n$ is the function u_n once the underlying scale is converted to the value function v_n. Accordingly, we call $\hat{u}_n$ the *value-scaled-utility* of $\overset{\vartriangle}{\succsim}^n$.

The next theorem establishes the connection between weak-SF-risk-aversion and concavity of the value-scaled-utilities of $\hat{u}_n$.

Theorem 1. *Assuming A1-A3, $\overset{\vartriangle}{\succsim}$ is weakly-SF-risk-averse if and only if all the valued-scaled-utilities $\hat{u}_n$ are concave.*

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) SF-Risk-Aversion and Strict Concavity.* We would have now wanted to claim that (strict) SF-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 value-scaled NM utility functions is bounded away from zero.

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

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

Theorem 2. *Assuming A1-A3, if $A_{\hat{u}_n}(x)$ is bounded away from 0, uniformly for all n and x,*[15] *then $\overset{\vartriangle}{\succsim}$ is SF-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

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

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

Set $\hat{\boldsymbol{u}} = (\hat{u}_1, \hat{u}_2, \ldots)$, the sequence of value-scaled-utilities. 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. *Assuming A1-A3,*

(a) If $RP_{\hat{\mathbf{u}}}(\epsilon) = \Omega(\epsilon^2)$ as $\epsilon \to 0$ then $\overset{\triangle}{\succsim}$ is SF risk averse.[16]
(b) If $RP_{\hat{\mathbf{u}}}(\epsilon) = O(\epsilon^{2+\beta})$ as $\epsilon \to 0$, for some $\beta > 0$, then $\overset{\triangle}{\succsim}$ is not SF risk averse.

The sufficient condition of (a) and the necessary one of (b) are not identical, but 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. *Assuming A1-A3, 1 and 3*

(a) Weak risk loving: $\overset{\triangle}{\succsim}$ is weakly SF risk loving if and only if all the all the valued-scaled-utilities $\hat{u}_n$ are convex.
(b) Risk loving
- *If $(-RP_{\hat{\mathbf{u}}}(\epsilon)) = \Omega(\epsilon^2)$ as $\epsilon \to 0$ then $\overset{\triangle}{\succsim}$ is SF risk loving.*
- *If $(-RP_{\hat{\mathbf{u}}}(\epsilon)) = O(\epsilon^{2+\beta})$ as $\epsilon \to 0$ (for some $\beta > 0$) then $\overset{\triangle}{\succsim}$ is not SF risk loving.*

(c) Risk Neutrality: $\overset{\triangle}{\succsim}$ is SF risk neutral if and only if $\hat{u}_n$ is linear for all n.

5. Finite Lottery Sequences

The definitions of Section 3 require an infinite sequence of time periods. In this section, we provide a definition for the setting with a finite number of periods.

The Space. We consider a space $\mathcal{S}$ and an independent partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$. In this case the factors $\mathcal{T}_i$ may either be different time periods, as in the previous sections, but also may represent any other independent partition of the consumption space.

5.1. **Risk Aversion.** When the number of factors is finite we cannot possibly expect a behavior "with probability one", as in the definitions of Section 3. Rather, for the finite case, risk aversion is defined as a preference wherein the probability that the realization of a lottery is preferred to its certainty equivalent is greater than the probability of the opposite, as follows.

[16] 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 $y < y_0$.

Weak Risk Aversion. Denote the certainty equivalent of a lottery $\boldsymbol{L}$ by $\boldsymbol{c}(\boldsymbol{L})$.[17] Consider a lottery sequence $\boldsymbol{L} = (L_1, \ldots, L_n)$, where $L_i \in \Delta(\mathcal{T}_i)$. The sequence is a *fair lottery sequence* if each L_i is a fair (50-50) over its respective factor.

Definition 3. $\overset{\vartriangle}{\precsim}$ *is* scale-free (SF) weakly-risk-averse *if for any fair lottery sequence* $\boldsymbol{L}$,

$$\Pr[\boldsymbol{c}(\boldsymbol{L}) \precsim \boldsymbol{\ell}] \geq \Pr[\boldsymbol{c}(\boldsymbol{L}) \succsim \boldsymbol{\ell}]. \tag{2}$$

Strong Weak Aversion. We would have wanted to define strong weak aversion as a preferences wherein (2) holds with a strong inequality. However, this cannot be a good definition, as in the case where all but one of the lotteries in the sequence are degenerate there are only two possible outcomes to the lottery, and the certainty equivalent would necessarily lay in between these two possible outcomes. So, the probability of both sides of (2) would equal 1/2. Similarly, if one of the lotteries in the sequence is "large" and all the other lotteries relatively "small", one would again get that the probabilities for both sites of (2) would be 1/2. So, what we need is two lotteries that are "of the same magnitude". Hence, the following definition:

Definition 4. *Fair lotteries* $L_i = \langle a_i, b_i : 0.5, 0.5\rangle$ *and* $L_j = \langle a_j, b_j : 0.5, 0.5\rangle$ *(in* $\Delta(\mathcal{T}_i)$ *and* $\Delta(\mathcal{T}_j)$ *respectively) are* of the same magnitude *if* $(a_i, b_j) \sim (b_i, a_j)$,[18]*. Lottery sequence* $\boldsymbol{L} = (L_i, L_j, \boldsymbol{d})$ *is a* repeated lottery sequence *if* L_i, L_j *are of the same magnitude and* $\boldsymbol{d} \in \mathcal{S}_{-\{i,j\}}$.

With this definition, strong risk aversion is defined as preferences wherein strict inequality holds for repeated lottery sequences.

Definition 5. $\overset{\vartriangle}{\precsim}$ *is* scale-free (SF) risk-averse *if for any non-degenerate repeated lottery sequence* $\boldsymbol{L}$

$$\Pr[\boldsymbol{c}(\boldsymbol{L}) \precsim \boldsymbol{\ell}] > \Pr[\boldsymbol{c}(\boldsymbol{L}) \succsim \boldsymbol{\ell}].$$

5.2. **Properties.** Definitions 3 and 5 implicitly assume some underlying partition (with respect to which the lottery sequences are defined). The following proposition establishes that if the definition holds for some partition then it holds for any partition.

Proposition 5.1. *If* $\overset{\vartriangle}{\precsim}$ *is weakly-SF-risk-averse with respect to some independent partition* $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$*, then it is also weakly-SF-risk-averse with respect to any independent partition. Similarly for (strong-)SF-risk-aversion.*

---

[17] If there are several certainty equivalents, then pick one arbitrarily.

[18] this is well defined, as the partition is *independent*, so the preferences on each pair of factors is independent of the others.

5.3. **The Quantitative Analogue.** As in Section 4, we now show that risk-aversion, as in definitions 3 and 5, corresponds to concavity of the NM utility with respect to the Debreu value function (if it exists).

Let $v^{\mathcal{T}_i} : \mathcal{T}_i \to \mathbb{R}$, be Debreu value functions for $\precsim$, with respect to the given partition. That is,

$$v(a_1, \ldots, a_n) = \sum_{i=1}^{n} v^{\mathcal{T}_i}(a_i)$$

represents $\precsim$.

As usual, let $\hat{u} = u \circ (v)^{-1}$ be the value-scaled-utility of $\overset{\vartriangle}{\precsim}$. We have,

Theorem 5. *For $\overset{\vartriangle}{\precsim}$ with value-scaled-utility $\hat{u}$, $\overset{\vartriangle}{\precsim}$ is weakly-SF-risk-averse if and only if $\hat{u}$ is concave, and (strong-)SF-risk-averse if and only if $\hat{u}$ is strictly-concave.*

So, once again, the axiomatic definition of risk aversion coincides with the Arrow-Pratt notion, once concavity is defined with respect to the Debreu value scale.

5.4. **Risk Loving and Risk Neutrality.** The definitions and theorems for risk-loving and risk neutrality are analogous.

Definition 6. $\overset{\vartriangle}{\precsim}$ *is* weakly-SF-risk-loving *if for any fair lottery sequence* $\boldsymbol{L}$*,*

$$\Pr[\boldsymbol{c}(\boldsymbol{L}) \precsim \boldsymbol{\ell}] \leq \Pr[\boldsymbol{c}(\boldsymbol{L}) \succsim \boldsymbol{\ell}]$$

and (strongly)-SF-risk-loving *if for any non-degenerate repeated lottery sequence* $\boldsymbol{L}$

$$\Pr[\boldsymbol{c}(\boldsymbol{L}) \precsim \boldsymbol{\ell}] < \Pr[\boldsymbol{c}(\boldsymbol{L}) \succsim \boldsymbol{\ell}].$$

$\overset{\vartriangle}{\precsim}$ *is* risk-SF-neutral *if for any fair lottery sequence* $\boldsymbol{L}$

$$\Pr[\boldsymbol{c}(\boldsymbol{L}) \precsim \boldsymbol{\ell}] = \Pr[\boldsymbol{c}(\boldsymbol{L}) \succsim \boldsymbol{\ell}]$$

Theorem 6. *For $\overset{\vartriangle}{\precsim}$ with value-scaled-utility $\hat{u}$, $\overset{\vartriangle}{\precsim}$ is weakly-SF-risk-loving if and only if $\hat{u}$ is convex, and (strongly-)-SF-risk-loving if and only if $\hat{u}$ is strictly-convex.*

$\overset{\vartriangle}{\precsim}$ is SF-risk-neutral by Definition 6 if and only if $\hat{u}$ is linear.

6. Multi-Commodity Risk Aversion

The seminal works of Arrow [2] and Pratt [28] defined risk aversion with respect to a single commodity – money. Ever since, researchers have considered how to extend the definition, and associated measures, to the multi-commodity setting (see [23, 32, 27, 10, 20, 29, 21, 26] 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 arises as to which scale should or can be used when measuring the concavity of the (multi-variable) utility function. Indeed, some papers (e.g. [10]) keep the multiple scales - in which case the measures of risk aversion become vectors (for the risk premium) and matrices (for the coefficient of absolute risk aversion).

Our approach here takes a different direction, which, in a way is the reverse. Rather than starting from the single commodity definition and extending it to multi-commodities, we 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 now show how, with this approach, the Arrow-Pratt framework carries over to the multi-commodity setting.

6.1. **DARA and CARA Preferences.** Arrow and Pratt introduced the *coefficient absolute risk aversion*, and the related notions of CARA - constant absolute risk aversion - and DARA - decreasing absolute risk aversion. They show that if (and only if) the utility function exhibits DARA then for a given lottery L, the associated *risk premium* decreases as the decision maker's wealth level increases. Similarly, the risk premium is independent of the wealth level if and only if the coefficient absolute risk aversion is constant. In particular, a decision maker accepting a lottery at some wealth level is guaranteed to accept the same lottery at a higher wealth level if and only if her utility function exhibits DARA or CARA.

The Arrow-Pratt coefficient of absolute risk aversion, and hence also the associated notions of CARA and DARA, are defined in a uni-commodity setting. As such, they do not carry over to the multi-commodity, or multi-period, setting, as exemplified next.

Consider two commodities, say the decision maker's property and her art collection, with $\$x$ being the value of the property and $\$y$ the value of the art work (and suppose that $x, y \geq 2$). Suppose that the decision maker's utility function is $u(x, y) = (\log_2 x + \log_2 y)^2$. Then, for any fixed y, the coefficient of absolute risk aversion of u with respect to x is DARA. However, with $y = 2$, the certainty equivalent of a fair lottery between $x = 2$ and $x = 4$ is $x \approx 2.93$, while with $y = 1000$ the certainty equivalent of the same lottery is ≈ 2.85, and with $y = 1,000,000$ it is ≈ 2.83. So, the certainty equivalent *decreases*, and the risk premium *increases*, as the wealth y grows. In particular, a decision maker with $x = 2.9$ will accept the fair lottery between $x = 2$ and $x = 4$ at a wealth level of $y = 2$, but would reject the same lottery at the higher wealth level $y = 1000$.

Thus, the Arrow-Pratt notions of DARA and CARA are limited to single commodities. We now show that, using our notion of risk aversion, these notions can be extended to the multi-commodity setting.

Consider a lottery preference $\overset{\scriptscriptstyle \Delta}{\precsim}$ with induced certainty preference $\precsim$. Let $\hat{u}$ be the value-scaled-utility of $\overset{\scriptscriptstyle \Delta}{\precsim}$ (that is $\hat{u} = u \circ v^{-1}$, where u and v are the NM utility and Debreu value functions representing $\stackrel{\vartriangle}{\precsim}$ and $\precsim$, respectively). We naturally define the *scale-free (SF) coefficient-of-absolute-risk-aversion of* $\stackrel{\vartriangle}{\precsim}$ as:

$$A_{\hat{u}} = -\left(\frac{\hat{u}''}{\hat{u}'}\right).$$

Note that this is the exactly the Arrow-Pratt coefficient of absolute risk aversion when using the v scale. SF-DARA and SF-CARA are naturally defined with respect to this coefficient. With this definition, once again, the risk premium of a lottery decreases with the wealth-level for DARA preferences. Specifically,

Proposition 6.1. *In the multi-commodity setting (with $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ an independent partition), the preference order $\stackrel{\vartriangle}{\precsim}$ exhibits SF-DARA if and only if for any factor $\mathcal{T}_i$, lottery L_i over $\mathcal{T}_i$, state $c_i \in \mathcal{T}_i$, and $\boldsymbol{d}_{-i}, \boldsymbol{d'}_{-i} \in \Omega_{-\{i\}}$, if $\boldsymbol{d}_{-i} \prec \boldsymbol{d'}_{-i}$ then*

$$(c_i, \boldsymbol{d}_{-i}) \stackrel{\vartriangle}{\precsim} (L_i, \boldsymbol{d}_{-i}) \Rightarrow (c_i, \boldsymbol{d'}_{-i}) \stackrel{\vartriangle}{\prec} (L_i, \boldsymbol{d'}_{-i}).$$

Similarly, $\stackrel{\vartriangle}{\precsim}$ exhibits SF-CARA if and only if for any commodity i, lottery L_i over $\mathcal{T}_i$, state $c_i \in \mathcal{T}_i$, and $\boldsymbol{d}_{-i}, \boldsymbol{d'}_{-i} \in \Omega_{-\{i\}}$

$$(c_i, \boldsymbol{d}_{-i}) \stackrel{\vartriangle}{\precsim} (L_i, \boldsymbol{d}_{-i}) \iff (c_i, \boldsymbol{d'}_{-i}) \stackrel{\vartriangle}{\precsim} (L_i, \boldsymbol{d'}_{-i})$$

So, once risk aversion is defined with respect to the value function, the DARA and CARA behavior carry over to the multi-commodity/multi-period setting.

6.2. **Correlation Aversion.** Richard [29] considered risk aversion in the multi-commodity setting, and introduced the following notion,

Definition 7. *Given an independent partition, $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$,*[19] *$\stackrel{\vartriangle}{\precsim}$ is* correlation-averse *with respect to factors $\mathcal{T}_i, \mathcal{T}_j$, $i \neq j$, if for any $a_i \prec b_i, a_j \prec b_j$, and $\boldsymbol{c} \in \mathcal{S}_{-\{i,j\}}$*

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

The preference is weakly correlation-averse *if* (3) *holds with a weak preference.*

This notion, which Richard [29] initially termed *multi-variate risk aversion*,[20] was introduced as "a new type of risk aversion unique to multivariate utility functions" [29]. A later paper states: "[Richard's definition] has nothing to do with what is generally known as risk aversion" [31].

We show that the two notions are deeply tied. We first show that a decision maker is (weakly) correlation averse if and only if she is (weakly) SF risk averse. This, in turn, establishes (by Theorem 5) that the decision maker is weakly correlation averse if and only if her NM utility function is concave with respect to the Debreu value function, and (strictly) correlation averse if and only if

[19] Technically, in [29] it is assumed that each $\mathcal{T}_i$ is a real interval. However, transformation into such an interval can always be obtained by means of [6].

[20] The now more familiar term - *correlation-aversion*, which we adopt here, was coined by Epstein and Tanny [13].

the utility function is strictly concave with respect to the value function. So, correlation aversion and Arrow-Pratt risk aversion are one and the same when using the value function scale.

Theorem 7. *If preference order* $\stackrel{\Delta}{\precsim}$ *is (weakly) SF risk averse then it is (weakly) correlation averse with respect to all pairs of factors* $\mathcal{T}_i, \mathcal{T}_j$. *Conversely, if* $\stackrel{\Delta}{\precsim}$ *is correlation averse with respect to some pair of factors* $\mathcal{T}_i, \mathcal{T}_j$, *then it is (weakly) SF risk averse*

Which, in turn establishes:

Corollary 6.1. *Let* $\stackrel{\Delta}{\precsim}$ *be a preference order over an independent partition* $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$, *with NM utility* u *and Debreu value function* v. *Then,* u *is concave with respect to* v *if and only if* $\stackrel{\Delta}{\precsim}$ *is weakly correlation averse with respect to any and all factor pairs* $\mathcal{T}_i, \mathcal{T}_j$. *Similarly,* u *is strictly concave with respect to* v *if and only if* $\stackrel{\Delta}{\precsim}$ *is correlation averse with respect to any and all factor pairs* $\mathcal{T}_i, \mathcal{T}_j$.

6.3. **Comparative Multi-Commodity Risk Aversion.** Kihlstrom and Mirman [23], observed that in the multi-commodity setting it is natural to limit risk aversion comparisons to decision makers agreeing on the certainty preferences.[21] 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. Consider two preference order $\stackrel{\Delta}{\precsim}_1, \stackrel{\Delta}{\precsim}_2$ on $\Delta(\mathcal{S})$ agreeing on the certainties. Let $\hat{u}_1$ and $\hat{u}_2$ be the respective value-scaled-utilities. For a lottery L let $c_j(L)$ be the certainty equivalent of L by $\hat{u}_j$ $(j = 1, 2)$. Then

$$c_1(L) \precsim c_2(L)$$

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.

7. Consumption-Saving Choice[22]

The following *portfolio selection* setting has been widely considered. Given a budget, a decision maker (DM) must distribute the funds amongst two assets, one secure and one risky, so as to maximize the expected utility. Pratt ([28] Theorem 7) shows that if DM-1 is more risk averse than

[21]Indeed, the same is also true for the classical framework, but it is implicitly assumed that all individuals agree on the certainty preferences: more money is better than less. Attempting to compare individuals who do not share this preference, e.g. comparing the risk attitude of Imelda Marcos with that of Saint Francis of Assisi, is meaningless also under the classic Arrow-Pratt framework.

[22]This is a new section and requires some additional work.

DM-2, then the amount DM-1 invests in the risky asset is less than the amount DM-2 invests in this asset. Arrow [2] shows that if the decision maker exhibits decreasing absolute risk aversion (DARA), then the amount invested in the risky asset increases with the overall budget, and if the decision maker exhibits decreasing relative risk aversion (DRRA), then the fraction invested in the risky asset increases with the overall budget (see also Diamond and Stiglitz [9], Corollary 2). These results, however, only hold for purely monetary investments, wherein, under certainty, a dollar of one asset is a perfect substitute for a dollar of the other asset. The results do not carry over to the multi-commodity setting, where different commodities may exhibit different and varying marginal utility for money. In particular, it is well known that the results do not (in general) apply to a two-period setting, where the decision maker needs to decide on the distribution of funds between sure consumption in the first period, and savings with a random return, for consumption in the second period.

Kihlstrom and Mirman [23] provide an extension of Pratt's result (on increases in risk aversion) to the two period setting (see also Diamond and Stiglitz [9]). Extending Arrow's analysis (on increases in wealth) has proved more challenging. In a subsequent work, Kihlstrom and Mirman [24] provide such an extension, but it is limited to homothetic preferences. Additionally, as emphasized in [24], their result extends that of DRRA preferences. To the best of our knowledge, there is no result to date that ties DARA preferences (under some definition of DARA) to investment behavior in the two period portfolio selection problem.

7.1. **The setting.** There are two periods, 1 and 2. Consumption of $\$c_1$ at time 1 and $\$c_2$ at time 2 yields an NM utility of $u(c_1, c_2)$. Funds for the second period are obtained via *savings* in some risky asset, with a rate of return $\tilde{z}$, which is a random variable. That is, upon saving of s, the available funds at time 2 are $\tilde{z} \cdot s$. For brevity, we shall consider $\tilde{z}$ with finite support. When z is non-random (that is, its entire support is in one value), then we write z instead of $\tilde{z}$.

A decision maker is endowed with wealth w, of which she consumes c_1 at time 1, and saves $s = w - c_1$ for time 2. Her expected utility is thus

$$E_{z\sim\tilde{z}}[u(w - s, zs)]. \tag{4}$$

She chooses $s^* = s^*(w, \tilde{z})$ that maximizes (4). We are interested in determining how $s^*(w, z)$ behaves as a function of w, and as a function of the level of risk aversion. For brevity, we write $s^*(w)$, when $\tilde{z}$ is fixed.

Suppose that the certainty preferences (induced by u) are additively separable. So, there exists Debreu value functions v_1, v_2, such that $v(c_1, c_2) = v_1(c_1) + v_2(c_2)$ represents this certainty preferences. We assume that v_1, v_2, are increasing (more money is always better). Let $\hat{u}$ be the value-scaled utility (that is, $\hat{u} = u \circ (v)^{-1}$). We assume that v_1, v_2, are differentiable, and $\hat{u}$ twice differentiable.

Note that in this setting it no longer holds that $s^*(w)$ is necessarily positive, even when $E[\tilde{z}] > 1$. As an example, consider $u(c_1, c_2) = 10c_1 + c_2$, and $\tilde{z} = 2$. Clearly, it is best to consume all funds in time 1. The following, however, is easy to prove

Proposition 7.1. *If $v_2'(0) \cdot E[\tilde{z}] > v_1'(w)$ then $s^*(w) > 0$.*

In particular, $s^*(w) > 0$ whenever $v_2'(0) \geq v_1'(w)$ and $E[\tilde{z}] > 0$. So, Proposition 7.1 is a generalization of Arrow's result that investment in a risky asset is non-zero whenever its expected return is positive (in the investment setting $v_2' = v_1' = 1$ everywhere).

7.2. **Increases in Wealth.** The following theorem establishes conditions under which $s^*(w)$ increases with the total budget w.

Theorem 8. *Provided that*

- *$v_1, v_2, \hat{u}$ are concave,*
- *$v_2 \circ \exp$ is convex,*

if $\hat{u}$ is DARA, then $s^*(w_-) \leq s^*(w_+)$ whenever $w_- < w_+$. If either v_1 is strictly concave or $v_2 \circ \exp$ strictly convex then $s^*(w_-) < s^*(w_+)$ (provided that $0 < s^*(w_-) > 0$ and $s^*(w_+) < w_+$).

The requirement that $v_2 \circ \exp$ is convex means that v_2 is convex when depicted on a logarithmic scale (of the wealth axis). The following proposition establishes that this is identical to stating that under certainty (that is - $\tilde{z}$ is non-random) s^* increases with z.

Proposition 7.2. *Assuming that v_1, v_2, and $\hat{u}$ are concave, then $v_2 \circ \exp$ is (strictly) convex (concave) if and only if for any w, $s^*(w, z)$ (res. strictly) increases (res. decreases) in z (provided that $0 < s^*(w, z) < w$).*

We note that Theorem 8 is a generalization of Arrow's result, as in Arrow's setting v_1, v_2, are both the identity, so v_1 is (weakly) concave and $v_2 \circ \exp$ strictly convex.

7.3. **Increases in Risk Aversion.** Next, we consider the affects of increasing the level of risk aversion. This is the setting considered by Kihlstrom and Mirman [23].

Previously (Section 6), we have shown that increased risk aversion coincides with an increased scale-free coefficient of absolute risk aversion, $A_{\hat{u}} = -(\hat{u}''/\hat{u}')$. Accordingly, we say that $\hat{u}$ is *more risk averse than* $\hat{u}_+$ if $A_{\hat{u}_+}(x) > A_{\hat{u}}(x)$ for all x. Let $s^*(w)$ be the optimal s by $\hat{u}$ and $s^*{}_+(w)$ the optimal s by $\hat{u}_+$.

Theorem 9. *Provided that $v_1, v_2, \hat{u}$ are concave, $\hat{u}_+$ is more risk averse than $\hat{u}$ then:*

- *$s^*(w) \geq s^*{}_+(w)$ for any w if$v_2 \circ \exp$ is convex, and $s^*(w) > s^*{}_+(w)$ if $v_2 \circ \exp$ is strictly convex (provided that $0 < s^*(w) < w$).*
- *$s^*(w) \leq s^*{}_+(w)$ for any w if$v_2 \circ \exp$ is concave, and $s^*(w) < s^*{}_+(w)$ if $v_2 \circ \exp$ is strictly concave (provided that $0 < s^*(w) < w$).*

Again, Theorem 9 is a generalization of Pratt's result.

Together with Proposition 7.2, Theorem 9 sates that s^* decreases with risk aversion if s^* increases with z, and increases with risk aversion if s^* increases with z. This is essentially the same result as obtained in [23].[23] The main contribution of Theorem 9 is in giving the conditions in purely functional form.

8. Temporal Discounting

Discounting of the future is a ubiquitous assumption in economic literature, and indeed, widely observed in practice. Scholars have identified several factors that may contribute to such discounting (see [16] for an excellent review). One of these factors is the uncertainty associated with the future - the decision maker may not live to enjoy the future. Other factors are more psychological in nature - e.g. a preference for immediate pleasure. The common discounted utility model bundles all factors in a single *discount rate*. We now show that the framework developed in this paper allows disentangling the former - discounting due to uncertainty associated with the future - from the other factors. We establish a *measure* for the rate of discounting associated with uncertainty of the future, and examine how (SF) risk aversion influences this discounting.

Consider a sequence of time periods $t = 1, \ldots, n$, and an associated consumption space $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$.[24] Suppose that while there are n *possible* time periods, actual consumption may be for a shorter duration, e.g. due to death. Specifically, for each $k = 1, \ldots, n$, there is a probability, s_k, that consumption is for only k time periods, with "death" occurring immediately following time k (here the term "death" should be interpreted in the wide sense, as any cause that may halt consumption - e.g. liquidation in case of a company). The vector $\boldsymbol{s} = (s_1, \ldots, s_n)$ is termed the *survival vector*. Clearly, the survival vector may influence the decision maker's preferences over the n-long sequences. For example, if the decision maker is sure to die after the first period, then only the preferences of the first time period matter; e.g. if 3 is preferred to 2 in the first period, then $(3, 2, \ldots)$ is preferred to say $(2, 100, \ldots)$. If, on the other hand, the decision maker is sure to live for both of the first time periods, then the preferences between these two time periods are determined by the risk-free inter-temporal rate of substitution between the two periods. At intermediate situations, when death between the first and the second time periods is possible but not certain, a combination of these two factors shall determine inter-temporal decision making, and, in particular, inter-temporal discounting. We now investigate the structure of the first of these factors; that is, inter-temporal discounting due to possible "death", and, in particular, the role of risk aversion in this discounting.

---

[23]Technically, we do not assume that $E[\tilde{z}] > 1$, which [23] does assume, but, on the other hand, [23] do not make our assumption that the certainty preferences are additively separable.

[24]Note that we do not suppose that the different $\mathcal{T}_k$'s are necessarily identical; that is, different time periods may offer different possible states.

Given $\boldsymbol{a} = (a_1, \dots, a_n) \in \mathcal{S}$, and survival vector $\boldsymbol{s} = (s_1, \dots, s_n)$, actual consumption is a lottery: outcome $(a_1, \dots, a_k)$ is obtained with probability s_k. So, for two such sequences, $\boldsymbol{a}$ and $\boldsymbol{b}$, the preference between the two is actually the preference between the associated lotteries. So, any survival vector $\boldsymbol{s}$ induces a preference order $\precsim_{\boldsymbol{s}}$ over $\mathcal{S}$, which reflects the preferences amongst the associated lotteries. Inter-temporal discounting in $\precsim_{\boldsymbol{s}}$ is determined by both risk-free inter-temporal preferences (over certainties), and by the probabilities in $\boldsymbol{s}$, and, as we will see, the decision maker's risk attitude.

For $k = 1, \dots, n$, denote $\mathcal{H}^k = \mathcal{T}_1 \times \cdots \mathcal{T}_k$ - the set of finite histories to time k. The set of all possible outcomes is $\mathcal{H}^* = \cup_{k=0}^n \mathcal{H}^k$. Let $\precsim^*$ and $\overset{\vartriangle}{\precsim}{}^*$ be agreeing preferences orders over $\mathcal{H}^*$ and $\Delta(\mathcal{H}^*)$, respectively. For each k, let $\precsim^k$ and $\overset{\vartriangle}{\precsim}{}^k$ be the induced certainty and lottery preferences on $\mathcal{H}^k$ and $\Delta(\mathcal{H}^k)$, respectively. We say that $\overset{\vartriangle}{\precsim}{}^*$ is risk averse if each $\overset{\vartriangle}{\precsim}{}^k$ is risk averse (in the sense of Definitions 3 and 5).

For the certainty preferences (but not necessarily the lottery preferences), suppose:

A1-Independence: $\mathcal{H}^k = \mathcal{T}_1 \times \cdots \times \mathcal{T}_k$ is an independent partition, with respect to $\precsim^k$, for all k.
A2-Consistency: for any k, the preference order induced by $\precsim^{k+1}$ on $\mathcal{H}^k$ is identical to $\precsim^k$.

Then, as in Proposition 4.1, we have

Proposition 8.1. *Assuming A1-A2, there exist Debreu value functions $v^{\mathcal{T}_i} : \mathcal{T}_i \to \mathbb{R}$, $i = 1, \dots, n$, such that for all k, $v_k = \sum_{i=1}^k v^{\mathcal{T}_i}$ represents $\precsim^k$.*

Let u^* be the NM utility representing $\overset{\vartriangle}{\precsim}{}^*$, and for each k, let u^k be the induced utility function on $\mathcal{H}^k$, representing $\overset{\vartriangle}{\precsim}{}^k$. Then, for a survival vector $\boldsymbol{s} = (s_1, \dots, s_n)$, the function $u_{\boldsymbol{s}}$ defined as

$$u_{\boldsymbol{s}}(a_1, \dots, a_n) = \sum_{k=1}^n s_k \cdot u^k(a_1, \dots, a_k). \tag{5}$$

represents $\precsim_{\boldsymbol{s}}$.

Assume that each $\mathcal{T}_i$ is some real interval I_i.[25] From (5), the *marginal rate of substitution* between the t-th and $t+1$ time periods, is

$$\frac{\frac{\partial u_{\boldsymbol{s}}}{\partial a_t}}{\frac{\partial u_{\boldsymbol{s}}}{\partial a_{t+1}}} = \frac{\sum_{k=1}^n \left(s_k \cdot \frac{\partial u^k}{\partial a_t} \right)}{\sum_{k=1}^n \left(s_k \cdot \frac{\partial u^k}{\partial a_{t+1}} \right)} \tag{6}$$

Let $\hat{u}_k$ be the value-scaled-utility of $\overset{\vartriangle}{\precsim}{}^k$; that is, $\hat{u}_k = u^k \circ (v_k)^{-1}$. Then, (6), at a given point $(a_1, \dots, a_n)$, can be re-written as

$$\frac{\sum_{k=1}^n \left(s_k \cdot \left.\frac{d\hat{u}_k}{dx}\right|_{x=v_k(a_1,\dots,a_k)} \cdot \left.\frac{\partial v_k}{\partial a_t}\right|_{(a_1,\dots,a_k)} \right)}{\sum_{k=1}^n \left(s_k \cdot \left.\frac{d\hat{u}_k}{dx}\right|_{x=v_k(a_1,\dots,a_k)} \cdot \left.\frac{\partial v_k}{\partial a_{t+1}}\right|_{(a_1,\dots,a_k)} \right)} \tag{7}$$

[25] This is only for technical reasons, so that the marginal-rate-of-substitution is well defined. Note, however, that we do not assume that the I_i's are identical.

By definition, $v_k(a_1,\ldots,a_k)=\sum_{i=1}^{k} v^{\mathcal{T}_i}(a_i)$. So for $k<t$, $\frac{\partial v_k}{\partial a_t}=0$, and for $k\geq t$, $\frac{\partial v_k}{\partial a_t}=\frac{dv^{\mathcal{T}_t}}{da_t}$, and similarly for $t+1$. So, (7) is

$$\frac{\sum_{k=t}^{n}\left(s_k\cdot\left.\frac{d\hat{u}_k}{dx}\right|_{x=v_k(a_1,\ldots,a_k)}\cdot\frac{dv^{\mathcal{T}_t}}{da_t}\right)}{\sum_{k=t+1}^{n}\left(s_k\cdot\left.\frac{d\hat{u}_k}{dx}\right|_{x=v_k(a_1,\ldots,a_k)}\cdot\frac{dv^{\mathcal{T}_{t+1}}}{da_{t+1}}\right)},$$

which is,

$$\left(\frac{\frac{dv^{\mathcal{T}_t}}{da_t}}{\frac{dv^{\mathcal{T}_{t+1}}}{da_{t+1}}}\right)\cdot\left(\frac{\sum_{k=t}^{n} s_k\cdot\left.\frac{d\hat{u}_k}{dx}\right|_{x=v_k(a_1,\ldots,a_k)}}{\sum_{k=t+1}^{n} s_k\cdot\left.\frac{d\hat{u}_k}{dx}\right|_{x=v_k(a_1,\ldots,a_k)}}\right) \tag{8}$$

So, the marginal rate of substitution is the product of two components. The first, $\frac{dv^{\mathcal{T}_t}}{da_t}/\frac{dv^{\mathcal{T}_{t+1}}}{da_{t+1}}$, is the standard, risk free marginal rate of substitution between the two time periods. This is the marginal rate of substitution that would be exhibited if there would be no risk of death. The second component, which is our subject of interest, is the component of the marginal rate of substitution directly associated with the risks of death. We call this second component, the *uncertainty discount rate*, denoted $UD_{\boldsymbol{s}}^{t,t+1}$.

Rewriting (8), we have,

$$UD_{\boldsymbol{s}}^{t,t+1}(a_1,\ldots,a_n)=1+\frac{s_t\cdot\left.\frac{d\hat{u}_t}{dx}\right|_{x=v_t(a_1,\ldots,a_t)}}{\sum_{k=t+1}^{n} s_k\cdot\left.\frac{d\hat{u}_k}{dx}\right|_{x=v_k(a_1,\ldots,a_k)}}. \tag{9}$$

Each of the functions $\hat{u}_k$ is monotone increasing, so the second addend is non-negative. So, $UD_{\boldsymbol{s}}^{t,t+1}\geq 1$, and it is indeed a *discount* rate. The magnitude of the discount is determined by several factors, which we now explore.

Dependency on $\boldsymbol{s}$. First note that $UD_{\boldsymbol{s}}^{t,t+1}$ approaches 1 as s_t goes to zero. This stands to reason, as $s_t=0$ means that there is no probability of death between time periods t and $t+1$, so there is no death associated discounting between the two. On the other hand, $UD_{\boldsymbol{s}}^{t,t+1}$ approaches infinity as $\sum_{k=t+1}^{n} s_k$ approaches zero. This again makes sense, as $\sum_{k=t+1}^{n} s_k=0$ means that there is no chance to live beyond the t-th time period, so it is better to move any possible (good) consumption to time t.

Dependency on $\boldsymbol{a}$. For a fixed $\boldsymbol{s}$, the magnitude of $UD_{\boldsymbol{s}}^{t,t+1}(a_1,\ldots,a_n)$ is determined by both the risk attitude, as captured by the curvature of the $\hat{u}_k$'s (that is, the curvature of the u^k with respect to v_k), and by the sequence $\boldsymbol{a}=(a_1,\ldots,a_n)$. The dependency on $\boldsymbol{a}$ can be divided into two: the dependency on the past - $(a_1,\ldots,a_t)$ - and the dependency on the future - $(a_{t+1},\ldots,a_n)$.

Dependency on the future. Suppose that u^* is risk averse, in the sense defined in this paper; that is, each $\hat{u}_k$ is concave. So, $\frac{d\hat{u}_k}{dx}$ is monotone decreasing for each k. So, for a fixed $a_1, \ldots, a_t$, the denominator of the second addend in (9) decreases with x. So, $UD_{\boldsymbol{s}}^{t,t+1}(a_1, \ldots, a_n)$ increases as $a_{t+1}, \ldots, a_n$, become more favorable. The explanation for this is that risk averters dislike big risks. As the "future" - $a_{t+1}, \ldots, a_n$ - becomes brighter, the risk of dying at time t becomes larger. So, a risk averter is more inclined to move some of the better consumption to the current time period, to better balance the risk. The opposite holds for risk lovers (again, as defined in this paper). For risk lovers the $\hat{u}_k$'s are convex. So, $UD_{\boldsymbol{s}}^{t,t+1}(a_1, \ldots, a_n)$ decreases as $a_{t+1}, \ldots, a_n$ becomes more favorable.

We thus obtain:

Proposition 8.2. *Assuming A1-A2. Given: a time t, survival vector $\boldsymbol{s}$ (with $s_k \neq 0$, $k \geq t$), past and present $a_1, \ldots, a_t$, and future sequences, $(b_{t+1}, \ldots, b_n), (c_{t+1}, \ldots, c_n)$, with $(b_{t+1}, \ldots, b_n) \prec (c_{t+1}, \ldots, c_n)$*[26]

- *if $\stackrel{\triangle}{\precsim}^*$ is SF-weakly-risk-averse then*

$$UD_{\boldsymbol{s}}^{t,t+1}(a_1, \ldots, a_t, b_{t+1}, \ldots, b_n) \leq UD_{\boldsymbol{s}}^{t,t+1}(a_1, \ldots, a_t, c_{t+1}, \ldots, c_n).$$

  and the inequality is strict if $\stackrel{\triangle}{\precsim}^*$ is (strictly-)SF-risk-averse.
- *if $\stackrel{\triangle}{\precsim}^*$ is weakly-SF-risk-loving then*

$$UD_{\boldsymbol{s}}^{t,t+1}(a_1, \ldots, a_t, b_{t+1}, \ldots, b_n) \geq UD_{\boldsymbol{s}}^{t,t+1}(a_1, \ldots, a_t, c_{t+1}, \ldots, b_n),$$

  and the inequality is strict if $\stackrel{\triangle}{\precsim}^*$ is (strictly-)SF-risk-loving.

So, risk averters discount more as the future becomes brighter, whereas risk lovers discount less.

Dependency on the past and present. The dependency on the past and present - $a_1, \ldots, a_t$ - exhibits a different behavior. Here, the determining factor is the directional *change* in the scale-free coefficient of absolute risk aversion - SF-DARA, SF-CARA, or SF-IARA (*Increasing Absolute Risk Aversion*). In addition, the prospects of the future, vis a vee death, are a determining factor. Proposition 8.3, which we state shortly, summarizes the relationship.

Before stating the proposition, we must generalize the notions DARA and IARA to our setting. Note that the space under consideration, $\mathcal{H}^*$, is a *union* of spaces, $\mathcal{H}^k$, one for each possible life duration. Each sub-space $\mathcal{H}^k$ has its own scale v_k. So, the regular notion of DARA is not well defined. The following definition generalizes the notions of DARA, CARA and IARA to this setting.

Definition 8. *Let $\stackrel{\triangle}{\precsim}^*$ be a preference order on $\Delta(\mathcal{H}^*)$, with NM utility u^*, and for each k, let $\hat{u}_k$ be the induced value-scaled-utility on $\mathcal{H}^k$. We say that $\stackrel{\triangle}{\precsim}^*$ is SF-DARA if for any $(a_1, \ldots, a_k), (b_1, \ldots, b_\ell)$,*

[26] this is well defined as $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ is an independent partition.

if $(a_1, \ldots, a_k) \overset{\vartriangle}{\prec}^* (b_1, \ldots, b_\ell)$*, then*

$$A_{\hat{u}_k}(v_k(a_1, \ldots, a_k)) \geq A_{\hat{u}_\ell}(v_\ell(b_1, \ldots, b_\ell)), \tag{10}$$

(where $A_{\hat{u}_j}$ is the coefficient of absolute risk aversion of the function $\hat{u}_j$), and strict-SF-DARA *if* (10) *holds with strict inequality.*[27] $\overset{\vartriangle}{\precsim}^*$ *is* (strict)-SF-IARA *if the inequality in* (10) *is (strict) in the opposite direction, and SF-CARA if* (10) *holds with equality.*

We say that future $(a_{t+1}, \ldots, a_n)$ is *better-than-death* if for all $k > t$, $(a_1, \ldots, a_t) \prec (a_1, \ldots, a_k)$; this future is *worse-than-death* if strict preference holds in the reverse direction, and *as-good-as-death* if there is indifference between the sides.

Proposition 8.3. *Assuming A1-A2. Given: a time t, survival vector $\boldsymbol{s}$ (with $s_k \neq 0$, $k \geq t$), and better-than-death future $c_{t+1}, \ldots, c_n$, for any past-to-present sequences $(a_1, \ldots, a_t) \prec (b_1, \ldots, b_t)$*

- *if each $\overset{\vartriangle}{\precsim}^*$ is SF-DARA then*

$$UD_{\boldsymbol{s}}^{t,t+1}(a_1, \ldots, a_t, c_{t+1}, \ldots, c_n) \geq UD_{\boldsymbol{s}}^{t,t+1}(b_1, \ldots, b_t, c_{t+1}, \ldots, c_n),$$

  and the inequality is strict if $\overset{\vartriangle}{\precsim}^*$ is strictly-SF-DARA.
- *if $\overset{\vartriangle}{\precsim}^*$ is SF-CARA then*

$$UD_{\boldsymbol{s}}^{t,t+1}(a_1, \ldots, a_t, c_{t+1}, \ldots, c_n) = UD_{\boldsymbol{s}}^{t,t+1}(b_1, \ldots, b_t, c_{t+1}, \ldots, b_n).$$

- *if $\overset{\vartriangle}{\precsim}^*$ is SF-IARA then*

$$UD_{\boldsymbol{s}}^{t,t+1}(a_1, \ldots, a_t, c_{t+1}, \ldots, c_n) \leq UD_{\boldsymbol{s}}^{t,t+1}(b_1, \ldots, b_t, c_{t+1}, \ldots, b_n),$$

  and the inequality is strict if $\overset{\vartriangle}{\precsim}^*$ is strictly-SF-IARA.

In all cases, the inequality is reversed in case the future is worse-than-death.

In words: for a better-than-death future, with SF-DARA preferences the discount rate decreases as the past and present are more favorable, with SF-IARA preferences the discount rate increases as the past and present are more favorable, and with SF-CARA preferences the discount rate is independent of the past and the present. The behavior is reversed for a worse-than-death future.

9. Related Work

Strength of Preference and Relative Risk Aversion. Dyer and Sarin [12] and Bell and Raiffa [5] 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

[27] Note that this is indeed a generalization of the case with a single space, as u is assumed to be monotone in the underlying scale (money in the Arrow-Pratt definition, v in our case).

$x_1 \longmapsto x_2$ 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* [11]) 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 [12] define the notion of *relative risk aversion*[28] as the concavity of the NM utility function u with respect to the measurable value function f. Bell and Raiffa [5] similarly define the notion of *intrinsic risk aversion.*

Bell and Raiffa [5] 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 [12, Theorem 1]). Thus, Theorem 5 establishes that, technically, risk aversion as of Definitions 3 and 5 coincides with the Dyer and Sarin notion of *relative risk aversion*, provided that the 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 [12] and [5]. 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, [12] and [5] do not provide an axiomatic justification for their change of scale.

Intertemporal Risk Aversion. An independent work of Traeger [33] is motivated by some of the same questions that motivate our work; namely - providing a scale-free, axiomatic framework for the theory of risk aversion.[29] Unlike our work, Traeger's work is set in the Kreps and Porteus [25] *temporal lottery framework*, wherein the timing of risk resolution is important. The core definition of [33] is that of *Intertemporal Risk Aversion (IRA)*, which is a version of correlation aversion wherein the non-correlated outcome is risk-free.[30] Intertemporal Risk Aversion is defined as a preference for the risk-free non-correlated alternative over the risky, correlated one. [31] Importantly, Theorem 2 of [33] establishes that a decision maker is intertemporal risk averse at time t if and only if $f_t \circ (g_t)^{-1}$ is concave, where f_t is the (time-t) risk aggregation function, and g_t is the (time-t) inter-temporal aggregation function. This is analogous to our Theorem 5 and Corollary 6.1. Thus, using a different axiomatic definition, and set in a different framework, Treager [33] reaches the same core conclusion of assessing risk aversion in terms of the concavity of risk utility function, with respect to the certainty aggregation function. In subsequent work, Treager [34] studies the impact of intertemporal risk aversion on discounting for an uncertain growth rate (using the Epstein-Zin [14] intertemporal model). This study is different from our study of discounting due to uncertainty in life duration (Section 8), but again emphasizes the importance of considering the more general form of risk aversion when analyzing multi-period risky settings.

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

[29] We are grateful to Jean-Christophe Vergnaud for calling our attention to this important work.

[30] See Lemma A.10 of the appendix, which relates the two definitions.

[31] We investigated the relationship between (standard) correlation aversion and our notion of SF-risk-aversion in Section 6.2.

10. Discussion

We presented two axiomatic 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. In the second definition, the probability of an inferior outcome is greater than the probability of a superior one. 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.

Inter-Commodity and Intra-Commodity Risk Aversion. We should stress that (scale-free) risk-aversion, as considered in this paper, does not relate only to lotteries involving multiple commodities or time periods, but also to lotteries *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.

Disentangling Risk Aversion from Diminishing Marginal Utility. It is well known that under the classical definition, risk aversion and diminishing marginal utility are 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 [35] 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. For example, suppose that for a two day setting, the utility of the decision maker is $u(x, y) = \ln(x) + \ln(y)$, where x and y are the total consumption levels of the decision maker, in kilograms, in day one and two, respectively. Then, the Debreu value in each period is logarithmic in the number of kilograms consumed, and the NM utility function is linear in the Debreu value. Under our interpretation, the logarithmic Debreu value functions correspond to diminishing marginal utility, with respect to kilograms of consumption, while the linearity of the NM utility with respect to the value function corresponds to risk neutrality.

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 (in Debreu value units) to the existing state. So, by our definition, the gamble would be rejected by risk neutral (or even some risk loving) individuals.

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, 30, 17]. By our definition, this corresponds to an assumption of risk neutrality. Accordingly, in a sequel work [4], we consider a theory of repeated games without this additivity assumption. We show that when players are risk averse - according to our scale-free definitions - new equilibria emerge, unaccounted for by the classic theory. In particular, in two person matching pennies games, if one player is risk averse and the other risk loving, then the resulting pure strategy equilibria are necessarily biased in favor of the risk loving player. Such biased equilibria are not possible in the classic 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] Yonatan Aumann and Erel Segal-Halevi. Repeated games with risk averse players. see https://sites.google.com/site/aumannbiu/home/publication-by-topic-1, 2016.

[5] 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.

[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] Peter A Diamond and Joseph E Stiglitz. Increases in risk and in risk aversion. *Journal of Economic Theory*, 8(3):337–360, 1974.

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

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

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

[13] 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.

[14] Larry G. Epstein and Stanley E. Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4):937–69, July 1989.

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

[16] Shane Frederick, George Loewenstein, and Ted O'Donoghue. Time discounting and time preference: A critical review. *Journal of Economic Literature*, 40(2):351–401, 2002.

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

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

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

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

[21] 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.

[22] 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.

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

[24] Richard E. Kihlstrom and Leonard J. Mirman. Constant, increasing and decreasing risk aversion with many commodities. *The Review of Economic Studies*, 48(2):271–280, 1981.

[25] David M. Kreps and Evan L. Porteus. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 46(1):185–200, 1978.

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

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

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

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

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

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

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

[33] Christian P. Traeger. Capturing intrinsic risk attitude. see https://are.berkeley.edu/ traeger/research.html, 2014.

[34] Christian P Traeger. Why uncertainty matters: discounting under intertemporal risk aversion and ambiguity. *Economic Theory*, 56(3):627–664, 2014.

[35] 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.

Throughout this appendix, the writing follows certain conventions that simplify the presentation:

- x, y, are real numbers, 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.

Proofs for Section 4.

Proposition 4.1. *Assuming A1-A2, 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

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

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

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 $\stackrel{\triangle}{\precsim}^n$, 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 there exists $s > 0$ with $\phi_i \pm s \in \mathcal{T}_i$, for all i.

Proofs for Section 4.2.1.

Lemma A.1. *Let $X_1, X_2, \dots$ be an infinite sequence of independent uniformly bounded random variables,*[32] *with $E(X_i) = 0$ for all i. Set $S_n = \sum_{i=1}^n X_i$. Then*

$$\Pr[S_n \geq 0 \text{ infinitely often}] > 0. \tag{11}$$

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. [19], 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. *Assuming A1-A3, $\stackrel{\Delta}{\succsim}$ is weakly-SF-risk-averse if and only if all the valued-scaled-utilities $\hat{u}_n$ are concave.*

Proof. $\stackrel{\Delta}{\succsim}$ is weakly SF risk averse $\Rightarrow$ all $\hat{u}_n$ are concave: 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

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

[32] 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.

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

$$\begin{aligned} &\hat{u}_m(x + \phi_{k+1} + \cdots + \phi_m + \beta) = \\ &\quad = \frac{1}{2}\left(\hat{u}_m(x + \phi_{k+1} + \cdots + \phi_m - \epsilon) + \hat{u}_m(x + \phi_{k+1} + \cdots + \phi_m + \epsilon)\right). \end{aligned} \tag{12}$$

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 (12), 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 (12) 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 $(L_1, L_2, \ldots)$.

<u>All $\hat{u}_n$ are concave $\Rightarrow \stackrel{*}{\succsim}$ is weakly SF risk averse:</u> 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 $risk\text{-}prem_f(x, L)$ 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*

$$risk\text{-}prem_{\hat{u}_n}(x, L) \geq risk\text{-}prem_{cara_\alpha}(x, L) \tag{13}$$

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

$$risk\text{-}prem_{\hat{u}_n}(x, \pm\epsilon) \geq \alpha\epsilon^2 \tag{14}$$

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

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

$$\begin{aligned} \frac{cara_\alpha(\epsilon) + cara_\alpha(-\epsilon)}{2} &= \frac{-e^{-\alpha\cdot\epsilon} - e^{\alpha\cdot\epsilon}}{2} \\ &= -\frac{1}{2}(1 - \alpha\epsilon + \frac{\alpha^2\epsilon^2}{2} + 1 + \alpha\epsilon + \frac{\alpha^2\epsilon^2}{2} + O(\epsilon^3)) \\ &= -(1 + \frac{\alpha^2\epsilon^2}{2} + O(\epsilon^3)) \end{aligned} \tag{15}$$

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,

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

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

$$risk\text{-}prem_{cara_\alpha}(x, \pm\epsilon) < \frac{2}{3}\alpha\epsilon^2, \tag{16}$$

for all x.

So, combining (14) and (16)

$$risk\text{-}prem_{\hat{u}_n}(x, \pm\epsilon) > risk\text{-}prem_{cara_\alpha}(x, \pm\epsilon), \tag{17}$$

for ϵ sufficiently small. But then, by Pratt [28], (17) holds for any lottery L.

Conversely, if $risk\text{-}prem_{\hat{u}_n}(x, \pm\epsilon) \geq risk\text{-}prem_{cara_\alpha}(x, \pm\epsilon)$ then by (15)

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

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

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,

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

By assumption, $\hat{\precsim}^k$ and $\hat{\precsim}^m$ 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,

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

Hence,

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

□

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, $1 \le i \le n$, n arbitrary, be independent random variables with $E[X_i] = 0$, $|X_i| \le C$, 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 $0 < a \le \delta \cdot \Sigma_n$*

$$\Pr[S_n > a\Sigma_n] < e^{-\frac{a^2}{2}(1-\gamma)}. \tag{18}$$

Lemma A.6. *Let $X_1, X_2, \ldots,$ be independent random variables with $E[X_i] = 0$, $|X_i| \le C$, 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 $\alpha > 0$*

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

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

Denote by A_k the event that there exists i, $n(k) < i \le n(k+1)$, 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 $a = \beta\Sigma_{n(k)}$

$$\Pr[S_{n(k)} > \beta\Sigma_{n(k)} \cdot \Sigma_{n(k)}] < e^{-\frac{\beta^2\Sigma_{n(k)}^2}{2}(1-\gamma)} \le e^{-\frac{\beta^2 k}{4}} \tag{19}$$

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,

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

So, by the Kolmogorov inequality

$$\Pr[\max_{n(k)<j\le n(k+1)}\{D_j\} \ge \beta\Sigma_{n(k)}^2] \le \frac{Var(D_{n(k+1)})}{(\beta\Sigma_{n(k)}^2)^2} \le \frac{1+C^2}{\beta^2 k^2}. \tag{20}$$

Combining (19)-(20), for any k

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

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 $n(k) < i \le n(k+1)$. So, $S_i > \alpha\Sigma_i^2$ infinitely often only if A_k occurs infinitely often, and the result follows. ∎

Theorem 3. *Assuming A1-A3,*

(a) If $RP_{\hat{\mathbf{u}}}(\epsilon) = \Omega(\epsilon^2)$ as $\epsilon \to 0$ then $\hat{\precsim}$ is SF risk averse.[33]

(b) If $RP_{\hat{\mathbf{u}}}(\epsilon) = O(\epsilon^{2+\beta})$ as $\epsilon \to 0$, for some $\beta > 0$, then $\hat{\precsim}$ is not SF 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 $|L_i| \leq C$ for all i.

By the Taylor expansion,

$$cara_\alpha(\epsilon) = -e^{-\alpha\epsilon} = -1 + \alpha\epsilon - \frac{\alpha^2\epsilon^2}{2} + O(\alpha^3\epsilon^3). \tag{21}$$

Let α_1 be such that the $O(\alpha^3\epsilon^3)$ term in (21) is small for $|\epsilon| \leq C$; that is,

$$cara_{\alpha_1}(\epsilon) \approx -1 + \alpha_1\epsilon - \frac{\alpha_1^2\epsilon^2}{2}, \tag{22}$$

for $|\epsilon| \leq C$.

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,

$$\left[-\alpha \cdot (\Sigma_n)^2 < S_n\right] \Rightarrow \left[\sum_{i=1}^n c_i < S_n\right] \Rightarrow \left[(c_1, \ldots, c_n) \prec (\ell_1, \ldots, \ell_n)\right]. \tag{23}$$

---

[33] 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

$$\Pr[S_n < -\alpha(\Sigma_n)^2 \text{ infinitely often}] = 0, \tag{24}$$

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

$$\textit{risk-prem}_{\hat{u}_i}(x, \pm\epsilon) \le \alpha \cdot \epsilon^{2+\beta}. \tag{25}$$

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 (25), for any j there exists i_j and x_j with

$$\textit{risk-prem}_{\hat{u}_{i_j}}(x_j, \pm a_j) \le \alpha \cdot a_j^{2+\beta}. \tag{26}$$

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 (26). Set $y_{n(j-1)} = c_1 + \cdots + c_{n(j-1)}$. By Lemma A.4 and (26), 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 $m+1$)[34],

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

which means that

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

[34]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

$$c_{m+1} \geq \phi_{m+1} - \alpha a_j^{2+\beta}. \tag{27}$$

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

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

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

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. $\square$

Theorem 4. *Assuming A1-A3, 1 and 3*

(a) Weak risk loving: $\stackrel{\Delta}{\precsim}$ is weakly SF risk loving if and only if all the all the valued-scaled-utilities $\hat{u}_n$ are convex.

(b) Risk loving

- *If $(-RP_{\hat{\mathbf{u}}}(\epsilon)) = \Omega(\epsilon^2)$ as $\epsilon \to 0$ then $\stackrel{\Delta}{\precsim}$ is SF 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 SF risk loving.*

(c) Risk Neutrality: $\stackrel{\Delta}{\precsim}$ is SF risk neutral if and only if $\hat{u}_n$ is linear 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. $\square$

Proofs for Section 5. The proof of Proposition 5.1 is easier with the aid of Theorem 5, so we start with proving the theorem and then come back to proving the proposition.

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 [15, Theorem 8.2]). Furthermore, since $\stackrel{\Delta}{\precsim}$ is continuous, so is u.

Theorem 5. *For $\stackrel{\triangle}{\precsim}$ with value-scaled-utility $\hat{u}$, $\stackrel{\triangle}{\precsim}$ is weakly-SF-risk-averse if and only if $\hat{u}$ is concave, and (strong-)SF-risk-averse if and only if $\hat{u}$ is strictly-concave.*

Proof. Set $\hat{u} = u \circ v^{-1}$.

Concavity $\Rightarrow$ Weak SF Risk Aversion. Suppose $\hat{u}$ is concave. Consider a fair lottery sequence $\boldsymbol{L} = (L_1, \ldots, L_n)$. Then,

$$\hat{u}(v(\boldsymbol{c}(\boldsymbol{L}))) = E_{\boldsymbol{\ell}\sim\boldsymbol{L}}[\hat{u}(v(\boldsymbol{\ell}))] \leq \hat{u}(E_{\boldsymbol{\ell}\sim\boldsymbol{L}}[v(\boldsymbol{\ell})]),$$

where the inequality is by concavity of $\hat{u}$. So, since $\hat{u}$ is monotone

$$v(\boldsymbol{c}(\boldsymbol{L})) \leq E_{\boldsymbol{\ell}\sim\boldsymbol{L}}[v(\boldsymbol{\ell})].$$

Since all the L_i's are fair, the distribution of $v(\boldsymbol{\ell})$ is symmetric around $E(v(\boldsymbol{\ell}))$. So,

$$\Pr[E[v(\boldsymbol{\ell})] \leq v(\boldsymbol{\ell})] = \Pr[E[v(\boldsymbol{\ell})] \geq v(\boldsymbol{\ell})]$$

So,

$$\begin{aligned}
\Pr[\boldsymbol{c}(\boldsymbol{L}) \precsim \boldsymbol{\ell}] &= \Pr[v(\boldsymbol{c}(\boldsymbol{L})) \leq v(\boldsymbol{\ell})] \\
&\geq \Pr[E[v(\boldsymbol{\ell})] \leq v(\boldsymbol{\ell})] \\
&= \Pr[E[v(\boldsymbol{\ell})] \geq v(\boldsymbol{\ell})] \\
&\geq \Pr[v(\boldsymbol{c}(\boldsymbol{L})) \geq v(\boldsymbol{\ell})] \geq \Pr[\boldsymbol{c}(\boldsymbol{L}) \succsim \boldsymbol{\ell}].
\end{aligned}$$

Strict Concavity $\Rightarrow$ Strong-SF-Risk-Aversion. Consider a non-degenerate repeated lottery sequence $\boldsymbol{L} = (L_1, L_2, \boldsymbol{d})$ (w.l.o.g. we may assume $i = 1, j = 2$). Suppose that L_1 is the fair lottery $\langle a_1^-, a_1^+ \rangle$ (the fair lottery between a_1^- and a_1^+), and similarly $L_2 = \langle a_2^-, a_2^+ \rangle$.

Lotteries L_1 and L_2 are of the same magnitude. That is,

$$(a_1^-, a_2^+) \sim (a_1^+, a_2^-).$$

So

$$v^{\mathcal{T}_1}(a_1^-) + v^{\mathcal{T}_2}(a_2^+) = v^{\mathcal{T}_1}(a_1^+) + v^{\mathcal{T}_2}(a_2^-). \tag{28}$$

The four possible, equi-probability realizations of $\boldsymbol{L}$ are:

(1) $\boldsymbol{\ell} = (a_1^-, a_2^-, \boldsymbol{d})$
(2) $\boldsymbol{\ell} = (a_1^-, a_2^+, \boldsymbol{d})$
(3) $\boldsymbol{\ell} = (a_1^+, a_2^-, \boldsymbol{d})$
(4) $\boldsymbol{\ell} = (a_1^+, a_2^+, \boldsymbol{d})$

So,

$$E_{\boldsymbol{\ell}\sim\boldsymbol{L}}[(v(\boldsymbol{\ell})] = \frac{v^{\mathcal{T}_1}(a_1^-) + v^{\mathcal{T}_1}(a_1^+) + v^{\mathcal{T}_2}(a_2^-) + v^{\mathcal{T}_2}(a_2^+)}{2} + \sum_{i=3}^{n} v^{\mathcal{T}_{d_i}}(d_i) = v(a_1^-, a_2^+, \boldsymbol{d})$$

where the second equality is by (28).

Suppose $\hat{u}$ is strictly concave. Then

$$\hat{u}(v(\boldsymbol{c}(\boldsymbol{L}))) = E_{\boldsymbol{\ell}\sim\boldsymbol{L}}[\hat{u}(v(\boldsymbol{\ell}))] < \hat{u}(E_{\boldsymbol{\ell}\sim\boldsymbol{L}}[(v(\boldsymbol{\ell})]) = \hat{u}(v(a_1^-, a_2^+, \boldsymbol{d})).$$

So,

$$\boldsymbol{c}(\boldsymbol{L}) \prec (a_1^-, a_2^+, \boldsymbol{d}).$$

So, of the four possible realization of $\boldsymbol{L}$ only $(a_1^-, a_2^-, \boldsymbol{d})$ is $\precsim \boldsymbol{c}(\boldsymbol{L})$. So,

$$\begin{aligned}\Pr[\boldsymbol{c}(\boldsymbol{L}) \precsim \boldsymbol{\ell}] &= 0.75\\ \Pr[\boldsymbol{c}(\boldsymbol{L}) \succsim \boldsymbol{\ell}] &= 0.25,\end{aligned}$$

as necessary.

Weak-SF-Risk-Aversion ⇒ Concavity: Conversely, suppose that $\hat{u}$ is not concave. Then, since it is continuous, it is strictly convex on some interval $(\underline{x}, \overline{x})$. Set $z = \frac{\underline{x}+\overline{x}}{2}$ and $\epsilon_1 = z - \underline{x}$. By definition, there exists some $(a_1, a_2, \boldsymbol{d})$ with $z = v(a_1, a_2, \boldsymbol{d})$. Since z is internal in $(\underline{x}, \overline{x})$, we may assume, w.l.o.g. that a_1, a_2 are not extreme in $\mathcal{T}_1, \mathcal{T}_2$, respectively (that is, there exists a', a'', with $a' \prec a_1 \prec a''$, and similarly for a_2). Choose $\epsilon < \epsilon_1$ sufficiently small so that there exists a_1^-, a_2^-, with $v^{\mathcal{T}_i}(a_i^-) = v^{\mathcal{T}_{a_i}} - \epsilon$, and a_1^+, a_2^+, with $v^{\mathcal{T}_i}(a_i^+) = v^{\mathcal{T}_{a_i}} + \epsilon$, $i = 1, 2$. Set $L_i = \langle a_i^-, a_i^+ \rangle$, for $= 1, 2$. Set $\boldsymbol{L} = (L_1, L_2, \boldsymbol{d})$. Then, $\boldsymbol{L}$ is a repeated lottery, and $\hat{u}$ is strictly convex on the domain of its realizations. So, by the reasoning above (for "strict-concavity ⇒ strict-risk-aversion"),

$$\Pr[\boldsymbol{c}(\boldsymbol{L}) \precsim \boldsymbol{\ell}] < \Pr[\boldsymbol{c}(\boldsymbol{L}) \succsim \boldsymbol{\ell}].$$

So, u is not weakly risk averse, in contradiction.

Strong-SF-Risk-Aversion ⇒ Strict Concavity: Suppose that $\hat{u}$ is not strictly concave. Then, it is convex on some interval. So, as above, we can construct a repeated lottery on this interval. For this lottery, by the above reasoning (for "concavity ⇒ weak-risk-aversion"),

$$\Pr[\boldsymbol{c}(\boldsymbol{L}) \precsim \boldsymbol{\ell}] \leq \Pr[\boldsymbol{c}(\boldsymbol{L}) \succsim \boldsymbol{\ell}].$$

So, $\overset{\vartriangle}{\precsim}$ is not strongly-SF-risk-averse. □

We now return to proving Proposition 5.1.

Lemma A.7. *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. 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 $T \neq \emptyset$.

Gorman [18, 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. $\square$

Proposition 5.1. *If $\overset{\vartriangle}{\precsim}$ is weakly-SF-risk-averse with respect to some independent partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$, then it is also weakly-SF-risk-averse with respect to any independent partition. Similarly for (strong-)SF-risk-aversion.*

Proof. If there is only one independent partition then there is nothing to prove. Otherwise, by Lemma A.7 there exists a Debreu value function representing $\precsim$. Note that aggregate value function is identical for the two partitions.

Suppose that $\overset{\vartriangle}{\precsim}$ is SF-risk-averse with respect to one partition. Then, by Theorem 5, u is concave with respect to v. So, again, by the same theorem, $\overset{\vartriangle}{\precsim}$ is SF-risk-averse with respect to any other partition. Similarly for weak-SF-risk-aversion, Sf-risk-loving and weak-SF-risk-loving, and SF-risk-neutrality. $\square$

Theorem 6. *For $\overset{\vartriangle}{\precsim}$ with value-scaled-utility $\hat{u}$, $\overset{\vartriangle}{\precsim}$ is weakly-SF-risk-loving if and only if $\hat{u}$ is convex, and (strongly-)-SF-risk-loving if and only if $\hat{u}$ is strictly-convex.*

$\overset{\vartriangle}{\precsim}$ is SF-risk-neutral by Definition 6 if and only if $\hat{u}$ is linear.

Proof. The proof for risk-loving is analogues to to that of risk-aversion. Risk neutrality then follows from the weak versions of risk-aversion and risk-loving. $\square$

Proofs for Section 6.

Proposition 6.1. *In the multi-commodity setting (with $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ an independent partition), the preference order $\overset{\vartriangle}{\precsim}$ exhibits SF-DARA if and only if for any factor $\mathcal{T}_i$, lottery L_i over $\mathcal{T}_i$, state $c_i \in \mathcal{T}_i$, and $\boldsymbol{d}_{-i}, \boldsymbol{d}'_{-i} \in \Omega_{-\{i\}}$, if $\boldsymbol{d}_{-i} \prec \boldsymbol{d}'_{-i}$ then*

$$(c_i, \boldsymbol{d}_{-i}) \overset{\vartriangle}{\precsim} (L_i, \boldsymbol{d}_{-i}) \Rightarrow (c_i, \boldsymbol{d}'_{-i}) \overset{\vartriangle}{\prec} (L_i, \boldsymbol{d}'_{-i}).$$

Similarly, $\overset{\vartriangle}{\succsim}$ exhibits SF-CARA if and only if for any commodity i, lottery L_i over $\mathcal{T}_i$, state $c_i \in \mathcal{T}_i$, and $\boldsymbol{d}_{-i}, \boldsymbol{d'}_{-i} \in \Omega_{-\{i\}}$

$$(c_i, \boldsymbol{d}_{-i}) \overset{\vartriangle}{\precsim} (L_i, \boldsymbol{d}_{-i}) \iff (c_i, \boldsymbol{d'}_{-i}) \overset{\vartriangle}{\precsim} (L_i, \boldsymbol{d'}_{-i})$$

Proof. Note that, technically, all claims and proofs in Pratt work [28], hold with respect to any scale one chooses to use. Now, using the v scale, the proposition follows directly from Sections 6-7 and Theorem 2 of [28]. $\square$

Theorem 7. *If preference order $\overset{\vartriangle}{\succsim}$ is (weakly) SF risk averse then it is (weakly) correlation averse with respect to all pairs of factors $\mathcal{T}_i, \mathcal{T}_j$. Conversely, if $\overset{\vartriangle}{\succsim}$ is correlation averse with respect to some pair of factors $\mathcal{T}_i, \mathcal{T}_j$, then it is (weakly) SF risk averse*

Proof. We prove the case for weak-SF-risk-aversion and weak correlation aversion. The proof for strict-SF-risk-aversion and strict correlation aversion is similar.

We first prove the theorem for the case where there exists a Debreu value function, v, representing $\precsim$ on the partition. In this case, the theorem is essentially a direct corollary of Theorem 4(a) of Epstein and Tanny[13], which states (using the notations of our paper):

> Let $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$, be an independent partition, and $\overset{\vartriangle}{\succsim}^{\mathcal{X}}$ 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 $\overset{\vartriangle}{\succsim}^{\mathcal{X}}$ is weakly correlation averse by 3 if and only if ϕ is concave.

Set $\phi = u \circ v^{-1}$. So, $u(a_1, \ldots, a_n) = \phi(\sum_{i=1}^{n} v^{\mathcal{T}_i}(a_i)))$.

Now, if $\overset{\vartriangle}{\succsim}$ is weakly-SF-risk-averse then u is concave with respect to v (Theorem 5). So, ϕ is concave. So, setting $\mathcal{X}_1 = \mathcal{T}_1$, and $\mathcal{X}_2 = \mathcal{T}_2 \times \cdots \times \mathcal{T}_n$, by the above Epstein and Tanny theorem, $\overset{\vartriangle}{\succsim}$ is weakly correlation averse with respect to $\mathcal{X}_1, \mathcal{X}_2$. In particular, for $a_1, b_1 \in \mathcal{T}_1$, $a_1 \prec b_1$, and $(a_2, \boldsymbol{c}), (b_2, \boldsymbol{c}) \in \mathcal{X}_2 = \mathcal{T}_2 \times \cdots \times \mathcal{T}_n$, with $(a_2, \boldsymbol{c}) \prec (b_2, \boldsymbol{c})$,

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

Since $\mathcal{T}_2$ is independent, $(a_2, \boldsymbol{c}) \prec (b_2, \boldsymbol{c})$ if and only if $a_2 \prec b_2$. So, $\overset{\vartriangle}{\succsim}$ is weakly correlation averse with respect to $\mathcal{T}_1, \mathcal{T}_2$. The same argument works for proving that $\overset{\vartriangle}{\succsim}$ is weakly correlation averse with respect to any $\mathcal{T}_i, \mathcal{T}_j$.

Conversely, suppose that $\overset{\vartriangle}{\succsim}$ is weakly correlation averse with respect to some $\mathcal{T}_i, \mathcal{T}_j$. For any fixed $\boldsymbol{c} \in \mathcal{S}_{-1,2}$, set $\mathcal{S}^{\boldsymbol{c}} = \mathcal{T}_i \times \mathcal{T}_2 \times \{\boldsymbol{c}\}$. Set $\mathcal{X}_1^{\boldsymbol{c}} = \mathcal{T}_i$ and $\mathcal{X}_1^{\boldsymbol{c}} = \mathcal{T}_j \times \{\boldsymbol{c}\}$. Then, $\overset{\vartriangle}{\succsim}$ is weakly correlation averse with respect to $\mathcal{X}_1^{\boldsymbol{c}}, \mathcal{X}_2^{\boldsymbol{c}}$. So, ϕ is concave on $v(\mathcal{S}^{\boldsymbol{c}})$. This holds for all $\boldsymbol{c}$. By continuity, each $v(\mathcal{S}^{\boldsymbol{c}})$ is an interval. Furthermore, the union of these intervals is $v(\mathcal{S})$, which is the entire domain of ϕ, and any point in the interior of $v(\mathcal{S})$ is in the interior of some $v(\mathcal{S}^{\boldsymbol{c}})$. So, ϕ is concave on its entire domain. So, $\overset{\vartriangle}{\succsim}$ is weakly-SF-risk-averse (Theorem 5).

It remains to prove the theorem for the case that the space is not additively separable. This can only happen when $n = 2$, as for $n \geq 3$ independence of the partition implies additive separability.

We first define a restricted notion of correlation aversion. For $\mathcal{S} = \mathcal{T}_1 \times \mathcal{T}_2$, and $\overset{\vartriangle}{\precsim}$ on $\Delta(\mathcal{S})$, we say that $\overset{\vartriangle}{\precsim}$ is *perfect weak correlation averse* if for any $a_1, b_1 \in \mathcal{T}_1$, $a_2, b_2 \in \mathcal{T}_2$, with $a_1 \overset{\vartriangle}{\prec} b_1$, $a_2 \overset{\vartriangle}{\prec} b_2$, and $(a_1, b_2) \sim (a_2, b_1)$

$$\langle (a_1, a_2), (b_1, b_2) \rangle \overset{\vartriangle}{\precsim} \langle (a_1, b_2), (b_1, a_2) \rangle . \tag{29}$$

The difference is that with perfect correlation aversion, (29) is only required when $(a_1, b_2) \sim (a_2, b_1)$. Lemma A.10, which we prove later, establishes that $\overset{\vartriangle}{\precsim}$ is (weak) correlation averse if and only if it is perfect (weak) correlation averse. So it suffices to prove the theorem for perfect (weak) correlation aversion.

Consider the $a_1, b_1 \in \mathcal{T}_1$, $a_2, b_2 \in \mathcal{T}_2$, with $a_1 \overset{\vartriangle}{\prec} b_1$, $a_2 \overset{\vartriangle}{\prec} b_2$, and $(a_1, b_2) \sim (a_2, b_1)$. Set $L_1 = \langle a_1, b_1 \rangle, L_2 = \langle a_2, b_2 \rangle$, and consider $\boldsymbol{L} = (L_1, L_2)$. So, starting from the definition of weak-SF-risk-aversion

$$\begin{aligned}
\Pr[\boldsymbol{c}(\boldsymbol{L}) \precsim \boldsymbol{\ell}] &\geq \Pr[\boldsymbol{c}(\boldsymbol{L}) \succsim \boldsymbol{\ell}] \iff \\
\Pr[u(\boldsymbol{L}) \leq u(\boldsymbol{\ell})] &\geq \Pr[u(\boldsymbol{L}) \geq u(\boldsymbol{\ell})] \iff \\
\Pr[u(\boldsymbol{L}) \leq u(\boldsymbol{\ell})] &\geq 1/2
\end{aligned} \tag{30}$$

The four possible realizations of $\boldsymbol{L}$ are

$$(a_1, a_2) \prec (a_1, b_2) \sim (b_1, a_2) \prec (b_1, b_2).$$

So, continuing (30),

$$\begin{aligned}
\Pr[u(\boldsymbol{L}) \leq u(\boldsymbol{\ell})] &\geq 1/2 \iff \\
u(\boldsymbol{L}) &\leq u(a_1, b_1) \iff \\
\frac{u(a_1, a_2) + u(a_1, b_2) + u(b_1, a_2) + u(b_1, b_2)}{4} &\leq u(a_1, b_1) \iff \qquad \text{(since } (a_1, b_1) \sim (b_1, a_2)\text{)}, \\
\frac{u(a_1, a_2) + u(b_1, b_2)}{2} &\leq \frac{u(a_1, b_1) + u(b_1, a_1)}{2}
\end{aligned}$$

establishing (29). $\square$

We now turn to proving that correlation aversion and perfect correlation aversion are equivalent, focusing on the case of a partition into two factors. Again, if the space is additively separable, the proof is easy. The more involved case is when it is not additively separable.

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, and and b, B, with or without subscripts or superscripts, for points in $\mathcal{A}$ and $\mathcal{B}$, respectively. By convention, $a \prec A$ and $b \prec B$.

Let $w^{\mathcal{A}} : \mathcal{A} \to \mathbb{R}$ be a continuous real valued 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.8. *$u \circ \boldsymbol{w}^{-1} : I_{\mathcal{A}} \times I_{\mathcal{B}} \to \mathbb{R}$ 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

$$0 < \hat{u}^{\mathcal{B}}_{x^*}(y') - \hat{u}^{\mathcal{B}}_{x^*}(y^*) < \frac{1}{2}\epsilon. \tag{31}$$

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

$$0 < \hat{u}^{\mathcal{A}}_{y'}(x') - \hat{u}^{\mathcal{A}}_{y'}(x^*) < \frac{1}{2}\epsilon. \tag{32}$$

Combining (31) and (32), 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.9. *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. □

Lemma A.10. *$\overset{\Delta}{\precsim}$ is (weakly) correlation averse if and only if it is perfect (weakly) correlation averse.*

Proof. (weak) correlation aversion $\Rightarrow$ perfect (weak) correlation aversion: The requirement of perfect risk aversion - (29) - is identical to that of correlation aversion - (3) - only limited to the case that $(a_1, b_2) \sim (b_2, a_1)$.

Perfect (weak) correlation aversion $\Rightarrow$ (weak) correlation aversion: We prove for the strong case. The weak case is similar.

Suppose that $\stackrel{\vartriangle}{\precsim}$ is perfect correlation averse on $\mathcal{S} = \mathcal{A} \times \mathcal{B}$. Let $a, A \in \mathcal{A}, b, B \in \mathcal{B}$, with $a \prec A$ and $b \prec B$. We need to show that

$$\langle (a, b), (A, B) \rangle \stackrel{\vartriangle}{\prec} \langle (A, b), (a, B) \rangle . \tag{33}$$

If $(a, B) \sim (A, b)$ then (33) holds by the definition of perfect correlation aversions.

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

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

We show that diff < 0, which establishes (33).

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

$$\begin{aligned} & w^{\mathcal{A}}(A_{\frac{1}{2}}) - w^{\mathcal{A}}(a_{\frac{1}{2}}) \leq \frac{1}{2}(w^{\mathcal{A}}(A) - w^{\mathcal{A}}(a)) \quad \text{or} \\ & w^{\mathcal{B}}(B_{\frac{1}{2}}) - w^{\mathcal{B}}(b_{\frac{1}{2}}) \leq \frac{1}{2}(w^{\mathcal{B}}(B) - w^{\mathcal{B}}(b)) \end{aligned} \tag{34}$$

and

$$\mathrm{diff} < u(a_{\frac{1}{2}}, b_{\frac{1}{2}}) + u(A_{\frac{1}{2}}, B_{\frac{1}{2}}) - u(a_{\frac{1}{2}}, B_{\frac{1}{2}}) - u(A_{\frac{1}{2}}, b_{\frac{1}{2}}). \tag{35}$$

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

$$(a^1, b) \sim (a, B). \tag{36}$$

Figure 2 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.9. 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

$$w^{\mathcal{A}}(A^1) \leq \frac{1}{2}(w^{\mathcal{A}}(a) + w^{\mathcal{A}}(A)), \tag{37}$$

or

$$w^{\mathcal{A}}(a^{\bar{i}}) \geq \frac{1}{2}(w^{\mathcal{A}}(a) + w^{\mathcal{A}}(A)). \tag{38}$$

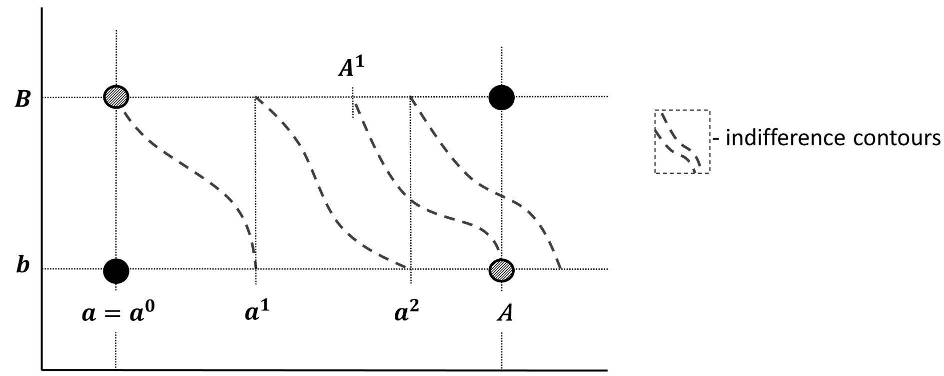

FIGURE 2. Illustration of the proof of Lemma A.10. The values a^i are calculated left-to-right, starting at $a = a^0$. Here $\bar{i} = 2$ and the point a^2 is such that $w^{\mathcal{A}}(a^2) \geq \frac{1}{2}(w^{\mathcal{A}}(A) + w^{\mathcal{A}}(a))$ (assuming the picture is scaled according to $w^{\mathcal{A}}$).

We consider each of these cases separately.

First, suppose that (37) 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,

$$\begin{aligned} & u(a, b) + u(A, B) - u(A, b) - u(a, B) = \\ & u(a, b) + u(A^1, B) - u(A^1, b) - u(a, B) + u(A^1, b) + u(A, B) - u(A^1, B) - u(A, b) < \\ & u(a, b) + u(A^1, B) - u(A^1, b) - u(a, B). \end{aligned} \tag{39}$$

Setting $a_{\frac{1}{2}} = a$, $A_{\frac{1}{2}} = A^1$, $b_{\frac{1}{2}} = b$ and $B_{\frac{1}{2}} = B$, by (37) and (39) we get (34) and (35).

Next, suppose that (38) 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 $\overset{\vartriangle}{\precsim}$ is ordinally risk averse,

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

for all i. So,

$$\frac{1}{2\bar{i}} \sum_{i=1}^{\bar{i}} \left(u(a^{i-1}, b) + u(a^i, B) \right) < \frac{1}{2\bar{i}} \sum_{i=1}^{\bar{i}} \left(u(a^{i-1}, B) + u(a^i, b) \right) ; \tag{40}$$

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,

$$\begin{aligned}
&u(a,b) + u(A,B) - u(A,b) - u(a,B) = \\
&u(a,b) + u(a^{\bar{i}}, B) - u(a^{\bar{i}}, b) - u(a,B) + u(a^{\bar{i}}, b) + u(A,B) - u(a^{\bar{i}}, B) - u(A,b) < \\
&u(a^{\bar{i}}, b) + u(A,B) - u(a^{\bar{i}}, B) - u(A,b). \qquad (41)
\end{aligned}$$

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 (38) and (41) we get (34) and (35).

Thus, we have established (34) and (35), and we now return to complete the proof that diff < 0. Set

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

Then,

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

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

$$w^{\mathcal{A}}(A_\delta) - w^{\mathcal{A}}(a_\delta) \le \delta \quad \text{or} \qquad (42)$$

$$w^{\mathcal{B}}(B_\delta) - w^{\mathcal{B}}(b_\delta) \le \delta \qquad (43)$$

and

$$\begin{aligned}
\text{diff}_{\frac{1}{2}} &< u(a_\delta, b_\delta) + u(A_\delta, B_\delta) - u(a_\delta, B_\delta) - u(A_\delta, b_\delta) = \\
&(u(A_\delta, B_\delta) - u(a_\delta, B_\delta)) + (u(a_\delta, b_\delta) - u(A_\delta, b_\delta)) = \qquad (44)\\
&(u(A_\delta, B_\delta) - u(A_\delta, b_\delta)) + (u(a_\delta, b_\delta) - u(a_\delta, B_\delta))\, . \qquad (45)
\end{aligned}$$

By Lemma A.8 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 (42) holds then (44) is $\le 2\epsilon$, and if (43) holds then (45) is $\le 2\epsilon$. Thus, $\text{diff}_{\frac{1}{2}} \le 0$, so diff < 0. □

Proofs for Section 7. NOTE: the proofs in this section require some additional work.

Let $z_1 < z_2 < \cdots < z_n$ be the support of $\tilde{z}$, and $p_i = \Pr[\tilde{z} = z_i]$.

For a given budget w, and savings level s, the expected utility is

$$u(w-s, \tilde{z}s) = \sum_{i=1}^{n} p_i u(w-s, z_i s) = \sum_{i=1}^{n} p_i \hat{u}(v_1(w-s) + v_2(z_i s)).$$

Differentiating by s,

$$\begin{aligned} \frac{\partial u(w-s, \tilde{z}s)}{\partial s} &= \sum_{i=1}^{n} p_i \frac{\partial \hat{u}(v_1(w-s) + v_2(z_i s))}{\partial s} \\ &= \sum_{i=1}^{n} p_i \cdot (-v_1'(w-s) + z_i v_2'(z_i s)) \cdot \hat{u}'(v_1(w-s) + v_2(z_i s)), \end{aligned} \tag{46}$$

(where f' is the derivative of f). The optimal $s^*(w)$ is obtained either at the boundaries, or when $\frac{\partial u(w-s,\tilde{z}s)}{\partial s} = 0$.

Lemma A.11. *If $v_1, v_2, \hat{u}$ are concave, and either: (i) v_1 or v_2 strictly concave, or (ii) $\hat{u}$ strictly concave and $\tilde{z} \neq v_1'/v_2'$, then:*

- *there exists at most one $s = s^*$ for which $\frac{\partial u(w-s,\tilde{z}s)}{\partial s} = 0$.*
- *$\frac{\partial u(w-s,\tilde{z}s)}{\partial s} > 0$, for $s < s^*$.*
- *$\frac{\partial u(w-s,\tilde{z}s)}{\partial s} < 0$, for $s > s^*$.*

Proof. For any z, set $h(s,z) = (-v_1'(w-s) + z v_2'(zs)) \cdot \hat{u}'(v_1(w-s) + v_2(zs))$. Then, $\frac{\partial u(w-s,\tilde{z}s)}{\partial s} = \sum_{i=1}^{n} p_i h(s, z_i)$. We prove that for any z, $h(s,z)$ is decreasing in s, and strictly decreasing if either (i) v_1 or v_2 strictly concave, or (ii) $\hat{u}$ strictly concave and $\tilde{z} \neq v_1'/v_2'$. So, $\frac{\partial u(w-s,\tilde{z}s)}{\partial s}$ is a sum of decreasing functions, at least one of which is strictly decreasing, which establishes the lemma.

Consider $h(s,z) = (-v_1'(w-s) + z v_2'(zs)) \cdot \hat{u}'(v_1(w-s) + v_2(zs))$. Since v_1 and v_2 are concave, $(-v_1'(w-s) + z v_2'(zs))$ decreases with s. Also note that $\hat{u}' > 0$, as $\hat{u}$ is increasing. So, there are three possible zones for $h(s,z)$:

- $(-v_1'(w-s) + z v_2'(zs)) > 0$: in this zone $v_1(w-s) + v_2(zs)$ - the argument of $\hat{u}$ - increases with s. So, since $\hat{u}$ is concave, $\hat{u}'(v_1(w-s)+v_2(zs))$ decreases with s. So, in all, in this zone $h(s,z)$ decreases with s. If one of: $v_1, v_2, \hat{u}$ is strictly concave then $h(s,z)$ strictly decreases.
- $(-v_1'(w-s) + z v_2'(zs)) < 0$: in this zone $v_1(w-s) + v_2(zs)$ - the argument of $\hat{u}$ - decreases with s. So, since $\hat{u}$ is concave, $\hat{u}'(v_1(w-s) + v_2(zs))$ increases with s. So, in all, in this zone $h(s,z)$ decreases with s, and if one of: $v_1, v_2, \hat{u}$ is strictly concave then $h(s,z)$ strictly decreases.
- $(-v_1'(w-s) + z v_2'(zs)) = 0$: in this zone $h(s,z) = 0$. This consists of at most one point unless both v_1 and v_2 are linear and $z = v_1'/v_2'$.

So, $h(s,z)$ decreases with s throughout, if either (i) v_1 or v_2 strictly concave, or (ii) $\hat{u}$ strictly concave and $\tilde{z} \neq v_1'/v_2'$. □

The following is simple generalization of Chebyshev's sum inequality.

Lemma A.12. *Let* $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n$*, and* $\beta_1 \leq \beta_2 \leq \cdots \leq \beta_n$*, be sequences of real numbers, and* $y_1, \ldots, y_n$ *a sequence of positive numbers. Then,*

$$\left(\sum_{i=1}^{n} \alpha_i \beta_i y_i\right)\left(\sum_{i=1}^{n} y_i\right) \geq \left(\sum_{i=1}^{n} \alpha_i y_i\right)\left(\sum_{i=1}^{n} \beta_i y_i\right), \tag{47}$$

and equality holds if and only if either all the α_i*'s or all the* β_j*'s are identical.*

Proof. For any i, j, $(\alpha_i - \alpha_j)$ and $(\beta_i - \beta_j)$ have (weakly) the same signs (that is, if one is positive the other cannot be negative). So,

$$\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j (\alpha_i - \alpha_j)(\beta_i - \beta_j) \geq 0,$$

and equality holds if and only if either all the α_i's or all the β_j's are identical. Opening the brackets we get,

$$2\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j \alpha_i \beta_i - 2\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j \alpha_i \beta_j \geq 0,$$

which, after rearrangement gives (47). □

Proposition 7.1. *If* $v_2'(0) \cdot E[\tilde{z}] > v_1'(w)$ *then* $s^*(w) > 0$.

Proof. At $s = 0$, we have

$$\begin{aligned}
\left.\frac{\partial u(w-s, \tilde{z}s)}{\partial s}\right|_{s=0} &= \sum_{i=1}^{n} p_i \cdot (-v_1'(w) + z_i v_2'(0)) \cdot \hat{u}'(v_1(w) + v_2(0)) \\
&= \left(-v_1'(w) + v_2'(0)\sum_{i=1}^{n} p_i z_i\right) \cdot \hat{u}'(v_1(w) + v_2(0)) \\
&= \left(-v_1'(w) + v_2'(0)E[\tilde{z}]\right) \cdot \hat{u}'(v_1(w) + v_2(0)).
\end{aligned}$$

since $\hat{u}'$ is always positive, the above is positive whenever $v_2'(0)E[\tilde{z}] > v_1'(w)$. If this is the case, the expected utility increases with s at $s = 0$, so, s^* cannot be zero. □

Theorem 8. *Provided that*

- $v_1, v_2, \hat{u}$ *are concave,*
- $v_2 \circ \exp$ *is convex,*

if $\hat{u}$ *is DARA, then* $s^*(w_-) \leq s^*(w_+)$ *whenever* $w_- < w_+$*. If either* v_1 *is strictly concave or* $v_2 \circ \exp$ *strictly convex then* $s^*(w_-) < s^*(w_+)$ *(provided that* $0 < s^*(w_-) > 0$ *and* $s^*(w_+) < w_+$*).*

Proof. We prove for the case that v_1 is strictly concave. The proof for the other cases case that $v_2 \circ \exp$ is strictly convex is similar.

For any w, $s^*(w)$ is obtained either at the boundaries, 0 and w, or when $\frac{\partial u(w-s,\tilde{z}s)}{\partial s}=0$. Consider the latter case. In this case, by Lemma A.11 for $s \le s^*(w)$,

$$\sum_{i=1}^{n} p_i \cdot (-v_1'(w-s) + z_i v_2'(z_i s)) \cdot \hat{u}'(v_1(w-s)+v_2(z_i s)) \ge 0.$$

Denoting $\hat{u}'_{w,s,i} = \hat{u}'(v_1(w-s)+v_2(z_i s)$, and some rearranging gives

$$\frac{\sum_{i=1}^{n} p_i z_i v_2'(z_i s)\hat{u}'_{w,s,i}}{\sum_{i=1}^{n} p_i \cdot \hat{u}'_{w,s,i}} \ge v_1'(w-s), \tag{48}$$

($\hat{u}$ is increasing, so the denominator is positive, and it was ok to divide by it). Since v_1 is strictly concave, the right-hand side of (48) decreases with w. We shortly show that left-hand side (weakly) *increases* with w. So, for any $s \le s^*$, equality between the two sides of (48) cannot hold for $w_+ > w$. So, it must be that $s^*(w_+) > s^*(w)$.

It remains to show that the left-hand side of (48) increases with w. Taking the derivative according to w, we have

$$\frac{\partial\left(\frac{\sum_{i=1}^{n} p_i z_i v_2'(z_i s)\hat{u}'_{w,s,i}}{\sum_{i=1}^{n} p_i \hat{u}'_{w,s,i}}\right)}{\partial w} =$$

$$\frac{v_1'(w-s)\cdot\left(\left(\sum_{i=1}^{n} p_i z_i v_2'(z_i s)\hat{u}''_{w,s,i}\right)\left(\sum_{i=1}^{n} p_i \hat{u}'_{w,s,i}\right) - \left(\sum_{i=1}^{n} p_i z_i v_2'(z_i s)\hat{u}'_{w,s,i}\right)\left(\sum_{i=1}^{n} p_i \hat{u}''_{w,s,i}\right)\right)}{(\sum_{i=1}^{n} p_i \hat{u}'_{w,s,i})^2} =$$

(49)
$$\frac{v_1'(w-s)\cdot\left(\left(\sum_{i=1}^{n} p_i z_i v_2'(z_i s)\frac{\hat{u}''_{w,s,i}}{\hat{u}'_{w,s,i}}\hat{u}'_{w,s,i}\right)\left(\sum_{i=1}^{n} p_i \hat{u}'_{w,s,i}\right) - \left(\sum_{i=1}^{n} p_i z_i v_2'(z_i s)\hat{u}'_{w,s,i}\right)\left(\sum_{i=1}^{n} p_i \frac{\hat{u}''_{w,s,i}}{\hat{u}'_{w,s,i}}\hat{u}'_{w,s,i}\right)\right)}{(\sum_{i=1}^{n} p_i \hat{u}'_{w,s,i})^2}.$$

The denominator of (49) is always positive, as is $v_i'(w-s)$. So, it remains to show that

$$\left(\sum_{i=1}^{n} p_i z_i v_2'(z_i s)\frac{\hat{u}''_{w,s,i}}{\hat{u}'_{w,s,i}}\hat{u}'_{w,s,i}\right)\left(\sum_{i=1}^{n} p_i \hat{u}'_{w,s,i}\right) - \left(\sum_{i=1}^{n} p_i z_i v_2'(z_i s)\hat{u}'_{w,s,i}\right)\left(\sum_{i=1}^{n} p_i \frac{\hat{u}''_{w,s,i}}{\hat{u}'_{x,s,i}}\hat{u}'_{w,s,i}\right) \tag{50}$$

is positive.

Denote:

- $\alpha_i = \frac{\hat{u}''_{w,s,i}}{\hat{u}'_{w,s,i}}$,
- $\beta_i = z_i v_2'(z_i s)$,
- $y_i = p_i \hat{u}'_{w,s,i}$.

With this notation, (50) is

$$\left(\sum_{i=1}^{n} \alpha_i \beta_i y_i\right)\left(\sum_{i=1}^{n} y_i\right) - \left(\sum_{i=1}^{n} \beta_i y_i\right)\left(\sum_{i=1}^{n} \alpha_i y_i\right). \tag{51}$$

Now, since $z_i \le z_{i+1}$, v_2 increasing, and $\hat{u}$ DARA,

$$\alpha_i = \frac{\hat{u}''_{w,s,i}}{\hat{u}'_{w,s,i}} = \frac{\hat{u}''(v_1(w-s)+v_2(z_i S))}{\hat{u}'(v_1(w-s)+v_2(z_i S))} \le \frac{\hat{u}''(v_1(w-s)+v_2(z_{i+1} S))}{\hat{u}'(v_1(w-s)+v_2(z_{i+1} S))} = \frac{\hat{u}''_{w,s,i+1}}{\hat{u}'_{w,s,i+1}} = \alpha_{i+1}.$$

Since $v_2 \circ \exp$ is convex, $(v_2 \circ \exp)'$ is increasing. That is, $(v_2(e^y))' = e^y \cdot v_2'(e^y)$ is increasing. So, for, $y = \ln(zs)$, $zsv_2(zs)$ is increasing in z (since $\ln(zs)$ increases in z). Hence, $zv_2(zs)$ increases with z.

$$\beta_i = z_i v_2'(z_i s) \leq z_{i+1} v_2'(z_{i+1} s) = \beta_{i+1}.$$

Finally, since $\hat{u}$ is increasing, y_i is positive. So, the conditions of Lemma A.12 hold, and (51) is positive. □

Proposition 7.2. *Assuming that v_1, v_2, and $\hat{u}$ are concave, then $v_2 \circ \exp$ is (strictly) convex (concave) if and only if for any w, $s^*(w, z)$ (res. strictly) increases (res. decreases) in z (provided that $0 < s^*(w, z) < w$).*

Proof. We prove for (weak) convexity. The other cases are similar.

Since $0 < s^*(w, z) < w$, either v_1 or v_2 are strictly concave. By Lemma A.11, for $s \leq s^*(w, z)$

$$\frac{\partial u(w - s, zs)}{\partial s} = (-v_1'(w - s) + zv_2'(zs)) \cdot \hat{u}'(v_1(w - s) + v_2(zs)) \geq 0, \tag{52}$$

and equality holds only at $s = s^*(w, z)$. Since $\hat{u}$ is increasing, $\hat{u}'$ is positive, so (52) is equivalent to

$$zv_2'(zs) \geq v_1'(w - s). \tag{53}$$

Now, the right-hand side is independent of z, and the left-hand side increases with z if and only if $zv_2'(zs)$ increases with z, which we have shown is equivalent to convexity of $v_2 \circ \exp$. So, (52) continues to hold as z increases if and only if $v_2 \circ \exp$ is convex. □

Theorem 9. *Provided that $v_1, v_2, \hat{u}$ are concave, $\hat{u}_+$ is more risk averse than $\hat{u}$ then:*

- *$s^*(w) \geq s^*{}_+(w)$ for any w if$v_2 \circ \exp$ is convex, and $s^*(w) > s^*{}_+(w)$ if $v_2 \circ \exp$ is strictly convex (provided that $0 < s^*(w) < w$).*
- *$s^*(w) \leq s^*{}_+(w)$ for any w if$v_2 \circ \exp$ is concave, and $s^*(w) < s^*{}_+(w)$ if $v_2 \circ \exp$ is strictly concave (provided that $0 < s^*(w) < w$).*

Proof. We prove for the case that $v_2 \circ \exp$ is strictly convex. The proof for the other cases is similar.

By Lemma A.11,and the analysis of in the beginning of the proof of Theorem 8, to prove that $s^*{}_+(w) < s^*(w)$ it suffices to prove that for any s,

$$\frac{\sum_{i=1}^n p_i z_i v_2'(z_i s) \hat{u}'(v_1(w - s) + v_2(z_i s))}{\sum_{i=1}^n p_i \cdot \hat{u}'(v_1(w - s) + v_2(z_i s))} > \frac{\sum_{i=1}^n p_i z_i v_2'(z_i s) \hat{u}_+'(v_1(w - s) + v_2(z_i s))}{\sum_{i=1}^n p_i \cdot \hat{u}_+'(v_1(w - s) + v_2(z_i s))}. \tag{54}$$

Denoting $\hat{u}_i' = \hat{u}'(v_1(w - s) + v_2(z_i s))$, $\hat{u}_{+,i}' = \hat{u}_+'(v_1(w - s) + v_2(z_i s))$, and some rearrangement, (54) gives

$$\left(\sum_{i=1}^n p_i z_i v_2'(z_i s) \hat{u}_i'\right) \left(\sum_{i=1}^n p_i \hat{u}_{+,i}'\right) > \left(\sum_{i=1}^n p_i z_i v_2'(z_i s) \hat{u}_{+,i}'\right) \left(\sum_{i=1}^n p_i \cdot \hat{u}_i'\right) \iff$$

$$\left(\sum_{i=1}^n p_i z_i v_2'(z_i s) \frac{\hat{u}_i'}{\hat{u}_{+,i}'} \hat{u}_{+,i}'\right) \left(\sum_{i=1}^n p_i \hat{u}_{+,i}'\right) > \left(\sum_{i=1}^n p_i z_i v_2'(z_i s) \hat{u}_{+,i}'\right) \left(\sum_{i=1}^n p_i \frac{\hat{u}_i'}{\hat{u}_{+,i}'} \hat{u}_{+,i}'\right). \tag{55}$$

Denote:

- $\alpha_i = \frac{\hat{u}'_i}{\hat{u}'_{+,i}}$,
- $\beta_i = z_i v'_2(z_i s)$,
- $y_i = p_i \hat{u}'_{+,i}$.

With this notation, (55) becomes

$$\left(\sum_{i=1}^{n} \alpha_i \beta_i y_i\right)\left(\sum_{i=1}^{n} y_i\right) > \left(\sum_{i=1}^{n} \beta_i y_i\right)\left(\sum_{i=1}^{n} \alpha_i y_i\right). \tag{56}$$

The fact that the sequence of β_i's is increasing was shown in the proof of Theorem 8. All y_i's are positive since $\hat{u}'_+$ is increasing.

By assumption $A_{\hat{u}_+}(x) > A_{\hat{u}}(x)$ for all x. So, $\frac{\hat{u}''(x)}{\hat{u}'(x)} > \frac{\hat{u}''_+(x)}{\hat{u}'_+(x)}$, for all x. Denote $x_i = v_1(w-s) + v_2(z_i s)$. Then,

$$\begin{aligned}
\ln(\alpha_{i+1}) =& \ln(\hat{u}'_{i+1}) - \ln(\hat{u}'_{+,i+1}) \\
=& \big(\ln(\hat{u}'_{i+1}) - \ln(\hat{u}'_{+,i+1}) - (\ln(\hat{u}'_i) - \ln(\hat{u}'_{+,i}))\big) + (\ln(\hat{u}'_i) - \ln(\hat{u}'_{+,i}) \\
=& \left(\int_{x_i}^{x_{i+1}} \big((\ln(\hat{u}'))'(x) - (\ln(\hat{u}'_+))'(x)\big)\, dx\right) + (\ln(\hat{u}'_i) - \ln(\hat{u}'_{+,i}) \\
>& ln(\hat{u}'_i) - \ln(\hat{u}'_{+,i}) = \ln(\alpha_i).
\end{aligned}$$

So, $\alpha_i < \alpha_{i+1}$, and the conditions of Lemma A.12 hold, which establishes 56. □

Proofs for Section 8. Proposition 8.2 was established in the main text.

Proposition 8.3. *Assuming A1-A2. Given: a time t, survival vector $\boldsymbol{s}$ (with $s_k \neq 0$, $k \geq t$), and better-than-death future $c_{t+1}, \ldots, c_n$, for any past-to-present sequences $(a_1, \ldots, a_t) \prec (b_1, \ldots, b_t)$*

- *if each $\stackrel{\triangleright}{\precsim}^*$ is SF-DARA then*

$$UD_{\boldsymbol{s}}^{t,t+1}(a_1, \ldots, a_t, c_{t+1}, \ldots, c_n) \geq UD_{\boldsymbol{s}}^{t,t+1}(b_1, \ldots, b_t, c_{t+1}, \ldots, c_n),$$

  and the inequality is strict if $\stackrel{\triangleright}{\precsim}^*$ is strictly-SF-DARA.
- *if $\stackrel{\triangleright}{\precsim}^*$ is SF-CARA then*

$$UD_{\boldsymbol{s}}^{t,t+1}(a_1, \ldots, a_t, c_{t+1}, \ldots, c_n) = UD_{\boldsymbol{s}}^{t,t+1}(b_1, \ldots, b_t, c_{t+1}, \ldots, b_n).$$

- *if $\stackrel{\triangleright}{\precsim}^*$ is SF-IARA then*

$$UD_{\boldsymbol{s}}^{t,t+1}(a_1, \ldots, a_t, c_{t+1}, \ldots, c_n) \leq UD_{\boldsymbol{s}}^{t,t+1}(b_1, \ldots, b_t, c_{t+1}, \ldots, b_n),$$

  and the inequality is strict if $\stackrel{\triangleright}{\precsim}^*$ is strictly-SF-IARA.

In all cases, the inequality is reversed in case the future is worse-than-death.

Proof. We prove for SF-DARA and a better-than-death future. The proof for the other cases is analogous. By definition of UD ((9))

$$
\begin{aligned}
UD_{\boldsymbol{s}}^{t,t+1}(a_1,\ldots,a_n) &= 1+\frac{s_t\cdot \hat{u}_t'(v_t(a_1,\ldots,a_t))}{\sum_{k=t+1}^{n} s_k\cdot \hat{u}_k'(v_k(a_1,\ldots,a_k))} \\
&= 1+\frac{s_t}{\sum_{k=t+1}^{n} s_k\cdot \frac{\hat{u}_k'(v_k(a_1,\ldots,a_k))}{\hat{u}_t'(v_k(a_1,\ldots,a_t))}}. \qquad (57)
\end{aligned}
$$

Set $x_t=v_t(a_1,\ldots,a_t)$, and for $k=t+1,\ldots,n$, set $y_k=\sum_{i=t+1}^{k} v^{\mathcal{T}_i}(a_i)$. Then $v_k=x_t+y_k$, for $k=t,\ldots,n$. So, for each k,

$$
\frac{\hat{u}_k'(v_k(a_1,\ldots,a_k))}{\hat{u}_t'(v_t(a_1,\ldots,a_t))}=\frac{\hat{u}_k'(x_t+y_k)}{\hat{u}_t'(x_t)}. \qquad (58)
$$

Differentiating by x_t, we have

$$
\frac{\partial\left(\frac{\hat{u}_k'(x_t+y_k)}{\hat{u}_t'(x_t)}\right)}{\partial x_t}=\frac{\hat{u}_k''(x_t+y_k)\cdot \hat{u}_t'(x_t)-\hat{u}_k'(x_t+y_k)\cdot \hat{u}_t''(x_t)}{(\hat{u}_t''(x_t))^2}. \qquad (59)
$$

Now, since $(a_{t+1},\ldots,a_n)$ is better-than-death, $(a_1,\ldots,a_t)\overset{\vartriangle}{\prec}^*(a_1,\ldots,a_k)$. So, since $\overset{\vartriangle}{\succsim}^*$ is DARA,

$$
-\left(\frac{\hat{u}_k''(x_t)}{\hat{u}_t'(x_t)}\right)\geq -\left(\frac{\hat{u}_k''(x_t+y_t)}{\hat{u}_t'(x_t+y_t)}\right).
$$

So, (59) is non-negative. So, each of the ratios in the denominator of (57) (weakly) increases with x_t, so the entire expression (weakly) decreases. □

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 $\overset{\vartriangle}{\succsim}=(\overset{\vartriangle}{\succsim}^1,\overset{\vartriangle}{\succsim}^2,\ldots)$, 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 $\overset{\vartriangle}{\precsim}^n$. 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 $c_1, c_2, \ldots$.