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

YONATAN AUMANN

Bar Ilan University
Ramat Gan, Israel

Abstract. The classic definition of risk aversion, which equates risk aversion with concavity of the utility function, is inherently scale-dependent, in the sense that it is not preserved under monotone (non-linear) transformations of underlying scale, most commonly taken to be money. This limits the notion to monetary, or liquid, goods. We introduce an axiomatic definition of risk aversion, based on the decision maker’s preference order alone, independent of any numerical scale. We then show that when cast in functional form this axiomatic definition coincides with the classic Arrow-Pratt definition once the latter is defined with respect to an appropriate “intrinsic” scale (which in general is not money). The applications of the theory are discussed, including, in particular, to disentangling risk aversion from diminishing marginal utility, to multi-dimensional risk aversion, and to the analysis of saving under uncertainty. The entire study is within the expected utility framework.

Keywords: Risk aversion, Decision theory, Multi-dimensional risk aversion, Saving under uncertainty, Diminishing marginal utility, Decreasing Risk Aversion.

E-mail address: aumann@cs.biu.ac.il.
Date: November 3, 2017.
† For the most updated version see https://sites.google.com/site/aumannbiu/home/publication-by-topic-1
1. Introduction

1.1. Risk Aversion - The Classic Definition. The concept of risk aversion is fundamental to economic theory. Classically, risk aversion is defined as a preference under which the certainty equivalent of a lottery is less than the lottery’s expected value. This definition is inherently scale-dependent; the notions of expectation and concavity are only defined with respect to a given scale. In particular, the definition is not preserved under monotone, non-linear transformation of the scale. So, a decision maker deemed risk averse using one scale, may be deemed risk loving by merely renaming the states.

This scale dependency brings about several issues. First, there is the matter of choosing the scale. In the seminal works of Arrow [2] and Pratt [27], 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 - health, joy, pain, duration of life, and the like.

The scale dependency of the classic definition is also an issue when moving to the multi-dimensional (i.e. multi-commodity/multi-period) setting. Here, each dimension may be equipped with its own scale, resulting in multiple “native” scales. Indeed, extending the theory of risk aversion to the multi-dimensional setting has proved challenging, and core problems remain open. In particular, as noted by Kihlstrom and Mirman [23]: “up to this point the theory of increasing, decreasing and constant absolute and relative risk aversion has not been satisfactorily extended to multidimensional utility functions.” To a great extent, the situation has not changed much since. In their paper, Kihlstrom and Mirman present a partial such extension, but it is limited in two respects: (a) it only applies to homothetic preferences, (b) it only extends the notion of increasing (decreasing/constant) relative risk aversion, not absolute risk aversion.

The difficulties in extending the notion of decreasing absolute risk aversion to the multi-dimensional setting have, in turn, hampered the analysis of the two-period saving-under-uncertainty problem, which is the multi-dimensional analogue of the portfolio selection problem analyzed by Arrow [2]. For the (uni-dimensional) portfolio selection case, Arrow proved that if the utility function exhibits DARA (decreasing absolute risk aversion) then investment in the risky asset increases with wealth. For the two-period case, no such result is known. In particular, to the best of our knowledge, the problem of establishing conditions under which savings increase with wealth has remained open to date.

Addressing these and other challenges regarding multi-dimensional risk aversion is a key goal of the present work.

The final question driving the current work is a conceptual one. Conceptually, the classic definition of risk aversion is based on an implicit presumption that the “natural value” of a lottery should be its expectation (if not for risk aversion). Hence, the difference between the expectation and the certainty equivalent is coined the “risk premium”. However, it is not exactly clear what is
the basis for this presumption. Evidently, it cannot rest on empirical evidence, as most people are risk averse, thus valuing a lottery at less than its expectation. So, the basis must be conceptual. But, on a conceptual level, it is not clear what reasoning dictates that a fair gamble$^1$ between, say, $100 and $200 “should” be worth $150. What makes the arithmetic mean the “appropriate” function to use? Providing a conceptual justification for basing the definition of risk aversion on the arithmetic mean is a key conceptual goal of this paper.

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 indifference curves. Having defined risk aversion in purely axiomatic terms, we then derive a functional form for these definitions. This functional form, we show, coincides with the classic Arrow-Pratt definition, once the latter is defined with respect to an appropriate, “intrinsic” scale. This scale, which in general is not money, applies to any type of good. We thus obtain a general framework for defining risk aversion for any type of good - monetary or not - and for determining a unified scale that can be used in the multi-dimensional setting. On a conceptual level, this provides a conceptual foundation for the classic definition, and for its reliance on the arithmetic mean.

**Axiomatic Definition I: 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 preferences, however, are ruled out by the von Neumann-Morgenstern (NM) axioms; the utility of a lottery must lie between the utilities of its possible outcomes. Interestingly, while such a situation 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 lotteries. We show that for some preference orders (agreeing with the NM axioms), repeatedly choosing the certainty equivalent of a lottery over the lottery itself results 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.

**Axiomatic Definition II: Finite Lottery 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

---

$^1$here and throughout, *fair gamble* means a gamble with two possible outcomes, each occurring with probability 0.5.
time periods (or alternatively, any other partition of the space into two or more separable factors). In the finite case, there is no (non-trivial) behavior “with probability 1”. So, we cannot require that choosing the certainty equivalent result in an inferior outcome with probability 1. Rather, in the finite case we define risk aversion as a preference wherein choosing the certainty equivalent most probably results in inferior outcome; that is, the probability that the realization is preferred to the certainty equivalent is greater than the probability that the certainty equivalent is preferred to the realization. The exact definition is provided in Section 5.

A Functional Form. Having established axiomatic definitions of risk-aversion, we show that these definitions can also be cast in functional form, using an appropriate scale. Such a scale, we show, is provided by the multi-attribute (additive) value function, pioneered by Debreu [8, 10], 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 structure of the 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 our axiomatic definitions of risk-aversion coincide with the Arrow-Pratt functional definition, once the latter is defined using the Debreu value function as the underlying scale. 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 axiomatic definitions.

1.3. Applications. Having defined risk aversion in axiomatic terms, and established its equivalent functional form, we apply this framework to address key issues in the theory of risk aversion, including those mentioned in the beginning of this introduction.

Non-monetary Goods. The framework offers a way to define risk aversion for non-monetary goods and goods with no natural scale, such as pain, pleasure, longevity. Indeed, in the definitions of this paper, externally defined scales (such as market value) do not play any role. The only scale of interest is the intrinsic Debreu value, which is determined by the structure of the indifference curves, which exist for monetary and non-monetary goods alike. We note, however, that our definition requires multi-dimensionality in the setting (e.g. at least two time periods or at least two commodities).

2 formally defined in Section 2
Disentangling Risk Aversion from Diminishing Marginal Utility. It is well known that under the classic, Arrow-Pratt definition of risk aversion, diminishing marginal utility and risk aversion are inseparable. 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 theory of non-expected utility\(^3\). We show that the approach offered in this paper provides a natural way to disentangle the two within the expected utility framework. We show that the curvature of the (indirect) NM utility function can be decomposed into two components: the curvature of the (indirect) Debreu value function, and the curvature of the NM utility function with respect to the Debreu value function. With this decomposition, the former, we show, is naturally associated with diminishing marginal utility, while the latter - with risk aversion. This decomposition is formally developed in Section 6.

Multi-dimensional Risk Aversion. The framework allows, for the first time, to extend the notions of decreasing/constant/increasing absolute risk aversion to the multi-dimensional setting. We show that with this extension the notions retain their characteristics, in the sense that the certainty equivalent of a lottery increases (/decreases/constant) with wealth according to whether the (multi-dimensional) utility function exhibits decreasing (/increasing/constant) risk aversion. We also consider a multi-dimensional analogue of the notion of a mean preserving spread, and prove that risk averse agents, as in our axiomatic definitions, are characterized by dislike of any such multi-dimensional mean-preserving spread. The details are developed in Section 7.

Saving Under Uncertainty. Consider a decision maker (DM) who needs to decide on how to split her funds between (sure) consumption “today”, and savings with a random return, for consumption “tomorrow”. As observed in [24, 23], this is the multi-dimensional analogue of the portfolio selection problem analyzed by Arrow [2], wherein the DM needs to split her budget between a risky asset and a non-risky one. As noted above, extending Arrow’s analysis to this two-period setting has proved challenging, in particular with respect to the wealth effects. Kihlstrom and Mirman [23] provide such an extension, but only for the case of homothetic preferences. We show that our framework allows for a much broader analysis, establishing conditions under which savings increase with wealth for general preferences. This analysis is carried out in Section 8.

1.4. Assumptions.

Multi-dimensionality. Our axiomatic definition rests upon a multi-dimensional structure of the state space; that is, the existence of more than one time period, or more than one commodity. We do note, however, that once the underlying state space is multi-dimensional, the theory does have

\(^3\)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.”
bearings on the uni-dimensional indirect utility function with respect to money, as discussed in Sections 6 and 8.

Separability. Separability is a key notion and frequent assumption throughout this work. Simply put, a time period (or set of commodities) is separable if preferences over bundles in this time period (set of commodities) are not affected by the state in the other time periods (commodities). Arguably, separability is a strong assumption; having eaten Chinese food today may affect one’s gastronomical preferences tomorrow. Nonetheless, separability is a common assumption in economic literature, and in particular with respect to time preferences; e.g. the standard discounted-utility model implies separability of any time period (and any collection of the time periods). We use the separability assumption not because we believe it is a 100% accurate representation of reality, but rather because we believe that it is a good enough approximation, which allows us to concentrate on and formalize other key notions. We note that it suffices to assume separability of sufficiently long time periods.

We stress that separability, if and when assumed, is only so assumed for the certainty (a.k.a. ordinal) preferences. Nowhere do we assume separability of the lottery preferences (a.k.a. utility independence).

Expected Utility. This work is entirely within the classic expected-utility (EU) framework. Extending the ideas to non-EU models is an interesting future research direction.

1.5. Plan of the Paper. The paper is structured as follows. Immediately following, Section 2 describes the model, terminology and notation used throughout. The remainder of the paper is divided into two parts: Part A, which focuses on the axiomatic foundations, and part B, which focuses on applications. Part A is divided into three sections: Section 3 presents the definition based on infinite lottery sequences, with its functional form developed in Section 4 and Section 5, which presents the finite sequences definition, together with its functional form. Part B is also divided into three sections. Section 6 considers the disentanglement of risk aversion from diminishing marginal utility. Section 7 considers multi-dimensional risk aversion, including the multi-dimensional notions of CARA and DARA. Savings under uncertainty are considered in Section 8. Some related work is discussed in Section 9, and the main body of the paper ends with concluding remarks in Section 10. All proofs are deferred to an appendix. For simpler orientation, theorems, propositions and corollaries all share the same running numbering (definitions are numbered separately).

2. Model, Terminology and Notation

The Spaces. Preferences are defined over consumption spaces of the form \( S = C_1 \times \cdots \times C_m \), where each \( C_i \) is a separable and arc-connected topological space, representing the consumption space of a specific commodity (at a specific time period). We call each \( C_i \) a commodity space.

---

4A formal definition is provided in the next section.
**Lotteries.** All lotteries considered are with finite support. The set of all such lotteries over a space $S$ is denoted by $\Delta(S)$. The notation $\langle s_1, \ldots, s_n : p_1, \ldots, p_n \rangle$ denotes the lottery wherein outcome $s_k$ occurs with probability $p_k$. The fair lottery between $s_1$ and $s_2$ is denoted $\langle s_1, s_2 \rangle$.

**Preference Orders.** For a space $S$, two preferences orders\footnote{A preference order is a complete, transitive and reflexive binary relation.} are considered:

- **certainty preferences:** a preference order $\preceq$ on $S$,
- **lottery preferences:** a preference order $\succeq$ on $\Delta(S)$.

It is assumed throughout that $\preceq$ and $\succeq$ agree on the preferences over the certainties - $S$ (for which $\succeq$ are the degenerate lotteries). As customary, $\prec$ denotes the strict preference order induced by $\preceq$, and $\sim$ the induced indifference relation; similarly $\prec$ and $\sim$ denote the relations induced by $\succeq$.

$\succeq$ is assumed to be continuous; that is, for any lottery $L$, the sets $\{s : s \succeq L\}$ and $\{s : s \succeq L\}$ are open (in $S$). Since $\prec$ and $\succeq$ agree on $S$, this implies that $\succeq$ is also continuous.

All commodity spaces $C_i$ are assumed to be strictly essential\footnote{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 $S = T_1 \times \cdots \times T_n$ does not correspond to the original order.} \cite{18}; that is, for each $i$ and $s_{-i} \in S_{-i}$ (the remaining commodities), there exist $s_i, s'_i \in C_i$ with $(s_i, s_{-i}) \not\sim (s'_i, s_{-i})$.

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

**Factors and Partitions.** The term factor refers to a single $C_i$ or a product of several $C_i$’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 $T_i$ is the product of all commodity spaces associated with consumption at time $i$. A partition of $S$ is a representation of $S$ as a product of factors $S = T_1 \times \cdots \times T_n$.\footnote{An element of $S$ (or of any factor) is called a bundle.} An element of $S$ (or of any factor) is called a bundle.

**Separability.** Separability is a key notion in our analysis. Simply put, a factor is separable 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 $S = T_1 \times \cdots \times T_n$, we say that factor $T_i$ is separable if there exists a preference order $\succeq_{T_i}$ on $T_i$ such that for any $a_i, b_i \in T_i$ and any $c \in S_{-i}$ (the remaining factors),

$$a_i \preceq_{T_i} b_i \iff (a_i, c) \preceq (b_i, c).$$

It is important to stress that here separability only refers to the certainty preferences; it does not state or imply that the preferences on lotteries in one factor are separable (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 $\preceq$ instead of $\preceq_{T_i}$; thus, when $a, a' \in T$, we may write $a \preceq a'$ instead of $a \preceq_{T} a'$. It is worth noting that the product of separable factors need not be separable\footnote{A simple example is the preference on $X \times Y \times Z = (\mathbb{R}^+)^3$ represented by the function $v(x, y, z) = xy + z$. Here, each commodity space is separable, but $Y \times Z$ is not.}.
A partition \( \mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n \) is a separable partition if the product of any subset of factors is separable. By Gorman [18], for \( n \geq 3 \), it suffices to assume that \( \mathcal{T}_i \times \mathcal{T}_{i+1} \) is separable for all \( i \), and the separability of all other products then follows.

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

\[
s \preceq s' \iff f(s) \leq f(s').
\]

The function \( f : \mathcal{S} \to \mathbb{R} \) is an NM utility of \( \preceq \) if for any \( L, L' \in \Delta(\mathcal{S}) \),

\[
L \preceq L' \iff E_L[f(s)] \leq E_{L'}[f(s)],
\]

where \( E_L[f(s)] \) is the expectation of \( f(s) \) when \( s \) is distributed according to \( L \). In this case we also say that \( f \) represents \( \preceq \).

**Notations.** Throughout, \( a_i, b_i, c_i \) represent elements of \( \mathcal{T}_i \), and \( L_i \) a lottery in \( \Delta(\mathcal{T}_i) \). For \( i, j \), we denote \( \mathcal{S}_{\{i,j\}} = \prod_{t \not= i, j} \mathcal{T}_t \). For \( c \in \mathcal{S}_{\{i,j\}} \), by a slight abuse of notation we denote

\[
(a_i, a_j, c) := (c_1, \ldots, c_{i-1}, a_i, c_{i+1}, \ldots, c_{j-1}, a_j, c_{j+1}, \ldots, c_n).
\]

For \( \underline{s}, \bar{s} \), we denote

\[
[\underline{s}, \bar{s}] = \{ s : \underline{s} \preceq s \preceq \bar{s} \}
\]

That is, \( [\underline{s}, \bar{s}] \) is the closed interval of bundles between \( \underline{s} \) and \( \bar{s} \). Hence, we call such an \( [\underline{s}, \bar{s}] \) a bundle interval, or simply interval.
Part A: An Axiomatic Foundation

3. Lottery Sequences: Scale-Free Risk Aversion

Our first definition of risk aversion is set in the context of lottery sequences. Conceptually, the 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.

The Space. We consider an infinite sequence of factors $T_1, T_2, \ldots$, where $T_i$ represents the consumption space at time $i$. We denote $H^n = T_1 \times \cdots \times T_n$ - the finite history up to time $n$. In the following, $a_i, b_i, c_i$, will always be taken to be in $T_i$, and lottery $L_i$ will be over $T_i$.

Preference Orders. While the number of factors is infinite, we only consider preferences on the finite history spaces $H^n$. We denote by $\succeq^n$ the preference order on $H^n$, and by $\succeq^\Delta$ the preference order on $\Delta(H^n)$. The superscript $n$ is frequently omitted when clear from the context. Each $T_i$ is assumed to be separable in every certainty preference order $\succeq^n$ (but not necessarily utility separable - in preference order $\succeq^\Delta$).

We call the sequence of preference orders $\sim = (\succeq^1, \succeq^2, \ldots)$ the preference policy.

Lottery Sequences. Let $L_1, L_2, \ldots$, be a sequence of lotteries, with $L_i$ over $T_i$. Denote by $(L_1, \ldots, L_n)$ the lottery over $H^n$ 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 $c = (c_1, c_2, \ldots)$ is the repeated certainty equivalent of $L = (L_1, L_2, \ldots)$ if $(c_1, \ldots, c_{n-1}, c_n)\succeq^n(c_1, \ldots, c_{n-1}, L_n)$ for all $n$.

---

8We do not assume that $T_i = T_j$, 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 consideration here is irrespective of any such nominal matters.
Ultimate Inferiority. Consider a sequence \( c = (c_1, c_2, \ldots) \) of sure states, and a sequence \( L = (L_1, L_2, \ldots) \) of lotteries. Let \( \ell_i \) be the realization of \( L_i \). We say that \( c \) is ultimately inferior to \( L \) if

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

Notably, here \( \prec^n \) denotes the preference over the sure states. Thus, if \( c \) is ultimately inferior to \( L \), 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, \( c \) is ultimately superior to \( L \) if

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

Bounded and Non-Vanishing Lottery Sequences. We now want to define risk aversion as a policy for which the repeated certainty equivalent of a lottery sequence is always ultimately inferior to the lottery sequence itself. However, as such, this definition cannot be a good one since in the case that the “magnitude” of the lotteries rapidly diminishes, the overall outcome will necessarily 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 it can be shown that for most preference policies one can construct a lottery sequence that is ultimately inferior to its repeated certainty equivalent\(^\text{[10]}\). 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 \( [a_j, b_j] \subseteq [a_i, b_i] \) if \( (a_i, b_j, c) \preceq (b_i, a_j, c) \) for all \( c \in S_{-\{i,j\}} \) (see Figure 1).

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

\(^\text{[10]}\)\text{See Appendix B.}
A sequence of intervals \([a_1, b_1], [a_2, b_2], \ldots,\) is \textit{bounded} if \([a_i, b_i] \subseteq [a_1, b_1],\) for all \(i > 1.\) The sequence is \textit{vanishing} if for any \([\tilde{a}_1, \tilde{b}_1],\) there exists a \(j_0\) such that \([a_j, b_j] \subseteq [\tilde{a}_1, \tilde{b}_1]\) for all \(j > j_0.\) That is, the intervals in the tail of the sequence are as small as desired.

A lottery sequence \(L = (L_1, L_2, \ldots)\) is \textit{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 \(L\) is \textit{non-vanishing} if it includes an infinite sub-sequence of fair lotteries, the support thereof is not entirely within any vanishing interval sequence.

\textit{Risk Averse Policies.} Equipped with these definitions, we can now give a scale-free definition of risk aversion:

\textbf{Definition 1.} We say that preference policy \(\sim\) 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.
- **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.

\subsection{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.

\textbf{Definition 2.} We say that preference policy \(\sim\) 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.

\section{Scale-Free Risk Aversion: A Functional 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 functional form. Specifically, we show that (under appropriate 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 separability assumptions such preferences can be represented in an additive form, as follows. Consider the space $H^n = T_1 \times \cdots \times T_n$ ($n \geq 2$), with preference order $\succeq^n$. Debreu [8] proves that, if the partition is separable and $n \geq 3$, then $\succeq^n$ is additively separable; that is, there exist functions $v_i : T_i \to \mathbb{R}$, such that for any $(a_1, \ldots, a_n), (b_1, \ldots, b_n)$

$$(a_1, \ldots, a_n) \succeq^n (b_1, \ldots, b_n) \iff \sum_{i=1}^{n} v_i(a_i) \leq \sum_{i=1}^{n} v_i(b_i).$$

For $n = 2$, the additional Thomsen condition is also required. Debreu 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 $v_i$ a (Debreu) value function for $T_i$, and the aggregate function $v^n = \sum_{i=1}^{n} v_i$ a (Debreu) value function for $H^n$. We note that Debreu [8] 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. It is important to note that the value functions are defined solely on the basis of the certainty preferences.

4.2. Risk Aversion and Concavity. We now show that our axiomatic 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 $H^n$’s. The exact conditions are now specified.

Certainty Preference. For the certainty preferences, assume:

A1. Separability of the certainty preferences: $H^n = T_1 \times \cdots \times T_n$ is a separable partition (with respect to $\succeq^n$), for all $n$.

A2. Consistency of the certainty preferences: for any $n' > n$, the preference order induced by $\succeq^{n'}$ on $H^n$ is identical to $\succeq^n$

These assumptions yield the existence of value functions, as follows:

Proposition 1. Assuming A1-A2, there exist Debreu value functions $v_i : T_i \to \mathbb{R}$, $i = 1, 2, \ldots$, such that for all $n$, $v^n(a_1, \ldots, a_n) = \sum_{i=1}^{n} v_i(a_i)$ represents $\succeq^n$. 

---

11see page 8
12The Thomsen condition is that for all $a_1, b_1, c_1 \in T_1$, and $a_2, b_2, c_2 \in T_2$, if $(a_1, b_2) \sim (b_1, a_2)$ and $(b_1, c_3) \sim (c_1, b_2)$ then $(a_1, c_2) \sim (c_1, a_2)$.
13This is a slight abuse of notation. More precisely, $v$ is the function on $H^n$ given by $v(a_1, \ldots, a_n) = \sum_{i=1}^{n} v_i(a_i)$.
14That is, each consecutive pair of factors $T_i \times T_{i+1}$ is separable.
Lottery Preferences. Whereas the factors are assumed separable, the lottery preferences thereupon need not be separable. That is, the preference order on $\Delta(\mathcal{H}^n)$ induced by $\succeq^n$ may depend on the state $a_{n+1}$ in $T_{n+1}$. We do assume, however, the following weak consistency.

**A3. Weak consistency of the lottery preferences:** for any $n$, there exists a $\phi_{n+1} \in T_{n+1}$ with

$$L \succeq^n L' \iff (L, \phi_{n+1}) \succeq^{n+1} (L', \phi_{n+1}),$$

and the $\phi_n$’s are bounded away from the edges of $T_n$’s.\(^{15}\)

That is, weak consistency states that the preference $\succeq^n$ on $\Delta(\mathcal{H}^n)$ is consistent with $\succeq^{n+1}$ for at least one possible state $\phi_{n+1}$ of $T_{n+1}$. We call $\phi_{n+1}$ the presumed future, and it is assumed that it is bounded away from the boundaries of the state space (conceptually this means that the presumed future is not “extreme”).

### 4.2.1. Weak Scale-Free Risk Aversion and (Weak) Concavity.

For each $n$, let $u^n$ be the NM utility function representing $\succeq^n$. Since $\succeq^n$ and $\succeq^n$ agree, $u^n$ is a monotone transformation of $v^n$. Set $\hat{u}^n = u^n \circ (v^n)^{-1}$; so, $\hat{u}^n(v^n(a)) = u^n(a)$, for any $a \in \mathcal{H}^n$. 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 $\succeq^n$.

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

**Theorem 2.** Assuming A1-A3, $\succeq^n$ is weakly scale-free risk averse if and only if all the valued-scaled-utilities $\hat{u}^n$ are concave.

Thus, Theorem 2 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 establishes that all value-scaled utility functions must be concave, not only from some $n$ on.

### 4.2.2. (Strict) Scale-Free Risk-Aversion and Strict Concavity.

We would have now wanted to claim that (strict) scale-free risk-aversion corresponds to strict concavity of the value-scaled NM utility functions. 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 in some sense “uniformly” strictly concave.

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:

$$\mathcal{A}_f(x) = -\frac{f''(x)}{f'(x)}.$$  

From here on we assume that $\hat{u}^n$ is twice differentiable for all $n$.\(^{15}\) that is, there exists an $\delta > 0$ with $v_n(\phi_n) \pm \delta \in v_n(T_n)$ for all $n$.\(^{13}\)
Theorem 3. Assuming A1-A3, if $A^n_u(x)$ is bounded away from 0, uniformly for all $n$ and $x$, then $\preceq$ is SF-risk-averse.

Theorem 3 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 $\hat{u}^n$ 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 lottery $\langle x + \epsilon, x - \epsilon \rangle$; that is, for $\pi = \text{risk-prem}_{\hat{u}^n}(x \pm \epsilon)$

$$
\hat{u}^n(x - \pi) = \frac{\hat{u}^n(x + \epsilon) + \hat{u}^n(x - \epsilon)}{2}.
$$

Set $\hat{u} = (\hat{u}^1, \hat{u}^2, \ldots)$, the sequence of value-scaled-utilities. For any $\epsilon$ (sufficiently small) define

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

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

**Theorem 4.** Assuming A1-A3,

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

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

Finally, we establish that the sufficient condition of Theorem 4-(a) and that of Theorem 3 are the same.

**Proposition 5.** $RP_{\hat{u}}(\epsilon) = \Omega(\epsilon^2)$ as $\epsilon \to 0$, if and only if $A^n_u(x)$ is bounded away from 0, uniformly for all $n$ and $x$.

4.3. Risk Loving and Risk Neutrality. In analogy to Theorems 2 and 4 we have:

**Theorem 6.** Assuming A1-A3, 2 and 4

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

(b) Risk loving

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

(c) Risk Neutrality: $\preceq$ 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 a finite number of periods.

---

16that is, there exists an constant $\alpha > 0$ such that $A^n_u(x) \geq \alpha$ for all $n$ and $x$.

17recall 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 $y < y_0$. 

The Space. We consider a space $S$ and a separable partition $S = T_1 \times \cdots \times T_n$. In this case the factors $T_i$ may either be different time periods, as in the previous sections, or any other separable 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 opting for the certainty equivalent most probably results in an inferior outcome; that is, the probability that the realization is (weakly) preferred to the certainty equivalent is greater than the probability that the certainty equivalent is (weakly) preferred to the realization. The details follow.

Weak Risk Aversion. Denote the certainty equivalent of a lottery $L$ by $c(L)$\(^{18}\) Consider a lottery sequence $L = (L_1, \ldots, L_n)$, where $L_i \in \Delta(T_i)$ (and $L$ represents the independent application of each $L_i$). The sequence is a fair lottery sequence if each $L_i$ is a fair (50-50) lottery.

Definition 3. Preference order $\succeq$ is scale-free (SF) weakly-risk-averse if for any fair lottery sequence $L$,

\[
\Pr[c(L) \succeq \ell] \geq \Pr[c(L) \succ \ell].
\]

(where $\ell$ is the realization of $L$).

In words: with weak risk aversion, the chances of benefiting from opting for the certainty equivalent are no more than the chances of losing from this choice.

Strong Weak Aversion. We would have wanted to define strong weak aversion as a preferences wherein \(2\) holds with strong inequality. This, however, cannot be a good definition, since in the case where there is one “big” (fair) lottery in the sequence, and all the rest are “small”, the outcome of the big lottery will dominate that of the entire sequence, and the probability on both sides of \(2\) would equal $1/2$. So, what we need is two lotteries that are “of the same magnitude”. Hence, the following definition:

Definition 4. Fair lotteries $L_i = (a_i, b_i)$ and $L_j = (a_j, b_j)$ (in $\Delta(T_i)$ and $\Delta(T_j)$ respectively) are of the same magnitude if $(a_i, b_j) \sim (b_i, a_j)$\(^{19}\). Lottery sequence $L = (L_i, L_j, d)$ is a repeated lottery sequence if $L_i, L_j$ are of the same magnitude and $d \in S - \{i, j\}$.

With this definition, strong risk aversion is defined as a preference wherein strict inequality in \(2\) holds for repeated lottery sequences.

\(^{18}\)If there are several certainty equivalents, then pick one arbitrarily.

\(^{19}\)this is well defined, as the partition is separable, so the preferences on each pair of factors is independent of the others.
Definition 5. Preference order \( \succeq \) is scale-free (SF) risk-averse if for any non-degenerate repeated lottery sequence \( L \)

\[
\Pr[c(L) \succeq \ell] > \Pr[c(L) \succeq \ell].
\]

In words: with (strict) risk aversion, the chances of benefiting from opting for the certainty equivalent are less than the chances of losing from this choice.

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 7. If \( \succeq \) is weakly-SF-risk-averse with respect to some separable partition \( S = T_1 \times \cdots \times T_n \), then it is also weakly-SF-risk-averse with respect to any separable partition. Similarly for (strict-)SF-risk-aversion.

5.3. A Functional Form. Similar to 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 \( v \) be a Debreu value function representing \( \succeq \), \( u \) an NM utility representing \( \succeq \), and \( \hat{u} = u \circ (v)^{-1} \) the value-scaled-utility. Then,

Theorem 8. \( \succeq \) is weakly scale-free risk-averse if and only if \( \hat{u} \) is concave, and (strict) scale-free 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. Preference order \( \preceq \) is weakly-SF-risk-loving if for any fair lottery sequence \( L \),

\[
\Pr[c(L) \preceq \ell] \leq \Pr[c(L) \succeq \ell]
\]

and SF-risk-loving if for any non-degenerate repeated lottery sequence \( L \)

\[
\Pr[c(L) \preceq \ell] < \Pr[c(L) \succeq \ell].
\]

Preference order \( \preceq \) is risk-SF-neutral if for any fair lottery sequence \( L \)

\[
\Pr[c(L) \preceq \ell] = \Pr[c(L) \succeq \ell]
\]

Theorem 9. \( \preceq \) is weakly-SF-risk-loving if and only if \( \hat{u} \) is convex, (strictly)-SF-risk-loving if and only if \( \hat{u} \) is strictly-convex, and SF-risk-neutral if and only if \( \hat{u} \) is linear.
It is well known that under the classic, Arrow-Pratt definition of risk aversion, diminishing marginal utility and risk aversion are inseparable. 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, we suggest, it is possible disentangle the two within the expected utility framework. Basically, we shall argue that diminishing marginal utility is captured by the concavity of the Debreu value function, while risk aversion is captured by the concavity of the NM utility function with respect to the Debreu value function. In the remainder of this section we make these notions formal, by way of the indirect utility function.

The Setting. The consumption space is \( S = C_1 \times \cdots \times C_m \), where \( C_j \) represents the consumption of commodity \( j \), possibly at a specific time (so \( C_j \) and \( C_k \) may both represent consumption of the same commodity, but at different times). Each \( C_j \) is the half-open ray \([0, \infty)\) representing the amount of the commodity consumed. Let \( \succeq \) and \( \succeq \) be agreeing certainty and lottery preferences on \( S \). Throughout this section we assume that \( \succeq \) is additively separable. That is, there exists a Debreu value function \( v \) representing \( \succeq \) (as in Section 4.1). All factors considered as separable.

6.1. Diminishing Marginal Utility. We start with a preference-based definition of the notion of diminishing marginal utility. Conceptually, diminishing marginal utility of say bread, means that consuming an extra gram of bread is less “valuable” if we have already consumed a lot of bread, as compared to if we have consumed little bread. In order to make such a statement meaningful, we must define what we mean by “valuable”. A natural way to define “value” is with respect to other goods. Accordingly, we define (see Figure 2):

**Definition 7.** Factor \( T_i \) (with \( t \) commodities) exhibits diminishing marginal utility if for any \( a_i \in T_i, \ b_{-i}, \ b_{-i} \in S_{-i}, \ d \in \mathbb{R}^t_+ \) and \( \epsilon > 0 \) if

\[
(a_i + \epsilon d, b_{-i}) \sim (a_i, b_{-i}),
\]

then for any \( \alpha > 0 \)

\[
(a_i + (\alpha + \epsilon) d, b_{-i}) \succeq (a_i + \alpha d, b_{-i}).
\]

Strict diminishing marginal utility is defined analogously with strict preference.
Conceptually, the definition uses the commodities in $S_{-i}$ as the yard-stick for measuring the “value” of changes in factor $T_i$; taking factor $T_i$ to be bread and apples to be some other commodity, then “diminishing marginal utility of bread” says that if, say, starting at one kilo of bread and one apple, getting an extra 100 grams of bread is indifferent to getting one extra apple, then starting at two kilos of bread and one apple, getting an extra 100 grams of bread is less preferable to getting one extra apple. Thus, the additional one apple is used as the yard-stick with which the “value” of the extra 100 grams of bread is measured.

Two aspects of this definition are worth noting:

- The definition is based on the certainty preferences, $\preceq$, alone. This, we believe, is appropriate, as the intuitive notion of diminishing marginal utility is not related to lotteries, but rather to the diminishing “value” of the certainties.
- The definition is not scale-free; rather it is with respect to an exogenous scale of the commodities of interest. This, we believe, is inherent to the notion of diminishing marginal utility, as the notion compares the value of adding “the same amount” of the goods at different consumption levels. The notion of “the same amount” requires a scale.

The following simple proposition establishes the connection between the above preference-based definition of diminishing marginal utility and concavity of the Debreu value function $v$.

**Proposition 10.** Factor $T_i$ exhibits (strict) diminishing marginal utility if and only if $v_i$, the Debreu value function associated with factor $T_i$, is (strictly) concave.

---

$^20$vector $d$ is the direction in which consumption grows, while $\epsilon$ is the amount.
6.2. The Indirect Utility Function. Up to now we have considered preferences and utilities on consumption. We now move to examining how these preferences reflect on the utility of money, by way of the indirect utility functions. For a price vector $p$ and money amount $w$, let $V(w, p)$ be the indirect utility associated with the Debreu value function, and $U(w, p)$ the indirect utility associated with the NM utility function. Also, for factor $T_i$, the function $V_i(w_i, p_i)$ is the indirect utility associated with Debreu value function $v_i$, with the price vector $p_i$ - for the commodities in $T_i$ - and budget $w_i$ allocated to these commodities. We have,

**Proposition 11.** If factor $T_i$ exhibits (strict) diminishing marginal utility then $V_i(w_i, p_i)$ is (strictly) concave with respect to $w_i$, for any $p_i \gg 0$.

Similarly, if all factors exhibit (strict) diminishing marginal utility then $V(w, p)$ is (strictly) concave with respect to $w$, for any $p \gg 0$.

6.3. Disentangling Risk Aversion from Diminishing Marginal Utility. Consider $U(w)$ - the indirect NM utility function (here the price vector $p$ is fixed and omitted). Since $u$ is a monotone increasing transformation of $v$, the optimal allocation according to $v$ is also optimal according to $u$; that is, $U(w) = \hat{u}(V(w))$, where $\hat{u}$ is the value scaled utility function introduced in the previous sections. So, the concavity of $U$ with respect to money $(w)$ can naturally be decomposed into two: the concavity of $V$ with respect to money, and the concavity of $\hat{u}$ with respect to $V$. Specifically, denoting the coefficient of absolute risk aversion of a function $f$ by $A_f$, a simple computation gives:

$$A_U(w) = A_V(w) + A_{\hat{u}}(V(w)) \cdot V'(w).$$

That is, the coefficient of absolute risk aversion of the (indirect) NM utility $U$ with respect to money is the sum of

- the coefficient of absolute risk aversion of the (indirect) Debreu value function $V$, which, as argued above, represents diminishing marginal utility, and

- the coefficient of absolute risk aversion of the value-scaled NM utility $\hat{u}$, which, as argued in the previous sections, represents risk aversion, scaled by the (local) “exchange rate” between the dollar scale and the Debreu value scale. We note that such a (re)scaling happens wherever one moves from one scale to another, e.g. from dollars to Euros. The only difference here is that the “exchange rate” is not fixed.

In particular, even if both $\hat{u}$ and $V$ exhibit constant absolute risk aversion, and are concave, the overall (indirect) NM utility will exhibit *decreasing* absolute risk aversion (DARA). This may explain the prevalence of DARA in the real world. It also corresponds to the intuitive explanation of DARA as having to do with diminishing marginal utility; a poor person is more averse to risking $10,000 than a rich person because this amount is “worth more” to the poor person than to the rich. This vague but intuitive explanation of DARA is made formal by equation (3).
The seminal works of Arrow and Pratt 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-dimensional setting (see Section 9 for some references). As indicated in the introduction, we believe that a key difficulty stems from the scale dependency of the classic definition. In this section we consider multi-dimensional risk aversion in light of the framework developed in this paper.

7.1. 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 any given lottery \( L \), the associated risk premium decreases as the decision maker’s wealth increases. Similarly, the risk premium is independent of wealth if and only if the coefficient of absolute risk aversion is constant (CARA).

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

Consider two commodities, say the decision maker’s salary and the size of her land plot, with \( x \) being the salary and \( y \) the area in acres (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 \( x \), the coefficient of absolute risk aversion of \( u \) with respect to \( y \) is DARA. However, with \( x = 2 \), the certainty equivalent of a fair lottery between \( y = 2 \) and \( y = 4 \) is \( y \approx 2.93 \), while with \( x = 1000 \) the certainty equivalent of the same lottery is \( \approx 2.85 \), and with \( x = 1,000,000 \) it is \( \approx 2.83 \). So, the certainty equivalent (in \( y \)) decreases, and the risk premium increases, as the wealth \( x \) grows - contrary to the characteristic DARA behavior.

The problem of providing a meaningful extension of the notions of DARA and CARA to the multi-dimensional setting was left an open problem in [23]. Using our definition of risk aversion, this extension follows naturally, we follows.

For a lottery \( L_i \) on factor \( i \), and state \( d_{-i} \) in the other factors, let \( c(L_i)|_{d_{-i}} \) be the certainty equivalent of \( L_i \) (within \( T_i \)), when the state in the other factors is \( d_{-i} \). Formally, \( (c, d_{-i}) \sim (L_i, d_{-i}) \), for \( c = c(L_i)|_{d_{-i}} \). We have:

**Proposition 12.** The value scaled utility \( \hat{u} \) exhibits DARA if and only if for any separable factor \( T_i \), lottery \( L_i \) over \( T_i \), and \( d_{-i} \prec \bar{d}_{-i} \),

\[
(4) \quad c(L_i)|_{d_{-i}} \sim c(L_i)|_{\bar{d}_{-i}}.
\]

The DARA behavior is strict if and only if the preference in (4) is strict.

Similarly, \( \hat{u} \) exhibits CARA if and only if \( c(L_i)|_{d_{-i}} \) is independent of \( d_{-i} \).
Figure 3. Illustration of a MaPS: \( L' \) (solid circles) is a MaPS of \( L \) (empty circles). \( L' \) has the same marginals as \( L \), but the two possible outcomes are more extreme.

In words: if \( \hat{u} \) is DARA, then the certainty equivalent in factor \( T_i \) (of a lottery within the factor) increases with the “wealth level” in the other factors. Similarly, if \( \hat{u} \) is CARA, then the certainty equivalent in any factor is independent of the wealth level in the other factors. So, once risk aversion is defined with respect to the value function, the characteristic behavior of DARA and CARA carries over to the multi-dimensional setting.

7.2. Mean Preserving Spreads and Correlation Aversion. Rothschild and Stiglitz [29] defined the notion of a mean preserving spread (MePS) and showed that risk averse agents can be characterized by their dislike of any MePS. The notion of a MePS, however, is scale dependent. We now consider the notion of MaPS, which, we show, is the scale-free analogue of MePS, and prove that dislike of MaPS’s characterizes scale-free risk-averse agents. Throughout this section we assume a separable partition \( S = T_1 \times \cdots \times T_n \).

Definition 8. Let \( L = \langle (a_i, \bar{a}_i), (\bar{a}_i, a_i) \rangle \) be a fair lottery with \( a_i \prec \bar{a}_i \) and \( \bar{a}_i \prec a_i \). Lottery \( L' = \langle (a_i, \bar{a}_i), (\bar{a}_i, a_i) \rangle \) is a marginals preserving spread (MaPS) of \( L \) (see Figure 3).

A MaPS retains the marginals in all factors, but the two possible outcomes are more extreme. The following establishes that dislike of MaPS characterizes scale-free risk aversion.

Theorem 13. \( \hat{\succ} \) is weakly scale-free risk averse if and only if \( L \hat{\succ} L' \) whenever \( L' \) is MaPS of \( L \). Similarly for (strict) scale-free risk aversion and strict preference.

We note that the theorem holds even if the preference \( \hat{\succ} \) is not additively separable. When the preference is additively separable, the following establishes that MaPS and MePS are the same when using the Debreu value scale:

Proposition 14. If \( L' \) is a MaPS of \( L \) then \( v(L') \) is a MePS of \( v(L) \).
Conversely, for any sufficiently small fair lottery \( \tilde{y} = \langle y, \overline{y} \rangle \) in the râge of \( v \)
if \( \tilde{y} \) is a MePS of \( \tilde{x} = \langle x, \overline{x} \rangle \) then there exist \( L, L' \in \Delta(S) \), with \( v(L) = \tilde{x}, v(L') = \tilde{y}, \) and \( L' \) a MaPS of \( L \).

Dislike of MaPS has been studied in the literature under the notion of correlation aversion (Epstein and Tanny [15]), and first introduced by Richard [28] under the name multi-variate risk aversion. Richard introduced this as “a new type of risk aversion unique to multivariate utility functions” [28]. A later paper states: “[Richard’s definition] has nothing to do with what is generally known as risk aversion” [31]. Our analysis shows that the two notions are deeply tied, as established by the following corollary:

**Corollary 15.** The value scaled NM utility \( \hat{u} \) is concave if and only if \( \preceq \) is weakly correlation averse (that is, \( L \preceq L' \) whenever \( L' \) is MaPS of \( L \)). Similarly, \( \hat{u} \) is strictly concave if and only if \( \preceq \) is strictly correlation averse (that is, \( L \prec L' \) whenever \( L' \) is MaPS of \( L \)).

So, Richard’s multi-variate risk aversion and Arrow’s risk aversion are one and the same when using the value function scale.

Corollary [15] also provides a simple functional characterization of correlation aversion for the case where a Debreu value function exists. For the general case, Richard [28] gave a characterization stating that all mixed partial derivatives are negative [23].

### 7.3. Comparative Multi-Dimensional Risk Aversion.

Kihlstrom and Mirman [24], argued that in the multi-dimensional setting it is natural to restrict comparisons of risk aversion to decision makers (DMs) agreeing on the certainty preferences. Taking this approach, for DMs agreeing on the certainty preferences, it is possible to compare their level of risk aversion using the coefficient of absolute risk aversion of their respective value-scaled utility functions, \( \hat{u} \), as follows.

Consider two preference orders \( \preceq^{(1)}, \preceq^{(2)} \) on \( \Delta(S) \) agreeing on the certainties. Assuming the certainty preferences are additively separable, let \( v \) be a value function representing these preferences. Let \( u^{(1)}, u^{(2)} \) be their respective NM utilities, and \( \hat{u}^{(1)}, \hat{u}^{(2)} \), the associated value-scaled utilities. For a lottery \( L \) let \( c^{(j)}(L) \) be the certainty equivalent of \( L \) by \( u^{(j)} \) \( j = 1, 2 \). Then

\[
c^{(1)}(L) \preceq c^{(2)}(L)
\]

for all lotteries \( L \), if and only if

\[
A_{\hat{u}^{(1)}}(x) \geq A_{\hat{u}^{(2)}}(x)
\]

formally: there exists a constant \( \alpha > 0 \) (depending on the partition of \( S \)), such that for any \( \tilde{y} = \langle y, \overline{y} \rangle \) with \( |\overline{y} - y| \leq \alpha \ldots

Indeed, Corollary [15] can be deduced directly from Richard’s characterization by noting that when \( u(x_1, \ldots, x_n) = u(\nu_1(x_1) + \cdots + \nu_n(x_n)) \), then \( \frac{\partial^2 u}{\partial x_i \partial x_j} = \hat{u}'' \nu'_i \nu'_j \), for \( i \neq j \). So, all mixed derivatives share the sign of \( \hat{u}'' \) (since all \( \nu_i \)'s are increasing).

Indeed, the same is also true for the classic, uni-dimensional 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.
for all $x$ (where $A_{i}(x)$ is the coefficient of absolute risk aversion of $u^{(i)}$ 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. We note, however, that for the purpose of such comparative stats, the spacial characteristics of the value scale are not important; any monotone transformation thereof can be used. The special benefits of using the value scale are exhibited when considering notions such as DARA and CARA, as shown above, and as exhibited in the next section.

8. Savings Under Uncertainty

The following portfolio selection problem has been widely considered. Given a budget, a decision maker must distribute her funds amongst two assets, one secure and one risky, so as to maximize her expected utility. Arrow and Pratt studied how the optimal distribution behaves as a function of the total wealth and the risk attitude. In particular, they show that the coefficients of absolute risk aversion and relative risk aversion determine this behavior.

The Arrow-Pratt 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. The results do not carry over to the multi-dimensional setting, wherein the indirect utility function may exhibit 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 her budget between consumption in the first period, and savings with a random return, for consumption in the second period. Extending the study to this two-period setting has proved challenging, in particular with respect to increases in wealth. Kihlstrom and Mirman [23] provide a partial such extension, but only for the special case of homothetic preferences, and only extending the results concerning the coefficient of relative risk aversion, not those concerning the coefficient of absolute risk aversion. We now show that using the framework of this paper, a simpler and broader extension is enabled.

8.1. The setting. The consumption space is $S = T_1 \times T_2$, where $T_1$ - represents the consumption at time 1 - “today” - and $T_2$ representing the consumption at time 2 - “tomorrow”. The decision maker is endowed with wealth $w$, which she splits between today and tomorrow. Funds allocated for today are used for consumption today, at market prices. Funds for tomorrow are set aside as savings, invested in an asset with a rate of return $\tilde{z}$, which is a random variable. That is, upon saving $s$, the available funds tomorrow are $\tilde{z}s$, which are then used for purchase of consumption at market prices. When the rate of return on savings is non-random, we write $z$ instead of $\tilde{z}$.

The decision maker has a preference $\succeq$ on $\Delta(S)$, with $\succeq$ being the induced order $S$. It is assumed that $\succeq$ is additively separable. The decision maker is faced with the following optimization problem: given a budget $w$, optimally distribute the funds between consumption today and savings for tomorrow, so as to maximize expected utility (ex ante).
Let \( u \) be an NM utility representing \( \approx \), and \( v \) the Debreu value function representing \( \approx \), with \( v_i \) the value function associated with factor \( T_i, i = 1, 2 \). Let \( \hat{u} \) be the value-scaled utility, \( \hat{u} = u \circ v^{-1} \).

Recall that \( V_i(w, p_i) \) is the indirect utility function associated with \( v_i \).

With these notations, the decision maker’s optimization problem is:

\[
\max_s E_{z \sim \tilde{z}} [\hat{u}(V_1(w - s, p_1) + V_2(z \cdot s, p_2))] .
\]

She chooses \( s^* = s^*(w, \tilde{z}) \) that maximizes (5). We are interested in determining how \( s^*(w, z) \) behaves as a function of \( w \) and as a function of her risk aversion. For brevity, we write \( s^*(w) \), when \( \tilde{z} \) is fixed. We are also interested in \( r^*(w) = \frac{s^*(w)}{w} \), the fraction of savings out of the total wealth.

From here on, we assume that the price vector \((p_1, p_2)\) is fixed, and omit reference to it. We note, however, that \( \tilde{z} \) can also be interpreted as (uniform) random price changes in \( p_2 \), that is - a random inflation rate - rather than a random interest rate. Since \( V \) is homogeneous of degree 0, all results carry over to this interpretation.

We assume that \( V_1, V_2, \) are increasing (more money is better), concave (diminishing marginal utility - see Section 6), and differentiable; \( \hat{u} \) is assumed to be concave (scale-free risk aversion), and twice differentiable. These concavity assumption guarantee a unique optimum. It is assumed throughout that \( \tilde{z} \gg 0 \).

For a real function \( f, f' \) denotes its derivative, and \( A_f(x) \) denotes the coefficient of absolute risk aversion of \( f \) at point \( x \).

8.2. Behavior Under Certainty. We first consider the behavior of \( s^*(w, z) \) under certainty, that is, when \( z \) is non-random.

Increases in Wealth. When the rate of return is non-random we have:

**Proposition 16.** For \( z \) non-random, \( s^*(w, z) \) increases with \( w \).

Increases in \( z \). The behavior of \( s^*(w, z) \) when \( z \) increases is determined by the curvature of \( V_2 \), as follows. Let \( R_{V_2}(x) = -\frac{xV_2''(x)}{V_2'(x)} \) be the coefficient of relative risk aversion of \( V_2 \) at \( x \).

We write \( R_{V_2} \leq 1 \) to denote that \( R_{V_2}(x) \leq 1 \) for all \( x \), and similarly for \(<,\geq \) and \(>\).

**Proposition 17.** For \( z \) non-random

- \( s^*(w, z) \) increases with \( z \) if \( R_{V_2} \leq 1 \).
- \( s^*(w, z) \) decreases with \( z \) if \( R_{V_2} \geq 1 \).

8.3. Condition for Strictly Positive Savings. Moving to the case where \( \tilde{z} \) is random, Arrow [2] showed that in his setting, investment in the risky asset is strictly positive, regardless of the risk attitude, provided that \( E[\tilde{z}] > 1 \). In our setting, this no longer holds. As an example, consider \( u(w_1, w_2) = 10w_1 + w_2 \), and \( z = 2 \). In this case it is best to consume all funds in time 1. The following, however, is an easy generalization of Arrow result:

\[ R_{V_2} \] has actually nothing to do with risk aversion - as \( V_2 \) is fully determined by the certainty preferences. It is merely a measure of the curvature of \( V_2 \).
Proposition 18. If $E[\tilde{z}] > \frac{V'_1(w)}{V'_2(0)}$ then $s^*(w, \tilde{z}) > 0$.

In Arrow’s setting $V'_1 = V'_2 = 1$, so Proposition 18 indeed generalizes Arrow’s result.

8.4. Increases in Wealth. We are interested in studying how $s^*$ and $r^*$ behave as the a function of $w$ (the wealth). For brevity, from here on, all statements are stated assuming that $s^*$ is not at the boundaries (0 or $w$). When $s^*$ is at boundaries the statements continue to hold in the weak sense.

Theorem 19. Assuming $V_1, V_2, \hat{u}$, are concave. If

- $\hat{u}$ is DARA and $R_{V_2} \leq 1$, or
- $\hat{u}$ is IARA and $R_{V_2} \geq 1$,

then $s^*(w)$ is increasing. If, in addition, either: the above are strictly so or $V_1$ is strictly concave, then $s^*(w)$ is strictly increasing.

Note that the first bullet in the theorem is a generalization of Arrow’s result, as in Arrow’s setting $V_1$ and $V_2$ are both the identity, so (weakly) concave and $R_{V_2} < 1$.

Since CARA is both weak-DARA and weak-IARA, for CARA preferences the condition on $V_2$ can be omitted:

Corollary 20. Assuming $V_1, V_2, \hat{u}$, are concave, if $\hat{u}$ is CARA then $s^*(w)$ is increasing. If $V_1$ is strictly concave, then $s^*(w)$ is strictly increasing.

For the ratio of savings, $r^*$, Arrow showed that in the uni-dimensional case, if $u$ is IRRA (increasing relative risk aversion) then $r^*$ decreases with wealth. This is generalized in the following Theorem.

Theorem 21. Assuming $V_1, V_2, \hat{u}$ are concave. If the following three hold

(1) the elasticity of $\hat{u}'$ of with respect to $z$, $\frac{\partial \ln(\hat{u}'(V_1(w(1-r)) + V_2(zw)))}{\partial \ln(z)}$, is decreasing in $w$,
(2) $R_{V_2} \leq 1$,
(3) the marginal rate of substitution $\frac{V'_1(c_1)}{V'_2(c_2)}$ is increasing in $w$ (for all $c_1, c_2$),

then $r^*(w)$ decreases with $w$. If either (1)+(2) or (3) are strictly so, then $r^*(w)$ is strictly decreasing. If (3) is reversed, and either (1) or (2) (but not both) also reversed, then $r^*(w)$ increases with $w$, and strictly so if the behavior in (1)+(2) or (3) is strict.

Condition (1) is a generalization of the IRRA condition. To see this note that in Arrow’s case the elasticity of $\hat{u}'$ with respect $z$ is:

$$\frac{\partial \ln(\hat{u}'(w(1-r) + zw))}{\partial \ln(z)} = \frac{\hat{u}''(w(1-r) + zw)}{\hat{u}'(w(1-r) + zw)} \cdot zw$$

$$= \frac{\hat{u}''(w(1-r) + zw)}{\hat{u}'(w(1-r) + zw)} \cdot (w(1-r) + zw) \cdot \frac{zw}{1-r + zw}$$

25
which is hence the negative of the coefficient of relative risk aversion, multiplied by a positive factor independent of $w$. So, this elasticity is decreasing when $\hat{u}$ is IRRA. This becomes more transparent in the following corollary:

**Corollary 22.** Assuming $V_1, V_2, \hat{u}$ are concave. If the certainty preferences exhibit constant elasticity of substitution (CES) with elasticity $\sigma \neq 1$, then

- if $\hat{u}$ is DRRA then $r^*(w)$ increases with $w$.
- if $\hat{u}$ is IRRA then $r^*(w)$ decreases with $w$.

If the certainty preference are Cobb-Douglas (CES, $\sigma = 1$) with (ordinal) utility representation $U(c_1, c_2) = c_1^{\beta_1} c_2^{\beta_2}$, then $r^* = \frac{\beta_1}{\beta_1 + \beta_2}$ for all $\hat{u}, w, \hat{z} \neq 0$.

8.5. **Increases in Risk Aversion.** Consider two decision makers, denoted $\oplus, \ominus$, with identical wealth, $w$, and agreeing on the certainty preferences, represented by $V_1, V_2$. DM $\oplus, \ominus$, have value-scaled utility functions $\hat{u}_\oplus, \hat{u}_\ominus$.

**Proposition 23.** Assuming $V_1, V_2, \hat{u}_\ominus, \hat{u}_\oplus$, are concave. If $\mathcal{A}_{\hat{u}_\ominus}(x) \geq \mathcal{A}_{\hat{u}_\oplus}(x)$, for all $x$, then

- if $\mathcal{R}_{V_2} \leq 1$ then $s^*_\ominus \leq s^*_\oplus$.
- if $\mathcal{R}_{V_2} \geq 1$ then $s^*_\ominus \geq s^*_\oplus$.

In both cases, the inequality is strict if the respective conditions are so.

Together with the analysis of Section 8.2, Proposition 23 states that $s^*$ decreases with risk aversion if $s^*$ increases with $z$, and increases with risk aversion if $s^*$ decreases with $z$. This is essentially the same result obtained in Kihlstrom and Mirman’s original paper [24] (assuming the certainty preferences are additively separable).\(^{26}\) The advantage of Proposition 23 is in stating the conditions in fully function form.

9. **Related Work**

**Multi-Dimensional Risk Aversion.** Multi-dimensional risk aversion has been extensively studied, both in the expected utility and the non-expected utility frameworks. It is out of the scope of this paper to review all this work. Here, we highlight some of the most relevant work (all within the expected utility framework).

Kihlstrom and Mirman [24] observed the difficulty in extending the notion of risk aversion to the multi-dimensional case, and were the first to suggest that risk attitude comparisons should be restricted to agents agreeing on the certainty preferences (see also Karni [21]). For such agents, they prove that “more risk averse” - in the sense of greater risk premium (in any direction) - corresponds

\(^{26}\) Technically, [24] also states the converse of the proposition, that is, that provided that $s^*(w, z)$ increases with $z$, if $s^*_\oplus \leq s^*_\ominus$ everywhere then $\mathcal{A}_{\hat{u}_\ominus}(x) \geq \mathcal{A}_{\hat{u}_\oplus}(x)$ for all $x$ (and similarly for the case that $s^*(w, z)$ decreases with $z$). However, this direction follows directly from the stated direction, as otherwise there is an interval wherein $\mathcal{A}_{\hat{u}_\ominus} < \mathcal{A}_{\hat{u}_\oplus}$ and the proposition can be applied in reverse on that interval.
to relative concavity of the respective NM utility functions. They mention, but leave open, the problem of providing a meaningful notion of DARA and CARA in the multi-dimensional case.

Levy and Levy [26] consider multivariate risk aversion by way of the probability premium. This allows them to define a “non-directional” index, though still only applicable to agents that agree on the certainty preferences.

Duncan [12] defines a matrix measure of multi-variate risk aversion, $R = \left[ -\frac{\partial^2 u}{\partial x_i \partial x_j} / \frac{\partial u}{\partial x_i} \right]$, and draws connections between this matrix and the risk-premium vectors. This measure does not directly lend itself to comparisons among different decision makers (as noted in [26]). We note that for additively separable certainty preferences, the off-diagonal entries of $R$ are all positive multiples of $A_\hat{u}$, so share the same sign.

A different approach is taken by Stiglitz [32] and Karni [20]. They compare the risk aversion in the multi-dimensional setting by way of the corresponding indirect utility function. Stiglitz studies the structure of the certainty preferences implied by risk aversion assumptions, such as risk neutrality, CARA, and so, of the indirect utility function. Karni introduces a matrix measure of risk aversion for the indirect utility function. He proves that this measure is everywhere greater for one indirect utility function, $U^{(1)}$, than for another, $U^{(2)}$ (in the sense that the difference is positive definite), if and only if the income risk premium according to $U^{(1)}$ is greater than according to $U^{(2)}$ - for any lottery on income and prices (where the income risk premium is defined as the reduction in income the decision maker is willing to suffer in order to avoid the risk). Importantly, by shifting attention to the indirect utility function, Karni is able to obtain meaningful comparisons even among decision makers that do not agree on the certainty preferences.

Yet another approach is taken by Richard [28], who defines multivariate risk aversion as aversion to correlated risks. The relation of Richard’s definition to ours is considered in Section 7.2. In particular, while Richard’s definition is introduced as “a new type of risk aversion unique to multivariate utility functions”, our analysis shows that it coincides with the Arrow-Pratt definition when using the value function scale.

Savings Under Uncertainty. Savings under uncertainty is another topic that has been extensively studied in the literature. Again, we only highlight a few related works.

Kihlstrom and Mirman [24] and Diamond and Stiglitz [11] considered the setting of Section 8.5 - distinct decision makers that agree on the certainty preferences, and with identical wealth. They prove a result similar to Proposition 23 (and without assuming the existence of Debreu value function). Bommier et al. [6] prove that the same holds under a generalized notion of increased risk aversion. Technically, our Proposition 23 follows from all of the above (once Proposition 17 is invoked). We included the proposition for completeness, and for its simple functional form.

The case of a single decision maker at different wealth levels was considered by Kihlstrom and Mirman in [23]. They offer a result only for the case of homothetic preferences. For such preferences they decompose the NM utility function, $u$, into two: the least concave representation of $u$ (see
which they denote by $u^*$, and $h = u \circ (u^*)^{-1}$ (so that $u(x) = h(u^*(x))$). With this decomposition, they obtain a result showing that $r^*$ - the fraction of savings - increases/decreases in wealth according to whether $h$ is DRRA or IRRA. This is similar to our Corollary 22. The authors note that “the concepts of increasing, decreasing and constant absolute risk aversion cannot be given analogous justification”. Using our framework, we are able both to consider general preferences, not only homothetic, and to establish the relevance of the notions of decreasing/increasing/constant absolute risk aversion to the savings-under-uncertainty setting.

**Strength of Preference and Relative Risk Aversion.** Dyer and Sarin [14] and Bell and Raiffa [5] suggest measuring risk aversion with respect to the strength of preference function. Roughly, the strength-of-preference theory assumes that decision makers have well defined preferences over differences between states; that is, the decision maker can state that she prefers the transition $x_1 \mapsto x_2$ over the the transition $y_1 \mapsto y_2$ (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 [13], that represents these preferences. Given such a function $F$, Dyer and Sarin [14] and Bell and Raiffa [5], consider a notion of relative risk aversion [28] which says that the $u$ - the NM utility - is more concave than the strength of function $F$ ([5] called this intrinsic risk aversion). We note that under some interpretations, the Debreu’s value-function can be used as a strength of preference function (see [5] and [14]). Under such an interpretation, the [14] and [5] relative risk aversion notion is related to our functional characterization of scale-free risk aversion. Conceptually, however, our approach is different from that of [14] and [5]. First, we do not suppose preferences over differences, but rather only over bundles. Second, [14] and [5] do not provide an axiomatic justification for their definitions or their choice of scale. Finally, on both the technical and applicative aspects, our study is different from that of [14] and [5].

**Intertemporal Risk Aversion.** Traeger [33] is motivated by some of the same conceptual questions that motivate our work; namely - providing a scale-free, axiomatic framework for the theory of risk aversion (see also [34]). Unlike our work, Traeger’s work is set in the Kreps and Porteus 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. Theorem 2 of [33] establishes that a decision maker is IRA 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 somewhat similar to our

---

27In fact, it can be shown that, provided that a a Debreu value exists, $h$ is DRRA (/IRRA) if and only if $\hat{u}$ is DRRA (/IRRA). So, the two results are closely related.

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

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

30See Lemma A.9 of the appendix, which relates the two definitions of correlation aversion.
Corollary 15. 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 concavity of the risk aggregation function with respect to the certainty aggregation function.

10. Concluding Remarks

We presented an axiomatic definition of risk aversion, based entirely on the internal structure of preferences of the decision maker; independent of money or any other units. We then showed that when cast in functional form, this axiomatic definition coincides 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. Several application of this framework have been presented. We conclude with a few remarks and mention of some additional applications.

Inter-Temporal and Intra-Temporal Risk Aversion. It should be stressed that scale-free risk-aversion, as considered in this paper, does not relate only to lotteries involving multiple time periods (or commodities), but also to lotteries within a single time (/commodity). It easy to see that (assuming standard expected utility) given the inter-temporal certainty preferences, inter-temporal lottery preferences determine intra-temporal lottery preferences, and vice versa. Thus, inter-temporal and intra-temporal risk attitudes are one and the same. We use the inter-temporal 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 functional form, based on the Debreu value function. The multi-dimensional setting merely provides us with the appropriate scale with which to measure risk aversion, both inter and intra-temporal.

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. The framework developed in this paper enables a formal way to isolate the contribution of this uncertainty to the overall discounting, obtaining a measure of the survival uncertainty discount rate. Furthermore, it can be shown that key stats of this discounting rate are directly related to the characteristics of the decision maker’s risk attitude - risk aversion, risk loving, DARA and CARA. These results shall be described in a separate publication.

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, even in two player matching pennies games there are multiple possible equilibria.
Acknowledgments. We are grateful to Itzhak Gilboa, Edi Karni, Uzi Segal, and Peyton Young for helpful discussions.

REFERENCES


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:

- \( S = T_1 \times \cdots \times T_n \) is a separable partition.
- \( a_i, b_i, \) and \( c_i \) are points in \( T_i \).
- \( L_i \) is a lottery over \( T_i \) and \( \ell_i \) is the realization of \( L_i \).
- \( v_i \) is the Debreu value function associated with factor \( T_i \).
- \( v \) is the aggregate value function for \( S \), and \( v^n \) the aggregate value function for \( H^n \).
- \( u \) is an NM utility on \( S \) and \( u^n \) an NM utility on \( H^n \).
- \( \hat{u} \) is the value scaled NM utility associated with \( u \), \( \hat{u} = u \circ (v)^{-1} \), and similarly \( \hat{u}^n = u^n \circ (v^n)^{-1} \).
- Generally, subscripts are used to denote the index of the factor in question (as in \( T_i, a_i, b_i, L_i, v_i \)), while superscripts are used to denote the number of factors (as in \( H^n, \preceq^n, u^n, v^n \), and so).
- \( x, y, z, \) are real numbers.
- \( \alpha, \beta, \delta \) - with or without indices or primes - are positive reals.
- 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 1. Assuming A1-A2, there exist Debreu value functions \( v_i : T_i \to \mathbb{R}, i = 1, 2, \ldots \), such that for all \( n \), \( v^n(a_1, \ldots, a_n) = \sum_{i=1}^n v_i(a_i) \) represents \( \preceq^n \).

Proof. Consider \( H^n \) for \( n \geq 3 \). By assumption A1, any product of the \( T_i \)’s is separable. Hence, there exist value functions \( v_1^n, \ldots, v_n^n \), with \( \sum_{i=1}^n v_i^n \) representing \( \preceq^n \). We now show that there is actually a single function \( v_i \), for each \( i \), that works for all the \( H^n \)’s.

For \( i = 1, 2, 3 \), set \( v_i := v_i^3 \). Suppose \( v_i \) has been defined for all \( i < n \); we inductively define \( v_n \). By the inductive hypothesis, \( \sum_{i=1}^{n-1} v_i \) represents \( \preceq^{n-1} \). By consistency (A2), the function \( \sum_{i=1}^{n-1} v_i^n \) also represents \( \preceq^{n-1} \). So, by uniqueness of the value functions, there exist constants \( \beta > 0, \xi_i \), such that \( v_i = \beta v_i^n + \xi_i \), for \( i = 1, \ldots, n-1 \). So, setting \( v_i = \beta v_i^n \), we have that

\[
\sum_{i=1}^n v_i = \sum_{i=1}^{n-1} (\beta v_i^n + \xi_i) + \beta v_n^n = \beta \sum_{i=1}^n v_i^n + \text{constant},
\]

which represents \( \preceq^n \), as required. \( \square \)

From now on we assume w.l.o.g. that the factors are already represented in units of the respective value functions; that is, \( v_i(a_i) = a_i \) for all \( i \). Then \( u^n \), the NM utility function representing \( \preceq^n \), is

\[31\]An NM utility for \( S \) and \( H^n \) necessarily exists since the NM axioms are assumed to hold, and we consider only lotteries with finite support (see Fishburn [16, Theorem 8.2]). Furthermore, since all lottery preferences are continuous, so are \( u, u^n \).
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 \( a_1 + \cdots + a_n = b_1 + \cdots + b_n \).

**Lemma A.1.** Let \( X_1, X_2, \ldots \) 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

\[
\text{Pr}[S_n \geq 0 \text{ infinitely often}] > 0. \tag{6}
\]

**Proof.** Denote \( \sigma_i^2 = \text{Var}(X_i) \), and \( \Sigma_n^2 = \sum_{i=1}^n \sigma_i^2 \). The \( X_i \)'s are independent, so \( \Sigma_n^2 = \text{Var}(S_n) \).

Now, either \( \Sigma_n^2 \to \infty \) or not. We consider each case separately.

If \( \Sigma_n^2 \to \infty \), applying the central limit theorem for uniformly bounded random variables (e.g. \(^{[19]} \), Theorem 9.5) we obtain that

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

In particular, \( \text{Pr}[S_n \geq 0 \text{ infinitely often}] > 0 \).

Next, suppose that \( \Sigma_n^2 \) does not go to infinity. Each \( \sigma_i^2 \) is non-negative. Hence, the \( \Sigma_i^2 \)’s form a monotonically non-decreasing and bounded sequence, and hence converge. So, for any \( \eta > 0 \) there exists an \( N_\eta \) with \( \sum_{i=N_\eta}^\infty \sigma_i^2 < \eta \). 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 \( \alpha > 0 \) with \( \text{Pr}(X_1 \geq \alpha) = q_\alpha > 0 \). Choose \( \eta < \alpha^2 \). By the Chebyshev inequality, for all \( n > N_\eta \),

\[
\text{Pr}\left[ \sum_{i=N_\eta}^n X_i < -\alpha \right] < \frac{\text{Var}(\sum_{i=N_\eta}^n X_i)}{\alpha^2} \leq \frac{\eta}{\alpha^2} < 1.
\]

Clearly, there is some probability \( p^+ > 0 \) for which \( \text{Pr}[\max_{n=2,\ldots,N_\eta} \{S_n - X_1\} \geq 0] \geq p^+ \). So for all \( n \),

\[
\text{Pr}[S_n \geq 0] \geq \text{Pr}[X_1 \geq \alpha] \cdot \text{Pr}\left[ \max_{n=2,\ldots,N_\eta} (S_n - X_1) \geq 0 \right] \cdot \text{Pr}\left[ \sum_{i=N_\eta}^n X_i \geq -\alpha \right] \geq q_\alpha \cdot p^+ \cdot (1 - \frac{\eta}{\alpha^2}) > 0.
\]

So, in particular, \( \text{Pr}[S_n \geq 0 \text{ infinitely often}] > 0. \)

Let \( (\phi_2, \phi_3, \ldots) \) be the presumed future sequence. By assumption, the presumed future is bounded away from the boundaries, that is, there exists \( \delta > 0 \) with \( \phi_i \pm \delta \in T_i \), for all \( i \).

**Theorem 2.** Assuming A1-A3, \( \bar{\zeta} \) is weakly scale-free risk averse if and only if all the valued-scaled-utilities \( \hat{u}^n \) are concave.

\(^{32}\)that is, the support of all the random variables is included in a real interval \([b, b]\), with \( b, b \) finite.
Proof. $\exists$ 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 $< \delta$. So, there exist $x$, $\epsilon \leq \delta$ and $0 < \beta < \epsilon$ with

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

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

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

(7)

We construct a recurring lottery sequence $L$ 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 $L = (L_1, L_2, \ldots)$ is defined as follows:

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

We now inductively determine the repeated certainty equivalent of $L = (L_1, L_2, \ldots)$, 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 $L_{k+1} = ((\phi_{k+1} - \epsilon), (\phi_{k+1} + \epsilon))$. So, by (7), its certainty equivalent is $\beta$ above the expectation; 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 (7) applies to the lottery at time $k+3$, which is $L_{k+3} = ((\phi_{k+3} - \epsilon), (\phi_{k+3} + \epsilon))$. 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, w.l.o.g. assume 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 $\ell_i$ be the realization of $L_i$. Then,

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

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

All $\hat{u}^n$ are concave $\Rightarrow$ $\mathcal{H}$ is weakly SF risk averse: Consider a lottery sequence $L = (L_1, L_2, \ldots)$. W.l.o.g. $E(L_i) = 0$ for all $i$. Denote by $c = (c_1, c_2, \ldots)$ the repeated certainty equivalent of $L$.

Since $\hat{u}^i$ is concave, $c_i \leq 0$ for all $i$ (recall that $u^i(x_1, \ldots, x_i) = \hat{u}^n(x_1 + \cdots + x_i)$). So, for any $n$,

$$(\ell_1, \ldots, \ell_n) \prec^n (c_1, \ldots, c_n) \Rightarrow \sum_{i=1}^{n} \ell_i < 0.$$
So, 
\[
\Pr[(\ell_1, \ldots, \ell_n) \prec^n (c_1, \ldots, c_n) \text{ from some } n \text{ on}] \leq (1 - \Pr\left[\sum_{i=1}^{n} \ell_i \geq 0 \text{ infinitely often}\right]) < 1.
\]
where the last inequality is by Lemma A.1. So, \((c_1, c_2, \ldots)\) is not ultimately superior to \((L_1, L_2, \ldots)\). \(\square\)

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

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

**Lemma A.3.** Let \(X_1, X_2, \ldots,\) be independent random variables with \(E[X_i] = 0, |X_i| \leq C.\) Set \(\Var(X_i) = \sigma_i^2, S_n = \sum_{i=1}^{n} X_i\) and \(\Sigma_n^2 = \sum_{i=1}^{n} \sigma_i^2.\) If \(\Sigma_n \to \infty,\) then for any \(\alpha > 0\)
\[
\Pr[S_n > \alpha\Sigma^2 \text{ infinitely often}] = 0.
\]

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

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

Set \(\gamma = 0.5,\) and let \(\eta\) be that provided by Theorem A.2. Set \(\beta = \min\{\eta, \alpha/2\}.\) Then, considering \(n(k),\) by Theorem A.2 setting \(a = \beta\Sigma_n(k)\)
\[
(9) \quad \Pr[S_{n(k)} > \beta\Sigma_n(k) \cdot \Sigma_n] < e^{-\frac{\beta^2\Sigma_n(k)}{2}(1-\gamma)} \leq e^{-\frac{\beta^2k}{4}}.
\]

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,
\[
\Var(D_{n(k+1)}) = \Sigma_n^2 - \Sigma_n^2 \leq (k + 1 + C^2) - k = 1 + C^2.
\]

So, by the Kolmogorov inequality
\[
(10) \quad \Pr\left[\max_{n(k) < j \leq n(k+1)} \{D_j\} \geq \beta\Sigma_n^2(k)\right] \leq \frac{\Var(D_{n(k+1)})}{(\beta\Sigma_n^2(k))^2} \leq \frac{1 + C^2}{\beta^2k^2}.
\]

Combining (9)-(10), for any \(k\)
\[
\Pr[A_k] = \Pr[\exists i, n(k) < i \leq n(k+1), S_i > \alpha\Sigma_i^2] \\
\leq \Pr[S_{n(k)} \geq \beta\Sigma_n^2(k)] + \Pr[\max_{n(k) < j \leq n(k+1)} \{D_j\} \geq \beta\Sigma_n^2(k)] \\
\leq e^{-\frac{\beta^2k}{4}} + \frac{1 + C^2}{\beta^2k^2}.
\]

35
So, \( \sum_{k=1}^{\infty} \Pr[A_k] < \infty \). So, by the Borel Cantelli lemma
\[
\Pr[A_k \text{ occurs infinitely often}] = 0.
\]
For any \( k \) there is only a finite number of \( i \)'s with \( n(k) < i \leq n(k+1) \). So, \( S_i > \alpha \Sigma_i^2 \) infinitely often only if \( A_k \) occurs infinitely often, and the result follows. \( \square \)

**Theorem 3.** Assuming A1-A3, if \( A_{\hat{u}}(x) \) is bounded away from 0, uniformly for all \( n \) and \( x \), then \( \hat{\succ} \) is SF-risk-averse.

*Proof.* Let \( \alpha \) be the uniform bound on \( A_{\hat{u}}(x) \), that is, \( A_{\hat{u}}(x) \geq \alpha \) for all \( n, x \). By Pratt [27], for sufficiently small lottery \( L \),
\[
\text{risk-prem}_{\hat{u}}(L) = \frac{A_{\hat{u}}(x) \text{ Var}(L)}{2} + o(\text{Var}(L)).
\]
So, there exists an \( 0 < \epsilon < 1 \), such that if \( \text{Var}(L) < \epsilon \), then
\[
\text{risk-prem}_{\hat{u}}(L) > \frac{\alpha \epsilon}{4}.
\]
(11)
Let \( L = (L_1, L_2, \ldots) \) be a bounded, non-vanishing lottery sequence. W.l.o.g. \( E(L_i) = 0 \) for all \( i \).
Set \( \sigma_i^2 = \text{Var}(L_i), S_n = \sum_{i=1}^{n} L_i \) and \( \Sigma_n^2 = \text{Var}(S_n) = \sum_{i=1}^{n} \sigma_i^2 \). Since \( L \) is non-vanishing \( \Sigma_n \to \infty \).
Since \( L \) is bounded, there exists a \( C \geq 1 \) such that \( \sigma_i^2 \leq C \) for all \( i \). Note that for \( L_i \) with \( \sigma_i^2 > \epsilon \) we still have the bound \( \text{risk-prem}_{\hat{u}}(L_i) > \frac{\alpha \epsilon}{4} \).

Set \( \alpha_1 = \frac{\alpha \epsilon}{4C} \). Then, for \( L_i \) with \( \sigma_i^2 > \epsilon \)
\[
\text{risk-prem}_{\hat{u}}(L_i) > \frac{\alpha \epsilon}{4} = \alpha_1 C \geq \alpha_1 \sigma_i^2.
\]
(12)
For \( L_i \) with \( \sigma_i^2 \leq \epsilon \)
\[
\text{risk-prem}_{\hat{u}}(L_i) > \frac{\alpha \sigma_i^2}{4} \geq \alpha_1 \sigma_i^2
\]
by (11). So, \( \text{risk-prem}_{\hat{u}}(L_i) > \alpha_1 \sigma_i^2 \) for all \( i \).
Let \( (c_1, c_2, \ldots) \) be the repeated certainty equivalent of \( L \). So,
\[ c_i < -\alpha_1 \sigma_i^2, \]
for all \( i \). So,
\[
[-\alpha_1 \Sigma_n^2 < S_n] \Rightarrow \left[ \sum_{i=1}^{n} c_i < S_n \right] \Rightarrow [(c_1, \ldots, c_n) \prec (\ell_1, \ldots, \ell_n)].
\]

\( ^{33} \)that is, there exists an constant \( \alpha > 0 \) such that \( A_{\hat{u}}(x) \geq \alpha \) for all \( n \) and \( x \).

\( ^{34} \)To see this, define \( L_i' = \sqrt{\epsilon} L_i. \) Then, \( \text{Var}(L_i') = \frac{\epsilon}{2} \text{Var}(L_i) \leq \epsilon \), and \( L_i \) is a mean preserving spread of \( L_i' \).
Hence, since \( \hat{u}^\alpha \) is everywhere concave,
\[
\text{risk-prem}_{\hat{u}}(L_i) \geq \text{risk-prem}_{\hat{u}}(L_i') \geq \frac{\alpha \epsilon}{4}.
\]
So, it is sufficient to prove that
\[ \Pr[S_n > -\alpha_1 \Sigma^2_n \text{ from some } n \text{ on}] = 1. \]
which is equivalent to saying that
\[ \Pr[S_n < -\alpha_1 \Sigma^2_n \text{ infinitely often}] = 0, \]
which is provided by Lemma A.3 (by symmetry).

Next, we skip to prove Proposition 5, and then return to prove Theorem 4.

**Proposition 5.** \(RP_{\hat{u}}(\epsilon) = \Omega(\epsilon^2)\) as \(\epsilon \to 0\), if and only if \(A_{\hat{u}}^n(x)\) is bounded away from 0, uniformly for all \(n\) and \(x\).

**Proof.** This follows from the known fact, proven by Pratt [27], that
\[ \text{risk-prem}_{\hat{u}}^n(x) = \frac{A_{\hat{u}}^n(x)\epsilon^2}{2} + o(\epsilon^2). \]
\[ \square \]

We now proceed to prove Theorem 4.

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.4.** For any \(m > k\), and \(\epsilon < \delta\)
\[ \text{risk-prem}_{\hat{u}}^m(x + \phi_{k+1} + \cdots + \phi_m \pm \epsilon) = \text{risk-prem}_{\hat{u}}^k(x \pm \epsilon). \]

**Proof.** Set \(\beta = \text{risk-prem}_{\hat{u}}^k(x \pm \epsilon)\). By definition
\[ \hat{u}^k(x - \beta) = \frac{1}{2}(\hat{u}^k(x - \epsilon) + \hat{u}^k(x + \epsilon)). \]
Let \(a_+, a_-, a_- \in \mathcal{H}^k\) be such that \(v^k(a_{+}) = x + \epsilon, v^k(a_{-}) = x - \epsilon, \) and \(v^k(a_{-\beta}) = x - \beta.\) So,
\[ (a_{-\beta})\hat{z}^k(a_{-\epsilon}, a_{+\epsilon}). \]
By assumption, \(\hat{z}^k\) and \(\hat{z}^m\) 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,
\[ (a_{-\beta}, \phi_{k+1}, \ldots, \phi_m)\hat{z}^m((a_{-\epsilon}, \phi_{k+1}, \ldots, \phi_m), (a_{+\epsilon}, \phi_{k+1}, \ldots, \phi_m)) \]
Hence,
\[ \hat{u}^m(x - \beta + \phi_{k+1} + \cdots + \phi_m) = \frac{1}{2}(\hat{u}^m(x - \epsilon + \phi_{k+1} + \cdots + \phi_m) + \hat{u}^m(x + \epsilon + \phi_{k+1} + \cdots + \phi_m)). \]
\[ \square \]
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 $y$.

**Lemma A.5.** For any $k, K, x, y$, with $x$ in the domain of $\hat{u}^k$ and $y$ in the domain of $\hat{u}^K$, and $\epsilon < \delta$ (where $\delta$ is the bound on the distance of $\phi_i$’s from the boundaries) there exist $m \geq \max\{k, K\}$ and $b_{K+1}, \ldots, b_m$, $b, \epsilon \in \mathcal{T}_i$, with

$$\text{risk-prem}_{\hat{u}^m}(y + b_{K+1} + \cdots + b_m \pm \epsilon) = \text{risk-prem}_{\hat{u}^k}(x \pm \epsilon).$$

**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 $\eta = y' - x, j = [\eta/\delta]$, and $m = K' + j + 1$. For $i = K' + 1, \ldots, m - 1$, set $b_i = \phi_i + \eta/j$, and $b_m = \phi_m$. Then, $m > \max\{k, K\}$, and $x + \phi_{k+1} + \cdots + \phi_m = y + b_{K+1} + \cdots + b_m$. The result now follows from Lemma A.4 (we added the extra $\phi_m$ at end to guarantee sufficient distance from the boundary to allow a ±$\epsilon$ lottery). \qed

**Theorem 4.** Assuming A1-A3,

(a) If $\text{RP}_{\hat{u}}(\epsilon) = \Omega(\epsilon^2)$ as $\epsilon \to 0$ then $\hat{z}$ is SF risk averse\(^{35}\)

(b) If $\text{RP}_{\hat{u}}(\epsilon) = O(\epsilon^{2+\beta})$ as $\epsilon \to 0$, for some $\beta > 0$, then $\hat{z}$ is not SF risk averse.

**Proof.** (a) follows immediately from Theorem 3 and Proposition 5.

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

$$\text{risk-prem}_{\hat{u}^i}(x \pm \epsilon) \leq \alpha \cdot \epsilon^{2+\beta}. \tag{15}$$

Set $\epsilon_1 = \min\{\epsilon_0^2, \delta^2\}$ (where $\delta$ is the bound on the distance of $\phi_i$’s from the boundaries). For $j = 1, 2, \ldots, \eta_1 = \frac{1}{\sqrt{\epsilon_1}} \cdot \frac{1}{\sqrt{\delta}}$, set $a_j$ as follows:

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

By (15), for any $j$ there exists $i_j$ and $x_j$ with

$$\text{risk-prem}_{\hat{u}^i_j}(x_j \pm a_j) \leq \alpha \cdot a_j^{2+\beta}. \tag{16}$$

We construct a bounded, non-vanishing lottery sequence $\mathbf{L} = (L_1, L_2, \ldots)$ that is not ultimately superior to its repeated certainty equivalent, which we denote by $(c_1, c_2, \ldots)$. The construction of $\mathbf{L}$ 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 ±$a_j$ lottery, with which the chunk ends.\(^{36}\)

We denote by $n(j)$ the index of the last lottery in the $j$-th chunk. The chunks are constructed as

\(^{35}\) 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 $y < y_0$.

\(^{36}\) The infrequent but reoccurring lotteries of ±$\sqrt{\epsilon_1}$ provide that the lottery is non-vanishing.
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 (16). Set \( y_{n(j-1)} = c_1 + \cdots + c_{n(j-1)} \). By Lemma A.5 and (16), there exists \( m > \max \{ n(j-1), i_j \} \) and \( b_{n(j-1)+1}, \ldots, b_m \), with

\[
\text{risk-prem}_{\hat{a}m}(y_{n(j-1)} + b_{n(j-1)+1} + \cdots + b_m \pm a_j) \leq \alpha a_j^{2+\beta}.
\]

Accordingly, set \( L_i = b_i \), for \( i = n(j-1) + 1, \ldots, m-1 \), and \( L_m = (b_m - a_j, b_m + a_j) \). Denote \( n(j) = m \); that is, \( n(j) \) is the index of the \( \pm a_j \) lottery.

By construction, \( c_i = b_i \) for \( i = n(j-1) + 1, \ldots, m-1 \), and

\[
c_m \geq b_m - \alpha a_j^{2+\beta}.
\]  

(17)

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 \( i \); that is \( b_i = 0 \) for all \( i \). So, we have that \( L_i = ((-\sigma_i), (\sigma_i)) \) with

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

and

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

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

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

On the other hand,

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

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

$$\Pr \left[ (\ell_1, \ldots, \ell_{n(3^k)}) \preceq (c_1, \ldots, c_{n(3^k)}) \right] = \Pr \left[ S_{n(3^k)} \leq \sum_{i=1}^{n(3^k)} c_i \right] \geq$$

$$\geq \Pr \left[ S_{n(3^k)} \leq -\gamma (D + k) \right] =$$

$$= \Pr \left[ \frac{S_{n(3^k)}}{\sqrt{\text{Var}(S_{n(3^k)})}} \leq -\gamma \frac{(D + k)}{\sqrt{\epsilon_1 \cdot k}} \right] \geq$$

$$\geq \Pr \left[ \frac{S_{n(3^k)}}{\sqrt{\text{Var}(S_{n(3^k)})}} \leq -\gamma \frac{2}{\sqrt{\epsilon_1}} \right] \approx$$

$$\approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{1-\gamma^{1/2}} e^{-x^2/2} dx = p > 0,$$

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

\[ \square \]

**Theorem 6.** Assuming A1-A3, 2 and 3

(a) Weak risk loving: $\preceq$ is weakly SF risk loving if and only if all the valued-scale-utilities $\hat{u}^n$ are convex.

(b) Risk loving

- If $(-RP\hat{u}(\epsilon)) = \Omega(\epsilon^2)$ as $\epsilon \to 0$ then $\preceq$ is SF risk loving.
- If $(-RP\hat{u}(\epsilon)) = O(\epsilon^{2+\beta})$ as $\epsilon \to 0$ (for some $\beta > 0$) then $\preceq$ is not SF risk loving.

(c) Risk Neutrality: $\preceq$ 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 4 and 2. (c) follows from combining Theorems 4 and 6. \[ \square \]

**Proofs for Section 5.** The proof of Proposition 7 is easier with the aid of Theorem 8, so we start with proving the theorem and then come back to proving the proposition.

**Theorem 8.** $\preceq$ is weakly scale-free risk-averse if and only if $\hat{u}$ is concave, and (strict) scale-free risk-averse if and only if $\hat{u}$ is strictly-concave.

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

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

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

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

(18)
Since all the $L_i$'s are fair, the distribution of $v(\ell)$ is symmetric around $E(v(L))$. So,

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

So,

$$Pr[c(L) \preceq \ell] = Pr[v(c(L)) \leq v(\ell)]$$

$$\geq Pr[E[v(\ell)] \leq v(\ell)]$$

$$= Pr[E[v(\ell)] \geq v(\ell)]$$

$$\geq Pr[v(c(L)) \geq v(\ell)]$$

$$= Pr[E[v(\ell)] \geq v(\ell)]$$

$$\geq Pr[v(c(L)) \geq v(\ell)]$$

$$= Pr[c(L) \succeq \ell]$$

by (18) and (19)

**Strict Concavity $\Rightarrow$ Strict-SF-Risk-Aversion.** Consider a non-degenerate repeated lottery sequence $L = (L_1, L_2, d)$. Suppose that $L_1$ is the fair lottery $\langle a_1, a_1 \rangle$, 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

$$v_1(a_1) + v_2(a_2) = v_1(a_2) + v_2(a_1).$$

The four possible, equi-probability realizations of $L$ are:

$$(a_1, a_2, d), (a_1, a_2, d), (a_1, a_2, d), (a_1, a_2, d)$$

So,

$$E_{\ell \sim L}[(v(\ell))] = \frac{v_1(a_1) + v_1(a_1) + v_2(a_2) + v_2(a_2)}{2} + \sum_{i=3}^{n} v_i(d_i) = v(a_1, a_2, d)$$

where the second equality is by (20).

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

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

So

$$c(L) \prec (a_1, a_2, d).$$

So, of the four possible realization of $L$ only $(a_1, a_2, d)$ is $\preceq c(L)$. So,

$$Pr[c(L) \preceq \ell] = 0.75$$

$$Pr[c(L) \succeq \ell] = 0.25,$$

as necessary.

**Weak-SF-Risk-Aversion $\Rightarrow$ Concavity:** Conversely, suppose that $\hat{u}$ is not concave. Then, since $\hat{u}$ is continuous, it is strictly convex on some interval $(\underline{x}, \overline{x})$. Set $z = \frac{\overline{x} + \underline{x}}{2}$ and $\epsilon_1 = z - \underline{x}$. By
definition, there exists some \((a_1, a_2, d)\) with \(z = v(a_1, a_2, d)\). Since \(z\) is internal in \((\bar{x}, \bar{v})\), 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 \(a' < a_1 < a''\), and similarly for \(a_2\)). Choose \(\epsilon < \epsilon_1\) sufficiently small so that there exists \(a_1, a_2\), with \(v_i(a_i) = v_i - \epsilon\), and \(\bar{a}_i, \bar{a}_2\), with \(v_i(\bar{a}_i) = v_i + \epsilon\), \(i = 1, 2\). Set \(L_i = (a_i, \bar{a}_i)\), for \(i = 1, 2\). Set \(L = (L_1, L_2, d)\). Then, \(L\) is a repeated lottery, and \(\hat{u}\) is strictly convex on the domain of its realizations. So, by the reasoning as above, for “strict-concavity \(\Rightarrow\) strict-risk-aversion”,

\[
\Pr[c(L) \preceq \ell] < \Pr[c(L) \succeq \ell].
\]

So, \(u\) is not weakly risk averse, in contradiction.

**Strong-SF-Risk-Aversion \(\Rightarrow\) Strict Concavity:** Suppose \(\hat{u}\) 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 \(\Rightarrow\) weak-risk-aversion”),

\[
\Pr[c(L) \preceq \ell] \leq \Pr[c(L) \succeq \ell].
\]

So, \(\hat{u}\) is not strongly-SF-risk-averse. \(\square\)

We now return to proving Proposition 7.

**Lemma A.6.** If there exist two non-identical separable partitions \(S = A \times B\) and \(S = C \times D\), then there exist value functions \(v^A, v^B, v^C,\) and \(v^D\) (for \(A, B, C, D\)), such that

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

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

Gorman [18, Theorem 1] proves that if two separable factors \(\mathcal{E}\) and \(\mathcal{F}\) overlap then \(\mathcal{E} \cap \mathcal{F}, \mathcal{E} \cap \mathcal{F}, \mathcal{E} - \mathcal{F}, \mathcal{F} - \mathcal{E},\) and \(\mathcal{E} \Delta \mathcal{F} = (\mathcal{E} - \mathcal{F}) \cup (\mathcal{F} - \mathcal{E})\) are all separable.

Set \(\mathcal{W} = A \cap C, \mathcal{X} = A \cap D, \mathcal{Y} = B \cap C,\) and \(\mathcal{Z} = B \cap D\). Then, by Gorman’s theorem, \(\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\) are separable, 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, \(S = \mathcal{W} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}\) is a separable partition with at least 3 factors. So, by Debreu [8], there are value functions \(v^W, v^X, v^Y,\) and \(v^Z\), with \(v^W + v^X + v^Y + v^Z\) representing \(\preceq\). So, the pair of functions \(v^A = v^W + v^X\) and \(v^B = v^Y + v^Z\) are value functions for the separable partition \(S = A \times B\). Similarly, the functions \(v^C = v^W + v^Y,\) and \(v^D = v^X + v^Z\) are value functions for the separable partition \(S = C \times D\), proving (1) and (2). Finally, (3) follows from (2) by the uniqueness of value functions. \(\square\)
Proposition 7. If $\succsim^*$ is weakly-SF-risk-averse with respect to some separable partition $S = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$, then it is also weakly-SF-risk-averse with respect to any separable partition. Similarly for (strict-)SF-risk-aversion.

Proof. If there is only one separable partition then there is nothing to prove. Otherwise, by Lemma A.6 there exists a Debreu value function representing $\succsim^*$. By Lemma A.6 the aggregate value function $v$ is identical for the two partitions.

Suppose that $\succsim^*$ is SF-risk-averse with respect to one partition. Then, by Theorem 8, $\hat{\mathbf{u}}$ is concave. So, again, by the same theorem, $\succsim^*$ 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. □

Theorem 9. $\succsim^*$ is weakly-SF-risk-loving if and only if $\hat{\mathbf{u}}$ is convex, (strictly)-SF-risk-loving if and only if $\hat{\mathbf{u}}$ is strictly-convex, and SF-risk-neutral if and only if $\hat{\mathbf{u}}$ is linear.

Proof. The proof for risk-loving is analogous to that of risk-aversion. Risk neutrality then follows from the weak versions of risk-aversion and risk-loving. □

Proofs for Section 6.

Proposition 10. Factor $\mathcal{T}_i$ exhibits (strict) diminishing marginal utility if and only if $v_i$, the Debreu value function associated with factor $\mathcal{T}_i$, is (strictly) concave.

Proof. We prove the strict version. The proof for the weak version is similar.

Consider $i = 1$. Let $S = \mathcal{T}_1 \times \mathcal{T}_2$ be an additively separable partition of $S$, with $\mathcal{T}_1 \subseteq \mathbb{R}^t_+$ (we may always combine all the “other” commodities into one factor - $\mathcal{T}_2$).

Diminishing Marginal Utility $\Rightarrow$ Concavity: Contrariwise, suppose that $v_1$ is not strictly concave. Then, there exists $c_1 \in \mathcal{T}_2$ and direction $d \in \mathbb{R}^t_+$, such that for $\epsilon > 0$ sufficiently small

$$\frac{v_1(c_1 + \epsilon d) + v_1(c_1 - \epsilon d)}{2} \geq v_1(c_1).$$

That is,

$$v_1(c_1 + \epsilon d) - v_1(c_1) \geq v_1(c_1) - v_1(c_1 - \epsilon d).$$

For $\epsilon$ sufficiently small, there exists $b_2 < \bar{b}_2$ (in $\mathcal{T}_2$), with $$(c_1, b_2) \sim (c_1 - \epsilon d, \bar{b}_2).$$

So,

$$v_1(c_1) + v_2(b_2) = v_1(c_1 - \epsilon d) + v_2(\bar{b}_2).$$

Adding with (21), we get

$$v_1(c_1 + \epsilon d) + v_2(b_2) \geq v_1(c_1) + v_2(\bar{b}_2).$$
So,

$$(c_1 + \epsilon d, b_2) \succsim (c_1, b_2).$$

Setting $a_1 = c_1 - \epsilon d$, and $\alpha = \epsilon$, we have that the requirements of Definition 7 (diminishing marginal utility) do not hold.

**Concavity ⇒ Diminishing Marginal Utility:** Suppose that $v_1$ is strictly concave. We show that $T_1$ exhibits diminishing marginal utility. Consider $a_1 \in T_1, b_2, b_2 \in T_2, d \in \mathbb{R}_+^t$, and $\epsilon > 0$, with

$$(a_1 + \epsilon d, b_2) \sim (a_1, b_2).$$

So,

$$v_1(a_1 + \epsilon d) + v_2(b_2) = v_1(a_1) + v_2(b_2).$$

Since $v_1$ is strictly concave, it is strictly concave along the direction $d$ out of $a_1$. So, for any $\alpha > 0$,

$$v_1(a_1 + (\alpha + \epsilon)d) - v_1(a_1 + \epsilon d) < v_1(a_1 + \alpha d) - v_1(a_1).$$

Adding with (22)

$$v_1(a_1 + (\alpha + \epsilon)d) + v_2(b_2) < v_1(a_1 + \alpha d) + v_2(b_2).$$

So,

$$(a_1 + (\alpha + \epsilon)d, b_2) \prec (a_1 + \alpha d, b_2),$$

as required. \qed

**Proposition 11.** If factor $T_i$ exhibits (strict) diminishing marginal utility then $V_i(w_i, p_i)$ is (strictly) concave with respect to $w_i$, for any $p_i \gg 0$.

Similarly, if all factors exhibit (strict) diminishing marginal utility then $V(w, p)$ is (strictly) concave with respect to $w$, for any $p \gg 0$.

**Proof.** We prove the strict version of the second claim. The proof for the other claims is similar.

By Proposition 10 all $v_i$’s are concave. Hence $v = \sum_i v_i$ is also concave.

Let $c^*(w)$ be the optimal allocation at $(w, p)$. Consider budgets $w_- < w_+$, and intermediate budget $w_\lambda = \lambda w_- + (1 - \lambda)w_+$. Then $\lambda c^*(w_-) + (1 - \lambda)c^*(w_+)$ is a feasible bundle with budget $w_\lambda$. So,

$$V(w_\lambda, p) = v(c^*(w_\lambda))$$

$$\geq v(\lambda c^*(w_-) + (1 - \lambda)c^*(w_+))$$

$$\geq \lambda v(c^*(w_-)) + (1 - \lambda)v(c^*(w_+)) \quad \text{by concavity of } v$$

$$= \lambda V(w_-, p) + (1 - \lambda)V(w_+, p).$$

\qed
Proofs for Section 7.

Proposition 12. The value scaled utility \( \hat{u} \) exhibits DARA if and only if for any separable factor \( \mathcal{T}_i \), lottery \( L_i \) over \( \mathcal{T}_i \), and \( d_{-i} < \bar{d}_{-i} \),

\[
\left. c(L_i) \right|_{d_{-i}} \sim \left. c(L_i) \right|_{\bar{d}_{-i}}.
\]

The DARA behavior is strict if and only if the preference in (4) is strict.

Similarly, \( \hat{u} \) exhibits CARA if and only if \( c(L_i)\big|_{d_{-i}} \) is independent of \( d_{-i} \).

Proof. This essentially follows directly from Arrow [2] and Pratt [27], once we move to the \( v \) scale.

Set \( y = \sum_{j \neq i} v_j(d_j), \ y = \sum_{j \neq i} v_j(d_j), \ z = v_i(c(L_i)\big|_{d_{-i}}), \ z = v_i(c(L_i)\big|_{\bar{d}_{-i}}), \) and \( \hat{x} = v_i(L_i) \) (the random variable with \( \Pr[\hat{x} = x] = \Pr[v_i(\ell_i) = x] \)). Then

\[
\hat{u}(z + y) = \hat{u}(\hat{x} + y), \quad \hat{u}(z + y) = \hat{u}(\hat{x} + y).
\]

Suppose \( \hat{u} \) is DARA, then by Arrow-Pratt, since \( y < y \), it must be that \( \hat{z} < \hat{z} \). So, \( c(L_i)\big|_{d_{-i}} \prec c(L_i)\big|_{\bar{d}_{-i}} \). Similarly for CARA.

We now turn to proving Theorem 13. As mentioned, the theorem holds even when the preference \( \preceq \) is not additively separable (when the preference is additively separable, the theorem follows directly from Proposition 14). In this case, we do not have a Debreu value function, and the proof is somewhat more involved.

In this proof we adopt the following notation, suited for the study of a two-factor setting. The separable partition is denoted \( S = A \times B \). We use \( a, A \), and and \( b, B \), with or without subscripts or superscripts, for points in \( A \) and \( B \), respectively. By convention, \( a < A \) and \( b < B \).

We first define a restricted notion of MaPS, which we call Perfect MaPS.

Definition 9. Lottery \( L' = \langle (a_i, a_{-i}), (\bar{a}_i, \bar{a}_{-i}) \rangle \) is a Perfect-MaPS of \( L = \langle (a_i, \bar{a}_{-i}), (\bar{a}_i, a_{-i}) \rangle \) if \( (a_i, \bar{a}_{-i}) \sim (\bar{a}_i, a_{-i}) \). See Figure 4.

The difference between a MaPS and a perfect-MaPS is that in the latter the two possible outcomes of \( L \) must be equivalent. The bulk of the proof is to prove that a decision maker dislikes all MaPS if and only if she dislikes all perfect-MaPS.

Let \( w^A : A \to \mathbb{R} \) be a continuous real valued function representing \( \preceq^A \), and similarly \( w^B \) a continuous real function representing \( \preceq^B \). We do not require that \( w \) be additive across the factors. So, it exist by Debreu [7] since \( \preceq^A \) and \( \preceq^B \) are continuous. Define \( w : A \times B \to \mathbb{R}^2 \) as \( w(a, b) = (w^A(a), w^B(b)) \). Let \( I_A \times I_B \subseteq \mathbb{R}^2 \) be the image of \( A \times B \) under \( w \).

Lemma A.7. \( u \circ w^{-1} : I_A \times I_B \to \mathbb{R} \) is well defined, increasing in each coordinate, and continuous.

\(^{13} \preceq^A \) and \( \preceq^B \) exist since \( S = A \times B \) is a separable partition, though not necessarily additively separable.
Figure 4. Illustration of a perfect MaPS: \( L' \) (the solid circles) is a perfect MaPS of \( L \) (the empty circles). In \( L \) the two possible outcomes are equivalent.

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

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

To prove continuity of \( \hat{u} \), we prove that the pre-images of the open rays \(( -\infty, r ) \) and \(( r, \infty ) \) 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 \( E_r = I_A \times I_B \) 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 \( I_A \) and \( y^* \) not maximal in \( I_B \) (the proof for the case that one of them is maximal is similar). The function \( \hat{u}^B_x \) is continuous. So, there exists some \( y' \) with

\[
0 < \hat{u}^B_x(y') - \hat{u}^B_x(y^*) < \frac{1}{2} \epsilon.
\]

Similarly, the function \( \hat{u}^A_y \) is continuous. Thus, there exists \( x' \) with

\[
0 < \hat{u}^A_y(x') - \hat{u}^A_y(x^*) < \frac{1}{2} \epsilon.
\]

Combining (23) and (24), we obtain

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

Set \( \delta = \min\{ x' - x^*, y' - y^* \} \). Then, for any \((x, y)\) if \( \|(x, y) - (x^*, y^*)\| < \delta \) then \( x < x' \) and \( y < y' \). So, by monotonicity of \( \hat{u} \), \( \hat{u}(x, y) < \hat{u}(x', y') < r \). So, the entire ball of size \( \delta \) around \(( x^*, y^* ) \) is contained in \( E_r \), as required.
Lemma A.8. Let $A \times B$ be a separable partition and $a \prec A$, $b \prec B$. Set $a^0 = a$, and while $(a^i, B) \succcurlyeq (A, b)$ 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 $i$ such that $(a^i, B) \succeq (A, b)$ (that is, the sequence $a^0, a^1, \ldots$ is finite).

Proof. Contrariwise, suppose there is no such $i$. Then, for $i = 1, 2, \ldots$, $(a^i, B) \prec (A, b)$, and hence $a^i \prec A$. Clearly, $a^i \preceq a^{i+1}$. 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 $b \prec B$ and $\succcurlyeq$ is strictly monotone in each factor. □

We are now ready to prove the main technical lemma, which establishes that dislike of perfect-MaPS implies dislike of all MaPS.

Lemma A.9. Suppose that $L' \succcurlyeq L$ whenever $L'$ is a perfect-MaPS of $L$, then $L' \succcurlyeq L$ whenever $L'$ is a MaPS of $L$. Similarly for weak preferences.

Proof. We prove for the strict case. The weak case is similar.

Suppose that $L' \succcurlyeq L$ whenever $L'$ is a perfect-MaPS of $L$. Let $a, A \in A, b, B \in B$, with $a \prec A$ and $b \prec B$. We need to show that

$$(a, b, (A, B)) \sim ((A, b), (a, B)).$$

(25)

If $(a, B) \sim (A, b)$ then $L'$ is a perfect-MaPS of $L$ and (25) follows by assumption.

Otherwise, let $u$ be an NM utility for $\succcurlyeq$. Set

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

We show that $\text{diff} < 0$, which establishes (25).

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

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

such that

$$w^A(A_{1 \over 2}) - w^A(a_{1 \over 2}) \leq \frac{1}{2} (w^A(A) - w^A(a)) \quad \text{or}$$

(26)

$$w^B(B_{1 \over 2}) - w^B(b_{1 \over 2}) \leq \frac{1}{2} (w^B(B) - w^B(b))$$

and

(27) $$\text{diff} < u(a_{1 \over 2}, b_{1 \over 2}) + u(A_{1 \over 2}, B_{1 \over 2}) - u(a_{1 \over 2}, B_{1 \over 2}) - u(A_{1 \over 2}, b_{1 \over 2}).$$
Figure 5. Illustration of the proof of Lemma A.9. 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 $w^A(a^2) \geq \frac{1}{2}(w^A(A) + w^A(a))$ (assuming the picture is scaled according to $w^A$).

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

\[(a^1, b) \sim (a, B).\]  

Figure 5 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.8. Then, $(a, B) \prec (A, b) \succsim (a^\bar{i}, B)$. Thus, there exists $A^1, a \prec A^1 \prec a^\bar{i}$, such that $(A^1, B) \sim (A, b)$. Clearly, $a^\bar{i} \succsim A$. Thus, either

\[(29) \quad w^A(A^1) \leq \frac{1}{2}(w^A(a) + w^A(A)),\]

or

\[(30) \quad w^A(a^\bar{i}) \geq \frac{1}{2}(w^A(a) + w^A(A)).\]

We consider each of these cases separately.

First, suppose that (29) holds. Then, by construction $(A^1, B) \sim (A, b)$. Hence, $\langle (A^1, b), (A, B) \rangle$ is a perfect-MaPS of $\langle (A^1, B), (A, b) \rangle$, and, by assumption

\[\langle (A^1, b), (A, B) \rangle \prec \langle (A^1, B), (A, b) \rangle.\]

So,

\[u(A^1, b) + u(A, B) - u(A^1, B) - u(A, b) < 0.\]
Hence,

\[ u(a, b) + u(A, B) - u(A, b) - u(a, B) = \]
\[ u(a, b) + u(A^1, B) - u(A^1, b) - u(a, B) + u(A^1, b) + u(A, B) - u(A^1, B) - u(A, b) < \]
\[ u(a, b) + u(A^1, B) - u(A^1, b) - u(a, B). \]

(31)

Setting \( a_{\frac{1}{2}} = a, A_{\frac{1}{2}} = A^1, b_{\frac{1}{2}} = b \) and \( B_{\frac{1}{2}} = B \), by (29) and (31) we get (26) and (27).

Next, suppose that (30) holds. Then, by construction, for \( i = 1, \ldots, \bar{i} \), \( (a^{i-1}, B) \sim (a^i, b) \). So, since \( \tilde{\prec} \) dislikes all perfect-MaPS,

\[ \langle (a^{i-1}, b), (a^i, B) \rangle \prec \langle (a^{i-1}, B), (a^i, b) \rangle, \]

for all \( i \). So,

(32)

\[ \frac{1}{2^i} \sum_{i=1}^{\bar{i}} (u(a^{i-1}, b) + u(a^i, B)) < \frac{1}{2^i} \sum_{i=1}^{\bar{i}} (u(a^{i-1}, B) + u(a^i, b)); \]

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,

\[ u(a, b) + u(A, B) - u(A, b) - u(a, B) = \]
\[ u(a, b) + u(a^\bar{i}, B) - u(a^\bar{i}, b) - u(a, B) + u(a^\bar{i}, b) + u(A, B) - u(a^\bar{i}, B) - u(A, b) < \]
\[ u(a^\bar{i}, b) + u(A, B) - u(a^\bar{i}, B) - u(A, b). \]

(33)

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 (30) and (33) we get (26) and (27).

Thus, we have established (26) and (27), and we now return to complete the proof that \( \text{diff} < 0 \).

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

(34) \[ w^A(A_\delta) - w^A(a_\delta) \leq \delta \quad \text{or} \]

(35) \[ w^B(B_\delta) - w^B(b_\delta) \leq \delta \]
By Lemma A.7 the function \( u \circ (w^A, w^B)^{-1} \) is continuous. So it is uniformly continuous on the rectangle \([w^A(a), w^A(A)] \times [w^B(b), w^B(B)]\). That is, for any \( \epsilon > 0 \), there exists a \( \delta \) such that if

\[
\|(w^A(a'), w^B(b')) - (w^A(a''), w^B(b''))\| < \delta
\]

then

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

In particular, if (34) holds then (36) is \( \leq 2\epsilon \), and if (35) holds then (37) is \( \leq 2\epsilon \). Thus, \( \text{diff}_2 \leq 2\epsilon \) for all \( \epsilon > 0 \). Hence, \( \text{diff}_2 \leq 0 \). So diff < 0.

**Theorem 13.** \( \preceq \) is weakly scale-free risk averse if and only if \( L \preceq L' \) whenever \( L' \) is MaPS of \( L \). Similarly for (strict) scale-free risk aversion and strict preference.

**Proof.** We prove the weak case. The proof for strict case is similar.

By Lemma A.9 it suffices to prove the claim with perfect MaPS instead of MaPS.

Consider \( a \succ A, b \prec B \), with \( (a, B) \sim (A, b) \). Set \( L = \langle (a, B), (A, b) \rangle, L' = \langle (a, b), (A, B) \rangle \). The four possible, equi-probable realizations of \( L \) are

\[
(a, b) \prec (a, B) \sim (A, b) \prec (A, B).
\]

So,

\[
\Pr[c(L) \preceq \ell] \geq \Pr[c(L') \preceq \ell] \iff c(L) \preceq (a, B) \iff u(L) \leq u(a, B) \iff \frac{u(a, b) + u(a, B) + u(B, a) + u(A, B)}{4} \leq u(a, B) \iff u(a, b) + u(A, B) \leq u(a, B) + u(A, b) \iff L \preceq L'
\]

as required. \( \Box \)

**Proposition 14.** If \( L' \) is a MaPS of \( L \) then \( v(L') \) is a MePS of \( v(L) \).

Conversely, for any sufficiently small fair lottery \( \tilde{y} = \langle \underline{y}, \overline{y} \rangle \) in the rage of \( v \) if \( \tilde{y} \) is a MePS of \( \tilde{x} = \langle \underline{x}, \overline{x} \rangle \) then there exist \( L, L' \in \Delta(S) \), with \( v(L) = \tilde{x}, v(L') = \tilde{y} \), and \( L' \) a MaPS of \( L \).

\[\text{formally: there exists a constant } \alpha > 0 \text{ (depending on the partition of } S\text{), such that for any } \tilde{y} = \langle \underline{y}, \overline{y} \rangle \text{ with } |\overline{y} - \underline{y}| \leq \alpha \ldots\]
Proof. By definition, for fair lotteries \( \tilde{x} = (x, \pi) \), \( x \leq \pi \), and \( \tilde{x}' = (x', \pi') \), \( \tilde{x}' \) is a MePS of \( \tilde{x} \) if and only if \( x' < x \leq \pi < \pi' \) and \( x + \pi = \pi' + x' \). Consider an additively separable partition \( S = A \times B \), with Debreu value functions \( v^A, v^B \).

Let \( L' = \langle (a, b), (A, B) \rangle \) be a MaPS of \( L = \langle (a, B), (A, b) \rangle \). So, \( a < A \) and \( b < B \). So,

\[
v(L) = \langle v^A(a) + v^B(b), v^A(A) + v^B(B) \rangle
\]

is a MePS of

\[
v(L) = \langle v^A(a) + v^B(b), v^A(A) + v^B(B) \rangle.
\]

Conversely. Choose some \( \alpha > 0 \), with \( \alpha < 1/2 \cdot \min\{|v^A(A)|, |v^B(B)|\} \). Consider \( \tilde{y} = \langle y, \overline{y} \rangle \) a MePS of \( \tilde{x} = (x, \pi) \), with \( y < x \leq \pi < \overline{y} \), all of which are in \( v(S) \), and \( \overline{y} - y \leq \alpha \). Set \( \delta_1 = x - y, \delta_2 = \overline{y} - x \). So, \( \delta_1 + \delta_2 = \overline{y} - y \).

Since \( y, \overline{y} \in v(S) \), there exits \( a, \pi \in v^A(A), b, \overline{b} \in v^B(B) \), with \( y = v^A(a) + v^B(b), \overline{y} = v^A(\pi) + v^B(\overline{b}) \). Since \( \delta_1 + \delta_2 = \overline{y} - y < \alpha < 1/2 \cdot |v^A(A)| \), then either \( v^A(a) + \delta_1 + \delta_2 \in v^A(A) \), or \( v^A(\pi) - \delta_1 - \delta_2 \in v^A(A) \). Suppose the former. The other case is analogous.

Now, either \( v^A(b) + \delta_2 \in v^B(B) \), or \( v^B(b) - \delta_2 \in v^B(B) \). Consider both cases:

- \( v^B(b) + \delta_2 \in v^B(B) \): then there exist \( A, B \), with \( \overline{v^A(A)} = v^A(a) + \delta_1, v^B(B) = v^B(b) + \delta_2 \). Set \( a = a, b = b, \) and \( L = \langle (a, b), (A, B) \rangle, L' = \langle (a, b), (A, B) \rangle \).

- \( v^A(b) - \delta_2 \in v^A(B) \): then there exist \( A, a, b \), with \( \overline{v^A(A)} = v^A(a) + \delta_1 + \delta_2, v^A(a) = v^A(a) + \delta_2, v^B(b) = v^B(b) - \delta_2 \). Set \( B = b, L = \langle (A, a, (a, B)), L' = \langle (a, (a, B)) \rangle \).

In both cases, \( L' \) is a MaPS of \( L \), and \( v(L') = \tilde{y}, v(L) = \tilde{x} \).

Corollary [15] follows directly from Theorems [8] and [13].

Proofs for Section 8.

**Proposition 16.** For \( z \) non-random, \( s^*(w, z) \) increases with \( w \).

**Proof.** Fix \( z \). With savings \( s \), the resulting Debreu value is:

\[
V_1(w - s) + V_2(zs).
\]

So, the first order conditions for \( s^* \) are

\[
V'_1(w - s^*) = zV'_2(zs).
\]

By Proposition [11], the left-hand side decreases with \( w \). So, the right-hand side must also decrease. This is only possible if \( s \) increases.

**Lemma A.10.** If \( R_{V_2} \leq 1 \) then \( zV'_2(zs) \) increases in \( z \) for all \( s \). Similarly, if \( R_{V_2} \geq 1 \) then \( zV'_2(zs) \) decreases in \( z \) for all \( s \). If the inequality is strict than the increasing/decreasing is so.
Proof. We prove the (weak) increasing case. The other cases are similar.

\[ zV'_2(zs) \text{ increases in } z \iff V'_2(zs) + zsV''_2(zs) \geq 0 \iff \text{since } V_2 \text{ is increasing } \]
\[ \frac{V'_2(zs) + zsV''_2(zs)}{V'_2(zs)} \geq 0 \iff 1 \geq -zsV''_2(zs) = \mathcal{R}_{V_2}(zs) \]

\[ \square \]

Proposition 17. For \( z \) non-random

- \( s^*(w, z) \) increases with \( z \) if \( \mathcal{R}_{V_2} \leq 1 \).
- \( s^*(w, z) \) decreases with \( z \) if \( \mathcal{R}_{V_2} \geq 1 \).

Proof. We prove the first bullet. The second is similar. As above, the first order condition for optimality of \( s^*(w, z) \) is

\[ V'_1(w - s^*) = zV'_2(zs^*) \]

The left-hand side is independent of \( z \). By assumption and Lemma A.10, the right-hand side increases with \( z \). So, in order to retain equality, we must increase \( s^* \), which, by Proposition 11 increases the term \( V'_1(w - s^*) \) and decreases the term \( V'_2(zs^*) \).

When \( \tilde{z} \) is random, let \( z_1 < z_2 < \cdots < z_n \) be the support of \( \tilde{z} \), and \( p_i = \Pr[\tilde{z} = z_i] \). For readability, we sometimes shorthand \( \hat{u}(w_1, w_2) \) for \( \hat{u}(v_1(w_1) + v_2(w_2)) \).

Given budget \( w \), and savings level \( s \), the expected utility is

\[ u(w - s, \tilde{z}s) = \sum_{i=1}^{n} p_i \hat{u}(V_1(w - s) + V_2(z_is)) \]

Differentiating by \( s \),

\[ \frac{\partial u(w - s, \tilde{z}s)}{\partial s} = \sum_{i=1}^{n} p_i \frac{\partial \hat{u}(V_1(w - s) + V_2(z_is))}{\partial s} \]
\[ = \sum_{i=1}^{n} p_i \cdot (V'_1(w - s) + z_i V'_2(z_i s)) \cdot \hat{u}'(V_1(w - s) + V_2(z_is)) \]

The optimal \( s^*(w) \) is obtained either at the boundaries, or when \( \frac{\partial u(w - s, \tilde{z}s)}{\partial s} = 0 \). By assumption, we are in the latter case. Since \( V_1, V_2 \), and \( \hat{u} \) are concave, \( \frac{\partial u(w - s, \tilde{z}s)}{\partial s} \) is decreasing, and there exists a unique solution to \( \frac{\partial u(w - s, \tilde{z}s)}{\partial s} = 0 \).

The following is simple variant of Chebyshev’s sum inequality.
Lemma A.11. 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, if
\[
\sum_{i=1}^{n} \alpha_i y_i = 0,
\]
then
\[
\sum_{i=1}^{n} \alpha_i \beta_i y_i \geq 0.
\]
and equality holds if and only if either all the $\alpha_i$’s or all the $\beta_j$’s are identical. If the $\beta_j$’s are decreasing, then the inequality is reversed.

Proof. We prove the the case that both sequences $\alpha_i$ and $\beta_i$ are increasing. 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,
\[
\sum_{j=1}^{n} \sum_{i=1}^{n} y_j y_i (\alpha_i - \alpha_j)(\beta_i - \beta_j) \geq 0,
\]
and equality holds if and only if either all the $\alpha_i$’s or all the $\beta_j$’s are identical. Opening the brackets we get,
\[
2 \sum_{j=1}^{n} \left( y_j \sum_{i=1}^{n} y_i \alpha_i \beta_i \right) - 2 \sum_{j=1}^{n} \left( \beta_j y_j \sum_{i=1}^{n} \alpha_i y_i \right) \geq 0,
\]
which, after rearrangement gives
\[
2 \left( \sum_{j=1}^{n} y_j \right) \left( \sum_{i=1}^{n} y_i \alpha_i \beta_i \right) \geq 2 \left( \sum_{j=1}^{n} \beta_j y_j \right) \left( \sum_{i=1}^{n} \alpha_i y_i \right) = 0,
\]
where the last equality is by assumption. Since all $y_j$’s are positive, the result follows. \qed

Proposition 18. If $E[\tilde{z}] > \frac{V_1'(w)}{V_2'(0)}$ then $s^*(w, \tilde{z}) > 0$.

Proof. At $s = 0$, we have
\[
\frac{\partial u(w - s, \tilde{z}s)}{\partial s} \bigg|_{s=0} = \sum_{i=1}^{n} p_i \cdot (-v'_1(w) + z_i v'_2(0)) \cdot \hat{u}'(v_1(w) + v_2(0))
\]
\[
= \left( -v'_1(w) + v'_2(0) \sum_{i=1}^{n} p_i z_i \right) \cdot \hat{u}'(v_1(w) + v_2(0))
\]
\[
= (-v'_1(w) + v'_2(0)E[\tilde{z}]) \cdot \hat{u}'(v_1(w) + v_2(0)).
\]
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. \qed

Theorem 19. Assuming $V_1, V_2, \hat{u}$, are concave. If

- $\hat{u}$ is DARA and $\mathcal{R}_V \leq 1$, or
• \( \hat{u} \) is IARA and \( R_{V_2} \geq 1 \),

then \( s^*(w) \) is increasing. If, in addition, either: the above are strictly so or \( V_1 \) is strictly concave, then \( s^*(w) \) is strictly increasing.

Proof. We prove the first set of conditions. The other cases are similar.

Consider budgets \( w_{\oplus} > w_{\ominus} \), and let \( s^*_{\oplus}, s^*_{\ominus} \) be the optimal saving levels at \( w_{\oplus}, w_{\ominus} \), respectively. For \( w, z, s \), denote \( x_{w,z,s} = V_1(w - s) + V_2(zs) \). The first order conditions for \( s^*_{\ominus} \) are

\[
\sum_{i=1}^{n} p_i \cdot (-V'_1(w_{\ominus} - s^*_{\ominus}) + z_i V'_2(z_i s^*_{\ominus})) \cdot \hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}}) = 0.
\]

Denote:

• \( \alpha_i = -V'_1(w_{\ominus} - s^*_{\ominus}) + z_i V'_2(z_i s^*_{\ominus}) \),

• \( \beta_i = \frac{\hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}})}{\hat{u}(x_{w_{\ominus},z_i,s^*_{\ominus}})} \),

• \( y_i = p_i \hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}}) \).

We show that the conditions of Lemma \[A.11\] holds for the \( \alpha_i \)’s, \( \beta_i \)’s and \( y_i \)’s.

Since \( R_{V_2} \leq 1 \), by Lemma \[A.10\], \( zV'_2(zs^*) \) increases with \( z \), so the \( \alpha_i \) are increasing. The \( y_i \) are positive since \( \hat{u} \) is monotone increasing.

For the \( \beta_i \)’s, we show that \( \frac{\hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}})}{\hat{u}(x_{w_{\ominus},z_i,s^*_{\ominus}})} \) increases with \( z \):

\[
\frac{\partial}{\partial z} \left( \frac{\hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}})}{\hat{u}(x_{w_{\ominus},z_i,s^*_{\ominus}})} \right) = s^*_{\ominus} V'_2(zs^*) \cdot \left( \frac{\hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}}) \cdot \hat{u}(x_{w_{\ominus},z_i,s^*_{\ominus}}) - \hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}}) \cdot \hat{u}(x_{w_{\ominus},z_i,s^*_{\ominus}})}{(\hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}}))^2} \right) \\
= s^*_{\ominus} V'_2(zs^*) \cdot \left( \frac{\hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}})}{\hat{u}(x_{w_{\ominus},z_i,s^*_{\ominus}})} \right) \cdot \left( \frac{\hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}}) \cdot \hat{u}(x_{w_{\ominus},z_i,s^*_{\ominus}}) - \hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}}) \cdot \hat{u}(x_{w_{\ominus},z_i,s^*_{\ominus}})}{(\hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}}))^2} \right) \geq 0,
\]

where that last inequality is since \( \hat{u} \) is DARA (and \( \hat{u}', V'_2 \) positive).

So, applying Lemma \[A.11\] to (40), we have

\[
0 \leq \sum_{i=1}^{n} p_i \cdot (-V'_1(w_{\ominus} - s^*_{\ominus}) + z_i V'_2(z_i s^*_{\ominus})) \cdot \hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}}) \cdot \frac{\hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}})}{\hat{u}(x_{w_{\ominus},z_i,s^*_{\ominus}})}
\]

\[
= \sum_{i=1}^{n} p_i \cdot (-V'_1(w_{\ominus} - s^*_{\ominus}) + z_i V'_2(z_i s^*_{\ominus})) \cdot \hat{u}'(x_{w_{\ominus},z_i,s^*_{\ominus}})
\]

(42)

(and strict equality holds when \( R_{V_2} < 1 \) and \( \hat{u} \) strictly DARA).

Since \( V_1 \) is concave

\[-V'_1(w_{\ominus} - s^*_{\ominus}) \leq -V'_1(w_{\ominus} - s^*_{\ominus}).\]
Theorem 21. Assuming $V_1, V_2, \hat{u}$ are concave. If the following three hold

1. the elasticity of $\hat{u}'$ of with respect to $z$, $\frac{\partial \ln(\hat{u}'(V_1(w(1-r))+V_2(zwr)))}{\partial \ln(z)}$, is decreasing in $w$,
2. $R_{V_2} \leq 1$,
3. the marginal rate of substitution $\frac{V_2'(w c_1)}{V_2'(w c_2)}$ is increasing in $w$ (for all $c_1, c_2$),

then $r^*(w)$ decreases with $w$. If either (1)+(2) or (3) are strictly so, then $r^*(w)$ is strictly decreasing. If (3) is reversed, and either (1) or (2) (but not both) also reversed, then $r^*(w)$ increases with $w$, and strictly so if the behavior in (1)+(2) or (3) is strict.

Proof. We prove for the case where (1)-(3) hold as is. The other cases are similar. The proof is along the same lines as that of Theorem 19.

Consider budgets $w_{\ominus} > w_{\ominus}$. Set $x_{w,z,r} = V_1(w(1-r)) + V_2(zwr)$. The first order conditions for optimality of $r^*_{\ominus}$ are

$$0 \leq \sum_{i=1}^{n} p_i \cdot (-V_1'(w_{\ominus}(1-r^*_{\ominus})) + z_i V_2'(z_i w_{\ominus} r^*_{\ominus})) \cdot \hat{u}'(x_{w_{\ominus},z_i,r^*_{\ominus}}) = 0.$$ 

Denote:

- $\alpha_i = -V_1'(w_{\ominus}(1-r^*_{\ominus})) + z_i V_2'(z_i w_{\ominus} r^*_{\ominus})$,
- $\beta_i = \frac{\hat{u}'(x_{w_{\ominus},z_i,r^*_{\ominus}})}{w'(x_{w_{\ominus},z_i,r^*_{\ominus}})}$,
- $y_i = p_i \cdot \hat{u}'(x_{w_{\ominus},z_i,r^*_{\ominus}})$.

We show that the conditions of Lemma A.11 holds for the $\alpha_i$’s, $\beta_i$’s and $y_i$’s.

Since $R_{V_2} \leq 1$, $z V_2'(z s^*_{\ominus})$ increases with $z$, so the $\alpha_i$’s are increasing. The $y_i$’s are positive since $\hat{u}$ is monotone increasing.

As for the $\beta_i$’s, we show that $\frac{\hat{u}'(x_{w_{\ominus},z,r})}{w'(x_{w_{\ominus},z,r})}$ increases with $z$ (for all $r$). By Condition (1) the following term decreases with $w$:

$$\frac{\partial \ln(\hat{u}'(x_{w,z,r}))}{\partial \ln(z)} = \frac{\hat{u}''(x_{w,z,r}) \cdot V_2'(zw_{\ominus}r) \cdot zw_{\ominus} \cdot \hat{u}'(x_{w,z,r}) - \hat{u}'(x_{w,z,r}) \cdot \hat{u}''(x_{w,z,r}) \cdot V_2'(zw_{\ominus}r) \cdot zw_{\ominus}}{\left(\hat{u}'(x_{w,z,r})\right)^2} =$$

$$\frac{1}{r} \cdot \hat{u}'(x_{w,z,r}) \cdot \left(\frac{\hat{u}''(x_{w,z,r})}{\hat{u}'(x_{w,z,r})} \cdot V_2'(zw_{\ominus}r) \cdot zw_{\ominus} - \frac{\hat{u}''(x_{w,z,r})}{\hat{u}'(x_{w,z,r})} \cdot V_2'(zw_{\ominus}r) \cdot zw_{\ominus} \right) > 0$$
where that last inequality is by $(44)$ (and $\hat{u}'$ positive).

So, with these $\alpha_i$, $\beta_i$, $y_i$, by $(43)$, applying Lemma A.11 we have

$$0 \leq \sum_{i=1}^{n} p_i \cdot (-V'_1(w(1-r^*_i)) + z_i V'_2(z_i w r^*_2)) \cdot \hat{u}'(x_{w, z_i, r^*_2}) \cdot \frac{\hat{u}'(x_{w, z_i, r^*_2})}{\hat{u}'(x_{w, z_i, r^*_2})}$$

$$= \sum_{i=1}^{n} p_i \cdot (-V'_1(w(1-r^*_i)) + z_i V'_2(z_i w r^*_2)) \cdot \hat{u}'(x_{w, z_i, r^*_2}),$$

(and strict equality holds when $R_{V^2} < 1$). So, multiplying by $\frac{V_1(w(1-r^*_i))}{V_1(w(1-r^*_i))}$,

$$0 \leq \sum_{i=1}^{n} p_i \cdot \left(-V'_1(w(1-r^*_i)) + z_i V'_2(z_i w r^*_2) \cdot \frac{V_1(w(1-r^*_i))}{V_1(w(1-r^*_i))}\right) \cdot \hat{u}'(x_{w, z_i, r^*_2})$$

By Condition (3)

$$\frac{V_1(w(1-r^*_i))}{V_1(w(1-r^*_i))} \leq \frac{V_2(z_i w r^*_2)}{V_2(z_i w r^*_2)}$$

for all $z_i$. So, from $(45)$,

$$0 \leq \sum_{i=1}^{n} p_i \cdot \left(-V'_1(w(1-r^*_i)) + z_i V'_2(z_i w r^*_2) \cdot \frac{V_2(z_i w r^*_2)}{V_2(z_i w r^*_2)}\right) \cdot \hat{u}'(x_{w, z_i, r^*_2})$$

$$= \sum_{i=1}^{n} p_i \cdot (-V'_1(w(1-r^*_i)) + z_i V'_2(z_i w r^*_2)) \cdot \hat{u}'(x_{w, z_i, r^*_2})$$

So, for $r^*_i$ to meet the first order conditions it must be that $r^*_i \geq r^*_i$. □

**Corollary 22.** Assuming $V_1, V_2, \hat{u}$ are concave. If the certainty preferences exhibit constant elasticity of substitution (CES) with elasticity $\sigma \neq 1$, then

- if $\hat{u}$ is DRRA then $r^*(w)$ increases with $w$.
- if $\hat{u}$ is IRRA then $r^*(w)$ decreases with $w$.

If the certainty preference are Cobb-Douglas (CES, $\sigma = 1$) with (ordinal) utility representation $U(c_1, c_2) = c_1^{\beta_1} c_2^{\beta_2}$, then $r^* = \frac{\beta_1}{\beta_1 + \beta_2}$ for all $\hat{u}$, $w$, and $\tilde{z} \neq 0$.

**Proof.** For elasticity $\sigma \neq 1$, the certainty preferences are represented by the (ordinal) utility function:

$$U(c_1, c_2) = \left(\beta_1 c_1^{d_1} + \beta_2 c_2^{d_2}\right)^{1/d},$$

for $d = (\sigma - 1)/\sigma$. So, $V_i(c_i) = \beta_i c_i^{d_i}$, are Debreu value functions representing the preference $(i = 1, 2)$. Since $V_1, V_2$ are concave $0 < d \leq 1$.

Suppose that $\hat{u}$ is IRRA. The other case is similar. We show that the requirements of Theorem 21 hold.
Condition (1). Denote \( x_{w,z,r} = V_1(w(1-r)) + V_2(zwr) = \beta_1(w(1-r))^d + \beta_2(zwr)^d \). Then

\[
\frac{\partial \ln(\hat{u}'(x_{w,z,r}))}{\partial \ln(z)} = \frac{\hat{u}''(x_{w,z,r})}{\hat{u}'(x_{w,z,r})} \cdot \beta_2 \cdot d(zwr)^{d-1} \cdot wr \cdot z = 
\left( \frac{\hat{u}''(x_{w,z,r})}{\hat{u}'(x_{w,z,r})} \cdot x_{w,z,r} \right) \frac{\beta_2(zr)^d}{\beta_1(1-r)^d + \beta_2(zr)^d}.
\]

So, this expression decreases with \( w \) when \( \hat{u} \) is IRRA (since the coefficient of relative risk aversion is the negative of the first term, and the second term is independent of \( w \)).

Condition (2).

\[
V_2'(x) \cdot x = \beta_2 d \cdot x^{d-1} \cdot x
\]

which is increasing for \( d > 0 \).

Condition (2).

\[
\frac{v'_1(wc_1)}{v'_2(wc_2)} = \frac{d\beta_1 \cdot w^{d-1} c_1^{d-1}}{d\beta_2 \cdot w^{d-1} c_2^{d-1}}
\]

which is constant.

So, the conditions of Theorem 21 hold and \( r^* \) decreases with \( w \).

Next, consider Cobb-Douglas preferences

\[
U(c_1, c_2) = c_1^{\beta_1} c_2^{\beta_2}.
\]

So,

\[
v(c_1, c_2) = \beta_1 \ln(c_1) + \beta_2 \ln(c_2),
\]

is an additive representation of the preferences. So, \( v_i(c_i) = \beta_i \ln(c_i) \) are the Debreu value functions.

Denote \( x_{w,z,r} = \beta_1 \ln(w(1-r)) + \beta_2 \ln(zwr) \). The first order conditions for \( r^* \) are:

\[
0 = \sum_{i=1}^n p_i \cdot \left( -\frac{\beta_1 w}{w(1-r^*)} + \frac{\beta_2 wz}{zwr^*} \right) \cdot \hat{u}'(x_{w,z,r^*}) = \left( -\frac{\beta_1}{1 - r^*} + \frac{\beta_2}{r^*} \right) \sum_{i=1}^n p_i \cdot \hat{u}'(x_{w,z,r^*}).
\]

So, conditions hold if and only if \( r^* = \frac{\beta_2}{\beta_1 + \beta_2} \) (as all addends on the summation are positive). \( \square \)

**Proposition 23.** Assuming \( V_1, V_2, \hat{u}_\ominus, \hat{u}_\boxplus \), are concave. If \( A_{\hat{u}_\ominus}(x) \geq A_{\hat{u}_\boxplus}(x) \), for all \( x \), then

- if \( R_{V_2} \leq 1 \) then \( s^*_\ominus \leq s^*_\boxplus \).
- if \( R_{V_2} \geq 1 \) then \( s^*_\ominus \geq s^*_\boxplus \).

In both cases, the inequality is strict if the respective conditions are so.

**Proof.** We prove the first case. The other cases are similar. The proof is similar to that of Theorem 19.
Denote \( x_{w,z,s} = V_1(w - s) + V_2(zs) \). The first order conditions for \( s^*_\bigcirc \) are

\[
(46) \quad \sum_{i=1}^{n} p_i \cdot (-V'_1(w - s^*_\bigcirc) + z_iV'_2(z_is^*_\bigcirc)) \cdot \hat{u}_\bigcirc(x_{w,z_i,s^*_\bigcirc}) = 0.
\]

Denote:

- \( \alpha_i = -V'_1(w - s^*_\bigcirc) + z_iV'_2(z_is^*_\bigcirc) \),
- \( \beta_i = \frac{\hat{u}'_\bigcirc(x_{w,z_i,s^*_\bigcirc})}{\hat{u}_\bigcirc(x_{w,z_i,s^*_\bigcirc})} \),
- \( y_i = p_i\hat{u}'_\bigcirc(x_{w,z_i,s^*_\bigcirc}) \).

The \( \alpha_i \)'s and \( \beta_i \)'s are increasing and the \( y_i \)'s positive as in the proof of Theorem 19.

So, applying Lemma \ref{lemma:A.11} to \( (46) \), we have

\[
0 \leq \sum_{i=1}^{n} p_i \cdot (-V'_1(w - s^*_\bigcirc) + z_iV'_2(z_is^*_\bigcirc)) \cdot \hat{u}_\bigcirc(x_{w,z_i,s^*_\bigcirc}) \cdot \frac{\hat{u}'_\bigcirc(x_{w,z_i,s^*_\bigcirc})}{\hat{u}_\bigcirc(x_{w,z_i,s^*_\bigcirc})} = \sum_{i=1}^{n} p_i \cdot (-V'_1(w - s^*_\bigcirc) + z_iV'_2(z_is^*_\bigcirc)) \cdot \hat{u}'_\bigcirc(x_{w,z_i,s^*_\bigcirc}).
\]

So, in order for \( s^*_\bigcirc \) to meet the first order conditions with \( \hat{u}_\bigcirc \) it must be that \( s^*_\bigcirc \geq s^*_\bigcirc \). \( \square \)

**Appendix B. Unbounded Lottery Sequences**

Here we show why in Definition \ref{dfn: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 \( \succsim = (\succsim^1, \succsim^2, \ldots) \), there exists a lottery sequence that is ultimately inferior to its repeated certainty equivalent.

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

Let \( u^n \) be an NM utility representing \( \succsim^n \). For each \( n \), let \( b_n \) be such that

\[
2^{-n} \cdot u^n(0, \ldots, 0, b_n) + (1 - 2^{-n}) u^n(0, \ldots, 0, -1) = u^n(0, \ldots, 0).
\]

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,

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

and hence

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