A CONCEPTUAL FOUNDATION FOR
THE THEORY OF RISK AVERSION

By

YONATAN AUMANN

Discussion Paper # 686   (June 2015)

THE FEDERMAN CENTER FOR
THE STUDY OF RATIONALITY

Feldman Building, Edmond J. Safra Campus,
Jerusalem 91904, Israel
PHONE: [972]-2-6584135    FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/
A CONCEPTUAL FOUNDATION FOR THE THEORY OF RISK AVERSION
(WORKING PAPER)

YONATAN AUMANN

Bar Ilan University
Ramat Gan, Israel

Abstract. Classically, risk aversion is equated with concavity of the utility function. In this work we explore the conceptual foundations of this definition. In accordance with neo-classical economics, we seek an ordinal definition, based on the decision maker’s preference order, independent of numerical values. We present two such definitions, based on simple, conceptually appealing interpretations of the notion of risk-aversion. We then show that when cast in quantitative form these ordinal definitions coincide with the classical Arrow-Pratt definition (once the latter is defined with respect to the appropriate units), thus providing a conceptual foundation for the classical definition. The implications of the theory are discussed, including, in particular, to the understanding of insurance. The entire study is within the expected utility framework.

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

1. Introduction

1.1. Risk Aversion - The Classical Approach. The concept of risk aversion is fundamental in economic theory. Classically, it is defined as an attitude under which the certainty equivalent of a gamble is less than the gamble’s expected value; e.g., if a decision maker prefers one unit with certainty over a fair gamble between three units and none, then she is deemed risk averse. Thus, the natural, or neutral, certainty equivalent of a gamble is presumed to be its expectation, and risk aversion is defined with respect to this natural certainty equivalent.

In this work we ask “why”? Why is the gamble’s expected value presumed to be its natural certainty equivalent? Clearly, this presumption cannot rest on empirical evidence, as most people are assumed to be risk averse. The justification must be conceptual. But what is the reasoning that dictates that a fair gamble between $100 and $200 “should” be worth $150? Why the arithmetic mean? Why not, say, the geometric mean? or some other function? Indeed, perhaps there is no...
one “natural” certainty equivalent for a gamble. Providing a conceptual justification for basing the
definition of risk aversion on the arithmetic mean is the main goal of this paper.

In addition, there is the matter of units. Clearly, the mean by one set of units may not be the
mean by another; so the choice of units upon which risk aversion is defined is pivotal. In the seminal
works of Arrow [1] and Pratt [22], risk aversion was defined with respect to money. However, it
is not totally clear why, or in what sense, this is the “right” or “appropriate” scale. In particular,
using this scale to measure risk aversion suggests that, if not for risk, the attitude towards money
should in some sense be linear in the money amount; but why? Also, basing the definition on
money limits the notion to fully liquid goods. Thus, for non-liquid goods, or in societies without
money, risk aversion is not well defined. This is somewhat troubling, as risk aversion would seem
to be associated with some attitude of the decision maker, not with the existence of an external
market.

Clearly, the above two questions are related. Without a well-founded justification for using the
arithmetic mean, there can be no rational way to reason about the appropriate units.

Finally, and perhaps most fundamentally, the classical definition of risk aversion is inherently
cardinal - both technically and conceptually. Technically, the definition is only invariant under
affine transformation of the underlying scale, but not under general monotone transformations
(e.g. if we measure the convexity of the utility function with respect, say, of the square root of the
money amount, instead of the amount itself, we will get a different definition). Conceptually, the
notion expectation and the associated risk premium are only meaningful in a cardinal framework
in which sizes are meaningful. But from a neoclassical perspective, where the preference order is
the core object of interest, this is troubling, or unaesthetic at the least. Can such a fundamental
notion be defined only in cardinal terms? Does it have no ordinal underpinnings?

1.2. An Ordinal Foundation. In order to establish a conceptual foundation for the theory of
risk aversion, we start by providing two new definitions of the term, independent of any units,
and making no use of arithmetic notions such as mean or expectation. Rather, our definitions
employ conceptually appealing interpretations of the term, based solely on the internal structure
of the decision maker’s preferences. Having defined risk aversion in purely ordinal terms, we then
show that it can also be cast in quantitative form, provided that the appropriate cardinal scale
exists. This quantitative form, we show, coincides with the Arrow-Pratt definition, once the latter
is defined with respect to this scale - which in general is not money. Thus, we provide the missing
conceptual justification for the use of the expectation as the baseline for defining risk aversion, and
determine the “right” units.

1.2.1. Disentangling Risk Attitude from Risk-Free Preferences. The starting point for our definition
is the realization that any convincing definition of risk aversion must disentangle the attitude
towards risk from the underlying attitude towards the risk-free states. Consider a person preferring
a 30 minute long shower with certainty over a fair gamble between a two-hour long shower and
no shower at all. What is the basis for this preference? Is it dislike of risk, or disregard for long
showers? or both? How can one tell? Is it at all possible to formalize the distinction between the two? Von Neumann-Morgenstern utilities hopelessly entangle the two.\(^1\) Indeed, as long as we consider preferences on showers alone, it is impossible to differentiate between the attitudes to risk and to risk-free states.

The key, we propose, is to consider the preferences within a wider context: that of commodity bundles, rather than a single commodity. Preferences on commodity bundles exhibit sufficient structure to allow disentangling risk from risk-free attitudes. This, in turn, allows us to provide elementary ordinal definition(s) of risk-aversion.

It is important to stress that here the term “commodities” may refer to different types of goods (e.g. apples and oranges), or to the same good at different times (e.g. oranges today and oranges tomorrow), or to any combination thereof (apples and oranges today and tomorrow).\(^2\) In the remainder we use the different times setting as the running example.

1.2.2. **Ordinal Definition I: Hedging.** Our first definition is similar to that of Richard \(^3\), and equates risk aversion with a preference for hedging bets, whenever possible.\(^3\) Consider two commodities (e.g. oranges today and oranges tomorrow), and assume that the (risk-free, certainty) preferences on each commodity separately are well defined (more oranges today is better than less, regardless of the amount of oranges tomorrow, and vice versa). Then, risk aversion is defined as follows:

Let \(a, A\), be two states of one commodity (e.g. 1 orange today and 10 oranges today), and \(b, B\), two states of the other commodity (e.g. 2 oranges tomorrow and 15 oranges tomorrow), such that the decision maker is indifferent between \((a, B)\) and \((A, b)\). Then, the decision maker is **risk-averse** if she prefers the fair gamble between \((A, b)\) and \((a, B)\) - a gamble that is fully hedged - over the non-hedged fair gamble between \((a, b)\) and \((A, B)\).

Note that this definition is fully ordinal; it uses only the ordinal preferences on commodity bundles, with no reference to any quantitative measure.\(^4\)

The above definition considers a setting with two commodities. A similar definition also applies to multi-commodity settings, wherein the commodities are partitioned into two groups and hedging takes place between the groups. In this case it might seem that the concept may depend on how the commodities are grouped: a person may, say, prefer hedging between today and tomorrow, but dislike hedging between work and pleasure. We show that this is not possible; regardless of how one chooses to partition the commodities into independent groups, a decision maker prefers

\(^1\)The separation of diminishing marginal utility from risk aversion is one of the earliest motivation for the non-expected utility (EU) literature, see Yaari \(^26\). In this work, however, we remain within the EU framework.

\(^2\)However, “commodities” does not refer to contingent commodities, as our use of the term specifically refers only to sure outcomes. Preferences over contingent commodities are determined by the lottery preferences.

\(^3\)The exact relationship to Richard’s definition is discussed in Section 3.1.

\(^4\)The definition does require independence of the commodities, but not additive separability. So, a cardinal representation is not assumed.
hedging according to one partition if and only if she prefers hedging according to any and all other partitions. Thus, this definition of risk aversion reflects an underlying attitude of the decision maker, not a particularity of the specific partition.

1.2.3. **Ordinal Definition II: Repeated Gambles.** Our second definition of risk aversion formulates the intuition that risk aversion is somehow related the law of large numbers, but does so in an ordinal framework.

Consider a gamble $L$ with certainty equivalent $c$. The most extreme form of risk aversion would be displayed if, with probability 1, the gamble provides a better outcome than its certainty equivalent; that is, the worst possible outcome of the gamble is better than its certainty equivalent. 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 an attitude, however, is ruled out by the von Neumann-Morgenstern (NM) axioms; the utility of a lottery must lie between the utilities of its possible outcomes. Interestingly, while such an attitude is indeed not possible for any single gamble, it is possible once we consider risk aversion as a *policy*, consistently adhered to over multiple gambles. For some preference orders (agreeing with the NM axioms), repeatedly choosing the certainty equivalent of a gamble over the gamble itself can result in an outcome that is, with probability 1, inferior to what would have been the outcome of the gambles. This is our second definition of risk aversion: a preference policy\(^5\) is deemed risk-averse if adhering to this policy over repeated gambles ultimately results in an inferior outcome, with probability 1. Importantly, here “inferior” is according to the decision maker’s own preference order (over the sure outcomes), not any external market-based criterion.

1.2.4. **The quantitative Perspective.** Having established the ordinal foundations for the theory of risk-aversion, we show that these ordinal notions can also be cast in quantitative form, using an appropriate scale - if and when it exists. Such a scale, we show, is provided by the *multi-attribute (additive) value function*, pioneered by Debreu [4, 5], and commonly used in the theory of multi-attribute decision theory (see [18]). Debreu proves that (under appropriate conditions) the preferences on commodity bundles can be represented by the sum of appropriately defined functions of the individual commodities. Importantly, these Debreu functions are defined solely on the basis of the internal preferences amongst the commodity bundles. Thus, unlike market value - which is determined by external market forces - the Debreu functions represent the decision maker’s own preferences. Also, the functions are defined using the preferences on sure outcomes alone, with no reference to gambles. Thus, they provide a natural, intrinsic yardstick with which risk-aversion can be measured.

We show that our ordinal definitions of risk-aversion coincide with the Arrow-Pratt cardinal definition, once the latter is defined with respect to the Debreu function. Specifically, 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 any of the two ordinal definitions.

\(^5\)Formally defined in Section 5.
1.2.5. **Inter-Commodity and Intra-Commodity Risk Aversion.** We should stress that risk-aversion, as defined above, does not relate only to gambles involving multiple commodities or times, but also to gambles within a single commodity/time. It may be seen that, given the multi-commodity certainty preferences, inter-commodity lottery preferences induce intra-commodity lottery preferences, and vice versa. Thus, inter-commodity and intra-commodity risk attitudes are one and the same. We use the inter-commodity setting as it provides an Archimedean vantage point from which the risk-attitude can be disentangled from the risk-free preferences, and “pure” risk aversion can be defined. Once defined, however, it applies to all manifestations of risk. This is highlighted by the quantitative form described above. The multi-commodity setting merely provides us with the appropriate *scale* with which to measure risk aversion, both inter and intra-commodity.

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

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

1.3. **Plan of the Paper.** The remainder of the paper is structured as follows. Immediately following, in Section 2, we present the terminology and notation used throughout. The first ordinal definition is then presented and analyzed in Section 3. The equivalent quantitative form of this definition is provided in Section 4. Section 5 presents the second definition, and its equivalent quantitative form. We conclude the main body of the paper with a discussion in Section 6. All proofs are deferred to an appendix.

---

6A formal definition is provided in the next section.
2. Terminology and Notation

The Commodity Spaces and Preference Orders. The setting is the following:

- A set $G_1, \ldots, G_n$ of commodity spaces, where each $G_i$ is a real interval,
- A preference order\footnote{A preference order is a complete, transitive and reflexive binary relation.} $\preceq$ on the product space $\Omega = G_1 \times G_2 \times \cdots \times G_n$,
- A continuous preference order $\succeq$ on $\Delta(\Omega)$ - the space of lotteries over $\Omega$ - which agrees with $\preceq$ on $\Omega$ (the sure outcomes).

As customary, $\prec$ denotes the strict preference order induced by $\preceq$, and $\sim$ the induced indifference relation; similarly $\prec\Delta$ and $\sim\Delta$ denote the relations induced by $\succeq\Delta$. Continuity of $\succeq\Delta$ means that for any lottery $L$, the sets $\{ \omega : \omega \prec\Delta L \}$ and $\{ \omega : \omega \succ\Delta L \}$ are open (in $\Omega$). Since $\succeq$ and $\succeq\Delta$ agree on $\Omega$, this implies that $\succeq$ is also continuous (that is, the sets $\{ \omega : \omega \prec\Delta \omega_1 \}$ and $\{ \omega : \omega \succ\Delta \omega_1 \}$ are open for all $\omega_1 \in \Omega$).

All commodity spaces $G_i$ are assumed to be strictly essential\footnote{For simplicity, we only consider lotteries with finite support.}; that is, for each $i$ and $\omega_{-i} \in \Omega_{-i}$ (the remaining commodities), there exist $\omega_i, \omega'_i \in G_i$ with $(\omega_i, \omega_{-i}) \not\sim (\omega'_i, \omega_{-i})$.

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

Lotteries. For commodity bundles $\omega_1, \ldots, \omega_m$, and probabilities $p_1, \ldots, p_m$, denote by $p_1\omega_1 \oplus \cdots \oplus p_m\omega_m$ the lottery whose outcome is $\omega_i$ with probability $p_i$.\footnote{For simplicity, we only consider lotteries with finite support.} Thus, $\frac{1}{2}\omega_1 \oplus \frac{1}{2}\omega_2$ denotes the fair lottery between $\omega_1$ and $\omega_2$.

Factors and Partitions. The term factor refers to a single $G_i$ or a product of several $G_i$’s; i.e., a factor is the product of one or more commodity spaces. A partition of $\Omega$ is a representation of $\Omega$ as a product of factors $\Omega = A_1 \times \cdots \times A_k$. An element of $\Omega$ (or of any factor) is called a bundle.

To represent bundles of two elements, we frequently use column vectors; e.g. $\begin{pmatrix} a \ b \end{pmatrix}$ is a bundle of the partition $\Omega = A \times B$.

Bundle Intervals. For $\omega \preceq \overline{\omega}$, we denote

$$[\omega, \overline{\omega}] = \{ \omega : \omega \preceq \omega \preceq \overline{\omega} \}$$

That is, $[\omega, \overline{\omega}]$ is the closed interval of bundles between $\omega$ and $\overline{\omega}$. Hence, we call such an $[\omega, \overline{\omega}]$ a bundle interval, or simply interval. Similarly, the open interval is defined with strict preferences.

Utility Representations. While preferences are the core objects we consider, it is often useful to represent the preferences with a quantitative function. Formally, a function $f : \Omega \to \mathbb{R}$ represents $\preceq$ if for any $\omega, \omega' \in \Omega$,

$$\omega \succeq \omega' \iff f(\omega) \leq f(\omega').$$

The function $f : \Omega \to \mathbb{R}$ is an NM utility of $\preceq$ if for any $L_1, L_2 \in \Delta(\Omega)$,

$$L_1 \preceq L_2 \iff E_{L_1}[f(\omega)] \leq E_{L_2}[f(\omega)].$$
where $E_{L_j}[f(\omega)]$ is the expectation of $f(\omega)$ when $\omega$ is distributed according to $L_j$. In that case we also say that $f$ represents $\preceq$.

**Independence.** Independence is a key notion in our analysis. Simply put, a factor is independent if the preferences on the factor are well defined; i.e., the preferences within the factor are independent of the state in other factors. Formally, for a partition $\Omega = A_1 \times \cdots \times A_k$, we say that factor $A_i$ is *independent* if there exists a preference order $\preceq_{A_i}$ on $A_i$ such that for any $a_1, a_2 \in A_i$ and any $b \in \Omega_{-i}$ (the remaining factors),

$$a_1 \preceq_{A_i} a_2 \iff \left(\frac{a_1}{b}\right) \preceq \left(\frac{a_2}{b}\right).$$

We use the phrases “$\preceq_{A_i}$ is well defined” and “$A_i$ is independent” synonymously. It is important to stress that independence only refers to the certainty preferences; it does not state or imply that the preferences on lotteries in the one factor are independent of the state in other factors. That would be a much stronger assumption, which we do not make.

When no confusion can result, we may write $\preceq$ instead of $\preceq_{A_i}$; thus, when $a, a' \in A$, we may write $a \preceq a'$ instead of $a \preceq_{A} a'$. It is worth noting that the product of independent factors need not be independent.\(^9\)

A partition $\Omega = \mathcal{A}_1 \times \cdots \times \mathcal{A}_k$ is an *independent partition* if each $\mathcal{A}_i$ is independent.

**Relative Convexity/Concavity.** Let $f, g : S \to \mathbb{R}$, for some space $S$, with $g(x) = g(y) \Rightarrow f(x) = f(y)$, for all $x, y \in S$. We say that $f$ is concave with respect to $g$ if $f \circ g^{-1} : \mathbb{R} \to \mathbb{R}$ is concave; i.e., if there is a concave function $h$ with $f = h \circ g$. Similarly for convexity, strict concavity, and strict convexity.

### 3. Ordinal Definitions I: Hedging

We start by providing the formal statement of the hedging based definition.

**Definition 1.** Consider an independent partition $\Omega = \mathcal{A} \times \mathcal{B}$, and $(\frac{a}{b}), (\frac{A}{B}) \in \mathcal{A} \times \mathcal{B}$. Say that $(\frac{a}{b}), (\frac{A}{B})$ are perfectly hedged if $(\frac{a}{b}) \sim (\frac{A}{B})$, but $(\frac{a}{b}) \not\sim (\frac{A}{B})$ (see Figure 1).

We say that $\preceq$ is ordinally risk-averse (with respect to the partition $\mathcal{A} \times \mathcal{B}$) if for any perfectly hedged $(\frac{a}{b}), (\frac{A}{B})$,

$$1\over 2 \left( \frac{a}{b} \right) \oplus 1\over 2 \left( \frac{A}{B} \right) \preceq 1\over 2 \left( \frac{a}{b} \right) \oplus 1\over 2 \left( \frac{A}{B} \right).$$

We say that $\preceq$ is weakly ordinally risk averse (with respect to the given partition) if (1) hold with a weak preference ($\preceq$) rather than strict.

\(^9\)A simple example is the preference on $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z} = (\mathbb{R}^+)^3$ represented by the function $v(x, y, z) = xy + z$. Here, each commodity space is independent, but $\mathcal{Y} \times \mathcal{Z}$ is not independent.
Similarly, $\preceq$ is ordinally risk-loving (with respect to the partition) if for any perfectly hedged $(a_B),(A)$,

\[
\frac{1}{2}\left(\frac{a}{b}\right) \oplus \frac{1}{2}\left(\frac{A}{B}\right) \succ \frac{1}{2}\left(\frac{a}{B}\right) \oplus \frac{1}{2}\left(\frac{A}{b}\right),
\]

and weakly ordinally risk loving if the preference in (2) is a weak one.

Finally, $\preceq$ is ordinally risk-neutral (with respect to the partition) if for any perfectly hedged $(a_B),(A)$,

\[
\frac{1}{2}\left(\frac{a}{b}\right) \oplus \frac{1}{2}\left(\frac{A}{B}\right) \sim \frac{1}{2}\left(\frac{a}{B}\right) \oplus \frac{1}{2}\left(\frac{A}{b}\right).
\]

Thus, a decision maker is ordinally risk-averse if she prefers to hedge her bets whenever possible. Note that the definition is fully ordinal, both conceptually and technically. Conceptually, the definition relates only to the preference order, not any numerical values. Technically, we do not assume that the preference order is additively separable,\footnote{The preference order $\preceq$ is additively separable if there exists a partition $\Omega = \mathcal{A} \times \mathcal{B}$ and functions $f_a : \mathcal{A} \to \mathbb{R}, f_b : \mathcal{B} \to \mathbb{R}$, such that the function $v$ on $\Omega$ defined as $v(a,b) = f_a(a) + f_b(b)$ represents $\preceq$.} so a cardinal scale need not exist.\footnote{We do assume that the partition is independent, but that alone does not provide additive separability.}

Definition 1 considers a specific partition $\Omega = \mathcal{A} \times \mathcal{B}$. However, there could possibly be more than one partition of the space into independent factors. In that case, it is conceivable that the decision maker prefers a hedge provided by one partition, while disliking another hedge provided by a different partition. In order for our definition of risk aversion to be coherent we must guarantee that it does not depend on the specific partition. This is established by the following theorem.

**Theorem 1.** If $\preceq$ is (weakly) ordinally risk averse with respect to some independent partition $\Omega = \mathcal{A} \times \mathcal{B}$, then it is also so with respect to any independent partition. Similarly for risk loving and risk neutral.
By Theorem 1, we may call \( \preceq \) ordinally risk averse if it is ordinally risk averse with respect to some partition, and so for all partitions. It then follows that the following three propositions are mutually exclusive: \( \preceq \) is ordinally risk-averse, \( \preceq \) is ordinally risk-loving, and \( \preceq \) is ordinally risk-neutral.

3.1. **Ordinal Risk Aversion and Correlation Aversion.** Richard [23] defined the following notion of correlation aversion (see also [10, 6]).\(^{12}\) Consider an independent partition \( \Omega = A \times B \), and \( a, A \in A, b, B \in B \), with \( a \prec A \) and \( b \prec B \). Then, \( \preceq \) is correlation averse if

\[
\frac{1}{2} \left( \frac{a}{b} \right) \oplus \frac{1}{2} \left( \frac{A}{B} \right) \prec \frac{1}{2} \left( \frac{a}{A} \right) \oplus \frac{1}{2} \left( \frac{b}{B} \right),
\]

for any such \( a, A, b, B \). Thus, correlation aversion requires the decision maker to prefer any reduction in correlation between the factors, not only perfect hedges. The following theorem establishes that ordinal risk aversion and correlation aversion are in fact equivalent.

**Theorem 2.** \( \preceq \) is ordinally risk-averse if and only if it is correlation averse.

We note that the theorem holds even if \( \preceq \) is not additively separable.

4. **A Quantitative Perspective**

The previous section provided a fully ordinal definition of risk aversion. We now show how this ordinal definition can also be cast in quantitative form. Specifically, we show that the ordinal definition of risk-aversion coincides with the Arrow-Pratt cardinal definition, once the latter is defined with respect to the appropriate scale, if and when this scale exists.

4.1. **Debreu Value Functions.** The theory of multi-attribute decision making considers certainty preferences over a multi-factor space, and establishes that under certain independence assumptions, such preferences can be represented by quantitative functions, as follows. Consider a partition \( \Omega = A_1 \times \cdots \times A_k \) \((k \geq 2)\). Debreu [4, 5] proves that, provided that the product of each sub-set of factors is independent,\(^{13}\) then \( \preceq \) is additively separable (with respect to this partition); that is, there exist functions \( v^{A_i} : A_i \to \mathbb{R} \), such that for any \((a_1, \ldots, a_k), (a_1', \ldots, a_k')\)

\[
(a_1, \ldots, a_k) \preceq (a_1', \ldots, a_k') \iff \sum_{i=1}^{k} v^{A_i}(a_i) \leq \sum_{i=1}^{k} v^{A_i}(a_i').
\]

In the case of two factors \((k = 2)\), the following Thomsen condition is also required: for all \(a^1, a^2, a^3 \in A_1\), and \(b^1, b^2, b^3 \in A_2\), if \((a^1) \sim (a^2)\) and \((a^3) \sim (a^4)\) then \((a^1) \sim (a^3)\).\(^{14}\) It is important to note that the functions are defined solely on the basis of the certainty preferences.

\(^{12}\)Actually, Richard used the term multivariate risk aversion. The now common term correlation aversion was later coined by Epstein and Tanny [10].

\(^{13}\)In fact, it suffice that each pair of consecutive factors, according to some ordering, is independent.

\(^{14}\)For \(k > 2\) the Thomsen condition is implied by the independence of all subsets.
Debreu’s theorem also establishes that the functions are unique up to similar positive affine transformations (that is, multiplication by identical positive constants and addition of possibly different constants).

We call the function \( v^{A_i} \) a (Debreu) value function for \( A_i \), and the aggregate function \( v = \sum_{i=1}^{k} v^{A_i} \) a (Debreu) value function for \( \Omega \).\(^{15}\) We note that Debreu [4] called these functions utility functions; but following Keeney and Raiffa [18], we use the term value functions, to distinguish them from the NM utility function. This is not, however, intended to suggest any inherent “value” interpretation to this function, beyond its representational capacity.

4.1.1. **Uniqueness of the Aggregate Debreu Value Function.** We now want to establish a relation between ordinal risk-aversion, on the one hand, and the aggregate Debreu value function, on the other. Before we can do so, however, we need to guarantee that the notion of “the” aggregate Debreu value function is well defined. Debreu’s theorem relates to a specific partition of the space, and asserts that the value functions are unique (up to similar positive affine transformations) for the given partition. It does not assert that a different function may not arise from a different partition. Thus, the notion of a single, unique value function for \( \Omega \) may not be well defined. The following theorem, which is of independent interest, shows that this is not the case; all disparate value functions that may arise from different partitions are identical.

**Theorem 3.** For any \( \Omega \), all (aggregate) Debreu value functions for \( \Omega \) are identical up to positive affine transformations.

4.2. **Risk Aversion and the Debreu Value Functions.** Assume that the appropriate conditions hold, and a Debreu value function exists. We now show that in this case, ordinal Definition 1 can be recast in quantitative form, using the Debreu value functions as the base units. First, we show that, when measured in terms of the value function, the (sure) value of a perfectly hedged lottery is exactly the expectation of the associated non-hedged lottery.

**Theorem 4.** Let \( v \) be a Debreu value function for \( \Omega \), and \( \left( \begin{array}{c} a \\ B \end{array} \right) \), \( \left( \begin{array}{c} \hat{a} \\ \hat{B} \end{array} \right) \) perfectly hedged. Then,

\[
v \left( \begin{array}{c} a \\ B \end{array} \right) = v \left( \begin{array}{c} A \\ b \end{array} \right) = \frac{1}{2} \left( v \left( \begin{array}{c} a \\ b \end{array} \right) + v \left( \begin{array}{c} A \\ B \end{array} \right) \right).
\]

Thus, (ordinal) risk aversion indeed corresponds to a preference for the expectation of a lottery over the lottery itself, once the expectation is taken in terms of the value function (rather than money).

\(^{15}\)This is a slight abuse of notation. More precisely, \( v \) is the function on \( \Omega \) given by \( v(a_1, \ldots, a_k) = \sum_{i=1}^{k} v^{A_i}(a_i) \).
This, in turn, establishes that ordinal risk aversion coincides with concavity of the NM utility, once the latter is defined with respect to the value function.

**Theorem 5.** For NM utility $u$ and Debreu value function $v$,

- **Risk aversion:**
  - $u$ is strictly concave with respect to $v$ if and only if $\succ$ is ordinally risk averse.
  - $u$ is concave with respect to $v$ if and only if $\succ$ is weakly ordinally risk averse.

- **Risk loving:**
  - $u$ is strictly convex with respect to $v$ if and only if $\prec$ is ordinally risk loving.
  - $u$ is convex with respect to $v$ if and only if $\preceq$ is weakly ordinally risk loving.

- **Risk neutrality:** $u$ is linear with respect to $v$ if and only if $\sim$ is ordinally risk-neutral.

In all, we obtain that ordinal risk aversion coincides with Arrow-Pratt risk aversion, if and when a Debreu value function exists and concavity is defined with respect to this function.

5. Ordinal Definition II: Repeated Lotteries

5.1. **The Ordinal Definition.** Our second ordinal definition of risk aversion is set in the context of repeated lotteries, and conceptually equates risk aversion with a policy that ultimately leads to an inferior outcome. More specifically, under a risk averse policy repeatedly choosing the certainty equivalent of a lottery over the lottery itself ultimately leads, with probability 1, to an inferior outcome. To make this definition concrete, we must first define the associated notions, including: policy, repeated lotteries, certainty equivalent of a repeated lottery, and ultimately inferior outcome.

5.1.1. **The Space.** We consider an infinite sequence of factors $A_1, A_2, \ldots$, where $A_i$ represents the consumption space at time $i$. Each factor $A_i$ may be a product of commodity spaces, as in the previous sections. We denote $H^k = A_1 \times \cdots \times A_k$.

5.1.2. **Certainty Preferences.** While the number of factors is infinite, we only consider preferences on the finite prefix spaces $H^k$. Assume that any finite product of factors is independent, and denote by $\succ^k$ the preference order on $H^k$. Naturally, the preference orders $\succ^k$ must be consistent, in the sense that for $k' > k$, the preference order induced on $H^{k'}$ by $\succ^{k'}$ is identical to $\succ^k$. For brevity, we omit the superscript $k$ from $\succ^k$ whenever clear from the context.

For $a_i, b_i \in A_i$ and $a_j, b_j \in A_j$, we denote $[a_i, b_i] \approx [a_j, b_j]$ if $(a_i, b_i) \sim (b_i, a_j)$ (see Figure 2). Conceptually this means that the “distance” between $a_i$ and $b_i$ along the $A_i$ axis is the same as the “distance” between $a_j$ and $b_j$ along the $A_j$ axis. A sequence $b = (b_j, b_{j+1}, \ldots)$ is internal if there exist $a_i \prec b_i \prec c_i$ for all $i \geq j$, with $[a_i, b_i] \approx [a_j, b_j] \approx [b_i, c_i] \approx [b_j, c_j]$. Conceptually this means that $b$ is “bounded away” from the boundary.

---

16We do not assume that $A_i = A_j$, 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, in nominal terms, discounting may or may not be applied between consecutive factors. Our discussion here is independent of any such nominal matters.
Figure 2. Illustration of \([a_i, b_i] \approx [a_j, b_j]\). Conceptually this means that the “distance” between \(a_i\) and \(b_i\) along the \(A_i\) axis is of the “same size” as the distance between \(a_j\) and \(b_j\) along the \(A_j\) axis.

5.1.3. Lottery Preferences. Denote by \(\succeq^k\) the preference order on \(\Delta(H^k)\). Whereas the factors are assumed independent, the lottery preferences thereupon need not be independent. That is, the preference order on \(\Delta(H^k)\) induced by \(\succeq^{k+1}\) may depend on the state \(a_{k+1}\) in \(A_{k+1}\). We do assume, however, a form of weak consistency, whereby there exists some \(\phi_{k+1} \in A_{k+1}\) with
\[
\ell_1 \succeq^k \ell_2 \iff (\ell_1, \phi_{k+1}) \succeq^{k+1} (\ell_2, \phi_{k+1});
\]
that is, the preferences on \(\Delta(H^k)\) are consistent with some possible future. We call the sequence \((\phi_2, \phi_3, \ldots)\) a presumed future, and assume that it is internal.\(^{17}\) Conceptually this means that the presumed future is not “extreme”. We call the sequence of preference orders \(\succeq = (\succeq^1, \succeq^2, \ldots)\), the risk policy.

5.1.4. Lottery Sequences. Let \(\ell_1, \ell_2, \ldots\), be a sequence of lotteries, with \(\ell_i\) a lottery over \(A_i\). We denote by \((\ell_1, \ldots, \ell_k)\) the lottery over \(H^k\) obtained by the independent application of each \(\ell_i\) on its associated factor. A lottery sequence \((\ell_1, \ell_2, \ldots)\) is bounded if there exist \((a_1, a_2, \ldots), (b_1, b_2, \ldots)\), with \([a_1, b_1] \approx [a_i, b_i]\) for all \(i\), and \(\ell_i \in \Delta([a_i, b_i])\). Conceptually, this means that the magnitude of the \(\ell_i\’s\) does not grow indefinitely.

Two fair lotteries, \(\ell_i = \frac{1}{2}a_i \oplus \frac{1}{2}b_i, \ell_j = \frac{1}{2}a_j \oplus \frac{1}{2}b_j\), are similar if \([a_i, b_i] \approx [a_j, b_j]\). A lottery sequence \(\ell = (\ell_1, \ell_2, \ldots)\) is a recurring lottery sequence if it is composed solely of infinitely many similar (non-degenerate) fair lotteries and sure states. Thus, in a recurring lottery sequence, the “same magnitude” fair gamble occurs again and again, possibly interleaved with sure states.

5.1.5. Certainty Equivalents. Suppose that at time \(t = 1\) the decision maker is offered the choice between lottery \(\ell_1\) and its certainty equivalent \(c_1\). Then, consistent with her risk policy, she may \(^{17}\)More precisely, we assume that there exists a presumed future that is internal.
choose $c_1$, which suppose she indeed does. Now, at time $t_2$, she is offered the choice between lottery $\ell_2$ and its certainty equivalent $c_2$. Again, consistent with her risk policy, she chooses $c_2$. Suppose that she is thus offered, in each time period, the choice between a lottery $\ell_i$ 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 $\ell = (\ell_1, \ell_2, \ldots)$ if $(c_1, \ldots, c_k) \prec (c_1, \ldots, c_{k-1}, \ell_k)$ for all $k$.

5.1.6. Ultimate Inferiority. Consider a sequence $c = (c_1, c_2, \ldots)$ of sure states ($c_i \in A_i$) and a sequence $\ell = (\ell_1, \ell_2, \ldots)$ of lotteries ($\ell_i \in \Delta(A_i)$). We say that $c$ is ultimately inferior to $\ell$ if

$$\Pr[(c_1, \ldots, c_k) \prec (\ell_1, \ldots, \ell_k) \text{ from some } k \text{ on}] = 1.$$ 

Notably, here $(\ell_1, \ldots, \ell_k)$ refers to the outcome of the lottery and $\prec$ denotes the preference over the sure states. Thus, if $c$ is ultimately inferior to $\ell$, then consistently choosing the sure state $c_i$ over the lottery $\ell_i$, will, with probability 1, eventually result in an inferior (sure) outcome, and continue doing so indefinitely.

Similarly, $\ell$ is ultimately inferior to $c$ if

$$\Pr[(\ell_1, \ldots, \ell_k) \prec (c_1, \ldots, c_k) \text{ from some } k \text{ on}] = 1.$$ 

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

**Definition 2.** We say that $\succeq$ is:

- risk averse if the repeated certainty equivalent of any recurring lottery sequence $\ell$ is ultimately inferior to $\ell$.
- weakly risk averse if no bounded lottery sequence is ultimately inferior to its repeated certainty equivalent.

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

From the definition itself, it is not immediately evident that risk aversion implies weak risk aversion. This is provided by the following theorem.

**Theorem 6.** If $\succeq$ is risk averse then it is weakly risk averse.

5.2. The Quantitative Perspective. The above definition relates only to the preference orders, with no mention of arithmetic. We now show that it, too, corresponds to concavity of the utility function with respect to the value function.

5.2.1. The Debreu Value Function. First, we show that the value function takes a simple form.

**Proposition 1.** There exist Debreu value functions $v^{A_i} : A_i \to \mathbb{R}$, $i = 1, 2, \ldots$, such that for all $k$, $v_k = \sum_{i=1}^k v^{A_i}$ represents $\succeq^k$. 

13
5.2.2. Uniform Strict Concavity. For each $k$, let $u^k$ be the NM utility function representing $\preceq \Delta^k$. We want to relate the risk aversion of $\preceq \Delta^k$ to strict concavity of $u^1, u^2, \ldots$, with respect to $v^1, v^2, \ldots$. Here, however, we do not consider a single lottery, but rather a sequence of recurring lotteries. Thus, it is not sufficient that each $u^k$ is strictly concave individually, but also that, together, they are uniformly strictly concave, in the sense that the “degree of concavity” on any recurring lottery-sequence is “bounded away from zero”. Our first objective is thus to define formally the above notions, starting with the notion of “degree of concavity”. The following discussion assumes that the functions under consideration are real-valued and monotonically increasing over a real interval (bounded or unbounded).

By the standard definition, a function $f$ is concave if
\[
f(x) > \frac{1}{2}(f(x + \epsilon) + f(x - \epsilon)),
\]
for any $x$ and $\epsilon$. So, a natural choice for a measure for concavity would seem to be the difference
\[
f(x) - \frac{1}{2}(f(x + \epsilon) + f(x - \epsilon)). \tag{3}
\]
This measure, however, depends not only on the function’s concavity, but also on its slope. Thus, to get a measure of the concavity, we need to divide (3) by the difference between the end points $f(x + \epsilon)$ and $f(x - \epsilon)$. We thus obtain the following measure of concavity:
\[
\text{concavity}_f(x, \epsilon) = \frac{f(x) - \frac{1}{2}(f(x + \epsilon) + f(x - \epsilon))}{\frac{1}{2}(f(x + \epsilon) - f(x - \epsilon))}. \tag{18}
\]
This measure is invariant under affine transformations of the underlying units (the $x$ axis). Note that the measure depends both on $x$ and on $\epsilon$; that is, it is a measure of the concavity with respect to some $[-\epsilon, +\epsilon]$ interval around the point $x$. This accords with our interest in lotteries, which necessarily span some interval.

Now, given a function $f$, or more generally, a family of functions $f = \{f_1, f_2, \ldots\}$, we can consider the least concavity on any $[-\epsilon, +\epsilon]$ interval, reaching the following definition:

**Definition 3.** A family of increasing functions $f = \{f_1, f_2, \ldots\}$, is uniformly strictly concave if for any $\epsilon > 0$, $\text{concavity}_{f_i}(x, \epsilon)$ is positive and bounded away from 0, uniformly for all $i$ and $x$.

The sequence $(u^1, u^2, \ldots)$ is uniformly strictly concave with respect to $(v^1, v^2, \ldots)$ if the family \{${u^1 \circ (v^1)^{-1}, u^2 \circ (v^2)^{-1}, \ldots}$\} is uniformly strictly concave.

Before moving on to describe the connection between uniform strict concavity and risk aversion, we elaborate some more on the meaning of this measure of concavity.

First, we show that this measure is closely related to the Arrow-Pratt coefficient of absolute risk aversion. For a function $f$, the coefficient of absolute risk aversion at point $x$ is defined as $A_f(x) = -\frac{f''(x)}{f'(x)}$. This coefficient, which depends only on $x$, is the scaled limit of our $\text{concavity}_f(x, \epsilon)$ measure, as established by the following proposition:

---

18In the denominator, we divide the difference between the end-points by 2 so that the measure is bounded between 1 and $-1$. 

14
Proposition 2. For an increasing function $f$, for any $x$

$$\lim_{\epsilon \to 0} \frac{\text{concavity}_f(x, \epsilon)}{\epsilon} = \frac{1}{2} A_f(x).$$

(assuming $f$ is twice differentiable - so that $A_f$ is defined).

Additionally, if the coefficient of absolute risk aversion is bounded away from zero, then in particular, the concavity is bounded away from zero, so the functions are uniformly strictly risk averse:

Proposition 3. For a family $f = \{f_1, f_2, \ldots\}$, if $A_{f_i}(x)$ is positive and bounded away from zero (uniformly for all $f_i$ and $x$), then $f$ is uniformly strictly concave.

5.2.3. Horizontal Concavity and the Risk Premium. We now present another measure of concavity, which has a natural economic interpretation, and relate it to the measure defined above. Consider again an increasing concave function, as depicted in Figure 3. The standard definition of concavity considers the vertically aligned points $C$ and $B$, and defines $f$ to be concave if $C$ is above $B$. One can, however, just as well consider the horizontally aligned points $A$ and $B$, and say that $f$ is concave if $A$ is to the left of $B$. So, in analogy to the concavity$_f(x, \epsilon)$ measure, we can associate a measure of concavity with this horizontal definition of concavity, taking the distance between $A$ and $B$ and dividing it by $\epsilon$ - half the horizontal distance between the ends-point. We thus obtain:

$$h\text{-concavity}_f(x, \epsilon) = \frac{1}{\epsilon} \left( x - f^{-1}(\frac{f(x+\epsilon) + f(x-\epsilon)}{2}) \right).$$

This measure has a natural economic interpretation. If $f$ is an NM utility function, then point $B$ represents the expected return of the fair gamble between $x + \epsilon$ and $x - \epsilon$, while point $A$ is the certainty equivalent of the gamble. Thus, $h\text{-concavity}_f(x, \epsilon)$ is the risk premium, scaled by the size of the gamble.
The following proposition establishes that our two measures of concavity are the same, up to a constant factor:

**Proposition 4.** For an increasing concave function \( f \),

\[
\frac{1}{2} \text{concavity}_f(x, \epsilon) \leq h\text{-concavity}_f(x, \epsilon) \leq 2 \cdot \text{concavity}_f(x, \epsilon)
\]

for any \( x, \epsilon \).

5.2.4. **Risk Aversion and Concavity.** We are now ready to characterize risk averse and weakly risk averse policies in terms of the concavity of their utility functions:

**Theorem 7.**

(a) \( \approx \) is weakly risk averse if and only if \( u^k \) is concave with respect to \( v^k \) for all \( k \).

(b) \( \approx \) is risk averse if and only if the sequence \( (u^1, u^2, \ldots) \) is uniformly strictly concave with respect to \( (v^1, v^2, \ldots) \).

In particular, the theorem implies that risk averse policies exist. For example, the CARA utility functions defined as \( u^k = -\exp(-v^k) \), for all \( k \), are uniformly strictly convex. Hence, the corresponding risk policy is risk averse.

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

**Definition 4.** We say that \( \approx \) is:

- risk loving if any recurring lottery sequence is ultimately inferior to its repeated certainty equivalent.
- weakly risk averse if no repeated certainty equivalent of a bounded lottery sequence is ultimately inferior to the sequence itself.
- risk neutral if it is both weakly risk loving and weakly risk averse.

Thus, the risk loving require an ultimately superior certainty equivalent to forgo their love of risk. In analogy to Theorem 6 we have,

**Theorem 8.** If \( \approx \) is risk loving then it is weakly risk loving.

5.3.1. **The Quantitative Perspective.**

**Definition 5.** A family of increasing functions \( f = \{f_1, f_2, \ldots\} \), is uniformly strictly convex if for any \( \epsilon > 0 \), \( \text{concavity}_{f_i}(x, \epsilon) \) is negative and bounded away from 0, uniformly for all \( i \) and \( x \).

The sequence \( (u^1, u^2, \ldots) \) is uniformly strictly convex with respect to \( (v^1, v^2, \ldots) \) if the family \( \{u^1 \circ (v^1)^{-1}, u^2 \circ (v^2)^{-1}, \ldots\} \) is uniformly strictly convex.

And finally,
Theorem 9.

(a) $\succeq$ is weakly risk loving if and only if $u^k$ is convex with respect to $v^k$ for all $k$.
(b) $\succeq$ is risk loving if and only if the sequence $(u^1, u^2, \ldots)$ is uniformly strictly convex with respect to $(v^1, v^2, \ldots)$.
(c) $\succeq$ is risk neutral if and only if $u^k$ is linear in $v^k$ for all $k$.

5.4. Relating the Two Ordinal Definitions. We provided two ordinal definition of risk aversion: Definition 1, based on hedging, and Definition 2, based on repeated lotteries. Technically, the two definitions relate to different mathematical objects: the first relates to a single preference order, while the latter relates to a risk policy, which is a sequence of preference orders. However, the two definitions are closely related, as established by Theorems 5 and 7: both definitions correspond to concavity of the NM utility function with respect to the Debreu value function. For weak risk aversion the concavity requirements in both theorem are identical - (weak) concavity. So, a risk policy is weakly risk averse, according to Definition 2, if and only if each of the preferences orders therein is weakly risk averse, according to the Definition 1. For (strict) risk aversion, the requirement in Theorem 7 is uniform strict concavity, whereas Theorem 5 only requires strict concavity. So, if the risk policy is (strictly) risk averse then so are all of the preference orders therein, but the opposite does not always hold. The reason is that since we are considering the behavior on recurring gambles we need the same (recurring) bound on the strict concavity in all the gambles.

6. Discussion

We presented fully ordinal definitions of risk aversion, based entirely on the internal structure of preferences of the decision maker; independent of money or any other units. Our definitions rest on two intuitively appealing interpretations of risk aversion. The first equates risk aversion with a preference for hedging bets. The second equates risk aversion with a policy under which, in order to avoid being in a risky situation, the decision maker’s choices necessarily lead her to an inferior outcome. We then show that when cast in numerical terms, these ordinal definitions coincide with the Arrow-Pratt definition, once the latter is defined with respect to the Debreu value function associated with the decision maker’s preferences over the sure outcomes. In particular, this provides the missing conceptual justification for the use of the arithmetic mean - or expectation - as the basis for defining risk aversion, and, at the same time, establishes the appropriate units to use.

Under the classical definition, risk aversion is synonymous with concavity of the utility function with respect to money. This has been the established definition of risk aversion for over half a century; but it frequently fails to accord with the plain meaning of the term. Consider, for example, a person offered the choice between a half-pound steak with certainty, and a fair gamble between a two pound steak and no steak at all. Clearly, she may prefer the certainty option not because of any
dislike of risk, but rather because she has little taste for more than half a pound of beef. Similarly, a person may prefer 1 billion dollars with certainty over a fair gamble between 10 billion dollars and bankruptcy, not because she dislikes risk, but rather because the extra 9 billion dollars provide her with little additional benefit (in some - perhaps not well defined - intuitive sense). Thus, equating risk aversion with concavity of the utility function frequently fails to convey the plain meaning of the term. We believe that our ordinal definitions (and their numerical equivalents) better accord with this plain, everyday meaning.

This new notion of risk aversion may have important implications for our understanding and interpretation of key economic behavior. Consider, for example, an aging, retired individual, comfortably living off her savings, who is offered a 50-50 gamble between tripling her savings and losing them all. Common sense has it that rejecting the gamble is a perfectly rational choice for all but the most risk loving individuals. Classical economic language, however, would deem such a rejection “risk aversion”. Our notion of ordinal risk aversion allows for a more convincing interpretation of the behavior. When measured in terms of the Debreu value function, which reflects the relative benefits provided by each of the possible outcomes, the 50-50 gamble is (most likely) actuarially inferior to the existing state. So, it should be rejected even by risk neutral, as well as some risk loving, individuals.

Interestingly, the same holds for insurance, as we show next.

6.1. Insurance. Buying insurance is a prime example of behavior classically and universally attributed to risk aversion. Indeed, the entire insurance industry is based on the fact that, in total, insurers pay back less than what they collect. Thus, in expectation, the insured pay more than they get, which, under the classical definition, equates with risk aversion. This, however, only holds when measured in dollar terms; once payments are measured in other units, the picture changes.

As an example, consider disability insurance. For this insurance, the industry’s typical loss-ratio is in the 70%-90% range; that is, on average, the insured get back only an expected 70%-90% of their investment. Classically, this would be interpreted as a clear indication of risk aversion. But this need not be so for ordinal risk aversion. In order to analyze the situation from an ordinal perspective, we must consider the payments, of both the insured and the insurers, in terms of the Debreu value function, rather than money.

The value function may vary from one individual to another, and its determination requires knowing the individual’s preferences across multiple commodity bundles. Thus, it is impossible to provide a simple universal analysis using the value function. However, the following provides an illustrative analysis of insurance in terms of these units.

\[\text{19} \text{For the sake of discussion we assume free disposal; that is, the decision maker can discard, at no additional cost, any surplus steak she may have.}\]

\[\text{20} \text{For simplicity, we ignore investment income (e.g. interest, dividends, capital gains) in our discussion here. In reality, investment income is an important component of the insurers’ revenues, but its inclusion would significantly complicate the discussion, without altering the core reasoning.}\]
Suppose that an employee’s compensation package is composed of: (i) an annual salary, and (ii) an annual number of vacation days. Thus, the compensation is a pair \((x, y)\), where \(x\) is the annual salary and \(y\) is the number of vacation days. Naturally, an individual prefers a higher salary and more vacation days. Suppose that, starting from a base salary of \(x\) and no vacation days, the employee is willing to forgo some fraction of the salary in return for getting some vacation days; e.g., she is willing to settle for 90% of the salary if the compensation includes one week of vacation, 85% of the salary if it includes 2 weeks of vacation, and so forth. Let \(g(y)\) be the fraction of salary that the decision maker is willing to settle for, if given \(y\) days of vacation; in the above example \(g(\text{one week}) = 90\%\) and \(g(2 \text{ weeks}) = 85\%\). So the individual is indifferent between the bundles \((x, 0)\) and \((g(y) \cdot x, y)\). Hence, the function

\[
f(x, y) = x \cdot g^{-1}(y),
\]

represents the preferences on the pairs \((x, y)\). Hence, so does the function

\[
v(x, y) = \ln(x) - \ln(g(y)).
\]

So, \(\ln(x)\) and \(-\ln(g(y))\) are Debreu value functions for this preference order. Since \(\ln(x)\) is concave, the value function is concave in the money amount. Thus, a negative expectation in dollar terms need not be negative in terms of the value function.

By way of example suppose that an individual earns $40k a year; there is a 2% chance of disability, which would lower her salary to $10k a year; insurance would bring the salary back to $40k; and the premium is $800 a year. Then in dollar terms, the expected return is:

\[
2\% \cdot 30,000 = 600,
\]

which represents an expected return rate of 75% on the investment of $800. Thus in dollar terms, such insurance provides an investment with a negative expected return.

Now, consider the situation is terms of the value function. The $800 paid as premium are worth

\[
\ln(40,000) - \ln(39,200) \approx 0.02
\]

*Debreu value units (DVU).* The insurance’s indemnification of the salary from $10k back to $40k is worth

\[
\ln(40,000) - \ln(10,000) \approx 1.39
\]

DVU’s. Thus in terms of DVUs, the expected return is

\[
2\% \cdot 1.39 \approx 0.028,
\]

\[21\]The assumption here is that the “value” of a vacation day is determined as a fraction of the salary, independent of the salary itself. This simplifies the analysis and, we believe, offers is a reasonable first approximation. Other functions, that do take into account the associated salary can also be used. These would change the details of the analysis, but not the essence of the argument.

\[22\]For simplicity, the presentation here considers each year separately. In practice, disability indemnification, as well as premiums, are paid over many years.
which represents an expected 140% return rate on the 0.02 DVUs invested as premium. So, the negative expectation in dollar terms translates into a positive expectation in terms of the Debreu value function. Thus, under the above assumptions, buying disability insurance is a perfectly rational choice even for (ordinally) risk loving individuals.

Indeed, the entire consideration of insurance primarily in terms of risk aversion seems misguided. A more instructive view of insurance, we suggest, is as a means for transferring funds from the well-off state of the individual to the less-well-off or poor state of the same individual - in which the funds are worth much more (in terms of the value function). If the poor state were sure to occur, this transfer of funds would simply take the form of a savings plan (e.g. a pension plan). Insurance comes into play when there is a probability that the poor state may not materialize, in which case it is wasteful to put aside the entire amount. Instead, insurance provides a mechanism by which only a fraction of the funds need to be set aside, in return for getting the full amount if the poor state occurs and getting nothing if it does not. Using this mechanism, i.e. buying insurance, may thus be a perfectly rational choice for (ordinally) risk neutral and even some (ordinally) risk loving individuals.

6.2. Measures of Risk Aversion. Arrow and Pratt defined concrete measures of risk aversion, namely the coefficient of absolute risk aversion at \( x \), defined as \(-\frac{u''(x)}{u'(x)}\), and the coefficient of relative risk aversion at \( x \), defined as \(-\frac{x u''(x)}{u'(x)}\). The measure of absolute risk aversion can naturally be converted to our definition of risk aversion, by simply considering the utility function with respect to the Debreu value function rather than with respect to (w.r.t.) money. The notion of relative-risk-aversion w.r.t. the value function, however, is not well defined, as the definition of relative risk aversion requires a well-defined zero point, and the value function is only defined up to an additive constant.\(^{23}\)

Once considered w.r.t. the value function, constant-absolute-risk-aversion (CARA) has a simple and intuitive meaning. A preference order is CARA w.r.t. the value function if and only if the preferences over lotteries in each individual factor are well defined and independent of the state in the other factors; preferences over apple lotteries are independent of the available amount of oranges and preferences over orange lotteries are independent of the available amount of apples (this is termed utility independence in \([23, 2, 18]\), where a proof of the equivalence can also be found).

In the economic literature, CRRA (constant relative risk aversion) rather than CARA, is the more prevalent model. CRRA, however, is assumed w.r.t. money. Once considered in terms of the value function, the observed CRRA w.r.t. money may actually reflect a combination of an

\(^{23}\)Indeed, we would argue that determining the zero point is a big problem, mostly overlooked, also when defining relative risk aversion w.r.t. money. What is the right zero point? no money in the bank? no material possessions (no house, no clothes, no food)? no money left after selling a kidney? Choosing any of these zero points results in very different relative risk aversion coefficients.
underlying CARA ordinal risk attitude superimposed on a value function that is logarithmic w.r.t.
money. This combination yields exactly the known CRRA family of functions:

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

Interestingly, this means that the utility functions \( \ln(x) \) and \( x^{\gamma} \) actually correspond to ordinal risk
neutrality and risk loving, not risk aversion.

### 6.3. Strength of Preference and Relative Risk Aversion

Dyer and Sarin [9] and Bell and Raiffa [2] have suggested measuring risk aversion with respect to the strength of preference function, rather than money. It is out of the scope of this paper to review the strength-of-preference theory, but generally speaking this theory assumes that not only do decision makers have a well defined preference order over sure states and lotteries, but also that they have a preference order over differences between states; that is, the decision maker can state that she prefers the transition \( x_1 \rightarrow x_2 \) over the the transition \( y_1 \rightarrow y_2 \) (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 [8]) that represents these preferences, in the sense that \( f(x_2) - f(x_1) > f(y_2) - f(y_1) \) if and only if the transition \( x_1 \rightarrow x_2 \) is preferred over the transition \( y_1 \rightarrow y_2 \). Given such a function, Dyer and Sarin [9] define the notion of relative risk aversion\(^{24}\) as the concavity of the NM utility function \( u \) with respect to the measurable value function \( f \). Bell and Raiffa [2] similarly define the notion of intrinsic risk aversion.

Bell and Raiffa [2] also show how the strength-of-preference function (assuming it exists) can be deduced and identified with a multi-attribute (Debreu) value function (see also [9, Theorem 1]). Thus, technically our notion of ordinal risk aversion coincides with the Dyer and Sarin notion of relative risk aversion, if the latter is computed using the Debreu value function. Conceptually, however, our approach is totally different from that of [9] and [2]. First, we do not suppose, technically or conceptually, any form of preferences over differences. Rather, we only use the standard preferences on bundles and lotteries thereof. Second, conceptually [9] and [2] follow the Arrow-Pratt framework, taking it as given that the “natural value” of a gamble “should be” its expectation. They differ from Arrow-Pratt only in using a different scale. Thus, at its core, their approach is also cardinal - attributing significance to cardinal amounts, not only to ordinal preferences. Our approach is the opposite. Our starting point, and all core definitions, are fully ordinal. The numerical representation is then mathematically derived from this ordinal theory.

### 6.4. Multi-Attribute Risk Aversion

There is a long and important line of research on multi-attribute risk aversion. Some of these works remain within the expected utility (EU) framework, while others venture out. For the works of the former type (see e.g. [19, 25, 21, 7, 15, 23, 16, 20]), the general approach is to start from the Arrow-Pratt definition of the single attribute case and

\(^{24}\)not to be confused with the Arrow-Pratt coefficient of relative risk aversion
explore how this definition can (or cannot) be extended to the multi-attribute case. Our approach is the reverse; we start from the multi-attribute case, and then derive the uni-attribute case as a quantitative representation of the former.

There is also much work on multi-attribute risk aversion in the non-EU framework (see e.g. [17, 13, 24] and references therein). As mentioned in the Introduction, the current work remains strictly within the EU framework.

6.5. Comparative Risk Aversion. Thus far we have considered ordinal risk aversion in the context of each decision maker separately. Is it also possible to compare the ordinal risk aversion levels of different individuals? Such a comparison is possible, but only between individuals who agree on the certainty preferences. For individuals agreeing on the certainty preferences, the comparative framework of Kihlstrom and Mirman [19] is applicable (see also [21, 20]). We note that Karni [15] developed a measure that allows comparing risk aversion among decision makers that do not agree on the certainty preferences. It is an interesting question to see if this theory can be extended to the framework suggested in this paper.

References


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

For readability, all theorems and propositions are restated in this appendix. Throughout, the following notation is used

- $v$ denotes a Debreu value function on $\Omega$, and $v^A$ a Debreu value function on a factor $A$.
- $u$ denotes an NM utility function on $\Omega$. An NM utility for $\Omega$ necessarily exists since the NM axioms are assumed to hold, and we consider only lotteries with finite support (see Fishburn [11, Theorem 8.2]). If the lottery preferences on a factor $A$ are well defined, then $u^A$ denotes an NM utility function on the factor.
- $w^A$ denotes a continuous function representing $\succeq^A$, for an independent factor $A$. Such a function exists by [3], since the preference order is continuous.

Proofs for Section 3. We start with some lemmas.

Lemma 1. $u$ is continuous.

Proof. It suffices to prove that the pre-images of the open rays $(-\infty, r)$ and $(r, \infty)$ are open, for all $r$ (these open rays constitute a subbase for the standard topology on the line). Consider $(-\infty, r)$ (the other case is analogous). If $u(\omega) \geq r$ for all $\omega \in \Omega$ then $u^{-1}(-\infty, r) = \emptyset$, which is open. Similarly, if $u(\omega) < r$ for all $\omega \in \Omega$ then $u^{-1}(-\infty, r) = \Omega$, which is open. Otherwise, there exist $s_1 < r \leq s_2$ and $\hat{\omega}_1, \hat{\omega}_2 \in \Omega$, with $u(\hat{\omega}_1) = s_1, u(\hat{\omega}_2) = s_2$. Set $\hat{p} = (r - s_1)/(s_2 - s_1)$. Then,
Consider an independent partition $\mathcal{A} \times \mathcal{B}$. Let $w^A : \mathcal{A} \to \mathbb{R}$ be a continuous real function representing $\succsim^A$, and similarly $w^B$ a continuous real function representing $\succsim^B$. Define $w : \mathcal{A} \times \mathcal{B} \to \mathbb{R}^2$ as $w(a,b) = (w^A(a), w^B(b))$. Let $I_A \times I_B \subseteq \mathbb{R}^2$ be the image of $\mathcal{A} \times \mathcal{B}$ under $w$.

**Lemma 2.** $u \circ w^{-1} : I_A \times I_B \to \mathbb{R}$ is well defined, increasing in each coordinate, and continuous.

**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 $w^A, w^B$ agree on the certainty preference.

To prove continuity, 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). Denote $\hat{u} = u \circ w^{-1}$. If $u(a,b) \geq r$ for all $(a,b)$ then the pre-image of $(-\infty, r)$ under $\hat{u}$ is empty, which is open. Otherwise, let $X_r = \{(x,y) : \hat{u}(x,y) < r\}$. Consider $(\hat{x}, \hat{y}) \in X_r$. We show that there is a neighborhood of $(\hat{x}, \hat{y})$ fully contained in $X_r$.

Suppose that $\hat{x}$ is not maximal in $I_A$ and $\hat{y}$ not maximal in $I_B$ (the proof for the case that one of them is maximal is similar). Let $(\hat{a}, \hat{b})$ be such that $w(\hat{a}, \hat{b}) = (\hat{x}, \hat{y})$. Fixing $\hat{a}$, the function $u_{\hat{a}} : \mathcal{B} \to \mathbb{R}$, defined by $u_{\hat{a}}(b) = u(\hat{a}, b)$ is continuous. Thus, there exists some $b'$ with

$$0 < u_{\hat{a}}(b') - u_{\hat{a}}(b) < \frac{1}{2} (r - u_{\hat{a}}(b)).$$

Similarly, the function $u_{\hat{y}} : \mathcal{A} \to \mathbb{R}$, defined by $u_{\hat{y}}(a) = u(a, \hat{b})$ is continuous. Thus, there exists $a'$ with

$$0 < u_{\hat{y}}(a') - u_{\hat{y}}(\hat{a}) < \frac{1}{2} (r - u_{\hat{y}}(a)).$$

Combining (4) and (5), we obtain

$$u(\hat{a}, \hat{b}) < u(a', b') < r.$$

Set $x' = w_A(a'), y' = w_B(b')$. By construction $x' > \hat{x}, y' > \hat{y}$ and $\hat{u}(\hat{x}, \hat{y}) < \hat{u}(x', y') < r$. Set $\delta = \min\{x' - \hat{x}, y' - \hat{y}\}$. Then, for any $(x,y)$ if $||(x,y) - (\hat{x}, \hat{y})|| < \delta$ 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 $(\hat{x}, \hat{y})$ is contained in $X_r$, as required. \qed

The following lemma is used in the proof of Theorem 2.

**Lemma 3.** Let $\mathcal{A} \times \mathcal{B}$ be an independent partition and $a \prec A$, $b \prec B$. Set $a^0 = a$, and while $(a^i_B) \not\sim (A)$ let $a^{i+1}$ be such that $(a^{i+1}) \sim (a^i_B)$ (such an $a^{i+1}$ exists by continuity). Then, there exists an $i$ such that $(a^i_B) \succeq (A)$ (that is, the sequence $a^0, a^1, \ldots$ is finite).
Proof. Contrariwise, suppose there is no such $\bar{i}$. Then, for $i = 1, 2, \ldots$, $(a^i_B) \prec (A^i_b)$, 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 $i$

$$
\left(\frac{a^i}{B}\right) \sim \left(\frac{a^{i+1}}{b}\right).
$$

Thus, by continuity,

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

which is impossible since $b \prec^B B$ and $\preceq$ is strictly monotone in each factor.

Theorem 1 will be proved after Theorem 5. We now prove Theorem 2. Note that the theorem holds even when $\preceq$ is not additively separable. When it is additively separable, the proof is much simpler.

**Theorem 2.** $\preceq$ is ordinally risk-averse if and only if it is correlation averse.

**Proof.** (if:) Suppose $\preceq$ is correlation averse. Consider a perfectly hedged pair $(a_B, A_b)$. By definition $(a_B) \not\sim (A_b)$. So, either $a \not\prec A$ or $b \not\prec B$. W.l.o.g. $a \prec A$. But $(a_B) \sim (A_b)$. So, $b \prec B$. Hence, by definition of correlation aversion

$$
\frac{1}{2} \left( \frac{a}{B} \right) \oplus \frac{1}{2} \left( \frac{A}{b} \right) \succ \frac{1}{2} \left( \frac{a}{b} \right) \oplus \frac{1}{2} \left( \frac{A}{B} \right).
$$

So, $\preceq$ is ordinally risk averse.

(only if:) Suppose that $\preceq$ is ordinally risk averse. Consider an independent partition $\Omega = A \times B$, and $a, A \in A, b, B \in B$, with $a \prec A$ and $b \prec B$. We show that

$$
\frac{1}{2} \left( \frac{a}{b} \right) \oplus \frac{1}{2} \left( \frac{A}{B} \right) \succ \frac{1}{2} \left( \frac{a}{b} \right) \oplus \frac{1}{2} \left( \frac{A}{B} \right).
$$

If $(a_B) \sim (A_b)$ then they are perfectly hedged and (6) holds by the definition of ordinal risk aversion.

Otherwise, let $u$ be an NM utility for $\preceq$. set

$$
diff = u \left( \frac{a}{b} \right) + u \left( \frac{A}{B} \right) - u \left( \frac{a}{B} \right) - u \left( \frac{A}{b} \right).
$$

We show that $\diff < 0$, which establishes (6).

Let $w^A$ be a continuous function representing $\preceq^A$ and $w^B$ a continuous function representing $\preceq^B$ (the certainty preferences). In order to prove that $\diff < 0$, we start out by proving that there exists $a_1^1, A_1^1, b_1^1, B_1^1$, with

$$
a \preceq a_1^1 \prec A_1^1 \preceq A, \quad \text{and} \quad b \preceq b_1^1 \prec B_1^1 \preceq B,
$$

such that

$$
w^A(A_1^1) - w^A(a_1^1) \leq \frac{1}{2} \left( w^A(A) - w^A(a) \right) \quad \text{or}
$$

$$
w^B(B_1^1) - w^B(b_1^1) \leq \frac{1}{2} \left( w^B(B) - w^B(b) \right)
$$

(7)
Figure 4. Illustration of the proof of Theorem 2. 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 $w^A$).

and

$\text{(8)} \quad \text{diff} < u\left(\frac{a_1}{b_2}\right) + u\left(\frac{A_1}{B_2}\right) - u\left(\frac{a_1}{B_2}\right) - u\left(\frac{A_1}{B_2}\right)$.

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

$\text{(9)} \quad \left(\frac{a^1}{b}\right) \sim \left(\frac{a}{B}\right)$.

Figure 4 illustrates the following argument. Set $a^0 = a$. Given $a^i$, let $a^{i+1}$ be such that $\left(\frac{a^{i+1}}{b}\right) \sim \left(\frac{a^i}{B}\right)$. Let $\bar{i}$ be the first index with $\left(\frac{a_{\bar{i}}}{B}\right) \succeq \left(\frac{A}{b}\right)$; such an $\bar{i}$ exists by Lemma 3. Then, $\left(\frac{a}{B}\right) < \left(\frac{A}{b}\right) \succeq \left(\frac{a_{\bar{i}}}{B}\right)$. Thus, there exists $A^1, a < A^1 \succeq a^i$, such that $\left(\frac{A^1}{B}\right) \sim \left(\frac{A}{b}\right)$. Clearly, $a^i \succeq A$. Thus, either

$\text{(10)} \quad w^A(A^1) \leq \frac{1}{2}(w^A(a) + w^A(A)),$

or

$\text{(11)} \quad w^A(a^i) \geq \frac{1}{2}(w^A(a) + w^A(A)).$

We consider each of these cases separately.

First, suppose that (10) holds. Then, by construction $\left(\frac{A^1}{b}\right) \sim \left(\frac{A}{b}\right)$, and they are perfectly hedged. Hence, by assumption,

$\frac{1}{2}\left(\frac{A^1}{b}\right) + \frac{1}{2}\left(\frac{A}{B}\right) \leq \frac{1}{2}\left(\frac{A}{b}\right) + \frac{1}{2}\left(\frac{A}{B}\right)$.

So,

$u\left(\frac{A^1}{b}\right) + u\left(\frac{A}{B}\right) - u\left(\frac{A}{b}\right) - u\left(\frac{A}{B}\right) < 0.$
Hence,
\[ u\left(\frac{a}{b}\right) + u\left(\frac{A}{B}\right) - u\left(\frac{A}{b}\right) - u\left(\frac{a}{B}\right) = \]
\[ u\left(\frac{a}{b}\right) + u\left(\frac{A_1}{B}\right) - u\left(\frac{A_1}{b}\right) - u\left(\frac{a}{B}\right) + u\left(\frac{A}{B}\right) - u\left(\frac{A_1}{B}\right) - u\left(\frac{A_1}{b}\right) < \]
\[ (12) \quad u\left(\frac{a}{b}\right) + u\left(\frac{A_1}{b}\right) - u\left(\frac{A_1}{b}\right) - u\left(\frac{a}{B}\right). \]

Setting \(a_\frac{1}{2} = a, \ A_\frac{1}{2} = A_1, \ b_\frac{1}{2} = b\) and \(B_\frac{1}{2} = B\), by (10) and (12) we get (7) and (8).

Next, suppose that (11) holds. Then, by construction, for \(i = 1, \ldots, \bar{i}\), \(\frac{a_i}{b_i} \sim \frac{a_i}{b_i}\), and each such pair is perfectly hedged. Since \(\prec\) is ordinally risk averse,
\[ \frac{1}{2} \left(\frac{a_i}{b_i}\right) \oplus \frac{1}{2} \left(\frac{a_i}{b_i}\right) \prec \frac{1}{2} \left(\frac{a_i}{b_i}\right) \oplus \frac{1}{2} \left(\frac{a_i}{b_i}\right), \]
for all \(i\). So,
\[ (13) \quad \frac{1}{2i} \bigoplus_{i=1}^{\bar{i}} \left(\frac{a_i}{b_i} \oplus \frac{a_i}{b_i}\right) \prec \frac{1}{2i} \bigoplus_{i=1}^{\bar{i}} \left(\frac{a_i}{b_i} \oplus \frac{a_i}{b_i}\right); \]
and
\[ \frac{1}{2} \left(\frac{a_0}{b}\right) \oplus \frac{1}{2} \left(\frac{a_0}{b}\right) \prec \frac{1}{2} \left(\frac{a_0}{b}\right) \oplus \frac{1}{2} \left(\frac{a_0}{b}\right); \]
so (as \(a_0 = a\))
\[ u\left(\frac{a}{b}\right) + u\left(\frac{a_\bar{i}}{b}\right) - u\left(\frac{a_\bar{i}}{b}\right) - u\left(\frac{a}{B}\right) < 0. \]

Hence,
\[ u\left(\frac{a}{b}\right) + u\left(\frac{A}{B}\right) - u\left(\frac{A}{b}\right) - u\left(\frac{a}{B}\right) = \]
\[ u\left(\frac{a}{b}\right) + u\left(\frac{a_\bar{i}}{b}\right) - u\left(\frac{a_\bar{i}}{b}\right) - u\left(\frac{a}{B}\right) + u\left(\frac{A}{B}\right) - u\left(\frac{a_\bar{i}}{B}\right) - u\left(\frac{a_\bar{i}}{b}\right) < \]
\[ (14) \quad u\left(\frac{a_\bar{i}}{b}\right) + u\left(\frac{A}{B}\right) - u\left(\frac{a_\bar{i}}{b}\right) - u\left(\frac{A}{b}\right). \]

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 (11) and (14) we get (7) and (8).

Thus, we have established (7) and (8), and we now return to complete the proof that \(\text{diff} < 0\).

Set
\[ \text{diff}_\frac{1}{2} = u\left(\frac{a_\frac{1}{2}}{b_\frac{1}{2}}\right) + u\left(\frac{A_\frac{1}{2}}{b_\frac{1}{2}}\right) - u\left(\frac{A_\frac{1}{2}}{B_\frac{1}{2}}\right) - u\left(\frac{A_\frac{1}{2}}{b_\frac{1}{2}}\right). \]

Then,
\[ \text{diff} < \text{diff}_\frac{1}{2}. \]
Applying the above halving procedure repeatedly, we obtain that for any $\delta > 0$ there exists $(a_\delta, b_\delta)$, such that

\begin{align*}
(15) \quad w^A(a_\delta) - w^A(b_\delta) & \leq \delta \\
(16) \quad w^B(b_\delta) - w^B(b_\delta) & \leq \delta
\end{align*}

and

\begin{align*}
diff_1 \leq \left| u\left(\begin{array}{c}
a_\delta \\
b_\delta
\end{array}\right) + u\left(\begin{array}{c}
a_\delta \\
b_\delta
\end{array}\right) - u\left(\begin{array}{c}
a_\delta \\
b_\delta
\end{array}\right) - u\left(\begin{array}{c}
a_\delta \\
b_\delta
\end{array}\right) =
\end{align*}

\begin{align*}
(17) \quad (u\left(\begin{array}{c}
a_\delta \\
b_\delta
\end{array}\right) - u\left(\begin{array}{c}
a_\delta \\
b_\delta
\end{array}\right)) + (u\left(\begin{array}{c}
a_\delta \\
b_\delta
\end{array}\right) - u\left(\begin{array}{c}
a_\delta \\
b_\delta
\end{array}\right)) =
\end{align*}

\begin{align*}
(18) \quad (u\left(\begin{array}{c}
a_\delta \\
b_\delta
\end{array}\right) - u\left(\begin{array}{c}
a_\delta \\
b_\delta
\end{array}\right)) + (u\left(\begin{array}{c}
a_\delta \\
b_\delta
\end{array}\right) - u\left(\begin{array}{c}
a_\delta \\
b_\delta
\end{array}\right)).
\end{align*}

By Lemma 2 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

\begin{align*}
||w^A(a') - w^A(a''), w^B(b') - w^B(b'')|| < \delta
\end{align*}

then

\begin{align*}
|u(a', b') - u(a'', b'')| < \epsilon.
\end{align*}

In particular, if (15) holds then (17) is $\leq 2\epsilon$, and if (16) holds then (18) is $\leq 2\epsilon$. Thus, $\text{diff}_1 \leq 0$, so $\text{diff} < 0$. \hfill \Box

**Proofs for Section 4.** Each factor $A$ is a product of some set of commodity spaces, $A = \prod_{i \in A} G_i$, for some index set $A$. For $A = \prod_{i \in A} G_i$ and $B = \prod_{i \in B} G_i$, by a slight abuse of notation, we write $A \cap B$ for $\prod_{i \in A \cap B} G_i$, $A - B$ for $\prod_{i \in A - B} G_i$, and $A \subseteq B$ if $A \subseteq B$. We say that $A$ and $B$ overlap if $A \cap B \neq \emptyset$ and neither is contained in the other; the factor $A$ is non-degenerate if $A \neq \emptyset$.

**Lemma 4.** If there exist two non-identical independent partitions $\Omega = A \times B$ and $\Omega = C \times D$, then:

1. there exist value functions $v^A, v^B, v^C, v^D$ (for $A, B, C, D$), with $v^A + v^B$ and $v^C + v^D$

\begin{align*}
eq\text{each representing } \preceq,
\end{align*}

2. $v^A + v^B = v^C + v^D$,

3. if $v^A, v^B$ are value functions for $A, B$, and $v^C, v^D$, are value functions for $C, D$, then $v^A + v^B$

\begin{align*}
is a positive affine transformation of $v^C + v^D$.
\end{align*}

**Proof.** Gorman [12, Theorem 1] proves that if two independent factors $E$ and $F$ overlap then $E \cup F, E \cap F, E - F, F - E$, and $E \Delta F = (E - F) \cup (F - E)$ are all independent.

Set $W = A \cap C$, $X = A \cap D$, $Y = B \cap C$, and $Z = B \cap D$. Then, by Gorman’s theorem, $W, X, Y, Z$ are independent, as is any product thereof. Since the partitions are not identical, at least three out of $W, X, Y, Z$ are non-degenerate. Thus, there are value functions $v^W, v^X, v^Y, v^Z$, with $v^W + v^X + v^Y + v^Z$ representing $\preceq$ (see Section 4). Thus, the pair of functions $v^A = v^W + v^X$, and
\[ v^B = v^Y + v^Z \] are value functions for the independent partition \( \Omega = A \times B \). Similarly, the functions \( v^C = v^W + v^Y \), and \( v^D = v^X + v^Z \) are value functions for the independent partition \( \Omega = C \times D \), proving (1) and (2). Finally, (3) follows from (2) by the uniqueness of value functions. \( \square \)

**Theorem 3.** For any \( \Omega \), all (aggregate) Debreu value functions for \( \Omega \) are identical up to positive affine transformations.

**Proof.** Suppose \( \Omega \) has two different independent partitions \( \Omega = A_1 \times \cdots \times A_k \) and \( \Omega = C_1 \times \cdots \times C_m \), with value functions \( v^{A_1}, \ldots, v^{A_k} \) and \( v^{C_1}, \ldots, v^{C_m} \), respectively. Since the two partitions are different, there must be some \( A_i \) for which there is no \( j \) with \( C_j = A_i \). W.l.o.g. this is \( A_1 \). Set \( B = A_2 \times \cdots \times A_k \) and \( v^B = \sum_{i=2}^k v^{A_i} \). Similarly, set \( D = C_2 \times \cdots \times C_j \) and \( v^D = \sum_{i=2}^j v^{C_i} \). Then, \( v^{A_1} + v^B \) represents \( v^C \), as does \( v^C + v^D \). Thus, by Lemma 4, \( \sum_{i=1}^k v^{A_i} = v^{A_1} + v^B \) is an affine transformation of \( \sum_{i=1}^m v^{C_i} = v^C + v^D \). \( \square \)

**Theorem 4.** Let \( v \) be a Debreu value function for \( \Omega \), and \( (\frac{a}{b}), (\frac{A}{B}) \) perfectly hedged. Then,

\[
v\left( \frac{a}{b} \right) = v\left( \frac{A}{B} \right) = \frac{1}{2} \left( v\left( \frac{a}{b} \right) + v\left( \frac{A}{B} \right) \right).
\]

**Proof.** Since there exists a value function for \( \Omega \), by definition, there exists an independent partition of \( \Omega \). Hence, by Lemma 4 there exist value functions for any independent partition. In particular, there exist value functions \( v^A, v^B \) for \( A, B \). Since \( (\frac{a}{b}), (\frac{A}{B}) \) are perfectly hedged

\[
v\left( \frac{a}{b} \right) = v\left( \frac{A}{B} \right) = v^A(a) + v^B(b) = v^A(A) + v^B(B)
\]

So,

\[
\frac{1}{2} \left( v\left( \frac{a}{b} \right) + v\left( \frac{A}{B} \right) \right) = \frac{1}{2} \left( v^A(a) + v^B(b) + v^A(A) + v^B(B) \right) = v\left( \frac{a}{b} \right) = v\left( \frac{A}{B} \right). \quad \square
\]

**Theorem 5.** For NM utility \( u \) and Debreu value function \( v \),

- **Risk aversion:**
  - \( u \) is strictly concave with respect to \( v \) if and only if \( \preceq \) is ordinally risk averse.
  - \( u \) is concave with respect to \( v \) if and only if \( \preceq \) is weakly ordinally risk averse.

- **Risk loving:**
  - \( u \) is strictly convex with respect to \( v \) if and only if \( \succeq \) is ordinally risk loving.
  - \( u \) is convex with respect to \( v \) if and only if \( \succeq \) is weakly ordinally risk loving.

- **Risk neutrality:** \( u \) is linear with respect to \( v \) if and only if \( \preceq \) is ordinally risk-neutral.

**Proof.** We prove the theorem for risk aversion and strict concavity. The proofs for the other cases are analogous.

(only if:) Suppose \( u \) is strictly concave with respect to \( v \). Let \( \Omega = A \times B \) be an independent partition, and \( (\frac{a}{b}), (\frac{A}{B}) \) perfectly hedged. By definition \( v\left( \frac{a}{b} \right) \neq v\left( \frac{A}{B} \right) \), and By Theorem 4 \( v\left( \frac{a}{b} \right) = \frac{1}{2} \left( v\left( \frac{a}{b} \right) + v\left( \frac{A}{B} \right) \right) \). Hence, since \( u \) is strictly concave with respect to \( v \),

\[
\frac{1}{2} \left( u\left( \frac{a}{b} \right) + u\left( \frac{A}{B} \right) \right) < u\left( \frac{a}{b} \right).
\]
Hence (since \( u \) is a representation of \( \preceq \))
\[
\frac{1}{2} \begin{pmatrix} a \\ b \end{pmatrix} \oplus \frac{1}{2} \begin{pmatrix} A \\ B \end{pmatrix} \preceq \begin{pmatrix} a \\ B \end{pmatrix},
\]
so \( \preceq \) is ordinally risk-averse.

(if:) Suppose \( \preceq \) is ordinally risk-averse with respect to some partition \( \Omega = A \times B \). Since there exists a value function \( v \) representing \( \preceq \) (based on some independent partition), by Lemma 4, there exist value functions \( v^A, v^B \), with \( v^A + v^B = v \).

Set \( I_A = v^A(A), I_B = v^A(B) \). Let \( \epsilon \) be such that both \( I_A \) and \( I_B \) are of size at least \( 2\epsilon \). We prove that \( u \circ v^{-1} \) is strictly concave on any interval of size \( 2\epsilon \), and hence strictly concave throughout.

Consider \( x \in I \). Then \( x = v^B(y) \) for some \( \hat{a}, \hat{b} \). Set \( x_A = v^A(\hat{a}) \) and \( x_B = v^B(\hat{b}) \). Assume that \((x_A + \epsilon, x_B + \epsilon) \in I_A \times I_B \) (the other cases are similar). We prove that \( u \circ v^{-1} \) is strictly concave on \([x, x + 2\epsilon]\).

Consider \( y, z \in [x, x + 2\epsilon] \) with \( y < z \). Then, \( y = x + \delta_y, z = x + \delta_z, \) with \( 0 \leq \delta_y < \delta_z \leq 2\epsilon \). Then there exist \( a, A, b, B \), with \( v^A(a) = x_A + \delta_y/2, v^B(b) = x_B + \delta_y/2, v^A(A) = x_A + \delta_z/2, v^B(B) = x_B + \delta_z/2 \). Then, \((a_B), (A)\) are perfectly hedged. Since \( \preceq \) is risk averse,
\[
\frac{1}{2} \begin{pmatrix} a \\ b \end{pmatrix} \oplus \frac{1}{2} \begin{pmatrix} A \\ B \end{pmatrix} \preceq \begin{pmatrix} a \\ B \end{pmatrix},
\]
So,
\[
\frac{1}{2} \left( u \begin{pmatrix} a \\ b \end{pmatrix} \right) \oplus \frac{1}{2} \left( u \begin{pmatrix} A \\ B \end{pmatrix} \right) \preceq \frac{1}{2} \left( u \begin{pmatrix} a \\ B \end{pmatrix} \right),
\]
i.e.
\[
\frac{1}{2} \left( u(v^{-1}(y)) + u(v^{-1}(z)) \right) \preceq \frac{1}{2} \left( u(v^{-1}(\frac{y + z}{2})) + u(v^{-1}(\frac{y + z}{2})) \right) = u(v^{-1}(\frac{y + z}{2})).
\]
So, \( u \circ v^{-1} \) is mid-point strictly concave and hence strictly concave.

\[\Box\]

**Theorem 1.** If \( \preceq \) is (weakly) ordinally risk averse with respect to some independent partition \( \Omega = A \times B \), then it is also so with respect to any independent partition. Similarly for risk loving and risk neutral.

**Proof.** This follows from Lemma 4 and Theorem 5. Suppose there are two non-identical independent partitions \( \Omega = A \times B \) and \( \Omega = C \times D \), and that \( \preceq \) is ordinally risk averse with respect to the former. Then, by Lemma 4 there exists a value function \( v \) representing \( \preceq \). So, by the “if” direction of Theorem 5, \( u \) is strictly concave with respect to \( v \). Hence, by the “only if” direction of Theorem 5, \( \preceq \) is ordinally risk averse with respect to any partition, in particular with respect to \( C \times D \). (Note that the proof of Theorem 5 does not rely on Theorem 1.) A similar argument holds for weak ordinal risk aversion.

\[\Box\]

**Proofs for Section 5.** The proofs in this section follow certain conventions that simplify the presentation:

- \( x \) is a real number, \( \alpha, \beta, \delta \) - with or without indices or primes - are positive reals.
• $a_i, b_i$, and $c_i$ are points in $A_i$.
• $f$ and $f_i$ are increasing real valued functions on a real interval (bounded or unbounded).
• variables not explicitly quantified are taken to be universally quantified, it being understood
  that the expressions in which they appear are defined.

The proof of Theorem 6 follows that of Theorem 7.

**Proposition 1.** There exist Debreu value functions $v^{A_i} : A_i \rightarrow \mathbb{R}$, $i = 1, 2, \ldots$, such that for all $k$, $v^k = \sum_{i=1}^{k} v^{A_i}$ represents $\succeq^k$.

**Proof.** Consider $\mathcal{H}^k$ for $k \geq 3$. By assumption, any product of the $A_i$'s is independent. Hence, there exist value functions $v^k_1, \ldots, v^k_{k-1}$, with $\sum_{i=1}^{k-1} v^k_i$ representing $\succeq^{k-1}$. We now show that there is actually a single function $v^{A_i}$, for each $i$, that works for all the $\mathcal{H}^k$'s.

For $i = 1, 2, 3$, set $v^{A_i} := v^3_i$. Suppose $v^{A_k}$ has been defined for all $i < k$; we inductively define $v^{A_k}$. By the induction hypothesis, $\sum_{i=1}^{k-1} v^{A_i}$ represents $\succeq^{k-1}$. By independence of $\mathcal{H}^{k-1}$ in $\succeq^k$, the function $\sum_{i=1}^{k-1} v^{A_i}$ also represents $\succeq^{k-1}$. So, by uniqueness of the value functions, there exist constants $\beta > 0, \xi_i$, such that $v^{A_i} = \beta v^{A_i} + \xi_i$, for $i = 1, \ldots, k-1$. So, setting $v^{A_k} = \beta v^{A_k}$, we have that
\[
\sum_{i=1}^{k} v^{A_i} = \sum_{i=1}^{k-1} (\beta v^{A_i} + \xi_i) + \beta v^{A_k} = \beta \sum_{i=1}^{k} v^i + \text{constant},
\]
which represents $\succeq^k$, as required. \qed

**Proposition 2.** For an increasing function $f$, for any $x$
\[
\lim_{\epsilon \rightarrow 0} \frac{\text{concavity}_f(x, \epsilon)}{\epsilon} = \frac{1}{2} A_f(x).
\]
(assuming $f$ is twice differentiable - so that $A_f$ is defined).

**Proof.**
\[
\lim_{\epsilon \rightarrow 0} \frac{\text{concavity}_f(x, \epsilon)}{\epsilon} = \lim_{\epsilon \rightarrow 0} \left( \frac{1}{\epsilon} \cdot \frac{f(x) - \frac{1}{2} (f(x + \epsilon) + f(x - \epsilon))}{\frac{1}{2} (f(x + \epsilon) - f(x - \epsilon))} \right) = \lim_{\epsilon \rightarrow 0} \left( \frac{2f(x) - f(x + \epsilon) - f(x - \epsilon)}{\epsilon^2} \cdot \left( \frac{f(x + \epsilon) - f(x)}{\epsilon} \right)^{-1} + \lim_{\epsilon \rightarrow 0} \frac{f(x - \epsilon) - f(x)}{-\epsilon} \right) = \frac{f''(x)}{f'(x) + f'(x)} = \frac{A_f(x)}{2} \quad \Box
\]

**Proposition 4.** For an increasing concave function $f$,
\[
\frac{1}{2} \text{concavity}_f(x, \epsilon) \leq h\text{-concavity}_f(x, \epsilon) \leq 2 \cdot \text{concavity}_f(x, \epsilon)
\]
for any $x, \epsilon$. 

31
Proof. Figure 5 depicts the elements of this proof. The solid (curved) line is the graph of $f$. Points are denoted by capital letters (A,B,...). For points A,B, we denote by $|AB|$ the length of the straight line from A to B. We have

$$
\frac{|A'B|}{|CB|} = \frac{|DE|}{|CE|}.
$$

So,

$$
(19) \quad \frac{|A'B|}{|DE|} = \frac{|CB|}{|CE|}.
$$

Since $f$ is concave, the graph of $f$ must lie above the chord CD. So, from (19) we get,

$$
h\text{-concavity}_f(x,\epsilon) = \frac{|AB|}{|DE|} \geq \frac{|A'B|}{|DE|} = \frac{|CB|}{|CE|} \geq \frac{|CB|}{2|BE|} = \frac{1}{2} \text{concavity}_f(x,\epsilon).
$$

Similarly,

$$
\frac{|C'B|}{|AB|} = \frac{|HG|}{|AG|}.
$$

So,

$$
(20) \quad \frac{|C'B|}{|HG|} = \frac{|AB|}{|AG|}.
$$

Since $f$ is concave, the graph of $f$ must lie above the chord AH. Hence, from (20) we get,

$$
\text{concavity}_f(x,\epsilon) = \frac{|CB|}{|HG|} \geq \frac{|C'B|}{|HG|} = \frac{|AB|}{|AG|} \geq \frac{|AB|}{2|BG|} = \frac{1}{2} h\text{-concavity}_f(x,\epsilon).
$$

□

Lemma 5. If $f$ is (increasing) concave then $\text{concavity}_f(x,\epsilon)$ is non-decreasing in $\epsilon$. 32
Proof. Let $\epsilon_1 > \epsilon_2 > 0$. By concavity
\[
f(x + \epsilon_2) \geq (1 - \frac{\epsilon_2}{\epsilon_1})f(x) + \frac{\epsilon_2}{\epsilon_1}f(x + \epsilon_1),
\]
so,
\[
f(x + \epsilon_2) - f(x) \geq \frac{\epsilon_2}{\epsilon_1}(f(x + \epsilon_1) - f(x)).
\]
Similarly,
\[
\frac{\epsilon_2}{\epsilon_1}(f(x) - f(x - \epsilon_1)) \geq f(x) - f(x - \epsilon_2).
\]
Set
\[
\begin{align*}
\delta_1^+ &= f(x + \epsilon_1) - f(x) \\
\delta_1^- &= f(x) - f(x - \epsilon_1) \\
\delta_2^+ &= f(x + \epsilon_2) - f(x) \\
\delta_2^- &= f(x) - f(x - \epsilon_2)
\end{align*}
\]
Then, multiplying (21) by (22) we obtain:
\[
\delta_2^+ \delta_1^- \geq \delta_1^+ \delta_2^-,
\]
which is equivalent to
\[
\frac{\delta_1^- - \delta_1^+}{\delta_1^- + \delta_1^+} \geq \frac{\delta_2^- - \delta_2^+}{\delta_2^- + \delta_2^+}.
\]
Hence,
\[
\text{concavity}_{f}(x, \epsilon_1) = \frac{f(x) - \frac{1}{2}(f(x + \epsilon_1) + f(x - \epsilon_1))}{} = \frac{\frac{1}{2}(\delta_1^- - \delta_1^+)}{\frac{1}{2}(\delta_2^- + \delta_2^+)} \geq \frac{\frac{1}{2}(\delta_2^- - \delta_2^+)}{\frac{1}{2}(\delta_2^- + \delta_2^+)} = \text{concavity}_{f}(x, \epsilon_2). \]

Proposition 3. For a family $f = \{f_1, f_2, \ldots\}$, if $A_{f_i}(x)$ is positive and bounded away from zero (uniformly for all $f_i$ and $x$), then $f$ is uniformly strictly concave.

Proof. The function $g(x) = -e^{-\alpha x}$ has absolute risk aversion coefficient $\alpha$ everywhere. For this $g$,
\[
\text{concavity}_{g}(x, \epsilon) = \frac{2g(x) - g(x + \epsilon) - g(x - \epsilon)}{g(x + \epsilon) - g(x - \epsilon)} = \frac{-2e^{-\alpha x} + e^{-\alpha(x+\epsilon)} + e^{-\alpha(x-\epsilon)}}{-e^{-\alpha(x+\epsilon)} + e^{-\alpha(x-\epsilon)}} = \frac{e^{-\alpha x}[-2 + e^{-\alpha \epsilon} + e^{\alpha \epsilon}]}{e^{-\alpha x}[-e^{-\alpha \epsilon} + e^{\alpha \epsilon}]} = \frac{-2 + e^{-\alpha \epsilon} + e^{\alpha \epsilon}}{-e^{-\alpha \epsilon} + e^{\alpha \epsilon}}.
\]
Thus, $\text{concavity}_{g}(x, \epsilon)$ is independent of $x$, and hence $g$ is uniformly strictly concave.\(^{26}\)

\(^{26}\)i.e. uniformly in $x$. Differently put, the family of functions defined by $g_i(x) = -e^{\alpha x}$ for all $i$ is uniformly strictly concave.
Now, consider a function $f_i$ with absolute risk aversion coefficient $A_{f_i}(x) \geq \alpha = A_g(x)$, for all $x$. Then, as shown by Pratt [22, Theorem 1], for any $x$ and $\epsilon$,
\[
\frac{g(x+\epsilon)-g(x)}{g(x)-g(x-\epsilon)} \geq \frac{f_i(x+\epsilon)-f_i(x)}{f_i(x)-f_i(x-\epsilon)},
\]
which after some rearrangement becomes
\[
(23) \quad \frac{(f_i(x) - f_i(x-\epsilon)) - (f_i(x+\epsilon) - f_i(x))}{(f_i(x) - f_i(x-\epsilon)) + (f_i(x+\epsilon) - f_i(x))} \geq \frac{(g(x) - g_i(x-\epsilon)) - (g(x+\epsilon) - g(x))}{(g(x) - g_i(x-\epsilon)) + (g(x+\epsilon) - g(x))}.
\]
The left-hand side of (23) is $\text{concavity}_{f_i}(x,\epsilon)$ while the right-hand side is $\text{concavity}_{g}(x,\epsilon)$. So,
\[
\text{concavity}_{f_i}(x,\epsilon) \geq \text{concavity}_{g}(x,\epsilon)
\]
for all $i, x,$ and $\epsilon$. Hence, $f = \{f_1, f_2, \ldots\}$ is also uniformly strictly concave. 

From now on we assume w.l.o.g. that the factors are already represented in units of the respective value functions; that is, $v^A_i(a_i) = a_i$ for all $i$ and $a_i \in \mathcal{A}_i$. Then $u^k$, the NM utility function representing $\hat{\sim}^k$, is actually only a function of the sum of its arguments; i.e. $u^k(a_1, \ldots, a_k) = u^k(b_1, \ldots, b_k)$ whenever $a_1 + \cdots + a_k = b_1 + \cdots + b_k$. Accordingly, denote by $\hat{u}^k$ the function such that $u^k(a_1, \ldots, a_k) = \hat{u}^k(a_1 + \cdots + a_k)$. Note that $\hat{u}^k = u^k \circ (v^k)^{-1}$. Thus, $(u^1, u^2, \ldots)$ is uniformly strictly concave with respect to $(v^1, v^2, \ldots)$ if and only if $(\hat{u}^1, \hat{u}^2, \ldots)$ is uniformly strictly concave.

Let $(\phi_2, \phi_3, \ldots)$ be the presumed future. By assumption $(\phi_2, \phi_3, \ldots)$ is internal.\(^27\) Thus, there exists some $s > 0$ with $\phi_i \pm s \in \mathcal{A}_i$, for all $i$.

As explained in the main text, the horizontal concavity measure - $h\text{-concavity}_{f}(x,\epsilon)$ - is directly associated with risk aversion, as it represents the magnitude of the risk premium scaled by the size of the lottery. Thus, from now on we focus on this measure. The following lemma establishes that any (horizontal) concavity exhibited by $\hat{u}^k$, for some $k$, is (re-)exhibited by all subsequent $\hat{u}^m$, for $m > k$.

**Lemma 6.** For any $m > k$,
\[
h\text{-concavity}_{\hat{u}^m}(x + \phi_{k+1} + \ldots, \phi_m, \epsilon) = h\text{-concavity}_{\hat{u}^k}(x,\epsilon).
\]

**Proof.** Set $\beta = h\text{-concavity}_{\hat{u}^k}(x,\epsilon) \cdot \epsilon$. By definition
\[
\hat{u}^k(x - \beta) = \frac{1}{2}(\hat{u}^k(x - \epsilon) + \hat{u}^k(x + \epsilon)).
\]
Let $a_+, a_-, a_- \beta \in \mathcal{H}^k$ be such that $v^k(a_+) = x + \epsilon$, $v^k(a_-) = x - \epsilon$, and $v^k(a_- \beta) = x - \beta$. So,
\[
(a_- \beta) \hat{\sim}^k \frac{1}{2}a_- \epsilon \oplus \frac{1}{2}a_+ \epsilon.
\]
By assumption, $\hat{\sim}^k$ and $\hat{\sim}^m$ agree on the preferences over $\Delta(\mathcal{H}^k)$ when fixing the state in $\mathcal{A}_{k+1} \times \cdots \times \mathcal{A}_m$ to the presumed future $(\phi_{k+1}, \ldots, \phi_m)$. So,
\[
(a_- \beta, \phi_{k+1}, \ldots, \phi_m) \hat{\sim}^m \frac{1}{2}(a_- \epsilon, \phi_{k+1}, \ldots, \phi_m) \oplus \frac{1}{2}(a_+ \epsilon, \phi_{k+1}, \ldots, \phi_m).
\]

\(^{27}\)More precisely, $(\phi_2, \phi_3, \ldots)$ is a presumed future that is internal, if there are several presumed futures.
Hence,
\[
\hat{u}^m(x - \beta + \phi_{k+1} + \cdots + \phi_m) = \\
\frac{1}{2}(\hat{u}^m(x - \epsilon + \phi_{k+1} + \cdots + \phi_m) + \hat{u}^m(x + \epsilon + \phi_{k+1} + \cdots + \phi_m)).
\]

The following lemma establishes that if \( \hat{u}^k \) exhibits some level of concavity, at some point \( x \), then not only is this concavity re-exhibited by all subsequent utility functions \( \hat{u}^m \) (\( m > k \)), but also that it is “reachable” from any state \( y \), of any period \( K \).

**Lemma 7.** For any \( k, K, x, y \), with \( x \) in the domain of \( \hat{u}^k \) and \( y \) in the domain of \( \hat{u}^K \), there exist \( m \geq \max\{k, K\} \) and \( b_{K+1}, \ldots, b_m, b_i \in \mathcal{A}_i \), with

\[
\text{h-concavity}_{\hat{u}^m}(y + b_{K+1} + \cdots + b_m, \epsilon) = \text{h-concavity}_{\hat{u}^k}(x, \epsilon).
\]

**Proof.** Set \( K' = \max\{k, K\} \). If \( K < k \) then for \( i = K + 1, \ldots, k \), let \( b_i \) be any point in \( \mathcal{A}_i \) and set \( y' = y + b_{K+1} + \cdots + b_k \). Otherwise \( (K \geq k) \) set \( y' = y \).

Let \( \delta = y' - x, j = \lfloor \delta/s \rfloor \), and \( m = K' + j \). For \( i = K' + 1, \ldots, m \), set \( b_i = \phi_i + \delta/j \). Then, \( m > \max\{k, K\} \), and \( x + \phi_{k+1} + \cdots + \phi_m = y + b_{K+1} + \cdots + b_m \). The result then follows from Lemma 6.

**Lemma 8.** Let \( X_1, X_2, \ldots \) be an infinite sequence of independent uniformly bounded random variables,\(^{28}\) with \( E(X_i) = 0 \) for all \( i \). Set \( S_k = \sum_{i=1}^{k} X_i \). Then

\[
(24) \quad \Pr[S_k \geq 0 \text{ infinitely often (i.o.)}] > 0.
\]

**Proof.** Denote \( v_i = \text{Var}(X_i) \), and \( V_k = \sum_{i=1}^{k} v_i \). The \( X_i \)'s are independent, so \( V_k = \text{Var}(S_k) \). Either \( V_k \rightarrow \infty \) or not. We consider each case separately.

If \( V_k \rightarrow \infty \), applying the central limit theorem for uniformly bounded random variables (e.g. [14], Theorem 9.5) we obtain that

\[
\lim_{k \to \infty} \Pr[\frac{S_k}{\sqrt{V_k}} > 0] = \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-x^2/2} dx > 0.
\]

In particular, \( \Pr[\frac{S_k}{\sqrt{V_k}} > 0 \text{ i.o.}] > 0 \). Since \( \sqrt{V_k} > 0 \), (24) follows.

Next, suppose that \( V_k \) does not go to infinity. Each \( v_i \) is non-negative. Hence, the \( V_i \)'s form a monotonically non-decreasing and bounded sequence, and hence converge. Thus, for any \( \delta > 0 \) there exists an \( N_\delta \) with \( \sum_{i=N_\delta}^{\infty} v_i < \delta \). W.l.o.g. \( X_1 \) is not identically 0. Thus there exists an \( x > 0 \) with \( \Pr(X_1 \geq x) = p_x > 0 \). Choose \( \delta < x^2 \). Then by the Chebyshev inequality, for all \( k > N_\delta \),

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

\(^{28}\)i.e., the support of all the random variables is included in a real interval \([\underline{b}, \bar{b}]\), with \( \underline{b}, \bar{b} \) finite.
In particular, for the fixed $N_\delta$ and varying $k$’s,

$$\Pr\left[ \sum_{i=N_\delta}^k X_i > -x \ i.o. \right] \geq 1 - \frac{\delta}{x^2} > 0.$$ 

Finally, for the fixed $N_\delta$, there exists some positive probability $p_+$, with

$$\Pr[\sum_{i=2}^{N_\delta-1} X_i \geq 0] = p_+.$$ 

So,

$$\Pr\left[ \sum_{i=1}^k X_i \geq 0 \ i.o. \right] = \Pr[X_1 \geq x] \cdot \Pr[\sum_{i=2}^{N_\delta-1} X_i \geq 0] \cdot \Pr[\sum_{i=N_\delta}^k X_i > -x \ i.o.] \geq p_x \cdot p_+ \cdot (1 - \frac{\delta}{x^2}) > 0. \quad \square$$

**Theorem 7.**

(a) $\hat{\zeta}$ is weakly risk averse if and only if $u^k$ is concave with respect to $v^k$ for all $k$.

(b) $\hat{\zeta}$ is risk averse if and only if the sequence $(u^1, u^2, \ldots)$ is uniformly strictly concave with respect to $(v^1, v^2, \ldots)$.

**Proof.** (a) $\hat{\zeta}$ is weakly risk averse $\Rightarrow$ all $\hat{u}^k$ are concave: Contrariwise, suppose that $\hat{u}^k$ is not concave, for some $k$. So, $\hat{u}^k$ is not concave on some interval of size $\leq s$. So, there exist $x$, $\epsilon \leq s$ and $0 < \beta < \epsilon$ with

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

So, by the same reasoning as in the proof of Lemma 6, also for any $m > k$,

$$\hat{u}^m(x + \phi_{k+1} + \cdots + \phi_m + \beta) =$$

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

We construct a recurring lottery sequence $\ell$ 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 $\ell = (\ell_1, \ell_2, \ldots)$ is defined as follows:

- for $i = 1, \ldots, k$: $\ell_i = b_i$;
- for $j$ odd: $\ell_{k+j} = \frac{1}{2}(\phi_{k+j} - \epsilon) \oplus \frac{1}{2}(\phi_{k+j} + \epsilon)$;
- for $j$ even: $\ell_{k+j} = \phi_{k+j} - \beta$.

We now inductively determine the repeated certainty equivalent of $\ell = (\ell_1, \ell_2, \ldots)$, 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 $\ell_{k+1} = \frac{1}{2}(\phi_{k+1} - \epsilon) \oplus \frac{1}{2}(\phi_{k+1} + \epsilon)$. So, by (25), its certainty equivalent is $\beta$ above the average; that is, $c_{k+1} = \phi_{k+1} + \beta$. The next lottery, at time $k + 2$, is the degenerate lottery $\ell_{k+2} = \phi_{k+2} - \beta$, 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 (25) applies to the lottery at time $k + 3$, which is $\ell_{k+3} = \frac{1}{2}(\phi_{k+1} - \epsilon) \oplus \frac{1}{2}(\phi_{k+1} + \epsilon)$. 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, assume w.l.o.g. that \( E(\ell_i) = 0 \) for all \( i \). Then, \( c_{k+j} = \beta \) for \( j \) odd and \( c_i = 0 \) for all other \( i \). Then,

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

where the last equality is by the law of large numbers. So, \((\ell_1, \ell_2, \ldots)\) is ultimately inferior to \((c_1, c_2, \ldots)\), and \( \sim \) is not weakly risk averse, contrary to our assumption.

All \( \hat{u}_k \) are concave ⇒ \( \hat{u} \) is weakly risk averse: This follows from Lemma 8. Consider a lottery sequence \( \ell = (\ell_1, \ell_2, \ldots) \). W.l.o.g. \( E(\ell_i) = 0 \) for all \( i \). Denote by \( e = (c_1, c_2, \ldots) \) the repeated certainty equivalent of \( \ell \). Since all \( \hat{u}_k \)'s are concave, also all the functions \( u^k \) are concave in each of their arguments. Thus, \( c_i \leq 0 \) for all \( i \). Thus,

\[
\Pr[ (\ell_1, \ldots, \ell_k) \prec (c_1, \ldots, c_k) \text{ from some } k \text{ on} ] \leq \Pr[ \sum_{i=1}^{k} \ell_i < 0 \text{ from some } k \text{ on} ] = 1 - \Pr[ \sum_{i=1}^{k} \ell_i \geq 0 \text{ i.o.} ] < 1
\]

where the last inequality is by Lemma 8.

(b) \( \hat{u} \) is risk averse ⇒ \( (\hat{u}^1, \hat{u}^2, \ldots) \) is uniformly strictly concave: Contrariwise, suppose that \( (\hat{u}^1, \hat{u}^2, \ldots) \) is not uniformly strictly concave. If there exists a \( k \) for which \( \hat{u}^k \) is not concave, then in the proof of (a) we constructed a recurring lottery sequence that is ultimately inferior to its repeated certainty equivalent.

Thus, suppose that all \( \hat{u}_k \) are concave. Since \( (\hat{u}^1, \hat{u}^2, \ldots) \) is not uniformly strictly concave, there exists an \( \epsilon \) such that for any \( \delta > 0 \) there exist \( k \) and \( x \) with concavity of \( (x, \epsilon) \leq \delta \). By Lemma 5 we may assume that \( \epsilon < s \). By Proposition 4 the same also holds for horizontal concavity. So, for any \( \delta > 0 \) there exist \( k = k_\delta \) and \( x = x_\delta \) with \( h\text{-concavity of } (x, \epsilon) \leq \delta \). In particular, setting \( \delta = 2^{-j} \), we have that for any \( j \), there exist \( k_j \) and \( x_j \) with

\[
(26) \quad h\text{-concavity of } (x_j, \epsilon) \leq 2^{-j}.
\]

We construct a recurring fair gamble sequence \( \ell = (\ell_1, \ell_2, \ldots) \) that is not ultimately superior to its repeated certainty equivalent, which we denote by \( (c_1, c_2, \ldots) \). The construction of \( \ell \) is inductive. Suppose \((\ell_1, \ldots, \ell_j)\) has been defined, and that its repeated certainty equivalent is \((c_1, \ldots, c_j)\). Let \( k_j, x_j \) be as in (26). Set \( y_j = c_1 + \cdots + c_j \). By Lemma 7, there exists \( m > \max\{j, k_j\} \) and \( b_{j+1}, \ldots, b_m \), with \n
\[
h\text{-concavity of } (y_j + b_{j+1} + \cdots + b_m, \epsilon) \leq 2^{-j}.
\]

Hence also (moving to \( m+1 \))\textsuperscript{29},

\[
h\text{-concavity of } (y + b_{j+1} + \cdots + b_m + \phi_{m+1}, \epsilon) \leq 2^{-j},
\]

\textsuperscript{29}We move to \( m+1 \) with \( \phi_{m+1} \) to guarantee sufficient distance from the boundaries to allow a \( \pm \epsilon \) lottery.
which means that
\[
\hat{u}^{m+1}(y + b_j + \cdots + b_m + \phi_{m+1} - \frac{\epsilon}{2j}) \geq \frac{1}{2}(\hat{u}^{m+1}(y + b_j + \cdots + b_m + \phi_{m+1} - \epsilon) + \hat{u}^{m+1}(y + b_j + \cdots + b_m + \phi_{m+1} + \epsilon)).
\]

Accordingly, set \( \ell_i = b_i \) for \( i = j + 1, \ldots, m \) and \( \ell_{m+1} = \frac{1}{2}(\phi_{m+1} - \epsilon) + \frac{1}{2}(\phi_{m+1} + \epsilon) \). Then, \( c_i = b_i \) for \( i = j + 1, \ldots, m \) and \( c_{m+1} \geq \phi_{m+1} - \epsilon/2j \).

We now show that \((c_1, c_2, \ldots)\) is not ultimately inferior to \((\ell_1, \ell_2, \ldots)\). W.l.o.g. \( E(\ell_i) = 0 \) for all \( i \). Thus, by construction \( c_i \geq -\epsilon/2j \). Thus, for any \( k \)
\[
\sum_{i=1}^{k} c_i \geq \sum_{i=1}^{\infty} -\frac{\epsilon}{2j-1} = -2\epsilon.
\]

Denote by \( n_k \) the number of non-degenerate lotteries in the first \( k \) time periods. Each such non-degenerate lottery is a fair \( \pm \epsilon \) lottery. So,
\[
\Pr[(c_1, \ldots, c_k) \prec (\ell_1, \ldots, \ell_k) \text{ from some } k \text{ on}] = \Pr\left[ \sum_{i=1}^{k} c_i < \sum_{i=1}^{k} \ell_i \text{ from some } k \text{ on} \right] \leq \Pr\left[ -2\epsilon < \sum_{i=1}^{k} \ell_i \text{ from some } k \text{ on} \right] \leq \lim_{k \to \infty} \Pr\left[ -2\epsilon < \sum_{i=1}^{k} \ell_i \right] = \lim_{k \to \infty} \Pr\left[ \frac{-2\epsilon}{n_k} < \frac{\sum_{i=1}^{k} \ell_i}{n_k} \right] < 1,
\]

where the last inequality is by the central limit theorem (since \( n_k \to \infty \) as \( k \to \infty \) and so \( \frac{-2\epsilon}{n_k} \) approaches 0).

\((\hat{u}^1, \hat{u}^2, \ldots)\) is uniformly strictly concave \( \Rightarrow \) is risk averse: Let \( \ell \) be a recurring lottery sequence. Let \( G = \{i_1, i_2, \ldots\} \), be the set of indices of the non-degenerate lotteries in \( \ell \). W.l.o.g. assume that \( E(\ell_i) = 0 \) for all \( i \). Then, there is some \( \epsilon \) with \( \ell_{i_j} = \frac{1}{2}(-\epsilon) + \frac{1}{2}\epsilon \) for \( i_j \in G \), and \( \ell_i = 0 \) for all other \( i ' s \). Let \( c = (c_1, c_2, \ldots) \) be the repeated certainty equivalent of \( \ell \). Since \((\hat{u}^1, \hat{u}^2, \ldots)\) is uniformly strictly concave, by definition and Proposition 4 there exists a \( \delta \) such that \( h\text{-concavity}_{\hat{u}k}(c_1 + \cdots + c_{k-1}, \epsilon) \geq \delta \), for all \( k \). Thus, \( c_{i_j} \leq -\delta \epsilon \), for \( i_j \in G \), and \( c_i = 0 \) otherwise.

Denote by \( n_k \) the number of non-degenerate lotteries in the first \( k \) time periods. Then,
\[
\Pr[(\ell_1, \ldots, \ell_k) \prec (c_1, \ldots, c_k) \text{ i.o.}] = \Pr[\sum_{i=1}^{k} \ell_i \leq \sum_{i=1}^{k} c_i \text{ i.o.}] = \Pr[\frac{\sum_{i=1}^{k} \ell_i}{n_k} \leq -\delta \epsilon \text{ i.o.}] = 0,
\]

where the last equality is by the law of large numbers.

\( \square \)

Theorem 6 is now a direct corollary of Theorem 7.

**Theorem 6.** If \( \hat{\sim} \) is risk averse then it is weakly risk averse.

**Proof.** Follows from Theorem 7 since any strictly concave function is also concave. \( \square \)
Theorem 9.

(a) $\succeq$ is weakly risk loving if and only if $u^k$ is convex with respect to $v^k$ for all $k$.

(b) $\succeq$ is risk loving if and only if the sequence $(u^1, u^2, \ldots)$ is uniformly strictly convex with respect to $(v^1, v^2, \ldots)$.

(c) $\succeq$ is risk neutral if and only if $u^k$ is linear in $v^k$ for all $k$.

Proof. The proof of (a) and (b) is analogous to that of Theorem 7. (c) follows from Theorem 7-(a) and 9-(a). □

Theorem 8. If $\succeq$ is risk loving then it is weakly risk loving.

Proof. Follows directly from Theorem 9. □