

A CONCEPTUAL FOUNDATION FOR THE THEORY OF RISK AVERSION

YONATAN AUMANN

Bar Ilan University

Ramat Gan, Israel

ABSTRACT. Classically, risk aversion is equated with concavity of the utility function. In this work we explore the conceptual foundations of this definition. In accordance with neo-classical economics, we seek an ordinal definition, based on the decisions 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 risk aversion on non-liquid goods, multi-commodity risk aversion, and 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? 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, and not some other function (say, the geometric mean)? 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.

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

Date: September 26, 2015.

In addition, there is the matter of units (or scale). Consider a decision maker having to choose between lotteries on the temperature-level in her office room. If she prefers 40° F with certainty over a fair gamble between 30° and 60° – should this be considered risk aversion? The Fahrenheit scale seems rather arbitrary in this case, but it is not clear what other scale should or can be used. In the seminal works of Arrow [2] and Pratt [22], risk aversion was defined with respect to money and the market value of the goods. This, however, limits the notion to fully liquid goods. Also, using the market value suggests that, if not for risk, the attitude towards money should in some sense be linear in the money amount; but why?

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 scale.

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 concavity of the utility function with respect to, say, the square root of the money amount, instead of the amount itself, we will get a different definition). Conceptually, the notion of expectation and the associated *risk premium* are only meaningful in a cardinal framework in which quantities 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 these definitions 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 “appropriate” units.

Ordinal Definition I: Repeated Gambles. 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 a preference, however, is ruled out by the von Neumann-Morgenstern (NM) axioms; the utility of a lottery must lie between the utilities of its possible outcomes. Interestingly, while such a preference is indeed not possible for any single gamble, it *is* possible once we consider *sequences* of gambles, and 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 inferior to what would have been the outcome of the gambles, with probability 1. This is thus our first definition of risk aversion: a preference order is deemed risk-averse if adhering to this preference order 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 sequences, not any external market-based criterion.

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

Ordinal Definition II: Hedging. Our second ordinal definition of risk aversion is set in the context of a multi-commodity space, and equates risk aversion with a preference for hedging bets, whenever possible.¹ 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) .

Thus, risk aversion is equated with a preference for hedging bets whenever possible. Note that this definition is fully ordinal; it uses only the ordinal preferences on commodity bundles, with no reference to any quantitative measure.²

The above definition considers a setting with two commodities. A similar definition also applies to multi-commodity settings, wherein the commodities are partitioned into separate *independent factors*³ and hedging takes place between two such factors. 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 factors, a

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

²The definition does require *independence* of the commodities, but not *additive separability*. So, a cardinal representation is not assumed.

³The exact definition of *independent factors* is provided in Section 2.

decision maker prefers 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.

A 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 [5, 6], and commonly used in the theory of multi-attribute decision theory (see [17]). 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. Essentially, we show that the NM utility function is concave with respect to the associated Debreu function if and only if the given preference order is risk averse, under either of the two ordinal definitions.

1.3. Implications. The approach we offer has several implication for the understanding of risk aversion, both conceptual and technical.

Non-Liquid Goods. Our approach provides, for the first time, a way to define risk aversion for non-liquid goods and goods with no natural scale, such as temperature, pain, and pleasure. Indeed, in our definition externally defined scales (such as market value) do not play any role. Rather, the only scale of interest is the intrinsic Debreu value, which reflects the decision maker’s own certainty preferences.

Multi-commodity Risk Aversion. Ever since the publications of Arrow [2] and Pratt [22], researchers have attempted to extend the definition and associated measures to the multi-commodity setting, and various approaches have been suggested (see [18, 25, 21, 8, 15, 23, 16, 19] for some references in the expected utility model). The basic problem is that in the multi-commodity setting each commodity has its own scale, so it is not clear what scale should be used when measuring the concavity of the utility function. Our approach here provides a simple and conceptually well-founded solution. In our approach the “native” scales of the different commodities are immaterial. Rather, we measure risk aversion with respect to the intrinsically defined Debreu value function. This value function is shared across all commodities, so there is only one relevant scale. We show that this approach extends the Arrow-Pratt framework to the multi-commodity setting, and even

allows comparisons amongst decision makers who do not agree on the certainty preferences, under certain conditions (see Section 8).

Financial Planning and Insurance. Perhaps the most fundamental implication of this work is that it gives us a new approach, and language, for understanding behavior in areas such as financial planning and insurance. Under classical economic language a person rejecting a fair gamble between tripling her savings and going broke would be deemed “risk averse”. But rejecting such a gamble seems to be common sense even without risk aversion; the upside of tripling the savings somehow seems “much less” than the down-side of going broke (in an intuitive, perhaps not well defined, sense). Classical economic language, however, lumps together all possible reasons for rejecting the gamble, thus entangling the risk-attitude with the certainty preferences. Our approach here allows disentangling the two.⁴ Under our definitions, rejecting the gamble may still be deemed risk neutral, if the Debreu utility is concave with respect to money. Interestingly, under our approach, insurance also need not be tied to risk aversion, as we argue in Section 9

1.4. Assumptions.

Independence. Independence is a key notion and assumption throughout this work. Simply put, a commodity, or set of commodities, is *independent* if the preference order over bundles of this set of commodities is independent of the state in other commodities.⁵ Arguably, independence is a strong assumption; having eaten Chinese food today may affect one’s gastronomical preferences tomorrow. Nonetheless, independence is a common assumption in economic literature, and in particular with respect to time preferences; in particular, 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.

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 nonetheless is a possible model; and one that is frequently used in real-world economic and financial applications. So, understanding the notion of risk aversion within this framework is of interest. Extending these ideas to non-EU models is an interesting future research direction.

⁴Disentangling diminishing marginal utility from risk aversion is one of the earliest motivations for the non-expected utility literature, see Yaari [26]. In this work, however, we remain within the expected utility framework.

⁵A formal definition is provided in the next section.

1.5. 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 presented in Section 3, and its quantitative form in Section 4. Section 5 presents the second definition, with its equivalent quantitative form in Section 6. The relationship between the two definitions is discussed in Section 7. The applications to multi-commodity risk aversion are discussed in Section 8. We conclude the main body of the paper with a discussion in Section 9. All proofs are deferred to an appendix.

2. TERMINOLOGY AND NOTATION

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

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

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

- *the certainty preferences:* a preference order \succsim on \mathcal{S} ,⁶
- *the lottery preferences:* a continuous preference order $\dot{\succsim}$ on $\Delta(\mathcal{S})$, which agrees with \succsim on the sure outcomes.

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

All commodity spaces \mathcal{C}_i are assumed to be *strictly essential* [13]; that is, for each i and $s_{-i} \in \mathcal{C}_{-i}$ (the remaining commodities), there exist $s_i, s'_i \in \mathcal{C}_i$ with $(s_i, s_{-i}) \not\sim (s'_i, s_{-i})$.

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

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

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

$$(1) \quad (a_i, a_j, \mathbf{c}) = (c_1, \dots, c_{i-1}, a_i, c_{i+1}, \dots, c_{j-1}, a_j, c_{j+1}, \dots, c_n).$$

⁶A *preference order* is a complete, transitive and reflexive binary relation.

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

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

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

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

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

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

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

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

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

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

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

When no confusion can result, we may write \precsim instead of $\precsim^{\mathcal{T}}$; thus, when $a, a' \in \mathcal{T}$, we may write $a \precsim a'$ instead of $a \precsim^{\mathcal{T}} a'$. It is worth noting that the product of independent factors need not be independent.⁷

A partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ is an *independent partition* if the product of any subset of factors is independent. By Gorman [13], for $n \geq 3$, it suffices to assume that $\mathcal{T}_i \times \mathcal{T}_{i+1}$ is independent for all i , and the independence of all other products then follows.

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

⁷A 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.

3. ORDINAL DEFINITION I: REPEATED LOTTERIES

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

The Space. We consider an infinite sequence of factors $\mathcal{T}_1, \mathcal{T}_2, \dots$, where \mathcal{T}_i represents the consumption space at time i .⁸ We denote $\mathcal{H}^n = \mathcal{T}_1 \times \dots \times \mathcal{T}_n$ - the finite history space up to time n . In the following, a_i, b_i, c_i , will always be taken to be in \mathcal{T}_i , and lottery L_i will be over \mathcal{T}_i .

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

We call the sequence of preference orders $\dot{\succsim} = (\dot{\succsim}^1, \dot{\succsim}^2, \dots)$ the *risk policy*.

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

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

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

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

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

⁸We do not assume that $\mathcal{T}_i = \mathcal{T}_j$, i.e. the state spaces need not be the same at different time periods. In particular, we do not assume any form of stationarity (though it is possible). Similarly, discounting may or may not be applied between consecutive factors. Our discussion here is independent of any such nominal matters.

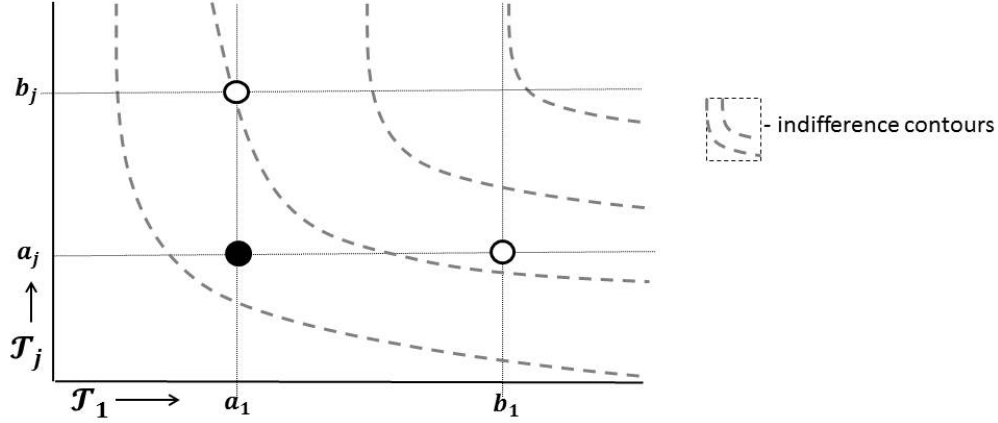


FIGURE 1. Illustration of $[a_j, b_j] \subseteq [a_1, b_1]$ (the factors of $\mathcal{S}_{-\{1,j\}}$ are not depicted).

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

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

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

Bounded and Non-Vanishing Lottery Sequences. We now want to define risk aversion as a policy for which the repeated certainty equivalent of a lottery sequence is always ultimately inferior to the lottery sequence itself. However, as such, this definition cannot be a good one since in the case that the “magnitude” of the lotteries rapidly diminishes the overall outcome will be dominated by that of the first lotteries, and we could never obtain an inferior outcome with probability 1. Similarly, if the “magnitude” of the lotteries can grow indefinitely, then for almost any risk policy one can construct a lottery sequence that is ultimately inferior to its repeated certainty equivalent.¹⁰ Hence, we now define the notions of a *bounded* lottery sequence and a *non-vanishing* lottery sequence.

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

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

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

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

¹⁰See Appendix B.

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

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

Definition 1. We say that risk policy \succsim is:

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

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

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

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

Definition 2. We say that risk policy \succsim is:

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

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

4. REPEATED LOTTERIES: THE QUANTITATIVE PERSPECTIVE

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

4.1. Debreu Value Functions. The theory of multi-attribute decision making considers certainty preferences over a multi-factor space, and establishes that under certain independence assumptions such preferences can be represented by quantitative functions, as follows. Consider the space $\mathcal{H}^n = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ ($n \geq 2$), with preference order \preceq^n . Debreu [5] proves that, if the partition is

independent¹¹ then \lesssim^n is *additively separable*;¹² that is, there exist functions $v^{\mathcal{T}_i} : \mathcal{T}_i \rightarrow \mathbb{R}$, such that for any $(a_1, \dots, a_n), (a'_1, \dots, a'_n)$

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

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

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

We call the function $v^{\mathcal{T}_i}$ a (*Debreu*) *value function* for \mathcal{T}_i , and the aggregate function $v_n = \sum_{i=1}^n v^{\mathcal{T}_i}$ a (*Debreu*) *value function* for \mathcal{H}^n .¹³ We note that Debreu [5] called these functions *utility functions*; but following Keeney and Raiffa [17], we use the term *value functions*, to distinguish them from the NM utility function.

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

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

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

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

$$L \stackrel{\Delta}{\lesssim} L' \iff (L, \phi_{n+1}) \stackrel{\Delta}{\lesssim}^{n+1} (L', \phi_{n+1});$$

that is, the preferences on $\Delta(\mathcal{H}^n)$ are consistent with *some* possible future. We call the sequence (ϕ_2, ϕ_3, \dots) a *presumed future*, and assume that it is *internal*,¹⁴ in the following sense. The sequence

¹¹see page 7.

¹²In the case of two factors ($n = 2$), the following *Thomsen condition* is also required: for all $a_1, B_1, c_1 \in \mathcal{T}_1$, and $a_2, b_2, c_3 \in \mathcal{T}_2$, if $(a_1, b_2) \sim (b_1, a_2)$ and $(b_1, c_3) \sim (c_1, b_2)$ then $(a_1, c_2) \sim (c_1, a_2)$. For $n > 2$ the Thomsen condition is implied by the independence of the pairs.

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

¹⁴More precisely, we assume that there exists a presumed future that is internal.

(ϕ_2, ϕ_3, \dots) is *internal* if there exists an $s > 0$ with $v^{\mathcal{T}_i}(\phi_i) \pm s \in v^{\mathcal{T}_i}(\mathcal{T}_i)$ for all i ; that is, the presumed future is bounded away from the boundaries of the \mathcal{T}_i 's.

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

Theorem 4.2. $\stackrel{\Delta}{\sim}$ is weakly risk averse if and only if u_n is concave with respect to v_n for all n .

Thus, Theorem 4.2 provides the missing conceptual justification for defining risk aversion by concavity of the utility function. It also establishes the appropriate scale - the Debreu value function.

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

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

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

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

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

Theorem 4.3. If $A_{\hat{u}_n}(x)$ is bounded away from 0, uniformly for all n and x ,¹⁵ then $\stackrel{\Delta}{\sim}$ is risk averse (assuming \hat{u}_n is twice differentiable for all n).

Theorem 4.3 establishes a sufficient condition for risk aversion. We now proceed to establish a necessary condition, which is “close” to being tight. To do so we need to consider the behavior of the functions \hat{u}_i , and the definition of $A_{\hat{u}_i}(\cdot)$, in a little more detail.

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

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

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

Now for any ϵ (sufficiently small) define

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

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

Theorem 4.4.

- (a) If $RP_{\hat{\mathbf{u}}}(\epsilon) = \Omega(\epsilon^2)$ as $\epsilon \rightarrow 0$ then $\hat{\mathbf{u}}$ is risk averse.¹⁶
- (b) If $RP_{\hat{\mathbf{u}}}(\epsilon) = O(\epsilon^{2+\beta})$ as $\epsilon \rightarrow 0$, for some $\beta > 0$, then $\hat{\mathbf{u}}$ is not risk averse.

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

Finally, we establish that the sufficient condition of Theorem 4.4-(a) and that of Theorem 4.3 are the same.

Proposition 4.5. $RP_{\hat{\mathbf{u}}}(\epsilon) = \Omega(\epsilon^2)$ as $\epsilon \rightarrow 0$, if and only if $A_{\hat{\mathbf{u}}_n}(x)$ is bounded away from 0, uniformly for all n and x (assuming $\hat{\mathbf{u}}_n$ is twice differentiable for all n).

4.3. Risk Loving and Risk Neutrality. In analogy to Theorems 4.2 and 4.4 we have:

Theorem 4.6. For v_n, u_n , and $\hat{\mathbf{u}}$ as in Theorems 4.2 and 4.4

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

5. ORDINAL DEFINITIONS II: HEDGING

5.1. The Definition. Consider a space \mathcal{S} and an independent partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$.¹⁷

Recall the notation $\mathcal{S}_{-\{i,j\}} = \prod_{t \neq i,j} \mathcal{T}_t$, and for a_i, a_j , and $\mathbf{c} \in \mathcal{S}_{-\{i,j\}}$, the slight abuse of notation (a_i, a_j, \mathbf{c}) for $(c_1, \dots, c_{i-1}, a_i, c_{i+1}, \dots, c_{j-1}, a_j, c_{j+1}, \dots, c_n)$.

Definition 3. For $a_i \prec b_i, a_j \prec b_j$, say that $(a_i, b_j), (b_i, a_j)$ are perfectly hedged if $(a_i, b_j) \sim (b_i, a_j)$ (see Figure 2).¹⁸

We say that $\hat{\mathbf{u}}$ is ordinally risk-averse (with respect to the partition $\mathcal{T}_1 \times \cdots \times \mathcal{T}_n$) if there exists $i \neq j$, such that for any perfectly hedged $(a_i, b_j), (b_i, a_j)$, and $\mathbf{c} \in \mathcal{S}_{-\{i,j\}}$

$$(2) \quad \langle (a_i, a_j, \mathbf{c}), (b_i, b_j, \mathbf{c}) \rangle \succ \langle (a_i, b_j, \mathbf{c}), (b_i, a_j, \mathbf{c}) \rangle.$$

¹⁶recall that $g(y) = \Omega(h(y))$ as $y \rightarrow 0$ if there exists a constant M and y_0 such that $g(y) > M \cdot h(y)$ for all $y < y_0$.

¹⁷Here we use the notation \mathcal{S} rather than \mathcal{H}^n since we will be considering one fixed space \mathcal{S} .

¹⁸The equivalence relation $(a_i, b_j) \sim (b_i, a_j)$ is well-defined as $\mathcal{T}_i \times \mathcal{T}_j$ is independent.

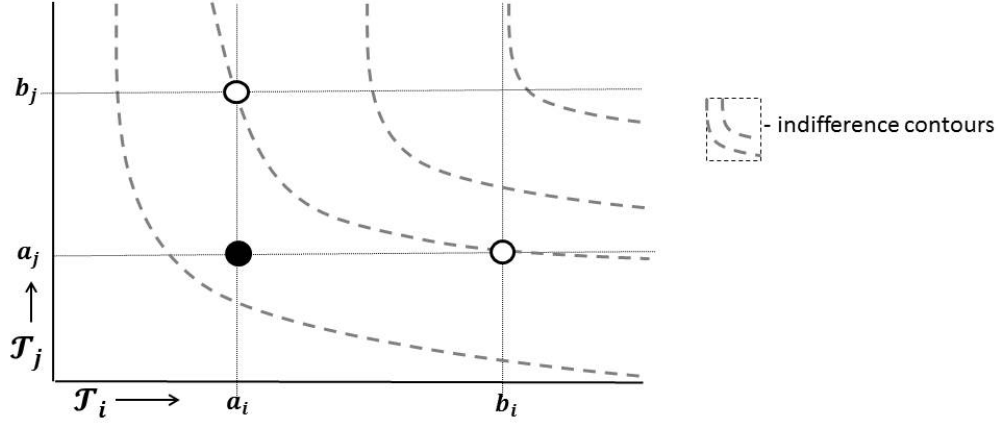


FIGURE 2. Illustration of a perfectly hedged pair.

The preference order is weakly ordinally risk averse (with respect to the given partition) if (2) holds with a weak preference (\succsim).

Thus, a decision maker is ordinally risk-averse if she prefers to hedge her bets whenever possible.

5.2. Properties. Definition 3 deems a decision maker risk averse if she prefers all hedges among *some* pair of factors $\mathcal{T}_i, \mathcal{T}_j$ ($i \neq j$). The following proposition establishes that in this case she prefers hedging among *any* pair of factors.

Proposition 5.1. *If \succsim is ordinally risk averse (by Definition 3) then (2) holds for all $i \neq j$ (and any fully hedged $(a_1, b_2), (b_1, a_2)$, and \mathbf{c}). Similarly for weak risk aversion.*

Definition 3 considers a specific partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$. 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 proposition.

Proposition 5.2. *If \succsim is (weakly) ordinally risk averse with respect to some independent partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$, then it is also so with respect to any independent partition.*

By Proposition 5.2, we may call \succsim ordinally risk averse if it is ordinally risk averse with respect to some partition, and so for all partitions.

5.3. Ordinal Risk Aversion and Correlation Aversion. Richard [23] defined the following notion of *correlation aversion* (see also [11, 7]):¹⁹ preference order \succsim is *correlation averse* with

¹⁹Actually, Richard used the term *multivariate risk aversion*. The now common term *correlation aversion* was later coined by Epstein and Tanny [11].

respect to $\mathcal{T}_i, \mathcal{T}_j$, if for any $a_i \prec b_i, a_j \prec b_j$, and $\mathbf{c} \in \mathcal{S}_{-\{i,j\}}$

$$\langle (a_i, a_j, \mathbf{c}), (b_i, b_j, \mathbf{c}) \rangle \succ^{\triangleleft} \langle (a_i, b_j, \mathbf{c}), (b_i, a_j, \mathbf{c}) \rangle.$$

The difference between this definition of correlation aversion and our definition of risk aversion is that correlation aversion does not require that $(a_i, b_j) \sim (b_i, a_j)$. Thus, correlation aversion requires that the decision maker 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 5.3. \succsim^{\triangleleft} is ordinally risk-averse if and only if it is correlation averse with respect to any and all $\mathcal{T}_i, \mathcal{T}_j$.

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

5.4. Risk Loving and Risk Neutrality.

Definition 4. We say that \succsim^{\triangleleft} is ordinally risk-loving if for any $i \neq j$, and any perfectly hedged $(a_i, b_j), (b_i, a_j)$, and any $\mathbf{c} \in \mathcal{S}_{-\{i,j\}}$

$$(3) \quad \langle (a_i, a_j, \mathbf{c}), (b_i, b_j, \mathbf{c}) \rangle \succ^{\triangleleft} \langle (a_i, b_j, \mathbf{c}), (b_i, a_j, \mathbf{c}) \rangle.$$

and weakly ordinally risk loving if the preference in (3) is a weak one.

Similarly, \succsim^{\triangleleft} is ordinally risk-neutral if for any $i \neq j$, and any perfectly hedged $(a_i, b_j), (b_i, a_j)$, and any $\mathbf{c} \in \mathcal{S}_{-\{i,j\}}$

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

Propositions 5.2 and 5.1 hold analogously for risk loving and risk neutrality.

6. HEDGING DEFINITION: THE QUANTITATIVE PERSPECTIVE

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

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

Theorem 6.1. *For any \mathcal{S} , all (aggregate) Debreu value functions for \mathcal{S} are identical up to positive affine transformations.*

6.2. Risk Aversion and the Debreu Value Functions. Assume that the conditions guaranteeing the existence of a Debreu value function for \mathcal{S} hold.²⁰ We now show that in this case, the ordinal Definition 3 coincides with concavity of the utility function with respect to the value function.

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 6.2. *Let v be a Debreu value function for \mathcal{S} , and $(a_i, b_j), (b_i, a_j)$ perfectly hedged. Then for any $\mathbf{c} \in \mathcal{S}_{-\{i,j\}}$.*

$$v(a_i, b_j, \mathbf{c}) = v(b_i, a_j, \mathbf{c}) = \frac{v(a_i, a_j, \mathbf{c}) + v(b_i, b_j, \mathbf{c})}{2}.$$

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.

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

- *Risk aversion:*
 - *u is strictly concave with respect to v if and only if \succsim is ordinally risk averse.*
 - *u is concave with respect to v if and only if \succsim is weakly ordinally risk averse.*
- *Risk loving:*
 - *u is strictly convex with respect to v if and only if \succsim is ordinally risk loving.*
 - *u is convex with respect to v if and only if \succsim is weakly ordinally risk loving.*
- *Risk neutrality: u is linear with respect to v if and only if \succsim 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.

7. RELATING THE TWO ORDINAL DEFINITIONS

We provided two ordinal definitions of risk aversion: Definition 1, based on repeated lotteries, and Definition 3, based on hedging. Technically, the two definitions relate to different mathematical objects: the first relates to a *risk policy*, which is a sequence of preference orders, while the latter relates to a single preference order. However, the two definitions are closely related, as established by Theorems 4.4 and 6.3: both definitions correspond to concavity of the NM utility function with respect to the Debreu value function (if and when such a value function exists). For weak risk

²⁰If there are three or more factors in the partition, then the existence of a value function is provided by the independence of the partition. If there are only two factors, the additional Thomsen condition is required (see Footnote 12).

aversion the concavity requirements in both theorem are identical – (weak) concavity. So, when a Debreu value exist, a risk policy is weakly risk averse, according to Definition 1, if and only if each of the preferences orders therein is weakly risk averse, according to the Definition 3. For (strict) risk aversion, the requirement in Theorem 4.4 is a coefficient of absolute risk aversion bounded away from zero, whereas Theorem 6.3 requires only 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 a “recurring” bound on the strict concavity in all the gambles.

8. MULTI-COMMODITY RISK AVERSION

The seminal works of Arrow [2] and Pratt [22] defined risk aversion with respect to a single commodity – money. Ever since, researchers have attempted to extend the definition, and associated measures, to the multi-commodity setting (see [18, 25, 21, 8, 15, 23, 16, 19] for some references in the expected utility model). It is out of the scope of this paper to review this extensive body of research, but the underlying problem addressed in these works is that in the multi-commodity setting each commodity has its own scale so the question is which scale should be used when measuring the concavity of the utility function. Indeed, the solution in several of these works was to keep the multiple scales - in which case the measures of risk aversion become vectors and matrices (see e.g. [8]).

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

We note, again, that our hedging based ordinal definition - Definition 3 - is very close in spirit to Richard’s definition of *multivariate risk aversion* [23] (see Section 5.3). Richard, however, viewed his definition as “a new type of risk aversion unique to multivariate utility functions” [23], and did not make the connection back to the classical definition (using the Debreu value function). Indeed, Scarsini, in a paper based on Richard’s definition, writes “[Richard’s definition] has nothing to do with what is generally known as risk aversion” [24]. We have shown that these two definitions are one and the same, once the appropriate scale is used.

We now show how, using our definition, the Arrow-Pratt framework carries over to the multi-commodity setting.

8.1. CARA Preferences. A (uni-scale) preference order is CARA (constant absolute risk aversion) if the coefficient of absolute risk aversion of its associated NM utility is constant. Arrow [2]

showed that a preference is CARA if and only if the preferences on lotteries are independent of the wealth level. Specifically, for wealth level x and lottery L denote by (L, x) the lottery that gives the random outcome L *in addition* to the sure outcome x . Then, Arrow shows that preference order $\overset{\circ}{\succsim}$ is CARA if and only if for all lotteries L, L' and wealth levels x, y

$$(L, x) \overset{\circ}{\succsim} (L', x) \iff (L, y) \overset{\circ}{\succsim} (L', y).$$

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

Theorem 8.1. *In the multi-commodity setting (with $\mathcal{S} = \mathcal{T}_1 \times \dots \times \mathcal{T}_n$ an independent partition), the NM utility function u has constant coefficient of absolute risk aversion when measured with respect to the Debreu value function v if and only if for any i , lotteries L, L' over \mathcal{T}_i , and $\mathbf{x}, \mathbf{y} \in \Omega_{-\{i\}}$*

$$(L, \mathbf{x}) \overset{\circ}{\succsim} (L', \mathbf{x}) \iff (L, \mathbf{y}) \overset{\circ}{\succsim} (L', \mathbf{y}).$$

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

In the literature, the condition that

$$(L, \mathbf{x}) \overset{\circ}{\succsim} (L', \mathbf{x}) \iff (L, \mathbf{y}) \overset{\circ}{\succsim} (L', \mathbf{y}).$$

for any $L, L' \in \Delta(\mathcal{T}_i)$, and $\mathbf{x}, \mathbf{y} \in \mathcal{S}_{-\{i\}}$ is termed *utility independence* of \mathcal{T}_i [17].

Proposition 8.2. *Let \succsim be an (additively separable) preference order on $\mathcal{S} = \mathcal{T}_1 \times \dots \times \mathcal{T}_n$, and g a real valued function on \mathcal{S} . Suppose that for any $\overset{\circ}{\succsim}$ the corresponding NM utility function u has constant coefficient of absolute risk aversion when measured with respect to g if and only if all factors \mathcal{T}_i are utility independent. Then g is a Debreu value function.*

8.2. Comparative Multi-Commodity Risk Aversion. One of the greatest benefits of the Arrow-Pratt framework is that it provides a tool for comparing risk attitudes among different decision makers, and establishes a concrete measure for doing so (the coefficient of absolute risk aversion). In the uni-scale setting if decision maker 1 rejects a lottery accepted by decision maker 2, then it is natural to say that 1 is more risk averse than 2 (on this lottery). When moving to the multi-commodity setting this no longer holds, as observed by Kihlstrom and Mirman [18]. In the multi-commodity setting the differences between 1 and 2 may stem from differences in their *certainty* preferences, whereby the lottery outcomes are superior in the eyes of 2 and inferior in the eyes of 1. So, risk attitude comparisons in the multi-commodity setting must also take the certainty preferences into account.

Some works have thus limited risk aversion comparisons to individuals agreeing on the certainty preferences (see [18, 21]). Such comparisons are also natural in our approach, as our definition of

risk aversion is always *with respect to* the certainty preferences. For individuals agreeing on the certainty preferences, using our approach the entire Arrow-Pratt framework carries over as is, once the underlying scale is converted to the associated (joint) Debreu value function. In particular, we have the following. Let v be the joint Debreu value function and \hat{u}_1 and \hat{u}_2 be the NM utility functions of players 1 and 2 - when measured with respect to v . For a lottery L let $ce_j(L)$ be the certainty equivalent of L by \hat{u}_j ($j = 1, 2$). Then

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

for all lotteries L if and only if

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

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

Non-identical Certainty Preferences. Interestingly, measuring risk aversion with respect to the Debreu value function also allows comparisons among individuals who *do not* fully agree on the certainty preferences, under some conditions.

Consider a space \mathcal{S} , and for $j = 1, 2$, consider

- \precsim_j - a preference order on \mathcal{S} ,
- $\dot{\precsim}_j$ - a preference order on $\Delta(\mathcal{S})$ agreeing with \precsim_j on \mathcal{S} ,

Consider an independent factor upon which \precsim_1 and \precsim_2 agree. For such a factor, it is possible to define the notion that one preference order is *more risk averse* than the other.

Definition 5. Let \mathcal{I} be an independent factor upon which $\dot{\precsim}_1$ and $\dot{\precsim}_2$ agree. We say that $\dot{\precsim}_2$ is more risk averse than $\dot{\precsim}_1$ on \mathcal{I} if for any lottery L over \mathcal{I} and any $\mathbf{b} \in \mathcal{S}_{-\mathcal{I}}$ (the remaining factor)

$$ce_2(L, \mathbf{b}) \precsim ce_1(L, \mathbf{b})$$

where $ce_j(L, \mathbf{b})$ is a certainty equivalent of (L, \mathbf{b}) according to $\dot{\precsim}_j$ wherein the state in $\mathcal{S}_{-\mathcal{I}}$ is \mathbf{b} , and \precsim denotes the preference agreed by both \precsim_1 and \precsim_2 .

The factor \mathcal{I} could be a single commodity, or any other independent factor.

We now wish to establish conditions, in terms of the value and utility functions, that provide that $\dot{\precsim}_2$ is more risk averse than $\dot{\precsim}_1$ on \mathcal{I} .

By way of example, suppose that $\mathcal{S} = \mathcal{A} \times \mathcal{B}$ is an independent partition (by both preference orders), and that the orders agree on the preferences on both \mathcal{A} and \mathcal{B} . However, the two preference orders do not agree on the preferences on the entire space $\mathcal{A} \times \mathcal{B}$. Suppose that the preference \precsim_1 is represented by the value function $v_1(a, b) = a + 2b$, and \precsim_2 is represented by $v_2(a, b) = \ln(a) + b$. Finally, suppose that the NM utility functions are $u_1(a, b) = (v_1(a, b))^{\frac{1}{2}}$ and $u_2(a, b) = (v_2(a, b))^{\frac{1}{3}}$.

Now, based on these value and utility functions, we would like to determine on which factors is $\dot{\precsim}_2$ more risk averse than $\dot{\precsim}_1$ (if such a factor exists). Note that with respect to the value function,

u_2 is more concave than u_1 . However, this does not necessarily guarantee that $\overset{\circ}{\succsim}_2$ is more risk averse (in the sense of Definition 5) as the value functions are different. So, we want to establish a condition, considering both the utility and the value functions, that does guarantee “more risk aversion”. The condition we will provide establishes that $\overset{\circ}{\succsim}_2$ is more risk averse on \mathcal{A} but not necessarily so on \mathcal{B} .

For $j = 1, 2$ consider

- v_j - an aggregate value function representing \succsim_j ,
- u_j - an NM utility representing $\overset{\circ}{\succsim}_j$,
- \hat{u}_j - the NM utility u_j when the scale is converted to u_j (that is $\hat{u}_j = u_j \circ v_j^{-1}$).

Note that \hat{u}_1 and \hat{u}_2 operate on (conceptually) different domains; \hat{u}_1 operates on the image of v_1 , while \hat{u}_2 operates on the image of v_2 . So, for any given x , $\hat{u}_1(x)$ and $\hat{u}_2(x)$ may give the utility corresponding to totally different points of \mathcal{S} .

Let $\mathcal{S} = \mathcal{T}_1 \times \dots \times \mathcal{T}_n$ be an independent, additively separable partition according to both \succsim_1 and \succsim_2 . Let v_j^i be a Debreu value function for \mathcal{T}_i under \succsim_j ($i = 1, \dots, n, j = 1, 2$). We say that v_2^i is a concave expansion of v_1^i if for any $a_i, b_i, c_i \in \mathcal{T}_i$ (with $v_1^i(a_i) < v_1^i(b_i) < v_1^i(c_i)$)

$$(4) \quad 1 \leq \frac{v_2^i(c_i) - v_2^i(a_i)}{v_1^i(c_i) - v_1^i(a_i)} \leq \frac{v_2^i(b_i) - v_2^i(a_i)}{v_1^i(b_i) - v_1^i(a_i)}.$$

The first inequality says that it is an expansion, and the second that it is concave. In the above example, v_2^A is a concave expansion of v_1^A , but v_2^B is not so of v_1^B ,

The following theorem says that on factors for which v_2^i is a concave expansion of v_1^i , we can use the coefficient of absolute risk aversion to obtain that $\overset{\circ}{\succsim}_2$ is more risk averse than $\overset{\circ}{\succsim}_1$.

Theorem 8.3. *If $A_{\hat{u}_2}(x) \geq A_{\hat{u}_1}(y) \geq 0$, for all x, y , then the following holds. Let $\mathcal{I} = \prod_{i \in I} \mathcal{T}_i$, for some $I \subseteq \{1, \dots, n\}$, be such that*

- \succsim_1 and \succsim_2 agree on the \mathcal{I} , and
- v_2^i is a concave expansion of v_1^i for all $i \in I$.

Then, $\overset{\circ}{\succsim}_2$ is more risk averse than $\overset{\circ}{\succsim}_1$ on \mathcal{I} .

9. DISCUSSION

We presented fully ordinal definitions of risk aversion, based entirely on the internal structure of preferences of the decision maker; independent of money or any other units. Our definitions rest on two intuitively appealing interpretations of risk aversion. The first equates risk aversion with a policy that, in the long run, necessarily leads to an inferior outcome. The second equates risk aversion with a preference for hedging bets. 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 as the basis for defining risk aversion, and, at the same time, establishes the appropriate scale 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.²¹ 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.

9.1. Insurance. Buying insurance is a prime example of behavior classically and universally attributed to risk aversion.²² Indeed, the entire insurance industry is based on the fact that, in total, insurers pay back less than what they collect.²³ 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 may change.

²¹For 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.

²²Arrow writes: “The risk aversion hypothesis owes its durability . . . to its success in giving a qualitative explanation of otherwise puzzling examples of economic behavior. The most obvious is insurance, which hardly needs elaboration” [2].

²³For 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.

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 under our ordinal definition. 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.

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.²⁴ 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 in terms of the value function. Suppose that the value function of money is logarithmic in the dollar amount.²⁵ Then the \$800 paid as premium are worth

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

²⁴For simplicity, the presentation here considers each year separately. In practice, disability indemnification, as well as premiums, are paid over many years.

²⁵An illustrative example of how such a logarithmic value function may emerge is as follows. 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. (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 a reasonable first approximation. Other functions, which 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.) 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 $(g(y)x, y)$. Hence, the function

$$f(x, y) = \frac{x}{g(y)},$$

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.

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

$$2\% \cdot 1.39 \approx 0.028,$$

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, costs money, so that the deal may not be actuarially favorable in terms of money. Still it may well be favorable in terms of the subjective benefits provided thereby (e.g. in terms of value function). So, buying insurance may very well be a perfectly rational choice for (ordinally) risk neutral and even some (ordinally) risk loving individuals.

9.2. CARA and CRRA. Arrow and Pratt defined concrete *measures* of risk aversion, namely the *coefficient of absolute risk aversion at x* , and the *coefficient of relative risk aversion at x* (defined as $-\frac{x \cdot u''(x)}{u'(x)}$). The measure of absolute risk aversion can naturally be converted to our definition of risk aversion, by simply considering the utility function *with respect to the Debreu value function*, as discussed in Section 8. The notion of relative-risk-aversion w.r.t. the value function, however, is not well defined, as the definition of relative risk aversion requires a well-defined zero point, and the value function is only defined up to an additive constant.²⁶

In Section 8 we proved that once considered w.r.t. the value function, constant-absolute-risk-aversion (CARA) has a simple and intuitive meaning. A preference order is CARA w.r.t. the value function if and only if the preferences over lotteries in each individual factor are well defined and independent of the state in the other factors; preferences over apple lotteries are independent of the

²⁶Indeed, 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.

available amount of oranges and preferences over orange lotteries are independent of the available amount of apples (this is termed *utility independence* in [23, 3, 17]).

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

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

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

9.3. Strength of Preference and Relative Risk Aversion. Dyer and Sarin [10] and Bell and Raiffa [3] 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 $x_1 \mapsto x_2$ over the the transition $y_1 \mapsto y_2$ (where x_1, x_2, y_1, y_2 are states). Assuming such preferences exist (and some additional technical conditions), the theory establishes that there exists a function f (termed *measurable value function* [9]) 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 $x_1 \mapsto x_2$ is preferred over the transition $y_1 \mapsto y_2$. Given such a function, Dyer and Sarin [10] define the notion of *relative risk aversion*²⁷ as the concavity of the NM utility function u with respect to the measurable value function f . Bell and Raiffa [3] similarly define the notion of *intrinsic risk aversion*.

Bell and Raiffa [3] 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 [10, Theorem 1]). Thus, technically our ordinal notion of risk aversion coincides with the Dyer and Sarin notion of *relative risk aversion*, if a Debreu value function exists and relative risk aversion is computed with respect to this function. Conceptually, however, our approach is totally different from that of [10] and [3]. 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 [10] and [3] 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

²⁷not to be confused with the Arrow-Pratt coefficient of relative risk aversion

all core definitions, are fully ordinal. The numerical representation is then mathematically *derived* from this ordinal theory.

REFERENCES

- [1] N. Alon and J.H. Spencer. *The Probabilistic Method*. Wiley Series in Discrete Mathematics and Optimization. Wiley, 2004.
- [2] Kenneth J. Arrow. *Essays in the theory of risk-bearing*. Markham Publishing Company, 1971.
- [3] David E Bell and Howard Raiffa. Marginal value and intrinsic risk aversion. In David E. Bell, Howard Raiffa, and Amos Tversky, editors, *Decision Making Descriptive, Normative and Prescriptive Interactions*, pages 384–397. Cambridge University Press, 1988.
- [4] G Debreu. Representation of a preference ordering by a numerical function. In R.M. Thrall, R. L. Davis, and C.H. Coombs, editors, *Decision Process*, pages 159–165. John Wiley, New York, 1954.
- [5] Gerard Debreu. Topological Methods in Cardinal Utility Theory. Technical Report 76, Cowles Foundation for Research in Economics, Yale University, 1959.
- [6] Gerard Debreu. Topological methods in cardinal utility theory. In *Mathematical Economics: Twenty Papers of Gerard Debreu*, chapter 9. Cambridge University Press, 1986.
- [7] Michel Denuit, Louis Eeckhoudt, and Batrice Rey. Some consequences of correlation aversion in decision science. *Annals of Operations Research*, 176(1):259–269, April 2010.
- [8] George T. Duncan. A matrix measure of multivariate local risk aversion. *Econometrica*, 45(4):895–903, 1977.
- [9] James S. Dyer and Rakesh K. Sarin. Measurable multiattribute value functions. *Operations Research*, 27(4):810–822, 1979.
- [10] James S. Dyer and Rakesh K. Sarin. Relative risk aversion. *Management Science*, 28(8):875–886, 1982.
- [11] Larry G. Epstein and Stephen M. Tanny. Increasing Generalized Correlation: A Definition and Some Economic Consequences. *Canadian Journal of Economics*, 13(1):16–34, February 1980.
- [12] Peter C. Fishburn. *Utility Theory for Decision Making*. Wiley, New York, 1970.
- [13] William M. Gorman. The structure of utility functions. *The Review of Economic Studies*, 35(4):367 – 390, October 1968.
- [14] Charles M. Grinstead and Laurie J. Snell. *Introduction to Probability*. American Mathematical Society, 2006.
- [15] Edi Karni. On multivariate risk aversion. *Econometrica*, 47(6):1391–1401, 1979.
- [16] Edi Karni. On the correspondence between multivariate risk aversion and risk aversion with state-dependent preferences. *Journal of Economic Theory*, 30(2):230 – 242, 1983.
- [17] R.L. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*. Wiley series in probability and mathematical statistics. Applied probability and statistics. Cambridge University Press, 1993.
- [18] Richard E. Kihlstrom and Leonard J. Mirman. Risk aversion with many commodities. *Journal of Economic Theory*, 8(3):361 – 388, 1974.
- [19] Dilip B. Madan. Measures of risk aversion with many commodities. *Economics Letters*, 11(12):93 – 100, 1983.
- [20] Richard F. Meyer. Some Notes on Discrete Multivariate Utility. Technical report, Mimeographed manuscript, Havard Business School, 1972.
- [21] Jacob Paroush. Risk premium with many commodities. *Journal of Economic Theory*, 11(2):283 – 286, 1975.
- [22] John W. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32(1/2):122 – 136, 1964.
- [23] Scott F. Richard. Multivariate risk aversion, utility independence and separable utility functions. *Management Science*, 22(1):12–21, 1975.
- [24] Marco Scarsini. Dominance conditions for multivariate utility functions. *Management Science*, 34(4):454–460, 1988.

- [25] Joseph E Stiglitz. Behavior towards risk with many commodities. *Econometrica*, pages 660–667, 1969.
- [26] Menahem E Yaari. The dual theory of choice under risk. *Econometrica*, 55(1):95–115, 1987.

APPENDIX A. PROOFS

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

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

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

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

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

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

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

which represents \precsim^n , as required. □

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

Let (ϕ_2, ϕ_3, \dots) be the presumed future. By assumption (ϕ_2, ϕ_3, \dots) is internal.²⁸ So, there exists $s > 0$ with $\phi_i \pm s \in \mathcal{T}_i$, for all i .

²⁸More precisely, (ϕ_2, ϕ_3, \dots) is a presumed future that is internal, if there are several presumed futures.

Proofs for Section 4.2.1.

Lemma A.1. *Let X_1, X_2, \dots be an infinite sequence of independent uniformly bounded random variables,²⁹ with $E(X_i) = 0$ for all i . Set $S_n = \sum_{i=1}^n X_i$. Then*

$$(5) \quad \Pr[S_n \geq 0 \text{ infinitely often}] > 0.$$

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

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

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

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

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

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

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

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

So, again, in particular, $\Pr[S_n \geq 0 \text{ infinitely often}] > 0$. □

Theorem 4.2. \succsim is weakly risk averse if and only if u_n is concave with respect to v_n for all n .

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

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

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

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

²⁹that is, the support of all the random variables is included in a real interval $[\underline{b}, \bar{b}]$, with \underline{b}, \bar{b} finite.

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

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

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

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

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

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

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

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

So,

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

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

Proofs for Section 4.2.2. Theorem 4.3 follows directly from Theorem 4.4 (a) and Proposition 4.5. So, proceed to prove this theorem and proposition.

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

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

$$(7) \quad risk-prem_{\hat{u}_n}(x, L) \geq risk-prem_{cara_\alpha}(x, L)$$

for all n, x and L .

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

$$(8) \quad risk-prem_{\hat{u}_n}(x, \pm\epsilon) \geq \alpha\epsilon^2$$

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

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

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

So, for ϵ sufficiently small

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

So,

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

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

$$(10) \quad risk-prem_{cara_\alpha}(x, \pm\epsilon) < \frac{2}{3}\alpha\epsilon^2,$$

for all x .

So, combining (8) and (10)

$$(11) \quad risk-prem_{\hat{u}_n}(x, \pm\epsilon) > risk-prem_{cara_\alpha}(x, \pm\epsilon),$$

for ϵ sufficiently small. But then, by Pratt [22], (11) holds for any lottery L .

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

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

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

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

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

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

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

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

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

$$(\mathbf{a}_{-\beta}) \succsim^k \langle \mathbf{a}_{-\epsilon}, \mathbf{a}_{+\epsilon} \rangle.$$

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

$$(\mathbf{a}_{-\beta}, \phi_{k+1}, \dots, \phi_m) \succsim^m \langle (\mathbf{a}_{-\epsilon}, \phi_{k+1}, \dots, \phi_m), (\mathbf{a}_{+\epsilon}, \phi_{k+1}, \dots, \phi_m) \rangle.$$

Hence,

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

□

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

So, by the Kolmogorov inequality

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

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

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

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

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

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

Theorem 4.4.

- (a) If $RP_{\mathbf{u}}(\epsilon) = \Omega(\epsilon^2)$ as $\epsilon \rightarrow 0$ then \succsim is risk averse.³⁰
(b) If $RP_{\mathbf{u}}(\epsilon) = O(\epsilon^{2+\beta})$ as $\epsilon \rightarrow 0$, for some $\beta > 0$, then \succsim is not risk averse.

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

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

By the Taylor expansion,

$$(15) \quad \text{cara}_{\alpha}(\epsilon) = -e^{-\alpha\epsilon} = -1 + \alpha\epsilon - \frac{\alpha^2\epsilon^2}{2} + O(\alpha^3\epsilon^3).$$

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

$$(16) \quad \text{cara}_{\alpha_1}(\epsilon) \approx -1 + \alpha_1\epsilon - \frac{\alpha_1^2\epsilon^2}{2},$$

for $|\epsilon| \leq C$.

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

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

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

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

So,

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

³⁰recall that $g(y) = \Omega(h(y))$ as $y \rightarrow 0$ if there exists a constant M and y_0 such that $g(y) > M \cdot h(y)$ for all $y < y_0$.

So, it is sufficient to prove that

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

which is equivalent to saying that

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

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

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

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

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

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

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

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

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

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

Hence also (moving to $m+1$)³¹,

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

which means that

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

³¹We move to $m+1$ with ϕ_{m+1} to guarantee sufficient distance from the boundaries to allow a $\pm a_j$ lottery.

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

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

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

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

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

and

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

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

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

On the other hand,

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

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

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

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

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

Proposition 4.5. $RP_{\hat{\mathbf{u}}}(\epsilon) = \Omega(\epsilon^2)$ as $\epsilon \rightarrow 0$, if and only if $A_{\hat{\mathbf{u}}_n}(x)$ is bounded away from 0, uniformly for all n and x (assuming $\hat{\mathbf{u}}_n$ is twice differentiable for all n).

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

Theorem 4.6. For v_n, u_n , and $\hat{\mathbf{u}}$ as in Theorems 4.2 and 4.4

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

Proof. The proofs of (a) and (b) are analogous to those of Theorems 4.4 and 4.2. (c) follows from combining Theorems 4.4 and 4.6. \square

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

Throughout, the following notation is used:

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

- u denotes an NM utility function on \mathcal{S} . An NM utility for \mathcal{S} necessarily exists since the NM axioms are assumed to hold, and we consider only lotteries with finite support (see Fishburn [12, Theorem 8.2]).

Preliminaries

Lemma A.7. u is continuous.

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

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

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

Proofs for Section 6.

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

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

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

Proof. Gorman [13, Theorem 1] proves that if two independent factors \mathcal{E} and \mathcal{F} overlap then $\mathcal{E} \cup \mathcal{F}, \mathcal{E} \cap \mathcal{F}, \mathcal{E} - \mathcal{F}, \mathcal{F} - \mathcal{E}$, and $\mathcal{E} \triangle \mathcal{F} = (\mathcal{E} - \mathcal{F}) \cup (\mathcal{F} - \mathcal{E})$ are all independent.

Set $\mathcal{W} = \mathcal{A} \cap \mathcal{C}, \mathcal{X} = \mathcal{A} \cap \mathcal{D}, \mathcal{Y} = \mathcal{B} \cap \mathcal{C}$, and $\mathcal{Z} = \mathcal{B} \cap \mathcal{D}$. Then, by Gorman's theorem, $\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are independent, as is any product thereof. Since the partitions are not identical, at least three out of $\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are non-degenerate. So, $\mathcal{S} = \mathcal{W} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ is an independent partition with at least 3 factors. So, by Debreu [5], there are value functions $v^{\mathcal{W}}, v^{\mathcal{X}}, v^{\mathcal{Y}}$, and $v^{\mathcal{Z}}$, with $v^{\mathcal{W}} + v^{\mathcal{X}} + v^{\mathcal{Y}} + v^{\mathcal{Z}}$ representing \succsim . So, the pair of functions $v^{\mathcal{A}} = v^{\mathcal{W}} + v^{\mathcal{X}}$ and $v^{\mathcal{B}} = v^{\mathcal{Y}} + v^{\mathcal{Z}}$ are value functions for the independent partition $\mathcal{S} = \mathcal{A} \times \mathcal{B}$. Similarly, the functions $v^{\mathcal{C}} = v^{\mathcal{W}} + v^{\mathcal{Y}}$, and $v^{\mathcal{D}} = v^{\mathcal{X}} + v^{\mathcal{Z}}$ are value functions for the independent partition $\mathcal{S} = \mathcal{C} \times \mathcal{D}$, proving (1) and (2). Finally, (3) follows from (2) by the uniqueness of value functions. \square

Theorem 6.1. *For any \mathcal{S} , all (aggregate) Debreu value functions for \mathcal{S} are identical up to positive affine transformations.*

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

Theorem 6.2. *Let v be a Debreu value function for \mathcal{S} , and $(a_i, b_j), (b_i, a_j)$ perfectly hedged. Then for any $\mathbf{c} \in \mathcal{S}_{-\{i,j\}}$.*

$$v(a_i, b_j, \mathbf{c}) = v(b_i, a_j, \mathbf{c}) = \frac{v(a_i, a_j, \mathbf{c}) + v(b_i, b_j, \mathbf{c})}{2}.$$

Proof. Since there exists a value function for \mathcal{S} , by definition, there exists an independent partition of \mathcal{S} . Hence, by Lemma A.8 there exist value functions for any independent partition. In particular, there exist value functions $v^{\mathcal{T}_i}$ for all i . Set $v^{-\{i,j\}}(\mathbf{c}) = \sum_{t \neq i,j} v^{\mathcal{T}_t}(\mathbf{c}_t)$. Since $(a_i, b_j) \sim (b_j, a_i)$,

$$v^{\mathcal{T}_i}(a_i) + v^{\mathcal{T}_j}(b_j) + v^{-\{i,j\}}(\mathbf{c}) = v(a_i, b_j, \mathbf{c}) = v(b_i, a_j, \mathbf{c}) = v^{\mathcal{T}_i}(b_i) + v^{\mathcal{T}_j}(a_j) + v^{-\{i,j\}}(\mathbf{c})$$

So,

$$\begin{aligned} \frac{v(a_i, a_j, \mathbf{c}) + v(b_j, b_i, \mathbf{c})}{2} &= \frac{v^{\mathcal{T}_i}(a_i) + v^{\mathcal{T}_j}(a_j) + v^{\mathcal{T}_i}(b_i) + v^{\mathcal{T}_j}(b_j) + 2v^{-\{i,j\}}(\mathbf{c})}{2} = \\ &= v(a_i, b_j, \mathbf{c}) = v(b_i, a_j, \mathbf{c}). \end{aligned} \quad \square$$

The following Lemma is essentially the “if” direction of Theorem 6.3, but with a weaker requirement.

Lemma A.9. *Let u be an NM utility and v a Debreu value function for \mathcal{S} . Let $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ be an independent partition. Suppose that for any $a_1 \prec b_1, a_2 \prec b_2$, with $(a_1, b_2) \sim (b_1, a_2)$, and any $\mathbf{c} \in \prod_{t=3}^n \mathcal{T}_t$,*

$$(22) \quad \langle (a_1, a_2, \mathbf{c}), (b_1, b_2, \mathbf{c}) \rangle \prec \langle (a_1, b_2, \mathbf{c}), (b_1, a_2, \mathbf{c}) \rangle.$$

Then, u is strictly concave with respect to v . If (22) holds with a weak preference then u is concave with respect to v ; if it holds with indifference then u is linear with respect to v ; if it holds with the preference reversed then u is strictly convex with respect to v .

Proof. We prove for the case that (22) holds as is. The other cases are analogous.

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

Set $I = v(\mathcal{S})$, $I_1 = v^{\mathcal{T}_1}(\mathcal{T}_1)$, and $I_2 = v^{\mathcal{T}_2}(\mathcal{T}_2)$. Let ϵ be such that both I_1 and I_2 are of size at least 2ϵ . We prove that $u \circ v^{-1}$ is strictly concave on any interval of size 2ϵ , and hence strictly concave throughout.

Consider $x \in I$. Then $x = v(\hat{c}_1, \hat{c}_2, \hat{c})$ for some $\hat{c}_1, \hat{c}_2, \hat{c}$. Set $x_1 = v^{\mathcal{T}_1}(\hat{c}_1)$ and $x_2 = v^{\mathcal{T}_2}(\hat{c}_2)$. Assume that $(x_1 + \epsilon, x_2 + \epsilon) \in I_1 \times I_2$ (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_1, b_1, a_2, b_2 , with $v^{\mathcal{T}_1}(a_1) = x_1 + \delta_y/2, v^{\mathcal{T}_2}(a_2) = x_2 + \delta_y/2, v^{\mathcal{T}_1}(b_1) = x_1 + \delta_z/2, v^{\mathcal{T}_2}(b_2) = x_2 + \delta_z/2$. Then, $(a_1, b_2) \sim (b_1, a_2)$. So, by assumption

$$\langle (a_1, a_2, \hat{c}), (b_1, b_2, \hat{c}) \rangle \prec \langle (a_1, b_2, \hat{c}), (b_1, a_2, \hat{c}) \rangle.$$

So,

$$\frac{1}{2} (u(a_1, a_2, \hat{c}) + u(b_1, b_2, \hat{c})) < \frac{1}{2} (u(a_1, b_2, \hat{c}) + u(b_1, a_2, \hat{c})),$$

so,

$$\frac{1}{2} (u(v^{-1}(y)) + u(v^{-1}(z))) < \frac{1}{2} \left(u(v^{-1}(\frac{y+z}{2})) + u(v^{-1}(\frac{y+z}{2})) \right) = u(v^{-1}(\frac{y+z}{2})).$$

So, $u \circ v^{-1}$ is mid-point strictly concave and hence strictly concave. \square

Lemma A.10. *If u is strictly concave with respect to v then for any independent partition, $\mathcal{S} = \mathcal{T}_1 \times \dots \times \mathcal{T}_n$, any $i \neq j$, perfectly hedged $(a_i, b_j), (b_i, a_j)$, and $\mathbf{c} \in \mathcal{S}_{-\{i,j\}}$*

$$(23) \quad \langle (a_i, a_j, \mathbf{c}), (b_i, b_j, \mathbf{c}) \rangle \stackrel{\Delta}{\succ} \langle (a_i, b_j, \mathbf{c}), (b_i, a_j, \mathbf{c}) \rangle.$$

Proof. W.l.o.g. $i = 1, j = 2$. By definition $v(a_1, a_2, \mathbf{c}) \neq v(b_1, b_2, \mathbf{c})$, and by Theorem 6.2 $v(b_1, a_2, \mathbf{c}) = v(a_1, b_2, \mathbf{c}) = \frac{1}{2}(v(a_1, a_2, \mathbf{c}) + v(b_1, b_2, \mathbf{c}))$. Hence, since u is strictly concave with respect to v ,

$$\frac{1}{2} (u(a_1, a_2, \mathbf{c}) + u(b_1, b_2, \mathbf{c})) < \frac{1}{2} (u(a_1, b_2, \mathbf{c}) + u(b_1, a_2, \mathbf{c})).$$

Hence (since u is a representation of $\stackrel{\Delta}{\succ}$)

$$\langle (a_1, a_2, \mathbf{c}), (b_1, b_2, \mathbf{c}) \rangle \stackrel{\Delta}{\succ} \langle (a_1, b_2, \mathbf{c}), (b_1, a_2, \mathbf{c}) \rangle. \quad \square$$

Theorem 6.3. *For NM utility u and Debreu value function v ,*

- *Risk aversion:*
 - u is strictly concave with respect to v if and only if $\stackrel{\Delta}{\succ}$ is ordinally risk averse.
 - u is concave with respect to v if and only if $\stackrel{\Delta}{\succ}$ is weakly ordinally risk averse.
- *Risk loving:*
 - u is strictly convex with respect to v if and only if $\stackrel{\Delta}{\succ}$ is ordinally risk loving.
 - u is convex with respect to v if and only if $\stackrel{\Delta}{\succ}$ is weakly ordinally risk loving.
- *Risk neutrality:* u is linear with respect to v if and only if $\stackrel{\Delta}{\succ}$ 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:) Lemma A.10.

(if:) Lemma A.9. \square

Proofs for Sections 5.

Proposition 5.1. *If \succsim^z is ordinally risk averse (by Definition 3) then (2) holds for all $i \neq j$ (and any fully hedged (a_1, b_2) , (b_1, a_2) , and \mathbf{c}). Similarly for weak risk aversion.*

Proof. Consider an independent partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$. If $n = 2$ there is nothing to prove. If $n \geq 3$, there necessarily is a Debreu value function v representing \succsim . By Lemma A.9 u is concave with respect to v . So, the result is established by Lemma A.10. \square

Proposition 5.2. *If \succsim^z is (weakly) ordinally risk averse with respect to some independent partition $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$, then it is also so with respect to any independent partition.*

Proof. This follows from Lemmas A.8, A.9 and Theorem 6.3. Suppose there are two non-identical independent partitions $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ and $\mathcal{S} = \mathcal{R}_1 \times \cdots \times \mathcal{R}_m$, and that \succsim^z is ordinally risk averse with respect to the former. Then, by Lemma A.8 there exists a value function v representing \succsim . So, by Lemma A.9, u is strictly concave with respect to v . Hence, by the “only if” direction of Theorem 6.3, \succsim^z is ordinally risk averse with respect to any partition, in particular with respect to $\mathcal{R}_1 \times \cdots \times \mathcal{R}_m$. A similar argument holds for weak ordinal risk aversion. \square

Proof of Theorem 5.3.

As it turns out, the proof of Theorem 5.3 is the most involved in the section. The challenge arises in the case that the partition is with only two factors, in which case a Debreu value function need not exist, and the tools of Section 6 do not apply.

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

Let $w^{\mathcal{A}} : \mathcal{A} \rightarrow \mathbb{R}$ be a continuous real function representing $\succsim^{\mathcal{A}}$, and similarly $w^{\mathcal{B}}$ a continuous real function representing $\succsim^{\mathcal{B}}$ (such functions exist by Debreu [4] since $\succsim^{\mathcal{A}}$ and $\succsim^{\mathcal{B}}$ are continuous). Define $\mathbf{w} : \mathcal{A} \times \mathcal{B} \rightarrow \mathbb{R}^2$ as $\mathbf{w}(a, b) = (w^{\mathcal{A}}(a), w^{\mathcal{B}}(b))$. Let $I_{\mathcal{A}} \times I_{\mathcal{B}} \subseteq \mathbb{R}^2$ be the image of $\mathcal{A} \times \mathcal{B}$ under \mathbf{w} .

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

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

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

To prove continuity of \hat{u} , we prove that the pre-images of the open rays $(-\infty, r)$ and (r, ∞) are open, for all r . Consider $(-\infty, r)$ (the other case is analogous). Set $E_r = \{(x, y) : \hat{u}(x, y) < r\}$.

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

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

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

$$(25) \quad 0 < \hat{u}_{y'}^{\mathcal{A}}(x') - \hat{u}_{y'}^{\mathcal{A}}(x^*) < \frac{1}{2}\epsilon.$$

Combining (24) and (25), we obtain

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

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

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

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

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

Thus, by continuity,

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

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

Theorem 5.3. $\stackrel{\Delta}{\sim}$ is ordinally risk-averse if and only if it is correlation averse with respect to any and all $\mathcal{T}_i, \mathcal{T}_j$.

Proof. (if:) Suppose $\stackrel{\Delta}{\sim}$ is correlation averse with respect to $\mathcal{T}_1, \mathcal{T}_2$. Consider a perfectly hedged pair $(a_1, b_2), (b_1, a_2)$. By definition $(a_1, b_2) \not\sim (b_1, a_2)$. So, either $a_1 \not\sim b_1$ or $a_2 \not\sim b_2$. W.l.o.g. $a_1 \prec b_1$. But $(a_1, b_2) \sim (b_2, a_1)$. So, $a_2 \prec b_2$. Hence, by definition of correlation aversion

$$\langle (a_1, a_2, \mathbf{c}), (b_1, b_2, \mathbf{c}) \rangle \stackrel{\Delta}{\prec} \langle (a_1, b_2, \mathbf{c}), (b_1, a_2, \mathbf{c}) \rangle$$

for all \mathbf{c} . Hence, by Lemma A.9 $\stackrel{\Delta}{\sim}$ is ordinally risk averse.

(only if:) Suppose that $\stackrel{\Delta}{\sim}$ is ordinally risk averse. First, consider the case that the independent partition is with three or more factors. That is, suppose that $\mathcal{S} = \mathcal{T}_1 \times \dots \times \mathcal{T}_n$, $n \geq 3$. Then, there are Debreu value functions for the partition. Let $v^{\mathcal{T}_i}$ be the value function of \mathcal{T}_i . By Theorem

6.3 u is concave with respect to $v = \sum_{i=1}^n v^{\mathcal{T}_i}$. Consider $a_1 \prec b_1, a_2 \prec b_2$, and $\mathbf{c} \in \mathcal{S}_{-\{1,2\}}$. Set $x_i = v^{\mathcal{T}_i}(a_i), y_i = v^{\mathcal{T}_i}(b_i)$ and $z = \sum_{i=2}^n v^{\mathcal{T}_i}(c_i)$. W.l.o.g. $x_1 = x_2 = z = 0$. Set $\lambda = \frac{y_1}{y_1 + y_2}$. Then, since u is concave with respect to v

(26)

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

$$(27) \quad (u \circ v^{-1})(\lambda \cdot 0 + (1 - \lambda)(y_1 + y_2)) = (u \circ v^{-1})(y_2) = u(a_1, b_2, \mathbf{c}).$$

Similarly,

$$(28) \quad (1 - \lambda)u(a_1, a_2, \mathbf{c}) + \lambda \cdot u(b_1, b_2, \mathbf{c}) < u(b_1, a_2, \mathbf{c}).$$

Combining (26) and (28)

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

and $\stackrel{\circ}{\succsim}$ is correlation averse.

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

$$(29) \quad \langle (a, b), (A, B) \rangle \stackrel{\circ}{\succ} \langle (A, b), (a, B) \rangle.$$

If $(a, B) \sim (A, b)$ then they are perfectly hedged and (29) holds by the definition of ordinal risk aversion.

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

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

We show that $\text{diff} < 0$, which establishes (29).

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

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

such that

$$(30) \quad \begin{aligned} w^{\mathcal{A}}(A_{\frac{1}{2}}) - w^{\mathcal{A}}(a_{\frac{1}{2}}) &\leq \frac{1}{2}(w^{\mathcal{A}}(A) - w^{\mathcal{A}}(a)) \quad \text{or} \\ w^{\mathcal{B}}(B_{\frac{1}{2}}) - w^{\mathcal{B}}(b_{\frac{1}{2}}) &\leq \frac{1}{2}(w^{\mathcal{B}}(B) - w^{\mathcal{B}}(b)) \end{aligned}$$

and

$$(31) \quad \text{diff} < u(a_{\frac{1}{2}}, b_{\frac{1}{2}}) + u(A_{\frac{1}{2}}, B_{\frac{1}{2}}) - u(a_{\frac{1}{2}}, B_{\frac{1}{2}}) - u(A_{\frac{1}{2}}, b_{\frac{1}{2}}).$$

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

$$(32) \quad (a^1, b) \sim (a, B).$$

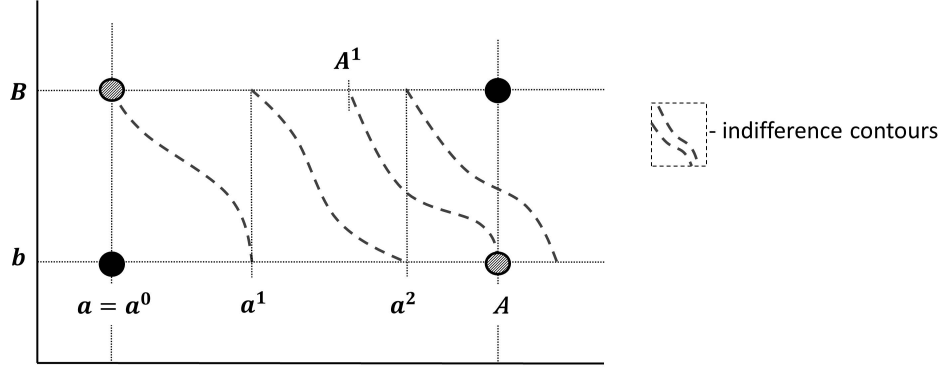


FIGURE 3. Illustration of the proof of Theorem 5.3. The values a^i are calculated left-to-right, starting at $a = a^0$. Here $\bar{i} = 2$ and the point a^2 is such that $w^A(a^2) \geq \frac{1}{2}(w^A(A) + w^A(a))$ (assuming the picture is scaled according to w^A).

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

$$(33) \quad w^A(A^1) \leq \frac{1}{2}(w^A(a) + w^A(A)),$$

or

$$(34) \quad w^A(a^{\bar{i}}) \geq \frac{1}{2}(w^A(a) + w^A(A)).$$

We consider each of these cases separately.

First, suppose that (33) holds. Then, by construction $(A^1, B) \sim (A, b)$, and they are perfectly hedged. Hence, by assumption,

$$\langle (A^1, b), (A, B) \rangle \stackrel{\Delta}{\succ} \langle (A^1, B), (A, b) \rangle.$$

So,

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

Hence,

$$(35) \quad \begin{aligned} & u(a, b) + u(A, B) - u(A, b) - u(a, B) = \\ & u(a, b) + u(A^1, B) - u(A^1, b) - u(a, B) + u(A^1, b) + u(A, B) - u(A^1, B) - u(A, b) < \\ & u(a, b) + u(A^1, B) - u(A^1, b) - u(a, B). \end{aligned}$$

Setting $a_{\frac{1}{2}} = a$, $A_{\frac{1}{2}} = A^1$, $b_{\frac{1}{2}} = b$ and $B_{\frac{1}{2}} = B$, by (33) and (35) we get (30) and (31).

Next, suppose that (34) holds. Then, by construction, for $i = 1, \dots, \bar{i}$, $(a^{i-1}, B) \sim (a^i, b)$, and each such pair is perfectly hedged. Since \succsim is ordinally risk averse,

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

for all i . So,

$$(36) \quad \frac{1}{2\bar{i}} \sum_{i=1}^{\bar{i}} (u(a^{i-1}, b) + u(a^i, B)) < \frac{1}{2\bar{i}} \sum_{i=1}^{\bar{i}} (u(a^{i-1}, B) + u(a^i, b));$$

and

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

so (as $a^0 = a$)

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

Hence,

$$(37) \quad \begin{aligned} & u(a, b) + u(A, B) - u(A, b) - u(a, B) = \\ & u(a, b) + u(a^{\bar{i}}, B) - u(a^{\bar{i}}, b) - u(a, B) + u(a^{\bar{i}}, b) + u(A, B) - u(a^{\bar{i}}, B) - u(A, b) < \\ & u(a^{\bar{i}}, b) + u(A, B) - u(a^{\bar{i}}, B) - u(A, b). \end{aligned}$$

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 (34) and (37) we get (30) and (31).

Thus, we have established (30) and (31), and we now return to complete the proof that $\text{diff} < 0$. Set

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

Then,

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

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

$$(38) \quad w^{\mathcal{A}}(A_\delta) - w^{\mathcal{A}}(a_\delta) \leq \delta \quad \text{or}$$

$$(39) \quad w^{\mathcal{B}}(B_\delta) - w^{\mathcal{B}}(b_\delta) \leq \delta$$

and

$$(40) \quad \begin{aligned} \text{diff}_{\frac{1}{2}} & < u(a_\delta, b_\delta) + u(A_\delta, B_\delta) - u(a_\delta, B_\delta) - u(A_\delta, b_\delta) = \\ & (u(A_\delta, B_\delta) - u(a_\delta, B_\delta)) + (u(a_\delta, b_\delta) - u(A_\delta, b_\delta)) = \end{aligned}$$

$$(41) \quad (u(A_\delta, B_\delta) - u(A_\delta, b_\delta)) + (u(a_\delta, b_\delta) - u(a_\delta, B_\delta)).$$

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

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

then

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

In particular, if (38) holds then (40) is $\leq 2\epsilon$, and if (39) holds then (41) is $\leq 2\epsilon$. Thus, $\text{diff}_{\frac{1}{2}} \leq 0$, so $\text{diff} < 0$. \square

Proofs for Section 8.

Proofs for Section 8.1.

Theorem 8.1. *In the multi-commodity setting (with $\mathcal{S} = \mathcal{T}_1 \times \dots \times \mathcal{T}_n$ an independent partition), the NM utility function u has constant coefficient of absolute risk aversion when measured with respect to the Debreu value function v if and only if for any i , lotteries L, L' over \mathcal{T}_i , and $\mathbf{x}, \mathbf{y} \in \Omega_{-\{i\}}$*

$$(L, \mathbf{x}) \succsim (L', \mathbf{x}) \iff (L, \mathbf{y}) \succsim (L', \mathbf{y}).$$

Proof. Let u be an NM utility representing \succsim . Meyer [20] (quoted in [23]) showed that all \mathcal{T}_i 's are utility independent if and only if there exist functions $u^i : \mathcal{T}_i \rightarrow \mathbb{R}$, $\beta > 0$ and α , such that one of the following holds:

$$(42) \quad u(a_1, \dots, a_n) = \sum_{i=1}^n u^i(a_i)$$

$$(43) \quad u(a_1, \dots, a_n) = \alpha + \beta \prod_{i=1}^n u^i(a_i), \text{ with } u^i(a_i) > 0$$

$$(44) \quad u(a_1, \dots, a_n) = \alpha - \beta \prod_{i=1}^n (-u^i(a_i)), \text{ with } u^i(a_i) < 0.$$

If (42) holds then the u^i 's are Debreu value functions (since \succsim agrees with \preceq). So u is linear with respect to v , and, in particular CARA.

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

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

is a Debreu value function representing \preceq . So,

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

is CARA w.r.t. v .

If (44) holds then setting $v^i = -\ln(-u^i)$ we have that

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

is a Debreu value function, and

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

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

Proposition 8.2. *Let \succsim be an (additively separable) preference order on $\mathcal{S} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$, and g a real valued function on \mathcal{S} . Suppose that for any \succsim the corresponding NM utility function u has constant coefficient of absolute risk aversion when measured with respect to g if and only if all factors \mathcal{T}_i are utility independent. Then g is a Debreu value function.*

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

$$(45) \quad v = \alpha + \beta e^{\gamma g},$$

for some α, β, γ .

Now consider another NM utility $u = e^v$. By Theorem 8.1 under this utility function each \mathcal{T}_i is utility independent. Hence, by assumption, this u must also be CARA in g . But, by (45), $u = e^v = e^{\alpha + \beta e^{\gamma g}}$, which is not CARA in g . \square

Proofs for Section 8.2.

Denote $\langle x_1, \dots, x_k : p_1, \dots, p_k \rangle$ the lottery obtaining value x_i with probability p_i , $i = 1, \dots, k$. Recall that $cara_\alpha$ is the function with $A_{cara_\alpha}(x) = \alpha$ for all x . For a function \hat{u} over a real interval, denote $\hat{ce}_{\hat{u}}(L)$ the certainty equivalent of the lottery L under the utility function \hat{u} . We will be using $\hat{ce}_{\hat{u}}(L)$ when considering the certainty equivalent in terms of the value function.

Lemma A.13. *For $\alpha \geq 0, x \geq 0, \beta \geq 1, p$ and $q = 1 - p$,*

$$\hat{ce}_{cara_\alpha} \langle 0, \beta x : p, q \rangle \leq \beta \cdot \hat{ce}_{cara_\alpha} \langle 0, x : p, q \rangle$$

Proof. For $\alpha = 0$, $cara_0$ is a linear function, so

$$\hat{ce}_{cara_\alpha} \langle 0, \beta x : p, q \rangle = q \cdot \beta x = \beta \cdot \hat{ce}_{cara_\alpha} \langle 0, x : p, q \rangle$$

For $\alpha > 0$, $cara_\alpha(x) = -e^{-\alpha x}$. Let γ be such that $\hat{ce}_{cara_\alpha} \langle 0, x : p, q \rangle = \gamma \cdot x$; that is

$$p + q \cdot (-e^{-\alpha \cdot x}) = -e^{-\alpha \cdot \gamma x}.$$

So,

$$p + q \cdot (-e^{-(\alpha/\beta) \cdot \beta x}) = -e^{-(\alpha/\beta) \cdot \gamma \beta x}.$$

So,

$$\hat{ce}_{cara_{\alpha/\beta}} \langle 0, \beta x : p, q \rangle = \gamma \beta x = \beta \cdot \hat{ce}_{cara_\alpha} \langle 0, x : p, q \rangle.$$

But $cara_\alpha$ is more risk averse than $cara_{\alpha/\beta}$ (since $\beta > 1$). So

$$\hat{ce}_{cara_\alpha} \langle 0, \beta x : p, q \rangle < \hat{ce}_{cara_{\alpha/\beta}} \langle 0, \beta x : p, q \rangle = \beta \cdot \hat{ce}_{cara_\alpha} \langle 0, x : p, q \rangle. \quad \square$$

Theorem 8.3. *If $A_{\hat{u}_2}(x) \geq A_{\hat{u}_1}(y) \geq 0$, for all x, y , then the following holds. Let $\mathcal{I} = \prod_{i \in I} \mathcal{T}_i$, for some $I \subseteq \{1, \dots, n\}$, be such that*

- \precsim_1 and \precsim_2 agree on the \mathcal{I} , and
- v_2^i is a concave expansion of v_1^i for all $i \in I$.

Then, \precsim_2 is more risk averse than \precsim_1 on \mathcal{I} .

Proof. Denote $v_j^{\mathcal{I}} = \sum_{i \in I} v_j^i$, and $v_j^{-\mathcal{I}} = \sum_{i \notin I} v_j^i$ ($j = 1, 2$). Then, $v_2^{\mathcal{I}}$ is a concave expansion of $v_1^{\mathcal{I}}$.

Let $L_x = \langle x_1, \dots, x_k : p_1, \dots, p_k \rangle$, $L_y = \langle x_1, \dots, x_{k-1}, y_k : p_1, \dots, p_{k-1}, p_k \rangle$ be two lotteries (over \mathcal{I}), with $x_1 \prec x_2 \prec \dots \prec x_k$ and $x_1 \prec y_k \precsim x_k$. We prove that for any $\mathbf{b} \in \mathcal{S}_{-I}$

$$(46) \quad ce_2(L_y, \mathbf{b}) \precsim ce_1(L_x, \mathbf{b}).$$

The result then follows when setting $L_y = L_x = L$.

Let α be such that

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

for all x, y . By assumption such an α exists and $\alpha \geq 0$.

W.l.o.g. assume that $v_1^{-\mathcal{I}}(\mathbf{b}) = v_2^{-\mathcal{I}}(\mathbf{b}) = 0$. Then, for any z

$$v_1(z, \mathbf{b}) = v_1^{\mathcal{I}}(z),$$

and

$$v_2(z, \mathbf{b}) = v_2^{\mathcal{I}}(z).$$

The proof of (46) is by induction on k (the number of points in the support of L). For $k = 1$ there is nothing to prove. For $k = 2$, assume, w.l.o.g. that $v_1^{\mathcal{I}}(x_1) = v_2^{\mathcal{I}}(x_1) = 0$, and consider the lotteries $L = \langle x_1, x : p, q \rangle$ and $L_y = \langle x_1, y : p, q \rangle$, $x_1 \prec x$ and $x_1 \precsim y \precsim x$.

Let $\beta = v_2^{\mathcal{I}}(x)/v_1^{\mathcal{I}}(x)$. Since $v_2^{\mathcal{I}}$ is an expansion, $\beta \geq 1$. Then,

$$(47) \quad \begin{aligned} v_2(ce_2(L_y, \mathbf{b})) &= \hat{ce}_{\hat{u}_2} \left\langle 0, v_2^{\mathcal{I}}(y) : p, q \right\rangle \leq \\ &\hat{ce}_{\hat{u}_2} \left\langle 0, v_2^{\mathcal{I}}(x) : p, q \right\rangle = \\ &\hat{ce}_{\hat{u}_2} \left\langle 0, \beta v_1^{\mathcal{I}}(x) : p, q \right\rangle \leq \\ &\hat{ce}_{cara_\alpha} \left\langle 0, \beta v_1^{\mathcal{I}}(x) : p, q \right\rangle \leq \\ &\beta \cdot \hat{ce}_{cara_\alpha} \left\langle 0, v_1^{\mathcal{I}}(x) : p, q \right\rangle \leq \\ &\beta \cdot \hat{ce}_{\hat{u}_1} \left\langle 0, v_1^{\mathcal{I}}(x) : p, q \right\rangle = \beta \cdot v_1(ce_1(L_x, \mathbf{b})). \end{aligned}$$

Since $v_2^{\mathcal{I}}$ a concave with respect to $v_1^{\mathcal{I}}$:

$$\frac{v_2(ce_1(L_x, \mathbf{b}))}{v_1(ce_1(L_x, \mathbf{b}))} \geq \frac{v_2^{\mathcal{I}}(x)}{v_1^{\mathcal{I}}(x)} = \beta.$$

So, from (47),

$$v_2(ce_2(L_y, \mathbf{b})) \leq \beta \cdot v_1(ce_1(L_x, \mathbf{b})) \leq v_2(ce_1(L_x, \mathbf{b})).$$

So, since v_2 and v_1 agree on \mathcal{I} ,

$$ce_2(L_y, \mathbf{b}) \preceq ce_1(L_x, \mathbf{b}).$$

Now, for $k > 2$, let $L_x^+ = \langle x_2, \dots, x_k : p'_2, \dots, p'_k \rangle$, $L_y^+ = \langle x_2, \dots, x_{k-1}, y_k : p'_1, \dots, p'_{k-1}, p'_k \rangle$, with $p'_i = p_i/(1 - p_1)$. Then, by the inductive hypothesis,

$$ce_2(L_y^+, \mathbf{b}) \preceq ce_1(L_x^+, \mathbf{b}).$$

Also,

$$(L_x, \mathbf{b}) \dot{\sim}_1 (\langle x_1, ce_1(L_x^+) : p_1, (1 - p_1) \rangle, \mathbf{b}).$$

and

$$(L_y, \mathbf{b}) \dot{\sim}_2 (\langle x_1, ce_2(L_y^+) : p_1, (1 - p_1) \rangle, \mathbf{b}).$$

So, again by the case $k = 2$,

$$ce_2(L_y, \mathbf{b}) \preceq ce_1(L_x, \mathbf{b}). \quad \square$$

APPENDIX B. UNBOUNDED LOTTERY SEQUENCES

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

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

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

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

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

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

So, by the Borel Cantelli lemma

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

So,

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

and hence

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

So, L_1, L_2, \dots is ultimately inferior to c_1, c_2, \dots