Dual theory of choice under multivariate risks

Alfred Galichon, Marc Henry

Available online at www.sciencedirect.com
ScienceDirect
Journal of Economic Theory 147 (2012) 1501–1516
JOURNAL OF Economic Theory
www.elsevier.com/locate/jet

Dual theory of choice with multivariate risks ☆

Alfred Galichon [a,*], Marc Henry [b,*]

[a] *Département d'économie, École polytechnique, 91128 Palaiseau, France*
[b] *Département de sciences économiques, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal QC H3C 3J7, Canada*

Received 24 February 2010; final version received 17 March 2011; accepted 3 May 2011

Available online 22 June 2011

Abstract

We propose a multivariate extension of Yaari's dual theory of choice under risk. We show that a decision maker with a preference relation on multidimensional prospects that preserves first order stochastic dominance and satisfies comonotonic independence behaves as if evaluating prospects using a weighted sum of quantiles. Both the notions of quantiles and of comonotonicity are extended to the multivariate framework using optimal transportation maps. Finally, risk averse decision makers are characterized within this framework and their local utility functions are derived. Applications to the measurement of multi-attribute inequality are also discussed.

JEL classification: D63; D81; C61

Keywords: Risk; Rank dependent utility theory; Multivariate comonotonicity; Optimal transportation; Multi-attribute inequality; Gini evaluation functions

☆ Both authors gratefully acknowledge support from the Chaire Axa "Assurance des Risques Majeurs" and the Chaire Société Générale "Risques Financiers". Galichon's research is partly supported by the Chaire EDF-Calyon "Finance and Développement Durable" and FiME, Laboratoire de Finance des Marchés de l'Energie (http://www.fime-lab.org). The authors thank Thibault Gajdos, John Weymark and the participants at the 2010 ParisTech–*Journal of Economic Theory* Symposium on Inequality and Risk for helpful discussions and comments. The authors are grateful to two anonymous referees and particularly an Associate Editor for their careful reading of the manuscript and insightful suggestions.
* Corresponding authors.
E-mail addresses: alfred.galichon@polytechnique.edu (A. Galichon), marc.henry@umontreal.ca (M. Henry).

0022-0531/$ – see front matter
doi:10.1016/j.jet.2011.06.002

0. Introduction

In his seminal paper [24], Menahem Yaari proposed a theory of choice under risk, which he called "dual theory of choice," where risky prospects are evaluated using a weighted sum of quantiles. The resulting utility is less vulnerable to paradoxes such as Allais' celebrated paradox [1]. The main ingredients in Yaari's representation are the preservation of first order stochastic dominance and insensitivity to hedging of comonotonic prospects. Both properties have strong normative and behavioral appeal once it is accepted that decision makers care only about the distribution of risky prospects. The preservation of stochastic dominance is justified by the fact that decision makers prefer risky prospects that yield higher values in all states of the world, whereas comonotonicity captures the decision maker's insensitivity to hedging comonotonic prospects, that is to say, the fact that the decision maker who is indifferent between two prospects that yield their higher and lower returns in the same states of the world, is also indifferent between any convex combination of those prospects. The dual theory has been used extensively as an alternative to expected utility in a large number of contexts. The main drawback of the dual theory is that it does not properly handle the case in which the prospects of consumptions of different natures are not perfect substitutes. The assumption of *law invariance* of the decision functional (called neutrality in [24] and by which the decision maker is insensitive to relabelings of the states of the world) is easier to substantiate when several dimensions of the risk are considered in the decision functional.

To handle these situations, we need to be able to express utility derived from monetary consumption in different *numéraires*, which is easily done with Expected Utility Theory, but so far not covered by Yaari's dual theory. Indeed, the latter applies only to risky prospects defined as univariate random variables, thereby ruling out choice among multidimensional prospects which are not perfect substitutes for each other, such as risks involving both a liquidity and a price risk, collection of payments in different currencies, payments at different dates, prospects involving different goods of different natures such as consumption and environmental quality, etc. Yaari [23] proposes a multivariate version of his dual theory, but it involves independence of the risk components and an axiom of separability (Axiom A in [23]), which essentially removes the multidimensional nature of the problem.

We propose to remove this constraint with a multivariate extension of the dual theory to risky prospects defined as random vectors that is applicable as such to the examples listed above. The main challenge in this generalization is the definition of quantile functions and comonotonicity in the multivariate setting. Another challenge is to preserve the simplicity of the functional representing preferences, so that they can be parameterized and can be computed as efficiently as in the univariate case. Both challenges are met with an appeal to optimal transportation maps that allow for the definition of "generalized quantiles," their efficient computation, and the extension of comonotonicity as a notion of distribution free perfect correlation. There are many ways of extending the notion of comonotonicity to a multivariate framework consistently with the univariate definition. Our proposed extension has the added property of preserving the equivalence between comonotonicity and Pareto efficiency of allocations (see [10] for the original result and [4] for the multivariate extension). With these notions of quantiles and comonotonicity in hand, we give a representation of a comonotonic independent preference relation as a weighted sum of generalized quantiles. The main difference between the univariate case and the multivariate case is that comonotonicity and generalized quantiles are defined with respect to an objective reference distribution, which features in the representation. The reference distribution is shown to be equal to

the distribution of equilibrium prices in an economy with at least one risk averse Yaari decision maker.

We then turn to the representation of a risk averse decision makers's preferences within this theory. Risk aversion is defined in the usual way as a preference for less risky prospects, where the notion of increasing risk is suitably generalized to multivariate risky prospects. We show, again in a direct generalization of the univariate case, that risk aversion is characterized by a special form of the quantile weights defined above: risk averse decision makers give more weight to low outcomes (low quantiles) and less weight to high outcomes (high quantiles). As a result, given the reference distribution with respect to which comonotonicity is defined, risk averse decision makers are characterized by further simple restrictions on their utility functionals, which makes this model as simple and as tractable as expected utility. A further advantage of our decision functional is the simple characterization of the local utility function and its close relation to the multivariate quantile function.

The risk averse Yaari decision functional is a version of the Weymark social evaluation function (in [22]) with a continuous state space. Indeed, the formal equivalence between the evaluation of risky prospects and the measurement of inequality noted in [2] and [9] allows us to draw implications of our theory for the measurement of inequality of allocations of multiple attributes, such as consumption, education, environment quality, etc. Seen as a social evaluation function, our decision functional provides a compromise between the approach of [7] and [19] in that it allows a flexible attitude to correlations between attributes without necessarily imposing correlation aversion and thereby circumventing the Bourguignon–Chakravarty [3] critique of the assumption that attributes are substitutes rather than complements.

The paper is organized as follows. The next section gives a short exposition of the dual theory. The following section develops the generalized notion of comonotonicity that is necessary for the multivariate extension, which is given in Section 3. Risk aversion is characterized in Section 4. The economic interpretation of the reference measure is given in Section 5 and the application to multi-attribute inequality measurement is discussed in Section 6. The final section concludes.

Notation and basic definitions

Let $(S, \mathcal{F}, \mathbb{P})$ be a non-atomic probability space. Let $X : S \to \mathbb{R}^d$ be a random vector. We denote the probability distribution of X by $\mathcal{L}_X$. $\mathbb{E}$ is the expectation operator with respect to $\mathbb{P}$. For x and y in $\mathbb{R}^d$, let $x \cdot y$ be the standard scalar product of x and y, and $\|x\|^2$ the Euclidean norm of x. We denote by $X =_d \mathcal{L}_X$ the fact that the distribution of X is $\mathcal{L}_X$ and by $X =_d Y$ the fact that X and Y have the same distribution. The *equidistribution class* of $X =_d \mathcal{L}_X$, denoted indifferently $\mathrm{equi}(\mathcal{L}_X)$ or $\mathrm{equi}(X)$, is the set of random vectors with distribution with respect to $\mathbb{P}$ equal to $\mathcal{L}_X$ (reference to $\mathbb{P}$ will be implicit unless stated otherwise). F_X denotes the cumulative distribution function of distribution $\mathcal{L}_X$. Q_X denotes its quantile function. In dimension 1, this is defined for all $t \in [0, 1]$ by $Q_X(t) = \inf_{x \in \mathbb{R}}\{\Pr(X \leqslant x) > t\}$. In larger dimensions, it is defined in Definition 3 of Section 2.1 below. We call L^2_d the set of random vectors X in dimension d such that $\mathbb{E}[\|X\|^2] < +\infty$. We denote by $\mathcal{D}$ the subset of L^2_d containing random vectors with a density relative to Lebesgue measure. A functional Φ on L^2_d is called upper semi-continuous (denoted u.s.c.) if for any real number α, $\{X \in L^2_d\colon\ \Phi(X) > \alpha\}$ is open. A functional Φ is lower semi-continuous (l.s.c.) if $-\Phi$ is upper semi-continuous. For a convex lower semi-continuous function $V : \mathbb{R}^d \mapsto \mathbb{R}$, we denote by ∇V its gradient (equal to the vector of partial derivatives). A doubly stochastic matrix is a square matrix of nonnegative real numbers, each of whose rows and columns sum to 1.

1. Dual theory of decision under risk

In this section, we first revisit Yaari's "Dual theory of choice under risk" presented in the eponymous paper [24]. As in [24], we consider a problem of choice among risky prospects as modeled by random variables defined on an underlying probability space. The risky prospect X is interpreted as a gamble or a lottery that a decision maker might consider holding and the realizations of X are interpreted as payments.

1.1. Representation

We suppose that the decision maker is characterized by a preference relation $\succsim$ on the set of risky prospects. $X \succsim Y$ indicates that the decision maker prefers prospect X to prospect Y, $X \succ Y$ stands for $X \succsim Y$ and not $Y \succsim X$, whereas $X \sim Y$ stands for $X \succsim Y$ and $Y \succsim X$. We first introduce the set of axioms satisfied by the preference relation that were proposed by Yaari in [24].

With the first axiom (which corresponds to Axioms A2 and A3 in Yaari [24]), we take the standard notion of preference as a continuous pre-order (reflexive and transitive binary relation) which is complete. Continuity of the preference relation is required relative to the topology of weak convergence: a sequence of random prospects X_n converges weakly to X if $\mathbb{E}f(X_n)$ converges to $\mathbb{E}f(X)$ for all continuous bounded functions f on $\mathbb{R}^d$. Then, $\succsim$ can be represented by a continuous real valued function γ in the sense that $X \succsim Y$ if and only if $\gamma(X) \geqslant \gamma(Y)$.

Axiom 1. The preference relation $\succsim$ is reflexive, transitive, complete and continuous relative to the topology of weak convergence.

A prospect X is said to first order stochastically dominate a prospect Y if there exist $\tilde{X} =_d X$ and $\tilde{Y} =_d Y$ such that $X(s) \geqslant Y(s)$ for almost all states of the world $s \in S$. The following axiom requires that whenever one prospect first order stochastically dominates a second, then the former is preferred to the latter. This is formally stated as follows.

Axiom 2. The preference $\succsim$ preserves first order stochastic dominance in the sense that if prospect X first order stochastically dominates prospect Y, then $X \succsim Y$, and if X strictly first order stochastically dominates prospect Y, then $X \succ Y$.

Two prospects with the same distribution first order stochastically dominate one another. Hence, Axiom 2 implies law invariance of the preference relation, or what [24] calls *neutrality*, i.e., $X =_d Y$ implies $X \sim Y$. Neutrality can be interpreted as the fact that the decision maker is indifferent to relabelings of the states of the world. Once neutrality is accepted, then Axiom 2 is reasonable as it is equivalent to requiring that the decision maker prefers prospects that yield a higher value in every state of the world. We shall see below that with suitable extensions of the concepts of monotonicity and stochastic dominance, this axiom remains reasonable in the multivariate extension of Yaari's representation theorem.

Finally, the third axiom is the crucial one in this framework, as it replaces independence by comonotonic independence. Recall that X and Y are comonotonic if $(X(s) - X(s'))(Y(s) - Y(s')) \geqslant 0$ for all $s, s' \in S$. The absence of a hedging opportunity between comonotonic prospects justifies the requirement below.

Axiom 3. If X, Y and Z are pairwise comonotonic prospects, then for any $\alpha \in [0,1]$, $X \succsim Y$ implies $\alpha X + (1-\alpha)Z \succsim \alpha Y + (1-\alpha)Z$.

We can now state Yaari's representation result.

Proposition 1 *(Yaari). A preference $\succsim$ on $[0,1]$-valued prospects satisfies Axioms* 1–3 *if and only if there exists a continuous non-decreasing function f defined on $[0,1]$, such that $X \succsim Y$ if and only if $\gamma(X) \geqslant \gamma(Y)$, where γ is defined for all X as $\gamma(X) = \int_0^1 f(1 - F_X(t))\,dt$.*

This result is interpretable in terms of weighting of outcomes (through the weighting of quantiles). Assume that each of the functions that we consider satisfy the invertibility and regularity conditions needed to perform the following operations. By integration by parts

$$\int_0^1 f\big(1 - F_X(t)\big)\,dt = \int_0^1 f(1-u)\,dQ_X(u) = \int_0^1 f(1-u)\frac{d}{du}Q_X(u)\,du$$
$$= \int_0^1 f'(1-u)Q_X(u)\,du.$$

Hence, calling $\phi(u) = f'(1-u)$, we have the representation of $\succsim$ with the functional $\int_0^1 \phi(t)Q_X(t)\,dt$. Hence, increasing f corresponds to positive ϕ, which can be interpreted as a weighting of the quantiles of the prospect X. As noted in [24], the functional γ satisfies $\gamma(\gamma(X)) = \gamma(X)$, so that $\gamma(X)$ is the certainty equivalent of X for the decision maker characterized by $\succsim$.

1.2. Risk aversion

We now turn to the characterization of risk averse decision makers among those satisfying Axioms 1–3. We define increasing risk as in Rothschild and Stiglitz [14]. The formulations in the first part of the definition below are equivalent by the Blackwell–Sherman–Stein Theorem (see, for instance, Chapter 7 of [16]).

Definition 1 *(Concave ordering, risk aversion).*

(a) A prospect Y is dominated by X in the *concave ordering*, denoted $Y \leqslant_{cv} X$, when the equivalent statements (i) or (ii) hold:
  (i) for all continuous concave functions, $\mathbb{E}f(Y) \leqslant \mathbb{E}f(X)$;
  (ii) Y has the same distribution as $\hat{Y}$ where $(X, \hat{Y})$ is a martingale, i.e., $\mathbb{E}(\hat{Y}|X) = X$ ($\hat{Y}$ is sometimes called a *mean-preserving spread* of X).

(b) The preference relation $\succsim$ is called *risk averse* if $X \succsim Y$ whenever $X \geqslant_{cv} Y$.

Notice that $x \to x$ and $x \to -x$ are both continuous and concave function. Therefore, condition (i) in the first part of the definition implies that $\mathbb{E}[X] = \mathbb{E}[Y]$ is necessary for a concave ordering relationship between X and Y to exist. With this definition, we can recall the characterization of risk averse preferences satisfying Axioms 1–3 as those with convex f (see Section 5 of [24] or Theorem 3.A.7 of [16]).

Proposition 2. *A preference relation $\succsim$ satisfying Axioms* 1–3 *is risk averse if and only if the function f in Theorem* 1 *is convex.*

This monotonicity of the derivative of f has the natural interpretation that risk averse decision makers evaluate prospects by giving high weights to low quantiles (corresponding to low values of the prospect) and low weights to high quantiles. Indeed, with the formulation $\gamma(X) = \int_0^1 \phi(u) Q_X(u)\, du$ and the identification $\phi(u) = f'(1-u)$, an increasing convex f corresponds to positive decreasing ϕ, and therefore to a weighting scheme in which low quantiles (corresponding to unfavorable outcomes) receive high weights and high quantiles (corresponding to favorable outcomes) receive low weights.

2. Multivariate quantiles and comonotonicity

The main ingredients in our multivariate representation theorem are the multivariate extensions of quantiles and comonotonicity. As we shall see, the two are intimately related.

2.1. Multivariate quantiles

We first note that the quantile of a random variable can be characterized as an increasing rearrangement of the latter. Hence, by classical rearrangement inequalities, quantiles are solutions to maximum correlation problems. More precisely, by the rearrangement inequality of Hardy, Littlewood and Pólya [8], we have the following well-known equality:

$$\int_0^1 t\, Q_X(t)\, d(t) = \max\bigl\{\mathbb{E}[XU]\colon\ U \text{ uniformly distributed on } [0,1]\bigr\}, \tag{2.1}$$

where the quantile function Q_X has been defined above. This variational characterization is crucial when generalizing Yaari's representation theorem to the multivariate setting. Indeed, consider now a random vector X on $\mathbb{R}^d$ and a reference distribution μ on $\mathbb{R}^d$, with U distributed according to μ. We introduce *maximum correlation functionals* to generalize the variational formulation of (2.1).

Definition 2 *(Maximal correlation functionals).* A functional $\varrho_\mu : L_d^2 \to \mathbb{R}$ is called a maximal correlation functional with respect to a reference distribution μ if for all $X \in L_d^2$,

$$\varrho_\mu(X) := \sup\bigl\{\mathbb{E}[X \cdot \tilde{U}]\colon\ \tilde{U} =_d \mu\bigr\}.$$

It follows from the theory of optimal transportation (see Theorem 2.12(ii), p. 66 of [20]) that if μ is absolutely continuous with respect to Lebesgue measure (which will be assumed throughout the rest of the paper), then there exists a convex lower semi-continuous function $V : \mathbb{R}^d \to \mathbb{R}$ and a random vector U distributed according to μ such that $X = \nabla V(U)$ holds μ-almost surely, and such that $\varrho_\mu(X) = \mathbb{E}[\nabla V(U) \cdot U]$. In that case, the pair (U, X) is said to achieve the *optimal quadratic coupling of* μ with respect to the distribution of X. The function V is called the *transportation potential* of X with respect to μ or the *transportation potential* from μ to the probability distribution of X.[1] This shows that the gradient of the convex function V thus

[1] V is convex and hence differentiable except on set of measure zero by Rademacher's Theorem (Theorem 2.4 in [13]), so that the expression $\mathbb{E}[\nabla V(U) \cdot U]$ above is well defined.

obtained satisfies the multivariate analogue of Eq. (2.1). We therefore adopt ∇V as our notion of a generalized quantile.

Definition 3 *(μ-quantile).* The μ-quantile function of a random vector X on $\mathbb{R}^d$ with respect to an absolutely continuous distribution μ on $\mathbb{R}^d$ is defined by $Q_X = \nabla V$, where V is the transportation potential of X with respect to μ.

This concept of a multivariate quantile is the counterpart of our definition of multivariate comonotonicity in the representation theorem, and the latter, introduced in the following section has strong economic underpinnings, as discussed in Section 5, where we give the economic interpretation of the reference measure μ.

2.2. Multivariate comonotonicity

Two univariate prospects X and Y are comonotonic if there is a prospect U and non-decreasing maps T_X and T_Y such that $Y = T_Y(U)$ and $X = T_X(U)$ almost surely or, equivalently, $\mathbb{E}[UX] = \max\{\mathbb{E}[\tilde{U}X]\colon\ \tilde{U} =_d U\}$ and $\mathbb{E}[UY] = \max\{\mathbb{E}[\tilde{U}Y]\colon\ \tilde{U} =_d U\}$. Comonotonicity is hence characterized by maximal correlation between the prospects over the equidistribution class. This variational characterization (where products will be replaced by scalar products) will be the basis for our generalized notion of comonotonicity.

Definition 4 *(μ-comonotonicity).* Let μ be a probability measure on $\mathbb{R}^d$ with finite second moments. A collection of random vectors $X_i \in L^2_d$, $i \in I$, are called μ-comonotonic if one has

$$\varrho_\mu\left(\sum_{i\in I} X_i\right) = \sum_{i\in I} \varrho_\mu(X_i).$$

When μ is absolutely continuous with respect to Lebesgue measure, it follows from the representation of ϱ_μ that the family X_i is μ-comonotonic if and only if there exists a vector U distributed according to μ such that $U \in \operatorname{argmax}_{\tilde{U}}\{\mathbb{E}[X_i \cdot \tilde{U}],\ \tilde{U} =_d \mu\}$ for all $i \in I$. In other words, the X_i's can be rearranged simultaneously so that they achieve maximal correlation with U. When the distributions of the random vectors are absolutely continuous with respect to Lebesgue measure, the concept of μ-comonotonicity is transitive.

Proposition 3. *Suppose that X and Y are μ-comonotonic and that Y and Z are μ-comonotonic, with the distribution of Y assumed to be absolutely continuous with respect to Lebesgue measure. Then X and Z are μ-comonotonic.*

Comonotonic allocations and Pareto efficiency It is worth discussing this definition of comonotonicity as generalizations of the classical univariate notion of comonotonicity are not unique. The main motivation for introducing it is to generalize the univariate equivalence between comonotonic and Pareto efficient allocations in a risk-sharing economy. Consider an Arrow–Debreu economy with n agents, and with an aggregate endowment which is a random vector X. Thus the i-th dimension of the realization $X^i(\omega)$ in state $\omega \in \Omega$ of this random vector is the quantity of good $i \in \{1, \ldots, d\}$ produced in this state. An *allocation* (or *risk-sharing allocation*) of X is a sharing rule of this aggregate endowment among the n agents, hence it is the specification of n random vectors $X_1, \ldots, X_n$ such that

$$\forall \omega \in \Omega, \quad \sum_{k=1}^{n} X_k(\omega) = X(\omega)$$

where $X_k^i(\omega)$ is the quantity of good i allocated to agent $k \in \{1, \ldots, n\}$ in state ω. An allocation is called (Pareto) *efficient* if no other allocation dominates the former, agent by agent, in the sense of the concave ordering (as defined in Proposition 4 below).

In dimension one, it is known since the seminal paper of Landsberger and Meilijson [10] that a risk-sharing allocation is Pareto efficient with respect to the concave order if and only if it is comonotonic. That is, given any comonotonic allocation, it is not possible to find another allocation such that each risky endowment would be preferred under the new allocation by every risk averse decision maker to the endowments in the original allocation. Multivariate generalization of this equivalence is not obvious, but it turns out that, as recently shown by Carlier, Dana and Galichon [4], this result can be extended to the multivariate case, with comonotonicity replaced by *multivariate comonotonicity*, if one defines an allocation to be comonotonic in the multivariate sense if and only if it is μ-comonotonic for some measure μ with enough regularity. In our view, this result strongly supports the claim that our notion of comonotonicity is in some sense the "natural" multivariate extension of comonotonicity.

Relation with other multivariate notions of comonotonicity Puccetti and Scarsini [12] have also applied the theory of optimal transportation to generalize the notion of comonotonicity to the multivariate setting. They review possible multivariate extensions of comonotonicity, including the notion of μ-comonotonicity that we propose. But the concept they favor differs from ours in the sense that according to their notion of multivariate comonotonicity (which they call *c-comonotonicity*), two vectors X and Y are c-comonotonic if and only if (X, Y) is an optimal quadratic coupling. That is, X and Y are c-comonotonic if and only if there is a convex function V such that $Y = \nabla V(X)$ holds almost surely. However, unlike μ-comonotonicity, c-comonotonicity is in general not transitive, and does not seem to be related to efficient risk-sharing allocations and equilibrium.

Schmeidler [15] introduces an internal notion of comonotonicity: if a decision maker evaluates prospects according to $\succsim$, then Schmeidler-comonotonicity of two prospects X and Y means that for all pairs of states of the world (s, t), $X(s) \succsim X(t)$ implies $Y(s) \succsim Y(t)$, i.e., prospects X and Y are more desirable in the same states of the world. In contrast, we extend the Weymark [22] – Yaari [24] motivation in our definition of comonotonicity, which can be related to the state prices in the economy, as explained in Section 5. The two notions have no obvious relation, as we see by considering two μ-comonotonic prospects X and Y and imposing Schmeidler-comonotonicity. By μ-comonotonicity, there exists $U =_d \mu$ and generalized quantile functions Q_X and Q_Y such that $X = Q_X(U)$ and $Y = Q_Y(U)$. Schmeidler-comonotonicity of X and Y would require that the univariate random variables $(\mathbb{E}U) \cdot X$ and $(\mathbb{E}U) \cdot Y$ are comonotonic in the usual sense. Although they are equivalent in dimension one, in higher dimensions, neither of these two concepts implies the other.

3. Multivariate representation theorem

Now that we have given a formalization of the notion of maximal correlation in a law invariant sense that is suitable for a multivariate extension of Yaari's dual theory, we can proceed to generalize Yaari's representation result to the case of a preference relation among multivariate prospects. We consider prospects, which are elements of L_d^2. Axiom $1'$ below is a mild

smoothness requirement for the preference relation. A functional γ is called Fréchet differentiable in X relative to the L^2_d metric if there is a linear functional L such that $|\gamma(X+h) - \gamma(X) - L(h)|/\sqrt{\mathbb{E}[h^2]} \to 0$. As in [5], the functional will not be Fréchet differentiable at all points; we only require differentiability at one point.

Axiom 1′. The preference $\succsim$ is represented by a continuous functional γ on L^2_d such that at least one point its Fréchet derivative exists and is non-zero.

Given sufficient regularity, first order stochastic dominance can be characterized equivalently by pointwise dominance of cumulative distribution functions or pointwise dominance of quantile functions. It is the latter that we adopt for our multivariate definition.

Definition 5 *(μ-first order stochastic dominance).* A prospect X μ-first order stochastically dominates prospect Y relative to the componentwise partial order $\geqslant$ on $\mathbb{R}^d$ if $Q_X(t) \geqslant Q_Y(t)$ for almost all $t \in \mathbb{R}^d$, where Q_X and Q_Y are the generalized quantiles of X and Y with respect to a distribution μ on $\mathbb{R}^d$.

For any $U =_d \mu$, we have $Q_X(U) =_d X$ and $Q_Y(U) =_d Y$. If X μ-first order stochastically dominates Y, then $Q_X(U) \geqslant Q_Y(U)$ almost surely. Hence, $\hat{X} \geqslant \hat{Y}$ almost surely for some $\hat{X} =_d X$ and $\hat{Y} =_d Y$, which is the "usual multivariate stochastic order" (see [16, p. 266]). The converse does not hold in general.

The remaining two axioms require fixing an absolutely continuous reference probability distribution μ on $\mathbb{R}^d$.

Axiom 2′. The preference $\succsim$ preserves μ-first order stochastic dominance in the sense that if prospect X μ-first order stochastically dominates prospect Y, then $X \succsim Y$, and if X μ-first order strictly stochastically dominates prospect Y, then $X \succ Y$.

The extension of the comonotonicity axiom is the key to the generalization of the dual theory to multivariate prospects. The statement of Axiom 3 is unchanged, but the concept of comonotonicity is now dependent on a reference distribution μ. The prospects X, Y and Z are comonotonic, or more precisely μ-comonotonic, if they are all *maximally correlated* in the law invariant sense of Definition 4 with a reference U (where U has distribution μ).

Axiom 3′. If X, Y and Z are μ-comonotonic prospects, then for any $\alpha \in [0,1]$, $X \succsim Y$ implies $\alpha X + (1-\alpha)Z \succsim \alpha Y + (1-\alpha)Z$.

We are now in a position to state the multivariate extension of Yaari's representation theorem.

Theorem 1 *(Multivariate representation). A preference relation on multivariate prospects in L^2_d satisfies Axioms $1'$, $2'$ and $3'$ relative to a reference probability measure μ if and only if there exists a function ϕ such that for $U =_d \mu$, $\phi(U) \in L^2_d$, $\phi(U) \in (\mathbb{R}_-)^d$ almost surely and such that for all pairs X, Y, $X \succsim Y$ if and only if $\gamma(X) \geqslant \gamma(Y)$, where γ is defined for all X by $\gamma(X) = \mathbb{E}[Q_X(U) \cdot \phi(U)]$, where Q_X is the μ-quantile of X.*

When $d = 1$, the representation is independent of μ and we recover the result of Proposition 1. As in the univariate case, the decision maker assesses prospects with a weighting scheme ϕ of

quantiles of the prospects. Because γ in Theorem 1 satisfies $\gamma(\gamma(X)) = \gamma(X)$, $\gamma(X)$ is the certainty equivalent of X as in the univariate case. Furthermore, $\succsim$ satisfies *linearity in payments*, i.e., for any positive real number a and any $b \in \mathbb{R}^d$ (identified with a constant multivariate prospect), $\gamma(aX + b) = a\gamma(X) + b$.

It should be noted that Choquet expected utility [15] handles multivariate prospects under Schmeidler-comonotonicity (defined in Section 2.2). As shown in [21], under Axiom 2, Choquet expected utility is identical to the functional of Proposition 1. Hence, when restricted to decision under risk, Choquet expected utility aggregates the multiple dimensions of the prospects with the utility function and then considers univariate quantiles of the resulting utility index. This is in contrast with the functional of Theorem 1, which directly evaluates multivariate quantiles of the prospects and thereby models attitudes to substitution risk between the dimensions of the prospect.

4. Risk aversion and the local utility function

In this section, we consider the question of representing those decision makers satisfying Axioms $1'$, $2'$ and $3'$ that are risk averse in the sense of Definition 1. We then show that the *local utility function* in the sense of [11] is easily computable and provides an interpretation of the reference distribution μ.

4.1. Risk aversion

For our characterization of risk averse Yaari decision makers, we need to generalize the concept of a mean-preserving spread to the multivariate setting.

Proposition 4 *(Concave ordering). For any prospects X and Y whose respective distributions are absolutely continuous with respect to Lebesgue measure, the following properties are equivalent*:

(a) *For every bounded concave function f on $\mathbb{R}^d$, $\mathbb{E}f(X) \geqslant \mathbb{E}f(Y)$.*
(b) *$Y =_d \hat{Y}$, with $\mathbb{E}[\hat{Y}|X] = X$.*
(c) *$\varrho_\mu(X) \leqslant \varrho_\mu(Y)$ for every probability measure μ.*
(d) *X belongs to the closure of the convex hull of the equidistribution class of Y.*
(e) *$\Phi(X) \geqslant \Phi(Y)$ for every u.s.c. law invariant concave functional $\Phi : L^2_d \to \mathbb{R}$.*

When any of the properties above hold, one says that Y is dominated by X in the concave ordering*, denoted $Y \leqslant_{cv} X$.*

Statements (a) and (b) are identical in the multivariate case as in [14]. The equivalence between the two is a classical result that can be traced back at least to [17] (see Appendix A.2 for details). The interpretation of the ordering as a preference ordering for all risk aversion expected utility maximizers (a) and as an ordering of mean-preserving spreads (b) also carry over to the multivariate case. Statement (d) is the continuous equivalent to multiplication by a doubly stochastic matrix.

As in the univariate case and Definition 1(b), risk averse decision makers will be defined by aversion to mean-preserving spreads. It turns out that imposing risk aversion on a preference relation that satisfies Axioms $1'$, $2'$ and $3'$ is equivalent to requiring the following property, sometimes called *preference for diversification*.

Axiom 4. For any two preference equivalent prospects X and Y (i.e., such that $X \sim Y$), convex combinations are preferred to either of the prospects (i.e., for any $\alpha \in [0,1]$, $\alpha X + (1-\alpha)Y \succsim X$).

This is formalized in the following theorem, which gives a representation for risk averse *Yaari decision makers*.

Theorem 2. *In dimension $d \geqslant 2$, for a preference relation satisfying Axioms $1'$, $2'$ and $3'$, the following statements are equivalent*:

(a) $\succsim$ *is risk averse, namely $X \succsim Y$ whenever $X \geqslant_{cv} Y$.*
(b) $\succsim$ *satisfies Axiom* 4.
(c) *The function ϕ involved in the representation of the preference relation in Theorem* 1 *satisfies $-\phi(u) = \alpha u + u_0$ for $\alpha > 0$ and $u_0 \in \mathbb{R}^d$.*

So, in the multivariate setting the functional γ is convex if and only if $\phi(x) = -\alpha x - x_0$ for some α real positive and $x_0 \in \mathbb{R}^d$. This is a major difference with dimension one, where the functional is convex if and only if $-\phi$ is a non-decreasing map. This implies that a multivariate Yaari risk averse decision maker is entirely characterized by the reference distribution μ.

4.2. Local utility function

Throughout the rest of the paper, we shall assume that the conditions in Theorem 2 are met. Hence, our discussion of local utility functions will be limited to the case of risk averse decision makers. By law invariance, we denote $\gamma(P) := \gamma(X)$, where $X =_d P$. Without loss of generality, we shall also assume that $\phi(u) = -u$, thus $\gamma(P) := -\mathbb{E}[\nabla V_P(U) \cdot U]$, where $V_P = V_X$ is the transportation potential (see Section 2.1 for the definition) from the reference probability distribution μ of U to the probability distribution P of X. As we have seen, the gradient ∇V_P of this transportation potential is the μ-quantile function of distribution P.

As shown in [11], when smoothness requirements are met, a local analysis can be carried out in which a (risk averse) non-expected utility function behaves for small perturbations around a fixed risk in that same way as a (concave) utility function. Formally, the local utility function is defined as $u(x|P) = D_P\gamma(x)$, where $D_P\gamma$ is the Fréchet derivative of γ at P (see Section 3 for the definition). Denoting by $V^*(x) = \sup_u [u \cdot x - V(u)]$ the *Legendre–Fenchel transform* of a convex lower semi-continuous function V, we have

$$\begin{aligned}\gamma(P) &= \mathbb{E}\big[\nabla V_P(U) \cdot U\big] = \max\big\{\mathbb{E}[X \cdot U]\colon\ X =_d P,\ U =_d \mu\big\} \\ &= \min\left\{\int V(u)\,d\mu(u) + \int V^*(x)\,dP(x)\colon\ V \text{ convex and l.s.c.}\right\} \\ &= \int V_P(u)\,d\mu(u) + \int V_P^*(x)\,dP(x)\end{aligned}$$

by the duality of optimal transportation (see, for instance, Theorem 2.9, p. 60 of [20]).

Defining $f(V, Q) = \int V\,d\mu + \int V^*\,dQ$, we have $\gamma(P) = -\inf_V f(V, P)$. Hence, an envelope theorem argument formally yields $u(x|P) = D_P\gamma(x) = -V_P^*(x)$. Therefore, *the local utility function is $-V_P^*$, the* (*negative of the*) *Legendre–Fenchel transform of the transportation potential V_P*. This point sheds light on the economic interpretation of this potential, thanks to Machina's theory of local utility. The function $-V_P^*$ is concave, which is consistent with the risk

aversion of a Yaari decision maker given the assumptions of Theorem 2. For univariate prospects, $u(x|P) = -V_P^*(x) = \int_{-\infty}^{x} F_X(z)\,dz$, so that we recover the fact that when $X =_d P$ is a mean-preserving spread of $Y =_d Q$, $u(z|P) \leqslant u(z|Q)$ for all z.

5. Economic interpretation of the reference measure

We now discuss the behavioral interpretation of μ. As we saw in Theorem 1, the generalization of the Yaari preferences to the multivariate case led us to define a utility functional γ over prospects such that $\gamma(X) = \mathbb{E}[X \cdot \phi(\tilde{U})]$ for some prospect $\tilde{U}$ which is correlated to X. $\tilde{U}$ is an index such that $X \cdot \tilde{U}$ measures how favorable the outcome is for the decision maker. $\phi(\tilde{U})$ is a weighting of the contingent outcome X, so that γ over- or under-weights prospects in each state using weights $\phi(\tilde{U})$. Hence, the dispersion of μ induces a departure from risk-neutrality. In the special case in which μ is the distribution of a constant u_0, $\gamma(X) = \mathbb{E}[X \cdot \phi(u_0)] = \mathbb{E}[X] \cdot \phi(u_0)$ and one recovers the case of a risk-neutral decision maker. On the contrary, when μ exhibits considerable dispersion, then the variance of $\phi(\tilde{U})$ is large in general, so that the "favorable" outcomes (in the sense that $X \cdot \tilde{U}$ is high) are weighted less, at least if $\phi(\tilde{U}) = -\alpha\tilde{U}$. This induces risk aversion. When $\phi(\tilde{U})$ differs from a rescaling of $\tilde{U}$, there may be some discrepancy between the weighting of a given state and how favorable it is. Hence, the variance of μ is no longer directly associated with risk aversion.

We now turn to the equilibrium implications of the reference measure μ and show how it is related to the distribution of the state prices in an economy in equilibrium when a decision maker with risk averse decision functional as in Theorem 2 is present in the economy. Consider an economy where one of the agents (whom we shall refer to as "Yaari") has preferences as in Theorem 2, with reference measure μ. Assume that there is a risk-sharing equilibrium in this economy, which is supported by the stochastic discount factor ξ, meaning that if the original risky endowment of the agent is X_0, then the agent's budget set is $\{X\colon \mathbb{E}[(X - X_0) \cdot \xi] = 0\}$. The demand for risk X of Yaari is therefore $\max_{\hat{X}} \gamma(\hat{X})$ subject to $\mathbb{E}[(\hat{X} - X_0) \cdot \xi] = 0$. Since Yaari is assumed risk averse, γ is concave, and the demand for risk X satisfies the local optimality condition $\max_{\hat{X}} \mathbb{E}[u(\hat{X}|P)]$ subject to $\mathbb{E}[(\hat{X} - X_0) \cdot \xi] = 0$. The first order conditions yield $\nabla u(X|P) = \lambda\xi$, where λ is the Lagrange multiplier associated with the budget constraint, where $\lambda \neq 0$ unless there is no trade in equilibrium. Now, as explained above, $u(x|P) = -V_P^*(x)$, hence $\nabla u(X|P) = -\nabla V_P^*(X)$. Now, by definition of the transportation potential V_p from μ to P, $\nabla V_P^*(X) =_d \mu$. Hence $-\lambda\xi =_d \mu$ which implies that μ is (up to scale) the distribution of the stochastic discount factor ξ. Therefore, when there is a Yaari decision maker with reference measure μ in the economy, the stochastic discount factor should be distributed according to μ. This result is an extension of the well-known result that states that when there is a risk-neutral decision maker in the economy, the stochastic discount factor should equal one, that is, the risk-neutral probability should coincide with the actuarial probability. To summarize, if a risk-sharing equilibrium exists with a Yaari risk averse decision maker with reference measure μ, then μ coincides with the distribution of the stochastic discount factor. Thereby, μ is related to the distribution of the state prices.

6. Relation with multi-attribute inequality measurement

The theory developed here has implications for inequality rankings of allocations of multiple attributes (such as income, education, environmental quality, etc.) in a population. Atkinson [2] recognized the relevance of stochastic orderings to the measurement of inequality and its foundation on principles such as the desirability of Pigou–Dalton transfers (also known as Pigou–Dalton

Majorization). Weymark [22] added to Pigou–Dalton Majorization a principle of comonotonic independence, which he interpreted as neutrality to the source of variation in income, and obtained a class of social evaluation functions, which he called generalized Gini evaluation functions. The functional form is identical to the decision functional derived independently on a continuous state space by [24]. Indeed, [9] notes the formal equivalence between the problem of decision under risk and the measurement of inequality. The random vector of risks or prospects that we consider in the present work can be interpreted as an allocation of multiple attributes over a continuum of individuals. With this interpretation, states of the world are identified with individuals in the population and the decision function γ is interpreted as a social evaluation function. Law invariance (Yaari neutrality, i.e., insensitivity to relabelings of the states of the world) of the decision functional is thus equivalent to anonymity of the social evaluation function. The ranking of ordinally equivalent allocations obtained through Pigou–Dalton Majorization (see [9]) corresponds to the concave ordering discussed in Proposition 4. More precisely, the mean-preserving spread characterization (b) in Proposition 4 is equivalent to (d), which is the infinite-dimensional analogue of multiplication by a doubly stochastic matrix. Hence, our risk averse multivariate Yaari decision functional can be interpreted as a social evaluation function for allocations of multiple attributes, which satisfies anonymity, monotonicity and Pigou–Dalton Majorization in the sense of Theorem 3 in [9].

The inequality literature achieves functional forms for social evaluation functions in the multi-attribute case by adding two distinct types of majorization principles that allow the comparison of non-ordinally equivalent social evaluations. Tsui [18,19] considers correlation increasing transfers. Gajdos and Weymark [7] extend generalized Gini social evaluation functions to the multivariate case with a comonotonic independence axiom. Two allocations are said to be comonotonic if all individuals are ranked identically in all attributes (i.e., the richest is also the most educated, etc.), and the ranking between two comonotonic allocations is not reversed by the addition of a comonotonic allocation. They use an attribute separability axiom (Axiom A in [23]) to reduce the dimensionality of the problem via independence of the attributes. Specifically, Theorem 4 of [7] is a special case of our Theorem 2 when the attribute vector X and the reference distribution μ both have independent marginals. Our representation can also incorporate trade-offs between attributes and attitudes to correlations between attributes of the kind that are entertained in [19], but is not restricted to the latter. Correlation aversion would correspond to perceived substitutability, but perceived complementarity can also be entertained in our approach, thereby circumventing Bourguignon and Chakravarty's critique of correlation increasing majorization (in [3]) based on the observation that "there is no a priori reason for a person to regard attributes as substitutes only. Some of the attributes can as well be complements" (p. 36).

7. Conclusion

We have developed concepts of quantiles and comonotonicity for multivariate prospects, thus allowing for the consideration of choice among vectors of payments in different currencies, at different times, in different categories of goods, etc. The multivariate concepts of quantiles and comonotonicity were used to generalize Yaari's dual theory of choice under risk, where decision makers that are insensitive to hedging of comonotonic risks are shown to evaluate prospects using a weighted sum of quantiles. Risk averse decision makers were shown to be characterized within this framework by a reference distribution, making the dual theory as readily applicable as expected utility. Risk attitudes were also analyzed from the point of view of a local utility function. Implications for the ranking of increasing risk aversion is the topic of further research.

Applications of the representation theorem to the measurement of multi-attribute inequality were also discussed. The flexibility in its handling of attitudes to correlation between attributes is a promising feature of the decision functional.

Appendix A. Proofs of results in the main text

A.1. Proof of Proposition 3

By definition, there are two convex lower semi-continuous functions V_1 and V_2 and a random vector $U =_d \mu$ such that $X = \nabla V_1(U)$ and $Y = \nabla V_2(U)$ almost surely. Similarly, there are convex functions V_3 and V_4 and a random vector $\tilde{U}$ such that $Y = \nabla V_3(\tilde{U})$ and $Z = \nabla V_4(\tilde{U})$. Now the assumptions on the absolute continuity of μ and the distribution of Y imply that ∇V_2 is essentially unique. Hence, $\nabla V_2 = \nabla V_3$ and, therefore, $U = \tilde{U}$ holds almost surely. It follows that X and Z are μ-comonotonic. □

A.2. Proof of Proposition 4

The equivalence between (a) and (b) is a famous result stated and extended by many authors, notably Hardy, Littlewood, Pólya, Blackwell, Stein, Sherman, Cartier, Fell, Meyer and Strassen. See Theorem 2 of [17] for an elegant proof. We now show that (b) implies (c). Suppose (b) holds. As explained in Section 2.1, there exists a map ζ such that $\varrho_\mu(X) = \mathbb{E}[\zeta(X) \cdot X]$ and $\zeta(X) =_d \mu$. Now, $\mathbb{E}[\zeta(X) \cdot X] = \mathbb{E}[\zeta(X) \cdot \mathbb{E}[\hat{Y}|X]] = \mathbb{E}[\zeta(X) \cdot \hat{Y}]$, which is less than $\varrho_\mu(Y)$. Next, we show that (c) implies (d). Indeed, the convex closure $\overline{co}(\mathrm{equi}(Y))$ of the equidistribution class of Y is a closed convex set and hence characterized by its support functional $\varrho_\mu(Y)$. Therefore, $X \in \overline{co}(\mathrm{equi}(Y))$ is equivalent to $\mathbb{E}[Z \cdot X] \leqslant \varrho_\mu(Y)$ for all Z, which in turn is equivalent to $\varrho_\mu(X) \leqslant \varrho_\mu(Y)$. Now, we show that (d) implies (e). Indeed, if $X \in \overline{co}(\mathrm{equi}(Y))$, then there is a sequence $(Y_k^n)_{k \leqslant n}$ of random vectors each distributed as Y and positive weights α_k^n such that $\sum_{k=1}^n \alpha_k^n = 1$ and $X = \lim_{n\to\infty} \sum_{k=1}^n \alpha_k^n Y_k^n$. Then, for any law invariant concave functional, we have $\Phi(\sum_{k=1}^n \alpha_k^n Y_k^n) \leqslant \sum_{k=1}^n \alpha_k^n \Phi(Y_k^n) = \Phi(Y)$ and the conclusion follows by upper semi-continuity. Finally, (e) implies (a) because when $\mathcal{L}_X$ is absolutely continuous with respect to Lebesgue measure, for any bounded concave function f, $X \mapsto \mathbb{E} f(X)$ is a law invariant concave upper semi-continuous functional. □

A.3. Proof of Theorem 1

Note first that γ defined for all prospects X by $\gamma(X) = \mathbb{E}[Q_X(U) \cdot \phi(U)]$ for a function ϕ such that $\phi(U) \in (\mathbb{R}_-)^d$ is Lipschitz and monotonic, so that Axioms $1'$ and $2'$ are satisfied for a preference relation represented by γ. Finally, comonotonic independence follows directly from the fact that for any two prospects X and Y, the generalized quantile functions Q_X, Q_Y and Q_{X+Y} satisfy $Q_{X+Y}(U) = Q_X(U) + Q_Y(U)$. We now show this fact. By the definition of the generalized quantile functions, we have $\mathbb{E}[Q_{X+Y}(U) \cdot U] = \sup_{\tilde{U}=_d U} \mathbb{E}[(X+Y) \cdot \tilde{U}] \leqslant \sup_{\tilde{U}=_d U} \mathbb{E}[X \cdot \tilde{U}] + \sup_{\tilde{U}=_d U} \mathbb{E}[Y \cdot \tilde{U}] = \mathbb{E}[Q_X(U) \cdot U] + \mathbb{E}[Q_Y(U) \cdot U]$. On the other hand, we also have $\mathbb{E}[Q_{X+Y}(U) \cdot U] = \sup_{\tilde{Z}=_d X+Y} \mathbb{E}[\tilde{Z} \cdot U] \geqslant \mathbb{E}[(Q_X(U) + Q_Y(U)) \cdot U]$ since by construction, $Q_X(U) =_d X$ and $Q_Y(U) =_d Y$, and the desired equality follows.

Conversely, we now prove that a preference relation $\succsim$ satisfying Axioms $1'$, $2'$ and $3'$ is represented by a functional γ defined for all prospect X by $\gamma(X) = \mathbb{E}[Q_X(U) \cdot \phi(U)]$ for a

function ϕ such that $\phi(U) \in (\mathbb{R}_-)^d$ almost surely. By Axiom 1′, there exists a functional γ representing $\succsim$ and there is a point $Z \in L^2_d$, where γ is Fréchet differentiable with non-zero gradient D. Let Q_Z be the generalized quantile of Z relative to μ. There exists a $U \in L^2_d$ with distribution μ such that $Z = Q_Z(U)$ almost surely. Let X and Y be two prospects in L^d_d with μ-quantile functions Q_X and Q_Y respectively. By the definition of μ-comonotonicity, $Q_X(U)$, $Q_Y(U)$ and $Z = Q_Z(U)$ are μ-comonotonic. By Axiom 2′, γ is law invariant, so that $\gamma(X) \geqslant \gamma(Y)$ is equivalent to $\gamma(Q_X(U)) \geqslant \gamma(Q_Y(U))$. Hence, by Axiom 3′, $\gamma(X) \geqslant \gamma(Y)$ implies that for any $0 < \epsilon \leqslant 1$, we have $\gamma(\epsilon Q_X(U) + (1-\epsilon)Z) \geqslant \gamma(\epsilon Q_Y(U) + (1-\epsilon)Z)$. Hence, $\gamma(Z + \epsilon(Q_X(U) - Z)) \geqslant \gamma(Z + \epsilon(Q_Y(U) - Z))$ and, therefore, $\gamma(Z) + \mathbb{E}[D \cdot \epsilon(Q_X(U) - Z)] \geqslant \gamma(Z) + \mathbb{E}[D \cdot \epsilon(Q_Y(U) - Z)] - o(\epsilon)$, or, finally, $\mathbb{E}[D \cdot Q_X(U)] \geqslant \mathbb{E}[D \cdot Q_Y(U)]$.

Suppose now that X and Y are two prospects such that $\mathbb{E}[D \cdot Q_X(U)] = \mathbb{E}[D \cdot Q_Y(U)]$. We shall show that $\gamma(Q_X(U)) = \gamma(Q_Y(U))$ and, hence, that $\gamma(X) = \gamma(Y)$, thereby concluding that the functional $X \mapsto \mathbb{E}[D \cdot Q_X(U)]$ represents $\succsim$. Indeed, suppose that $\mathbb{E}[D \cdot Q_X(U)] = \mathbb{E}[D \cdot Q_Y(U)]$. We will show shortly that there exists a function ϕ such that $\mathbb{E}[D \cdot \nabla\phi(U)] > 0$ and, hence, that $\mathbb{E}[D \cdot (Q_X(U) + \epsilon\nabla\phi(U))] > \mathbb{E}[D \cdot Q_Y(U)]$ and $\mathbb{E}[D \cdot (Q_X(U) - \epsilon\nabla\phi(U))] < \mathbb{E}[D \cdot Q_Y(U)]$ for any $\epsilon > 0$. Using the result above yields $\gamma(Q_X(U) + \epsilon\nabla\phi(U)) \geqslant \gamma(Q_Y(U))$ and $\gamma(Q_X(U) - \epsilon\nabla\phi(U)) \leqslant \gamma(Q_Y(U))$. Hence, $\gamma(Q_X(U)) = \gamma(Q_Y(U))$ by the continuity of γ. Let us now show that $\mathbb{E}[D \cdot \nabla\phi(U)] = 0$ for all gradient functions $\nabla\phi$ yields a contradiction. Calling V_Z the convex function such that $Z = Q_Z(U) = \nabla V_Z(U)$ almost surely, D is the Fréchet derivative of γ at $Z = \nabla V_Z(U)$. Hence, $\mathbb{E}[D \cdot \nabla\phi(U)] = 0$ implies that $\gamma(\nabla(V_Z(U) + \epsilon\phi(U))) = \gamma(\nabla V_Z(U)) + o(\epsilon)$. This is true for all gradient functions $\nabla\phi$ and, in particular, for $\phi = (V_\epsilon - V_Z)/\epsilon$, where V_ϵ is such that $Z_\epsilon = \nabla V_\epsilon(U)$ converges to Z in L^2. We then have $\gamma(Z_\epsilon) - \gamma(Z) = o(\epsilon)$ and, hence, $D = 0$, which contradicts Axiom 1′. We have shown that $\succsim$ is represented by the functional $X \mapsto \mathbb{E}[D \cdot Q_X(U)]$. As μ is absolutely continuous with respect to Lebesgue measure, D can be written as $\phi(U)$ for some function ϕ which takes values in $(\mathbb{R}_-)^d$ by Axiom 2′. □

A.4. Proof of Theorem 2

That (a) implies (b) follows from Proposition 4. We now show that (b) implies (c). Axiom 4 implies that $\gamma(Q_X + \alpha(\tilde{X} - Q_X)) \geqslant \gamma(X)$ for all $\alpha \in (0, 1]$ and all $\tilde{X}$ in the equidistribution class of X. The representation of Theorem 1 implies the differentiability of γ at $Q_X(U)$ for any X, call D_X its gradient. This implies that $\mathbb{E}[\tilde{X} \cdot D_X] \leqslant \mathbb{E}[Q_X(U) \cdot D_X]$ for all $\tilde{X} =_d X$. Hence, $\gamma(X) = -\varrho_{\mathcal{L}_{\phi(U)}}(X)$. Thus, by Axiom 3′, comonotonicity with respect to μ implies comonotonicity with respect to $\mathcal{L}_{\phi(U)}$. By Lemma 10 in [6], this implies that $\phi(U) = -\alpha U - x_0$ for some $\alpha > 0$ and $x_0 \in \mathbb{R}^d$, and the result follows. Finally, we show that (c) implies (a). Assume (c), in which case for all $X \in \mathbb{L}^2_d$, $-\gamma(X) = \alpha\mathbb{E}[Q_X(U) \cdot U] + u_0 \cdot \mathbb{E}[X]$. Thus, $-\gamma(X) = \alpha\varrho_\mu(X) + u_0 \cdot \mathbb{E}[X]$. Therefore, by Proposition 4, $X \geqslant_{cv} Y$ implies $\varrho_\mu(X) \leqslant \varrho_\mu(Y)$, and so $\gamma(X) \geqslant \gamma(Y)$. □

References

[1] M. Allais, Le comportement de l'homme rationnel devant le risque: Critiques des postulats et axiomes de l'école américaine, Econometrica 21 (1953) 503–546.

[2] A. Atkinson, On the measurement of inequality, J. Econ. Theory 2 (1970) 244–263.

[3] F. Bourguignon, S. Chakravarty, The measurement of multidimensional poverty, J. Econ. Inequality 1 (2003) 25–49.

[4] G. Carlier, R.-A. Dana, A. Galichon, Pareto efficiency for the concave order and multivariate comonotonicity, unpublished manuscript, Université Paris-Dauphine and École polytechnique, 2009.

[5] S. Chew, E. Karni, Z. Safra, Risk aversion in the theory of expected utility with rank dependent probabilities, J. Econ. Theory 42 (1987) 370–381.
[6] I. Ekeland, A. Galichon, M. Henry, Comonotone measures of multivariate risks, Math. Finance, forthcoming.
[7] T. Gajdos, J. Weymark, Multidimensional generalized Gini indices, Econ. Theory 26 (2005) 471–496.
[8] G. Hardy, J. Littlewood, G. Pólya, Inequalities, Cambridge University Press, Cambridge, 1952.
[9] S.-C. Kolm, Multidimensional egalitarianisms, Quart. J. Econ. 91 (1977) 1–13.
[10] M. Landsberger, I. Meilijson, Comonotone allocations, Bickel Lehmann dispersion and the Arrow–Pratt measure of risk aversion, Ann. Oper. Res. 52 (1994) 97–106.
[11] M. Machina, 'Expected utility' analysis without the independence axiom, Econometrica 50 (1982) 277–323.
[12] G. Puccetti, M. Scarsini, Multivariate comonotonicity, J. Multivariate Anal. 101 (1) (2010) 291–304.
[13] D. Preiss, Differentiability of Lipschitz functions on Banach spaces, J. Funct. Anal. 91 (1990) 312–345.
[14] M. Rothschild, J. Stiglitz, Increasing risk: I. A definition, J. Econ. Theory 2 (1970) 225–243.
[15] D. Schmeidler, Subjective probability and expected utility without additivity, Econometrica 57 (1989) 571–587.
[16] M. Shaked, J. Shantikumar, Stochastic Orders, Springer, New York, 2007.
[17] V. Strassen, The existence of probability measures with given marginals, Ann. Math. Stat. 36 (1965) 423–439.
[18] K.-Y. Tsui, Multidimensional generalizations of the relative and absolute inequality indices: the Atkinson–Kolm–Sen approach, J. Econ. Theory 67 (1995) 251–265.
[19] K.-Y. Tsui, Multidimensional inequality and multidimensional generalized entropy measures: an axiomatic derivation, Soc. Choice Welfare 16 (1999) 145–157.
[20] C. Villani, Topics in Optimal Transportation, American Mathematical Society, Providence, 2003.
[21] P. Wakker, Under stochastic dominance Choquet-expected utility and anticipated utility are identical, Theory Dec. 29 (1990) 119–132.
[22] J. Weymark, Generalized Gini inequality indices, Math. Soc. Sci. 1 (1981) 409–430.
[23] M. Yaari, Univariate and multivariate comparisons of risk aversion: a new approach, in: W. Heller, R. Starr, D. Starrett (Eds.), Essays in Honor of Kenneth Arrow, Cambridge University Press, 1986, pp. 173–187.
[24] M. Yaari, The dual theory of choice under risk, Econometrica 55 (1987) 95–115.