Fear of Loss, Inframodularity, and Transfers

Alfred Müller
FB 6 Mathematik
Universität Siegen
D–57068 Siegen, Germany
mueller@mathematik.uni-siegen.de

Marco Scarsini*
Dipartimento di Scienze Economiche e Aziendali
LUISS
Viale Romania 12
I–00197 Roma, Italy
and HEC, Paris
marco.scarsini@luiss.it

March 8, 2010

Abstract

There exist several characterizations of concavity for univariate functions. One of them states that a function is concave if and only if it has non-increasing differences. This definition provides a natural generalization of concavity for multivariate functions, called inframodularity. This paper shows that a finite lottery is preferred to another by all expected utility maximizers with an inframodular utility if and only if the first measure can be obtained from the second via a sequence of suitable transfers. This result is a natural multivariate generalization of Rothschild and Stiglitz's construction based on mean preserving spreads.

Journal of Economic Literature classification numbers: D81.

Keywords: mean preserving spread, integral stochastic orders, risk aversion, ultramodularity, dual cones.

*Partially supported by MIUR-COFIN

1 Introduction

A risk averse decision maker prefers to enjoy a sure wealth w rather than $w + \varepsilon$, where ε is a fair (i.e., zero-mean) random variable: a non-degenerate fair random variable involves possible losses that, in the preference of a risk averter, are not compensated by possible gains. It is well known that in a von Neumann-Morgenstern expected utility context risk aversion coincides with concavity of the decision maker's utility function. A random variable Y is riskier than a random variable X if all risk averters prefer X to Y. Therefore Y is riskier than X if $\mathbb{E}[u(X)] \geqslant \mathbb{E}[u(Y)]$ for all concave functions u (this implies that X and Y have the same mean). This comparative concept of riskier depends only on the distributions of X and Y.

Given a random variable Y, one way to make it less risky is to consider a bounded interval and transfer some probability mass in the distribution of Y from outside the interval to inside, keeping the mean fixed. It is well known that in a sense this is the only way to make Y less risky, since Y is riskier than X if and only if the law of X can be obtained from the law of Y via a sequence of transfers of this type.

Often decisions involve several commodities that are not necessarily priceable. For instance, when comparing two job offers, a person takes into account the salary, the type of job, the working environment, the commuting time from home, etc. and most of these quantities involve randomness of some sort, so that a truly multivariate evaluation is necessary.

The generalization of risk comparison to the multivariate case seems immediate, and it has generally taken to be so. A random vector $\boldsymbol{Y}$ is riskier than a random vector $\boldsymbol{X}$ if all risk averters prefer $\boldsymbol{X}$ to $\boldsymbol{Y}$. Given a non-degenerate zero-mean random vector $\boldsymbol{\varepsilon}$, a risk averter prefers a sure amount $\boldsymbol{w}$ rather than $\boldsymbol{w} + \boldsymbol{\varepsilon}$. Or does she? In $\mathbb{R}^d$ the natural order is only partial, so it is possible to have $\mathbb{E}[\boldsymbol{\varepsilon}] = \mathbf{0}$ even if $\boldsymbol{\varepsilon}$ is never negative. Therefore this definition of risk aversion departs from the univariate rationale of fear of losses, and it just embodies the idea of aversion to randomness.

This paper formalizes a concept of *fear of loss* and shows that it corresponds to a class of utility functions called *inframodular*. A stochastic order for random vectors is then defined, according to which $\boldsymbol{X}$ induces less fear of loss than $\boldsymbol{Y}$ if $\mathbb{E}[u(\boldsymbol{X})] \geqslant \mathbb{E}[u(\boldsymbol{Y})]$ for all inframodular functions u. This is shown to happen if and only if the distribution of $\boldsymbol{X}$ can be obtained from the distribution of $\boldsymbol{Y}$ via a sequence of suitable transfers that naturally generalize the ones studied in the univariate case.

1.1 Existing literature

The classical results of Pratt (1964); Arrow (1970) provide a comparative study of risk aversion in terms of local and global conditions on the decision maker's univariate utility function. Rothschild and Stiglitz (1970, 1971, 1972) study the dual problem of comparison of risks. To do this, they use some balayage results previously unknown in the economic literature (the reader is referred to Hardy, Littlewood, and Polya, 1929;

Hardy, Littlewood, and Pólya, 1988; Sherman, 1951; Blackwell, 1951, 1953; Cartier, Fell, and Meyer, 1964; Strassen, 1965, for the classical comparison results). Most importantly Rothschild and Stiglitz focus on the idea of mean preserving spread, i.e., a transfer of probability mass from inside a finite interval to outside the interval, that does not alter the mean of a distribution. They show that mean preserving spreads are the building blocks of distribution comparison, since a risk X is preferred to another risk Y by all risk averters if and only if the distribution of Y can be obtained from the distribution of X via a sequence of mean preserving spreads (their results are framed in a more general and precise way by Müller (1996) and Machina and Pratt (1997)). Rothschild and Stiglitz's papers have had a tremendous impact on the literature, reducing many comparisons to analyzing the effect of a single mean preserving spread. The reader can find some useful reference on the duality between risk and risk aversion in Scarsini (1994).

de Finetti (1952) is the first to consider a form of bivariate risk aversion that involves the comparison of two lotteries having each two equally probable bi-dimensional outcomes. The two lotteries involve the same quantities of two commodities, and differ only for the way the items are combined: in the first lottery one possible outcome is a small quantity of one commodity combined with a small quantity of the other, and the other outcome is a large quantity of one commodity combined with a large quantity of the other; in the other lottery the small quantity of one commodity is combined with the large quantity of the other. Preference of the second lottery over the first is a form of bivariate risk aversion. These results are re-discovered more than twenty years later by Richard (1975). Epstein and Tanny (1980) use the framework proposed by Richard (1975) to prove comparison results in terms of generalized correlation. Even if they don't use the term, these authors frame multivariate risk aversion in terms of submodular utility functions. The relevance of supermodularity/submodularity in economic theory is widespread (see Topkis, 1998, for an extended analysis of its theory and applications). Comparison of distribution functions in terms of the supermodular order is an important tool to study positive dependence (see, e.g, Joe, 1997; Müller and Scarsini, 2000).

Kihlstrom and Mirman (1974) propose a multivariate generalization of the Arrow-Pratt theory of risk aversion, when cardinal utility functions represent the same ordinal preferences. Kihlstrom and Mirman (1981) extend these results and study monotone multivariate risk aversion when preferences are homothetic.

Building on Richard's results, Duncan (1977) defines a matrix measure of multivariate local risk aversion and studies its properties in terms of multivariate risk premiums. Karni (1979) relates local and global concepts of multivariate risk aversion and achieves comparative results in the spirit of Arrow (1970); Pratt (1964).

Multivariate utility functions have been recently studied in the management science literature and their construction based on lotteries that combine good and bad outcomes has been examined (see, e.g.,. Eeckhoudt, Rey, and Schlesinger, 2007; Tsetlin and Winkler, 2009; Denuit, Eeckhoudt, and Rey, 2010).

Elton and Hill (1992) prove a result à la Rothschild and Stiglitz (1970) for measures on separable Banach spaces, that includes as a particular case Euclidean spaces. They show that if one measure dominates another one in terms of the convex order, then the first can be transformed into the second via a sequence of fusions. Fusions are basically the reverse operation of mean preserving spread, but in a more general abstract setting. Elton and Hill (1998) give an elementary proof of their result for purely atomic measures with a finite number of atoms in $\mathbb{R}^n$. Although these articles have no economic motivation, they provide very useful tools that are used in this paper.

1.2 Fear of loss and inframodularity

This paper focuses on multivariate transfers that naturally generalize the concept that is so fruitfully used in the univariate case. Aversion to risk represents preference for a sure amount of money w versus a random amount having expectation w. In different words a risk averse decision maker does not like to add to her sure wealth w a random variable ε having zero mean. What's the reason for disliking randomness? Obviously it has to see with the possibility of ending up with less than the initial wealth w once ε is realized. In fact any non-degenerate random variable with zero mean can assume negative values with positive probability, therefore can give rise to a loss. A risk averter fears losses, and the possibility of getting a positive gain does not compensate for these possible losses. In the univariate case *fear of loss* coincides with risk aversion, which coincides with concavity of the agent's utility function.

It is commonly assumed that this is the case also in the multivariate case. An agent who prefers any sure multivariate wealth $\boldsymbol{w}$ to a random vector $\boldsymbol{w} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon}$ has mean vector $\mathbf{0}$ is risk averse, so her utility function is concave. But does this embody the same rationale that exists in the univariate case? Since the natural order on $\mathbb{R}^d$ is only partial, a random vector $\boldsymbol{\varepsilon}$ can have zero expectation even if it never assumes values that are smaller than zero. For instance in $\mathbb{R}^2$ a random vector that assumes with equal probability the values $(1, -1)$ and $(-1, 1)$ has zero mean. Is there any good reason to fear this random variable? This paper argues that the spirit of the univariate case is best kept by limiting aversion to zero-mean vectors that involve possible losses and gains, for instance a vector that assumes with equal probability the values $(-1, -1)$ and $(1, 1)$.

To do this consider a transfer that mimics the mean preserving spread as described by Müller (1996) and Machina and Pratt (1997), i.e., a transfer that moves mass from inside a (multidimensional) intervals to the sets below and above the interval. A loss fearful individual dislikes this transfer. A decision maker who dislikes any such transfer has an inframodular utility function. The central result of this paper is the converse of this statement. If a random vector $\boldsymbol{X}$ is preferred to another random vector $\boldsymbol{Y}$ by all decision makers with an inframodular utility function, then the distribution of $\boldsymbol{Y}$ can be obtained from the distribution of $\boldsymbol{X}$ via a sequence of such transfers.

A concave function f on $\mathbb{R}$ is characterized by having non-increasing differences: $f(x+\varepsilon) - f(x)$ is non-increasing in x for all positive ε. A function $g : \mathbb{R}^d \to \mathbb{R}$ has non-increasing differences if and only if it is inframodular. Therefore inframodularity is a natural generalization of concavity to the multivariate setting. Any inframodular function is just the negative of a ultramodular function. Ultramodular functions have been studied and used by many authors, often under different names. Marinacci and Montrucchio (2005) examine this class of functions in detail and provide several relevant references.

As mentioned in Müller and Scarsini (2001), a concave function of a positive linear combination of variables is inframodular. Therefore the analysis developed here allows the comparison of portfolios of commodities for any given price vector.

The main results in this paper are proved using functional analytical tools of duality. Duality theory has been used before to prove stochastic comparison results (see, among others, Brumelle and Vickson, 1975; Ziegler, 1968; Border, 1991; Castagnoli and LiCalzi, 1997; Castagnoli and Maccheroni, 2000; Dubra, Maccheroni, and Ok, 2004).

This paper is organized as follows. Section 2 describes different types of transfers. Section 3 states the main results. Section 4 contains the proofs.

2 Transfers

This section introduces a general definition of transfer. To do this the definition of some useful classes of functions is needed. The following notation is used:

$$\boldsymbol{x} \vee \boldsymbol{y} := (\max\{x_1, y_1\}, \dots, \max\{x_d, y_d\}),$$
$$\boldsymbol{x} \wedge \boldsymbol{y} := (\min\{x_1, y_1\}, \dots, \min\{x_d, y_d\}).$$

Definition 2.1. (a) Let $A \subset \mathbb{R}^d$ be convex. A function $f : A \to \mathbb{R}$ is *convex* if for all $\boldsymbol{x}, \boldsymbol{y} \in A$ and all $\alpha \in [0,1]$

$$f(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leqslant \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}). \tag{2.1}$$

A function is *concave* if the reverse inequality holds.

(b) Let $A \subset \mathbb{R}^d$ be convex. A function $f : A \to \mathbb{R}$ is *component-wise convex* if (2.1) holds for all $\boldsymbol{x}, \boldsymbol{y} \in A$ such that $x_j = y_j$ for $j \neq i$, for some $i \in \{1, \dots, d\}$. A function is *component-wise concave* if the reverse inequality holds.

(c) Let $B \subset \mathbb{R}^d$ be a lattice. A function $f : B \to \mathbb{R}$ is *supermodular* if for all $\boldsymbol{x}, \boldsymbol{y} \in B$

$$f(\boldsymbol{x}) + f(\boldsymbol{y}) \leqslant f(\boldsymbol{x} \vee \boldsymbol{y}) + f(\boldsymbol{x} \wedge \boldsymbol{y}).$$

A function is *submodular* if the reverse inequality holds.

(d) Let $C \subset \mathbb{R}^d$ be a convex lattice. A function $f : C \to \mathbb{R}$ is *ultramodular* if for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{z} \in C$ such that $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{z} + \boldsymbol{w}$ and $\boldsymbol{w} \leqslant \boldsymbol{x}, \boldsymbol{y} \leqslant \boldsymbol{z}$

$$f(\boldsymbol{x}) + f(\boldsymbol{y}) \leqslant f(\boldsymbol{z}) + f(\boldsymbol{w}).$$

A function is *inframodular* if the reverse inequality holds.

Topkis (1998) is the classical reference for properties and applications of supermodular functions. The term "ultramodular" has been coined by Marinacci and Montrucchio (2005), who provide a thorough analysis of this class of functions, previously known under different, sometime misleading, names, such as "directionally convex."

Let $S \subset \mathbb{R}^d$ be compact, and let $\mathscr{S}$ be the Borel-σ-algebra on S. For a signed measure μ on $(S, \mathscr{S})$, its positive and negative parts are denoted μ^+ and μ^-, respectively, $|\mu| = \mu^+ + \mu^-$ is the total variation, and $\|\mu\| := \mu^+(S) + \mu^-(S)$ is the total variation norm. Denote by $\mathbb{M}$ the set of all signed measures on S with finite total variation norm $\|\mu\| < \infty$ and with the property that $\mu^+(S) = \mu^-(S)$. Notice that for any two probability measures P, Q the difference $Q - P \in \mathbb{M}$, and that in fact $\mathbb{M}$ is the linear space spanned by the differences of probability measures.

A degenerate probability measure on $\boldsymbol{x}$ is denoted $\delta_{\boldsymbol{x}}$. Given two probability measures P, Q supported on a finite subset of $\mathbb{R}^d$, call the signed measure $Q - P$ a *transfer* from Q to P. If

$$(Q - P)^- = \sum_{i=1}^{n} \beta_i \delta_{\boldsymbol{y}_i} \quad \text{and} \quad (Q - P)^+ = \sum_{i=1}^{m} \alpha_i \delta_{\boldsymbol{x}_i},$$

then the transfer $Q - P$ removes probability mass β_i from points $\boldsymbol{y}_i$, $i = 1, \ldots, n$ and adds probability mass α_i to $\boldsymbol{x}_i$, $i = 1, \ldots, m$. To indicate this transfer write

$$\sum_{i=1}^{n} \beta_i \delta_{\boldsymbol{y}_i} \to \sum_{i=1}^{m} \alpha_i \delta_{\boldsymbol{x}_i}.$$

Definition 2.2. Consider a set $M \subset \mathbb{M}$ of transfers and the class $\mathscr{F} \subset \mathscr{C}$ of continuous functions f such that

$$\sum_{i=1}^{n} \beta_i f(\boldsymbol{y}_i) \geqslant \sum_{i=1}^{m} \alpha_i f(\boldsymbol{x}_i)$$

whenever $\mu \in M$, where

$$\mu := \sum_{i=1}^{n} \beta_i \delta_{\boldsymbol{y}_i} - \sum_{i=1}^{m} \alpha_i \delta_{\boldsymbol{x}_i}.$$

The class $\mathscr{F}$ is said to be *induced* by M.

This definition has the following economic interpretation. Any decision maker using expected utility theory with a utility function $u \in \mathscr{F}$ will prefer Q to P if $Q - P \in M$, i.e. if Q is obtained from P by a transfer in M.

Next comes the definition of simple transfers that induce the classes of functions of Definition 2.1. Here all probability measures are supported on a finite subset of $\mathbb{R}^d$, and all transfers involve a mass $0 \leqslant \eta \leqslant 1$. In the following definition the terminology of Elton and Hill (1998) is adopted.

Definition 2.3. Given a measure P with finite support on $\mathbb{R}^d$, call $\|P\|$ its total mass and

$$\operatorname{bar}(P) := \|P\|^{-1} \int_{\mathbb{R}^d} \boldsymbol{x} \, \mathrm{d}P(\boldsymbol{x})$$

the *barycenter* of P.

For a discrete measure $P = \sum_{i=1}^m \alpha_i \delta_{\boldsymbol{x}_i}$ this simplifies to

$$\operatorname{bar}(P) = \frac{1}{\sum_{i=1}^m \alpha_i} \sum_{i=1}^m \alpha_i \boldsymbol{x}_i$$

2.1 Simple transfers

A simple transfer μ has the form

$$\mu = \beta_1 \delta_{\boldsymbol{y}_1} + \beta_2 \delta_{\boldsymbol{y}_2} - \alpha_1 \delta_{\boldsymbol{x}_1} - \alpha_2 \delta_{\boldsymbol{x}_2},$$

where it is possible that $\boldsymbol{y}_1 = \boldsymbol{y}_2$ or $\boldsymbol{x}_1 = \boldsymbol{x}_2$. Therefore a simple transfer involves the move of some probability mass from at most two points to at most two other points. In the sequel only simple transfers that preserve the barycenter are considered.

Simple convex transfer

Given $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{z} \in \mathbb{R}^d$ and $\alpha, \beta, \gamma, \varepsilon \in [0, 1]$ such that

$$\begin{aligned} &\boldsymbol{z} = \alpha \boldsymbol{x} + (1 - \alpha) \boldsymbol{y}, \quad \boldsymbol{w} = \beta \boldsymbol{y} + (1 - \beta) \boldsymbol{x}, \\ &\gamma \boldsymbol{x} + (1 - \gamma) \boldsymbol{y} = \varepsilon \boldsymbol{z} + (1 - \varepsilon) \boldsymbol{w}, \end{aligned}$$

a simple transfer $\eta \left(\varepsilon \delta_{\boldsymbol{z}} + (1 - \varepsilon) \delta_{\boldsymbol{w}} \right) \to \eta \left(\gamma \delta_{\boldsymbol{x}} + (1 - \gamma) \delta_{\boldsymbol{y}} \right)$ is called *convex*. The reverse transfer is called *concave*. When $\alpha = \beta$, hence $\gamma = \varepsilon = 1/2$, the transfer is called symmetric. Notice that, if $\alpha = 1 - \beta$, then $\boldsymbol{w} = \boldsymbol{z}$.

Figure 1 about here.

Simple component-wise convex transfer

Given $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{z} \in \mathbb{R}^d$ and $\alpha, \beta, \gamma, \varepsilon \in [0,1]$ such that $x_j = y_j$ for all $j \neq i$ and

$$\begin{aligned}
&\boldsymbol{z} = \alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}, \quad \boldsymbol{w} = \beta \boldsymbol{y} + (1-\beta)\boldsymbol{x}, \\
&\gamma \boldsymbol{x} + (1-\gamma)\boldsymbol{y} = \varepsilon \boldsymbol{z} + (1-\varepsilon)\boldsymbol{w},
\end{aligned}$$

a simple transfer $\eta\left(\varepsilon\delta_{\boldsymbol{z}} + (1-\varepsilon)\delta_{\boldsymbol{w}}\right) \to \eta\left(\gamma\delta_{\boldsymbol{x}} + (1-\gamma)\delta_{\boldsymbol{y}}\right)$ is called *component-wise convex*. The reverse transfer is called *component-wise concave*. As before, when $\alpha = \beta$, hence $\gamma = \varepsilon = 1/2$, the transfer is called symmetric.

Figure 2 about here.

Simple supermodular transfer

Given $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{z} \in \mathbb{R}^d$ such that

$$\boldsymbol{x} = \boldsymbol{z} \wedge \boldsymbol{w}, \quad \boldsymbol{y} = \boldsymbol{z} \vee \boldsymbol{w},$$

A simple transfer $\eta\left(\frac{1}{2}\delta_{\boldsymbol{z}} + \frac{1}{2}\delta_{\boldsymbol{w}}\right) \to \eta\left(\frac{1}{2}\delta_{\boldsymbol{x}} + \frac{1}{2}\delta_{\boldsymbol{y}}\right)$ is called *supermodular*. The reverse transfer is called *submodular*.

Figure 3 about here.

Simple ultramodular transfer

Given $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{z} \in \mathbb{R}^d$ and $\gamma, \varepsilon \in [0,1]$ such that $\boldsymbol{x} \leqslant \boldsymbol{w}, \boldsymbol{z} \leqslant \boldsymbol{y}$ and

$$\gamma \boldsymbol{x} + (1-\gamma)\boldsymbol{y} = \varepsilon \boldsymbol{z} + (1-\varepsilon)\boldsymbol{w},$$

a simple transfer $\eta\left(\varepsilon\delta_{\boldsymbol{z}} + (1-\varepsilon)\delta_{\boldsymbol{w}}\right) \to \eta\left(\gamma\delta_{\boldsymbol{x}} + (1-\gamma)\delta_{\boldsymbol{y}}\right)$ is called *ultramodular*. The reverse transfer is called *inframodular*. When $\gamma = \varepsilon = 1/2$, the transfer is called symmetric.

Figure 4 about here.

Component-wise convex and supermodular transfers are particular cases of ultramodular transfers. The following proposition shows a stronger property.

Proposition 2.4. *Any simple ultramodular transfer can be obtained by combining simple supermodular and component-wise convex transfers.*

It is immediate to see that the classes of convex, concave, component-wise convex, component-wise concave, supermodular, submodular, ultramodular, inframodular functions are induced by the set of simple symmetric transfers with the same name.

General (non-simple) transfers are obtained by iterating simple transfers. In dimension 1 a convex transfer is nothing else than a mean-preserving spread, as studied by Rothschild and Stiglitz (1970, 1971, 1972). In dimension d concave transfers are related to fusions (see Elton and Hill, 1992).

Figure 5 about here.

The generalization of a mean preserving spread to $\mathbb{R}^d$ requires some care. Given a convex set $A \subset \mathbb{R}^d$, one may think that a transfer of mass from A to A^c that preserves the barycenter is a convex transfer, i.e., can be obtained as a sequence of simple convex transfers. This is in general not the case, as the following counterexample easily shows. Take P, Q probability measures on $\mathbb{R}^2$ defined as

$$P = \frac{1}{4}\left(\delta_{\boldsymbol{e}_1} + \delta_{-\boldsymbol{e}_1} + \delta_{\boldsymbol{e}_2} + \delta_{-\boldsymbol{e}_2}\right) \quad \text{and} \quad Q = \frac{1}{4}\left(\delta_{\boldsymbol{x}} + \delta_{-\boldsymbol{x}} + \delta_{\boldsymbol{y}} + \delta_{-\boldsymbol{y}}\right)$$

where $\boldsymbol{e}_i$ is the i-th element of the canonical base, $\boldsymbol{x} = (2/3, 2/3)$, and $\boldsymbol{y} = (2/3, -2/3)$. The two measures have the same barycenter $(0,0)$. It is clear that $\operatorname{supp}(Q) \subseteq [\operatorname{conv}(\operatorname{supp}(P))]^c$ and $\operatorname{supp}(P) \subseteq [\operatorname{conv}(\operatorname{supp}(Q))]^c$. However, to have the convex ordering it would be necessary that the convex hull of the support of one probability measure be included in the support of the other, but $\operatorname{conv}(\operatorname{supp}(P)) \nsubseteq \operatorname{conv}(\operatorname{supp}(Q))$ and $\operatorname{conv}(\operatorname{supp}(Q)) \nsubseteq \operatorname{conv}(\operatorname{supp}(P))$. Therefore neither Q can be obtained from P via a sequence of simple convex transfers, nor vice versa.

Figure 6 about here.

Simple supermodular transfers and their iterations have been studied by Tchen (1980).

In all the situations examined in this paper transfers are reversible, so if a probability measure P is obtained from Q via a sequence of transfers of some type, then Q is obtained from P via a sequence of transfers of the reverse type. Reversibility is used in the proof of some results. In general reversibility of transfers does not hold, for instance fusions are not always reversible, as Elton and Hill (1992) show.

3 Main results

3.1 General ultramodular transfers

For univariate distributions Müller (1996) and Machina and Pratt (1997) show that mean-preserving spreads correspond to taking mass from some bounded interval and moving it above and below this interval, without affecting the mean.

Figures 7 and 8 about here.

The following theorem shows that something similar holds for ultramodular transfers in the multivariate case. The following notation is used: given $\boldsymbol{x} \in \mathbb{R}^d$ define the *upper set* and *lower set*

$$U(\boldsymbol{x}) := \{\boldsymbol{z} \in \mathbb{R}^d : \boldsymbol{z} \geqslant \boldsymbol{x}\} \quad \text{and} \quad L(\boldsymbol{x}) := \{\boldsymbol{z} \in \mathbb{R}^d : \boldsymbol{z} \leqslant \boldsymbol{x}\},$$

and for two ordered points $\boldsymbol{x} \leqslant \boldsymbol{y}$ define the *interval between* $\boldsymbol{x}$ *and* $\boldsymbol{y}$:

$$B(\boldsymbol{x}, \boldsymbol{y}) := \{\boldsymbol{z} \in \mathbb{R}^d : \boldsymbol{x} \leqslant \boldsymbol{z} \leqslant \boldsymbol{y}\}.$$

Theorem 3.1. *Let P, Q two discrete probability measures on $\mathbb{R}^d$ with* $\operatorname{bar}(P) = \operatorname{bar}(Q)$ *such that for some* $\boldsymbol{x} \leqslant \boldsymbol{y}$

$$\operatorname{supp}(P) \subset B(\boldsymbol{x}, \boldsymbol{y}), \quad \operatorname{supp}(Q) \subset L(\boldsymbol{x}) \cup U(\boldsymbol{y}).$$

Then Q can be obtained from P via a sequence of simple ultramodular transfers.

Figure 9 about here.

Theorem 3.1 justifies the following definition.

Definition 3.2. Given $\boldsymbol{x} \leqslant \boldsymbol{y}$, a transfer

$$\mu := \sum_{i=1}^{n} \beta_i \delta_{\boldsymbol{z}_i} - \sum_{i=1}^{m} \alpha_i \delta_{\boldsymbol{w}_i}.$$

is called *ultramodular* if

$$\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n \in B(\boldsymbol{x}, \boldsymbol{y}), \quad \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m \in L(\boldsymbol{x}) \cup U(\boldsymbol{y}), \quad \text{and} \quad \sum_{i=1}^{n} \beta_i \boldsymbol{z}_i = \sum_{i=1}^{m} \alpha_i \boldsymbol{w}_i.$$

The reverse transfer is called *inframodular*.

It is interesting to notice that the concept of inframodular (or ultramodular) transfer involves both the vector space and the order structure of $\mathbb{R}^d$, whereas the concave (or convex) transfer is based only on the vector space structure of $\mathbb{R}^d$. An ultramodular transfer moves probability mass from some points in an interval to other points that are either smaller or larger than all points in the interval. A convex transfer just moves mass away from a point. In a univariate setting the difference between the two concepts disappears, but in the multivariate case they represent two different attitudes towards randomness.

3.2 Integral orders and transfers

Definition 3.3. A probability measure P is dominated by a probability measure Q with respect to the integral order $\leqslant_{\mathscr{F}}$ (denoted $P \leqslant_{\mathscr{F}} Q$) if

$$\int u \, \mathrm{d}P \leqslant \int u \, \mathrm{d}Q \quad \text{for all} \quad u \in \mathscr{F}.$$

The economic meaning of this definition is that any expected utility maximizer with a utility function $u \in \mathscr{F}$ prefers the lottery Q to the lottery P.

For the general theory of stochastic orders the reader is referred to Müller and Stoyan (2002); Shaked and Shanthikumar (2007). Arlotto and Scarsini (2009) study a family of integral orders $\leqslant_{\mathscr{F}}$ where $\mathscr{F}$ can be, among others, any of the classes in Definition 2.1.

Rothschild and Stiglitz (1970) prove (under some regularity conditions) that if a measure P on $\mathbb{R}$ dominates Q in terms of the concave order, then Q can be obtained from P via a sequence of mean preserving spreads. Machina and Pratt (1997) refine the result using a more general definition of mean preserving spread. Elton and Hill (1998) prove an analogous theorem for measures on $\mathbb{R}^d$. The following theorem proves a similar result for the inframodular order.

Theorem 3.4. *Let $\mathscr{F}$ be the class of inframodular functions, and let P and Q be two measures on $\mathbb{R}^d$ with finite support. Then the following statements are equivalent:*

(a) *$P \leqslant_{\mathscr{F}} Q$,*

(b) *P can be obtained from Q by a finite number of simple inframodular transfers,*

(c) *P can be obtained from Q by a finite number of inframodular transfers as in Definition 3.2.*

4 Proofs

4.1 General transfers

A set $S \subset \mathbb{R}^d$ is called *comonotonic* if it is totally ordered in the natural componentwise order of $\mathbb{R}^d$. Given a convex set $A \in \mathbb{R}^d$, the set of its extreme points is denoted by $\mathrm{Ex}(A)$.

Lemma 4.1. *Let P be any measure supported on $B(\boldsymbol{x}, \boldsymbol{y})$, and call P^* a probability measure supported on $\mathrm{Ex}(B(\boldsymbol{x}, \boldsymbol{y}))$, such that $\mathrm{supp}(P^*)$ is comonotonic, and $\mathrm{bar}(P^*) = \mathrm{bar}(P)$. Then P^* can be obtained from P via a sequence of ultramodular transfers.*

Proof. First of all existence of P^* is shown. Call $P_1, \ldots, P_d$ the univariate marginals of P. For each P_i there exists a measure P_i^* supported on the extreme points x_i, y_i and such that $\mathrm{bar}(P_i^*) = \mathrm{bar}(P_i)$. Consider the upper Fréchet bound of d-variate measures with marginals $P_1^*, \ldots, P_d^*$. This is P^*.

Take each point $\boldsymbol{z} \in \mathrm{supp}(P)$ and split its mass into the two points $(x_1, z_2, \ldots, z_d)$ and $(y_1, z_2, \ldots, z_d)$ in such a way that the barycenter is preserved (there is only one way to do this). Now all the points in the support of the new measure have their first coordinate equal to either x_1 or y_1. Repeat the operation for all the remaining coordinates. Now the new measure $\tilde{P}$ is supported only on extreme points of $B(\boldsymbol{x}, \boldsymbol{y})$. For any pair of points $\boldsymbol{s}, \boldsymbol{t} \in \mathrm{supp}(\tilde{P})$, move as much mass as possible to $\boldsymbol{s} \wedge \boldsymbol{t}$ and $\boldsymbol{s} \vee \boldsymbol{t}$, keeping the barycenter fixed. In the end the obtained measure is exactly P^*. □

Lemma 4.1 says that using a sequence of ultramodular transfers any measure on a compact interval can be trnaformed into the unique measure whose univariate marginals are maximal with respect to the convex order (therefore are supported on the extreme points of the interval), and whose joint distribution is the upper Fréchet bound in the class of d-variate distributions with these marginals.

Corollary 4.2. *If $\mathrm{bar}(P) = \alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y}$, then $\mathrm{supp}(\tilde{P}) = \{\boldsymbol{x}, \boldsymbol{y}\}$.*

Proof of Proposition 2.4. The proof is similar to the one of Lemma 4.1 and therefore omitted. □

Proof of Theorem 3.1. Consider each point in $\mathrm{supp}(Q) \cap L(\boldsymbol{x})$ one by one and move its mass along the first coordinate upwards towards x_1, while moving the mass of points in $\mathrm{supp}(Q) \cap U(\boldsymbol{y})$ downwards towards y_1, all this without changing the barycenter. Stop when no mass can be moved further, namely when either all mass in $L(\boldsymbol{x})$ rests on points having first coordinate equal to x_1, or all mass in $U(\boldsymbol{y})$ rests on points having first coordinate y_1. Call the obtained probability measure Q_1. Repeating the same procedure with the other coordinates yields a probability measure Q_d with the property that there is an index set $I \subseteq \{1, \ldots, d\}$ such that for all $\boldsymbol{z} \in \mathrm{supp}(Q_d)$ it is either $z_i = x_i$ for all $i \in I$ (if $\boldsymbol{z} \in L(\boldsymbol{x})$) or $z_i = y_i$ for all $i \notin I$ (if $\boldsymbol{z} \in U(\boldsymbol{y})$).

In light of Lemma 4.1 the proof can be finished by showing that Q_d can be obtained from P^* via a sequence of ultramodular transfers. To do so it is sufficient to show that for a fixed $\boldsymbol{z} \in \mathrm{supp}(Q_d)$ a measure $P_{\boldsymbol{z}}$ can be obtained from P^* via a sequence of ultramodular transfers, where $P_{\boldsymbol{z}}$ is comonotone and has the same mass in $\boldsymbol{z}$ as Q_d and

$$\mathrm{supp}(P_{\boldsymbol{z}}) \subseteq \{\boldsymbol{z}\} \cup \bigtimes_{i=1}^{d} \{x_i, y_i\}.$$

The proof then follows by induction.

Without loss of generality assume $\boldsymbol{z} \in L(\boldsymbol{x})$ and distinguish two cases.

If $\delta := Q_d(\{\boldsymbol{z}\}) \leqslant P^*(\{\boldsymbol{x}\})$ then $P_{\boldsymbol{z}}$ can be obtained from P^* by moving the mass δ from the point $\boldsymbol{x}$ to the point $\boldsymbol{z}$ using a sequence of ultramodular transfers indexed over $j \notin I$ that move mass along the j-th coordinate from x_j to $z_j < x_j$ and at the same time move mass from some point $\boldsymbol{s}$ in the support of P^* with $s_j = x_j$ along the same coordinate from x_j to y_j. As a consequence the j-th marginal is transformed from the one of P^* (supported on x_j and y_j) to the one of $P_{\boldsymbol{z}}$ (supported on z_j, x_j and y_j). Once this is done some supermodular transfers within $\times_{i=1}^{d}\{x_i, y_i\}$ may be necessary to get the comonotone probability measure $P_{\boldsymbol{z}}$.

In the other case $\eta := P^*(\{\boldsymbol{x}\}) < Q_d(\{\boldsymbol{z}\})$. Then move all the mass η from the point $\boldsymbol{x}$ to the point $\boldsymbol{z}$ as above. Then continue moving mass from the smallest point $\boldsymbol{x}' \geqslant \boldsymbol{x}$ with $\boldsymbol{x}' \in \operatorname{supp}(P^*)$ to the point $\boldsymbol{z}$ in the same fashion. Iterate this as long as necessary to move mass δ to the point $\boldsymbol{z}$. Again, at the end of this procedure some supermodular transfers within $\times_{i=1}^{d}\{x_i, y_i\}$ may be necessary to obtain the comonotone probability measure $P_{\boldsymbol{z}}$. $\square$

The proof of Theorem 3.4 requires some known results from functional analysis, and some theory of discrete ultramodularity. These results are described in the next subsections.

4.2 Duality theory

For $S \subset \mathbb{R}^d$ compact, denote by $\mathscr{C}$ the set of a continuous functions on S. By the compactness assumption on S these functions are all bounded and therefore integrable with respect to any $\mu \in \mathbb{M}$.

Integrals are often written as a bilinear form $\langle f, \mu \rangle = \int f \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu^+ - \int f \, \mathrm{d}\mu^-$.

Some results from functional analysis are presented. The details can be found, e.g., in Choquet (1969, §22).

A pair (E, F) of vector spaces is said to be in *duality*, if there is a bilinear mapping $\langle \cdot, \cdot \rangle : E \times F \to \mathbb{R}$. The duality is said to be *strict*, if for each $0 \neq x \in E$ there is a $y \in F$ with $\langle x, y \rangle \neq 0$ and if for each $0 \neq y \in F$ there is an $x \in E$ with $\langle x, y \rangle \neq 0$.

Unfortunately the duality $(\mathbb{M}, \mathscr{C})$ is not strict, as $\langle f, \mu \rangle = 0$ for all $\mu \in \mathbb{M}$ only implies f to be constant. But strict duality can be obtained by identifying functions which differ only by a constant. Formally, define an equivalence relation $f \sim g$ if $f - g$ is constant. Equivalently, fix some $s_0 \in S$ and require $f(s_0) = 0$. With utility functions in mind, it is quite natural to identify functions that differ only by a constant, as they lead to the same preference relation. Denote the corresponding quotient space by $\mathscr{C}_\sim$.

Lemma 4.3. $\mathbb{M}$ *and* $\mathscr{C}_\sim$ *are in strict duality under the bilinear mapping*

$$\begin{aligned}\langle\cdot,\cdot\rangle &: \mathbb{M}\times\mathscr{C}_\sim \to \mathbb{R},\\ \langle\mu, f\rangle &= \int f \,\mathrm{d}\mu.\end{aligned}$$

A crucial role in our further investigations is played by the bipolar theorem for convex cones. The notion of polars is introduced following the notation of Choquet (1969).

The *polar* M° of a set $M \subset E$ (in the duality (E, F)) is defined as

$$M^\circ = \{y \in F : \langle x, y\rangle \geqslant -1 \text{ for all } x \in M\}. \tag{4.1}$$

The polar of a set $N \subset F$ is defined analogously.

Given a vector space V, a subset $K \subset V$ is called a *cone* if $x \in K$ implies $\alpha x \in K$ for all $\alpha \geqslant 0$. Given any subset $M \subset V$, the convex cone $\mathrm{co}(M)$ generated by M is the smallest convex cone that contains M.

Define the *dual cone* of an arbitrary set $M \subset E$ by

$$M^* = \{y \in F : \langle x, y\rangle \geqslant 0 \text{ for all } x \in M\}.$$

It is easy to see that M^* is a convex cone. Moreover, notice that for a convex cone K the polar and dual cones coincide: $K^\circ = K^*$.

For any duality (E, F) define the *weak topology* $\sigma(E, F)$ on E as the weakest topology on E such that the mappings $x \mapsto \langle x, y\rangle$ are continuous for all $y \in F$. Now the bipolar theorem for convex cones can be stated as follows (see Choquet, 1969, Corollary 22.10).

Theorem 4.4. *Suppose* E *and* F *are in strict duality and* $X \subset E$ *is an arbitrary set. Then* X^{**} *is the weak closure of the convex cone generated by* X*.*

4.3 Discrete inframodularity and concavity

Now consider the classes of functions of Definition 2.1 when their domain is a suitable finite set.

First recall the definition of discrete concavity for functions defined on a finite subset of the real line. Let $S = \{x_1, x_2, \ldots, x_n\} \subset \mathbb{R}$ be a finite set, where the elements are ordered, i.e. $x_1 < x_2 < \ldots < x_n$. For a function $f : S \to \mathbb{R}$ define the difference operator

$$\Delta f(x_i) := \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}, \quad x_i \in \{x_1, x_2, \ldots, x_{n-1}\}.$$

A function $f : S \to \mathbb{R}$ is said to be *discrete concave* if $x_i \mapsto \Delta f(x_i)$ is decreasing. This is equivalent to requiring that for any three consecutive points $x_i, x_{i+1}, x_{i+2} \in S$

$$f(x_{i+1}) \geqslant \alpha f(x_{i+2}) + (1-\alpha) f(x_i),$$

where $\alpha = (x_{i+1} - x_i)/(x_{i+2} - x_i)$, i.e., α is such that $x_{i+1} = \alpha x_{i+2} + (1-\alpha)x_i$.

This definition of discrete concavity is consistent with the usual definition of concavity for functions $f : \mathbb{R} \to \mathbb{R}$, as the following lemma shows.

Lemma 4.5. (i) *The restriction of any concave function $f : \mathbb{R} \to \mathbb{R}$ to the finite subset S is discrete concave.*

(ii) *Any discrete concave function $f : S \to \mathbb{R}$ can be extended to a concave function $f : \mathbb{R} \to \mathbb{R}$.*

Proof. Property (i) is obvious, and to show property (ii) one can use the linear interpolation in the intervals $[x_i, x_{i+1}]$, and outside of $[x_1, x_n]$ one can use the linear extension

$$f(x_n + t) = f(x_n) + \Delta f(x_{n-1})t \quad \text{and} \quad f(x_1 - t) = f(x_1) - \Delta f(x_1)t, \quad t > 0.$$

□

A similar definition of discrete inframodular function on a finite lattice $S \subset \mathbb{R}^d$ is now given. Assume that

$$S := \bigtimes_{i=1}^{d} S_i := \bigtimes_{i=1}^{d} \{x_{i,1}, \ldots, x_{i,n_i}\},$$

is a finite lattice, where, as before, the elements of S_i are ordered, i.e., $x_{i,1} < \ldots < x_{i,n_i}$. Define the difference operator in direction i computed at point $\boldsymbol{x} = (x_{1,k_1}, \ldots, x_{d,k_d})$ as

$$\Delta_i f(\boldsymbol{x}) := \frac{f(x_{1,k_1}, \ldots, x_{i-1,k_{i-1}}, x_{i,k_i+1}, x_{i+1,k_{i+1}}, \ldots, x_{d,k_d}) - f(\boldsymbol{x})}{x_{i,k_i+1} - x_{i,k_i}}, \quad \boldsymbol{x} \in S,\ k_i < n_i.$$

The function $f : S \to \mathbb{R}$ is *discrete component-wise concave*, if $x_i \mapsto \Delta_i f(x_1, \ldots, x_i, \ldots, x_d)$ is decreasing for all $i = 1, \ldots, d$, for any fixed $x_j \in S_j, j \neq i$. As in the univariate case this is equivalent to requiring for any three consecutive points $x_{i,k_i}, x_{i,k_i+1}, x_{i,k_i+2}$ and for any fixed $x_j \in S_j,\ j \neq i$ that

$$f(x_1, \ldots, x_{i,k_i+1}, \ldots, x_d) \geqslant \alpha f(x_1, \ldots, x_{i,k_i+2}, \ldots, x_d) + (1-\alpha) f(x_1, \ldots, x_{i,k_i}, \ldots, x_d),$$

where $\alpha = (x_{i,k_i+1} - x_{i,k_i})/(x_{i,k_i+2} - x_{i,k_i})$, i.e., α is such that $x_{i,k_i+1} = \alpha x_{i,k_i+2} + (1-\alpha)x_{i,k_i}$.

The classical definition of submodularity is valid on any lattice and is equivalent to the requirement that

$$\Delta_i \Delta_j f(\boldsymbol{x}) \leqslant 0 \quad \text{for all } i \neq j, \text{ and all } (x_{1,k_1}, \ldots, x_{d,k_d}) \in S$$

with $k_i < n_i, k_j < n_j$.

Lemma 4.6. *The following conditions are equivalent:*

(a) *The function f is inframodular.*

(b) $\Delta_i f(\boldsymbol{x})$ *is a decreasing function of* $\boldsymbol{x}$.

(c) *The function $f : S \to \mathbb{R}$ is submodular and component-wise concave.*

For the proof of the above lemma see, e.g. Marinacci and Montrucchio (2005). The following consistency result holds.

Lemma 4.7. (a) *if $f : \mathbb{R}^d \to \mathbb{R}$ is inframodular, then its restriction to S is discrete inframodular.*

(b) *any discrete inframodular function $f : S \to \mathbb{R}$ can be extended to an inframodular function $f : \mathbb{R}^d \to \mathbb{R}$.*

Proof. Part (a) is obvious. For (b) the extension has to be defined. Between grid points this is done in a component-wise linear fashion. For $\boldsymbol{z} \in \times_{i=1}^{d}[x_{i,k_i}, x_{i,k_{i+1}}]$ write the coordinates as $z_i = \alpha_i x_{i,k_{i+1}} + (1-\alpha_i)x_{i,k_i}$, $i = 1, \dots, d$. The coordinate-wise linear extension can then be defined iteratively, starting with

$$f(z_1, x_{2,k_2}, \dots, x_{d,k_d}) := \alpha_1 f(x_{1,k_1+1}, x_{2,k_2}, \dots, x_{d,k_d}) + (1-\alpha_1) f(x_{1,k_1}, x_{2,k_2}, \dots, x_{d,k_d}).$$

In step i define

$$\begin{aligned} f(z_1, \dots, z_{i-1}, z_i, x_{i+1,k_{i+1}}, \dots, x_{d,k_d}) := \;& \alpha_i f(z_1, \dots, z_{i-1}, x_{i,k_i+1}, \dots, x_{d,k_d}) \\ & + (1-\alpha_i) f(z_1, \dots, z_{i-1}, x_{i,k_i}, \dots, x_{d,k_d}). \end{aligned}$$

Thus a piecewise linear extension of $f : S \to \mathbb{R}$ to $\operatorname{conv}(S) = \times_{i=1}^{d}[x_{i1}, x_{in_i}]$ has been obtained. It is straightforward to see that this piecewise linear extension is component-wise concave and submodular (this construction is similar to the extension of a subcopula to a copula: see Schweizer and Sklar, 1983). Outside of $\operatorname{conv}(S)$ extend the function by component-wise linear extrapolation as in Müller and Scarsini (2001, proof of Theorem 2.7), which leads to a function that is inframodular on the entire $\mathbb{R}^d$. $\square$

The two properties of Lemma 4.7 together imply that the set of discrete inframodular functions $f : S \to \mathbb{R}$ is equivalent to the set of restrictions of inframodular functions $f : \mathbb{R}^d \to \mathbb{R}$ to S, if S is a finite lattice. It follows therefore that for probability measures P and Q with finite support in $\mathbb{R}^d$ the following statements are equivalent:

(i) $\int f \, \mathrm{d}P \leqslant \int f \, \mathrm{d}Q$ for all inframodular functions $f : \mathbb{R} \to \mathbb{R}$;

(ii) $\int f \, \mathrm{d}P \leqslant \int f \, \mathrm{d}Q$ for all discrete inframodular functions $f : S \to \mathbb{R}$, where S is the smallest finite lattice containing the supports of P and Q.

4.4 Stochastic orders and transfers

Using the properties of the two previous subsections Theorem 3.4 can now be proved.

Proof of Theorem 3.4. The equivalence of (b) and (c) follows from Theorem 3.1, and it is clear that (b) implies (a). Thus it remains to show that (a) implies (b). Hence assume that P and Q are probability measures on $\mathbb{R}^d$ with finite support fulfilling $\int f \, \mathrm{d}P \leqslant \int f \, \mathrm{d}Q$ for all inframodular functions $f : \mathbb{R}^d \to \mathbb{R}$. It follows from Lemma 4.7 that this is equivalent to the statement that $\int f \, \mathrm{d}P \leqslant \int f \, \mathrm{d}Q$ for all inframodular functions $f : S \to \mathbb{R}$, where S is the smallest lattice containing the supports of P and Q. Using the terminology of duality theory as described in Subsection 4.2 the condition can be rewritten as $Q - P \in \mathscr{F}^*$, where $\mathscr{F}$ is the set of all inframodular functions $f : S \to \mathbb{R}$. The fact that $\mathscr{F}$ is induced by the set M of inframodular transfers can be rewritten as $\mathscr{F} = M^*$, thus $Q - P \in M^{**}$.

Therefore it follows from Theorem 4.4 that $Q - P$ is in the weak closure of the convex cone generated by M. As S is finite, the set M of inframodular transfers on this set is also finite, and therefore the convex cone generated by M is weakly closed. Thus $Q - P = \sum_{i=1}^{n} \gamma_i \mu_i$ with $\gamma_i > 0$ and $\mu_i \in M$. As P and Q are probability measures, it is possible to choose $\gamma_i \leqslant 1$. But this means that Q can be obtained from P by a finite number of inframodular transfers $\gamma_i \mu_i$. $\square$

Acknowledgment

The authors thank Alessandro Arlotto for interesting discussions and help with the bibliography.

References

Arlotto, A. and Scarsini, M. (2009) Hessian orders and multinormal distributions. *J. Multivariate Anal.* **100**, 2324–2330.

Arrow, K. J. (1970) *Essays in the Theory of Risk-Bearing.* North-Holland Publishing Co., Amsterdam.

Blackwell, D. (1951) Comparison of experiments. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, 93–102. University of California Press, Berkeley and Los Angeles.

Blackwell, D. (1953) Equivalent comparisons of experiments. *Ann. Math. Statistics* **24**, 265–272.

Border, K. C. (1991) Functional analytic tools for expected utility theory. In *Positive Operators, Riesz Spaces, and Economics (Pasadena, CA, 1990)*, 69–88. Springer, Berlin.

Brumelle, S. and Vickson, R. (1975) A unified approach to stochastic dominance. In *Stochastic Optimization Models in Finance*, 101–113. Academic Press, New York.

Cartier, P., Fell, J. M. G., and Meyer, P.-A. (1964) Comparaison des mesures portées par un ensemble convexe compact. *Bull. Soc. Math. France* **92**, 435–445.

Castagnoli, E. and LiCalzi, M. (1997) Direct and indirect duality for dominance relations. In *Decisions, Games, and Markets*, 3–30. Kluwer, Boston, MA.

Castagnoli, E. and Maccheroni, F. (2000) Restricting independence to convex cones. *J. Math. Econom.* **34**, 215–223.

Choquet, G. (1969) *Lectures on Analysis. Vol. II: Representation Theory.* Edited by J. Marsden, T. Lance and S. Gelbart. W. A. Benjamin, Inc., New York-Amsterdam.

Denuit, M., Eeckhoudt, L., and Rey, B. (2010) Some consequences of correlation aversion in decision science. *Annals of Operations Research* Forthcoming.

Dubra, J., Maccheroni, F., and Ok, E. A. (2004) Expected utility theory without the completeness axiom. *J. Econom. Theory* **115**, 118–133.

Duncan, G. T. (1977) A matrix measure of multivariate local risk aversion. *Econometrica* **45**, 895–903.

Eeckhoudt, L., Rey, B., and Schlesinger, H. (2007) A good sign for multivariate risk taking. *Management Sci.* **53**, 117–124.

Elton, J. and Hill, T. P. (1992) Fusions of a probability distribution. *Ann. Probab.* **20**, 421–454.

Elton, J. and Hill, T. P. (1998) On the basic representation theorem for convex domination of measures. *J. Math. Anal. Appl.* **228**, 449–466.

Epstein, L. G. and Tanny, S. M. (1980) Increasing generalized correlation: a definition and some economic consequences. *Canad. J. Econom.* **13**, 16–34.

de Finetti, B. (1952) Sulla preferibilità. *Giorn. Econ. Ann. Econ.* **11**, 685–709.

Hardy, G., Littlewood, J., and Polya, G. (1929) Some simple inequalities satisfied by convex functions. *Messenger Math.* **58**, 310.

Hardy, G. H., Littlewood, J. E., and Pólya, G. (1988) *Inequalities.* Cambridge Mathematical Library. Cambridge University Press, Cambridge. Reprint of the 1952 edition.

JOE, H. (1997) *Multivariate Models and Dependence Concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

KARNI, E. (1979) On multivariate risk aversion. *Econometrica* **47**, 1391–1401.

KIHLSTROM, R. E. and MIRMAN, L. J. (1974) Risk aversion with many commodities. *J. Econom. Theory* **8**, 361–388.

KIHLSTROM, R. E. and MIRMAN, L. J. (1981) Constant, increasing and descreasing risk aversion with many commodities. *Rev. Econom. Stud.* **48**, 271–280.

MACHINA, M. and PRATT, J. W. (1997) Increasing risk: some direct constructions. *Journal of Risk and Uncertainty* **14**, 103–27.

MARINACCI, M. and MONTRUCCHIO, L. (2005) Ultramodular functions. *Math. Oper. Res.* **30**, 311–332.

MÜLLER, A. (1996) Orderings of risks: a comparative study via stop-loss transforms. *Insurance Math. Econom.* **17**, 215–222.

MÜLLER, A. and SCARSINI, M. (2000) Some remarks on the supermodular order. *J. Multivariate Anal.* **73**, 107–119.

MÜLLER, A. and SCARSINI, M. (2001) Stochastic comparison of random vectors with a common copula. *Math. Oper. Res.* **26**, 723–740.

MÜLLER, A. and STOYAN, D. (2002) *Comparison Methods for Stochastic Models and Risks*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester.

PRATT, J. W. (1964) Risk aversion in the small and the large. *Econometrica* **32**, 122–136.

RICHARD, S. F. (1975) Multivariate risk aversion, utility independence and separable utility functions. *Management Sci.* **22**, 12–21.

ROTHSCHILD, M. and STIGLITZ, J. E. (1970) Increasing risk. I. A definition. *J. Econom. Theory* **2**, 225–243.

ROTHSCHILD, M. and STIGLITZ, J. E. (1971) Increasing risk. II. Its economic consequences. *J. Econom. Theory* **3**, 66–84.

ROTHSCHILD, M. and STIGLITZ, J. E. (1972) Addendum to: "Increasing risk. I. A definition". *J. Econom. Theory* **5**, 306.

SCARSINI, M. (1994) Comparing risk and risk aversion. In SHAKED, M. and SHANTHIKUMAR, J. G. (eds.), *Stochastic Orders and their Applications*, 351–378. Academic Press, Boston, MA.

Schweizer, B. and Sklar, A. (1983) *Probabilistic Metric Spaces*. North-Holland Series in Probability and Applied Mathematics. North-Holland Publishing Co., New York.

Shaked, M. and Shanthikumar, J. G. (2007) *Stochastic Orders*. Springer Series in Statistics. Springer, New York.

Sherman, S. (1951) On a theorem of Hardy, Littlewood, Polya, and Blackwell. *Proc. Nat. Acad. Sci. U. S. A.* **37**, 826–831; errata: 38, 382 (1952).

Strassen, V. (1965) The existence of probability measures with given marginals. *Ann. Math. Statist.* **36**, 423–439.

Tchen, A. H. (1980) Inequalities for distributions with given marginals. *Ann. Probab.* **8**, 814–827.

Topkis, D. M. (1998) *Supermodularity and Complementarity*. Frontiers of Economic Research. Princeton University Press, Princeton, NJ.

Tsetlin, I. and Winkler, R. L. (2009) Multiattribute utility satisfying a preference for combining good with bad. *Management Sci.* **55**, 1942–1952.

Ziegler, Z. (1968) On the characterization of measures of the cone dual to a generalized convexity cone. *Pacific J. Math.* **24**, 603–626.

5 Figures

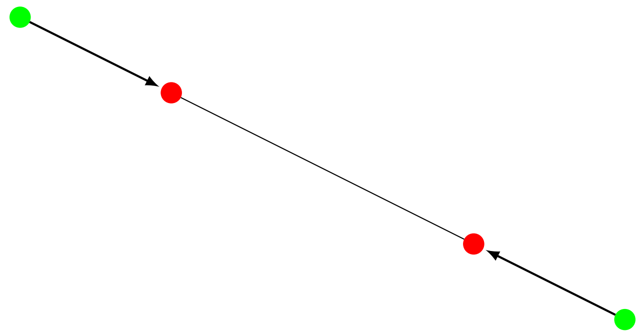

Figure 1: Simple symmetric concave transfer

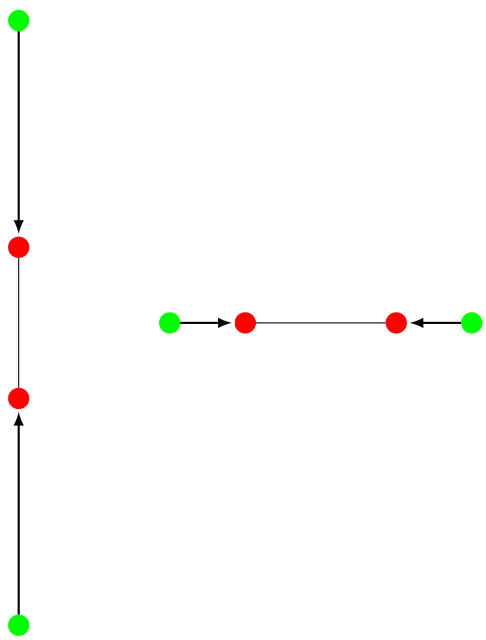

Figure 2: Simple symmetric component-wise concave transfers

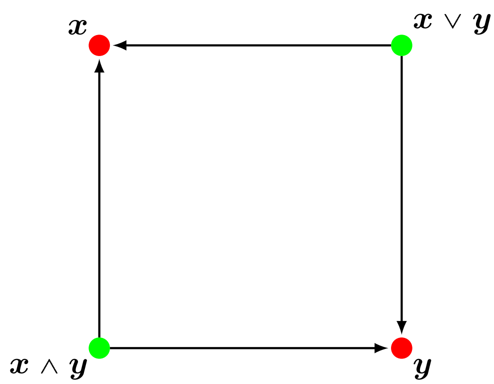

Figure 3: Simple submodular transfer

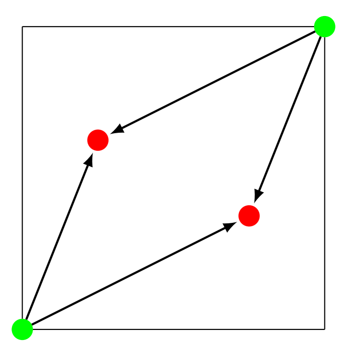

Figure 4: Simple symmetric inframodular transfer

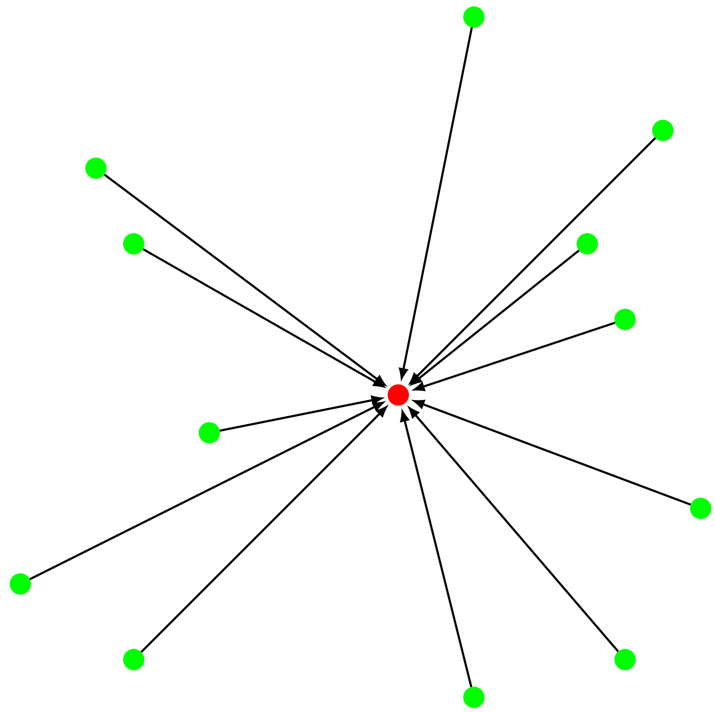

Figure 5: General concave transfer (fusion)

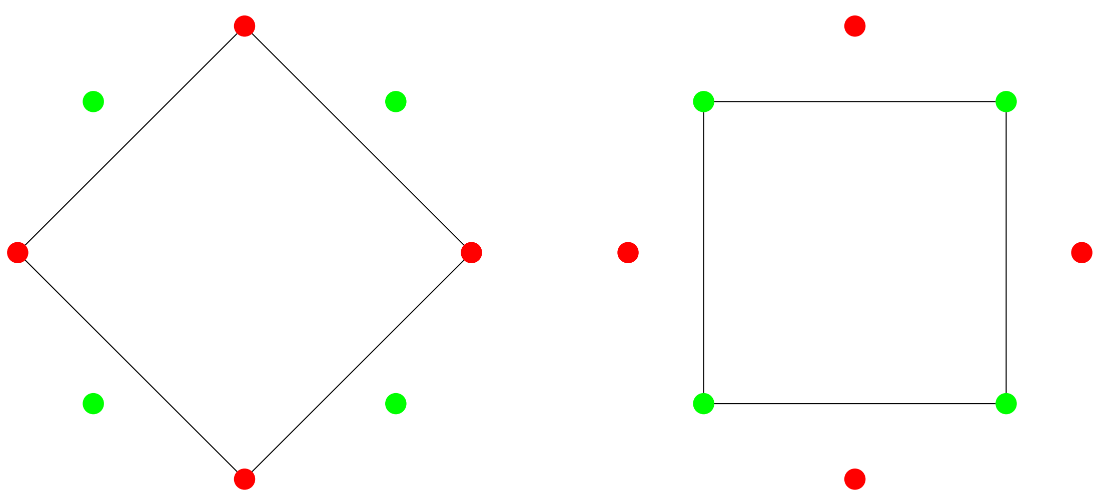

Figure 6: These are not convex transfers

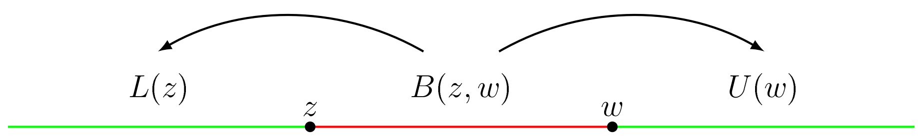

Figure 7: General mean preserving spread

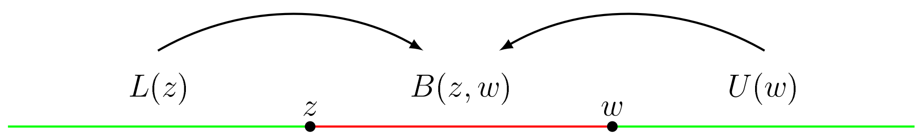

Figure 8: General mean preserving contraction

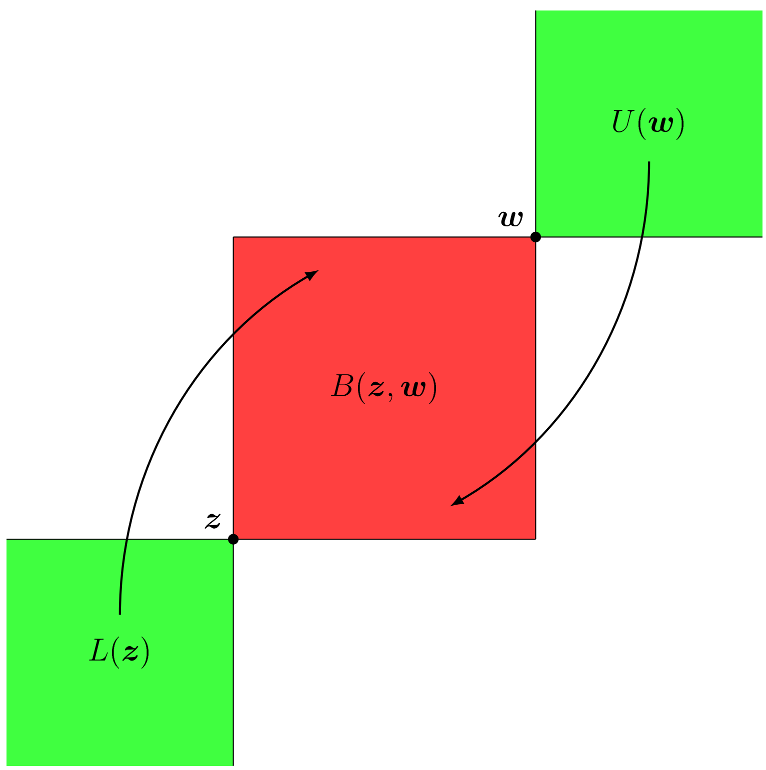

Figure 9: General inframodular transfer