Discussion Paper Series – CRC TR 224

Discussion Paper No. 567
Project B 02

Monotone Decision Rules and Supermodularity

Gregorio Curello[1]

June 2024

[1]University of Bonn, Email: gcurello@uni-bonn.de

Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Collaborative Research Center Transregio 224 - www.crctr224.de
Rheinische Friedrich-Wilhelms-Universität Bonn - Universität Mannheim

Monotone Decision Rules and Supermodularity*

Gregorio Curello†

June 13, 2024

Abstract

We study decision problems under uncertainty involving the choice of a rule mapping states into actions. We show that for any rule, there exists an increasing rule generating higher expected value for *all* payoff functions that are supermodular in action and state. We present applications to problems of taxation, betting, and price-discrimination in markets with demand externalities. We then consider rules mapping noisy signals of the state into actions. Under some conditions, optimal rules are increasing when (a) several agents are constrained to choose a single rule or (b) the relationship between signal and state is ambiguous. Moreover, standard informativeness criteria apply.

Keywords: monotone comparative statics, rearrangement, optimal taxation, price discrimination, uncertainty, informativeness.

JEL Classification: C61, D71, D81.

1 Introduction

Many decision problems entail the choice of a decision rule; that is, a map from observables (or states) into actions. We seek general conditions on the primitives of the problem ensuring that some increasing decision rule is optimal.

Sufficient conditions are known for problems that are separable across states. Suppose, for instance, that the state y has distribution F with support $Y \subseteq \mathbb{R}$, and the decision maker earns payoff $u(x, y)$ from action $x \in X \subseteq \mathbb{R}$ when the state is $y \in Y$. Assuming expected-utility preferences, the decision maker picks a map

*I thank John Quah, Péter Eső, Meg Meyer, Bruno Strulovici, and the seminar participants at the University of Oxford for helpful comments. Support by the German Research Foundation (DFG) through CRC TR 224 (Project B02) is gratefully acknowledged.

†Institute for Microeconomics, University of Bonn. E-mail: gcurello@uni-bonn.de.

$\psi : Y \to X$ in order to maximise

$$\int_Y u(\psi(y), y)\mathrm{d}F(y).^{1} \tag{1}$$

The main result of Milgrom and Shannon (1994) implies that if u has single-crossing differences in (x, y) then, under mild assumptions, for any rule ψ there exists an increasing rule ϕ such that $u(\phi(y), y) \geq u(\psi(y), y)$ for all $y \in Y$.[2] In particular,

$$\int_Y u(\phi(y), y)\mathrm{d}F(y) \geq \int_Y u(\psi(y), y)\mathrm{d}F(y) \tag{2}$$

so the decision maker prefers ϕ to ψ. Then, (1) is maximised by an increasing ψ.

Many decision problems, however, do not amount to the maximisation of an objective of the form given in (1). This could be due one or more constraints delimiting the set of admissible rules (Applications 1 and 4), to externalities (Application 3), or non-expected-utility preferences (Application 5). Yet, as the applications below show, we can often conclude that the decision maker prefers a rule ϕ to a rule ψ if (2) holds *simultaneously* for all functions u within a sufficiently large set $\mathcal{U}$. If this is the case, we say that ϕ $\mathcal{U}$-dominates ψ. Thus, in these cases, some increasing rule is optimal as long as any rule ψ is $\mathcal{U}$-dominated by some increasing rule ϕ.

Under what conditions on $\mathcal{U}$ can one find, for any rule ψ, an increasing rule ϕ that $\mathcal{U}$-dominates ψ? Our main result is that, if the decision maker is allowed to take a random action x after observing y, then this is the case if $\mathcal{U}$ is the set of all supermodular functions (denoted $\mathcal{U}_{\text{SPM}}$).[3] The rule ϕ is constructed by rearranging ψ without altering the ex-ante distribution of actions induced by F; we call ϕ the *increasing rearrangement* of ψ *with respect to* F.[4] Thus, for example, a decision maker with objective (1) who does not know the functional form of her payoffs u – merely that they are supermodular, can improve on any given non-increasing rule ψ, as long as she knows F. Conversely, if $\mathcal{U}$ is any set of payoffs u such that, for any F, any rule ψ is $\mathcal{U}$-dominated by its increasing rearrangement with respect to F, then all elements of $\mathcal{U}$ must be supermodular.

The result that any decision rule is $\mathcal{U}_{\text{SPM}}$-dominated by its increasing rearrangement has many interesting applications. We present five problems in which relatively weak assumptions ensure that optimal rules are monotone.

Applications

1. A social planner wishes to implement a subsidy based on a household-specific

1. An optimal ψ exists if X is compact, $u : X \times Y \to \mathbb{R}$ is measurable, $u(\cdot, y)$ is bounded for all $y \in Y$, and $\sup_{x \in X} u(x, \cdot)$ is F-integrable. These assumptions will be maintained throughout.

2. It is sufficient that $\arg\max_{x \in X} u(x, y)$ in non-empty and has a maximal element for all $y \in Y$.

3. A random rule ψ is increasing if the support of $\psi(y')$ lies above that of $\psi(y)$ whenever $y' > y$. If randomisation is not allowed, then any rule ψ is $\mathcal{U}_{\text{SPM}}$-dominated by some increasing rule ϕ if and only if F is a continuous distribution or a discrete uniform distribution.

4. If X is multi-dimensional, the components of ψ are rearranged in such a way that, in each state, the action taken is *comonotone*; that is, any two elements of its support are ordered component-wise.

measure of need for financial support, subject to a budget constraint. If the marginal utility of the transfer is increasing in need, the planner may restrict attention rules that transfer more resources to households who are more in need. Importantly, the planner may improve on any other rule without knowing the quantitative aspects of the households' utility (which are hardly ever known in practice), if she knows the distribution of need in the population.

2. Two agents place monetary bets over a random outcome. If one of the agents is more optimistic about the outcome in a MLR-sense then, under any Pareto optimal bet, higher outcomes induce larger transfers to her. Moreover, this holds whether agents maximise expected utility or *cautious* expected utility.

3. A monopolist price-discriminates based on buyers' wealth levels in a market with a positive demand externality. If production costs are negligible, demand is normal and wealthier buyers are less price-sensitive, then a profit-maximising seller would charge them more. In particular, shifting the weight of high prices on wealthier buyers without altering the overall distribution of prices increases profits. The reason is twofold: first, since wealthier buyers are less price-sensitive, this boosts the demand externality; second, since the good is normal, this increases the average price per unit sold.

4. An insurer sells a single product to a group of buyers with varying risk-aversion who are subject to different shocks. If she can only condition insurance payments on a noisy signal of the buyers' shocks, she should offer an insurance that pays more when signals indicating larger shocks are observed. Moreover, if the insurer gains access to a signal that is more accurate in Lehmann's sense, she may achieve higher profits whilst making all agents better-off.

5. A firm has access to a noisy signal of its marginal cost, which is unobservable. Before observing the signal, the firm commits to a signal-contingent level of output. The firm is ambiguity-averse, and does not know the joint distribution of signal and marginal cost – it merely knows the marginal distribution of the signal. We show that the firm should produce more after observing signals that indicate lower costs, and derive a condition on the structure of the ambiguity enabling the firm to achieve higher ex-ante utility.

A number of authors studied other problems in economics using techniques similar to those we draw on here. Landsberger and Meilijson (1994) show that for any non-comonotone allocation of a random endowment to a group of risk-averse agents, there exists a comonotone allocation raising the expected utility of all agents, regardless of their specific utility functions. Arieli and Babichenko (2019) consider a problem of Bayesian persuasion in which a single sender sends private messages to multiple receivers. They show that, if the state is binary, the same signalling policy is optimal for any sender with a payoff function that is supermodular in the receivers' actions. Moreover, the signals are increasing in the state. Meyer (2018) obtains a general result that applies to constrained problems in which the marginal distributions for each component of the (multi-dimensional) action are predetermined. She links the degree of heterogeneity among these marginals to the value of the (unique) increasing rule when payoffs are supermodular and symmetric. Her result applies to the setting studied by Arieli and Babichenko (2019); she shows that heterogeneity in the persuadability of receivers is harmful to the sender. In the context of statistical estimation, Chernozhukov, Fernandez-Val, and Galichon (2009) show that the monotone rearrangement of an estimate of a monotone function improves upon the original estimate whenever the latter is not monotone. Similarly, rearranging the upper and lower bound of a confidence band for monotone functions reduces its length and increases its coverage probability.

The rest of the paper is divided in two sections. Section 2 contains the theoretical results, which apply to a general decision environment. Applications $1-3$ are discussed in Sections $3.1-3.3$, respectively. Sections 3.4 and 3.5 contain specific results about choice under uncertainty and informativeness, respectively. Application 4 (resp. 5) is discussed in Example 1 (resp. 2) of both sections.

2 Theory

This section contains the theoretical results, expressed in their most general form. A series of applications are presented in Section 3.

Consider a decision maker who takes an action x after observing the state of the world y. Suppose that y is distributed according to some (cumulative) distribution function $F : \mathbb{R} \to [0,1]$ with support $Y \subseteq \mathbb{R}$, and that, for some $n \in \mathbb{N}$, $x \in X = X_1 \times \cdots \times X_n$ where $X_i \subset \mathbb{R}$ is compact for all $i \leq n$. We allow the decision maker to randomise over X after observing any signal $y \in Y$. A *(random) decision rule* is a collection $\psi = \{\psi(\cdot|y)\}_{y\in Y}$ of distribution functions $\psi(\cdot|y) : \mathbb{R}^n \to [0,1]$ with support within X, indexed by $y \in Y$. The value $\psi(x|y)$ is the probability that the decision maker takes an action lower or equal to x (in the component-wise order) after observing state y; ψ is *deterministic* if, for all $y \in Y$, $\psi(\cdot|y)$ is degenerate (that is, it assigns probability 1 to some action $x \in X$). In this case, denote by $\psi(y)$ the action taken after observing state y.[5] Let D_F be the set of decision rules, and let $\hat{D}_F \subset D_F$ be the set of deterministic decision rules. Throughout the discussion, we regard any two decision rules ψ and ψ' such that $\psi(\cdot|y) = \psi'(\cdot|y)$ for F-almost all y's as being essentially the same rule.[6]

5. If Y is uncountable, we require that $\psi(x|\cdot)$ is a measurable function over Y for any $x \in X$. This implies that if ψ is deterministic, it may be described by a measurable map $\psi : Y \to X$. Throughout the discussion, we assume that all functions are measurable.

6. Formally, D_F is the quotient space of collections $\{\psi(\cdot|y)\}_{y\in Y}$ by the equivalence relation $\sim$ whereby $\psi \sim \psi'$ if and only if $\psi(\cdot|y) = \psi'(\cdot|y)$ for F-almost all y's. $\hat{D}_F$ is the set of equivalence classes in D_F that contain a deterministic rule.

Definition 1. A decision rule ψ is *increasing* if, for any $(x, y) \in \mathbb{R}^n \times Y$ the following holds. If $\psi(x|y) < 1$, for any $y' > y$, we obtain $\psi(x|y') = 0$.

Intuitively, ψ is increasing if the support of $\psi(\cdot|y')$ lies above that of $\psi(\cdot|y)$ whenever $y' > y$. Note that, if ψ is deterministic, the definition reduces to the requirement that ψ is a (weakly) increasing function from Y to X.

An n-dimensional distribution G is *comonotone* if, for any two points $z, z' \in \mathbb{R}^n$ within its support, either $z \geq z'$ or $z' \geq z$. Equivalently, G is comonotone if and only if $G(z) = \min_{i \leq n} G_i(z_i)$ where G_i is the i^{th} one-dimensional marginal distribution of G for all $i \leq n$.[7] A decision rule $\psi \in D_F$ is comonotone if $\psi(\cdot|y)$ is a comonotone distribution for all $y \in Y$.[8] Note that any deterministic decision rule is comonotone and, if $n = 1$, then any rule is comonotone.

Given $\psi \in D_F$ and $i \leq n$, let $\psi_i = \{\psi_i(\cdot|y)\}_{y \in Y}$ be such that $\psi_i(\cdot|y)$ is the marginal distribution of $\psi(\cdot|y)$ over X_i. Intuitively, ψ_i is the 'projection' of ψ onto X_i.[9] Any rule $\psi \in D_F$ induces a distribution of action components $x_i \in X_i$ given by

$$M_{\psi_i F}(x_i) = \int_Y \psi_i(x_i|y) \mathrm{d}F(y). \tag{3}$$

Definition 2. A *rearrangement of* a rule $\psi \in D_F$ with respect to F is a rule $\phi \in D_F$ such that $M_{\phi_i F} = M_{\psi_i F}$ for all $i \leq n$.

Lemma 1 below shows that any rule ψ admits a unique increasing and comonotone rearragement with respect to F, and gives conditions on F and ψ ensuring that the rearrangement is deterministic. The distribution F is *continuous* whenever $F : \mathbb{R} \to [0, 1]$ is a continuous function; F is a *discrete uniform* distribution if it puts mass $1/m$ on each of m points for some $m \in \mathbb{N}$.

Lemma 1. *For any $\psi \in D_F$, there exists a unique increasing and comonotone $\phi \in D_F$ that is a rearrangement of ψ with respect to F. Moreover, $\phi \in \hat{D}_F$ if (a) F is continuous, or (b) $\psi \in \hat{D}_F$ and F is a discrete uniform distribution. If $\phi \in \hat{D}_F$ then ϕ is the unique increasing rearrangement of ψ in $\hat{D}_F$.*

Proof of Lemma 1. We first show that ϕ exists. We start by defining the marginal distributions $\phi_i(\cdot|y)$ for all $i \leq n$ and $y \in Y$. Fix $y \in Y$. If F is continuous at y, then let

$$\phi_i(x_i|y) = \begin{cases} 0 & \text{if } M_{\psi_i F}(x_i) < F(y) \\ 1 & \text{if } M_{\psi_i F}(x_i) \geq F(y). \end{cases}$$

7. See Dhaene et al. (2002) for a theoretical analysis of comonotonicity and a proof of this equivalence.

8. In view of the definition of D_F (c.f. footnote 6), if ψ is increasing (resp. comonotone) then any ψ' such that $\psi'(\cdot|y) = \psi(\cdot|y)$ for F-almost all y's is considered increasing (resp. comonotone), too.

9. If ψ is deterministic (and, as a consequence, may be viewed as a map $\psi : Y \to X$) then ψ_i may be viewed as the function mapping $y \in Y$ to the i^{th} coordinate of $\psi(y)$.

If instead F has an atom of mass $p > 0$ at y, let

$$\phi_i(x_i|y) = \begin{cases} 0 & \text{if } M_{\psi_i F}(x_i) < F(y) - p \\ \frac{1}{p}[M_{\psi_i F}(x_i) - (F(y) - p)] & \text{if } F(y) - p \leq M_{\psi_i F}(x_i) < F(y) \\ 1 & \text{if } F(y) \leq M_{\psi_i F}(x_i). \end{cases}$$

Since $M_{\psi_i F}(x_i)$ is a distribution function and X_i is compact, then $\phi_i(\cdot|y)$ is a well-defined distribution for all $y \in Y$.[10] Let $\phi \in D_F$ be given by

$$\phi(x|y) = \min_{i \leq n} \phi_i(x_i|y) \quad \text{for all} \quad x \in \mathbb{R}^n, y \in Y.$$

Then, ϕ is a well-defined decision rule, and it is comonotone.

To complete the proof of existence, it remains to show that ϕ is increasing and that it is a rearrangement of ψ. To show that ϕ is increasing, fix $y \in Y$ and $x \in \mathbb{R}^n$ such that $\phi(x|y) < 1$. Then $\phi_i(x_i|y) < 1$ for some $i \leq n$. This implies that $F(y) > M_{\psi_i F}(x_i)$. Then $\phi_i(x_i|y') = 0$ for any $y' > y$. If F does not have an atom at y', this holds since $F(y') \geq F(y) > M_{\psi_i F}(x_i)$. If F puts mass $p > 0$ on y' this holds since $M_{\psi_i F}(x_i) - F(y') \leq F(y) - F(y') \leq -p$. Hence $\phi(x|y') = 0$. Therefore ϕ is increasing.

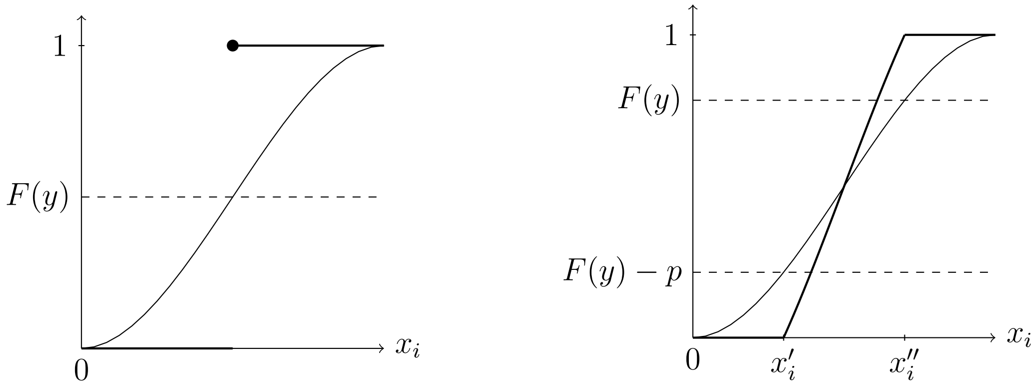
To show that ϕ is a rearrangement of ψ, fix $i \leq n$ and $x_i \in \mathbb{R}$. If $M_{\psi_i F}(x_i) = 0$, then $\phi_i(x_i|y) = 0$ for all $y \in Y$ such that $F(y) > 0$, so that $M_{\phi_i F}(x_i) = \int_Y \phi_i(x_i|y)\mathrm{d}F(y) = 0 = M_{\psi_i F}(x_i)$. Otherwise, let $y_0 = \inf\{y \in Y : M_{\psi_i F}(x_i) \leq F(y)\}$. Then, $\phi_i(x_i|y) = 1$ for $y < y_0$. Moreover, $\phi_i(x_i|y) = 0$ for any y such that $F(y) > F(y_0)$. Hence $\phi_i(x_i|y) = 0$ for F-almost all y's such that $y > y_0$. Let $p \geq 0$ be the probability mass that F puts on y_0. Then $M_{\psi_i F}(x_i) \in [F(y_0) - p, F(y_0)]$. Therefore $M_{\phi_i F}(x_i) = \int_Y \phi_i(x_i|y)\mathrm{d}F(y) = F(y_0) - p + p\phi_i(x_i|y_0) = M_{\psi_i F}(x_i)$. Since i and $x_i \in \mathbb{R}$ are arbitrary, ϕ is a rearrangement of ψ with respect to F. This proves existence. See the Appendix for a proof that ϕ is unique.

If F is continuous, it is clear from the definition of ϕ that $\phi \in \hat{D}_F$. Suppose that F is a discrete uniform distribution. Label $y_1 < \cdots < y_m$ the elements of Y. Then, to prove that $\phi \in \hat{D}_F$, it suffices to show that $\phi_i(\cdot|y_j)$ is degenerate for all $i \leq n$ and $j \leq m$. Fix i and j, and note that $\phi_i(x_i|y_j) = (1 + (mM_{\psi_i F}(x_i) - j) \wedge 0) \vee 0$. Moreover, since $\psi \in \hat{D}_F$, the image of $M_{\psi_i F}$ is within $\{0, 1/m, \ldots, (m-1)/m, 1\}$. It follows that $\phi_i(\cdot|y_j)$ is degenerate with value $\min\{x_i \in X_i : j/m \leq M_{\psi_i F}(x_i)\}$.

To prove the last part, note that any deterministic rule is comonotone. Therefore the uniqueness of the increasing rearrangement follows from the uniqueness of the increasing and comonotone rearrangement. □

The construction of $\phi_i(\cdot|y)$ is illustrated in Figure 1. If F is continuous at $y \in Y$, then $\phi_i(\cdot|y)$ is degenerate and puts all mass on the smallest x_i such that $M_{\psi F}(x_i)$

10. That is, $\phi_i(\cdot|y)$ is increasing, right-continuous and satisfies $\lim_{x_i \to -\infty} \phi_i(x_i|y) = 0$ and $\lim_{x_i \to \infty} \phi_i(x_i|y) = 1$.



(a) F is continuous at y (b) F has an atom of size p at y

Figure 1: $\phi_i(\cdot|y)$ (thick line) and $M_{\psi_i F}$ (thin line) as functions of x_i.

lies above $F(y)$. If F has an atom of size $p > 0$ at y, then the support of $\phi_i(\cdot|y)$ lies within $[x_i', x_i'']$ where x_i' (resp. x_i'') is the smallest x_i such that $M_{\psi_i F}(x_i)$ lies above $F(y) - p$ (resp. $F(y)$). More specifically, $\phi_i(\cdot|y)$ is an affine transformation of $M_{\psi_i F}$ on $[x_i', x_i'']$.

Denote the i^{th} entry of a vector $z \in \mathbb{R}^m$ by z_i. Given $z, z' \in \mathbb{R}^m$, let $z \wedge z' = (\min\{z_1, z_1'\}, \ldots, \min\{z_m, z_m'\})$ and $z \vee z' = (\max\{z_1, z_1'\}, \ldots, \max\{z_m, z_m'\})$.

Definition 3. A function $u : \mathbb{R}^m \to \mathbb{R}$ is *supermodular* if for any $z, z' \in \mathbb{R}^m$

$$u(z \wedge z') + u(z \vee z') \geq u(z) + u(z'); \tag{4}$$

u is *strictly supermodular* if (4) is strict unless $\{z \wedge z', z \vee z'\} = \{z, z'\}$; v is *(strictly) submodular* if $-v$ is (strictly) supermodular.

A *payoff function* is a map $u : X \times Y \to \mathbb{R}$ such that the map $y \mapsto \sup_{x \in X} |u(x, y)|$ is F-integrable. Let $\mathcal{U}_{\text{SPM}}$ (resp. $\mathcal{U}_{\text{SSPM}}$) be the set of all (strictly) supermodular payoff functions.

If the decision maker obtains payoff $u(x, y)$ from choosing action $x \in X$ when the state is $y \in Y$, choosing a decision rule $\psi \in D_F$ prior to observing the state yields the expected payoff

$$\int_Y \int_X u(x, y) \mathrm{d}\psi(x|y) \mathrm{d}F(y). \tag{5}$$

Definition 4. For a set $\mathcal{U}$ of payoff functions and $\phi, \psi \in D_F$, ϕ *$\mathcal{U}$-dominates* ψ if

$$\int_Y \int_X u(x, y) \mathrm{d}\phi(x|y) \mathrm{d}F(y) \geq \int_Y \int_X u(x, y) \mathrm{d}\psi(x|y) \mathrm{d}F(y) \tag{6}$$

for all $u \in \mathcal{U}$; ϕ *strictly $\mathcal{U}$-dominates* ψ if (6) is strict for all $u \in \mathcal{U}$.

In words, ϕ (strictly) $\mathcal{U}$-dominates ψ if ϕ induces (strictly) higher expected payoffs than ψ for *all* payoff functions $u \in \mathcal{U}$.

Theorem 1. *Let $\psi \in D_F$ and let ϕ be the increasing and comonotone rearrangement of ψ with respect to F. Then*

1. *ϕ $\mathcal{U}_{SPM}$-dominates ψ*
2. *no other increasing and comonotone rule $\mathcal{U}_{SPM}$-dominates ψ*
3. *ϕ strictly $\mathcal{U}_{SSPM}$-dominates ψ unless ψ is increasing and comonotone.*

Theorem 1 allows to identify qualitative properties of the solutions to decision problems without the need to know their quantitative features. In Section 3, we give conditions on a range of problems ensuring that (a) some optimal decision rule is increasing or (b) all optimal decision rules are increasing.[11] Among other things, we show that a known informativeness criterion applies to decision problems in which it had not been used before (c.f. Section 3.5).

More specifically, qualitative assumptions on the primitives ensure that, in each case, the decision maker prefers a rule ϕ to an alternative ψ whenever ϕ $\mathcal{U}_{\text{SPM}}$-dominates ψ. Then, if ψ is optimal and ϕ is the increasing and comonotone rearrangement of ψ with respect to F, Part 1 of Theorem 1 implies that ϕ is optimal, too. In particular, some optimal decision rule is increasing. Under further minimal assumptions, the decision maker strictly prefers ϕ to ψ if ϕ $\mathcal{U}_{\text{SPM}}$-dominates *and* strictly $\mathcal{U}_{\text{SSPM}}$-dominates ψ. Then, unless ψ is increasing, Part 3 of Theorem 1 implies that the decision maker strictly prefers the increasing and comonotone rearrangement of ψ to ψ itself. In particular, ψ is not optimal. Thus, all optimal decision rules must be increasing.

Theorem 1 is proved in the Appendix. It follows by combining Lemma 1 with the 'inequality for rearrangements' due to Lorentz (1953). Essentially, Lorentz's result states that Theorem 1 holds if F is the uniform distribution on $[0, 1]$ and D_F is replaced by $\hat{D}_F$. We generalise the result by allowing F to be arbitrary and the decision rules to be non-deterministic.

2.1 Limitations of deterministic rules

In general, restricting attention to deterministic decision rules entails a loss of generality. In particular, unless specific restrictions are imposed on F, one can find a non-inceasing deterministic rule ψ that is not $\mathcal{U}_{\text{SPM}}$-dominated by any increasing deterministic rule ϕ. In this case, Theorem 1 implies that the increasing and comonotone rearrangement of ψ with respect to F is not deterministic, even though ψ itself is. Counter-example 3 of Section 3.1 provides one such ψ, as well as a *random* increasing rule dominating it.

11. For simplicity, actions are one-dimensional in each application, so that all rules are trivially comonotone. In settings with multi-dimensional actions, then same arguments imply that (a) some optimal decision rule is increasing and comonotone, and (b) all optimal rules are increasing and comonotone.

We showed in Lemma 1 that if F is continuous then the increasing (and comonotone) rearrangement of *any* rule ψ is deterministic.[12] We also showed that, if F is a discrete uniform distribution, the increasing (and comonotone) rearrangement of any deterministic rule is deterministic. Hence, in these cases, Theorem 1 holds even if attention is restricted to deterministic rules.

Corollary 1 (Deterministic rules)**.** *Let F be a continuous distribution or a discrete uniform distribution. Let $\psi \in \hat{D}_F$ and let $\phi \in \hat{D}_F$ be the increasing rearrangement of ψ w.r.t. F. Then ϕ is the unique increasing deterministic rule that $\mathcal{U}_{SPM}$-dominates ψ. Moreover, ϕ strictly $\mathcal{U}_{SSPM}$-dominates ψ unless ψ is increasing.*

However, as the next result shows, this restriction on F is also *necessary* for any deterministic rule to admit a increasing rearrangement that is deterministic.

Proposition 1. *Suppose that X has more than one element and that, for any $\psi \in \hat{D}_F$, $\hat{D}_F$ contains an increasing rearrangement of ψ with respect to F. Then F is a continuous distribution or a discrete uniform distribution.*

Proposition 1 is proved in the Appendix. The result implies that, unless F is continuous or a discrete uniform distribution, one may find a non-increasing $\psi \in \hat{D}_F$ that is not $\mathcal{U}_{\text{SPM}}$-dominated by any increasing $\phi \in \hat{D}_F$. This, together with Corollary 1, implies the following.

Corollary 2. *If X has more than one element, the following are equivalent:*

1. *For any $\psi \in \hat{D}_F$, there exists an increasing $\phi \in \hat{D}_F$ that $\mathcal{U}_{SPM}$-dominates ψ.*
2. *F is a continuous distribution or a discrete uniform distribution.*

Proof. To prove that 1 implies 2, fix $\psi \in \hat{D}_F$ and let $\phi \in \hat{D}_F$ be increasing and $\mathcal{U}_{\text{SPM}}$-dominate ψ. Since ϕ is deterministic, it is comonotone. Then Theorem 1 implies that ϕ is the increasing and comonotone rearrangement of ψ w.r.t. F. In particular, ϕ is a rearrangement of ψ w.r.t. F. Since ψ is arbitrary, result follows from Proposition 1. The fact that 2 implies 1 follows directly from Corollary 1. □

An analogue of Theorem 1 for deterministic rules holds if attention is restricted to payoff functions that are concave in the action. Let $\mathcal{U}_{\text{C}}$ be the set of payoff functions $u : X \times Y \to \mathbb{R}$ such that $u(\cdot, y)$ is concave for F-almost all $y \in Y$.

Corollary 3 (Concave payoffs)**.** *Given $\psi \in D_F$, there exists an increasing $\phi \in \hat{D}_F$ such that ϕ $\mathcal{U}_C \cap \mathcal{U}_{SPM}$-dominates ψ and, unless ψ is increasing and comonotone, ϕ strictly $\mathcal{U}_C \cap \mathcal{U}_{SSPM}$-dominates ψ.*

Proof. Fix $\psi \in D_F$ and let $\phi' \in D_F$ be the (possibly random) increasing and comonotone rearrangement of ψ with respect to F. Let $\phi \in \hat{D}_F$ be such that $\phi(y) = \int_X x \mathrm{d}\phi'(x|y)$ for all $y \in Y$. Then, Jensens' inequality implies that ϕ $\mathcal{U}_{\text{C}}$-dominates ϕ'. Result follows from Theorem 1. □

12. Hence, for the purposes of Theorem 1, restricting attention to deterministic rules *is* without loss of generality if F is continuous.

2.2 The necessity of supermodularity

Theorem 1 implies that, for the purposes of simultaneous maximisation of supermodular payoffs, attention may be restricted to increasing (and comonotone) decision rules. From this perspective, it may be viewed as a result in monotone comparative statics. In this section, we show that Theorem 1 does not hold if supermodularity is weakened to the notion of single-crossing differences in (x, y), defined below. Moreover, we prove that supermodularity is tight in a specific sense.

Definition 5. A function $u : \mathbb{R}^2 \to \mathbb{R}$ has *single-crossing differences in* (x, y) if for any $x, x', y, y' \in \mathbb{R}$ such that $x' > x$ and $y' > y$,

$$u(x', y) \geq u(x, y) \Rightarrow u(x', y') \geq u(x, y')$$
$$u(x', y) > u(x, y) \Rightarrow u(x', y') > u(x, y').$$ [13]

Suppose that $X \subseteq \mathbb{R}$ is compact and let $u : X \times Y \to \mathbb{R}$ be such that $u(\cdot, y)$ is continuous for all $y \in Y$. As noted in the introduction, Theorem 4 of Milgrom and Shannon (1994) implies that, if u has single-crossing differences in (x, y), then for any $\psi \in \hat{D}_F$ there exists an increasing $\phi \in \hat{D}_F$ such that $u(\phi(y), y) \geq u(\psi(y), y)$ for all $y \in Y$.[14] Therefore, (2) holds for any distribution F. Nevertheless, given payoffs u_1 and u_2 with single-crossing differences in (x, y), there may exist some non-increasing $\psi \in D_F$ that is not $\{u_1, u_2\}$-dominated by any increasing $\phi \in D_F$.

Counter-example 1. Let $Y \subset \mathbb{R}$ be bounded, $X = \mathbb{R}$, and consider $u_1(x, y) = -x^2/2$ and $u_2(x, y) = e^{-(x-y)^2/2}$. Note that u_1 is supermodular and u_2 has single-crossing differences in (x, y) but it is not supermodular.[15] Let F be any distribution with support Y, and let $D_F^* = \arg\max_{\varphi \in D_F} \int_Y U(\varphi(y), y) \mathrm{d}F(y)$ where $U(x, y) = u_1(x, y) + u_2(x, y)$. Note that, for each $y \in \mathbb{R}$, $U(\cdot, y)$ has a unique maximum.[16] Let $\psi : Y \to \mathbb{R}$ be such that $\psi(y) = \arg\max_{x \in \mathbb{R}} U(x, y)$ for all $y \in Y$. Since Y is bounded, it follows that $D_F^* = \{\psi\}$. Suppose that, for some F, ψ is $\{u_1, u_2\}$-dominated by some $\phi \in D_F$. Then $\int_Y U(\phi(y), y) \mathrm{d}F(y) \geq \int_Y U(\psi(y), y) \mathrm{d}F(y)$. Since $\psi \in D_F^*$, it follows that $\phi \in D_F^*$. Hence $\phi = \psi$. Therefore no rule $\{u_1, u_2\}$-dominates ψ other than itself. Moreover, one can show that ψ is decreasing on $\left(-\infty, -1 - e^{-1/2}\right]$, increasing on $\left[-1 - e^{-1/2}, 1 + e^{1/2}\right]$, and decreasing on

13. Milgrom and Shannon (1994) introduced this concept as the 'single-crossing property'. We adopt the name used by Quah and Strulovici (2009).

14. Let ϕ be such that $\phi(y) = \sup D^*(y)$ for all $y \in Y$ where $D^*(y) = \arg\max_{x \in X} u(x, y)$. Since $u(\cdot, y)$ is continuous and X is compact, then $D^*(y)$ is non-empty and compact. Hence $\phi(y) \in D^*(y)$, so $u(\phi(y), y) \geq u(\psi(y), y)$. Since u has single-crossing differences in (x, y), Theorem 4 of Milgrom and Shannon (1994) implies that $D^*(y)$ is increasing in y in the strong set order. Therefore ϕ is increasing.

15. A smooth function $v : \mathbb{R}^2 \to \mathbb{R}$ is supermodular if and only if its cross-derivative is positive. In particular, the function $v(x, y) = -(x - y)^2$ is supermodular. Then u_2 is a increasing transformation of a supermodular function. Hence it has single-crossing differences in (x, y). However, u_2 is not supermodular since $\frac{\partial^2 u_2}{\partial x \partial y} = u_2(x, y) \left[1 - (x - y)^2\right]$ is positive if and only if $|x - y| \leq 1$.

16. That is because $U(\cdot, y)$ is strictly concave and $U(x, y) \to -\infty$ if $x \to \pm\infty$.

$\left[1+e^{1/2},\infty\right)$.[17]

The next result shows that, unless specific assumptions are imposed on F, supermodularity is necessary for Theorem 1 to hold. The proof is in the Appendix.

Proposition 2. *Let $\mathcal{U}$ be an arbitrary set of payoff functions $u : X \times Y \to \mathbb{R}$. Suppose that, for any distribution F with support within Y, and any $\psi \in D_F$, some increasing rearrangement of ψ with respect to F $\mathcal{U}$-dominates ψ. Then $\mathcal{U} \subseteq \mathcal{U}_{SPM}$.*

3 Applications

To simplify exposition, we restrict attention to scalar deterministic decision rules throughtout this section. That is, we impose $n = 1$ so that $X \subset \mathbb{R}$, and constrain the decision maker to pick maps $\psi : Y \to X$ (i.e. elements of the set $\hat{D}_F$). In order to be able to apply our main result (c.f. Corollary 1), we assume that the distribution F is continuous. However, as noted in Section 2, the result holds for arbitrary distributions F if the decision maker is allowed to pick random decision rules (c.f. Theorem 1). Accordingly, all the results proved in this section continue to hold for arbitrary distributions F if the decision maker may pick random rules.

3.1 Welfare Transfers

Consider a social planner wishing to implement a welfare transfer based on a household-specific measure $y \in Y$ of need for financial support. Let $v(x, y)$ be the utility of households with need $y \in Y$ who receive a transfer $x \in X$, where the function $v : X \times Y \to \mathbb{R}$ is increasing in x and supermodular. This means that households who receive larger transfers are better-off, and that the marginal utility of transfers is increasing in need.[18] Let F describe the distribution of need in the population. A *transfer* is a function $\psi \in \hat{D}_F$ whereby $\psi(y)$ is transferred to households with need $y \in Y$. A planner with a budget $b \in \mathbb{R}$ solves

$$\max_{\psi \in D_F} \int_Y v(\psi(y), y)\mathrm{d}F(y) \quad \text{subject to} \quad \int_Y \psi(y)\mathrm{d}F(y) \leq b.$$

This can be viewed as an optimal taxation problem, similar to those studied by Ramsey (1927) and Mirrlees (1971).[19] Note that $\int_Y v(\psi(y), y)\mathrm{d}F(y)$ is the average *welfare* generated by the transfer, whereas $\int_Y \psi(y)\mathrm{d}F(y)$ is its *cost*. To ensure that the problem has a solution, suppose that Y is compact and that v is continuous.

17. Since $U(\cdot, y)$ is strictly concave, $\psi(y)$ satisfies the first-order condition $\frac{\partial U}{\partial x}(\psi(y), y) = 0$. Total differentiation with respect to y shows that $\psi'(y)$ has the same sign as $\frac{\partial^2 u_2}{\partial x \partial y}(\psi(y), y)$. The first-order condition implies that $\psi'(y) = 0$ if and only if $y = \pm\left(1 + e^{-1/2}\right)$.

18. For example, pick $v(x, y) = g(x - y)$ for some increasing and concave function $g : [0, \infty) \to \mathbb{R}$ and view y as the negative of wealth. Alternatively, y can be a poverty index that is not based on wealth alone (such as the Global Multi-dimensional Poverty Index).

19. Note that the households have no outside option and they are assumed to report y truthfully regardless of incentive compatibility.

The following result states that any non-increasing transfer can be rearranged into an increasing transfer that produces higher welfare, without altering its cost. Crucially, this rearrangement is not sensitive to the functional form of the households' utility u, as long as u is supermodular. Thus, a planner facing a population of households with supermodular utility may restrict attention to increasing transfers even if, as it almost always is the case in practice, she does not know the quantitative aspects of the households' utility function.

Proposition 3. *For any non-increasing $\psi \in \hat{D}_F$, the increasing rearrangement of ψ with respect to F has same cost and generates higher welfare for any supermodular v. If v is* strictly *supermodular, then all optimal transfers are increasing.*

Proof. Fix $\psi \in \hat{D}_F$ and let ϕ be its increasing rearrangement with respect to F. Since F is continuous, Corollary 1 implies that ϕ $\mathcal{U}_{\text{SPM}}$-dominates ψ. Since $v \in \mathcal{U}_{\text{SPM}}$, ϕ produces higher welfare than ψ. Since ϕ is a rearrangement of ψ then ϕ and ψ induce the same *distribution* of transfers x. In particular, $\int_Y \phi(y)\mathrm{d}F(y) = \int_Y \psi(y)\mathrm{d}F(y)$; that is, ϕ and ψ have the same cost. To prove the second part, suppose that ψ is not increasing. Then Corollary 1 implies that ϕ strictly $\mathcal{U}_{\text{SSPM}}$-dominates ψ. Since $v \in \mathcal{U}_{\text{SSPM}}$, then ϕ induces strictly higher welfare than ψ. Hence ψ can't be optimal. □

The following examples show that the (unique) optimal transfer may be *decreasing* if v is not supermodular, or if F is not continuous. However, in the latter case, this only holds under the constraint of deterministic transfers.

Counter-example 2: *if v is not supermodular.* Let $X = [0, 1]$, and let $v(x, y) = (x - y)^2$ where $y \leq 0$ denotes the negative of wealth. This means that the marginal revenue to transfers is *increasing* in households' wealth. Note that v has single-crossing differences in (x, y) since it is strictly increasing in x. However, it is strictly *sub*modular. It is easy to show that in this case the optimal rule takes the form $\psi(y) = \mathbb{I}\{F(y) \leq b\}$. That is, the planner should implement the largest feasible transfers and target all households that are sufficiently wealthy. In particular, the optimal transfer is *decreasing*.[20]

Counter-example 3: *if F is not continuous.* Let $Y = \{1, 2\}$ where 1 (resp. 2) denotes low (resp. high) need and assume that a proportion $\rho \in (1/2, 1)$ of the households is in high need.[21] Suppose that $u(x, y) = x^2 y$ so that the marginal return to transfers is increasing in their size. It is easy to show that, for any $b > 0$, the planner should transfer all resources to the population in low need if $\rho > 2/3$. By targeting the minority, the planner is able to transfer a higher amount to each household. Since $u(\cdot, 1)$ is sufficiently convex, this is optimal.[22]

20. More generally, if v is submodular, then rearranging *any* transfer into a decreasing transfer increases welfare without altering its cost; if v is strictly submodular, then all optimal transfers are decreasing. This can be seen by adapting the argument in the proof of Proposition 3.

21. If $\rho \in \{1/2, 1\}$, then F is a discrete uniform distribution and Proposition 3 holds.

22. Corollary 3 implies that if $u(\cdot, y)$ is concave for all $y \in Y$ then there exists an optimal increasing transfer, whether or not F is continuous.

If the planner is not constrained to transfer the same amount of resources to households with the same level of need, then an increasing transfer is optimal.[23] In particular, if $X = [0, x_0]$, then it is optimal to transfer x_0 to a fraction $b/(x_0\rho)$ of the households in high need, and to transfer nothing to the rest of the population. This enables the planner to achieve strictly higher welfare than if she were constrained to use a deterministic transfer. However, *horizontal equity* is violated since, among the households in high need, some receive higher transfers than others.[24]

3.2 Bets with subjective beliefs

We analyse a setting in which supermodular payoffs arise naturally from the combination of their monotonicity and the monotonicity of the likelihood ratio of events.

Consider two agents who agree on a bet over an outcome $y \in Y$. A *bet* is a function $\psi : Y \to X$ whereby $\psi(y)$ is transferred from Agent 2 to Agent 1 if the outcome is y. For $i = 1, 2$, let $u_i(x)$ be the utility to Agent i if the net transfer is $x \in X$, where the function $u_1 : X \to \mathbb{R}$ is increasing in x and $u_2 : X \to \mathbb{R}$ is decreasing. For $i = 1, 2$, Agent i believes that the outcome y has distribution F_i, which admits a density f_i with support Y. Then, the value to Agent i of a bet ψ is

$$U_i(\psi) = \int_Y u_i[\psi(y)]\mathrm{d}F_i(y).$$

Suppose that Agent 1 is more optimistic about the outcome in the sense that F_1 is a MLR-shift of F_2. The following result states that the Pareto optimal expected utility pairs can be reached by transferring more to Agent 1 when the outcome is higher. Moreover, if F_1 is a strict MLR-shift of F_2 and u_1 is strictly increasing, any Pareto optimal bet must be of this form.[25]

Proposition 4. *Any Pareto optimal expected utility pair can be achieved by an increasing bet. If F_1 is a* strict *MLR-shift of F_2 and u_1 is strictly increasing, then all Pareto optimal bets are increasing.*

Proof. Fix a bet ψ and let ϕ be the increasing rearrangement of ψ with respect to F_2. Then ϕ and ψ induce the same distribution of transfers under F_2, so $U_2(\phi) = U_2(\psi)$. Let $u(x, y) = u_1(x)f_1(y)/f_2(y)$. Since u_1 is increasing and F_1 is a MLR-shift of F_2, then u is supermodular. From Corollary 1, if $F = F_2$ then ϕ $\mathcal{U}_{\mathrm{SPM}}$-dominates ψ. Moreover, for any transfer φ, $\int_Y u(\varphi(y), y)\mathrm{d}F_2(y) = U_1(\varphi)$. Therefore $U_1(\phi) \geq U_1(\psi)$. Hence, if ψ is Pareto optimal, then so is ϕ. This proves the first part. To prove the last part, note that if u_1 is strictly increasing and F_1 is a strict MLR-shift

23. A transfer of this kind is an example of a random decision rule. They are defined in Section 2. In particular, see Definition 1 of increasing random decision rules.

24. Stiglitz (1982) argues that welfare maximisation and even Pareto optimality are sometimes at odds with the principle of horizontal equity, particularly in the context of indirect taxation. He describes economically relevant settings in which payoffs are not concave and random taxation is optimal.

25. F_1 is a strict MLR-shift of F_2 if $f_1(y)/f_2(y)$ is strictly increasing on Y.

of F_2, then u is strictly supermodular. Suppose that ψ is not increasing. Then Corollary 1 implies that, if $F = F_2$, ϕ strictly $\mathcal{U}_{\text{SSPM}}$-dominates ψ. From above, it follows that $U_1(\phi) > U_1(\psi)$. Hence ψ is not Pareto optimal. $\square$

Proposition 4 can be extended to the case where agents maximise *cautious expected utility* (c.f. Cerreia-Vioglio, Dillenberger, and Ortoleva 2015). These preferences satisfy a weaker form of the independence axiom which captures a leaning towards certain prospects. If agents are cautious expected utility maximisers, the value of a bet ψ to Agent i is

$$C_i(\psi) = \inf_{u \in \mathcal{U}_i} u^{-1}\left[\int_Y u[\psi(y)]\mathrm{d}F_i(y)\right]$$

where $\mathcal{U}_1$ (resp. $\mathcal{U}_2$) is a compact set of strictly increasing (resp. decreasing) utility functions. Note that these preferences reduce to expected utility whenever U_i's are singletons. We showed in the proof of Proposition 4 that, under expected utility, rearranging any bet ψ with respect to F_2 into an increasing bet ϕ induces a Pareto improvement. Furthermore, Since the construction of ϕ is not sensitive to the particular choice of u, we are able to deduce that

$$u^{-1}\left[\int_Y u[\phi(y)]\mathrm{d}F_i(y)\right] > u^{-1}\left[\int_Y u[\psi(y)]\mathrm{d}F_i(y)\right] \tag{7}$$

holds for *all* $u \in \mathcal{U}_i$ and $i = 1, 2$. Hence, Proposition 4 extends to the cautious expected utility case.[26]

Proposition 5. *Any Pareto optimal cautious expected utility pair can be achieved by an increasing bet. If F_1 is a* strict *MLR-shift of F_2, then all Pareto optimal bets are increasing.*

Proof. Let ψ be a bet and let ϕ be its increasing rearrangement with respect to F_2. The proof of Proposition 4 implies that $U_2(\phi) = U_2(\psi)$ for *any* function u_2. If u_2 is strictly decreasing, $U_2(\phi) = U_2(\psi)$ is equivalent to $u_2^{-1}[U_2(\phi)] = u_2^{-1}[U_2(\psi)]$. Since elements of $\mathcal{U}_2$ are strictly decreasing, $C_2(\phi) = C_2(\psi)$. Similarly, $U_1(\phi) \geq U_1(\psi)$ for any increasing function u_1. If u_1 is strictly increasing then $U_1(\phi) \geq U_1(\psi)$ is equivalent to $u_1^{-1}[U_1(\phi)] \geq u_1^{-1}[U_1(\psi)]$. Since elements of $\mathcal{U}_1$ are strictly increasing, $C_1(\phi) \geq C_1(\psi)$. Thus, if ψ is Pareto optimal, then so is ϕ. This proves the first part. The second part follows by a similar reasoning. $\square$

26. In fact, Proposition 4 holds for the larger class of *convex* preferences, introduced by Cerreia-Vioglio (2009). The value of ψ under these preferences may be represented as $\inf_{u \in U_i} V_i(\int_Y u[\psi(y)]\mathrm{d}F_i(y), u)$ for a set U_i of increasing functions u and a family $\{V_i(\cdot, u)\}_{u \in U_i}$ of increasing functions. This representation reduces to cautious expected utility if the elements of U_i are strictly increasing and $V_i(\cdot, u) = u^{-1}$ for all $u \in U_i$. The argument used in the proof of Proposition 5 extends naturally to this case.

3.3 Price discrimination and demand externality

Consider a market for a single good with a continuum of buyers and a single seller. Let $q(x, y, z)$ be the demand of a buyer with wealth y facing price x when the average demand for the good is z. Suppose that q is increasing in z. Thus, the good exhibits a positive demand externality.[27] Assume also that q is increasing in wealth y, so that the good is normal, and that q is supermodular in prices and wealth (x, y). That is, for any $x' < x''$ and any z, $q(x', y, z) - q(x'', y, z)$ is decreasing in y. Hence the demand of wealthier buyers is less sensitive to changes in prices.

As an example, consider the demand for a particular software package or operating system. The positive externality is due to network effects and wealthier buyers, such as large firms, have higher demand (than small firms or individuals, say) as they have more employees and therefore require more licenses. Large firms are also less price-sensitive as their expenditure on the good is a smaller share of their total costs. The variable y may also be interpreted as the buyers' *brand preference* for the company selling the good. Buyers with a stronger preference demand more of the good, all other things being equal, and are less sensitive to changes in its price.

Suppose that the seller is able to price discriminate based on individual wealth.[28] In the market for software packages, the fact that firms offer a range of vertically differentiated products which all share the same (small) marginal cost of production may be viewed as an attempt to price discriminate. Let the distribution of wealth in the population be described by F. A *pricing rule* for the seller is a function $\psi \in D_F$ such that buyers with wealth $y \in Y$ are charged $\psi(y) \in X$.

As noted at the start of Section 3, our main result applies if F is continuous or if the seller is able to implement a random pricing rule. The former assumption requires the seller to have sufficient information about buyers' wealth to be able to price discriminate finely among them. The latter requires the seller to be able to charge different prices to buyers with the same wealth (using e.g. lotteries for discounts). Given a pricing rule ψ, average demand, denoted by z_ψ, solves

$$z_\psi = \int_Y q(\psi(y), y, z_\psi) \mathrm{d}F(y). \tag{8}$$

To ensure that (8) has a unique solution for any ψ, suppose that Y is compact, that q is strictly positive and that, for any (x, y), $q(x, y, \cdot)$ is concave and $q(x, y, z) < z$ for z large enough.[29] Suppose that the seller can produce at no cost. Then her

27. The externality is *global*, as it depends on average demand. Farrell and Saloner (1985) and Katz and Shapiro (1985) study models of global demand externalities, but do not allow price discrimination.

28. There is a network literature studying price discrimination by a single seller in a market with a positive demand externality; however, the externality is local and arises from a network of social interactions. In the models of Candogan, Bimpikis, and Ozdaglar (2012) and Bloch and Quérou (2013) as well as in ours, the monopolist has full information. Sundararajan (2007) and Fainmesser and Galeotti (2015) limit the information that the seller has about the network. Using the results of Section 3.4, our framework can be extended to the case where the seller only observes a noisy signal of individual wealth.

29. Note that, since q is decreasing in x and increasing in y, the function $g_\psi(z) = \int_Y q(\psi(y), y, z) \mathrm{d}F(y)$

profit-maximisation problem is

$$\max_{\psi \in D_F} \int_Y \psi(y) q(\psi(y), y, z_\psi) \mathrm{d}F(y). \tag{9}$$

To ensure that the problem has a solution, suppose further that q is continuous. The following result states that it is optimal for the seller to charge wealthier buyers (who are also less price-sensitive) more.

Proposition 6. *For any pricing rule ψ, the increasing rearrangement of ψ with respect to F induces higher revenue. If $q(x, y, z)$ is strictly supermodular in (x, y) for all $z \geq 0$, then the optimal pricing rules are increasing.*

Proof. Fix $\psi \in D_F$. Since F is a continuous and q is supermodular in (x, y) for all $z \geq 0$, Corollary 1 implies that the increasing rearrangement of ψ with respect to F (ϕ, say) $\mathcal{U}$-dominates ψ where $\mathcal{U} = \{q(\cdot, z) : z \geq 0\}$. This implies that $z_\phi \geq z_\psi$.[30] Let $v(x, y) = xq(x, y, z_\psi)$. Since $q(x, y, z_\psi)$ is increasing in y and supermodular in (x, y), then v is supermodular. Hence Corollary 1 implies that

$$\int_Y \phi(y) q(\phi(y), y, z_\psi) \mathrm{d}F(y) \geq \int_Y \psi(y) q(\psi(y), y, z_\psi) \mathrm{d}F(y). \tag{10}$$

Then, since q is increasing in z and $z_\phi \geq z_\psi$, $xq(x, y, z_\phi) \geq xq(x, y, z_\psi)$ for all (x, y). Hence

$$\int_Y \phi(y) q(\phi(y), y, z_\phi) \mathrm{d}F(y) \geq \int_Y \phi(y) q(\phi(y), y, z_\psi) \mathrm{d}F(y). \tag{11}$$

Combining (10) and (11) implies that ϕ induces weakly higher revenue than ψ. To prove the last part, suppose that ψ is not increasing. If $q(x, y, z)$ is strictly supermodular in (x, y) for all $z \geq 0$, then v is strictly supermodular. Then, Corollary 1 implies that (10) is strict. The previous reasoning shows that ψ is not optimal. □

Intuitively, rearranging the pricing rule into an increasing rule shifts the weight of high prices onto wealthier consumers without altering the overall distribution of prices. Since revenue is supermodular in prices and wealth, other things being equal, this increases revenue. However, the shift also affects average demand. Indeed, since demand is supermdoular in prices and wealth, it increases too. This further increases revenue. For this reason, the revenue-maximisation problem has a monotone solution even though it involves a fixed-point.

has image in $[0, z_0]$ for any z_0 such that $q(\min X, \max Y, z_0) < z_0$. To prove existence, apply Tarski's theorem to g_ψ on $[0, z_0]$. To prove uniqueness, note that g_ψ is concave and $g_\psi(0) > 0$. Moreover, for any $z' \in (0, z_\psi)$, $z' = (z'/z_\psi) g_\psi(z_\psi) < (z'/z_\psi) g_\psi(z_\psi) + (1 - z'/z_\psi) g_\psi(0) \leq g_\psi(z')$, where the last inequality follows from the concavity of g_ψ. Hence z' is not a fixed-point of g_ψ. The same holds if $z' > z_\psi$.

30. Note that, since ϕ $\mathcal{U}$-dominates ψ, $g_\phi(z) \geq g_\psi(z)$ for all $z \in [0, z_0]$ (c.f. footnote 29). Since z_ψ is the unique fixed-point of g_ψ on $[0, z_0]$, $z_\psi = \inf\{z \in [0, z_0] | z \geq g_\psi(z)\}$. Then $z < g_\psi(z) \leq g_\phi(z)$ for all $z < z_\psi$. Since z_ϕ is the unique fixed-point of g_ϕ on $[0, z_0]$, $z_\phi = \inf\{z \in [0, 1] | z \geq g_\phi(z)\}$. Hence $z_\phi \geq z_\psi$.

3.4 Choice under uncertainty

Consider a decision maker who has to take an action $x \in X$ before observing the state of the world $s \in S$, where $S \subseteq \mathbb{R}$. Let $u(x,s)$ be her utility if she takes action x and the state is s. Suppose that she has access to a signal y of the state with marginal distribution F and support Y. Let $P(\cdot|y)$ be the posterior distribution of the state s conditional on observing the signal realisation $y \in Y$. The family $P = \{P(\cdot|y)\}_{y\in Y}$ is (strictly) *FOSD-ordered* if for any $y'' > y'$, $P(\cdot|y'')$ is a (strict) FOSD-shift of $P(\cdot|y')$.[31] Intuitively, this means that the decision maker expects the state to be higher when she observes higher signals. A *decision rule* is a function $\psi \in D_F$ whereby the agent takes action $\psi(y)$ when she observes the signal y. The (ex-ante) *value* of ψ is

$$\int_Y \int_S u(\psi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y). \tag{12}$$

Given $\psi, \phi \in D_F$, a set $\mathcal{U}$ of payoff functions u and a set $\mathcal{P}$ of families P of posteriors, ϕ *$\mathcal{U} \times \mathcal{P}$-dominates* ψ if

$$\int_Y \int_S u(\phi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y) \geq \int_Y \int_S u(\psi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y) \tag{13}$$

for all $(u, P) \in \mathcal{U} \times \mathcal{P}$. Moreover, ϕ *strictly* $\mathcal{U} \times \mathcal{P}$-dominates ψ if (13) is strict for all $(u, P) \in \mathcal{U} \times \mathcal{P}$.

Several known monotonicity conditions on u and P guarantee the existence of an increasing decision rule that maximises ex-ante value.[32] However, the optimal rule generally depends on the choice of u and P. The following result shows that rearranging a decision rule with respect to F into an increasing rule increases ex-ante value for *all* supermodular payoff functions and *all* FOSD-ordered families of posteriors simultaneously.[33] Let $\mathcal{U}_{\text{SPM}}$ (resp. $\mathcal{U}_{\text{SSPM}}$) be the set of all (strictly) supermodular payoff functions, and let $\mathcal{P}_{\text{FOSD}}$ (resp. $\mathcal{P}_{\text{SFOSD}}$) be the set of all (strictly) FOSD-ordered families of posteriors.

Proposition 7. *Let $\psi \in D_F$ and let ϕ be its increasing rearrangement with respect to F. Then ϕ $\mathcal{U}_{SPM} \times \mathcal{P}_{FOSD}$-dominates ψ. Moreover ϕ strictly $\mathcal{U}_{SSPM} \times \mathcal{P}_{SFOSD}$-dominates ψ unless ψ is increasing.*

31. $P(\cdot|y'')$ is a strict FOSD-shift of $P(\cdot|y')$ if $P(s|y'') \leq P(s|y')$ for all $s \in S$ and the inequality is strict for some $s \in S$.

32. Athey and Levin (1998) derive a joint condition on u and P that is satisfied both if u is supermodular and P is FOSD-ordered, and if u has single-crossing differences in (x,s) and P is MLR-ordered (that is, $P(\cdot|y'')$ is a MLR-shift of $P(\cdot|y')$ whenever $y'' > y'$). From Karlin and Rubin (1956), it suffices that $\{u(\cdot, y)\}_{y\in Y}$ is a quasiconcave family with increasing peaks (QCIP) and P is MLR-ordered. From Quah and Strulovici (2009), it suffices that P is MLR-ordered and u has the interval-dominance order property, a weaker notion than both single-crossing differences in (x,s) and QCIP.

33. The complete class theorems of Karlin and Rubin (1956) and Quah and Strulovici (2009) give conditions on payoffs $u(x,s)$ and experiments $G = \{G(s|y)\}_{s\in S}$ ensuring that, for any ψ, there exists an increasing ϕ such that $\int_Y u(\phi(y), s) \mathrm{d}G(y|s) \geq \int_Y u(\psi(y), s) \mathrm{d}G(y|s)$ for all $s \in S$ simultanesouly. However, ϕ depends on the choice of u and G.

Proof. Let u be a (strictly) supermodular payoff function and let P be a (strictly) FOSD-ordered family. Then, the interim expected utility $v_{u,P}(x,y) = \int_S u(x,s)\mathrm{d}P(s|y)$ after observing signal y and taking action x is a (strictly) supermodular function of (x,y). Since F is continuous, Corollary 1 implies that ϕ $\mathcal{U}$-dominates ψ where $\mathcal{U} = \{v_{u,P} : u \in \mathcal{U}_{\text{SPM}}, P \in \mathcal{P}_{\text{FOSD}}\}$. Moreover, unless ψ is increasing, ϕ $\mathcal{U}'$-dominates ψ where $\mathcal{U}' = \{v_{u,P} : u \in \mathcal{U}_{\text{SSPM}}, P \in \mathcal{P}_{\text{SFOSD}}\}$. Result follows since $\int_Y \int_S u(\varphi(y), s)\mathrm{d}P(s|y)\mathrm{d}F(y) = \int_Y v_{u,P}(\varphi(y), y)\mathrm{d}F(y)$ for any $\varphi \in D_F$, payoff function u and family of posteriors P. □

Example 1 – Insurance for heterogeneous agents. A group of agents, indexed by i, are subject to shocks s_i. A seller offers insurance payments conditioned on noisy signals y_i of s_i, where y_i's have identical marginal distribution F.[34] An *insurance* contract is a rule $\psi \in D_F$ whereby agent i receives a net transfer $\psi(y_i)$ if the insurer observes $y_i \in Y$.[35] Let $v_i(x_i - s_i)$ be the utility of Agent i if she buys the insurance, the value of the shock is s_i, and the insurance transfer is x_i, where v_i is increasing and concave. Let $\bar{v}_i$ be the utility of Agent i if she does not buy the insurance. Suppose that the insurer sells a single product and wants to maximise profits subject to all agents purchasing it. Then, she solves

$$\min_{\psi \in D_F} \int_Y \psi(y)\mathrm{d}F(y) \quad \text{s.t.} \quad \int_Y \int_S v_i(\psi(y_i) - s_i)\mathrm{d}P_i(s_i|y_i)\mathrm{d}F(y_i) \geq \bar{v}_i \quad \forall i$$

where P_i is the FOSD-ordered family of distributions of s_i conditional on y_i.[36]

I claim that it is optimal to offer an insurance that pays more whenever the observed signal is higher. Formally, there exists an optimal ψ that is increasing. Given any ψ, let ϕ be the increasing rearrangement of ψ with respect to F. For all i, since v_i is concave, then $u_i(x_i, s_i) = v_i(x_i - s_i)$ is supermodular. Then Proposition 7 implies that ϕ induces higher ex-ante utility than ψ to all agents. Thus, any agent purchasing ψ also buys ϕ. Moreover, as ϕ is a rearrangement of ψ, $\int_Y \phi(y)\mathrm{d}F(y) = \int_Y \psi(y)\mathrm{d}F(y)$. Hence, if ψ is optimal, then so is ϕ. This proves the claim. A similar reasoning shows that, if P_i's are strictly FOSD-ordered and v_i's are strictly concave for all i, then any optimal insurance must be increasing.

Example 2 – Ambiguity over the family of posteriors. Consider a firm facing uncertainty about its marginal cost, parametrised by s. Suppose that a noisy signal y of s is available and that the firm commits to a decision rule ψ whereby it supplies $\psi(y)$ units if the signal realisation is y. Let $u(x,s)$ be the firm's profits when producing x units. Suppose that marginal cost is decreasing in s, so

34. The fact that y_i is contractible whereas s_i is not may be due to legal restrictions or to the fact that the insurer does not observe s_i. The fact that y_i's have identical marginal distribution is justified if y_i is determined by the risk profiles of agent i and the insurer is able to offer different products to agents with different risk profiles.

35. The net transfer may be decomposed as $t(y_i) - p$ where $t(y_i) \geq 0$ is the insurance payment and $p \geq 0$ is the price of insurance.

36. If the insurer were able to offer different insurance contracts to each agent, then the problem would be separable across i's and could be solved using standard techniques.

that u is supermodular.[37] If the firm knew the marginal distribution F of y and the family P of posteriors of s given y, its profit-maximisation problem would be

$$\max_{\psi \in D_F} \int_Y \int_S u(\psi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y).$$

Suppose instead that the firm knows F but does not know P.[38] Moreover, it is nevertheless confident that P belongs to some set $\mathcal{P}$ of FOSD-ordered posterior families.[39] Following the theory of Gilboa and Schmeidler (1989), the firm is a *ambiguity-averse* if her ex-ante utility from $\psi \in D_F$ is $\inf_{P \in \mathcal{P}} \int_Y \int_S u(\psi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y)$. Then, the firm solves

$$\max_{\psi \in D_F} \inf_{P \in \mathcal{P}} \int_Y \int_S u(\psi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y). \tag{14}$$

I claim that it is optimal to supply more after observing higher realisations of y (which signals lower marginal costs). Formally, there exists an optimal rule that is increasing. Given any ψ, let ϕ be its increasing rearrangement with respect to F. Since u is supermodular and elements of $\mathcal{P}$ are FOSD-ordered, Proposition 7 implies ϕ $\{u\} \times \mathcal{P}$-dominates ψ. Hence

$$\inf_{P \in \mathcal{P}} \int_Y \int_S u(\phi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y) \geq \inf_{P \in \mathcal{P}} \int_Y \int_S u(\psi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y).$$

Thus, if ψ is optimal, then so is ϕ.

Note that, in the absence of commitment, the firm may not be dynamically consistent in this setting. That is, a decision rule ψ may be optimal before observing the signal but $\psi(y)$ may not be an optimal action after observing the signal realisation y. Nevertheless, an ambiguity-averse firm who picks supply x *after* observing the signal y in order to maximise *interim* expected profit $v(x, y) = \inf_{P \in \mathcal{P}} \int_S u(x, s) \mathrm{d}P(s|y)$ should also adopt an increasing rule. Indeed, a result from Dziewulski and Quah (2016) implies that, for any rule ψ, there exists an increasing rule ϕ such that $v(\phi(y), y) \geq v(\psi(y), y)$ for any $y \in Y$.[40] However, as noted above, ϕ need not improve on ψ from an ex-ante perspective.[41]

37. This holds if $u(x, s) = p(x)x - c(x, s)$ where $p(x)$ is the inverse demand and c is the cost function.

38. For instance, suppose that the firm has access to a large number of i.i.d. realisations of y but only observed few i.i.d. realisations of (y, s), e.g. because it has only been active for a short period of time. Then, the firm can estimate F with confidence but does not have a trustworthy estimate of P.

39. For instance, the firm might conjecture that, for any $y \in Y$, $s = \alpha + \beta y + \epsilon$, where $\epsilon \sim N(0, \sigma^2)$ and pick $\mathcal{P} = \{N(\alpha + \beta y, \sigma^2) : (\alpha, \beta, \sigma^2) \in A\}$ where $A \subset \mathbb{R}^2 \times \mathbb{R}_+$ is a confidence interval for $(\alpha, \beta, \sigma^2)$. In this case, the elements of $\mathcal{P}$ are FOSD-ordered if and only if $\beta \geq 0$ for any $(\alpha, \beta, \sigma^2) \in A$.

40. Dziewulski and Quah (2016) show that if u is supermodular and, for any $y \in \mathbb{R}$, $(\Lambda(y), \geq)$ is a set of one-dimensional distributions endowed with the FOSD-order such that $\Lambda(y'') \geq \Lambda(y')$ in the strong set order whenever $y'' > y'$, then $w(x, y) = \inf_{\lambda \in \Lambda(y)} \int_{\mathbb{R}} u(x, s) d\lambda(s)$ is supermodular. Note that $v = w$ if $\Lambda(y) = \{P(\cdot|y) : P \in \mathcal{P}\}$.

41. In fact, there is more than one plausible way for the firm to update its beliefs about s after observing y. See Epstein and Le Breton (1993), Gilboa and Schmeidler (1993), Epstein and Schneider (2003) and Hanany and Klibanoff (2007) for a discussion.

Example 3 – Knightian uncertainty. Proposition 7 can also be applied to problems of choice under uncertainty with incomplete preferences. Consider a decision maker with utility function u who does not know the joint distribution J of signal y and state s. Suppose that she knows that J belongs to some (closed, convex) set $\mathcal{J}$. Following Bewley (2002), the decision maker has *Knightian preferences* if, for any pair of decision rules ψ, ϕ she prefers ϕ to ψ if and only if

$$\int_{Y \times S} u(\phi(y), s) \mathrm{d}J(y, s) \geq \int_{Y \times S} u(\psi(y), s) \mathrm{d}J(y, s) \quad \forall J \in \mathcal{J}.$$ [42]

Note that these preferences are incomplete over the set of all decision rules. Suppose that u is supermodular and that elements of $\mathcal{J}$ share the same marginal distribution F over Y and induce FOSD-ordered posteriors over S. Then Proposition 7 implies that, for any $\psi \in D_F$, the decision maker has a preference for the increasing rearrangement of ψ with respect to F over ψ itself. Therefore, an agent with a supermodular utility function who faces knightian uncertainty over a FOSD-ordered family of posteriors may restrict attention to increasing decision rules.

3.5 Informativeness

Consider the problem of choosing an action $x \in X$ based on a noisy signal $y \in Y$ of the state of the world $s \in S$, described in the previous section. Given two families P and P' of posteriors, P' is *more informative* than P for a decision maker with utility function $u(x, s)$ and signal distribution F if, for any $\psi \in D_F$, there exists $\phi \in D_F$ such that

$$\int_Y \int_S u(\phi(y), s) \mathrm{d}P'(s|y) \mathrm{d}F(y) \geq \int_Y \int_S u(\psi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y). \tag{15}$$

Intuitively, a more informative family of posteriors gives a decision maker higher ex-ante utility if she picks the right decision rule.

It is well-known that, if P is FOSD-ordered, then P' is more informative than P for any supermodular u and any F such that the follwing holds. The joint distribution over (y, s) induced by P' and F dominates the one induced by P and F in the *supermodular stochastic order* (written $P' \geq_F P$), i.e.

$$\int_Y \int_S v(y, s) \mathrm{d}P'(s|y) \mathrm{d}F(y) \geq \int_Y \int_S v(y, s) \mathrm{d}P(s|y) \mathrm{d}F(y) \tag{16}$$

for all supermodular functions v.[43]

However, given ψ, the particular ϕ for which (15) holds is generally sensitive

42. Such preferences may arise in the presence of a status quo. For instance, if ψ is the status quo and ϕ is a potential alternative, ϕ is preferred to ψ if and only if it has higher value in all conceivable scenarios.

43. The supermodular stochastic order admits a useful characterisation. See Epstein and Tanny (1980), Meyer and Strulovici (2012) and Meyer and Strulovici (2015).

to the choice of u, P and P'. The next result states that rearranging a rule ψ with respect to F into an increasing rule ϕ raises the ex-ante utility of *all* decision makers with supermodular utility functions who switch from *any* posterior family P to *any* other family P' that satifies the above informativeness criterion.

Proposition 8. *Let $\psi \in D_F$ and let ϕ be the increasing rearrangement of ψ with respect to F. Then (15) holds for any supermodular u and any posterior families P' and P such that P is FOSD-ordered and $P' \geq_F P$.*

Proof. Fix $\psi \in D_F$, P and u. Since P is FOSD-ordered then, Proposition 7 implies that (13) holds. Moreover, since ϕ is increasing and u is supermodular, then the function $v(y,s) = u(\phi(y),s)$ is supermodular. Since $P' \geq_F P$ then (16) holds. Result follows by combining the inequalities. □

Proposition 8 improves on existing results in statistical decision theory. A *experiment* $G = \{G(\cdot|s)\}_{s\in S}$ is a collection of distributions $G(\cdot|s)$ of the signal y conditional on the value of the state s.[44] An experiment H is *more informative* than G for a decision maker with utility u if for any rule ψ there exists another rule ϕ such that $\int_Y u(\phi(y),y)dH(y|s) \geq \int_Y u(\psi(y),y)dG(y|s)$ *for all* $s \in S$.[45] In contrast, Proposition 8 states that, given F, for any ψ, there exists ϕ such that (15) holds for any P' such that $P' \geq_F P$. Results about informativeness of experiments cannot be used to derive Proposition 8 since, given P_1, P_1', P_2, P_2' and F such that $P_i' \geq_F P_i$ for $i = 1,2$, there might not exist experiments G and H such that H is more informative than G, and priors Λ_1, Λ_2 over S, such that G and Λ_i give rise to P_i and F and H and Λ_i give rise to P_i' and F for $i = 1,2$.

Example 1 – continued: Recall that a seller offers an insurance $\psi \in D_F$ inducing a net transfer $\psi(y_i)$ to agent i, where y_i is the realisation of a noisy signal of the shock s_i. The insurer minimises $\int_Y \psi(y)\mathrm{d}F(y)$ subject to the constraint that all agents are willing to buy the insurance. For agent i, this holds whenever $\int_Y \int_S u_i(\psi(y_i), s_i)\mathrm{d}P_i(s_i|y_i)\mathrm{d}F(y_i) \geq \bar{v}_i$ where u_i is supermodular.

Assume that $P_i = P$ for all i, where P is some FOSD-ordered family. This means that the agents are exposed to the same shock. However, risk attitutes (described by the curvatures of u_i's) may still differ across agents. Assume that P and F are obtained by the insurer from some experiment G and a prior Λ over S. Suppose that the insurer gains access to an experiment H that is more accurate than G in Lehmann's sense.[46] Without loss of generality, we may as-

44. Note that there exists a one-to-one mapping between pairs (Λ, G), where Λ is a prior distribution over S, and pairs (P,F), which preserves the induced joint distribution over $Y \times S$.

45. Important informativeness criteria are due to Blackwell (1953) and Lehmann (1988). Blackwell's criterion captures informativeness when no substantial restrictions are imposed on experiments and utility functions, whereas Lehmann's criterion applies to statistical experiments that are MLR-ordered and quasiconcave families of utility functions with increasing peaks. Quah and Strulovici (2009) proved that Lehman's criterion applies to utilities satisfying the (weaker) interval dominance order property.

46. The experiment H is more accurate than G in Lehmann's sense whenever there exists a map $T : Y \times S \to Y$ such that $T(y,\cdot)$ is increasing for all $y \in Y$, and $H(T(y,s)|s) = G(y|s)$ for all $y \in Y$ and $s \in S$.

sume that H and Λ give rise to the same marginal distribution F for some posterior family P'.[47] Moreover, one can show that $P' \geq_F P$.[48] Then, Proposition 8 implies that, for any ψ, the increasing rearrangement of ψ with respect to F solves $\int_Y \int_S u_i(\phi(y), s) \mathrm{d}P'(s|y) \mathrm{d}F(y) \geq \int_Y \int_S u_i(\psi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y)$ for all i. Thus, agents who buy ψ when the insurer has access to G will also buy ϕ when the insurer has access to H. Moreover, since ϕ is a rearrangement of ψ, $\int_Y \phi(y) \mathrm{d}F(y) = \int_Y \psi(y) \mathrm{d}F(y)$. It follows that, if the insurer has access to H, she is able to sell an insurance to all agents inducing (weakly) lower expected transfers.[49]

Example 2 – continued: Recall that an ambiguity-averse firm commits to a decision rule ψ before observing a signal y of the state s. The firm knows that y has marginal distribution F, and that the family P of posteriors of s given y belongs to some set $\mathcal{P}$ of FOSD-ordered families. When the firm optimlises, its ex-ante utility $V(\mathcal{P})$ is given by (14).

Given two sets $\mathcal{P}$ and $\mathcal{P}'$, when is it the case that $V(\mathcal{P}') \geq V(\mathcal{P})$? We claim that this holds if, for any $P' \in \mathcal{P}'$ there exists $P \in \mathcal{P}$ such that $P' \geq_F P$. If this is so, then Proposition 8 implies that, for any ψ, there exists an increasing ϕ such that, for any $P' \in \mathcal{P}'$, (15) holds for some $P \in \mathcal{P}$. Since $P' \in \mathcal{P}'$ is arbitrary, then

$$\inf_{P' \in \mathcal{P}'} \int_Y \int_S u(\phi(y), y) \mathrm{d}P'(s|y) \mathrm{d}F(y) \geq \inf_{P \in \mathcal{P}} \int_Y \int_S u(\psi(y), y) \mathrm{d}P(s|y) \mathrm{d}F(y).$$

Since ψ is arbitrary, then $V(\mathcal{P}') \geq V(\mathcal{P})$.

4 Conclusion

In this paper, we showed that it is possible to rearrange any decision rule into an increasing rule in a way that does not alter the induced distribution of actions. Moreover, doing so increases the expected payoff induced by any supermodular utility function. In particular, for the purposes of the simultaneous maximisation of supermodular expected payoffs, the decision maker may restrict attention to increasing decision rules.

47. Note that H and Λ give rise to some posterior family P' and some marginal F' which need not equal F. However, if F' is continuous, then the set of distributions over (x, s) that the insurer is able to induce is the same whether she has access to F' or to F.

48. See Milgrom (1981) for a proof.

49. This example is close to Example 6 in Quah and Strulovici (2009), in which a fund manager allocates the wealth of a group of heterogeneous investors between a safe and a risky asset, based on a signal of its return. They show that access to a signal that is more accurate in Lehamnn's sense allows the manager to increase the utility of all investors, irrespective of their increasing utility functions and priors. Their reasoning cannot be used to derive our result because it requires the utility functions to be increasing transformations of each other for any given s. In our setting, this is not the case because the insurer's utility is decreasing in x whereas the agents' utilities are increasing in x.

Appendix

Note that any $\psi \in D_F$ induces a distribution over $X \times Y$ given by

$$J_{\psi F}(x,y) = \int_Y \mathbb{I}\{y' \leq y\}\psi\left(x|y'\right) \mathrm{d}F\left(y'\right). \tag{17}$$

The following facts will be used in the proof of Lemma 1.

Lemma 2. *Let $\psi \in D_F$ be increasing and comonotone. Then $J_{\psi F}$ is comonotone.*

Proof. Pick $(x,y),(\hat{x},\hat{y}) \in X \times Y$ in the support of $J_{\psi F}$. If $x = \hat{x}$, then either $(x,y) \leq (\hat{x},\hat{y})$ or $(x,y) \geq (\hat{x},\hat{y})$ and result holds. Therefore, suppose without loss of generality that $\hat{x}_1 > x_1$. Then it suffices to show that $(\hat{x},\hat{y}) \geq (x,y)$. Since (x,y) is in the support of $J_{\psi F}$, $J_{\psi F}$ assigns positive probability to all neighbourhoods of (x,y). Hence, there exists a sequence $(x^m,y^m)_m \subset X \times Y$ converging to (x,y) such that x^m is in the support of $\psi(\cdot|y^m)$ for all $m \geq 0$. Similarly, there exists a sequence $(\hat{x}^m,\hat{y}^m)_m \subset X \times Y$ converging to $(\hat{x},\hat{y})$ such that $\hat{x}^m$ is in the support of $\psi(\cdot|\hat{y}^m)$ for all $m \geq 0$. If, for some $i \neq 1$, $\hat{x}_i < x_i$, then there would exist $m \geq 0$ such that $\hat{x}_i^m < x_i^m$ and $\hat{x}_1^m > x_1^m$. But then $\psi(\cdot|y^m)$ would not be comonotone. Hence $\hat{x} \geq x$. If $\hat{y} < y$, there would exist $m \geq 0$ such that $\hat{y}^m < y^m$ and $\hat{x}_1^m > x_1^m$. Then $\psi(x'|\hat{y}^m) < 1$ and $\psi(x'|y^m) > 0$, where $x'_1 = (\hat{x}_1^m + \hat{x}^m)/2$ and $x'_i = \max X_i$ for $i > 1$. This contradicts the fact that ψ is increasing. Hence $\hat{y} \geq y$. Result follows. □

Proof of Lemma 1. I prove that ϕ is unique. Suppose that ϕ' is another increasing and comonotone rearrangement of ψ with respect to F. Then it suffices to show that $\phi(\cdot|y) = \phi'(\cdot|y)$ for F-almost all y's. Since ϕ and ϕ' are increasing and comonotone, then Lemma 2 implies that $J_{\phi F}$ and $J_{\phi' F}$ are comonotone. Moreover, since ϕ and ϕ' are rearrangements of ψ, $J_{\phi F}$ and $J_{\phi' F}$ have identical one-dimensional marginal distributions. Then, since $J_{\phi F}$ and $J_{\phi' F}$ are comonotone, $J_{\phi F} = J_{\phi' F}$. Suppose that there exists a set $Y_0 \subset Y$ of positive F-measure such that $\phi(\cdot|y) \neq \phi'(\cdot|y)$ for all $y \in Y_0$ and seek a contradiction. If this is so then, without loss of generality, for any $y \in Y_0$ there exists $x \in X$ such that $\phi(x|y) > \phi'(x|y)$. Then there exists $Y_1 \subseteq Y_0$ and $\epsilon > 0$ such that, for all $y \in Y_1$ there exists $x \in X$ such that $\phi(x|y) - \phi'(x|y) > \epsilon$. Let $g : Y_1 \to X$ be given by $g(y) = \inf\{x \in X : \phi(x|y) - \phi'(x|y) > \epsilon\}$. Since $X = X_1 \times \cdots \times X_n$ where $X_i \subset \mathbb{R}$ is compact for all i, g is well-defined. Since $\phi(\cdot|y)$ and $\phi'(\cdot|y)$ are right-continuous, $\phi(g(y)|y) - \phi'(g(y)|y) \geq \epsilon$ for all $y \in Y_1$. Moreover, it is clear that g is measurable. It follows that $\int_{Y_1} \phi(g(y)|y)\mathrm{F}(y) - \int_{Y_1} \phi'(g(y)|y)\mathrm{F}(y) \geq \epsilon \int_{Y_1} \mathrm{d}F(y) > 0$. That is, $J_{\phi F}$ puts larger mass on $\{(x,y) \in X \times Y_1 : x \leq g(y)\}$ than $J_{\phi' F}$. This contradicts the fact that $J_{\phi F} = J_{\phi' F}$. □

Theorem 1 follows from a version of Lorentz's inequality proved by Burchard and Hajaiej (2006). Proposition 9 below is a corollary of their result.

Proposition 9 (Corollary of Burchard and Hajaiej 2006)**.** *Given $\varphi : (0,1] \to \mathbb{R}^m$, let $\varphi' : (0,1] \to \mathbb{R}^m$ be such that*

$$\varphi_i'(\omega) = \inf\{z \in \mathbb{R} : \mu(\{\omega' \in (0,1] : \varphi_i(\omega') \leq z\}) \geq \omega\} \tag{18}$$

for all $i \leq m$. Then

$$\int_0^1 v\left[\varphi'(\omega)\right] d\omega \geq \int_0^1 v[\varphi(\omega)]d\omega \tag{19}$$

for any supermodular function $v : \mathbb{R}^m \to \mathbb{R}$ such that

$$\int_0^1 \sup_{z \in \mathbb{R}^{m-1}} |v(z, \varphi_m(\omega))|d\omega < \infty. \tag{20}$$

Moreover, unless

$$[\varphi_i(\omega) - \varphi_i(\omega')][\varphi_j(\omega) - \varphi_j(\omega')] \geq 0 \tag{21}$$

for all $i, j \leq m$ and almost all $\omega, \omega' \in (0,1]$, (19) is strict for any strictly supermodular v such that (20) holds.[50]

The *generalised inverse* of a one-dimensional distribution G is the function $G^- : (0,1] \to \mathbb{R}$ given by

$$G^-(\omega) = \inf\{z \in \mathbb{R} : G(z) \geq \omega\}.$$ [51]

The following fact will be used in the proof of Theorem 1.

Lemma 3. *Let G be a comontone distribution and let $G_1^-, \ldots, G_m^-$ be the generalised inverses of its one-dimensional marginals. Let $\varphi : (0,1] \to \mathbb{R}^m$ be given by $\varphi(\omega) = [G_1^-(\omega), \ldots, G_m^-(\omega)]$. Then φ is a random vector with distribution G.*

Proof. Note that $\varphi_0 : (0,1] \to \mathbb{R}$ given by $\varphi_0(\omega) = \omega$ is a random vector with uniform distribution over $[0,1]$. Then result follows from the implication (3) $\Rightarrow$ (1) of Theorem 2 of Dhaene et al. (2002). □

Proof of Theorem 1. To prove the first part, fix $\psi \in D_F$ and let ϕ be the increasing and comonotone rearrangement of ψ with respect to F. Then it suffices to show that (6) holds for any $u \in \mathcal{U}_{\text{SPM}}$ such that $\int_Y \sup_{x \in X} |u(x,y)|\mathrm{d}F(y) < \infty$. Fix u and let $\varphi : (0,1] \to X \times Y$ be a random vector with distribution $J_{\psi F}$, defined in (17). Then (20) holds. Hence Proposition 9 implies that (19) holds, where $\varphi' : (0,1] \to \mathbb{R}^{n+1}$ is defined by (18). Note that, since φ has distribution $J_{\psi F}$

50. Proposition 9 is a special case of a corollary of Theorem 1 of Burchard and Hajaiej (2006), as expressed in equation (2.2). The authors use a different notation: our φ (resp. φ', v) corresponds to their u (resp. $u^\#$, F). Moreover, they require that φ is positive as they impose an integrability condition on v (expressed in equation 2.1) that is weaker than (20). Given our stronger assumption, the requirement that φ is positive is void as one may replace φ_i with e^{φ_i} and v with $v[\log(\cdot), \ldots, \log(\cdot)]$. Note also that, since the domain of φ is restricted to $(0,1]$, φ trivially vanishes at infinity. Finally, the second part of Proposition 9 is not stated explicitly in Burchard and Hajaiej (2006), but it follows easily from the second part of their Theorem 1.

51. If G is invertible, then its inverse equals G^-.

then $\int_Y \int_X u(x,y)\mathrm{d}\psi(x|y)\mathrm{d}F(y) = \int_0^1 u[\varphi(\omega)]\mathrm{d}\omega$. Then, it suffices to show φ' is a random vector with distribution $J_{\phi F}$. Since ϕ is increasing and comonotone, then Lemma 2 implies that $J_{\phi F}$ is comonotone. Moreover, for any $z \in \mathbb{R}$, $\mu(\{\omega' \in (0,1] : \varphi_i(\omega') \leq z\}) = M_{\psi_i F}(z)$ for all $i \leq n$, and $\mu(\{\omega' \in (0,1] : \varphi_{n+1}(\omega') \leq z\}) = F(z)$. Then φ'_i is the generalised inverse of $M_{\psi_i F}$ for all $i \leq n$, and φ_{n+1} is the generalised inverse of F. Finally, since ϕ is a rearrangement of ψ, the one-dimensional marginal distributions of $J_{\psi F}$ and $J_{\phi F}$ coincide. Hence Lemma 3 implies that φ is a random vector with distribution $J_{\phi F}$. Result follows.

To prove the second part, note that, from Lemma 1, the increasing and comonotone rearrangement of ψ with respect to F is unique. Therefore, it suffices to show that, if ψ is $\mathcal{U}_{\mathrm{SPM}}$-dominated by some rule $\phi \in D_F$, then ϕ is a rearrangement of ψ with respect to F. Fix $\phi, \psi \in \hat{D}_F$ such that ϕ is not a rearrangement of ψ with respect to F. Then it remains to show that ϕ does not $\mathcal{U}_{\mathrm{SPM}}$-dominate ψ, i.e. that (6) fails for some supermodular payoff u. Since ϕ is not a rearrangement of ψ, $M_{\phi_i F}(x'_i) \neq M_{\psi_i F}(x'_i)$ for some $i \leq n$ and $x'_i \in X_i$. If $M_{\phi_i F}(x'_i) > M_{\psi_i F}(x'_i)$, then pick $u(x,y) = \mathbb{I}\{x_i > x'_i\}$. If $M_{\phi_i F}(x'_i) < M_{\psi_i F}(x'_i)$, then pick $u(x,y) = \mathbb{I}\{x_i \leq x'_i\}$. Note that u is supermodular in either case. Hence ϕ does not $\mathcal{U}_{\mathrm{SPM}}$-dominate ψ. Result follows.

To prove the last part, suppose that ψ is not strictly $\mathcal{U}_{\mathrm{SSPM}}$-dominated by its increasing and comonotone rearrangement with repsect to F, ϕ. It suffices to show that ψ is increasing and comonotone. Note that there exists a strictly supermodular payoff function u such that (6) is not strict. The argument to prove the first part of the theorem implies that (19) is not strict. Then, Proposition 9 implies that (21) holds for all $i, j \leq n+1$ and almost all $\omega, \omega' \in (0,1]$. Since φ has distribution $J_{\psi F}$, it follows that $J_{\psi F}$ is comonotone.[52] Then, the argument used to prove uniqueness in Lemma 1 implies that $\phi(\cdot|y) = \psi(\cdot|y)$ for F-almost all y's. Hence ψ is increasing and comonotone. □

Proof of Proposition 1. Since X has more than one element, X_i has more than one element for some $i \leq n$. Without loss of generality, assume that $\{0,1\} \in X_i$. Given $y_0 < y_1$, let $\psi \in D_F$ be such that $\psi_i(y) = 1 - \mathbb{I}\{y_0 < y \leq y_1\}$. Since there exists an increasing rearrangement ϕ of ψ with respect to F, $F(y_1) - F(y_0) = M_{\psi_i F}(0) = M_{\phi_i F}(0) \in \overline{F(\mathbb{R})}$, where the inclusion follows from the fact that ϕ is increasing. Then, it suffices to show that, unless F is a continuous or a discrete uniform distribution, $F(y_1) - F(y_0) \notin \overline{F(\mathbb{R})}$ for some $y_0 < y_1$. Suppose that F does not have an atom at the bottom of its support. Formally, suppose that for all $\epsilon > 0$, there exists $y_0 \in \mathbb{R}$ such that $F(y_0) \in (0,\epsilon)$. Since F is not a continuous distribution, it has an atom $y_1 \in \mathbb{R}$. Then there exists $y_0 \in \mathbb{R}$ such that $F(y_0) \in (0,p)$ where p is the probability mass at y_1. Hence $y_0 < y_1$ and $F(y_1) - F(y_0) \notin \overline{F(\mathbb{R})}$. Suppose instead that F has an atom y_0 at the bottom of its support. Since F is not a discrete uniform distribution, there exists $y_1 \in \mathbb{R}$ such

52. This follows from the definition of comonotonicty, see Section 2.

that $F(y_1) \notin \{mF(y_0) : m \in \mathbb{N}_0, m \leq F(y_0)^{-1}\}$. Then $y_1 > y_0$. Moreover, y_1 can be chosen such that $F(y_1) - F(y_0) \notin \overline{F(\mathbb{R})}$. □

Proof of Proposition 2. Suppose that $u \in \mathcal{U}$ is not supermodular. It suffices to find a distribution F and $\psi \in D_F$ such that (6) fails if $\phi \in D_F$ is an increasing rearrangement of ψ with respect to F. Since u is not supermodular, there exists $(x^1, y^1), (x^2, y^2) \in X \times Y$ such that $y^2 \geq y^1$ and $u(x^1 \wedge x^2, y^1) + u(x^1 \vee x^2, y^2) < u(x^1, y^1) + u(x^2, y^2)$. Let F be the uniform distribution over $\{y_1, y_2\}$ and let $\psi \in \hat{D}_F$ be such that $\psi(y^j) = x^j$ for $j = 1, 2$. Then, ψ admits a unique increasing rearrangement with respect to F, given by $\phi(y^1) = x^1 \wedge x^2$ and $\phi(y^2) = x^1 \vee x^2$. So

$$\int_Y u(\phi(y), y)\mathrm{d}F(y) - \int_Y u(\psi(y), y)\mathrm{d}F(y)$$
$$= u(x^1 \wedge x^2, y^1) - u(x^1, y^1) + u(x^1 \vee x^2, y^2) - u(x^2, y^2) < 0.$$

Therefore (6) fails. □

References

Arieli, Itai, and Yakov Babichenko. 2019. "Private bayesian persuasion." *Journal of Economic Theory* 182:185–217.

Athey, S., and J. Levin. 1998. "The value of information in monotone decision problems." MIT Working Paper.

Bewley, T. 2002. "Knightian decision theory. Part I." *Decisions in economics and finance* 25 (2): 79–110.

Blackwell, D. 1953. "Equivalent comparisons of experiments." *The annals of mathematical statistics* 24 (2): 265–272.

Bloch, F., and N. Quérou. 2013. "Pricing in social networks." *Games and economic behavior* 80:243–261.

Burchard, A., and H. Hajaiej. 2006. "Rearrangement inequalities for functionals with monotone integrands." *Journal of Functional Analysis* 233 (2): 561–582.

Candogan, O., K. Bimpikis, and A. Ozdaglar. 2012. "Optimal pricing in networks with externalities." *Operations Research* 60 (4): 883–905.

Cerreia-Vioglio, S., D. Dillenberger, and P. Ortoleva. 2015. "Cautious expected utility and the certainty effect." *Econometrica* 83 (2): 693–728.

Cerreia-Vioglio, Simone. 2009. "Maxmin expected utility on a subjective state space: Convex preferences under risk."

Chernozhukov, V., I. Fernandez-Val, and A. Galichon. 2009. "Improving point and interval estimators of monotone functions by rearrangement." *Biometrika* 96 (3): 559–575.

Dhaene, Jan, Michel Denuit, Marc J Goovaerts, Rob Kaas, and David Vyncke. 2002. "The concept of comonotonicity in actuarial science and finance: theory." *Insurance: Mathematics and Economics* 31 (1): 3–33.

Dziewulski, P., and J. Quah. 2016. "Supermodular Correspondences." University of Oxford Working Paper, no. 795.

Epstein, L., and M. Le Breton. 1993. "Dynamically consistent beliefs must be Bayesian." *Journal of Economic theory* 61 (1): 1–22.

Epstein, L., and M. Schneider. 2003. "Recursive multiple-priors." *Journal of Economic Theory* 113 (1): 1–31.

Epstein, L., and S. Tanny. 1980. "Increasing generalized correlation: a definition and some economic consequences." *Canadian Journal of Economics,* 16–34.

Fainmesser, I., and A. Galeotti. 2015. "Pricing network effects." *The Review of Economic Studies.*

Farrell, J., and G. Saloner. 1985. "Standardization, compatibility, and innovation." *The RAND Journal of Economics,* 70–83.

Gilboa, I., and D. Schmeidler. 1989. "Maxmin expected utility with non-unique prior." *Journal of mathematical economics* 18 (2): 141–153.

———. 1993. "Updating Ambiguous Beliefs." *Journal of Economic Theory* 59 (1): 33–49.

Hanany, E., and P. Klibanoff. 2007. "Updating preferences with multiple priors." *Theoretical Economics* 2 (3): 261–298.

Karlin, S., and H. Rubin. 1956. "The theory of decision procedures for distributions with monotone likelihood ratio." *The Annals of Mathematical Statistics,* 272–299.

Katz, M. L., and C. Shapiro. 1985. "Network externalities, competition, and compatibility." *The American economic review* 75 (3): 424–440.

Landsberger, M., and I. Meilijson. 1994. “Co-monotone allocations, Bickel-Lehmann dispersion and the Arrow-Pratt measure of risk aversion.” *Annals of Operations Research* 52 (2): 97–106.

Lehmann, E. L. 1988. “Comparing Location Experiments.” *The Annals of Statistics,* 521–533.

Lorentz, G. 1953. “An inequality for rearrangements.” *The American Mathematical Monthly* 60 (3): 176–179.

Meyer, M. 2018. “Choosing Joint Distributions: Theory and Application to Information Design.”

Meyer, M., and B. Strulovici. 2015. *Beyond correlation: Measuring interdependence through complementarities.* Technical report. Working Paper.

———. 2012. “Increasing interdependence of multivariate distributions.” *Journal of Economic Theory* 147 (4): 1460–1489.

Milgrom, P. 1981. “Good news and bad news: Representation theorems and applications.” *The Bell Journal of Economics,* 380–391.

Milgrom, P., and C. Shannon. 1994. “Monotone comparative statics.” *Econometrica,* 157–180.

Mirrlees, J. A. 1971. “An exploration in the theory of optimum income taxation.” *The review of economic studies* 38 (2): 175–208.

Quah, J. K.-H., and B. Strulovici. 2009. “Comparative statics, informativeness, and the interval dominance order.” *Econometrica* 77 (6): 1949–1992.

Ramsey, F. 1927. “A Contribution to the Theory of Taxation.” *The Economic Journal* 37 (145): 47–61.

Stiglitz, J. 1982. “Utilitarianism and horizontal equity: The case for random taxation.” *Journal of Public Economics* 18 (1): 1–33.

Sundararajan, A. 2007. “Local network effects and complex network structure.” *The BE Journal of Theoretical Economics* 7 (1).