Supplement to "Belief heterogeneity in the Arrow-Borch-Raviv insurance model"

Ghossoub, Mario

University of Montreal

10 June 2010

SUPPLEMENT TO "BELIEF HETEROGENEITY IN THE ARROW-BORCH-RAVIV INSURANCE MODEL"

MARIO GHOSSOUB

UNIVERSITÉ DE MONTRÉAL

FIRST DRAFT: JUNE 10, 2010

THIS DRAFT: MARCH 22, 2012

LATEST DRAFT: https://sites.google.com/site/marioghossoub

Abstract. This paper is a supplement to Ghossoub [11]. In this supplement, some of the results of Ghossoub [11], as well as the techniques used to obtain these result are extended to a more general problem of demand for contingent claims with belief heterogeneity. Moreover, a general problem of monotone comparative statics under heterogeneous uncertainty is examined, and I show how the idea of vigilance can be used to obtain a monotone comparative statics result in this case.

8. The Demand for Contingent Claims under Heterogeneous Uncertainty

This section gives an extension of the insurance model considered in Ghossoub [11] to a general setting of demand for claims that pay contingent on the realization of some underlying random variable. This setting can encompass, for instance, models of demand for derivative securities, that is, financial securities that pay contingent on the realization of the (random) price of some underlying stock. The analysis is purposefully kept at a general level, but it can easily be applied to different practical situations.

As in Section 4, S is the set of states of the world and $\mathcal{G}$ is a σ-algebra of events on S. Assume that a decision maker (DM) faces a fundamental uncertainty that affects her wealth and consumption. This uncertainty will be modelled as a (henceforth fixed) element X of $B^+(\mathcal{G})$ with a closed range $[0, M] := X(S)$, where $M := \|X\|_{sup} < +\infty$. The DM wishes to purchase from a claim issuer (CI) a claim that pays contingent on the realizations of the underlying uncertainty. For instance, in problems of demand for insurance the uncertainty X can be seen as the underlying insurable loss against which the DM seeks an insurance coverage $I \circ X$. In problems of optimal debt contracting, the uncertainty X can be seen as the interest on a loan, and $I \circ X$ as the repayment scheme. Hereafter, I will denote by Σ the σ-algebra $\sigma\{X\}$ of subsets of S generated by X.

8.1. **Preferences and Utilities.**

The DM's decision process is assumed to consist in choosing a certain *act* among a collection of given *acts* whose realization, in each state of the world s, depends on the value $X(s)$ of the uncertainty X is the state s. Formally, the DM and the CI have preferences over acts in a framework à la Savage. Here, the set of consequences (or prizes) is taken to be $\mathbb{R}$. Let $\mathcal{F}$ denote the collection of all $\mathcal{G}$-measurable functions $f : S \to \mathbb{R}$. The elements of choice (or *acts*) are taken to be the elements of $B^+(\Sigma) \subset \mathcal{F}$. The nature of the problem makes this a natural assumption. Indeed, the goal here is to determine the optimal function of the uncertainty, that is, the optimal claim $Y := I \circ X \in B^+(\Sigma)$, for some Borel-measurable map $I : X(S) \to \mathbb{R}^+$, that will satisfy a certain set of requirements (constraints). The DM's preferences $\succcurlyeq_{DM}$ over $B^+(\Sigma)$ and the CI's preferences $\succcurlyeq_{CI}$ over $B^+(\Sigma)$ determine their subjective beliefs. These beliefs are represented by subjective probability measures μ and ν, respectively, on the measurable space (S, Σ). Moreover, I will assume the following representations for the preferences:

Assumption 8.1. *The DM's preference $\succcurlyeq_{DM}$ admits a representation of the form:*

$$Y_1 \succcurlyeq_{DM} Y_2 \iff \int \mathcal{U}(X, Y_1)\ d\mu \geqslant \int \mathcal{U}(X, Y_2)\ d\mu \tag{8.1}$$

where for each act $Y \in B^+(\Sigma)$, the mapping

$$\begin{aligned} \mathcal{U}(X, Y) : S &\to \mathbb{R} \\ s &\mapsto \mathcal{U}(X(s), Y(s)) \end{aligned} \tag{8.2}$$

is understood to be the DM's utility of wealth (associated with the act Y), and where the mapping

$$\begin{aligned} \mathcal{U}(X, .) : B^+(\Sigma) &\to B(\Sigma) \\ Y &\mapsto \mathcal{U}(X, Y) \end{aligned} \tag{8.3}$$

is (uniformly) bounded and sequentially continuous in the topology of pointwise convergence.

Similarly, the CI's preference $\succcurlyeq_{CI}$ admits a representation of the form:

$$Y_1 \succcurlyeq_{CI} Y_2 \iff \int \mathcal{V}(Y_1)\ d\nu \geqslant \int \mathcal{V}(Y_2)\ d\nu \tag{8.4}$$

where for each act $Y \in B^+(\Sigma)$, the mapping

$$\begin{aligned} \mathcal{V}(Y) : S &\to \mathbb{R} \\ s &\mapsto \mathcal{V}(Y(s)) \end{aligned} \tag{8.5}$$

is understood to be the CI's utility of wealth (associated with the act Y), and where the mapping

$$\begin{aligned} \mathcal{V} : B^+(\Sigma) &\to B(\Sigma) \\ Y &\mapsto \mathcal{V}(Y) \end{aligned} \tag{8.6}$$

is (uniformly) bounded and sequentially continuous in the topology of pointwise convergence.

For instance, if for each $Y = I \circ X \in B^+(\Sigma)$ one has $\mathcal{U}(X, Y) := u(a - X + Y)$, for some $a \in \mathbb{R}$ and some continuous bounded utility function $u : \mathbb{R} \to \mathbb{R}$, then the mapping $\mathcal{U}(X, .) : B^+(\Sigma) \to B(\Sigma)$ is (uniformly) bounded and sequentially continuous in the topology of pointwise convergence. Also, if for each $Y = I \circ X \in B^+(\Sigma)$ one has $\mathcal{V}(Y) := v(b - Y)$, for some $b \in \mathbb{R}$ and some continuous bounded utility function $v : \mathbb{R} \to \mathbb{R}$, then the mapping $\mathcal{V} : B^+(\Sigma) \to B(\Sigma)$ is (uniformly) bounded and sequentially continuous in the topology of pointwise convergence.

Given Assumption 8.1, the DM's problem here is choosing the optimal *act* $Y^* \in B^+(\Sigma)$ that will maximize her expected utility of wealth, with respect to her subjective probability measure μ.

8.2. **Subjective Beliefs and Vigilance.** The subjectivity of the beliefs of each of the DM and the CI is reflected in the different subjective probability measure that each has over the measurable space (S, Σ). I will also make the assumption that the uncertainty X (with closed range $[0, M]$) has a nonatomic law induced by the probability measure μ.

Assumption 8.2. *The DM's beliefs are represented by the countably additive*[1] *probability measure μ on (S, Σ), and the CI's beliefs are represented by the countably additive probability measure ν on (S, Σ). Moreover, $\mu \circ X^{-1}$ is nonatomic.*

Definition 8.3. The probability measure ν is said to be (μ, X)-*vigilant* if for any $Y_1, Y_2 \in B^+(\Sigma)$ such that

(i) Y_1 and Y_2 have the same distribution under μ, i.e. $\mu \circ Y_1^{-1}(B) = \mu \circ Y_2^{-1}(B)$ for each Borel set B,

(ii) Y_2 and X are comonotonic, i.e. $\Big[Y_2(s) - Y_2(s')\Big]\Big[X(s) - X(s')\Big] \geqslant 0$, for all $s, s' \in S$,

the following holds:

$$Y_2 \succcurlyeq_{CI} Y_1, \text{ that is, } \int \mathcal{V}(Y_2)\ d\nu \geqslant \int \mathcal{V}(Y_1)\ d\nu$$

Clearly, μ is (μ, X)-vigilant. In Section 9, I show that in the specific setting where the DM and the CI assign different probability density functions to the uncertainty X with range $[0, M]$, the assumption of *vigilance* is weaker than the assumption of a *monotone likelihood ratio*.

8.3. **The DM's Problem.** The DM seeks the contingent claim that will maximize her expected utility of wealth, under her subjective probability measure, subject to the CI's participation constraint and to some constraints on the claim. Specifically, the DM's problem is the following:

Problem 8.4.

$$\sup_{Y \in B^+(\Sigma)} \left\{ \int \mathcal{U}(X, Y)\ d\mu \right\} :$$
$$\begin{cases} 0 \leqslant Y \leqslant X \\ \int \mathcal{V}(Y)\ d\nu \geqslant R \end{cases}$$

Remark 8.5. By Assumption 8.1, if Problem 8.4 has a nonempty feasibility set then the supremum in Problem 8.4 is finite. Indeed, there is $N < +\infty$ such that for any feasible $Y \in B^+(\Sigma)$, one has $\mathcal{U}(X, Y)(s) \leqslant N$, for all $s \in S$. Consequently, $\int_D \mathcal{U}(X, Y)\ d\mu \leqslant N\mu(D)$, for each $D \in \Sigma$.

Denote by $\mathcal{F}_{SB}$ the feasibility set Problem 8.4:

[1] Countable additivity of the subjective probability measure representing preferences can be obtained by assuming that preferences satisfy the Arrow-Villegas *Monotone Continuity* axiom [2, 8, 25].

Definition 8.6. Let $\mathcal{F}_{SB} := \left\{ Y \in B^+(\Sigma) : 0 \leqslant Y \leqslant X \text{ and } \int \mathcal{V}(Y)\ d\nu \geqslant R \right\}$.

In the following, I will assume that this feasibility set is nonempty:

Assumption 8.7. *$\mathcal{F}_{SB} \neq \varnothing$.*

The following result shows that *vigilance* is sufficient for the existence of a monotone solution to the DM's problem, that is, a solution which is comonotonic with the underlying uncertainty X. The proof is given in Appendix H.

Theorem 8.8. *Under Assumptions 8.1, 8.2, and 8.7, if ν is (μ, X)-vigilant and if $\mathcal{U}(X, Y)$ is supermodular, then Problem 8.4 admits a solution which is comonotonic with X.*

In Appendix H.2, I give a general algorithm that can be used to characterize a monotone solution to Problem 8.4. The general procedure is based on the following idea:

(1) Lebesgue's Decomposition Theorem [9, Theorem 4.3.1] suggests a decomposition of the measure ν with respect to the measure μ, whereby one can write ν as a sum of two measures, one of which is absolutely continuous with respect to μ, and the other is mutually singular with μ;

(2) This decomposition then suggests a splitting of the initial problem into three subproblems;

(3) A solution of the initial problem is then obtained from the solutions of the other subproblems, combined in an appropriate way.

9. Vigilance and Monotone Likelihood Ratios

The purpose of this subsection is to show that the assumption of *vigilance of beliefs* is weaker than the assumption of a *monotone likelihood ratio* in a setting where the DM and the insurer assign a different probability density function (pdf) to the random loss on its range. Needless to say, this presupposes the existence of such pdf-s. Suppose then that the DM's subjective probability measure μ on (S, Σ) is such that the law $\mu \circ X^{-1}$ is absolutely continuous with respect to the Lebesgue measure, with a Radon-Nikodým derivative f, where $f(t)$ is interpreted as the pdf that the DM assigns to the loss X. Similarly, suppose that the insurer's subjective probability measure ν on (S, Σ) is such that the law $\nu \circ X^{-1}$ is absolutely continuous with respect to the Lebesgue measure, with a Radon-Nikodým derivative g, where $g(t)$ is interpreted as the pdf that the insurer assigns to the loss X. Then $f(t)$ and $g(t)$ are both continuous functions with support $[0, M]$.

Definition 9.1. The *likelihood ratio* is the function $LR : [0, M] \to \mathbb{R}^+$ defined by

$$LR(t) := g(t)/f(t) \tag{9.1}$$

for all $t \in [0, M]$ such that $f(t) \neq 0$.

Now, define the map $Z : S \to \mathbb{R}^+$ by $Z := LR \circ X$. Then Z is nonnegative and Σ-measurable, and LR is a nondecreasing (resp. nonincreasing) function on its domain if and only if Z is comonotonic (resp. anti-comonotonic) with X. Consider the following two conditions that one might impose.

Condition 9.2 (Monotone Likelihood Ratio). *LR is a nonincreasing function on its domain.*

Condition 9.3 (Vigilance). *ν is (μ, X)-vigilant.*

The following proposition shows that the *vigilance* condition is implied by the *monotone likelihood ratio* condition in this particular setting, and under a mild assumption.

Proposition 9.4. *Suppose that $\mathcal{V}$ is such that the induced mapping $\mathcal{V}(.) : \mathbb{R}^+ \to \mathbb{R}$ is a nonincreasing function*[2] *of the parameter y. If Condition 9.2 (Monotone Likelihood Ratio) holds and if the map $\mathcal{V}(I \circ X)\, LR(X) : S \to \mathbb{R}$ is μ-integrable for each $I \circ X \in B^+(\Sigma)$, then condition 9.3 (Vigilance) holds.*

Proof. First note that since the mapping $\mathcal{V}(.) : \mathbb{R}^+ \to \mathbb{R}$ is a nonincreasing function of the parameter y, it follows from Condition 9.2 and Definition B.3 that the map $L : [0, M] \times [0, M] \to \mathbb{R}$ defined by $L(x, y) := \mathcal{V}(y)\, LR(x)$ is supermodular (see Example B.4 (4)).

Suppose that Condition 9.2 holds. To show that Condition 9.3 is implied, choose $Y_1, Y_2 \in B^+(\Sigma)$ such that Y_1 and Y_2 have the same distribution under μ, and Y_2 is comonotonic with X. Then by the μ-a.s. uniqueness of the nondecreasing μ-rearrangement, Y_2 is μ-a.s. equal to $\widetilde{Y}_1$, where $\widetilde{Y}_1$ is the nondecreasing μ-rearrangement of Y_1 with respect to X, that is, $Y_2 = \widetilde{Y}_1$, μ-a.s. Since the function $L(x, y)$ is supermodular, as observed above, then Lemma B.5 yields $\int L\left(X, \widetilde{Y}_{1,\mu}\right)\, d\mu \geqslant \int L\Big(X, Y_1\Big)\, d\mu$, that is, $\int \mathcal{V}\left(\widetilde{Y}_{1,\mu}\right)\, Z\, d\mu \geqslant \int \mathcal{V}(Y_1)\, Z\, d\mu$, where Z is as defined above. Since $Y_2 = \widetilde{Y}_1$, μ-a.s., one then has $\int \mathcal{V}(Y_2)\, Z\, d\mu \geqslant \int \mathcal{V}(Y_1)\, Z\, d\mu$, which yields (by two "changes of variable"[3], and using the definition of f and g as Radon-Nikodým derivatives of $\mu \circ X^{-1}$ and $\nu \circ X^{-1}$, respectively, with respect to the Lebesgue measure) the following:

$$\int \mathcal{V}(Y_2)\, d\nu \geqslant \int \mathcal{V}(Y_1)\, d\nu$$

as required. Condition 9.3 hence follows from Condition 9.2. □

10. Monotone Comparative Statics under Heterogeneous Uncertainty

This section gives a monotone comparative statics result for a class of demand problems under uncertainty, where this uncertainty is perceived differently by the parties involved, in that they assign different likelihoods to its realizations. The uncertainty is taken as given, and the decision maker's (DM) choice variable is a function of this uncertainty. What guarantees that the DM's optimal choice is a nondecreasing function of this underlying uncertainty?

[2] For instance, if $\mathcal{V}(Y) = v(b - Y)$, where v is a nondecreasing utility function and $b \in \mathbb{R}$, then the induced mapping $\mathcal{V}(.) : \mathbb{R} \to \mathbb{R}$ (defined by $\mathcal{V}(t) := v(b - t)$) is a nonincreasing function of the parameter y. This situation occurs most often in contracting problems, and simply says that the CI has an increasing utility function, and his wealth is a nonincreasing function of the claim Y that he issues.

[3] As in [1, Theorem 13.46], and since the map $\mathcal{V}(I \circ X)\, LR(X) : S \to \mathbb{R}$ is μ-integrable for each $I \circ X \in B^+(\Sigma)$.

This problem is an abstraction of many common problems in economic theory that were hitherto only considered in a framework of complete homogeneity of beliefs about the realizations of an underlying uncertainty. It can be formulated as

$$\sup_{Y \in \Theta} V(X,Y) := \int \mathcal{U}(X,Y)\ dP \tag{10.1}$$

where X is a given random variable on a probability space (S, Σ, P), $B(\Sigma)$ is the linear space of all bounded and Σ-measurable functions on S, $\Theta \subset B(\Sigma)$ is a given non-empty constraint set, and $\mathcal{U}(X,Y)$ is bounded and Σ-measurable for each $Y \in \Theta$. When Θ contains another party's individual rationality constraint (participation constraint), one can distinguish between two types of problems, depending on how the underlying uncertainty is perceived by both parties: (i) either both parties agree on the distribution of this uncertainty (which will hence be induced by the probability measure P), or (ii) they have different perceptions of such randomness. The first type of problem is one where uncertainty can be called *homogeneous*, whereas the second type is a problem in which uncertainty can be referred to as being *heterogeneous*.

Surprisingly, the literature is mostly silent on problems of the form (10.1) where the uncertainty is heterogeneous, whereas problems of the form (10.1) with homogeneous uncertainty are abundant. For example, the vast majority of problems of optimal insurance design, or demand for insurance coverage are based on the classical formulation of Arrow [2], Borch [6], and Raviv [20], and are usually stated as a problem of the form (10.1) with homogeneous uncertainty. That is, both the insurer and the insured share the same beliefs about the realizations of some underlying insurable loss random variable X. As discussed in Ghossoub [11], monotonicity of an optimal insurance contract Y is typically desired because such contracts can avoid *ex-post* moral hazard that might arise from a voluntary downward misrepresentation of the loss by the insured.

Problems of debt contracting between investors (lenders) and entrepreneurs (borrowers), such as the ones studied by Gale and Hellwig [10], Townsend [24], or Williamson [27], are also usually stated as a problem of the form (10.1) with homogeneous uncertainty. In this case, a contract specifies the repayment Y that the borrower makes to the lender as a function of the (uncertain) return X on the project being financed. The monotonicity of an optimal contract as a function of the return on investment is a coveted feature since such contracts will be *de facto* truthtelling, and will avoid any misrepresentation of the profitability of the project by the borrower.

Principal-agent problems have also been traditionally stated as problems of the form (10.1) with homogeneous uncertainty, as in the work of Grossman and Hart [12], Holmstrom [13], Mirrlees [17], Page [18], or Rogerson [21], for instance. In that setting, a contract specifies the wage Y that an agent receives from the principal, as a function of the (uncertain) outcome, or output X that occurs as a result of the agent's activity. Since the work of Rogerson [21], monotonicity of the optimal wage contract in the observed output is usually sought after[4].

Numerous other problems can be formulated as in (10.1), such as problems of demand for financial securities given a pricing or budgeting constraint, and where monotonicity of an optimal security Y might reflect hedging against X, for example. Whatever the nature of the problem might be, it is interesting to examine under what conditions an optimal choice of the choice variable Y is monotone in the underlying variable X, and the theory of monotone comparative statics has usually been very fruitful in answering questions of this sort.

[4]For various reasons that are beyond the scope of this paper.

10.1. **The Theory of Monotone Comparative Statics and its Limitations.** The importance of monotone comparative statics analyses in economic theory is well-understood. One can even say that at the core of the motivation behind a sizeable collection of problems in economic theory, very often lies the question of whether or not a quantity is a monotone function of a parameter, or whether a variable output changes monotonically with a variable input. This is even more so significant if the monotonicity of an optimal such output as a function of an input parameter is desired, and indeed, monotone comparative statics techniques have proven to be very fruitful (see [23, 26]). Such techniques can be, and have been used in consumer theory, theory of production, portfolio choice theory, financial economics, and contract theory to answer some basic and intuitive questions.

The theory of monotone comparative statics is typically concerned with the behavior of a solution to a given optimization problem when a primitive of the problem changes. Specifically, let $(L, \geqslant_L)$ be a lattice, $B \subseteq L$ a choice set, $(T, \geqslant_T)$ a partially ordered set interpreted as a set of parameters, and $f : L \times T \to \mathbb{R}$ a given objective function. For the problem of choosing an $x \in B$ that maximizes the objective function given a value t of the parameter, the chief concern is the isotonicity of an optimal choice $x^*(t)$ of x as a function of t, that is,

$$t_1 \geqslant_T t_2 \implies x^*(t_1) \geqslant_L x^*(t_2) \tag{10.2}$$

The classical theory of monotone comparative statics [15, 16, 22, 23] seeks conditions on the function f that guarantee that eq. (10.2) holds.

Athey [3, 4, 5] examined a problem of monotone comparative statics in the presence of uncertainty, where the objective function is an integral of some function with respect to some measure. Specifically, let $(L, \geqslant_L)$ be a lattice, $B \subseteq L$ a choice set, $S = \prod_{i=1}^{m} S_i$ with $S_i \subseteq \mathbb{R}$ for $i = 1, \ldots, m$, $\Theta \subseteq \mathbb{R}$ a set of parameters, μ a finite nonnegative product measure on S, and $u : L \times S \to \mathbb{R}$ and $\psi : S \times \Theta \to \mathbb{R}$ given bounded measurable functions. Define the objective function $\Phi : L \times \Theta \to \mathbb{R}$ by

$$\Phi(x, \theta) = \int_S u(x, s)\, \psi(s, \theta)\, d\mu(s)$$

For the problem of choosing an $x \in B$ that maximizes the objective function given a value θ of the parameter, the problem of monotone comparative statics in this situation of uncertainty is to find conditions on the primitives u and ψ so that an optimal choice $x^*(\theta)$ of x is a nondecreasing function of θ, that is,

$$\theta_1 \geqslant \theta_2 \implies x^*(\theta_1) \geqslant_L x^*(\theta_2) \tag{10.3}$$

In particular, in both situations of certainty and uncertainty, the interest is in the variation of the optimal solution with respect to the lattice order $\geqslant_L$ on L, given a variation of the parameter (t or θ, respectively) in the order on the parameter set ($\geqslant_T$ or the usual order on $\mathbb{R}$, respectively). Often, however, these notions of order are too strong for the problem under consideration. For instance, in problems of the form (10.1), conditions on the primitive $\mathcal{U}$ for the optimal choice Y^* of Y to be monotone in X are desired. Specifically, under what conditions on $\mathcal{U}$ do we have that for all $s, s' \in S$, $X(s) \geqslant X(s') \Rightarrow Y^*(s) \geqslant Y^*(s')$? The classical techniques of monotone comparative statics are of no help in these situations since the lattice order on $B(\Sigma)$ is not adequate here. This order $\geqslant_L$ on $L = B(\Sigma)$ is defined by

$$Y_1 \geqslant_L Y_2 \quad \text{if and only if} \quad Y_1(s) \geqslant Y_2(s), \text{ for all } s \in S$$

10.2. **A Class of Monotone Comparative Statics Problems.** The problem that will be examined here takes the form

$$\sup_{Y=I\circ X}\left\{\int \mathcal{U}\left(X,Y\right)\ dP:\ 0\leqslant Y\leqslant X,\ V\left(Y\right)\geqslant R\right\} \tag{10.4}$$

where X is a given underlying uncertainty on the measurable space (S,Σ,P), $Y=I\circ X$ is a claim contingent on this uncertainty, $\int \mathcal{U}\left(X,I\circ X\right)\ dP$ is a DM's expected utility of wealth with respect to the probability measure P, $V:B^{+}\left(\Sigma\right)\rightarrow\mathbb{R}$ is some given mapping, and $R\in\mathbb{R}$ is fixed.

The first constraint is standard in many problems in economic theory. In the insurance framework [2, 11, 20], this constraint says that an indemnity is nonnegative and cannot exceed the loss itself. In a framework of debt contracting [10], this constraint is a limited liability constraint. The second constraint is interpreted as some "aggregation constraint". For instance, in problems of insurance demand, R would be the insurer's reservation utility, and $V\left(I\circ X\right)$ would be the insurer's expected utility of wealth with respect to his probability measure. The "aggregation constraint" is then simply the insurer's participation constraint, or individual rationality constraint.

The mapping $V:B^{+}\left(\Sigma\right)\rightarrow\mathbb{R}$ need not be law-invariant with respect to[5] P. When V is not law-invariant with respect to P, this creates some heterogeneity in the perception of the uncertainty X, and poses some important mathematical complications. For instance, in the insurance framework, it might be that the DM and the insurer assign different "distributions" to the underlying uncertainty.

This section's main result (Theorem 10.3) is that when the mapping V satisfies a property that will be called *Vigilance* (Definition 10.2) and a property that will be called the *Weak DC-Property* (Definition 10.1), supermodularity of the function $\mathcal{U}:\mathbb{R}^2\rightarrow\mathbb{R}$ (Definition B.3) is sufficient for an optimal choice of $Y=I\circ X$ to be a nondecreasing function of the underlying uncertainty X. Roughly speaking, vigilance of the operator V can be understood as a (weak) preference for comonotonicity with X (Definition 4.1), on the collection of all functions that are identically distributed for the probability measure P. Given two elements Y_1 and Y_2 of $B^{+}\left(\Sigma\right)$ that have the same distribution with respect to the probability measure P, vigilance of an operator $V:B^{+}\left(\Sigma\right)\rightarrow\mathbb{R}$ means that if any one of Y_1 or Y_2 is a nondecreasing function of X, it will assigned a higher value by V than the other function. The Weak DC-Property of an operator roughly means that the operator preserves dominated convergence. This property is satisfied by a large class of operators on $B^{+}\left(\Sigma\right)$, such as the Lebesgue integral or the Choquet integral (Appendix G).

Here, the definition of vigilance is extended from the notion of vigilant beliefs, introduced by Ghossoub [11], to the concept of a vigilant real-valued mapping ρ on the collection of functions Y over which a decision maker (DM) has a given preference. When $\rho\left(Y\right)=\int Y\ dP$, one will recover Ghossoub's [11] definition of vigilant beliefs as a special case of the definition of vigilance given here (Definition 10.2).

10.3. **The Setting.** As in Section 4, let S denote the set of states of the world, and suppose that $\mathcal{G}$ is a σ-algebra of subsets of S, called events. Denote by $B\left(\mathcal{G}\right)$ the supnorm-normed Banach

[5]A mapping $\Psi:B\left(\Sigma\right)\rightarrow\mathbb{R}$ is said to be *P-law-invariant*, or *law-invariant with respect to P*, if for any $\phi_1,\phi_2\in B\left(\Sigma\right)$, $\Psi\left(\phi_1\right)=\Psi\left(\phi_2\right)$ whenever ϕ_1 and ϕ_2 have the same distribution according to P.

space of all bounded, $\mathbb{R}$-valued and $\mathcal{G}$-measurable functions on $(S, \mathcal{G})$, and denote by $B^+(\mathcal{G})$ the collection of all $\mathbb{R}^+$-valued elements of $B(\mathcal{G})$. For any $f \in B(\mathcal{G})$, the suprnorm of f is given by $\|f\|_{sup} := \sup\{|f(s)| : s \in S\} < +\infty$. For $C \subseteq S$, denote by $\mathbf{1}_C$ the indicator function of C. For any $A \subseteq S$ and for any $B \subseteq A$, denote by $A \backslash B$ the complement of B in A.

For any $f \in B(\mathcal{G})$, denote by $\sigma\{f\}$ the σ-algebra of subsets of S generated by f, and denote by $B(\sigma\{f\})$ the linear space of all bounded, $\mathbb{R}$-valued and $\sigma\{f\}$-measurable functions on $(S, \mathcal{G})$. Then by Doob's measurability theorem [1, Theorem 4.41], for any $g \in B(\sigma\{f\})$ there exists a Borel-measurable map $\zeta : \mathbb{R} \to \mathbb{R}$ such that $g = \zeta \circ f$. Denote by $B^+(\sigma\{f\})$ the cone of nonnegative elements of $B(\sigma\{f\})$.

For any $f \in B(\mathcal{G})$, if $\mathcal{A}$ is any sub-σ-algebra of $\mathcal{G}$ such that $\sigma\{f\} \subseteq \mathcal{A}$, and if P is any probability measure on the measurable space $(S, \mathcal{A})$, it will be said that f is a continuous random variable for P when the law $P \circ f^{-1}$ of f is a nonatomic Borel probability measure. Recall that a finite nonnegative measure η on a measurable space $(\Omega, \mathcal{A})$ is said to be *nonatomic* if for any $A \in \mathcal{A}$ with $\eta(A) > 0$, there is some $B \in \mathcal{A}$ such that $B \subsetneq A$ and $0 < \eta(B) < \eta(A)$.

10.4. **Vigilant Operators and the Weak DC-Property.** Let P be a given probability measure on the measurable space $(S, \mathcal{G})$. In many situations of choice under uncertainty, the elements of choice are the elements of $B^+(\mathcal{G})$, as in the problem that will be examined in this paper. Often, a problem of choice involving these elements is stated as an optimization problem subject to some constraints. In an abstract form, some of these constraints can be stated in terms of operators $\rho : B^+(\mathcal{G}) \to \mathbb{R}$, and might be called "aggregation constraints". Here I will define two special kinds of these operators.

Definition 10.1 (Weak DC-Property)**.** An operator $\rho : B^+(\mathcal{G}) \to \mathbb{R}$ is said to have the *Weak DC-Property* if for any $Y^* \in B^+(\mathcal{G})$ and for any sequence $\{Y_n\}_{n \geqslant 1} \subset B^+(\mathcal{G})$ such that

(1) $\lim_{n \to +\infty} Y_n = Y^*$ (pointwise), and
(2) there is some $Z \in B^+(\mathcal{G})$ such that $Y_n \leqslant Z$, for each $n \geqslant 1$,

the following holds:

$$\lim_{n \to +\infty} \rho(Y_n) = \rho(Y^*)$$

When ρ is defined as a Lebesgue integral with respect to P, i.e. $\rho(Y) = \int Y \, dP$ for each $Y \in B^+(\mathcal{G})$, then Lebesgue's Dominated Convergence Theorem [9, Th. 2.4.4] implies that ρ has the Weak DC-Property. More generally, if ρ is a Choquet integral (Appendix G) with respect to some continuous capacity ν on $(S, \mathcal{G})$ (Definition G.2), i.e. $\rho(Y) = \oint Y \, d\nu$ for each $Y \in B^+(\mathcal{G})$, then when seen as an operator on $B^+(\mathcal{G})$, ρ has the Weak DC-Property. This is a consequence of [19, Th. 7.16].

Definition 10.2 (Vigilance)**.** Let X be a given element of $B^+(\mathcal{G})$, and recall that P is a probability measure on $(S, \mathcal{G})$. Denote by Σ the σ-algebra $\sigma\{X\}$ of subsets of S generated by X. An operator $\rho : B^+(\Sigma) \to \mathbb{R}$ is said to be (P, X)-*vigilant* if for any $Y_1, Y_2 \in B^+(\Sigma)$ such that

(i) Y_1 and Y_2 have the same distribution under P, i.e. $P \circ Y_1^{-1} = P \circ Y_2^{-1}$, and,
(ii) Y_2 is a nondecreasing function of X, i.e. Y_2 and X are comonotonic,

the following holds:

$$\rho(Y_2) \geqslant \rho(Y_1)$$

Clearly, if ρ is P-law invariant then it is (P, X)-vigilant. This covers a large collection of operators on $B^+(\Sigma)$ such that a Lebesgue integral with respect to P, a Choquet integral with respect to a distortion of P (Appendix G), and so on. When ρ is not P-law invariant, the same intuition as that behind Ghossoub's [11] definition of vigilance applies here.

10.5. **A Monotone Comparative Statics Result.** Let X be a given element of $B^+(\mathcal{G})$ with closed range $X(S) = [0, M]$, where $M := \|X\|_{sup} < +\infty$. Denote by Σ the σ-algebra $\sigma\{X\}$ of subsets of S generated by X, and let P be a probability measure on $(S, \mathcal{G})$. Let $\mathcal{U} : \mathbb{R}^2 \to \mathbb{R}$ be a given function, and let $V : B^+(\Sigma) \to \mathbb{R}$ be a given operator. The random variable X is fixed, and the objects P, $\mathcal{U}$, and V are considered to be the primitives of the following problem:

$$\sup_{Y \in B(\Sigma)} \left\{ \int \mathcal{U}(X, Y)\ dP : \ 0 \leqslant Y \leqslant X,\ V(Y) \geqslant R \right\} \tag{10.5}$$

where $R \in \mathbb{R}$ is fixed. The following theorem gives sufficient conditions for the optimal choice Y^* of the choice variable Y to be a nondecreasing function of X. The proof of the theorem is given in Appendix I.

Theorem 10.3. *If the following hold:*

1. *Problem 10.5 has a nonempty feasibility set,*
2. *The Borel probability measure $P \circ X^{-1}$ is nonatomic,*
3. *The mapping $\mathcal{U} : \mathbb{R}^2 \to \mathbb{R}$ is supermodular,*
4. *The mapping $\mathcal{E} : B^+(\Sigma) \to B(\Sigma)$ defined by $\mathcal{E}(Y) = \mathcal{U}(X, Y)$ is uniformly bounded and sequentially continuous in the topology of pointwise convergence*[6]*,*
5. *V is (P, X)-vigilant, and,*
6. *V has the Weak DC-Property,*

then Problem 10.5 admits a solution Y^* which is comonotonic with X. Moreover, any other solution Z^* which is comonotonic with X and identically distributed as Y^* under P is such that $Y^* = Z^*$, P-a.s.

A few comments on the assumptions in Theorem 10.3 are in order. First, the assumption of nonemptiness of the feasibility set of Problem 10.5 is made simply to rule out trivial cases where no solution can exist. The assumption of nonatomicity of the law of X is a technical requirement, and it means that the random variable X is diffused enough. This is a very common assumption in many instances, such as when it is assumed that a probability density function for X exists.

[6]That is, (i) there exists some $N < +\infty$ such that $\| \mathcal{E}(Y) \|_{sup} \leqslant N$ for each $Y \in B^+(\Sigma)$; and, (ii) if $\{Y_n\}_n$ is a sequence in $B^+(\Sigma)$ that converges pointwise to some $Y \in B^+(\Sigma)$, then the sequence $\{\mathcal{E}(Y_n)\}_n$ converges pointwise to $\mathcal{E}(Y)$.

The assumption of supermodularity of the mapping $\mathcal{U} : \mathbb{R}^2 \to \mathbb{R}$ is not a strong assumption by any means. It is usually given in many situations by the very nature of the problem considered. This happens for instance when $\mathcal{U}(X,Y) = u(a - X + Y)$, for a concave utility function u and some $a \in \mathbb{R}$. See Example B.4 (1). Assumption (4) in Theorem 10.3 is typically obtained whenever $\mathcal{U}(X,Y) = u(a - X + Y)$, for some continuous and bounded utility function u, and some $a \in \mathbb{R}$. Assumptions (5) and (6) were discussed in Section 10.4.

Appendix G. Related Analysis

G.1. A Useful Result.

Lemma G.1. *If $(f_n)_n$ is a uniformly bounded sequence of nondecreasing real-valued functions on some closed interval $\mathcal{I}$ in $\mathbb{R}$, with bound N (i.e. $|f_n(x)| \leqslant N$, $\forall x \in \mathcal{I}$, $\forall n \geqslant 1$), then there exists a nondecreasing real-valued bounded function f^* on $\mathcal{I}$, also with bound N, and a subsequence of $(f_n)_n$ that converges pointwise to f^* on $\mathcal{I}$.*

Proof. [7, Lemma 13.15]. □

G.2. Capacities and the Choquet Integral.

Definition G.2. A (normalized) *capacity* on a measurable space (S, Σ) is a set function $\nu : \Sigma \to [0,1]$ such that

(1) $\nu(\varnothing) = 0$;
(2) $\nu(S) = 1$; and,
(3) ν is monotone: for any $A, B \in \Sigma$, $A \subseteq B \Rightarrow \nu(A) \leqslant \nu(B)$.

The capacity ν is said to be

(1) *Continuous from above* if for any sequence $\{A_n\}_n$ in $\mathcal{G}$ such that $A_{n+1} \subseteq A_n$ for each $n \geqslant 1$, one has $\lim\limits_{n \to +\infty} \nu(A_n) = \nu\left(\bigcap_{n=1}^{+\infty} A_n\right)$.
(2) *Continuous from below* if for any sequence $\{A_n\}_n$ in $\mathcal{G}$ such that $A_n \subseteq A_{n+1}$ for each $n \geqslant 1$, one has $\lim\limits_{n \to +\infty} \nu(A_n) = \nu\left(\bigcup_{n=1}^{+\infty} A_n\right)$.
(3) *Continuous* if it is both continuous from above and continuous from below.

For instance, if P is a probability measure on (S, Σ) and $T : [0,1] \to [0,1]$ is increasing, with $T(0) = 0$ and $T(1) = 1$, then the set function $\nu := T \circ P$ is a capacity on (S, Σ). Such a function T is usually called a *probability distortion*, and the capacity $T \circ P$ is usually called a *distorted probability measure*. If, moreover, the function T is continuous, then the set function $\nu := T \circ P$ is a capacity on (S, Σ) which is continuous. This is an immediate consequence of the continuity of the measure P for monotone sequences [9, Prop. 1.2.3] and the continuity of T. In particular, any probability measure is continuous.

Definition G.3. For a given capacity ν on (S, Σ) and a given $\phi \in B(\Sigma)$, the *Choquet integral* of ϕ with respect to ν is defined by

$$\oint \phi \, d\nu := \int_0^{+\infty} \nu\left(\{s \in S : \phi(s) \geqslant t\}\right) \, dt + \int_{-\infty}^0 \left[\nu\left(\{s \in S : \phi(s) \geqslant t\}\right) - 1\right] \, dt$$

where the integrals are taken in the sense of Riemann.

The Choquet integral with respect to a measure is simply the usual Lebesgue integral with respect to that measure [14, p. 59]. Unlike the Lebesgue integral, however, the Choquet integral is not an additive operator on $B(\Sigma)$. However, the Choquet integral is additive on comonotonic functions (Definition 4.1). For more about capacities and Choquet integrals, I refer to Marinacci and Montrucchio [14].

Appendix H. Proofs of the Results of Section 8

H.1. **Proof of Theorem 8.8.** By Assumption 8.7,

$$\mathcal{F}_{SB} := \left\{Y \in B^+(\Sigma) : 0 \leqslant Y \leqslant X \text{ and } \int \mathcal{V}(Y) \, d\nu \geqslant R\right\} \neq \varnothing$$

Let $\mathcal{F}_{SB}^{\uparrow} := \left\{Y = I \circ X \in \mathcal{F}_{SB} : I \text{ is nondecreasing}\right\}$ denote the collection of all feasible $Y \in B^+(\Sigma)$ for Problem 8.4 which are also comonotonic with X.

Lemma H.1. *If ν is (μ, X)-vigilant, then $\mathcal{F}_{SB}^{\uparrow} \neq \varnothing$.*

Proof. Since $\mathcal{F}_{SB} \neq \varnothing$, choose any $Y = I \circ X \in \mathcal{F}_{SB}$, and let $\widetilde{Y}_\mu$ denote the nondecreasing μ-rearrangement of Y with respect to X. Then (i) $\widetilde{Y}_\mu = \widetilde{I} \circ X$ where $\widetilde{I}$ is nondecreasing, and (ii) $0 \leqslant \widetilde{Y}_\mu \leqslant X$, by Lemma B.6. Furthermore, since ν is (μ, X)-vigilant, it follows that $\int \mathcal{V}\left(\widetilde{Y}_\mu\right) \, d\nu \geqslant \int \mathcal{V}(Y) \, d\nu$, by definition of (μ, X)-vigilance. But $\int \mathcal{V}(Y) \, d\nu \geqslant R$ since $Y \in \mathcal{F}_{SB}$. Therefore, $\int \mathcal{V}\left(\widetilde{Y}_\mu\right) \, d\nu \geqslant R$. Thus, $\widetilde{Y}_\mu \in \mathcal{F}_{SB}^{\uparrow}$, and so $\mathcal{F}_{SB}^{\uparrow} \neq \varnothing$. ☐

Definition H.2. If $Y_1, Y_2 \in \mathcal{F}_{SB}$, Y_2 is said to be a Pareto improvement of Y_1 (or is Pareto-improving) when the following hold:

(1) $\int \mathcal{U}(X, Y_2) \, d\mu \geqslant \int \mathcal{U}(X, Y_1) \, d\mu$; and,

(2) $\int \mathcal{V}(Y_2) \, d\nu \geqslant \int \mathcal{V}(Y_1) \, d\nu$.

Lemma H.3. *Suppose that ν is (μ, X)-vigilant and that $\mathcal{U}(X, Y)$ is supermodular*[7]*. If $Y \in \mathcal{F}_{SB}$, then there is some $Y^* \in \mathcal{F}_{SB}^{\uparrow}$ which is Pareto-improving.*

[7] This happens for instance when $\mathcal{U}(X, Y) = u(a - X + Y)$, or $\mathcal{U}(X, Y) = u(a + X - Y)$, for a concave utility function u and some $a \in \mathbb{R}$. See Example B.4 (1) and (2).

Proof. First note that by Lemma H.1 $\mathcal{F}_{SB}^{\uparrow} \neq \varnothing$. Choose any $Y \in \mathcal{F}_{SB}$, and let $Y^* := \widetilde{Y}_\mu$, where $\widetilde{Y}_\mu$ denotes the nondecreasing μ-rearrangement of Y with respect to X. Then $Y^* \in \mathcal{F}_{SB}^{\uparrow}$, as in the proof of Lemma H.1. Moreover, since $\mathcal{U}(X,Y)$ is supermodular, it follows from Lemma B.5 that $\int \mathcal{U}(X,Y^*)\ d\mu \geqslant \int \mathcal{U}(X,Y)\ d\mu$. Finally, since ν is (μ,X)-vigilant, it follows from the definition of (μ,X)-vigilance that $\int \mathcal{V}(Y^*)\ d\nu \geqslant \int \mathcal{V}(Y)\ d\nu$. Therefore, $Y^* \in \mathcal{F}_{SB}^{\uparrow}$ is a Pareto improvement of $Y \in \mathcal{F}_{SB}$. □

Hence, by Lemma H.3, one can choose a maximizing sequence $\{Y_n\}_n$ in $\mathcal{F}_{SB}^{\uparrow}$ for Problem 8.4. That is, $\lim_{n\to+\infty} \int \mathcal{U}(X,Y_n)\ d\mu = N := \sup_{Y \in B^+(\Sigma)} \left\{\int \mathcal{U}(X,Y)\ d\mu\right\} < +\infty$. Since $0 \leqslant Y_n \leqslant X \leqslant M := \|X\|_{sup}$, the sequence $\{Y_n\}_n$ is uniformly bounded. Moreover, for each $n \geqslant 1$ one has $Y_n = I_n \circ X$, with $I_n : [0,M] \to [0,M]$. Consequently, the sequence $\{I_n\}_n$ is a uniformly bounded sequence of nondecreasing Borel-measurable functions. Thus, by Lemma G.1, there is a nondecreasing function $I^* : [0,M] \to [0,M]$ and a subsequence $\{I_m\}_m$ of $\{I_n\}_n$ such that $\{I_m\}_m$ converges pointwise on $[0,M]$ to I^*. Hence, I^* is also Borel-measurable, and so $Y^* := I^* \circ X \in B^+(\Sigma)$ is such that $0 \leqslant Y^* \leqslant X$. Moreover, the sequence $\{Y_m\}_m$, defined by $Y_m = I_m \circ X$, converges pointwise to Y^*. Thus, by Assumption 8.1 and by Lebesgue's Dominated Convergence Theorem, $Y^* \in \mathcal{F}_{SB}^{\uparrow}$. Now, by Assumption 8.1 and by Lebesgue's Dominated Convergence Theorem, one has

$$\int \mathcal{U}(X,Y^*)\ d\mu = \lim_{m\to+\infty} \int \mathcal{U}(X,Y_m)\ d\mu = \lim_{n\to+\infty} \int \mathcal{U}(X,Y_n)\ d\mu = N$$

Hence Y^* solves Problem 8.4. □

H.2. **Characterization of a Monotone Solution to Problem 8.4.** By Lebesgue's decomposition theorem [9, Theorem 4.3.1] there exists a unique pair (ν_{ac}, ν_s) of (nonnegative) finite measures on (S,Σ) such that $\nu = \nu_{ac} + \nu_s$, $\nu_{ac} << \mu$, and $\nu_s \perp \mu$. That is, for all $B \in \Sigma$ with $\mu(B) = 0$, one has $\nu_{ac}(B) = 0$, and there is some $A \in \Sigma$ such that $\mu(S\backslash A) = \nu_s(A) = 0$. It then also follows that $\nu_{ac}(S\backslash A) = 0$ and $\mu(A) = 1$. Note also that for all $Z \in B^+(\Sigma)$, $\int Z\ d\nu = \int_A Z\ d\nu_{ac} + \int_{S\backslash A} Z\ d\nu_s$. Furthermore, by the Radon-Nikodým theorem [9, Theorem 4.2.2] there exists a μ-a.s. unique Σ-measurable and μ-integrable function $h : S \to [0,+\infty)$ such that $\nu_{ac}(C) = \int_C h\ d\mu$, for all $C \in \Sigma$. Consequently, for all $Z \in B^+(\Sigma)$, $\int Z\ d\nu = \int_A Zh\ d\mu + \int_{S\backslash A} Z\ d\nu_s$. Moreover, since $\nu_{ac}(S\backslash A) = 0$, it follows that $\int_{S\backslash A} Z\ d\nu_s = \int_{S\backslash A} Z\ d\nu$. Thus, for all $Z \in B^+(\Sigma)$, $\int Z\ d\nu = \int_A Zh\ d\mu + \int_{S\backslash A} Z\ d\nu$. In particular, $\int Y\ d\nu = \int_A Yh\ d\mu + \int_{S\backslash A} Y\ d\nu$. In the following, the Σ-measurable set A on which μ is concentrated (and $\nu_s(A) = 0$) is assumed to be fixed all throughout. Finally, since $A \in \Sigma$ and since $X(S) = [0,M]$, $X(A)$ is a Borel subset of $[0,M]$, as previously discussed.

Lemma H.4. *Let Y^* be an optimal solution for Problem 8.4, and suppose that ν is (μ,X)-vigilant and that $\mathcal{U}(X,Y)$ is supermodular. Let $\widetilde{Y}^*_\mu$ be the nondecreasing μ-rearrangement of Y^* with respect to X. Then:*

(1) *$\widetilde{Y}^*_\mu$ is optimal for Problem 8.4; and,*

(2) *$\widetilde{Y}^*_\mu = \widetilde{Y}^*_{\mu,A}$, μ-a.s., where $\widetilde{Y}^*_{\mu,A}$ is the nondecreasing μ-rearrangement of Y^* with respect to X on A.*

Proof. Optimality of $\widetilde{Y}^*_\mu$ for Problem 8.4 is an immediate consequence of Lemmata B.5 and B.6, and of the (μ, X)-vigilance of ν. Now, let $\widetilde{Y}^*_{\mu,A}$ be the nondecreasing μ-rearrangement of Y^* with respect to X on A. Since $\mu(A) = 1$, then by Lemma B.2 one has that $\widetilde{Y}^*_\mu = \widetilde{Y}^*_{\mu,A}$, μ-a.s. □

Lemma H.5. *Let an optimal solution for Problem 8.4 be given by:*

$$Y^* = Y_1^* \mathbf{1}_A + Y_2^* \mathbf{1}_{S \setminus A} \tag{H.1}$$

for some $Y_1^*, Y_2^* \in B^+(\Sigma)$. Let $\widetilde{Y}^*_\mu$ be the nondecreasing μ-rearrangement of Y^* with respect to X, and let $Y^*_{1,\mu}$ be the nondecreasing μ-rearrangement of Y_1^* with respect to X. Then $\widetilde{Y}^*_\mu = \widetilde{Y}^*_{1,\mu}$, μ-a.s.

Proof. Let $\widetilde{Y}^*_{\mu,A}$ be the nondecreasing μ-rearrangement of Y^* with respect to X on A. Since $\mu(A) = 1$, then by Lemma B.2 one has $\widetilde{Y}^*_\mu = \widetilde{Y}^*_{\mu,A}$, μ-a.s. Similarly, let $\widetilde{Y}^*_{1,\mu,A}$ be the nondecreasing μ-rearrangement of Y_1^* with respect to X on A. Then $\widetilde{Y}^*_{1,\mu} = \widetilde{Y}^*_{1,\mu,A}$, μ-a.s. Therefore, it suffices to show that $\widetilde{Y}^*_{\mu,A} = \widetilde{Y}^*_{1,\mu,A}$, μ-a.s. Since both $\widetilde{Y}^*_{\mu,A}$ and $\widetilde{Y}^*_{1,\mu,A}$ are nondecreasing functions of X on A, then by the μ-a.s. uniqueness of the nondecreasing rearrangement, it remains to show that they are μ-equimeasurable with Y^* on A. Now, for each $t \in [0, M]$,

$$\begin{aligned}
\mu\Big(\{s \in A : \widetilde{Y}^*_{\mu,A}(s) \leqslant t\}\Big) &= \mu\Big(\{s \in A : Y^*(s) \leqslant t\}\Big) = \mu\Big(\{s \in A : Y_1^*(s) \leqslant t\}\Big) \\
&= \mu\Big(\{s \in A : \widetilde{Y}^*_{1,\mu,A}(s) \leqslant t\}\Big)
\end{aligned}$$

where the first equality follows from the definition of $\widetilde{Y}^*_{\mu,A}$ (equimeasurabiltiy), the second equality follows from equation (H.1), and the third equality follows from the definition of $\widetilde{Y}^*_{1,\mu,A}$ (equimeasurabiltiy). □

Consider now the following two problems:

Problem H.6. *For a given $\beta \in \mathbb{R}$,*

$$\sup_{Y \in B^+(\Sigma)} \left\{ \int_A \mathcal{U}(X, Y) \; d\mu \right\} :$$
$$\begin{cases} 0 \leqslant Y \mathbf{1}_A \leqslant X \mathbf{1}_A \\ \int_A \mathcal{V}(Y) \; d\nu = \beta \end{cases}$$

Problem H.7.

$$\sup_{Y \in B^+(\Sigma)} \left\{ \int_{S \setminus A} \mathcal{U}(X, Y) \; d\mu \right\} :$$
$$\begin{cases} 0 \leqslant Y \mathbf{1}_{S \setminus A} \leqslant X \mathbf{1}_{S \setminus A} \\ \int_{S \setminus A} \mathcal{V}(Y) \; d\nu \geqslant R - \beta, \; \textit{for the same } \beta \textit{ as in Problem H.6} \end{cases}$$

Remark H.8. By Remark 8.5, the supremum of each of the above two problems is finite when their feasibility sets are nonempty.

Definition H.9. For a given $\beta \in \mathbb{R}$, let:

(1) $\Theta_{A,\beta}$ be the feasibility set of Problem H.6 with parameter β. That is,

$$\Theta_{A,\beta} := \left\{ Y \in B^{+}(\Sigma) : 0 \leqslant Y\mathbf{1}_A \leqslant X\mathbf{1}_A,\ \int_A \mathcal{V}(Y)\ d\nu = \beta \right\}$$

(2) $\Theta_{S\setminus A,\beta}$ be the feasibility set of Problem H.7 with parameter β. That is,

$$\Theta_{S\setminus A,\beta} := \left\{ Y \in B^{+}(\Sigma) : 0 \leqslant Y\mathbf{1}_{S\setminus A} \leqslant X\mathbf{1}_{S\setminus A},\ \int_{S\setminus A} \mathcal{V}(Y)\ d\nu \geqslant R - \beta \right\}$$

Denote by Γ the collection of all β for which the feasibility sets $\Theta_{A,\beta}$ and $\Theta_{S\setminus A,\beta}$ are nonempty:

Definition H.10. Let $\Gamma := \left\{ \beta \in \mathbb{R} : \Theta_{A,\beta} \neq \varnothing,\ \Theta_{S\setminus A,\beta} \neq \varnothing \right\}$

Lemma H.11. *$\Gamma \neq \varnothing$.*

Proof. By Assumption 8.7, there is some $Y \in B^{+}(\Sigma)$ such that $0 \leqslant Y \leqslant X$, and $\int \mathcal{V}(Y)\ d\nu \geqslant R$. Let $\beta_Y := \int_A \mathcal{V}(Y)\ d\nu$. Then, by definition of β_Y, and since $0 \leqslant Y \leqslant X$, one has $Y \in \Theta_{A,\beta_Y} \cap \Theta_{S\setminus A,\beta_Y}$, and so $\Theta_{A,\beta_Y} \neq \varnothing$ and $\Theta_{S\setminus A,\beta_Y} \neq \varnothing$. Consequently, $\beta_Y \in \Gamma$. It then follows that $\Gamma \neq \varnothing$. ☐

Now, consider the following problem:

Problem H.12.

$$\sup_{\beta} \left\{ F_A^{*}(\beta) + F_A^{*}(R-\beta) : \beta \in \Gamma \right\} :$$

$$\begin{cases} F_A^{*}(\beta) \ \text{ is the supremum of Problem H.6, for a fixed } \beta \in \Gamma \\ F_A^{*}(R-\beta) \ \text{ is the supremum of Problem H.7, for the same fixed } \beta \in \Gamma \end{cases}$$

Lemma H.13. *If β^* is optimal for Problem H.12, Y_3^* is optimal for Problem H.6 with parameter β^*, and Y_4^* is optimal for Problem H.7 with parameter β^*, then $Y_2^* := Y_3^*\mathbf{1}_A + Y_4^*\mathbf{1}_{S\setminus A}$ is optimal for Problem 8.4.*

Proof. Feasibility of Y_2^* for Problem 8.4 is immediate. To show optimality of Y_2^* for Problem 8.4, let $\widetilde{Y}$ be any other feasible solution for Problem 8.4, and define $\alpha := \int_A \mathcal{V}\left(\widetilde{Y}\right)\ d\nu$. Then α is feasible for Problem H.12, and $\widetilde{Y}\mathbf{1}_A$ (resp. $\widetilde{Y}\mathbf{1}_{S\setminus A}$) is feasible for Problem H.6 (resp. Problem H.7) with parameter α. Hence

$$\begin{cases} F_A^{*}(\alpha) \geqslant \int_A \mathcal{U}\left(X, \widetilde{Y}\right)\ d\mu \\ F_A^{*}(R-\alpha) \geqslant \int_{S\setminus A} \mathcal{U}\left(X, \widetilde{Y}\right)\ d\mu \end{cases}$$

Now, since β^* is optimal for Problem H.12, it follows that

$$F_A^*(\beta^*) + F_A^*(R-\beta^*) \geqslant F_A^*(\alpha) + F_A^*(R-\alpha)$$

However,

$$\begin{cases} F_A^*(\beta^*) = \int_A \mathcal{U}(X, Y_3^*)\ d\mu \\ F_A^*(R-\beta^*) = \int_{S\backslash A} \mathcal{U}(X, Y_4^*)\ d\mu \end{cases}$$

Therefore, $\int \mathcal{U}\Big(X, Y_2^*\Big)\ d\mu \geqslant \int \mathcal{U}\Big(X, \widetilde{Y}\Big)\ d\mu$. Hence, Y_2^* is optimal for Problem 8.4. □

By Lemma H.13, one can restrict the analysis to solving Problems H.6 and H.7 with a parameter $\beta \in \Gamma$. By Lemmata H.4, H.5, and H.13, if ν is (μ, X)-vigilant, $\mathcal{U}(X,Y)$ is supermodular, β^* is optimal for Problem H.12, Y_1^* is optimal for Problem H.6 with parameter β^*, and Y_2^* is optimal for Problem H.7 with parameter β^*, then $\widetilde{Y}_\mu^*$ is optimal for Problem 8.4, and $\widetilde{Y}_\mu^* = \widetilde{Y}_{1,\mu}^*$, μ-a.s., where $\widetilde{Y}_\mu^*$ (resp. $\widetilde{Y}_{1,\mu}^*$) is the μ-a.s. unique nondecreasing μ-rearrangement of $Y^* := Y_1^*\mathbf{1}_A + Y_2^*\mathbf{1}_{S\backslash A}$ (resp. of Y_1^*) with respect to X.

Solving Problems H.6 and H.7. Since $\mu(S\backslash A) = 0$, it follows that, for all $Y \in B^+(\Sigma)$, one has $\int_{S\backslash A} \mathcal{U}(X,Y)\ d\mu = 0$. Consequently, any Y which is feasible for Problem H.7 with paramter β is also optimal for Problem H.7 with parameter β. Now, for a fixed parameter $\beta \in \Gamma$, Problem H.6 will be solved "statewise", as follows:

Lemma H.14. *If $Y^* \in B^+(\Sigma)$ satisfies the following:*

(1) $0 \leqslant Y^*(s) \leqslant X(s)$, *for all* $s \in A$;
(2) $\int_A \mathcal{V}(Y^*)\, h\ d\mu = \beta$; *and,*
(3) *There exists some* $\lambda \geqslant 0$ *such that for all* $s \in A$,

$$Y^*(s) = \underset{0 \leqslant y \leqslant X(s)}{\arg\max} \left[\mathcal{U}\Big(X(s), y\Big) - \lambda \mathcal{V}(y)\, h(s)\right] \tag{H.2}$$

Then the function Y^* solves Problem H.6 with parameter β.

Proof. Suppose that $Y^* \in B^+(\Sigma)$ satisfies (1), (2), and (3) above. Then Y^* is clearly feasible for Problem H.6. To show optimality of Y^* for Problem H.6 note that for any other $Y \in B^+(\Sigma)$ which is feasible for Problem H.6 with parameter β, one has, for all $s \in A$,

$$\mathcal{U}\Big(X(s), Y^*(s)\Big) - \mathcal{U}\Big(X(s), Y(s)\Big) \ \geqslant\ \lambda\Big[h(s)\,\mathcal{V}\Big(Y^*(s)\Big) - h(s)\,\mathcal{V}\Big(Y(s)\Big)\Big]$$

Consequently,

$$\int_A \mathcal{U}\Big(X, Y^*\Big)\ d\mu - \int_A \mathcal{U}\Big(X, Y\Big)\ d\mu \ \geqslant\ \lambda\Big[\beta - \beta\Big] = 0,$$

which completes the proof. □

The application of the general algorithm presented above depends on the specific forms of the functions $\mathcal{U}$ and $\mathcal{V}$. Depending on the nature of the problem considered, these functions can take different forms, and the algorithm can be carried out further.

Appendix I. Proof of Theorem 10.3

Suppose that $\mathcal{H} := \Big\{Y \in B(\Sigma) : 0 \leqslant Y \leqslant X \text{ and } V(Y) \geqslant R\Big\} \neq \varnothing$, $P \circ X^{-1}$ is nonatomic, V is (P, X)-vigilant and has the Weak DC-Property, the mapping $\mathcal{U} : \mathbb{R}^2 \to \mathbb{R}$ is supermodular, and the mapping $\mathcal{E} : B^+(\Sigma) \to B(\Sigma)$ defined by $\mathcal{E}(Y) = \mathcal{U}(X, Y)$ is uniformly bounded and sequentially continuous in the topology of pointwise convergence.

Lemma I.1. *For each $Y \in \mathcal{H}$ there is $Y^* \in \mathcal{H}$ such that:*

(1) *Y^* is comonotonic with X, i.e. Y^* is of the form $I^* \circ X$ where $I^* : [0, M] \to [0, M]$ is nondecreasing;*
(2) $\int \mathcal{U}\Big(X, Y^*\Big)\, dP \geqslant \int \mathcal{U}\Big(X, Y\Big)\, dP$*; and,*
(3) $V(Y^*) \geqslant V(Y)$.

Proof. By Assumption, $\mathcal{H} \neq \varnothing$. Choose any $Y = I \circ X \in \mathcal{H}$, and let $Y^* := \widetilde{Y}_P$, where $\widetilde{Y}_P$ denotes the nondecreasing P-rearrangement of Y with respect to X. Then (i) $Y^* = \widetilde{I} \circ X$ where $\widetilde{I} : [0, M] \to [0, M]$ is nondecreasing, bounded, and Borel-measurable; and, (ii) $0 \leqslant Y^* \leqslant X$, by Lemma B.6. Furthermore, since V is (P, X)-vigilant, it follows that $V(Y^*) \geqslant V(Y)$, by definition of (P, X)-vigilance. But $V(Y) \geqslant R$ since $Y \in \mathcal{H}$. Therefore, $V(Y^*) \geqslant R$. Thus, $Y^* \in \mathcal{H}$ is comonotonic with X. Moreover, since the function $\mathcal{U}$ is supermodular, it follows from Lemma B.5 that $\int \mathcal{U}\Big(X, Y^*\Big)\, dP \geqslant \int \mathcal{U}\Big(X, Y\Big)\, dP$. □

Now, let $\mathcal{H}^{\uparrow}$ denote the collection of all elements of $\mathcal{H}$ that are comonotonic with X. Then $\mathcal{H}^{\uparrow} \neq \varnothing$, by Lemma I.1. Also, by Lemma I.1, one can choose a maximizing sequence $\{Y_n\}_n$ in $\mathcal{H}^{\uparrow}$ for Problem 10.5. That is, $\lim\limits_{n \to +\infty} \int \mathcal{U}(X, Y_n)\, dP = N := \sup_{Y \in B^+(\Sigma)} \Big\{\int \mathcal{U}(X, Y)\, dP\Big\} < +\infty$. Since $0 \leqslant Y_n \leqslant X \leqslant M := \|X\|_{sup}$, the sequence $\{Y_n\}_n$ is uniformly bounded. Moreover, for each $n \geqslant 1$ one has $Y_n = I_n \circ X$, with $I_n : [0, M] \to [0, M]$. Consequently, the sequence $\{I_n\}_n$ is a uniformly bounded sequence of nondecreasing Borel-measurable functions. Thus, by Lemma G.1, there is a nondecreasing function $I^* : [0, M] \to [0, M]$ and a subsequence $\{I_m\}_m$ of $\{I_n\}_n$ such that $\{I_m\}_m$ converges pointwise on $[0, M]$ to I^*. Hence, I^* is also Borel-measurable, and so $Y^* := I^* \circ X \in B^+(\Sigma)$ is such that $0 \leqslant Y^* \leqslant X$. Moreover, the sequence $\{Y_m\}_m$, defined by $Y_m = I_m \circ X$, converges pointwise to Y^*. Thus, by the assumption that V has the *Weak DC-Property*, $Y^* \in \mathcal{H}^{\uparrow}$. Now, by the assumption of uniform boundedness and sequentially continuity of the map $\mathcal{U}(X, .)$ in the topology of pointwise convergence, and by Lebesgue's Dominated Convergence Theorem one has

$$\int \mathcal{U}(X, Y^*)\, dP = \lim_{m \to +\infty} \int \mathcal{U}(X, Y_m)\, dP = \lim_{n \to +\infty} \int \mathcal{U}(X, Y_n)\, dP = N$$

Hence Y^* solves Problem 10.5. The P-a.s. uniqueness of Y^* is a consequence of Proposition B.1. This concludes the proof of Theorem 10.3. □

References

[1] C.D. Aliprantis and K.C. Border. *Infinite Dimensional Analysis - 3rd edition.* Springer-Verlag, 2006.
[2] K.J. Arrow. *Essays in the Theory of Risk-Bearing.* Chicago: Markham Publishing Company, 1971.

[3] S. Athey. Characterizing Properties of Stochastic Objective Functions. mimeo (1999).
[4] S. Athey. Comparative Statics under Uncertainty: Single Crossing Properties and Log-Supermodularity. mimeo (1996).
[5] S. Athey. Monotone Comparative Statics under Uncertainty. *Quarterly Journal of Economics*, 117(1):187–223, 2002.
[6] K.H. Borch. *The Mathematical Theory of Insurance: An Annotated Selection of Papers on Insurance Published 1960-1972*. Lexington Books, 1974.
[7] N.L. Carothers. *Real Analysis*. Cambridge University Press, 2000.
[8] A. Chateauneuf, F. Maccheroni, M. Marinacci, and J.M. Tallon. Monotone Continuous Multiple Priors. *Economic Theory*, 26(4):973–982, 2005.
[9] D.L. Cohn. *Measure Theory*. Birkhauser, 1980.
[10] D. Gale and M. Hellwig. Incentive-Compatible Debt Contracts: The One-Period Problem. *The Review of Economic Studies*, 52(4):647–663, 1985.
[11] M. Ghossoub. Belief Heterogeneity in the Arrow-Borch-Raviv Insurance Model. mimeo (2011).
[12] S.J. Grossman and O.D. Hart. An Analysis of the Principal-Agent Problem. *Econometrica*, 51(1):7–45, 1983.
[13] B. Holmstrom. Moral Hazard and Observability. *The Bell Journal of Economics*, 10(1):74–91, 1979.
[14] M. Marinacci and L. Montrucchio. Introduction to the Mathematics of Ambiguity. In *I. Gilboa (ed.), Uncertainty in Economic Theory*. Routledge, London, 2004.
[15] P. Milgrom and J. Roberts. The Economics of Modern Manufacturing: Technology, Strategy, and Organization. *The American Economic Review*, 80(3):511–528, 1990.
[16] P. Milgrom and C. Shannon. Monotone Comparative Statics. *Econometrica*, 62(1):157–180, 1994.
[17] J.A. Mirrlees. The Theory of Moral Hazard and Unobservable Behaviour: Part I. *Review of Economic Studies*, 66(1):3–21, 1999.
[18] F.H. Page. The Existence of Optimal Contracts in the Principal-Agent model. *Journal of Mathematical Economics*, 16(2):157–167, 1987.
[19] E. Pap. *Null-Additive Set Functions*. Kluwer Academic Publishers, 1995.
[20] A. Raviv. The Design of an Optimal Insurance Policy. *The American Economic Review*, 69(1):84–96, 1979.
[21] W.P. Rogerson. The First-Order Approach to Principal-Agent Problems. *Econometrica*, 53(6):1357–1367, 1985.
[22] D.M. Topkis. Minimizing a Submodular Function on a Lattice. *Operations Research*, 26(2):305–321, 1978.
[23] D.M. Topkis. *Supermodularity and Complementarity*. Princeton University Press, 1998.
[24] R. Townsend. Optimal Contracts and Competitive Markets with Costly State Verification. *Journal of Economic Theory*, 21(2):265–293, 1979.
[25] C. Villegas. On Qualitative Probability σ-Algebras. *The Annals of Mathematical Statistics*, 35(4):1787–1796, 1964.
[26] X. Vives. *Oligopoly Pricing: Old Ideas and New Tools*. MIT Press, 2001.
[27] S.D. Williamson. Costly Monitoring, Financial Intermediation, and Equilibrium Credit Rationing. *Journal of Monetary Economics*, 18(2):159–179, 1986.

Université de Montréal – Département de Sciences Économiques – C.P. 6128, succursale Centre-ville – Montréal, QC, H3C 3J7 – Canada

E-mail address: mario.ghossoub@umontreal.ca