AMBIGUITY ON THE INSURER'S SIDE: THE DEMAND FOR INSURANCE

MASSIMILIANO AMARANTE
UNIVERSITÉ DE MONTRÉAL AND CIREQ

MARIO GHOSSOUB
IMPERIAL COLLEGE LONDON

EDMUND PHELPS
CENTER ON CAPITALISM AND SOCIETY, COLUMBIA UNIVERSITY

THIS DRAFT: OCTOBER 21, 2014

Corresponding Author: Mario Ghossoub, m.ghossoub@imperial.ac.uk.

Key Words and Phrases: Optimal Insurance, Deductible, Ambiguity, Choquet Integral, Distorted Probabilities.

JEL Classification: G22.

We thank Enrico Biffis, Daniel Gottlieb, John Quah, and an anonymous referee for their comments and suggestions. Mario Ghossoub acknowledges financial support from the Social Sciences and Humanities Research Council of Canada.

Abstract. Empirical evidence suggests that ambiguity is prevalent in insurance pricing and underwriting, and that often insurers tend to exhibit more ambiguity than the insured individuals (e.g., [23]). Motivated by these findings, we consider a problem of demand for insurance indemnity schedules, where the insurer has ambiguous beliefs about the realizations of the insurable loss, whereas the insured is an expected-utility maximizer. We show that if the ambiguous beliefs of the insurer satisfy a property of compatibility with the non-ambiguous beliefs of the insured, then there exist optimal monotonic indemnity schedules. By virtue of monotonicity, no *ex-post* moral hazard issues arise at our solutions (e.g., [25]). In addition, in the case where the insurer is either ambiguity-seeking or ambiguity-averse, we show that the problem of determining the optimal indemnity schedule reduces to that of solving an auxiliary problem that is simpler than the original one in that it does not involve ambiguity. Finally, under additional assumptions, we give an explicit characterization of the optimal indemnity schedule for the insured, and we show how our results naturally extend the classical result of Arrow [5] on the optimality of the deductible indemnity schedule.

1. Introduction

The classical formulation of the problem of demand for insurance indemnity schedules is due to Arrow [5]: a risk-averse Expected-Utility (EU) maximizing individual faces an insurable random loss X, against which he seeks an insurance coverage; and, a risk-neutral EU-maximizing insurer is willing to insure this individual against the realizations of the random loss, in return for an upfront premium payment. The insured seeks an indemnity schedule that maximizes his expected utility of final wealth, subject to the given premium Π determined by the insurer. Arrow's [5] classical theorem states that in this case, the optimal insurance indemnity schedule takes the form of full insurance above a constant positive deductible. That is, there exists a constant $d \geqslant 0$ such that the optimal insurance indemnity schedule is of the form

$$Y^* = \max\Big(0, X - d\Big).$$

This is a pure risk-sharing result: both parties have the same probabilistic beliefs, and the need for insurance is a consequence only of their different attitudes toward risk. Ghossoub [17] extended Arrow's result to the case of heterogeneous beliefs. Ghossoub's [17] result and *a fortiori* Arrow's [5] result apply, however, only to those situations where both parties have rich information about the relevant uncertainty, so as to be able to reduce that uncertainty to risk and form a probabilistic assessment. In contrast, empirical evidence suggests that ambiguity (as opposed to risk) is prevalent in insurance pricing and underwriting, and that often insurers tend to exhibit more ambiguity than the insured individuals (e.g., [23]). Motivated by these findings, we re-examine the classical insurance demand problem of Arrow [5] in a setting where the insurer has ambiguous beliefs (in the sense of Schmeidler [41]) about the realizations of the insurable loss, whereas the insured is an EU-maximizer.

Formally, we examine a problem similar to that of Arrow [5], with the sole difference that the beliefs of the insurer are represented by a capacity (Appendix A, Def. A.1) rather than a probability measure. Our results are as follows. First, we present a general result (Theorem 4.6), which states that if the parties' beliefs satisfy a certain compatibility condition (Def. 4.5), then optimal indemnity schedules exist and are monotonic. Here, monotonicity means that the optimal indemnity schedule is a nondecreasing function of the realizations of the insurable loss random variable. As it is well-known, this property rules out *ex post* moral hazard issues that could arise from the possibility that the insurer could misreport the actual amount of loss suffered (Huberman, Mayers and Smith [25]). This result complements a similar result that we obtained in [3] for a slightly different setting (which involves some minor technical differences).

We then consider the case where the insurer is either ambiguity-seeking or ambiguity-averse in the sense of Schmeidler [41]. We show that in both cases, an optimal indemnity schedule can be replicated by an optimal indemnity obtained from an insurance problem in which both the insured and the insurer are EU-maximizers, but have different beliefs about the realizations of the insurable random loss (Proposition 5.1 and Proposition 6.1). Such problems have been recently studied by Ghossoub [17].

Finally, under additional assumptions, we obtain an explicit characterization of the optimal indemnity schedule as a function of the underlying data. In the case of an ambiguity-seeking insurer whose capacity is a distortion of the probability measure of the insured, we show that the optimal indemnity schedule takes the form

$$Y^* = \min\left[X, \max\Big(0, X - d\,(T)\Big)\right],$$

where T is the concave probability distortion function of the insurer (see Appendix A), and $d\,(T)$ is a state-contingent deductible that depends on the state of the world only through the function T (Theorem 5.4). In the case of an ambiguity-averse insurer whose capacity has a core (Appendix A, Def. A.2) consisting of probability measures with the monotone likelihood ratio (MLR) property, we show that the optimal indemnity schedule is a state-contingent deductible of the form

$$Y^* = \min\left[X, \max\Big(0, X - d\,(LR)\Big)\right],$$

where LR denotes a function of the likelihood ratios of the probabilities in the core of the supermodular capacity over the probability of the insured (Corollary 6.3). In both cases, we determine the state-contingent deductible d explicitly. Arrow's solution obtains as a limit case from both settings: when the distortion function T becomes the identity function in the ambiguity-seeking case and when the core collapses to the probability measure of the insured in the ambiguity-averse case.

Related Literature. The literature on ambiguity in insurance design can be split into two main streams: (i) ambiguity on only one side of the insurance problem, and (ii) ambiguity on both sides. In the former category, all of the work that has been done has invariably assumed that the ambiguity is on the side of the insured. As such, it is very different from what we do in this paper. For instance, Alary et al. [1] consider an insured who is ambiguity-averse in the sense of Klibanoff, Marinacci, and Mukerji [31], and assume that the ambiguity is concentrated only in the probability that a loss occurs. Conditional on a loss occurring, the distribution of the loss severity is unambiguous. Under these assumptions, they show that the optimal indemnity is a straight deductible. Gollier [20] also focuses on the case of an insured who is ambiguity-averse in the sense of Klibanoff, Marinacci, and Mukerji [31]. He shows that if the collection of priors can be ordered according to the MLR property, then the optimal indemnity schedule contains a disappearing deductible. Jeleva [27] considers an insurance model in which the insurer is Choquet-Expected Utility (CEU) maximizer [41]. She specifies *ex ante* that the insurance contract is of the co-insurance type, and she then examines the optimal co-insurance factor. Young [46] and Bernard et al. [6] examine the case where the insured is a Rank-Dependent Expected-Utility maximizer [37, 45]. Doherty and Eeckhoudt [15] study the optimal level of deductible under Yaari's Dual Theory [45]. Karni [30] and Machina [32] consider a setting where the preferences of the insured have a non-EU representation that satisfies certain differentiability criteria. The former shows that a deductible indemnity schedule is optimal; whereas he latter examines the optimal level of co-insurance and optimal level of deductible. Schlesinger [39] examines the optimal co-insurance level in a situation where the preferences of the insured are not necessarily EU preferences, but they are risk-averse in the sense of disliking men-preserving increases in risk.

In the second stream of the literature on ambiguity in insurance design, which contemplates ambiguity on both sides, Carlier et al. [10] consider the case in which both parties' beliefs are epsilon-contaminations of a given prior, and they show that the optimal indemnity contains a deductible for high values of the loss. Anwar and Zheng [4] allow for both two-sided ambiguity and belief heterogeneity but restrict to a model with only two states of the world. As such, the scope of their inquiry is limited because, in general, financial and insurance risks are not binary risks (as they would necessarily be in a two-state model). Moreover, the shape of an optimal indemnity schedule cannot be determined in a two-state model where the loss X can take only two values[1]: L with probability p, and 0 with probability $1 - p$.

More general problems that are directly relevant to the insurance problem considered here have been examined by Carlier and Dana [7, 8, 9] and Chateauneuf et al. [12]. However, none of these studies provide a full characterization of the optimal insurance indemnity schedule, which is one of the main goals of the present paper.

2. Setup

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 $B(\mathcal{G})$ the linear 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 $B(\mathcal{G})$. Any $f \in B(\mathcal{G})$ is bounded, and we define its supnorm by $\|f\|_{sup} := \sup\{|f(s)| : s \in S\} < +\infty$.

Suppose that an individual has initial wealth W_0 and is facing an insurable random loss X, against which he seeks insurance. This random loss is 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. Then by Doob's measurability theorem [2, Theorem 4.41], for any $Y \in B(\Sigma)$ there exists a Borel-measurable map $I : \mathbb{R} \to \mathbb{R}$ such that $Y = I \circ X$. Denote by $B^+(\Sigma)$ the cone of nonnegative elements of $B(\Sigma)$. Let P be a probability measure on (S, Σ). We will make the following assumption all throughout.

Assumption 2.1. *The random loss X is a continuous random variable*[2] *on the probability space (S, Σ, P). That is, the Borel probability measure $P \circ X^{-1}$ is nonatomic*[3].

[1] At least if one imposes, as it is customary, the constraint that the indemnity be non-negative and not larger than the loss itself.

[2] This is a standard assumption, and it holds in many instances, such as when it is assumed that a probability density function for X exists.

[3] 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 $0 < \eta(B) < \eta(A)$.

The insured seeks an insurance coverage against this random loss X. He has access to an insurance market in which he can purchase an insurance indemnity $Y = I(X)$, which pays the amount $I(X(s)) \geqslant 0$, in the state of the world $s \in S$. By Doob's measurability theorem, we can identify $B^+(\Sigma)$ with the collection of all possible indemnity schedules. The price of this insurance indemnity schedule is called the *insurance premium*, and it is denoted by Π. The premium is determined by the insurer, based on his beliefs about the realizations of X (and hence of Y).

3. Insurance Demand with No Ambiguity: The Classical Case

In the classical insurance model of Arrow [5], both the insurer and the insured have non-ambiguous beliefs about the realizations of X and share the same probabilistic beliefs. Both individuals are EU-maximizers. The insurer is risk-neutral, having a linear utility function[4], whereas the insured is risk-averse, having a concave increasing utility function. Let u denote the utility function of the insured. In each state of the world $s \in S$, the wealth of the insured is given by $W(s) = W_0 - \Pi - X(s) + Y(s)$, where $Y = I \circ X$ is an indemnity function. The problem of the insured is that of determining which indemnity schedule maximizes his expected utility of wealth, for a given premium $\Pi > 0$. Specifically, the problem of the insured is that of choosing Y in $B^+(\Sigma)$ so as to maximize

$$\int u(W_0 - \Pi - X + Y)\, dP,$$

subject to the classical constraint that the indemnity function is nonnegative and does not exceed the loss, that is, $0 \leqslant Y \leqslant X$, and subject to the participation constraint of the insurer

$$\int \left(W_0^{ins} + \Pi - Y - c(Y)\right) dP \geqslant W_0^{ins},$$

where W_0^{ins} is the initial wealth of the insurer and $c(Y) = \rho.Y$ is a (proportional) cost function, with $\rho \geqslant 0$. The participation constraint of the insurer can then be re-written as

[4]Since utility functions are defined up to a positive linear transformation, it is usually assumed, without loss of generality, that the linear utility function of the insurer is simply the identity function.

the following *premium constraint*:

$$\Pi \geqslant (1+\rho) \int Y dP.$$

Arrow's [5] classical result states that in this case, the optimal insurance indemnity schedule is a deductible insurance schedule:

Theorem 3.1 (**Arrow**). *The optimal indemnity schedule is a deductible contract, with deductible level $d \geqslant 0$. That is, the optimal indemnity schedule is*

$$Y^* = \max\Big(0, X - d\Big).$$

Moreover, d is such that $\int Y^* dP = \frac{\Pi}{1+\rho}$.

4. Insurance Demand with Insurer Ambiguity

The assumption that both parties reduce uncertainty to risk severely limits the applicability of results such as Arrow's. Hogarth and Kunreuther [23] have shown that ambiguity is prevalent in insurance pricing and underwriting, and that the insurers tend to exhibit more ambiguity than the insured. Because of this, we are going to modify Arrow's model by assuming that the insurer has ambiguous beliefs about the realization of the random loss X, while we maintain the assumption that the insured is an EU-maximizer. Precisely, we assume that the insured is a risk-averse EU-maximizer with a concave utility function u that satisfies the following assumption.

Assumption 4.1. *The utility function u of the insured satisfies Inada's [26] conditions*[5]*:*

[5]Note the following:

Remark 4.2. The strict concavity and the continuous differentiability of u imply that first derivative u' is both continuous and strictly decreasing. The latter implies that $(u')^{-1}$ is continuous and strictly decreasing, by the Inverse Function Theorem [38, pp. 221-223]. Moreover, the continuity of u implies that u is bounded on every closed and bounded subset of $\mathbb{R}$.

(1) $u(0) = 0$;

(2) u *is strictly increasing and strictly concave;*

(3) u *is continuously differentiable; and,*

(4) *The first derivative satisfies* $u'(0) = +\infty$ *and* $\lim_{x \to +\infty} u'(x) = 0$.

In contrast, the insurer is a CEU-maximizer, as in Schmeidler[6] [41]. In this case, for a given indemnity schedule $Y \in B^+(\Sigma)$, the participation constraint of the insurer is given by

$$\int \left(W_0^{ins} + \Pi - (1+\rho)\, Y\right) d\nu \geqslant W_0^{ins}.$$

By letting $R = \frac{\Pi}{1+\rho} > 0$, this can be rewritten (see Proposition A.6) as the following premium constraint[7]:

$$R \geqslant -\int -Y d\nu \geqslant 0.$$

Thus the problem of determining the optimal indemnity schedules becomes as follow. The insured seeks an indemnity schedule that maximizes his expected utility of terminal wealth, subject to the above premium constraint, and to the classical constraint that the indemnity is nonnegative and does not exceed the full amount of the loss:

$$\sup_{Y \in B(\Sigma)} \left\{ \int u\left(W_0 - \Pi - X + Y\right)\ dP \ \middle|\ 0 \leqslant Y \leqslant X,\ -\int -Y\ d\nu \leqslant R \right\} \tag{4.1}$$

Remark 4.3. For any $Y \in B(\Sigma)$ which is feasible for Problem (4.1), one has $0 \leqslant Y \leqslant X$. Therefore, by monotonicity of the Choquet integral (Proposition A.6 (2)), it follows that

$$0 \leqslant \int Y\ d\nu \leqslant \int X\ d\nu \quad \text{and} \quad 0 \leqslant -\int -Y\ d\nu \leqslant -\int -X\ d\nu.$$

[6]All necessary background material about capacities and Choquet integration is given in Appendix A.
[7]Since the Choquet integral is only positively homogeneous (Proposition A.6 (3)), the premium constraint cannot be written as $R \geqslant \int Y d\nu$.

Now, there are two cases: either $R \geqslant -\int -X d\nu$ or $0 < R < -\int -X d\nu$. It is easily seen that when $R \geqslant -\int -X d\nu$, Problem (4.1) is solved at once, and that the solution Y^* is full insurance. Thus, we are going to focus on the remaining case, that is we are going to assume that

Assumption 4.4. $0 < R < -\int -X \, d\nu$.

In a Bayesian setting (i.e., with no ambiguity) monotonicity properties of optimal solutions are typically obtained by imposing the Monotone Likelihood Ratio (MLR) property (see Milgrom [34]). Unfortunately, the MLR property cannot be formulated in the case of capacities because of the lack of density functions. Ghossoub [17], however, has shown that in a Bayesian setting a property (strictly) weaker than the MLR property suffices. This has been extended to the case of a capacity in Amarante, Ghossoub, and Phelps [3], who refer to it as *vigilance*. This property expresses a form of compatibility between the probabilistic beliefs of the insured and the non-additive beliefs of the insurer. Here, we use a slightly different formulation from that of Amarante, Ghossoub and Phelps [3].

Definition 4.5 (Belief Compatiblity). Let ν be a capacity on Σ, let P be a probability measure on the same σ-algebra and let X be a random variable on (S, Σ). We say that ν is compatible with P (or P-compatible) if for any $Y_1, Y_2 \in B^+(\Sigma)$ such that

(i) Y_1 and Y_2 have the same distribution under P; and

(ii) Y_2 and X are comonotonic[8] (Y_2 is a nondecreasing function of X),

the following holds

$$-\int -Y_2 d\nu \leqslant -\int -Y_1 d\nu.$$

Clearly, any capacity $\nu = T \circ P$ that is a distortion of the probability measure P (see Appendix A) is compatible with P. In particular, any probability measure P is clearly

[8] See Definition A.5.

P-compatible. We refer to Amarante, Ghossoub, and Phelps [3] and to Appendix B for a list of examples of capacities that are P-compatible.

We can now state our main result. Its proof is given in Appendix D.

Theorem 4.6 (Existence and Monotonicity of Optimal Indemnity Schedules). *If ν is continuous (Definition A.3) and compatible with P, then Problem (4.1) admits a solution Y^* which is comonotonic with X; that is, there exists an optimal indemnity schedule which is a nondecreasing function of the random loss X. Moreover, any optimal indemnity schedule is necessarily comonotonic with X, except possibly on a set of probability zero under P.*

Theorem 4.6 is a general result. It says that when the ambiguous beliefs of the insurer are compatible with the non-ambiguous beliefs of insured in the sense of Definition 4.5, then the problem of existence of an optimal indemnity schedule always has a solution. Moreover, any such solution has to be a nondecreasing function of the insurable loss (except maybe on an event to which the insured assigns zero likelihood), thus ruling any *ex-post* moral hazard issues that might occur.

5. Insurance Demand with an Ambiguity-Seeking Insurer

In Theorem 4.6 we do not make any assumption on the insurer's attitude toward ambiguity. In this section, we focus on an ambiguity-seeking insurer. We are going to show that the problem of determining the optimal indemnity schedule reduces to that of solving an auxiliary problem which is simpler than the original one, in that it does not involve ambiguity. Furthermore, we are going to explicitly solve for the optimal indemnity schedule in the case where the beliefs of the insurer are represented by a distortion of a probability measure. In the next section, we will study the case of an ambiguity-averse insurer.

When the insurer is ambiguity-seeking, his capacity ν is submodular. A classical result of Schmeidler [40] states that there exists a non-empty, weak*-compact, and convex set $\mathcal{AC}$

of probability measures on (S, Σ) (called the anticore of ν) such that for all $Y \in B(\Sigma)$,

$$\int Y d\nu = \max_{Q \in \mathcal{AC}} \int Y dQ.$$

Hence, for each $Y \in B^+(\Sigma)$,

$$-\int -Y d\nu = -\max_{Q \in \mathcal{AC}} \int -Y dQ = \min_{Q \in \mathcal{AC}} \int Y dQ.$$

Consequently, the premium constraint of the ambiguity-seeking insurer can be written as $R \geqslant \min_{Q \in \mathcal{AC}} \int Y dQ$, which implies that for a given indemnity schedule Y, R has to be just enough to cover the lowest possible expectation of Y under the probability measures Q in the anticore of ν.

In this case, the problem of the insured becomes the following:

$$\sup_{Y \in B(\Sigma)} \left\{ \int u\left(W_0 - \Pi - X + Y\right) \, dP \,\middle|\, 0 \leqslant Y \leqslant X, \ \min_{Q \in \mathcal{AC}} \int Y dQ \leqslant R \right\}, \tag{5.1}$$

where $\mathcal{AC}$ is the anticore of the submodular capacity ν.

We now show that the solution of this problem reduces to the solution of a simpler problem, one that does not involve ambiguity but only involves heterogeneity of beliefs. For each $Q \in \mathcal{AC}$, consider the problem

$$\sup_{Y \in B(\Sigma)} \left\{ \int u\left(W_0 - \Pi - X + Y\right) \, dP \,\middle|\, 0 \leqslant Y \leqslant X, \ \int Y dQ \leqslant R \right\}, \tag{5.2}$$

That is, (5.2) is a problem similar to (5.1) but (ideally) involves an insurer who is an EU-maximizer, with $Q \in \mathcal{AC}$ being the probability representing the non-ambiguous beliefs of the insurer. If Q is compatible with P, then by Theorem 4.6, Problem (5.2) admits a solution that is comonotonic with X. Let us denote by $Y^*(Q)$ this optimal solution.

Proposition 5.1. *If the capacity ν is submodular with anticore $\mathcal{AC}$, and if every $Q \in \mathcal{AC}$ is compatible with P, then there exists a $Q^* \in \mathcal{AC}$ such that $Y^*(Q^*)$ is a solution to Problem* (5.1).

The proof of Proposition 5.1 is given in Appendix E. It is important to note that Proposition 5.1 does *not* say that an ambiguity-seeking insurer with a submodular capacity ν behaves just like an EU-maximizer with subjective probability measure Q^*: when presented with an indemnity schedule $Y \neq Y^*$, the insurer will evaluate Y using a probability measure $Q \neq Q^*$. Equivalently, Proposition 5.1 is only a statement that a maximum is obtained, but this simple observation buys us a lot of mileage. We refer the reader to Ghossoub [17] for an extensive inquiry into the properties of solutions to problems of optimal insurance design with non-ambiguous but heterogeneous beliefs.

The Case of a Probability Distortion. We now find an explicit solution for the optimal indemnity schedule. We do so under the additional assumption that the capacity of the insurer is a distortion of the probability measure of the insured, that is ν is of the form $\nu = T \circ P$, for a distortion function that satisfies the following properties.

Assumption 5.2. *The distortion function $T : [0,1] \to [0,1]$ is such that:*

- *T is concave;*
- *T is increasing and twice differentiable; and,*
- *$T(0) = 0$ and $T(1) = 1$.*

We also make the following assumption.

Assumption 5.3. *The insured has initial wealth $W_0 - \Pi$ such that $X \leqslant W_0 - \Pi$, P-a.s. That is, $P\left(\left\{s \in S : X(s) > W_0 - \Pi\right\}\right) = 0$.*

Assumption 5.3 simply states that the insured is well-diversified so that the particular exposure to X is sufficiently small, with respect to the total wealth of the insured.

The problem of the insured is the following:

$$\sup_{Y \in B(\Sigma)} \left\{ \int u\left(W_0 - \Pi - X + Y\right)\, dP \;\middle|\; 0 \leqslant Y \leqslant X,\ -\int -Y\, dT \circ P \leqslant R \right\}, \tag{5.3}$$

The following Theorem gives a full characterization of the optimal indemnity schedule for the insured. Its proof is given in Appendix F.

Theorem 5.4. *If assumptions 2.1, 4.1, 4.4, 5.2, and 5.3 hold, then the function Y^* defined by*

$$Y^* = \min\left[X, \max\Big(0, X - d\left(T\right)\Big)\right], \tag{5.4}$$

is an optimal solution for Problem (5.3), where

$$d\left(T\right) = W_0 - \Pi - \left(u'\right)^{-1}\Big(\lambda^* T'\left(U\right)\Big), \tag{5.5}$$

$U = F_X\left(X\right)$ *and* $\lambda^* \geqslant 0$ *is such that the second constraint of Problem 5.3 is binding at the optimum (whenever* $\lambda^* > 0$*).*

Equations (5.4)-(5.5) are sufficient for an optimal solution. The λ^* appearing in Theorem 5.4 is the Lagrange multiplier associated with the premium constraint (the participation constraint of the insurer), but in a restated version of Problem 5.3 (all details are in Appendix F).

Note that in Arrow's theorem (Theorem 3.1), the optimal indemnity takes the form $Y^* = \max\left(0, X - d\right)$, where $d \geqslant 0$. The positivity of d then implies that the optimal indemnity can be written as $Y^* = \min\Big[X, \max\left(0, X - d\right)\Big]$. Theorem 5.4 asserts that when the beliefs of the insurer are a concave distortion of the probability measure of the insured, the optimal indemnity takes a similar form, but in which the deductible is variable.

As one would expect, the state-contingent deductible $d(T)$ given in equation (5.5) depends also on the initial wealth level of the insured. The higher the initial wealth of the insured, the higher the deductible level, *ceteris paribus*. Therefore, the higher initial wealth of the risk-averse insured, the lower the amount of indemnity that the insured receives (given by equation (5.4)). In this sense, insurance is an inferior good, which is a well-known result in the classical theory of insurance. Similar findings in the classical insurance model with non-ambiguous beliefs are discussed by Moffet [35] and Mossin [36]. Moreover, the higher the premium paid by the risk-averse insured, the lower the deductible level, and the higher the amount of indemnification received by the insured. This is intuitive: the more a risk-averse individual pays for insurance, the higher the insurance coverage that the individual expects to receive.

Proposition 5.5 below further characterizes the optimal indemnity schedule. Its proof is given in Appendix G.

Proposition 5.5. *If the assumptions of Theorem 5.4 hold, then:*

(1) *If Y is a solution to Problem (5.3), then Y is comonotonic with X, except possibly on a set of probability zero under P.*

(2) *The optimal indemnity schedule Y^* given in eq. (5.4)-(5.5) is a nondecreasing function of the loss random variable X.*

(3) *If Z^* is any other indemnity schedule which is a nondecreasing function of X and is identically distributed as Y^*, then $Z^* = Y^*$, P-a.s.*

The final part of Proposition 5.5 is noteworthy. We refer to Gollier and Schlesinger [21] for the importance of characterizing the distribution of an optimal indemnity schedule rather than its actual shape.

The optimal indemnity schedule given in Theorem 5.4 satisfies another useful property. Recall that in the classical setting of Arrow, the optimal indemnity schedule is a deductible schedule of the form $Y = \max(X - d, 0)$, for some constant $d \geqslant 0$, the deductible level. In this case, if s_1 and s_2 are two states of the world such that $X(s_1) > d$ and $X(s_2) \leqslant d$, then it follows immediately that $X(s_1) > X(s_2)$. This kind of ordering is preserved in our more general setting:

Proposition 5.6. *Under the assumptions of Theorem 5.4, the optimal indemnity schedule Y^* given in eq. (5.4)-(5.5) satisfies the following property: If $s_1, s_2 \in S$ are such that $d(T)(s_1) < X(s_1)$ and $X(s_2) \leqslant d(T)(s_2)$, then it follows that $X(s_2) < X(s_1)$.*

The proof of Proposition 5.6 is given in Appendix H. When the insurer does not distort probabilities, the function T is the identity function, and we recover Arrow's [5] result:

Corollary 5.7. *If the insurer is a risk-neutral EU-maximizer and if the insured is a risk-averse EU-maximizer such that assumptions 2.1, 4.1, 5.3 and 4.4 hold, then the function Y^* defined by*

$$Y^* = \min\left[X, \max\Big(0, X - d\Big)\right] = \max\Big(0, X - d\Big),$$

is an optimal indemnity schedule for the insured, where

$$d = W_0 - \Pi - (u')^{-1}(\lambda^*) > 0,$$

and λ^* is chosen so that $\int Y^* \, dP = R = \frac{\Pi}{1+\rho}$.

The λ^* appearing in Corollary 5.7 is obtained similarly to the one appearing in Theorem 5.4. The proof of Corollary 5.7 is given in Appendix I.

6. Insurance Demand with an Ambiguity-Averse Insurer

This section mirrors the previous one but with an ambiguity-averse insurer replacing the ambiguity-seeking insurer. We are going to show that the problem of determining the optimal indemnity schedule reduces, just like before, to that of solving an auxiliary problem that does not involve ambiguity. Moreover, we are going to find an explicit solution for the optimal indemnity schedule by imposing an additional assumption.

Suppose now that the insurer is ambiguity-averse, that is, his capacity ν is supermodular. By a result of Schmeidler [40], there exists a non-empty, weak*-compact, and convex set $\mathcal{C}$ of probability measures on (S, Σ) (called the core of ν) such that for all $Y \in B(\Sigma)$,

$$\int Y \, d\nu = \min_{Q \in \mathcal{C}} \int Y \, dQ.$$

Hence, for each $Y \in B^+(\Sigma)$,

$$-\int -Y \, d\nu = -\min_{Q \in \mathcal{C}} \int -Y \, dQ = \max_{Q \in \mathcal{C}} \int Y \, dQ.$$

Consequently, the premium constraint of the ambiguity-averse insurer can be written as $R \geqslant \max_{Q \in \mathcal{C}} \int Y \, dQ$, which implies that for a given indemnity schedule Y, R has to be enough to cover the highest possible expectation of Y under the probability measures Q in the core of ν.

In this case, the problem of the insured becomes the following:

$$\sup_{Y \in B(\Sigma)} \left\{ \int u \left(W_0 - \Pi - X + Y\right) \, dP \;\middle|\; 0 \leqslant Y \leqslant X, \; \max_{Q \in \mathcal{C}} \int Y \, dQ \leqslant R \right\}, \tag{6.1}$$

where $\mathcal{C}$ is the core of the supermodular capacity ν.

We now give a characterization of the solution to Problem 6.1 similar to that of Section 5. For each $Q \in \mathcal{C}$, consider the following problem

$$\sup_{Y \in B(\Sigma)} \left\{ \int u \left(W_0 - \Pi - X + Y\right) \, dP \;\middle|\; 0 \leqslant Y \leqslant X, \; \int Y \, dQ \leqslant R \right\}, \tag{6.2}$$

That is, (6.2) is a problem similar to (6.1) but (ideally) involves an insurer who is an EU-maximizer, with $Q \in \mathcal{C}$ being the probability representing the non-ambiguous beliefs of the insurer. If Q is compatible with P, then by Theorem 4.6, Problem (6.2) admits a solution that is comonotonic with X. Let us denote by $Y^*(Q)$ this optimal solution.

Proposition 6.1. *If the capacity ν is supermodular with core $\mathcal{C}$, and if every $Q \in \mathcal{C}$ is compatible with P, then there exists a $Q^* \in \mathcal{C}$ such that $Y^*(Q^*)$ is a solution to Problem (6.1).*

The proof of Proposition 6.1 is given in Appendix J. Again, Proposition 6.1 shows that one can replicate the solution of Problem 6.1 by the solution of a problem with non-ambiguous but heterogeneous beliefs, such as the one studied by Ghossoub [17].

The MLR Case. Suppose now that the supermodular capacity ν with core $\mathcal{C}$ is such that all elements of $\mathcal{C}$ are absolutely continuous[9] with respect to the probability measure P. By the Radon-Nikodým Theorem [13, Th. 4.2.2], the core $\mathcal{C}$ of ν is isometrically isomorphic to the collection Ξ of corresponding Radon-Nikodým derivatives, where:

$$\Xi := \left\{ \phi_Q \;\middle|\; \phi_Q \circ X = \frac{dQ}{dP},\ Q \in \mathcal{C} \right\}. \tag{6.3}$$

Definition 6.2 (Monotone Likelihood Ratio). The core $\mathcal{C}$ of ν is said to have the *Monotone Likelihood Ratio (MLR) property with respect to P* if each $\phi_Q \in \Xi$ is a nonincreasing function.

The following result states that whenever the core of ν has the MLR property with respect to P, the optimal indemnity schedule is a state-dependent deductible.

Corollary 6.3. *Suppose that assumptions 2.1, 4.1, and 5.3 hold, and suppose that the capacity ν is supermodular with core $\mathcal{C}$. Let Ξ denote the set of corresponding Radon-Nikodým derivatives. If $\mathcal{C}$ has the MLR property with respect to P, and if all functions*

[9] A probability measure Q on the measurable space (S, Σ) is said to be absolutely continuous with respect to the probability measure P on (S, Σ) if for any $A \in \Sigma$, $P(A) = 0 \Longrightarrow Q(A) = 0$.

$\phi_Q \in \Xi$ *are continuous, then there exists* $Q^* \in \mathcal{C}$ *such that the function*

$$Y^* = \begin{cases} X & \text{if } R \geqslant \int X dQ^*, \\ \min\left[X, \max\left(0, X - d\left(\phi_{Q^*}\right)\right)\right] & \text{if } R < \int X dQ^*, \end{cases} \tag{6.4}$$

is a solution to Problem (6.1)*, where:*

- $d\left(\phi_{Q^*}\right) = W_0 - \Pi - \left(u'\right)^{-1}\left(\lambda^*\ \phi_{Q^*} \circ X\right)$*;*
- $\phi_{Q^*} \circ X = dQ^*/dP$*; and,*
- $\lambda^* \geqslant 0$ *is chosen such that* $\int Y^*\ dQ^* = R$*.*

Equation (6.4) is sufficient for an optimal solution. The proof of Corollary 6.3 is given in Appendix K. Obviously, one recovers Arrow's result from Theorem 6.3 when $\mathcal{C}$ consists of only the probability measure P. By an argument similar to that of Proposition 5.5, we also have the following result.

Proposition 6.4. *If the assumptions of Corollary 6.3 hold, then:*

(1) *If* Y *is a solution to Problem* (6.1)*, then* Y *is comonotonic with* X*, except possibly on a set of probability zero under* P*.*

(2) *The optimal indemnity schedule* Y^* *given in eq.* (6.4) *is a nondecreasing function of the loss random variable* X*.*

(3) *If* Z^* *is any other indemnity schedule which is a nondecreasing function of* X *and is identically distributed as* Y^**, then* $Z^* = Y^*$*,* P*-a.s.*

Finally, Proposition 6.5 below parallels Proposition 5.6.

Proposition 6.5. *Under the assumptions of Corollary 6.3, the optimal indemnity schedule* Y^* *given in eq.* (6.4) *satisfies the following property: If* $s_1, s_2 \in S$ *are such that* $d\left(\phi_{Q^*}\right)(s_1) < X(s_1)$ *and* $X(s_2) \leqslant d\left(\phi_{Q^*}\right)(s_2)$*, then it follows that* $X(s_2) < X(s_1)$*.*

7. Conclusion

In this paper, we re-examined Arrow's [5] classical problem of insurance demand by introducing ambiguity on the side of the insurer. We showed that if the ambiguous beliefs of the insurer satisfy a property of compatibility with the non-ambiguous beliefs of the insured, then there exists an optimal indemnity schedule. Moreover, optimal indemnity schedules are nondecreasing functions of the insurable random loss. We also showed that if the insurer is either ambiguity-seeking or ambiguity-averse in the sense of Schmeidler [41], then an optimal indemnity schedule can be replicated by one arising from a problem where both parties have non-ambiguous but heterogeneous beliefs. Under additional assumptions, we found an explicit form for the optimal indemnity schedule – both the in case of ambiguity-aversion and of ambiguity-seeking – and showed that our results are natural extensions of those of Arrow [5].

Appendix A. Capacities and Choquet Integration

Definition A.1. 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:

- *supermodular* if $\nu(A \cup B) + \nu(A \cap B) \geqslant \nu(A) + \nu(B)$, for all $A, B \in \Sigma$; and,
- *submodular* if $\nu(A \cup B) + \nu(A \cap B) \leqslant \nu(A) + \nu(B)$, for all $A, B \in \Sigma$.

For instance, if (S, Σ, P) is a probability space and $T : [0,1] \to [0,1]$ is an increasing function, such that $T(0) = 0$ and $T(1) = 1$, then the set function $\nu := T \circ P$ is a capacity on (S, Σ) called a *distorted probability measure*. The function T is usually called a *probability distortion*. If, moreover, the distortion function T is convex (resp. concave), then the capacity $\nu = T \circ P$ is supermodular (resp. submodular) [14, Ex. 2.1].

Definition A.2. Let ν_1 be a supermodular capacity and ν_2 a submodular capacity on (S, Σ).

- The *core* of ν_1, denoted by *core* (ν_1), is the collection of all probability measures Q on (S, Σ) such that $Q(A) \geqslant \nu(A), \forall A \in \Sigma$.
- The *anti-core* of ν_2, denoted by *acore* (ν), is the collection of all probability measures Q on (S, Σ) such that $Q(A) \leqslant \nu(A), \forall A \in \Sigma$.

Definition A.3. A capacity ν on (S, Σ) is continuous from above (resp. below) if for any sequence $\{A_n\}_{n \geqslant 1} \subseteq \Sigma$ such that $A_{n+1} \subseteq A_n$ (resp. $A_{n+1} \supseteq A_n$) for each n, it holds that

$$\lim_{n \to +\infty} \nu(A_n) = \nu\left(\bigcap_{n=1}^{+\infty} A_n\right) \quad \left(\text{resp. } \lim_{n \to +\infty} \nu(A_n) = \nu\left(\bigcup_{n=1}^{+\infty} A_n\right)\right)$$

A capacity that is continuous both from above and below is said to be continuous.

For instance, if ν is a distorted probability measure of the form $T \circ P$ where T is a continuous function, then ν is a continuous capacity.

Definition A.4. Let ν be a capacity on (S, Σ). The *Choquet integral* of $Y \in B(\Sigma)$ with respect to ν is defined by

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

where the integrals are taken in the sense of Riemann.

Definition A.5. Two functions $Y_1, Y_2 \in B(\Sigma)$ are said to be comonotonic if

$$\Big[Y_1(s) - Y_1(s')\Big]\Big[Y_2(s) - Y_2(s')\Big] \geqslant 0, \text{ for all } s, s' \in S.$$

For instance any $Y \in B(\Sigma)$ is comonotonic with any $c \in \mathbb{R}$. Moreover, if $Y_1, Y_2 \in B(\Sigma)$, and if Y_2 is of the form $Y_2 = I \circ Y_1$, for some Borel-measurable function I, then Y_2 is comonotonic with Y_1 if and only if the function I is nondecreasing.

The Choquet integral with respect to a (countably additive) measure is the usual Lebesgue integral with respect to that measure [33, p. 59]. Unlike the Lebesgue integral, the Choquet integral is not an additive operator on $B(\Sigma)$. However, the Choquet integral is additive over comonotonic functions.

Proposition A.6. *Let ν be a capacity on (S, Σ).*

(1) *If $\phi_1, \phi_2 \in B(\Sigma)$ are comonotonic, then $\int (\phi_1 + \phi_2)\, d\nu = \int \phi_1\, d\nu + \int \phi_2\, d\nu$.*
(2) *If $\phi_1, \phi_2 \in B(\Sigma)$ are such that $\phi_1 \leqslant \phi_2$, then $\int \phi_1\, d\nu \leqslant \int \phi_2\, d\nu$.*
(3) *For all $\phi \in B(\Sigma)$ and all $c \geqslant 0$, then $\int c\phi\, d\nu = c \int \phi\, d\nu$.*
(4) *If ν is submodular, then for any $\phi_1, \phi_2 \in B(\Sigma)$, $\int (\phi_1 + \phi_2)\, d\nu \leqslant \int \phi_1\, d\nu + \int \phi_2\, d\nu$.*

Appendix B. P-Compatible Capacities

In robust statistics, capacities have a long history (e.g. [24]). An important class of capacities used in robust statistics is the class of *symmetric* capacities introduced by Wasserman and Kadane [44] and Kadane and Wasserman [29].

Definition B.1. Let (S, Σ, P) be a probability space. A capacity ν on (S, Σ) is said to be:

(1) *Wasserman-Kadane P-symmetric* if for any two random variables Z_1 and Z_2 on (S, Σ, P) that are identically distributed for P, one has $\int Z_1 d\nu = \int Z_2 d\nu$.
(2) *Weakly P-symmetric* if for any $A, B \in \Sigma$, one has: $P(A) = P(B) \Rightarrow \nu(A) = \nu(B)$.

The probability measure P is clearly Wasserman-Kadane P-symmetric and weakly P-symmetric. Another example of a weakly P-symmetric capacity is a distorted probability measure of the form $T \circ P$.

All probability measures on the measurable space (S, Σ) that are either Wasserman-Kadane P-symmetric or weakly P-symmetric are indeed P-compatible:

Proposition B.2. *Let (S, Σ, P) be a probability space. If ν is a capacity on (S, Σ) such that ν is either Wasserman-Kadane P-symmetric or weakly P-symmetric, then ν is compatible with P. In particular, the probability measure P, and any distortion $T \circ P$ of the probability measure P are compatible with P.*

Proof. Let $Y_1, Y_2 \in B^+(\Sigma)$ be identically distributed under P. Then so are $-Y_1$ and $-Y_2$. In particular, $P[-Y_1 \geqslant t] = P[-Y_2 \geqslant t]$, for all $t \in \mathbb{R}$. Then, if ν is weakly P-symmetric, we also have $\nu[-Y_1 \geqslant t] = \nu[-Y_2 \geqslant t]$, for all $t \in \mathbb{R}$. Hence, $\int -Y_2 d\nu = \int -Y_1 d\nu$, and so $-\int -Y_2 d\nu = -\int -Y_1 d\nu$. If ν is Wasserman-Kadane P-symmetric, then we have $\int -Y_2 d\nu = \int -Y_1 d\nu$. In both cases, we have $-\int -Y_2 d\nu = -\int -Y_1 d\nu$. Therefore, in both cases, we have

$$-\int -Y_2 d\nu = -\int -Y_1 d\nu.$$

Consequently, ν is compatible with P. Finally, since any distorted probability measure of the form $T \circ P$ is a P-symmetric capacity, it follows that any distortion $T \circ P$ of the probability measure P is compatible with P. □

Another important class of capacities that has also been used in robust statistics is the class of *coherent* capacities, or *upper probabilities* (see Walley [42] and Kadane and Wasserman [29]).

Definition B.3. Let (S, Σ) be a measurable space. A capacity ν on (S, Σ) is said to be *coherent* if there exists a nonempty collection $\mathcal{C}$ of probability measures on (S, Σ) such that

$$\nu(A) = \sup_{Q \in \mathcal{C}} Q(A), \quad \text{for all} \quad A \in \Sigma \tag{B.1}$$

By a result of Schmeidler [40], a subclass of coherent capacities is the collection of all capacities that are submodular: any submodular capacity can be represented as coherent capacities of the form (B.1), where the set $\mathcal{C}$ is the *anticore* of ν.

Proposition B.4. *If ν is a submodular (and hence coherent) capacity on (S, Σ, P) such that each element of the anticore of ν is either weakly P-symmetric or Wasserman-Kadane P-symmetric, then ν is compatible with P.*

Proof. Denote by $\mathcal{C}$ the anticore of ν. Let $Y_1, Y_2 \in B^+(\Sigma)$ be identically distributed under P. In particular, $P[Y_1 \geqslant t] = P[Y_2 \geqslant t]$, for all $t \in \mathbb{R}$. Then:

(1) If each element of the anticore of ν is weakly P-symmetric, then we also have $Q[Y_1 \geqslant t] = Q[Y_2 \geqslant t]$, for all $t \in \mathbb{R}$ and for all $Q \in \mathcal{C}$. Hence, $\int Y_2 dQ = \int Y_1 dQ$, for all $Q \in \mathcal{C}$.

(2) If each element of the anticore of ν is Wasserman-Kadane P-symmetric, then we have $\int Y_2 dQ = \int Y_1 dQ$, for all $Q \in \mathcal{C}$.

In both cases, we have $\int -Y_2 dQ = \int -Y_1 dQ$, for all $Q \in \mathcal{C}$. Therefore, since ν is submodular, with anticore $\mathcal{C}$, we have

$$\int -Y_2 d\nu = \max_{Q \in \mathcal{C}} \int -Y_2 dQ = \max_{Q \in \mathcal{C}} \int -Y_1 dQ = \int -Y_1 d\nu.$$

Therefore, ν is compatible with P. □

A similar result holds for supermodular capacities, an important example of *lower probabilities*. By a result of Schmeidler [40], if ν is a supermodular capacity with core $\mathcal{C}$, then

$$\nu(A) = \inf_{Q \in \mathcal{C}} Q(A), \quad \text{for all} \quad A \in \Sigma.$$

Proposition B.5. *If ν is a convex capacity on (S, Σ, P) such that each element of the core of ν is either weakly P-symmetric or Wasserman-Kadane P-symmetric, then ν is compatible with P.*

The proof of Proposition B.5 is similar to that of Proposition B.4.

Appendix C. Rearrangements and Supermodularity

Here, the idea of an equimeasurable rearrangement of a random variable with respect to another random variable is discussed. All proofs, additional results and references to the literature may be found in Ghossoub [18, 19].

C.1. **The Nondecreasing Rearrangement.** Consider the setting of Section 2, and let ζ be the probability law of X defined by $\zeta(B) := P \circ X^{-1}(B) = P\Big(\Big\{s \in S : X(s) \in B\Big\}\Big)$, for any Borel subset B of $\mathbb{R}$.

Definition C.1. For any Borel-measurable map $I : [0, M] \to \mathbb{R}$, define the *distribution function* of I as the map $\zeta_I : \mathbb{R} \to [0, 1]$ defined by

$$\zeta_I(t) := \zeta\Big(\Big\{x \in [0, M] : I(x) \leqslant t\Big\}\Big). \tag{C.1}$$

Then ζ_I is a nondecreasing right-continuous function.

Definition C.2. Let $I : [0, M] \to [0, M]$ be any Borel-measurable map and define the function $\widetilde{I} : [0, M] \to \mathbb{R}$ by

$$\widetilde{I}(t) := \inf\Big\{z \in \mathbb{R}^+ \;\Big|\; \zeta_I(z) \geqslant \zeta\big([0, t]\big)\Big\}. \tag{C.2}$$

The following proposition gives some useful properties of the map $\widetilde{I}$ defined above.

Proposition C.3. *Let $I : [0, M] \to [0, M]$ be any Borel-measurable map and let $\widetilde{I}$: $[0, M] \to \mathbb{R}$ be defined as in equation (C.2). Then the following hold:*

(1) *$\widetilde{I}$ is left-continuous, nondecreasing and Borel-measurable;*
(2) *$\widetilde{I}(0) = 0$ and $\widetilde{I}(M) \leqslant M$. Therefore $\widetilde{I}([0, M]) \subseteq [0, M]$;*
(3) *If $I_1, I_2 : [0, M] \to [0, M]$ are such that $I_1 \leqslant I_2$, ζ-a.s., then $\widetilde{I}_1 \leqslant \widetilde{I}_2$;*

(4) *$\widetilde{I}$ is ζ-equimeasurable with I, in the sense that for any Borel set B,*

$$\zeta\left(\left\{t \in [0, M] : I(t) \in B\right\}\right) = \zeta\left(\left\{t \in [0, M] : \widetilde{I}(t) \in B\right\}\right); \tag{C.3}$$

(5) *If $\overline{I} : [0, M] \to \mathbb{R}^+$ is another nondecreasing, Borel-measurable map which is ζ-equimeasurable with I, then $\overline{I} = \widetilde{I}$, ζ-a.s.*

$\widetilde{I}$ is called the nondecreasing ζ-rearrangement of I. Now, define $Y := I \circ X$ and $\widetilde{Y} := \widetilde{I} \circ X$. Then:

(1) $Y, \widetilde{Y} \in B^+(\Sigma)$, since I and $\widetilde{I}$ are Borel-measurable mappings of $[0, M]$ into itself;

(2) $\widetilde{Y}$ is a nondecreasing function of X:
$\left[X(s) \leqslant X(s')\right] \Rightarrow \left[\widetilde{Y}(s) \leqslant \widetilde{Y}(s')\right]$, for all $s, s' \in S$; and,

(3) Y and $\widetilde{Y}$ have the same distribution under P (i.e., they are P-equimeasurable):
$P\left(\left\{s \in S : Y(s) \leqslant \alpha\right\}\right) = P\left(\left\{s \in S : \widetilde{Y}(s) \leqslant \alpha\right\}\right)$, for any $\alpha \in [0, M]$.

Call $\widetilde{Y}$ a nondecreasing P-rearrangement of Y with respect to X and denote it by $\widetilde{Y}_P$. Then $\widetilde{Y}_P$ is P-a.s. unique. Note also that if Y_1 and Y_2 are P-equimeasurable; and, for any Borel-measurable function ψ, $\psi(Y_1)$ is P-integrable if and only if $\psi(Y_2)$ is P-integrable, in which case we have $\int \psi(Y_1)\ dP = \int \psi(Y_2)\ dP$.

Lemma C.4. *Fix $Y \in B^+(\Sigma)$ and let $\widetilde{Y}_P$ denote the nondecreasing P-rearrangement of Y with respect to X. If $0 \leqslant Y \leqslant X$, P-a.s., then $0 \leqslant \widetilde{Y}_P \leqslant X$.*

C.2. **Supermodularity.**

Definition C.5. A function $L : \mathbb{R}^2 \to \mathbb{R}$ is *supermodular* if for any $x_1, x_2, y_1, y_2 \in \mathbb{R}$ with $x_1 \leqslant x_2$ and $y_1 \leqslant y_2$, one has

$$L(x_2, y_2) + L(x_1, y_1) \geqslant L(x_1, y_2) + L(x_2, y_1). \tag{C.4}$$

A function $L : \mathbb{R}^2 \to \mathbb{R}$ is called *strictly supermodular* if for any $x_1, x_2, y_1, y_2 \in \mathbb{R}$ with $x_1 < x_2$ and $y_1 < y_2$, one has

$$L(x_2, y_2) + L(x_1, y_1) > L(x_1, y_2) + L(x_2, y_1). \tag{C.5}$$

Lemma C.6. *A function $L : \mathbb{R}^2 \to \mathbb{R}$ is supermodular (resp. strictly supermodular) if and only if the function $\eta(y) := L(x + h, y) - L(x, y)$ is nondecreasing (resp. increasing) on $\mathbb{R}$, for any $x \in \mathbb{R}$ and $h \geqslant 0$ (resp. $h > 0$).*

Example C.7. If $g : \mathbb{R} \to \mathbb{R}$ is concave and $a \in \mathbb{R}$, then the function $L_1 : \mathbb{R}^2 \to \mathbb{R}$ defined by $L_1(x, y) = g(a - x + y)$ is supermodular. If, moreover, g is strictly concave then L_1 is strictly supermodular.

Lemma C.8 (Hardy-Littlewood). *Fix $Y \in B^+(\Sigma)$ and let $\widetilde{Y}_P$ denote the nondecreasing P-rearrangement of Y with respect to X. If L is supermodular, then (assuming integrability) we have*

$$\int L\Big(X, Y\Big)\, dP \leqslant \int L\left(X, \widetilde{Y}_P\right)\, dP.$$

Moreover, if L is strictly supermodular then equality holds if and only if $Y = \widetilde{Y}_P$, P-a.s.

Appendix D. Proof of Theorem 4.6

Let us denote by $\mathcal{F}_{SB}$ the feasibility set for Problem (4.1) (which we assume nonempty to rule out trivial situations): $\mathcal{F}_{SB} = \Big\{Y \in B(\Sigma) : 0 \leqslant Y \leqslant X \text{ and } -\int -Y\, d\nu \leqslant R\Big\}$. Let $\mathcal{F}_{SB}^{\uparrow}$ be the set of all the $Y \in \mathcal{F}_{SB}$ which, in addition, are comonotonic with X: $\mathcal{F}_{SB}^{\uparrow} = \Big\{Y = I \circ X \in \mathcal{F}_{SB} : I \text{ is nondecreasing}\Big\}$.

Lemma D.1. *If ν is compatible with P, then for each $Y \in \mathcal{F}_{SB}$ there exists a $\widetilde{Y} \in \mathcal{F}_{SB}$ such that:*

1. *$\widetilde{Y}$ is comonotonic with X;*
2. *$\int u\left(W_0 - \Pi - X + \widetilde{Y}\right) dP \geqslant \int u\Big(W_0 - \Pi - X + Y\Big) dP$; and,*
3. *$-\int -\widetilde{Y}\, d\nu \leqslant -\int -Y\, d\nu$.*

Proof. Choose any $Y = I \circ X \in \mathcal{F}_{SB}$, and let $\widetilde{Y}_P$ denote the nondecreasing P-rearrangement of Y with respect to X. Then (i) $\widetilde{Y}_P = \widetilde{I} \circ X$ where $\widetilde{I}$ is nondecreasing, and (ii) $0 \leqslant \widetilde{Y}_P \leqslant X$, by Lemma C.4. Furthermore, since ν is compatible with P, it follows that $-\int -\widetilde{Y}_P d\nu \leqslant -\int -Y d\nu$. But $-\int -Y d\nu \leqslant R$ since $Y \in \mathcal{F}_{SB}$. Hence, $\widetilde{Y}_P \in \mathcal{F}_{SB}^{\uparrow}$. Moreover, since the utility u is concave (Assumption 4.1), the function $\mathcal{U}(x, y) = u(W_0 - \Pi - x + y)$ is supermodular (as in Example C.7 (1)). Then, by Lemma C.8, $\int u\left(W_0 - \Pi - X + \widetilde{Y}\right) dP \geqslant \int u\Big(W_0 - \Pi - X + Y\Big) dP$. □

Now, by Lemma D.1, we can choose a maximizing sequence $\{Y_n\}_n$ in $\mathcal{F}_{SB}^{\uparrow}$ for Problem (4.1). That is,

$$\lim_{n \to +\infty} \int u\left(W_0 - \Pi - X + Y_n\right) dP = N \equiv \sup_{Y \in B^+(\Sigma)} \int u\left(W_0 - \Pi - X + Y\right) dP < +\infty.$$

Since $0 \leqslant Y_n \leqslant X \leqslant M \equiv \|X\|_\infty$, the sequence $\{Y_n\}_n$ is uniformly bounded. Moreover, for each $n \geqslant 1$ we have $Y_n = I_n \circ X$, 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 Helly's First Theorem [11, Lemma 13.15] (a.k.a. Helly's Compactness Theorem), 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 $\{Y_m\}_m$, $Y_m = I_m \circ X$, converges pointwise to Y^*. Thus, the sequence $\{-Y_m\}_m$ is uniformly bounded and converges pointwise to $-Y^*$. By the assumption that ν is continuous, it follows from a Dominated Convergence-type Theorem [43, Theorem 11.9] that

$$\lim_{m \to +\infty} -\int -Y_m d\nu = -\int -Y^* d\nu \leqslant R,$$

and so $Y^* \in \mathcal{F}_{SB}^{\uparrow}$. Now, since the function u is continuous (Assumption 4.1), it is bounded on any closed and bounded subset of $\mathbb{R}$. Therefore, since the range of X is closed and bounded, Lebesgue's Dominated Convergence Theorem [2, Theorem 11.21] implies that

$$\begin{aligned} \int u\left(W_0 - \Pi - X + Y^*\right) dP &= \lim_{m \to +\infty} \int u\left(W_0 - \Pi - X + Y_m\right) dP \\ &= \lim_{n \to +\infty} \int u\left(W_0 - \Pi - X + Y_n\right) dP = N. \end{aligned}$$

Hence Y^* solves Problem (4.1).

Finally, suppose that Y is a solution to Problem (4.1). If Y is comonotonic with X, then the proof of Theorem 4.6 is complete. Suppose now that Y is not comonotonic with X, that is, there exist some $s_1, s_2 \in S$ such that $X(s_1) < X(s_2)$ but $Y(s_1) > Y(s_2)$. Let $\widetilde{Y}$ denote the (P-a.s. unique) nondecreasing rearrangement of Y with respect to X.

Suppose that $P[Y \neq \widetilde{Y}] > 0$. Since Y is optimal for Problem (4.1), it is, in particular feasible. Therefore, $0 \leqslant Y \leqslant X$ and $-\int -Y d\nu \leqslant R$. By Lemma C.4, we have that $0 \leqslant \widetilde{Y} \leqslant X$. Moreover, since ν is compatible with P, we have that and $-\int -\widetilde{Y} d\nu \leqslant -\int -Y d\nu \leqslant R$. Hence, in particular, $\widetilde{Y}$ is feasible for Problem (4.1). Moreover, since the function u is strictly concave (Assumption 4.1), it follows that the function $L(x, y) := u(W_0 - \Pi - x + y)$ is strictly supermodular (see Example C.7). Therefore, Lemma C.8 implies that $\int u\left(W_0 - \Pi - X + \widetilde{Y}\right) dP > \int u\left(W_0 - \Pi - X + Y\right) dP$, contradicting the fact that Y is optimal for Problem (4.1).

Therefore, it must be that $P[Y \neq \widetilde{Y}] = 0$, that is, $Y = \widetilde{Y}$, P-a.s. Hence, Y is comnonotonic with X, except on a set of probability zero under P. This concludes the proof of Theorem 4.6. □

Appendix E. Proof of Proposition 5.1

Let $\mathcal{AC}$ denote the anticore of ν. Since each $Q \in \mathcal{AC}$ is compatible with P, it follows that ν is compatible with P. Hence, by Theorem 4.6, there exists a solution Y^{**} to Problem (5.1). Fix $Q \in \mathcal{AC}$ arbitrarily, and let $Y^*(Q)$ be an optimal solution of Problem (5.2) for this given $Q \in \mathcal{AC}$. The existence of $Y^*(Q)$ follows from the fact that Q is compatible with P, in light of Theorem 4.6. Then, $Y^*(Q)$ satisfies $0 \leqslant Y^*(Q) \leqslant X$ and $\int Y^*(Q)\, dQ \leqslant R$.

Hence,

$$\min_{\mu \in \mathcal{AC}} \int Y^* (Q)\, d\mu \leqslant \int Y^* (Q)\, dQ \leqslant R,$$

which shows that $Y^* (Q)$ is feasible for Problem (5.1). Since Y^{**} solves Problem (5.1), we must have that

$$\int u\left(W_0 - \Pi - X + Y^{**}\right) dP \geqslant \int u\left(W_0 - \Pi - X + Y^* (Q)\right) dP. \tag{E.1}$$

Therefore, to conclude the proof, it suffices to find some $Q^{**} \in \mathcal{AC}$ such that inequality (J.2) holds as an equality. Suppose, by the way of contradiction, that no such Q^{**} exists. Then, for all $Q \in \mathcal{AC}$ it holds that

$$\int u\left(W_0 - \Pi - X + Y^{**}\right) dP > \int u\left(W_0 - \Pi - X + Y^* (Q)\right) dP. \tag{E.2}$$

Since, by definition, $Y^*(Q)$ solves the problem of type (5.2) defined by Q, inequality (E.2) implies that Y^{**} must not be feasible for any problem of the type (5.2). That is, for all $Q \in \mathcal{AC}$,

$$\int Y^{**} dQ > R.$$

However, by the feasibility of Y^{**} for Problem (5.1), we have that for all $Q \in \mathcal{AC}$,

$$\int Y^{**} dQ > R \geqslant \min_{\mu \in \mathcal{AC}} \int Y^{**} d\mu,$$

which, since $Y^{**} \in B(\Sigma)$, contradicts the fact that $\mathcal{AC}$ is weak*-compact and convex. □

Appendix F. Proof of Theorem 5.4

Suppose that assumptions 2.1, 4.1, 4.4, 5.2, and 5.3 all hold. Let $\mathcal{H} := \left\{ Y \in B(\Sigma) \;\middle|\; 0 \leqslant Y \leqslant X \text{ and } -\int -Y \, dT \circ P \leqslant R \right\}$ denote the feasibility set of Problem (5.3). Then it is easy to verify that $\mathcal{H} \neq \varnothing$. For each $Y \in B^+(\Sigma)$, let $F_Y(t) := P\big(\{s \in S : Y(s) \leqslant t\}\big)$ denote the cumulative distribution function (cdf) of Y with respect to the probability measure P, and let $F_Y^{-1}(t)$ be the left-continuous inverse of the cdf F_Y (i.e., the *quantile function* of Y), defined by

$$F_Y^{-1}(t) := \inf \left\{ z \in \mathbb{R}^+ \;\middle|\; F_Y(z) \geqslant t \right\}, \;\; \forall t \in [0,1]. \tag{F.1}$$

Lemma F.1. *We have*

1. $U := F_X(X)$ *is a random variable on the probability space* (S, Σ, P) *with a uniform distribution on* $(0,1)$*; and,*
2. $X = F_X^{-1}(U)$, P*-a.s.*

Moreover, for each $Y \in \mathcal{H}$*, the function* Y^* *defined by* $Y^* := F_Y^{-1}(F_X(X))$ *is such that:*

(1) $Y^* \in \mathcal{H}$;

(2) Y^* *is comonotonic with* X*;*

(3) $\int u\left(W_0 - \Pi - X + Y^*\right)\ dP \geqslant \int u\left(W_0 - \Pi - X + Y\right)\ dP$*; and,*

(4) $-\int -Y^*\ dT \circ P = -\int -Y\ dT \circ P$.

Proof. Since, by assumption, X is a continuous random variable on the probability space (S, Σ, P), the Borel probability measure $\zeta := P \circ X^{-1}$ is nonatomic. For each $Y \in B^+(\Sigma)$, one can then define the P-a.s. unique nondecreasing P-rearrangement $\widetilde{Y}$ of Y with respect to X, as in Appendix C.

Fix some $Y \in B^+(\Sigma)$, and let $Y^* := F_Y^{-1}\left(F_X\left(X\right)\right)$. Then Y can be written as $\phi \circ X$ for some nonnegative Borel-measurable and bounded map ϕ on $X(S)$. Define the mapping $\widetilde{\phi} : [0, M] \to [0, M]$ as in Appendix C (see equation (C.2) on p. 24) to be the nondecreasing ζ-rearrangement of ϕ, that is,

$$\widetilde{\phi}\left(t\right) := \inf \left\{ z \in \mathbb{R}^+ \;\middle|\; \zeta\big(\{x \in [0, M] : \phi\left(x\right) \leqslant z\}\big) \geqslant \zeta\big(\, [0, t]\,\big) \right\}.$$

Then, as in Appendix C, $\widetilde{Y} = \widetilde{\phi} \circ X$. Therefore, for each $s_0 \in S$,

$$\widetilde{Y}\left(s_0\right) = \widetilde{\phi}\left(X\left(s_0\right)\right) = \inf \left\{ z \in \mathbb{R}^+ \;\middle|\; \zeta\big(\{x \in [0, M] : \phi\left(x\right) \leqslant z\}\big) \geqslant \zeta\big(\, [0, X\left(s_0\right)]\,\big) \right\}.$$

However, for each $s_0 \in S$,

$$\zeta\big(\, [0, X\left(s_0\right)]\,\big) = P \circ X^{-1}\big(\, [0, X\left(s_0\right)]\,\big) = F_X\left(X\left(s_0\right)\right) := F_X\left(X\right)\left(s_0\right).$$

Moreover,

$$\begin{aligned}\zeta\big(\{x \in [0, M] : \phi\left(x\right) \leqslant z\}\big) &= P \circ X^{-1}\big(\{x \in [0, M] : \phi\left(x\right) \leqslant z\}\big) \\ &= P\big(\{s \in S : \phi\left(X\left(s\right)\right) \leqslant z\}\big) = F_Y\left(z\right).\end{aligned}$$

Consequently, for each $s_0 \in S$,

$$\widetilde{Y}\left(s_0\right) = \inf \left\{ z \in \mathbb{R}^+ \;\middle|\; F_Y\left(z\right) \geqslant F_X\left(X\right)\left(s_0\right) \right\} = F_Y^{-1}\left(F_X\left(X\left(s_0\right)\right)\right) := F_Y^{-1}\left(F_X\left(X\right)\right)\left(s_0\right).$$

That is,

$$\widetilde{Y} = F_Y^{-1}\left(F_X\left(X\right)\right),$$

where F_Y^{-1} is the left-continuous inverse of F_Y, as defined in equation (F.1). Hence the function Y^* coincides with the function $\widetilde{Y}$, the equimeasurable nondecreasing P-rearrangement of Y with respect to X.

In particular, the P-a.s. unique equimeasurable nondecreasing P-rearrangement of X with respect to itself is given by $F_X^{-1}\left(F_X\left(X\right)\right)$. Since X is a nondecreasing function of X and P-equimeasurable with X, it follows from the P-a.s. uniqueness of the equimeasurable nondecreasing P-rearrangement (see Proposition C.3) that $X = F_X^{-1}\left(F_X\left(X\right)\right)$, P-a.s. (see also [16, Lemma A.21]). Moreover, since $\zeta = P \circ X^{-1}$ is nonatomic, it follows that $U := F_X\left(X\right)$ has a uniform distribution over $(0, 1)$ [16, Lemma A.21], that is, $P\big(\{s \in S : F_X\left(X\right)\left(s\right) \leqslant t\}\big) = t$ for each $t \in (0, 1)$.

Now, let $Y \in \mathcal{H}$ be given and let $\widetilde{Y}$ denote the nondecreasing P-rearrangement of Y with respect to X. Then $\widetilde{Y} = F_Y^{-1}(F_X(X)) = Y^*$, Y^* is comonotonic with X, and $0 \leqslant Y^* \leqslant X$ since $0 \leqslant Y \leqslant X$ (see Lemma C.4). Also, since Y and $\widetilde{Y}$ are P-equimeasurable, we have $\int Y^* \, dT \circ P = \int Y \, dT \circ P$ and $-\int -Y^* \, dT \circ P = -\int -Y \, dT \circ P$. Moreover, since the function u is concave, the function $\mathcal{U}\Big(X, Y\Big) = u(W_0 - \Pi - X + Y)$ is supermodular (Example C.7). Therefore, by Lemma C.8, it follows that $\int u(W_0 - \Pi - X + Y^*) \, dP \geqslant \int u(W_0 - \Pi - X + Y) \, dP$. ☐

Therefore, Lemma F.1 implies that for each $Y \in \mathcal{H}$ the following holds:

(i) $\int u\left(W_0 - \Pi - F_X^{-1}(U) + F_Y^{-1}(U)\right) \, dP \geqslant \int u(W_0 - \Pi - X + Y) \, dP$;
(ii) $\int F_Y^{-1}(U) \, dT \circ P = \int Y \, dT \circ P$; and,
(iii) $-\int -F_Y^{-1}(U) \, dT \circ P = -\int -Y \, dT \circ P$.

Hence, by Lemma F.1, one can look for a solution to Problem (5.3) of the form $F^{-1}(U)$, where F is the cdf of a function $Z \in B^+(\Sigma)$ such that $0 \leqslant Z \leqslant X$ and $-\int -Z \, dT \circ P \leqslant R$.

Now, for any given $Y \in B^+(\Sigma)$, if $Y^* = F_Y^{-1}(F_X(X)) = F_Y^{-1}(U)$, then one can write[10]

$$\begin{aligned}\int Y^* \, dT \circ P &= \int F_Y^{-1}(U) \, dT \circ P \\ &= \int_0^1 T'(1-t) F_Y^{-1}(t) \, dt = \int T'(1-U) F_Y^{-1}(U) \, dP,\end{aligned}$$

and

$$\begin{aligned}-\int -Y^* \, dT \circ P = -\int -F_Y^{-1}(U) \, dT \circ P &= -\int F_{-Y}^{-1}(1-U) \, dT \circ P \\ &= -\int_0^1 T'(1-t) F_{-Y}^{-1}(t) \, dt = -\int_0^1 -T'(1-t) F_Y^{-1}(1-t) \, dt \\ &= \int T'(1-U) F_Y^{-1}(1-U) \, dP = \int T'(U) F_Y^{-1}(U) \, dP,\end{aligned}$$

and

$$\begin{aligned}\int u(W_0 - \Pi - X + Y^*) \, dP &= \int u\left(W_0 - \Pi - F_X^{-1}(U) + F_Y^{-1}(U)\right) \, dP \\ &= \int_0^1 u\left(W_0 - \Pi - F_X^{-1}(t) + F_Y^{-1}(t)\right) \, dt.\end{aligned}$$

Definition F.2. Let $\mathcal{Q}$ denote the collection of all quantile functions. That is,

$$\mathcal{Q} := \Big\{ f : (0,1) \to \mathbb{R} \;\Big|\; f \text{ is nondecreasing and left-continuous} \Big\}. \tag{F.2}$$

[10]This is a standard exercise. See, for instance, [28, p. 418] or [22, p. 213].

Let $\mathcal{Q}^*$ denote the collection of all quantile functions $f \in \mathcal{Q}$ of the form F^{-1}, where F is the cdf of some function $Y \in B^+(\Sigma)$ such that $0 \leqslant Y \leqslant X$. That is,

$$\mathcal{Q}^* = \left\{ f \in \mathcal{Q} \;\middle|\; 0 \leqslant f(z) \leqslant F_X^{-1}(z), \text{ for each } 0 < z < 1 \right\}. \tag{F.3}$$

Now, let $U = F_X(X)$ and consider the following problem.

$$\sup_{f \in \mathcal{Q}^*} \left\{ \int u\left(W_0 - \Pi - F_X^{-1}(U) + f(U)\right) \, dP \;\middle|\; \int T'(U) f(U) \, dP \leqslant R \right\}. \tag{F.4}$$

Proposition F.3 (**Quantile Characterization**)**.** *If f^* is optimal for Problem (F.4), then the function $f^*(U)$ is optimal for Problem (5.3).*

Proof. By Lemma F.1, $X = F_X^{-1}(U)$, P-a.s. Therefore, for any quantile function $f \in \mathcal{Q}$,

$$\int u\left(W_0 - \Pi + f(U) - F_X^{-1}(U)\right) \, dP = \int u\left(W_0 - \Pi + f(U) - X\right) \, dP.$$

Let $f^* \in \mathcal{Q}^*$ is optimal for Problem (F.4), and let $Z^* = f^*(U)$. Then, by Lemma F.1, Z^* is feasible for Problem (5.3). To show optimality, let Z be any feasible solution for Problem (5.3) and let F be the cdf of Z. Then, by Lemma F.1, the function $\widetilde{Z} := F^{-1}(U)$ is feasible for Problem (5.3), comonotonic with X and satisfies:

- $\int u\left(W_0 - \Pi - X + \widetilde{Z}\right) \, dP \geqslant \int u\left(W_0 - \Pi - X + Z\right) \, dP$; and,
- $-\int -\widetilde{Z} \, dT \circ P = -\int -Z \, dT \circ P \leqslant R$.

Moreover, $\widetilde{Z}$ has also F as a cdf. To show optimality of $Z^* = f^*(U)$ for Problem (5.3) it remains to show that

$$\int u\left(W_0 - \Pi - X + Z^*\right) \, dP \geqslant \int u\left(W_0 - \Pi - X + \widetilde{Z}\right) \, dP.$$

Now, let $f := F^{-1}$, so that $\widetilde{Z} = f(U)$. Since $\widetilde{Z}$ is feasible for Problem (5.3), it follows that

$$R \geqslant \int T'(U) f(U) \, dP.$$

Hence, f is feasible for Problem (F.4). Since f^* is optimal for Problem (F.4), it follows that

$$\int u\left(W_0 - \Pi - X + Z^*\right) \, dP \geqslant \int u\left(W_0 - \Pi - X + \widetilde{Z}\right) \, dP.$$

Therefore, $Z^* = f^*(U)$ is optimal for Problem (5.3). ☐

Note that Proposition F.3 holds for any distortion function T, and the concavity of T has not been used yet. Proposition F.3 allows us to focus on solving Problem (F.4), which we now address.

Lemma F.4. *If $f^* \in \mathcal{Q}^*$ satisfies the following:*

(1) $\int_0^1 T'(t)\, f^*(t)\ dt = R$,
(2) *There exists $\lambda \geqslant 0$ such that for all $t \in (0,1)$,*

$$f^*(t) = \underset{0 \leqslant y \leqslant F_X^{-1}(t)}{\arg\max} \left[u\Big(W_0 - \Pi + y - F_X^{-1}(t) \Big) - \lambda T'(t)\, y \right],$$

then f^* solves Problem (F.4), and therefore, the function $f^*(U)$ is optimal for Problem (5.3), where $U = F_X(X)$.

Proof. Suppose that $f^* \in \mathcal{Q}^*$ satisfies conditions (1) and (2) above. Then, in particular, f^* is feasible for Problem (F.4). To show optimality of f^* for Problem (F.4), let f by any other feasible solution for Problem (F.4). Then $\int_0^1 T'(t)\, f(t)\ dt \leqslant R$ and, for all $t \in (0,1)$,

$$u\Big(W_0 - \Pi + f^*(t) - F_X^{-1}(t)\Big) - \lambda T'(t)\, f^*(t) \geqslant u\Big(W_0 - \Pi + f(t) - F_X^{-1}(t)\Big) - \lambda T'(t)\, f(t),$$

that is,

$$\left[u\Big(W_0 - \Pi + f^*(t) - F_X^{-1}(t)\Big) - u\Big(W_0 - \Pi + f(t) - F_X^{-1}(t)\Big)\right] \geqslant \lambda T'(t) \left[f^*(t) - f(t)\right].$$

Integrating yields

$$\int_0^1 u\Big(W_0 - \Pi + f^*(t) - F_X^{-1}(t)\Big)\ dt - \int_0^1 u\Big(W_0 - \Pi + f(t) - F_X^{-1}(t)\Big)\ dt$$
$$\geqslant \lambda \left[R - \int_0^1 T'(t)\, f(t)\ dt\right] \geqslant 0,$$

or,

$$\int u\Big(W_0 - \Pi + f^*(U) - F_X^{-1}(U)\Big)\ dP = \int_0^1 u\Big(W_0 - \Pi + f^*(t) - F_X^{-1}(t)\Big)\ dt$$
$$\geqslant \int_0^1 u\Big(W_0 - \Pi + f(t) - F_X^{-1}(t)\Big)\ dt = \int u\Big(W_0 - \Pi + f(U) - F_X^{-1}(U)\Big)\ dP,$$

as required. The rest follows from Proposition F.3. □

Lemma F.4 suggests that in order to find a solution for Problem (F.4), one can start by solving the problem

$$\max_{0 \leqslant f_\lambda(t) \leqslant F_X^{-1}(t)} \left[u\Big(W_0 - \Pi + f_\lambda(t) - F_X^{-1}(t)\Big) - \lambda T'(t)\, f_\lambda(t)\right], \tag{F.5}$$

for a given $\lambda \geqslant 0$ and for a fixed $t \in (0,1)$.

Lemma F.5. *Let $U = F_X(X)$. An optimal solution for Problem (5.3) takes the form:*

$$\mathcal{Y}^* = \max\left[0, \min\left\{F_X^{-1}(U), (u')^{-1}(\lambda^* T'(U)) + F_X^{-1}(U) - W_0 + \Pi\right\}\right], \tag{F.6}$$

where λ^* *is chosen so that* $\int T'(U)\,\mathcal{Y}^*\,dP = R$.

Proof. For a given $\lambda \geqslant 0$ and for a fixed $t \in (0,1)$ consider the problem:

$$\max_{f_\lambda(t)} \left[u\Big(W_0 - \Pi + f_\lambda(t) - F_X^{-1}(t) \Big) - \lambda T'(t)\, f_\lambda(t) \right]. \tag{F.7}$$

By Assumption 4.1, the first-order conditions are sufficient for an optimum for Problem (F.7) and they imply that the function

$$f_\lambda^*(t) := (u')^{-1}(\lambda T'(t)) + F_X^{-1}(t) - W_0 + \Pi$$

solves Problem (F.7). Concavity of u and T imply that the function $f_\lambda^* : (0,1) \to \mathbb{R}$ is nondecreasing, since F_X^{-1} is a nondecreasing function. Assumption 5.2 implies that the function T' is continuous. Assumption 4.1 implies that the function $(u')^{-1}$ is continuous and strictly decreasing (see Remark 4.2). This yields the left-continuity of f_λ^*. Consequently, $f_\lambda^* \in \mathcal{Q}$, the set of all quantile functions.

Now, define the function f_λ^{**} by

$$f_\lambda^{**}(t) := \max \left[0, \min \Big\{ F_X^{-1}(t), f_\lambda^*(t) \Big\} \right]. \tag{F.8}$$

It is then easy to check that $f_\lambda^{**} \in \mathcal{Q}$, since $f_\lambda^* \in \mathcal{Q}$ and since F_X^{-1} is a nondecreasing function. Moreover, $0 \leqslant f_\lambda^{**}(z) \leqslant F_X^{-1}(z)$, for each $z \in (0,1)$. Therefore, $f_\lambda^{**} \in \mathcal{Q}^*$. Finally, it is easily seen that $f_\lambda^{**}(t)$ solves Problem (F.5) for the given λ and t, since the concavity of u yields the concavity of the function $z \mapsto u\Big(W_0 - \Pi + z - F_X^{-1}(t)\Big) - \lambda T'(t)\, z$, for each $t \in (0,1)$. Hence, in view of Lemma F.4, it remains to show that there exists a $\lambda^* \geqslant 0$ such that $\int_0^1 T'(t)\, f_{\lambda^*}^{**}(t)\ dt = R$.

Let ψ be the function of the parameter $\lambda \geqslant 0$ defined by

$$\psi(\lambda) := \int_0^1 T'(t) \max \left[0, \min \Big\{ F_X^{-1}(t), (u')^{-1}(\lambda T'(t)) + F_X^{-1}(t) - W_0 + \Pi \Big\} \right] dt.$$

It then suffices to show that there exists a $\lambda^* \geqslant 0$ such that $\psi(\lambda^*) = R$. Since X is bounded and since F_X^{-1} is nondecreasing, it follows that for each $t \in [0,1]$,

$$\min \Big\{ F_X^{-1}(t), (u')^{-1}(\lambda T'(t)) + F_X^{-1}(t) - W_0 + \Pi \Big\} \leqslant F_X^{-1}(t) \leqslant F_X^{-1}(1) \leqslant \|X\|_{sup} < +\infty.$$

Moreover, since T is concave and increasing, T' is nonincreasing and nonnegative, and so for each $t \in [0,1]$, $0 \leqslant T'(t) \leqslant T'(0)$. But Assumption 5.2 implies that the function T' is continuous, and hence bounded on every closed and bounded subset of $\mathbb{R}$. Therefore, $T'(0) < +\infty$. Hence, for each $t \in [0,1]$,

$$\min \Big\{ F_X^{-1}(t), (u')^{-1}(\lambda T'(t)) + F_X^{-1}(t) - W_0 + \Pi \Big\}\, T'(t) \leqslant F_X^{-1}(1)\, T'(0) < +\infty.$$

Hence, Lebesgue's Dominated Convergence Theorem [2, Theorem 11.21] implies that ψ is a continuous function of λ. Moreover, ψ is a nonincreasing function of λ (by concavity of u and by monotonicity of the Lebesgue integral).

Now, Assumption 4.1 implies that

$$\psi(0) = \int_0^1 T'(t)\, F_X^{-1}(t)\ dt = -\int -X\ dT \circ P,$$

and that

$$\begin{aligned}\lim_{\lambda\to+\infty} \psi(\lambda) &= \int_0^1 T'(t) \max\left[0, \min\left\{F_X^{-1}(t), F_X^{-1}(t) - W_0 + \Pi\right\}\right] dt \\ &= \int_0^1 T'(t) \max\left[0, F_X^{-1}(t) - W_0 + \Pi\right] dt.\end{aligned}$$

Furthermore, by Assumption 5.3, $F_X(W_0 - \Pi) = 1$. Therefore, for all $t \in (0,1)$, $F_X^{-1}(t) \leqslant W_0 - \Pi$, and so

$$\lim_{\lambda\to+\infty} \psi(\lambda) = 0.$$

Consequently (recall Remark 4.3 and Assumption 4.4),

$$0 = \lim_{\lambda\to+\infty} \psi(\lambda) \leqslant R < -\int -X\ dT \circ P = \int_0^1 T'(t)\, F_X^{-1}(t)\ dt = \psi(0).$$

Hence, by the Intermediate Value Theorem [38, Theorem 4.23], there exists a $\lambda^* \geqslant 0$ such that $\psi(\lambda^*) = R$. This concludes the proof of Lemma F.5. □

Now, since X is a continuous random variable for P, it follows from Lemma F.1 that $X = F_X^{-1}(U)$, P-a.s., where $U = F_X(X)$. Therefore,

$$\begin{aligned}&\max\left[0, \min\left\{F_X^{-1}(U), (u')^{-1}(\lambda^* T'(U)) + F_X^{-1}(U) - W_0 + \Pi\right\}\right] \\ &\quad = \max\left[0, \min\left\{X, (u')^{-1}(\lambda^* T'(U)) + X - W_0 + \Pi\right\}\right],\ P\text{-a.s.}\end{aligned}$$

Therefore, letting $\mathcal{Y}_2^* = \max\left[0, \min\left\{X, (u')^{-1}(\lambda^* T'(U)) + X - W_0 + \Pi\right\}\right]$, it follows that

$$\int T'(U)\, \mathcal{Y}_2^*\ dP = \int T'(U)\, \mathcal{Y}^*\ dP = R,$$

and that

$$\int u(W_0 - \Pi - X + \mathcal{Y}_2^*)\ dP = \int u(W_0 - \Pi - X + \mathcal{Y}^*)\ dP.$$

In other words, $\mathcal{Y}_2^*$ is also optimal for Problem (5.3). Define Y^* by

$$Y^* = \min\left[X, \max\Big(0, X - d(T)\Big)\right],$$

where

$$d(T) = W_0 - \Pi - (u')^{-1}\Big(\lambda^* T'(U)\Big).$$

It can then easily be verified that $Y^* = \mathcal{Y}_2^*$. □

Appendix G. Proof of Proposition 5.5

(1) The proof is similar to that of that last part of Theorem 4.6.
(2) Since $d(T) = W_0 - \Pi - (u')^{-1}\Big(\lambda^* T'(F_X(X))\Big)$, the result follows immediately from the fact that T' is nonincreasing (T is concave) and $(u')^{-1}$ is decreasing (u is concave and increasing).
(3) The P-a.s. uniqueness property results from the property of the equimeasurable nondecreasing rearrangement (Appendix C). □

Appendix H. Proof of Proposition 5.6

First, recall that $d(T) = W_0 - \Pi - (u')^{-1}(\lambda^* T'(F_X(X)))$, where:

- $\lambda^* \geqslant 0$;
- T' is nonincreasing (T is concave); and,
- $(u')^{-1}$ is decreasing (u is concave and increasing).

Lemma H.1. *The function* $\Phi : x \mapsto W_0 - \Pi - (u')^{-1}(\lambda^* T'(F_X(x)))$ *is nonincreasing.*

Proof. This is a direct consequence of the fact that T' is nonincreasing and $(u')^{-1}$ is decreasing. □

Now, let $s_1, s_2 \in S$ be such that

$$0 < d(T)(s_1) < X(s_1) \quad \text{and} \quad X(s_2) \leqslant d(T)(s_2), \tag{H.1}$$

To show that $X(s_2) < X(s_1)$, suppose, by way of contradiction, that $X(s_2) \geqslant X(s_1)$. Then, Lemma H.1 implies that

$$d(T)(s_2) = \Phi(X(s_2)) \leqslant \Phi(X(s_1)) = d(T)(s_1).$$

Therefore, since $X(s_2) \geqslant X(s_1)$ by assumption, we have (using eq. (H.1)),

$$d(T)(s_2) \geqslant X(s_2) \geqslant X(s_1) > d(T)(s_1).$$

Consequently,

$$d(T)(s_2) > d(T)(s_1),$$

hence contradicting the fact that $d(T)(s_2) \leqslant d(T)(s_1)$. Therefore, $X(s_2) < X(s_1)$. □

Appendix I. Proof of Corollary 5.7

By a proof identical to that of Theorem 5.4, an optimal indemnity schedule is given by

$$Y^* = \min\left[X, \max\left(0, X - d\right)\right],$$

where $d = W_0 - \Pi - (u')^{-1}(\lambda^*)$, and λ^* is such that $\int Y^* dP = R = \frac{\Pi}{1+\rho}$. It only remains to show that, in this case, $d > 0$. Suppose, by way of contradiction, that $d \leqslant 0$. Then, $\max\left(0, X-d\right) = X-d \geqslant X$, and so $Y^* = \min\left[X, \max\left(0, X-d\right)\right] = X$. Therefore, $R = \int Y^* dP = \int X dP$, contradicting the fact that $R < -\int -X \, dP = \int X dP$ (by Assumption 4.4, and since T is the identity function in this case). $\square$

Appendix J. Proof of Proposition 6.1

For each $Y \in B(\Sigma)$, let

$$\widetilde{u}(Y) = \int u\left(W_0 - \Pi - X + Y\right) dP.$$

Let $\mathcal{C}$ denote the core of ν. Since each $Q \in \mathcal{C}$ is compatible with P, it follows that ν is compatible with P. Hence, by Theorem 4.6, there exists a solution Y^{**} to Problem (6.1). Let

$$\overline{V} = \widetilde{u}\left(Y^{**}\right)$$

denote the value of Problem 6.1.

For a given $Q \in \mathcal{C}$, let $\mathcal{M}(Q)$ denote Problem (6.2) for this given $Q \in \mathcal{C}$. Fix $Q \in \mathcal{C}$ arbitrarily, and let $Y^*(Q)$ be an optimal solution of Problem $\mathcal{M}(Q)$. The existence of $Y^*(Q)$ follows from the fact that Q is compatible with P, in light of Theorem 4.6. For any $Q \in \mathcal{C}$, let

$$V(Q) = \widetilde{u}\left(Y(Q)\right)$$

denote the value of Problem $\mathcal{M}(Q)$.

Since Y^{**} solves Problem (6.1), it is, in particular, feasible for Problem $\mathcal{M}(Q)$, for any $Q \in \mathcal{C}$. Therefore,

$$\overline{V} \leqslant V(Q), \quad \forall Q \in \mathcal{C}.$$

Hence, in order to conclude the proof of Proposition 6.1, it suffices to show that there exists some $Q^* \in \mathcal{C}$ such that

$$\overline{V} \geqslant V\left(Q^*\right).$$

To do this, we will show that the function

$$\begin{aligned} V : \mathcal{C} &\longrightarrow \mathbb{R} \\ Q &\longmapsto V(Q) = \widetilde{u}\left(Y(Q)\right) \end{aligned}$$

attains its minimum on $\mathcal{C}$, and that $\overline{V} \geqslant \min_{Q \in \mathcal{C}} V(Q)$.

The fact that V attains its minimum on $\mathcal{C}$ is a consequence of the Maximum Theorem [2, Theorem 17.31] applied to the correspondence $\Gamma : \mathcal{AC} \rightrightarrows B(\Sigma)$ defined by

$$\Gamma : Q \rightrightarrows \left\{ Y \in B(\Sigma) \;\middle|\; 0 \leqslant Y \leqslant X \text{ and } \int Y dQ \leqslant R \right\}.$$

Indeed, since $\mathcal{C}$ is weak*-compact [33, Proposition 4.2], it suffices to show that V is weak*-continuous. When $\mathcal{C}$ and $B(\Sigma)$ are endowed with their weak* topologies, it is easy to see that the function $Y \longmapsto \widetilde{u}(Y)$ is weak*-continuous (since P is a charge), and that Γ is continuous and compact-valued. The weak*-continuity of V then results from the Maximum Theorem [2, Theorem 17.31]. Hence, $\arg\min V \neq \varnothing$.

Lemma J.1. *If $Y(Q_x)$ solves problem $\mathcal{M}(Q_y)$, then $Y(Q_x) = Y(Q_y)$, P-a.s.*

Proof. By assumption, both $Y(Q_x)$ and $Y(Q_y)$ are feasible for $\mathcal{M}(Q_y)$. Therefore, any strict convex combination of $Y(Q_x)$ and $Y(Q_y)$ is also feasible for $\mathcal{M}(Q_y)$. By strict concavity of u and since $\widetilde{u}(Y(Q_y)) = \widetilde{u}(Y(Q_x))$, we have

$$\int \Big[u(W_0 - \Pi - X + \alpha Y(Q_x) + (1-\alpha) Y(Q_y)) - u(W_0 - \Pi - X + Y(Q_y)) \Big] dP > 0$$

when $P(|Y(Q_x) - Y(Q_y)| \neq 0) > 0$. □

Now, since $\arg\min V \neq \varnothing$, choose some $Q_0 \in \arg\min V$. Then $V(Q_0) = \min_{Q \in \mathcal{C}} V(Q)$. If $\int Y(Q_0) dQ \leqslant R$ for every $Q \in \mathcal{C}$, then $Y(Q_0)$ is feasible for Problem 6.1, and the proof of Proposition 6.1 is done. So suppose that there exists $Q_1 \in \mathcal{C}$ such that

$$\int Y(Q_0) dQ_1 > R. \tag{J.1}$$

Then, in particular, $Q_1 \neq Q_0$. Let $Y(Q_1)$ be a solution to problem $\mathcal{M}(Q_1)$. Then there are two cases to consider:

(a) $\int Y(Q_1) dQ_0 > R$; and,

(b) $\int Y(Q_1) dQ_0 \leqslant R$.

We are going to show that (a) does not obtain. Suppose, by way of contradiction, that we are in case (a). Then since $\int Y(Q_0) dQ_0 \leqslant R$, it must be that $Y(Q_1) \neq Y(Q_0)$, P-a.s. Moreover, by (a) and eq. (J.1) we have

$$\int Y(Q_0) dQ_0 \leqslant R, \quad \int Y(Q_0) dQ_1 > R, \quad \int Y(Q_1) dQ_1 \leqslant R, \quad \text{and} \quad \int Y(Q_1) dQ_0 > R,$$

which implies that

$$\int [Y(Q_0) - Y(Q_1)] dQ_1 > 0 > \int [Y(Q_0) - Y(Q_1)] dQ_0.$$

Hence, there exists $\overline{\alpha} \in (0,1)$ such that

$$\overline{\alpha} \int \left[Y\left(Q_0\right) - Y\left(Q_1\right)\right] dQ_1 + \left(1 - \overline{\alpha}\right) \int \left[Y\left(Q_0\right) - Y\left(Q_1\right)\right] dQ_0 = 0. \tag{J.2}$$

Notice that eq. (J.2) implies that

$$R \geqslant \int Y(Q_0)dQ_0 = \int Y(Q_1)dQ_0 + \bar{\alpha}\left[\int \left[Y(Q_0) - Y(Q_1)\right] dQ_0 - \int \left[Y(Q_0) - Y(Q_1)\right] dQ_1\right],$$

and hence

$$\int \left[Y(Q_0) - Y(Q_1)\right] dQ_0 = \overline{\alpha}\left[\int \left[Y(Q_0) - Y(Q_1)\right] dQ_0 - \int \left[Y(Q_0) - Y(Q_1)\right] dQ_1\right]. \tag{J.3}$$

Now, for each $\varepsilon > 0$, let $\widetilde{Y}\left(\varepsilon\right) = \varepsilon m \mathbb{1}_{B_M}$, where $m \leqslant \left\|Y\left(Q_0\right)\right\|_{sup}$ and $\mathbb{1}_{B_M}$ is the indicator function of the set $B_M = \{s \in S : Y\left(Q_0\right)(s) \geqslant m\}$. Then, by using eq. (J.3), we can write for each $\beta \in (0,1)$, for each m, and for each $\varepsilon > 0$

$$\begin{aligned}
\int &(\beta Y(Q_0) + (1-\beta)Y(Q_1) - \widetilde{Y}\left(\varepsilon\right))dQ_0 \\
&= \int Y(Q_1)dQ_0 \\
&\quad + \beta\overline{\alpha}\left[\int \left[Y(Q_0) - Y(Q_1)\right] dQ_0 - \int \left[Y(Q_0) - Y(Q_1)\right] dQ_1\right] \\
&\quad - \int \widetilde{Y}\left(\varepsilon\right) dQ_0.
\end{aligned}$$

Then, for a given $\beta \in (0,1)$ and m, ε can be chosen so that $\beta Y(Q_0) + (1-\beta)Y(Q_1) - \widetilde{Y}\left(\varepsilon\right)$ is feasible for Problem $\mathcal{M}\left(Q_0\right)$.

Next observe that by strict concavity of u and by $Y(Q_1) \neq Y(Q_0)$ P-a.s., as observed above, we have

$$\begin{aligned}
\lim_{\varepsilon \to 0} \widetilde{u}\left(\beta Y(Q_0) + (1-\beta)Y(Q_1) - \widetilde{Y}(\varepsilon)\right) &= \widetilde{u}\Big(\beta Y(Q_0) + (1-\beta)Y(Q_1)\Big) \\
&> \beta\widetilde{u}\left(Y(Q_0)\right) + (1-\beta)\widetilde{u}\left(Y(Q_1)\right) \\
&\geqslant \widetilde{u}\left(Y(Q_0)\right) = V\left(Q_0\right) = \min_{\mu \in \mathcal{C}} V\left(\mu\right).
\end{aligned}$$

Hence, there exists $\varepsilon > 0$ such that

$$\widetilde{u}\left(\beta Y(Q_0) + (1-\beta)Y(Q_1) - \widetilde{Y}\left(\varepsilon\right)\right) > \widetilde{u}\left(Y(Q_0)\right),$$

which contradicts the fact that $Y\left(Q_0\right)$ is a solution for Problem $\mathcal{M}(Q_0)$.

In sum, for any $Q_0 \in \arg\min V$, if there exists $Q_1 \in \mathcal{C}$ such that $\int Y(Q_0)dQ_1 > R$, then $\int Y(Q_1)dQ_0 \leqslant R$. We can re-formulate this as the following result.

Lemma J.2. *Let $Q_0 \in \arg\min V$, and let*

$$\mathcal{I}(Q_0) = \left\{ Q \in \mathcal{C} \;\middle|\; \int Y(Q_0) dQ > R \right\}.$$

Then:

(a) *If $Q \in \mathcal{I}(Q_0)$, then $Y(Q)$ is feasible for both Problem $\mathcal{M}(Q_0)$ and Problem $\mathcal{M}(Q)$;*

(b) $\mathcal{I}(Q_0) \subset \arg\min V$.

Proof. Part (a) of the Lemma is the argument preceding it plus the obvious observation that, by definition, $Y(Q)$ is feasible for Problem $\mathcal{M}(Q)$. To show part (b), note that since $Y(Q)$ is feasible for Problem $\mathcal{M}(Q_0)$ (by part (a)), we have $V(Q_0) = \widetilde{u}\,(Y(Q_0)) \geqslant \widetilde{u}\,(Y(Q)) = V(Q) \geqslant \min_{\mu \in \mathcal{C}} V\,(\mu) = V\,(Q_0)$. Hence, $Q \in \mathcal{I}\,(Q_0) \Longrightarrow Q \in \arg\min V$. □

Let $Q_1 \in \mathcal{I}\,(Q_0) \subset \arg\min V$, and let

$$\mathcal{I}_1\,(Q_1) = \left\{ Q \in \mathcal{I}\,(Q_0) \;\middle|\; \int Y\,(Q_1)\,dQ > R \right\}.$$

Lemma J.3. *If $Q \in \mathcal{I}_1(Q_1)$, then $Y(Q)$ is feasible for problems $\mathcal{M}(Q)$, $\mathcal{M}(Q_0)$, and $\mathcal{M}(Q_1)$.*

Proof. The feasibility of $Y(Q)$ for problems $\mathcal{M}(Q)$ and $\mathcal{M}(Q_0)$ follows from Lemma J.2. An argument similar to the one used in proving part (a) of Lemma J.2 also shows feasibility of $Y(Q)$ for Problem $\mathcal{M}(Q_1)$. □

Lemma J.4. *There exists $Q^* \in \mathcal{C}$ such that $\int Y(Q_0)dQ^* > R$ and $Y\,(Q^*)$ is feasible for (and hence solves) Problem $\mathcal{M}\,(Q)$ for all $Q \in \mathcal{I}\,(Q_0) \cup \{Q_0\}$.*

Proof. By transfinite induction (Zorn's Lemma), using Lemma J.3, and using a construction similar to that of $\mathcal{I}_1\,(Q_1)$ and $\mathcal{I}\,(Q_0)$. □

Lemma J.5. *There exists $Q^* \in \arg\min V$ such that $Y\,(Q^*)$ is feasible for (and hence solves) Problem $\mathcal{M}\,(Q)$ for all $Q \in \arg\min V$.*

Proof. Let Q^* be the probability measure given in Lemma J.4, and let $Q \in \arg\min V$ be chosen arbitrarily. Then, either

(i) $Q \in \mathcal{I}(Q_0) \cup \{Q_0\}$; or

(ii) $\int Y(Q_0)dQ \leqslant R$ and $Q \neq Q_0$.

In case (i), Lemma J.4 implies that $Y(Q^*)$ solves Problem $\mathcal{M}\,(Q)$.

In case (ii), since $Q \in \arg\min V$, it follows that $Y(Q_0)$ is also a solution for Problem $\mathcal{M}(Q)$. Note also that since $Q \neq Q_0$, Lemma J.1 implies that $Y(Q_0) = Y(Q)$ P-a.s.

Moreover, by an argument similar to that used in the proof of part (a) of Lemma J.2, $Y(Q^*)$ solves Problem $\mathcal{M}(Q_0)$, and $Q^* \neq Q_0$ (because $\int Y(Q_0)dQ_0 \leqslant R < \int Y(Q_0)dQ^*$). Therefore, by Lemma J.1 we have $Y(Q_0) = Y(Q^*)$ P-a.s. Consequently, letting

$$A = \left\{ s \in S : Y\left(Q_0\right)(s) = Y\left(Q^*\right)(s) \right\} \quad \text{and} \quad B = \left\{ s \in S : Y\left(Q_0\right)(s) = Y\left(Q\right)(s) \right\},$$

we have $P(A) = P(B) = 1$, and

$$\int Y(Q^*)dQ = \int_A Y(Q^*)dQ = \int_A Y\left(Q_0\right)dQ = \int_{A\cap B} Y\left(Q_0\right)dQ = \int_{A\cap B} Y\left(Q\right)dQ$$
$$= \int Y(Q)dQ \leqslant R.$$

That is, $Y(Q^*)$ is feasible for (and hence solves) Problem $\mathcal{M}(Q)$.

Since Q was chosen arbitrarily, it follows that $Y(Q^*)$ is feasible for (and hence solves) Problem $\mathcal{M}(Q)$ for all $Q \in \arg\min V$. □

Proposition There exists $Q^* \in \mathcal{C}$ such that $Y\left(Q^*\right)$ solves Problem 6.1.

Proof of the Proposition: Let Q^* be given from Lemma J.5. We are going to show that

$$\int Y(Q^*)dQ \leqslant R, \quad \forall Q \in \mathcal{C}.$$

Suppose, by way of contradiction, that there exists $\widetilde{Q} \in \mathcal{C}$ such that $\int Y(Q^*)d\widetilde{Q} > R$. By Lemma J.2, $\widetilde{Q} \in \arg\min V$. By Lemma J.5, $Y(Q^*)$ solves Problem $\mathcal{M}(\widetilde{Q})$. Hence, in particular $\int Y(Q^*)d\widetilde{Q} \leqslant R$, a contradiction. It follows that $Y(Q^*)$ is feasible in all problems $\mathcal{M}(Q)$, and hence $Y(Q^*)$ is feasible for Problem 6.1. This implies that $\overline{V} \geqslant \widetilde{u}\left(Y(Q^*)\right) = \min_{\mu\in\mathcal{C}} V(\mu)$. □

Appendix K. Proof of Corollary 6.3

By a result of Ghossoub [17], the MLR property implies compatibility with P: since every element of the core of ν has a nonincreasing Radon-Nikodým derivative with respect to P, every element of the core of ν is compatible with P [17]. Therefore, the supermodular capacity ν is itself compatible with P. Hence, by Proposition 6.1, there exists a $Q^* \in \mathcal{C}$ such that a solution to Problem (6.1) is given by a solution to the following problem:

$$\text{(K.1)} \qquad \sup_{Y\in B(\Sigma)} \left\{ \int u\left(W_0 - \Pi - X + Y\right)\, dP \;\middle|\; 0 \leqslant Y \leqslant X,\ \int Y dQ^* \leqslant R \right\}.$$

If $R \geqslant \int X dQ^*$, then it is easy to verify that X is optimal for Problem (K.1).

Now, suppose that $R < \int X dQ^*$. By the Radon-Nikodým Theorem [13, Th. 4.2.2], for each $Q \in \mathcal{C}$ there exists a P-a.s. unique Σ-measurable and P-integrable function $h_Q : S \to$

$[0, +\infty)$ such that $Q(B) = \int_B h_Q dP$, for all $B \in \Sigma$. The function h_Q is called the Radon-Nikodým derivative of Q with respect to P and it is denoted by dQ/dP. Moreover, since $\Sigma = \sigma\{X\}$ and since $h_Q : S \to [0, +\infty)$ is Σ-measurable and P-integrable, there exists a Borel-measurable and $P \circ X^{-1}$-integrable map $\phi_Q : X(S) \to [0, +\infty)$ such that $h_Q = dQ/dP = \phi_Q \circ X$. Denoting dQ^*/dP by $\phi_{Q^*} \circ X$, Problem (K.1) is equivalent to the following problem:

$$\text{(K.2)} \qquad \sup_{Y \in B(\Sigma)} \left\{ \int u\left(W_0 - \Pi - X + Y\right) \, dP \;\middle|\; 0 \leqslant Y \leqslant X, \; \int Y \phi_{Q^*} \circ X dP \leqslant R \right\}.$$

Let $\mathcal{H} := \left\{ Y \in B(\Sigma) \;\middle|\; 0 \leqslant Y \leqslant X \text{ and } \int Y \phi_{Q^*} \circ X dP \leqslant R \right\}$ denote the feasibility set of Problem (K.2).

Lemma K.1. *For each $Y \in \mathcal{H}$, the function Y^* defined by $Y^* := F_Y^{-1}(F_X(X))$ is such that:*

(1) *$Y^* \in \mathcal{H}$;*

(2) *Y^* is comonotonic with X;*

(3) *$\int u(W_0 - \Pi - X + Y^*) \, dP \geqslant \int u(W_0 - \Pi - X + Y) \, dP$; and,*

(4) *$\int Y^* \phi_{Q^*} \circ X dP \leqslant \int Y \phi_{Q^*} \circ X dP$.*

Proof. The proofs of (1), (2), and (3) are similar to what was done in the proof of Lemma F.1. The proof of (4) is an immediate consequence of the fact that Q^* is compatible with P. □

Hence, by Lemma K.1, one can look for a solution to Problem (K.2) of the form $F^{-1}(U)$, where F is the cdf of a function $Z \in B^+(\Sigma)$ such that $0 \leqslant Z \leqslant X$ and $\int Z \phi_{Q^*} \circ X dP$, $U = F_X(X)$ is a a random variable on the probability space (S, Σ, P) with a uniform distribution on $(0, 1)$ (Lemma F.1).

Recall the definition of the sets of quantiles $\mathcal{Q}$ and $\mathcal{Q}^*$ given in equations (F.2)-(F.3), and recall from Lemma F.1 that $X = F_X^{-1}(U)$, P-a.s. Consider the following problem.

$$\text{(K.3)} \qquad \sup_{f \in \mathcal{Q}^*} \left\{ \int u\left(W_0 - \Pi - F_X^{-1}(U) + f(U)\right) \, dP \;\middle|\; \int f(U) \, \phi_{Q^*} \circ F_X^{-1}(U) \, dP \leqslant R \right\}.$$

Proposition K.2 (Quantile Characterization). *If f^* is optimal for Problem (K.3), then the function $f^*(U)$ is optimal for Problem (K.2).*

The proof of Proposition K.2 is similar to that of Proposition F.3. Proposition K.2 allows us to focus on solving Problem (K.3), which we now address.

Lemma K.3. *If $f^* \in \mathcal{Q}^*$ satisfies the following:*

(1) $\int_0^1 f(t)\,\phi_{Q^*} \circ F_X^{-1}(t)\ dt = R$,

(2) *There exists* $\lambda \geqslant 0$ *such that for all* $t \in (0,1)$,

$$f^*(t) = \underset{0 \leqslant y \leqslant F_X^{-1}(t)}{\arg\max} \left[u\Big(W_0 - \Pi + y - F_X^{-1}(t)\Big) - \lambda \phi_{Q^*} \circ F_X^{-1}(t)\, y \right],$$

then f^* *solves Problem* (K.3), *and therefore, the function* $f^*(U)$ *is optimal for Problem* (K.2), *where* $U = F_X(X)$.

The proof of Proposition K.3 is similar to that of Proposition F.4. Lemma K.3 suggests that in order to find a solution for Problem (K.3), one can start by solving the problem

$$\max_{0 \leqslant f_\lambda(t) \leqslant F_X^{-1}(t)} \left[u\Big(W_0 - \Pi + f_\lambda(t) - F_X^{-1}(t)\Big) - \lambda \phi_{Q^*} \circ F_X^{-1}(t)\, f_\lambda(t) \right], \tag{K.4}$$

for a given $\lambda \geqslant 0$ and for a fixed $t \in (0,1)$.

Lemma K.4. *Let* $U = F_X(X)$. *An optimal solution for Problem* (K.2) *takes the form:*

$$\mathcal{Y}^* = \max\left[0, \min\Big\{F_X^{-1}(U), (u')^{-1}\left(\lambda^* \phi_{Q^*} \circ F_X^{-1}(U)\right) + F_X^{-1}(U) - W_0 + \Pi\Big\}\right], \tag{K.5}$$

where λ^* *is chosen so that* $\int \phi_{Q^*} \circ F_X^{-1}(U)\, \mathcal{Y}^*\ dP = R$.

Proof. For a given $\lambda \geqslant 0$ and for a fixed $t \in (0,1)$ consider the problem:

$$\max_{f_\lambda(t)} \left[u\Big(W_0 - \Pi + f_\lambda(t) - F_X^{-1}(t)\Big) - \lambda \phi_{Q^*} \circ F_X^{-1}(t)\, f_\lambda(t) \right]. \tag{K.6}$$

By Assumption 4.1, the first-order conditions are sufficient for an optimum for Problem (K.6) and they imply that the function

$$f_\lambda^*(t) := (u')^{-1}\left(\lambda \phi_{Q^*} \circ F_X^{-1}(t)\right) + F_X^{-1}(t) - W_0 + \Pi$$

solves Problem (K.6). The fact that u is concave and ϕ_{Q^*} is nonincreasing imply that the function $f_\lambda^* : (0,1) \to \mathbb{R}$ is nondecreasing, since F_X^{-1} is a nondecreasing function. Since ϕ_{Q^*} and $(u')^{-1}$ are continuous, f_λ^* is left-continuous. Consequently, $f_\lambda^* \in \mathcal{Q}$, the set of all quantile functions.

Now, define the function f_λ^{**} by

$$f_\lambda^{**}(t) := \max\left[0, \min\Big\{F_X^{-1}(t), f_\lambda^*(t)\Big\}\right]. \tag{K.7}$$

It is then easy to check that $f_\lambda^{**} \in \mathcal{Q}$, since $f_\lambda^* \in \mathcal{Q}$ and since F_X^{-1} is a nondecreasing function. Moreover, $0 \leqslant f_\lambda^{**}(z) \leqslant F_X^{-1}(z)$, for each $z \in (0,1)$. Therefore, $f_\lambda^{**} \in \mathcal{Q}^*$. Finally, it is easily seen that $f_\lambda^{**}(t)$ solves Problem (F.5) for the given λ and t, since the concavity of u yields the concavity of the function $z \mapsto u\Big(W_0 - \Pi + z - F_X^{-1}(t)\Big) - \lambda \phi_{Q^*} \circ$

$F_X^{-1}(t)\, z$, for each $t \in (0,1)$. Hence, in view of Lemma K.3, it remains to show that there exists a $\lambda^* \geqslant 0$ such that $\int_0^1 \phi_{Q^*} \circ F_X^{-1}(t)\, f^{**}_{\lambda^*}(t)\ dt = R$.

Let ψ be the function of the parameter $\lambda \geqslant 0$ defined by

$$\psi(\lambda) := \int_0^1 \phi_{Q^*} \circ F_X^{-1}(t) \max\left[0, \min\left\{F_X^{-1}(t), (u')^{-1}\left(\lambda \phi_{Q^*} \circ F_X^{-1}(t)\right) + F_X^{-1}(t) - W_0 + \Pi\right\}\right] dt.$$

It then suffices to show that there exists a $\lambda^* \geqslant 0$ such that $\psi(\lambda^*) = R$. Since X is bounded and since F_X^{-1} is nondecreasing, it follows that for each $t \in [0,1]$,

$$\min\left\{F_X^{-1}(t), (u')^{-1}\left(\lambda \phi_{Q^*} \circ F_X^{-1}(t)\right) + F_X^{-1}(t) - W_0 + \Pi\right\} \leqslant F_X^{-1}(t) \leqslant F_X^{-1}(1) \leqslant \|X\|_{sup} < +\infty.$$

Moreover, since ϕ_{Q^*} is nonincreasing and nonnegative, we have for each $t \in [0,1]$, $0 \leqslant \phi_{Q^*} \circ F_X^{-1}(t) \leqslant \phi_{Q^*}\left(F_X^{-1}(0)\right)$. But since the function ϕ_{Q^*} is continuous, it is bounded on every closed and bounded subset of $\mathbb{R}$. Therefore, $\phi_{Q^*}\left(F_X^{-1}(0)\right) < +\infty$. Hence, for each $t \in [0,1]$,

$$\begin{aligned}\min\left\{F_X^{-1}(t), (u')^{-1}\left(\lambda \phi_{Q^*} \circ F_X^{-1}(t)\right) + F_X^{-1}(t) - W_0 + \Pi\right\} \phi_{Q^*} \circ F_X^{-1}(t) \\ \leqslant F_X^{-1}(1)\, \phi_{Q^*}\left(F_X^{-1}(0)\right) < +\infty.\end{aligned}$$

Hence, Lebesgue's Dominated Convergence Theorem [2, Theorem 11.21] implies that ψ is a continuous function of λ. Moreover, ψ is a nonincreasing function of λ (by concavity of u and by monotonicity of the Lebesgue integral).

Now, Assumption 4.1 implies that

$$\begin{aligned}\psi(0) &= \int_0^1 \phi_{Q^*} \circ F_X^{-1}(t)\, F_X^{-1}(t)\ dt = \int F_X^{-1}(U)\, \phi_{Q^*}\left(F_X^{-1}(U)\right)\ dP \\ &= \int X \phi_{Q^*}(X)\ dP = \int X\ dQ^*,\end{aligned}$$

and that

$$\begin{aligned}\lim_{\lambda \to +\infty} \psi(\lambda) &= \int_0^1 \phi_{Q^*} \circ F_X^{-1}(t) \max\left[0, \min\left\{F_X^{-1}(t), F_X^{-1}(t) - W_0 + \Pi\right\}\right] dt \\ &= \int_0^1 \phi_{Q^*} \circ F_X^{-1}(t) \max\left[0, F_X^{-1}(t) - W_0 + \Pi\right] dt = 0,\end{aligned}$$

where the last equality follows from the fact that $F_X(W_0 - \Pi) = 1$ (Assumption 5.3), and hence $F_X^{-1}(t) \leqslant W_0 - \Pi$, for all $t \in (0,1)$. Consequently,

$$0 = \lim_{\lambda \to +\infty} \psi(\lambda) \leqslant R < \int X\ dQ^* = \psi(0).$$

Hence, by the Intermediate Value Theorem [38, Theorem 4.23], there exists a $\lambda^* \geqslant 0$ such that $\psi(\lambda^*) = R$. This concludes the proof of Lemma K.4. □

Now, since $X = F_X^{-1}(U)$, P-a.s.,

$$\max\left[0, \min\left\{F_X^{-1}(U), (u')^{-1}\left(\lambda^* \phi_{Q*} \circ F_X^{-1}(U)\right) + F_X^{-1}(U) - W_0 + \Pi\right\}\right]$$
$$= \max\left[0, \min\left\{X, (u')^{-1}(\lambda^* \phi_{Q*} \circ X) + X - W_0 + \Pi\right\}\right], \ P\text{-a.s.}$$

Therefore, letting $\mathcal{Y}_2^* = \max\left[0, \min\left\{X, (u')^{-1}(\lambda^* \phi_{Q*} \circ X) + X - W_0 + \Pi\right\}\right]$, it follows that

$$\int \phi_{Q*} \circ X \mathcal{Y}_2^* \, dP = \int \mathcal{Y}^* \phi_{Q*} \circ X \, dP = \int \mathcal{Y}^* \, dQ^* = R,$$

and that

$$\int u\left(W_0 - \Pi - X + \mathcal{Y}_2^*\right) \, dP = \int u\left(W_0 - \Pi - X + \mathcal{Y}^*\right) \, dP.$$

In other words, $\mathcal{Y}_2^*$ is also optimal for Problem (5.3). Define Y^* by

$$Y^* = \min\left[X, \max\Big(0, X - d\left(\phi_{Q*}\right)\Big)\right], \quad \text{where} \quad d\left(\phi_{Q*}\right) = W_0 - \Pi - (u')^{-1}\Big(\lambda^* \phi_{Q*} \circ X\Big).$$

It can then easily be verified that $Y^* = \mathcal{Y}_2^*$. $\square$

References

[1] D. Alary, C. Gollier, and N. Treich. The Effect of Ambiguity Aversion on Insurance and Self-Protection. *The Economic Journal*, 123(573):1188–1202, 2013.

[2] C.D. Aliprantis and K.C. Border. *Infinite Dimensional Analysis - 3^{rd} edition*. Springer-Verlag, 2006.

[3] M. Amarante, M. Ghossoub, and E.S. Phelps. The Entrepreneurial Economy I: Contracting under Knightian Uncertainty. Columbia University, Center on Capitalism and Society, Working Paper No. 68 (April 2011).

[4] S. Anwar and M. Zheng. Competitive Insurance Market in the Presence of Ambiguity. *Insurance: Mathematics and Economics*, 50(1):79–84, 2012.

[5] K.J. Arrow. *Essays in the Theory of Risk-Bearing*. Chicago: Markham Publishing Company, 1971.

[6] C. Bernard, X. He, J.A. Yan, and X.Y. Zhou. Oprimal Insurance Design under Rank-Dependent Expected Utility. *Mathematical Finance*, Forthcoming.

[7] G. Carlier and R.A. Dana. Insurance Contracts with Deductibles and Upper Limits. Preprint, Ceremade, Universite Paris Dauphine (2002).

[8] G. Carlier and R.A. Dana. Core of Convex Distortions of a Probability. *Journal of Economic Theory*, 113(2):199–222, 2003.

[9] G. Carlier and R.A. Dana. Two-persons Efficient Risk-sharing and Equilibria for Concave Law-invariant Utilities. *Economic Theory*, 36(2):189–223, 2008.

[10] G. Carlier, R.A. Dana, and N. Shahidi. Efficient Insurance Contracts under Epsilon-Contaminated Utilities. *The GENEVA Papers on Risk and Insurance-Theory*, 28:59–71, 2003.

[11] N.L. Carothers. *Real Analysis*. Cambridge University Press, 2000.

[12] A. Chateauneuf, R.A. Dana, and J.M. Tallon. Optimal Risk-sharing Rules and Equilibria with Choquet-expected-utility. *Journal of Mathematical Economics*, 34(2):191–214, 2000.

[13] D.L. Cohn. *Measure Theory*. Boston: Birkhauser, 1980.

[14] D. Denneberg. *Non-Additive Measure and Integral*. Kluwer Academic Publishers, 1994.

[15] N.A. Doherty and L. Eeckhoudt. Optimal Insurance without Expected Utility: The Dual Theory and the Linearity of Insurance Contracts. *Journal of Risk and Uncertainty*, 10(2):157–179, 1995.

[16] H. Föllmer and A. Schied. *Stochastic Finance: An Introduction in Discrete Time – 3^{rd} ed.* Walter de Gruyter, 2011.

[17] M. Ghossoub. Arow's Theorem of the Deducible with Heterogeneous Beliefs. mimeo (2013).

[18] M. Ghossoub. Equimeasurable Rarrangements with Capacities. *Mathematics of Operations Research*, Forthcoming.

[19] M. Ghossoub. *Contracting under Heterogeneous Beliefs*. PhD thesis, University of Waterloo – Dept. of Statistics & Actuarial Science, May 2011.

[20] C. Gollier. Optimal Insurance Design of Ambiguous Risks. mimeo (2012).

[21] C. Gollier and H. Schlesinger. Arrow's Theorem on the Optimality of Deductibles: A Stochastic Dominance Approach. *Economic Theory*, 7:359–363, 1996.

[22] X.D. He and X.Y. Zhou. Portfolio Choice via Quantiles. *Mathematical Finance*, 21(2):203–231, 2011.

[23] R.M. Hogarth and H. Kunreuther. Risk, Ambiguity, and Insurance. *Journal of Risk and Uncertainty*, 2(1):5–35, 1989.

[24] P.J. Huber and E.M. Ronchetti. *Robust Statistics - Second Edition*. John Wiley and Sons, 2011.

[25] G. Huberman, D. Mayers, and C.W. Smith Jr. Optimal Insurance Policy Indemnity Schedules. *The Bell Journal of Economics*, 14(2):415–426, 1983.

[26] K. Inada. On a Two-Sector Model of Economic Growth: Comments and a Generalization. *The Review of Economic Studies*, 30(2):119–127, 1963.

[27] M. Jeleva. Background Risk, Demand for Insurance, and Choquet Expected Utility Preferences. *The GENEVA Papers on Risk and Insurance-Theory*, 25(1):7–28, 2000.

[28] H. Jin and X.Y. Zhou. Behavioral Portfolio Selection in Continous Time. *Mathematical Finance*, 18(3):385–426, 2008.

[29] J.B. Kadane and L. Wasserman. Symmetric, Coherent, Choquet Capacities. *The Annals of Statistics*, 24(3):1250–1264, 1996.

[30] E. Karni. Optimal Insurance: A Nonexpected Utility Analysis. In *G. Dionne (ed.), Contributions to Insurance Economics*. Kluwer Academic publishers, 1992.
[31] P. Klibanoff, M. Marinacci, and S. Mukerji. A Smooth Model of Decision Making under Ambiguity. *Econometrica*, 73(6):1849–1892, 2005.
[32] M.J. Machina. Non-Expected Utility and the Robustness of the Classical Insurance Paradigm. *The GENEVA Papers on Risk and Insurance-Theory*, 20(1):9–50, 1995.
[33] M. Marinacci and L. Montrucchio. Introduction to the Mathematics of Ambiguity. In *I. Gilboa (ed.), Uncertainty in Economic Theory: Essays in Honor of David Schmeidlers 65th Birthday*, pages 46–107. Routledge, London, 2004.
[34] P. Milgrom. Good News and Bad News: Representation Theorems and Applications. *The Bell Journal of Economics*, 12(2):380–391, 1981.
[35] D. Moffet. Optimal Deductible and Consumption Theory. *Journal of Risk and Insurance*, pages 669–682, 1977.
[36] J. Mossin. Aspects of Rational Insurance Purchasing. *The Journal of Political Economy*, 76(4):553–568, 1968.
[37] J. Quiggin. A Theory of Anticipated Utility. *Journal of Economic Behavior*, 3(4):323–343, 1982.
[38] W. Rudin. *Principles of Mathematical Analysis – 3^{rd} ed.* New York: McGraw-Hill Book Company, 1976.
[39] H. Schlesinger. Insurance Demand without the Expected-Utility Paradigm. *Journal of Risk and Insurance*, 64(1):19–39, 1997.
[40] D. Schmeidler. Integral Representation without Additivity. *Proceedings of the American Mathematical Society*, 97(2):255–261, 1986.
[41] D. Schmeidler. Subjective Probability and Expected Utility without Additivity. *Econometrica*, 57(3):571–587, 1989.
[42] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall, 1991.
[43] Z. Wang and G.J. Klir. *Generalized Measure Theory*. New York: Springer, 2009.
[44] L. Wasserman and J.B. Kadane. Symmetric Upper Probabilities. *The Annals of Statistics*, 20(4):1720–1736, 1992.
[45] M. Yaari. The Dual Theory of Choice under Risk. *Econometrica*, 55(1):95–115, 1987.
[46] V.R. Young. Optimal Insurance under Wang's Premium Principle. *Insurance Mathematics and Economics*, 25(2):109–122, 1999.

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

E-mail address: massimiliano.amarante@umontreal.ca

Mario Ghossoub: Imperial College London – South Kensington Campus – London, SW7 2AZ – United Kingdom

E-mail address: m.ghossoub@imperial.ac.uk

Edmund Phelps: Columbia University – International Affairs Building – 420 W. 118th St. – New York, NY, 10027 – USA

E-mail address: esp2@columbia.edu