

Article

Optimal Insurance for a Minimal Expected Retention: The Case of an Ambiguity-Seeking Insurer

Massimiliano Amarante ¹ and Mario Ghossoub ^{2,*}

¹ Département de Sciences Économiques, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, QC H3C 3J7, Canada; massimiliano.amarante@umontreal.ca

² Imperial College London, South Kensington Campus, London SW7 2AZ, UK

* Correspondence: m.ghossoub@imperial.ac.uk; Tel.: +44-020-7594-9130

Academic Editor: Alexander Szimayer

Received: 22 December 2015; Accepted: 9 March 2016; Published: 21 March 2016

Abstract: In the classical expected utility framework, a problem of optimal insurance design with a premium constraint is equivalent to a problem of optimal insurance design with a minimum expected retention constraint. When the insurer has ambiguous beliefs represented by a non-additive probability measure, as in Schmeidler, this equivalence no longer holds. Recently, Amarante, Ghossoub and Phelps examined the problem of optimal insurance design with a premium constraint when the insurer has ambiguous beliefs. In particular, they showed that when the insurer is ambiguity-seeking, with a concave distortion of the insured's probability measure, then the optimal indemnity schedule is a state-contingent deductible schedule, in which the deductible depends on the state of the world only through the insurer's distortion function. In this paper, we examine the problem of optimal insurance design with a minimum expected retention constraint, in the case where the insurer is ambiguity-seeking. We obtain the aforementioned result of Amarante, Ghossoub and Phelps and the classical result of Arrow as special cases.

Keywords: optimal insurance; deductible; minimum retention; ambiguity; Choquet integral; probability distortion

JEL: G22

1. Introduction

In the classical problem of optimal insurance design, it is well known since the work of Arrow [1] that when the insured is a risk-averse expected utility (EU) maximizer and the insurer is a risk-neutral EU maximizer, the indemnity schedule that maximizes the insured's expected utility of terminal wealth subject to a premium constraint is a deductible indemnity of the form $Y^* = \max(0, X - d)$, where X is the loss random variable and $d \geq 0$ is a given constant deductible.

For a given indemnity schedule Y , the premium constraint is a constraint of the form:

$$\Pi \geq (1 + \rho) \int Y dP$$

where $\Pi \geq 0$ is the premium paid and $\rho \geq 0$ is a loading factor. It requires that the premium received by the insurer is enough to cover the expected indemnity payments loaded up for expenses and profits. By letting $R_0 = \int X dP - \frac{\Pi}{1+\rho}$, the premium constraint can be written as:

$$\int (X - Y) dP \geq R_0 \tag{1}$$

The quantity $X - Y$ is a random variable that represents the portion of the loss that is actually suffered (or retained) by the insured after the insurance indemnification payment, and it is called the retention random variable. The constraint given in Equation (1) is often referred to as a minimum expected retention constraint. It requires that the indemnity schedule Y be such that the associated retention random variable has a minimum pre-specified expectation under the insurer's beliefs. Since a premium constraint and a minimum expected retention constraint are equivalent in the classical EU case, the optimal indemnity schedule with a minimum expected retention constraint is then also a deductible indemnity schedule, with a constant positive deductible level.

Both the premium constraint and the minimum expected retention constraint are a simple restatement of the risk-neutral EU-maximizing insurer's participation constraint. The insurer and the insured share the same probabilistic beliefs about the realizations of the loss random variable, and hence, Arrow's [1] classical result is a pure risk-sharing result. The applicability of this result is limited, nonetheless. Indeed, the classical theory of optimal insurance design due to Arrow is based on the classical theory of choice under uncertainty [2–4], which follows the Bayesian paradigm, whereby the uncertainty that a decision maker faces in a given decision problem is described by a probability measure over a space of contingencies or states of the world. In other words, the economic agent has a clear probabilistic assessment of the underlying uncertainty that he faces.

However, the seminal work of Knight [5] suggested that there might be situations where the information available to a decision maker is too coarse for him or her to be able to formulate an additive probability measure over the list of contingencies. These occurrences are typically referred to as situations of decision under *Knightian uncertainty*, or *ambiguity*. Yet, this did not penetrate the mainstream theory of choice until Ellsberg's [6] famous thought experiments, which can be seen as an indication of people's aversion to unknown unknowns, or vagueness in beliefs about likelihoods. There is now a substantial body of empirical evidence for the pervasiveness of ambiguity in situations of choice under uncertainty, and we refer to Camerer [7] for a still timely review.

Largely motivated by the Ellsberg paradox, modern decision theory has developed several models of decision under ambiguity that rationalize behaviours, such as the ones depicted by Ellsberg [6]. For example, in Schmeidler [8], ambiguity is represented by a non-additive subjective "probability" measure, called a *capacity*, and preferences are aggregated using an integral defined with respect to capacities: the *Choquet integral* (we review capacities and Choquet integration in Section 3.1 and in Appendix A). Schmeidler's [8] seminal work and his model of decision under ambiguity, which came to be known as Choquet expected utility (CEU), can be seen as the starting point of decision theoretic investigations of models of choice under ambiguity.

After Schmeidler's work, many axiomatic models of decision under ambiguity were introduced. In Gilboa and Schmeidler [9], ambiguity is described by a collection of additive priors (a set of probability measures), and preferences are aggregated using the minimum value of the usual (Lebesgue) integral over this collection. Ghirardato, Maccheroni and Marinacci [10] propose a general model of decision under ambiguity that includes that of [9]. Amarante [11] introduces a model of decision under ambiguity that includes the aforementioned ones. We refer to the recent survey of Gilboa and Marinacci [12] for more on this topic, including other models of decision under ambiguity and applications of these models to several problems in economic theory.

In particular, Amarante [11] shows that most models of decision under ambiguity can be represented as models where the objects of choice are evaluated by a Choquet integral with respect to some capacity. That is, Choquet integration is a wide enough aggregation concept for preferences that it can encompass most models of decision under ambiguity and, in particular, the most popular ones. As a result, we focus in this paper on Choquet integration as an aggregation concept for decision-making under ambiguity, and we consider the CEU model of Schmeidler [8] as a model of decision-making under ambiguity.

1.1. Ambiguity in Optimal Insurance Design

Empirical evidence suggests that ambiguity, rather than risk, is prevalent in insurance pricing and underwriting and that often, insurers tend to exhibit more ambiguity than the insured individuals (e.g., [13]). Motivated by these findings, Amarante, Ghossoub and Phelps (AGP) [14] re-examined the classical insurance demand problem of Arrow [1] in a setting where the insurer has ambiguous beliefs (in the sense of Schmeidler [8]) about the realizations of the insurable loss, whereas the insured is an EU maximizer. Specifically, they examined the problem of a risk-averse EU-maximizing insured who wishes to choose the indemnity schedule that maximizes his or her expected utility of terminal wealth, subject to a premium constraint based on the insurer's ambiguous beliefs and that takes the form:

$$\frac{\Pi}{1 + \rho} \geq - \int -Y dv \geq 0 \quad (2)$$

where integration is in the sense of Choquet (Section 3.1 and Appendix A). This premium constraint is a restatement of the insurer's participation constraint. In particular, AGP [14] showed that when the insurer is ambiguity-seeking, with a concave distortion of the insured's probability measure, then the optimal indemnity schedule is a state-contingent deductible schedule, in which the deductible depends on the state of the world only through the insurer's distortion function. Arrow's result then obtains as a special case when the insurer does not distort probabilities.

Because of the properties of the Choquet integral, the premium constraint in Equation (2) cannot be re-written as a minimum expected retention constraint. Therefore, in the case where the insurer is a Choquet-expected utility maximizer as in Schmeidler [8], a problem of insurance design given a premium constraint is not equivalent to a problem of insurance design given a minimum expected retention constraint, where the expectation is in the sense of Choquet, that is it is an ambiguous (or non-additive) expectation.

1.2. Related Literature

There is a growing literature devoted to the study of ambiguity in insurance design and, more broadly, in situations of contracting or risk sharing. For example, Alary *et al.* [15] examine a problem of optimal insurance design in which the insured is ambiguity-averse in the sense of Klibanoff, Marinacci and Mukerji [16]. The authors assume that, conditional on a non-zero loss occurring, the loss severity distribution is not ambiguous; however, the probability that a loss occurs is ambiguous. The authors then show that, in this situation, the optimal indemnity is a straight deductible. Gollier [17] also examines a problem of optimal insurance design in which the insured is ambiguity-averse in the sense of Klibanoff, Marinacci and Mukerji [16], and he shows that the optimal indemnity schedule contains a disappearing deductible when the insured's collection of (additive) priors can be ordered according to the monotone-likelihood ratio property.

Jeleva [18] considers an insurance model in which the insured behaves according to Schmeidler's CEU model. She assumes that the optimal insurance contract is of the co-insurance type and then looks for the optimal co-insurance factor. Young [19] and Bernard *et al.* [20] examine a problem of optimal insurance design in which the insured is a rank-dependent expected utility maximizer [21,22]. Doherty and Eeckhoudt [23] study the optimal level of deductible under Yaari's dual theory [22]. Karni [24] considers an insurance model in which the preferences of the insured can accommodate ambiguous beliefs, but they satisfy certain differentiability criteria. He shows that a deductible indemnity schedule is optimal in that case. Machina [25] considers a similar setting where the preferences of the insured have a non-EU representation that satisfies certain differentiability criteria, and he examines the optimal level of co-insurance and the optimal level of deductible. Schlesinger [26] 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 mean-preserving increases in risk.

The papers mentioned above all assume that the ambiguity is on the side of the insured. Motivated by empirical evidence suggesting that insurers tend to exhibit more ambiguity than the insured individuals (e.g., [13]), AGP [14] study the problem of optimal insurance design in a setting where the insurer has ambiguous beliefs about the realizations of the insurable loss (and behaves according to Schmeidler's CEU model), whereas the insured does not.

Carlier *et al.* [27] examine ambiguity on both sides: the insurer's and the insured's. Specifically, they 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 [28] also examine ambiguity on both sides, but they consider a model with only two states of the world: a no-loss state and a loss state. As such, this is of limited scope, since the shape of an optimal indemnity schedule cannot be determined in a two-state world.

Carlier and Dana [29–31] and Chateauneuf *et al.* [32] examine more general problems of risk sharing and contracting that are relevant to the insurance problem considered here. However, they do not give a full analytical characterization of the optimal risk sharing rule, which is one of the main goals of the present paper in the context of an insurance design problem.

Finally, we would like to mention the recent work of Balbas *et al.* [33] who consider ambiguity on the side of both the insurer and the insured¹. The authors assume that both parties behave according to the maxmin expected utility (MEU) model of Gilboa and Schmeidler [9], that is their ambiguity is described by a collection of additive priors (a set of probability measures), and preferences are aggregated using the minimum value of the usual (Lebesgue) integral over this collection of priors. The authors do not consider the retention random variable as a decision variable, but rather its sensitivity (*i.e.*, mathematical derivative) with respect to total claims. This is meant to prevent some potential moral hazard issues. The authors then show that the optimal indemnity schedule is often such that the sensitivity of the retention random variable to the total claims saturates the problem's constraints (a bang-bang solution).

1.3. This Paper's Contribution

This paper is meant to supplement the literature and provide a framework for scholars or practitioners who are used to considering retention as a decision variable, but would like to examine the effect of ambiguity. Specifically, we examine the problem of optimal insurance design for a given minimum expected retention level, when the insurer is ambiguity-seeking in the sense of Schmeidler [8]. We focus on the case of an ambiguity-seeking rather than an ambiguity-averse insurer, since recent work suggests that situations of ambiguity-seeking are empirically relevant (e.g., [34–36]). We examine two cases:

- (1) We first examine the problem of insurance design when the insurer's beliefs are a concave distortion of a probability measure that differs from the insurer's probability measure. We show that there is an event A to which the insurer assigns full probability and on which the optimal indemnity schedule is a state-contingent deductible schedule, with a state-contingent deductible that is a function of the insurer's distortion function and of a likelihood ratio between the two parties' beliefs. On the complement of the event A , the optimal indemnity is full insurance. However, the insurer assigns zero probability to the complement of A .
- (2) As a special case of the above, we examine the problem of insurance design when the insurer's beliefs are a concave distortion of the insured's probability measure. We show that in this case, the result of AGP [14] mentioned above still holds: the optimal indemnity schedule is a state-contingent deductible schedule, in which the deductible depends on the state of the world only through the insurer's distortion function. Arrow's result then obtains as a special case when the insurer does not distort probabilities.

¹ We thank one of the reviewers for bringing this to our attention.

1.4. Outline

The rest of this paper is organized as follows. In Section 2, we review the classical optimal insurance design problem. In Section 3, we examine the problem of optimal insurance design given a minimum expected retention constraint, in the case of an insurer who is ambiguity-seeking in the sense of Schmeidler [8]. Proofs and related analysis are given in the [Appendices](#).

2. Optimal Insurance: The Classical Case

2.1. Setup and Preliminaries

Consider a collection S of states of the world, and endow S with a σ -algebra \mathcal{F} of events. Let $B(\mathcal{F})$ denote the linear space of all bounded, \mathbb{R} -valued and \mathcal{F} -measurable functions on (S, \mathcal{F}) . Let $B^+(\mathcal{F})$ denote the collection of all \mathbb{R}^+ -valued elements of $B(\mathcal{F})$. Any $f \in B(\mathcal{F})$ is bounded, and we define its sup norm by $\|f\|_{sup} := \sup\{|f(s)| : s \in S\} < +\infty$.

In the classical problem of optimal insurance design given a premium constraint, a risk-averse EU-maximizing individual with initial wealth $W_0 > 0$ faces an insurable random loss, against which he or she seeks insurance. This random loss is represented by a bounded, nonnegative random variable X on the state space. That is, the random loss is a given element of $B^+(\mathcal{F})$ 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, for any $Y \in B(\Sigma)$, there exists a Borel-measurable map $I : \mathbb{R} \rightarrow \mathbb{R}$, such that $Y = I \circ X$ ([37], Theorem 4.41). Let $B^+(\Sigma)$ denote the collection of nonnegative elements of $B(\Sigma)$, and let P be a probability measure on (S, Σ) . We will make the following assumption all throughout.

Assumption 1. The random loss X is a continuous random variable² on the probability space (S, Σ, P) . That is, the Borel probability measure $P \circ X^{-1}$ is nonatomic³.

The individual seeking an insurance coverage against this random loss X has the possibility of purchasing an insurance indemnity schedule $Y = I(X)$, which pays the amount $I(X(s)) \geq 0$, in the state of the world $s \in S$. That is, $Y \in B^+(\Sigma)$, and we can hence 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 $\Pi \geq 0$. The premium is determined by the insurer, based on the insurer's beliefs about the realizations of X (and hence, of Y).

2.2. The Insurance Design Problem

In the classical insurance model of Arrow [1], both the insurer and the insured are EU maximizers, having the same non-ambiguous beliefs about the realizations of X . The insurer is assumed to be a risk-neutral EU maximizer, with a linear utility function v , which we assume to be the identity function. This is without loss of generality, as utility functions are defined up to a positive linear transformation. The insured is risk-averse, having a concave increasing utility function u .

After purchasing an indemnity schedule $Y = I \circ X$, the insured's wealth is the random variable W defined by $W(s) = W_0 - \Pi - X(s) + Y(s)$, in each state of the world $s \in S$. The insured's problem is that of choosing an indemnity schedule that maximizes his or her expected utility of wealth,

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 (Ω, \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 $0 < \eta(B) < \eta(A)$.

for a given premium Π . Specifically, the insured's problem 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 $0 \leq Y \leq X$, and subject to a *premium constraint* of the form:

$$\Pi \geq (1 + \rho) \int Y dP \quad (3)$$

where $\rho \geq 0$ is a given loading factor.

Arrow's [1] classical result states that in this case, the optimal insurance indemnity schedule is a deductible insurance schedule:

Theorem 1 (Arrow). *The optimal indemnity schedule is a deductible schedule given by:*

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

where $d \geq 0$ is a constant, such that $\int Y^* dP = \frac{\Pi}{1+\rho}$.

2.3. Premium Constraint vs. Minimal Expected Retention Constraint

In this classical framework, it is easy to see how the premium constraint given in Equation (3) is equivalent to a constraint for the form:

$$\int (X - Y) dP \geq R_0 \quad (4)$$

for $R_0 = \int X dP - \frac{\Pi}{1+\rho}$. For a given indemnity schedule $Y \in B^+(\Sigma)$, the random variable

$$R = X - Y$$

is the amount of loss that is retained by the insured individual and is called the retention random variable. The constraint given in Equation (4) is called a *minimum retention constraint*.

Therefore, the classical problem of demand for insurance is equivalent to the following problem of optimal insurance design with a minimal expected retention:

For a given $R_0 \geq 0$

$$\begin{aligned} & \sup_{Y \in B(\Sigma)} \int u(W_0 - \Pi - X + Y) dP \\ & \text{s.t. } 0 \leq Y \leq X, \\ & \int (X - Y) dP \geq R_0 \end{aligned} \quad (5)$$

By the monotonicity of u , if $R_0 = 0$, then full insurance is optimal, that is $Y^* = X$ is an optimal solution for Problem (5). When $R_0 > 0$, Theorem 1 implies that the optimal solution is a deductible indemnity schedule with a constant, positive deductible.

One should note the equivalence between a premium constraint and a minimum expected retention constraint holds only under the expected value premium principle, *i.e.*, under Expected Utility Theory. Insurance pricing, risk measurement and premium principles under non-Expected Utility models or models of Ambiguity have been recently examined in the literature (see, e.g. Pichler [38] and Ghossoub [39]). Under such non-linear premium principles, the equivalence between a premium constraint and a minimum retention constraint does not hold.

3. The Case of an Ambiguity-Seeking Insurer

3.1. Preliminaries: Capacities, Choquet Integration and the CEU Model

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

- (1) $v(\emptyset) = 0$;
- (2) $v(S) = 1$; and,
- (3) v is monotone: for any $A, B \in \Sigma$, $A \subseteq B \Rightarrow v(A) \leq v(B)$.

The capacity v is said to be:

- *supermodular* (or convex) if $v(A \cup B) + v(A \cap B) \geq v(A) + v(B)$, for all $A, B \in \Sigma$; and,
- *submodular* (or concave) if $v(A \cup B) + v(A \cap B) \leq v(A) + v(B)$, for all $A, B \in \Sigma$.

For instance, if (S, Σ, P) is a probability space and $T : [0, 1] \rightarrow [0, 1]$ is an increasing function, such that $T(0) = 0$ and $T(1) = 1$, then the set function $v := T \circ P$ 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 (respectively concave), then the capacity $v = T \circ P$ is supermodular (respectively submodular) ([40] Ex. 2.1).

Definition 2. Let v_1 be a supermodular capacity and v_2 a submodular capacity on (S, Σ) .

- The *core* of v_1 , denoted by $core(v_1)$, is the collection of all probability measures Q on (S, Σ) , such that $Q(A) \geq v_1(A), \forall A \in \Sigma$.
- The *anti-core* of v_2 , denoted by $acore(v_2)$, is the collection of all probability measures Q on (S, Σ) , such that $Q(A) \leq v_2(A), \forall A \in \Sigma$.

Both $core(v_1)$ and $acore(v_2)$ are weak*-compact and convex collections of probability measures on the space (S, Σ) .

Definition 3. Let v be a capacity on (S, Σ) . The *Choquet integral* of $Y \in B(\Sigma)$ with respect to v is defined by:

$$\int Y dv := \int_0^{+\infty} v(\{s \in S : Y(s) \geq t\}) dt + \int_{-\infty}^0 [v(\{s \in S : Y(s) \geq t\}) - 1] dt$$

where the integrals are taken in the sense of Riemann.

The Choquet integral with respect to a (countably additive) measure is the usual a Lebesgue integral with respect to that measure ([41] p. 59). By a classical result of Schmeidler [42], we can represent Choquet integrals with respect to a supermodular (respectively submodular) capacity as a lower (respectively upper) envelope of Lebesgue integrals:

Proposition 1 (Schmeidler [42]). *Let v be a capacity on (S, Σ) , and let $Y \in B(\Sigma)$:*

- *If v is supermodular, then $\int Y dv = \min \left\{ \int Y dP : P \in core(v_1) \right\}$;*
- *If v is submodular, then $\int Y dv = \max \left\{ \int Y dP : P \in acore(v_2) \right\}$.*

Definition 4. Two functions $Y_1, Y_2 \in B(\Sigma)$ are said to be comonotonic if:

$$\left[Y_1(s) - Y_1(s') \right] \left[Y_2(s) - Y_2(s') \right] \geq 0, \text{ for all } s, s' \in S$$

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. Appendix A provides some additional information about capacities and Choquet integration.

In the CEU model of Schmeidler [8], the ambiguous beliefs of a decision maker are represented by a non-additive probability measure (a capacity) on the state space, and preferences are aggregated using the concept of Choquet integration.

The decision maker’s preferences over alternatives induce a unique non-additive probability measure ν over the state space and a utility function u , such that an alternative Y is evaluated by $\int u(Y) d\nu$, where integration is in the sense of Choquet. Moreover, in the CEU model, ambiguity aversion is equivalent to the convexity of the capacity, and ambiguity loving is equivalent to the concavity of the capacity.

3.2. The Insurance Design Problem

Here, we assume that the insurer has ambiguous beliefs about the realizations of the loss random variable X . His or her ambiguous beliefs are represented by a non-additive probability measure (a capacity) on the state space, as in the CEU model of Schmeidler [8].

In this case, an indemnity schedule Y will be deemed acceptable by the insurer if the associated retention $R = X - Y$ yields a non-additive expectation of at least $R_0 > 0$ under the insurer’s ambiguous beliefs. In other words, the minimum retention constraint in this context is given by:

$$\int (X - Y) d\nu \geq R_0 \tag{6}$$

where integration is in the sense of Choquet. It is important to note that since the Choquet integral is not additive in general and only positively homogeneous (Proposition A1), the constraint given in Equation (6) cannot be written as a premium constraint of the form $\int Y d\nu \leq \frac{\Pi}{1+\rho}$ for some $\rho > 0$. This is a fundamental point of departure from the classical case.

Recently, AGP [14] studied the problem of optimal insurance design given a premium constraint, when the insurer has ambiguous beliefs in the sense of the CEU model of Schmeidler [8]. Specifically, they study the problem:

$$\begin{aligned} & \sup_{Y \in B(\Sigma)} \int u(W_0 - \Pi - X + Y) dP \tag{7} \\ \text{s.t. } & 0 \leq Y \leq X, \\ & - \int -Y d\nu \leq \frac{\Pi}{1 + \rho} \end{aligned}$$

The constraint $- \int -Y d\nu \leq \frac{\Pi}{1+\rho}$ is a restatement of the insurer’s participation constraint:

$$\int (W_0^{ins} + \Pi - (1 + \rho) Y) d\nu \geq W_0^{ins}$$

where W_0^{ins} is the insurer’s initial wealth.

Here, for a given $R_0 \geq 0$, the insured’s problem is the following:

$$\begin{aligned} & \sup_{Y \in B(\Sigma)} \int u(W_0 - \Pi - X + Y) dP \tag{8} \\ \text{s.t. } & 0 \leq Y \leq X, \\ & \int (X - Y) d\nu \geq R_0 \end{aligned}$$

As mentioned above, due to the properties of the Choquet integral, problems Equations (7) and (8) are not equivalent.

Assumption 2. The insured’s utility function u satisfies the following properties:

- (1) $u(0) = 0$;
- (2) u is strictly increasing and strictly concave;
- (3) u is continuously differentiable.
- (4) The first derivative satisfies $u'(0) = +\infty$ and $\lim_{x \rightarrow +\infty} u'(x) = 0$.

Thus, in particular, we assume that the insured is risk-averse. We also make the assumption 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:

Assumption 3. $X \leq W_0 - \Pi$, P -a.s. That is, $P\left(\left\{s \in S : X(s) > W_0 - \Pi\right\}\right) = 0$.

In this paper, we consider the case of an insurer who is an ambiguity-seeking insurer in the sense of Schmeidler [8]. That is, the capacity ν is submodular (concave; Definition 1). Specifically, we consider the case where the insurer’s capacity ν is of the form $\nu = T \circ Q$, for some probability measure $Q \neq P$ on (S, Σ) and some function $T : [0, 1] \rightarrow [0, 1]$, which is increasing, concave and continuous with $T(0) = 0$ and $T(1) = 1$. Then, $T \circ Q$ is a submodular capacity on (S, Σ) .

Assumption 4. We assume that $\nu = T \circ Q$, where:

- (1) Q is a probability measure on (S, Σ) , such that $Q \circ X^{-1}$ is nonatomic;
- (2) $T : [0, 1] \rightarrow [0, 1]$ is increasing, concave and twice differentiable; and
- (3) $T(0) = 0$ and $T(1) = 1$.

Henceforth, we examine the following problem, for a given $R_0 \geq 0$:

$$\begin{aligned} & \sup_{Y \in B(\Sigma)} \int u(W_0 - \Pi - X + Y) dP & (9) \\ & \text{s.t. } 0 \leq Y \leq X, \\ & \int (X - Y) dT \circ Q \geq R_0 \end{aligned}$$

3.3. A Characterization of the Optimal Indemnity Schedule

For each $Z \in B^+(\Sigma)$, let $F_Z(t) = Q(\{s \in S : Z(s) \leq t\})$, denote the *distribution function* of Z with respect to the probability measure Q , and let $F_Z^{-1}(t)$ be the left-continuous inverse of the distribution function F_Z (that is, the *quantile function* of Z), defined by:

$$F_Z^{-1}(t) = \inf \{z \in \mathbb{R}^+ : F_Z(z) \geq t\}, \quad \forall t \in [0, 1] \tag{10}$$

Definition 5. Denote by \mathcal{AQuant} the collection of all quantile functions f of the form F^{-1} , where F is the distribution function of some $Z \in B^+(\Sigma)$, such that $0 \leq Z \leq X$.

That is, \mathcal{AQuant} is the collection of all quantile functions f that satisfy the following properties:

- (1) $f(z) \leq F_X^{-1}(z)$, for each $0 < z < 1$;
- (2) $f(z) \geq 0$, for each $0 < z < 1$.

Denoting by $Quant = \left\{ f : (0, 1) \rightarrow \mathbb{R} \mid f \text{ is nondecreasing and left-continuous} \right\}$ the collection of all quantile functions, we can then write $\mathcal{A}Quant$ as follows:

$$\mathcal{A}Quant = \left\{ f \in Quant : 0 \leq f(z) \leq F_X^{-1}(z), \text{ for each } 0 < z < 1 \right\} \tag{11}$$

By Lebesgue’s decomposition theorem ([37], Th. 10.61) there exists a unique pair (P_{ac}, P_s) of (nonnegative) finite measures on (S, Σ) , such that:

- $P = P_{ac} + P_s$;
- $P_{ac} \ll Q$ (P_{ac} is absolutely continuous with respect to Q); and,
- $P_s \perp Q$ (P_s and Q are mutually singular).

That is, for all $B \in \Sigma$, $P_{ac}(B) = 0$, whenever $Q(B) = 0$. Moreover, there exists some $A \in \Sigma$, such that $Q(S \setminus A) = P_s(A) = 0$, which then implies that $P_{ac}(S \setminus A) = 0$ and $Q(A) = 1$. Note also that for all $Z \in B(\Sigma)$, $\int Z dP = \int_A Z dP_{ac} + \int_{S \setminus A} Z dP_s$. Furthermore, by the Radon–Nikodým Theorem ([43], Th. 4.2.2), there exists a Q -a.s. unique Σ -measurable and Q -integrable function $h : S \rightarrow [0, +\infty)$, such that $P_{ac}(C) = \int_C h dQ$, for all $C \in \Sigma$. Hence, for all $Z \in B(\Sigma)$, $\int Z dP = \int_A Zh dQ + \int_{S \setminus A} Z dP_s$. Furthermore, since $P_{ac}(S \setminus A) = 0$, it follows that $\int_{S \setminus A} Z dP_s = \int_{S \setminus A} Z dP$. Thus, for all $Z \in B(\Sigma)$, $\int Z dP = \int_A Zh dQ + \int_{S \setminus A} Z dP$.

Moreover, since $h : S \rightarrow [0, +\infty)$ is Σ -measurable and Q -integrable, there exists a Borel-measurable and $Q \circ X^{-1}$ -integrable map $\phi : X(S) \rightarrow [0, +\infty)$, such that $h = dP_{ac}/dQ = \phi \circ X$. We will also make the following assumption, which can be interpreted as a kind of monotone likelihood ratio property.

Assumption 5. The Σ -measurable function $h = \phi \circ X = dP_{ac}/dQ$ is comonotonic with X , i.e., ϕ is nondecreasing.

Since $Q \circ X^{-1}$ is nonatomic (by Assumption 4), it follows that $F_X(X)$ has a uniform distribution over $(0, 1)$ ([44], Lemma A.21), that is $Q(\{s \in S : F_X(X)(s) \leq t\}) = t$ for each $t \in (0, 1)$. Letting $U := F_X(X)$, it follows that U is a random variable on the probability space (S, Σ, Q) with a uniform distribution on $(0, 1)$. Consider the following quantile problem:

$$\begin{aligned} & \sup_f \int u(W_0 - \Pi - f(U)) \phi(F_X^{-1}(U)) dQ & (12) \\ \text{s.t. } & f \in \mathcal{A}Quant, \\ & \int T'(1 - U) f(U) dQ = R_0 \end{aligned}$$

The following theorem characterizes the solution of the insured’s problem (Problem (9)) in terms of the solution of the relatively easier quantile problem given in Problem (12), provided the previous assumptions hold. The proof is given in Appendix C.

Theorem 2. Under the previous assumptions, if f^* is optimal for Problem (12), then the function:

$$Y^* = (X - f^*(U))\mathbf{1}_A + X\mathbf{1}_{S \setminus A}$$

is optimal for Problem (9).

In particular, $Y^* = X - f^*(U)$, Q -a.s. That is, the set E of states of the world s , such that $Y^*(s) \neq (X - f^*(U))(s)$, has probability zero under the probability measure Q (and hence, $v(E) = T \circ Q(E) = 0$). The optimal indemnity schedule will be seen by the insurer to be almost surely equal to the function $X - f^*(U)$.

Another immediate implication of Theorem 2 is that the collection of states of the world in which the optimal indemnity schedule is a full insurance rule is a set of states to which the insurer assigns a zero likelihood. On the set of all other states of the world, the optimal indemnity schedule deviates from the full insurance rule by the function $f^*(U)$.

Under the following two assumptions, it is possible to fully characterize the shape of an optimal insurance contract. This is done in Corollary 1.

Assumption 6. The Σ -measurable function $h = \phi \circ X = dP_{ac}/dQ$ is such that ϕ is left-continuous.

Assumption 7. the function $t \mapsto \frac{T'(1-t)}{\phi(F_X^{-1}(t))}$, defined on $t \in (0, 1) \setminus \{t : \phi \circ F_X^{-1}(t) = 0\}$, is nondecreasing.

Assumption 7 is also a monotone likelihood ratio type assumption. Similar assumptions have been used in Jin and Zhou [45] in their study of portfolio choice under prospect theory, in He and Zhou [46] in their study of a portfolio choice problem under Yaari’s [22] dual theory of choice, in Jin and Zhou [47] in their study of greed and leverage within a portfolio choice problem under prospect theory and in Carlier and Dana [48] in their study of the demand for contingent claims under rank-dependent expected utility [21]. Furthermore, note that when both the insured and the insurer are Bayesian decision makers, that is when the insurer does not distort probabilities, then Assumption 7 and Assumption 5 are equivalent⁴.

When the previous assumptions hold, we can give an explicit characterization of an optimal indemnity schedule, as follows.

Corollary 1 (The Shape of an Optimal Indemnity Schedule). Under the previous assumptions, the function Y^* defined below is an optimal solution for Problem (9):

$$Y^* = \min \left[X, \max \left(0, X - d(T, \phi) \right) \right] \mathbf{1}_A + X\mathbf{1}_{S \setminus A}$$

where:

$$d(T, \phi) = W_0 - \Pi - (u')^{-1} \left(\frac{-\lambda^* T'(1-U)}{\phi(X)} \right)$$

⁴ Indeed, suppose that $T(t) = t$, the identity function. Then, $T' = 1$, and so, $\frac{T'(1-t)}{\phi(F_X^{-1}(t))} = \frac{1}{\phi(F_X^{-1}(t))}$. Moreover, the function $t \mapsto F_X^{-1}(t)$ is nondecreasing, since F_X^{-1} is a quantile function. Therefore, the function $t \mapsto \frac{T'(1-t)}{\phi(F_X^{-1}(t))}$ is nondecreasing if and only if the function ϕ is nonincreasing.

U is a random variable on the probability space (S, Σ, Q) with a uniform distribution on $(0, 1)$, and $\lambda^* \leq 0$ is chosen, so that:

$$\int T'(1 - U) \max \left[0, \min \left\{ X, W_0 - \Pi - (u')^{-1} \left(\frac{-\lambda^* T'(1 - U)}{\phi(X)} \right) \right\} \right] dQ = R_0$$

The proof of Corollary 1 is given in Appendix D. Note that if Assumption 5 holds, then Assumption 6 is a weak assumption⁵.

3.4. A Special Case

Finally, we consider the special case where $Q = P$, in which the insurer’s submodular capacity v is a concave distortion of the insured’s probability measure P of the form $v = T \circ P$. In this case, the insured’s problem becomes:

$$\begin{aligned} & \sup_{Y \in B(\Sigma)} \int u(W_0 - \Pi - X + Y) dP & (13) \\ \text{s.t. } & 0 \leq Y \leq X, \\ & \int (X - Y) dT \circ P \geq R_0 \end{aligned}$$

where the distortion function $T : [0, 1] \rightarrow [0, 1]$ is increasing, concave, twice differentiable and satisfies $T(0) = 0$ and $T(1) = 1$.

Since $P = Q$, we have:

- $P = P_{ac} = Q$;
- $h = \phi \circ X = dP_{ac}/dQ$ is the constant function equal to one. Hence, Assumptions 5 and 6 trivially hold;
- $A = S$, and hence, $S \setminus A = \emptyset$;
- $U = F_X(X)$ is a random variable on the probability space (S, Σ, P) with a uniform distribution on $(0, 1)$.

Since, moreover, the random variable $V = 1 - U$ is also a random variable on the probability space (S, Σ, P) with a uniform distribution on $(0, 1)$, Corollary 1 then becomes:

Corollary 2. *If Assumptions 2, 3 and 7 hold and if $P \circ X^{-1}$ is nonatomic, then the function Y^* defined below is an optimal solution for Problem (13):*

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

where:

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

V is a random variable on the probability space (S, Σ, P) with a uniform distribution on $(0, 1)$, and $\lambda^* \geq 0$ is chosen, so that:

$$\int T'(V) \max \left[0, \min \left\{ X, W_0 - \Pi - (u')^{-1} (\lambda^* T'(V)) \right\} \right] dP = R_0$$

⁵ Any monotone function is Borel-measurable and, hence, “almost continuous”, in view of Lusin’s Theorem ([49] Theorem 7.5.2). Furthermore, any monotone function is almost surely continuous for the Lebesgue measure.

Corollary 2 states that when the insurer's ambiguous beliefs are a concave distortion of the insured's non-ambiguous beliefs, then an indemnity schedule that solves the optimal insurance design problem given a *minimum retention constraint* is a state-contingent deductible schedule, with a state-contingent deductible given by:

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

for an appropriately-chosen $\lambda^* \geq 0$. This is an identical result to Theorem 5.4 of AGP [14], which states that when the insurer's ambiguous beliefs are a concave distortion of the insured's non-ambiguous beliefs, then an indemnity schedule that solves the optimal insurance design problem given a *premium constraint* is a state-contingent deductible schedule, with a state-contingent deductible given by:

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

for an appropriately-chosen $\lambda^* \geq 0$.

Hence, when the insurer's ambiguous beliefs are a concave distortion of the insured's non-ambiguous beliefs, a problem of optimal insurance design with a *minimum retention constraint* is equivalent to a problem of optimal insurance design with a *premium constraint*. The optimal indemnity schedule is given by a state-contingent deductible schedule⁶, in which the state-contingent deductible $d(T)$ depends on the state of the world only through the insurer's distortion function T . Clearly, when the insurer does not distort probabilities, so that the function T is the identity function, one recovers Arrow's result as a special case.

Acknowledgments: We thank Daniel Bauer, Enrico Biffis, Glenn Harisson, Stefan Trautmann, George Zanjani, two anonymous reviewers and the audiences at Georgia State University and the 2014 CEAR/MRIC Behavioural Insurance Workshop for their comments and suggestions.

Author Contributions: Both authors have contributed equally to all aspects of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. More on Capacities and Choquet Integration

Definition A1. A capacity v on (S, Σ) is continuous from above (respectively below) if for any sequence $\{A_n\}_{n \geq 1} \subseteq \Sigma$, such that $A_{n+1} \subseteq A_n$ (respectively $A_{n+1} \supseteq A_n$) for each n , it holds that:

$$\lim_{n \rightarrow +\infty} v(A_n) = v\left(\bigcap_{n=1}^{+\infty} A_n\right) \quad \left(\text{respectively } \lim_{n \rightarrow +\infty} v(A_n) = v\left(\bigcup_{n=1}^{+\infty} A_n\right)\right)$$

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

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

The Choquet integral with respect to a (countably additive) measure is the usual Lebesgue integral with respect to that measure ([41] 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 A1. Let v be a capacity on (S, Σ) .

(1) If $\phi_1, \phi_2 \in B(\Sigma)$, then in general, $\int (\phi_1 + \phi_2) dv \neq \int \phi_1 dv + \int \phi_2 dv$.

⁶ Note that this variable deductible is anti-comonotonic with the loss X , since T is concave, u is increasing and concave (and hence, $(u')^{-1}$ is decreasing by the inverse function theorem ([50] pp. 221–223)), $X = F_X^{-1}(V)$, $P - a.s.$, and F_X is a nondecreasing function. However, the indemnity schedule is comonotonic with X .

- (2) If $\phi_1, \phi_2 \in B(\Sigma)$ are comonotonic, then $\int (\phi_1 + \phi_2) dv = \int \phi_1 dv + \int \phi_2 dv$.
- (3) If $\phi_1, \phi_2 \in B(\Sigma)$ are such that $\phi_1 \leq \phi_2$, then $\int \phi_1 dv \leq \int \phi_2 dv$.
- (4) For all $\phi \in B(\Sigma)$ and all $c \geq 0$, then $\int c\phi dv = c \int \phi dv$.
- (5) If v is submodular, then for any $\phi_1, \phi_2 \in B(\Sigma)$, $\int (\phi_1 + \phi_2) dv \leq \int \phi_1 dv + \int \phi_2 dv$.

Appendix B. 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 [51,52].

Appendix B.1. The Nondecreasing Rearrangement

Consider the setting of Section 2.1, and let ζ be the probability law of X defined by $\zeta(B) := P \circ X^{-1}(B) = P(\{s \in S : X(s) \in B\})$ for any Borel subset B of \mathbb{R} .

Definition A2. For any Borel-measurable map $I : [0, M] \rightarrow \mathbb{R}$, define the *distribution function* of I as the map $\zeta_I : \mathbb{R} \rightarrow [0, 1]$ defined by:

$$\zeta_I(t) := \zeta(\{x \in [0, M] : I(x) \leq t\}) \tag{A1}$$

Then, ζ_I is a nondecreasing right-continuous function.

Definition A3. Let $I : [0, M] \rightarrow [0, M]$ be any Borel-measurable map, and define the function $\tilde{I} : [0, M] \rightarrow \mathbb{R}$ by:

$$\tilde{I}(t) := \inf \{z \in \mathbb{R}^+ \mid \zeta_I(z) \geq \zeta([0, t])\} \tag{A2}$$

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

Proposition A2. Let $I : [0, M] \rightarrow [0, M]$ be any Borel-measurable map, and let $\tilde{I} : [0, M] \rightarrow \mathbb{R}$ be defined as in Equation (A2). Then, the following hold:

- (1) \tilde{I} is left-continuous, nondecreasing and Borel-measurable;
- (2) $\tilde{I}(0) = 0$ and $\tilde{I}(M) \leq M$. Therefore, $\tilde{I}([0, M]) \subseteq [0, M]$;
- (3) If $I_1, I_2 : [0, M] \rightarrow [0, M]$ are such that $I_1 \leq I_2$, ζ -a.s., then $\tilde{I}_1 \leq \tilde{I}_2$;
- (4) \tilde{I} is ζ -equimeasurable with I , in the sense that for any Borel set B ,

$$\zeta(\{t \in [0, M] : I(t) \in B\}) = \zeta(\{t \in [0, M] : \tilde{I}(t) \in B\}) \tag{A3}$$

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

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

- (1) $Y, \tilde{Y} \in B^+(\Sigma)$, since I and \tilde{I} are Borel-measurable mappings of $[0, M]$ into itself;
- (2) \tilde{Y} is a nondecreasing function of X :
 $[X(s) \leq X(s')] \Rightarrow [\tilde{Y}(s) \leq \tilde{Y}(s')]$, for all $s, s' \in S$; and
- (3) Y and \tilde{Y} have the same distribution under P (i.e., they are P -equimeasurable):
 $P(\{s \in S : Y(s) \leq \alpha\}) = P(\{s \in S : \tilde{Y}(s) \leq \alpha\})$, for any $\alpha \in [0, M]$.

Call \tilde{Y} a nondecreasing P -rearrangement of Y with respect to X and denote it by \tilde{Y}_P . Then, \tilde{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 A1. Fix $Y \in B^+(\Sigma)$, and let \tilde{Y}_P denote the nondecreasing P -rearrangement of Y with respect to X . If $0 \leq Y \leq X$, P -a.s., then $0 \leq \tilde{Y}_P \leq X$.

Appendix B.2. Supermodularity

Definition A4. A function $L : \mathbb{R}^2 \rightarrow \mathbb{R}$ is *supermodular* if for any $x_1, x_2, y_1, y_2 \in \mathbb{R}$ with $x_1 \leq x_2$ and $y_1 \leq y_2$, one has:

$$L(x_2, y_2) + L(x_1, y_1) \geq L(x_1, y_2) + L(x_2, y_1) \tag{A4}$$

A function $L : \mathbb{R}^2 \rightarrow \mathbb{R}$ 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:

$$L(x_2, y_2) + L(x_1, y_1) > L(x_1, y_2) + L(x_2, y_1) \tag{A5}$$

Lemma A2. A function $L : \mathbb{R}^2 \rightarrow \mathbb{R}$ is *supermodular* (respectively *strictly supermodular*) if and only if the function $\eta(y) := L(x+h, y) - L(x, y)$ is nondecreasing (respectively increasing) on \mathbb{R} , for any $x \in \mathbb{R}$ and $h \geq 0$ (respectively $h > 0$).

Example A1. If $g : \mathbb{R} \rightarrow \mathbb{R}$ is concave and $a \in \mathbb{R}$, then the function $L_1 : \mathbb{R}^2 \rightarrow \mathbb{R}$ 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 A3 (Hardy–Littlewood). Fix $Y \in B^+(\Sigma)$, and let \tilde{Y}_P denote the nondecreasing P -rearrangement of Y with respect to X . If L is supermodular, then (assuming integrability) we have:

$$\int L(X, Y) dP \leq \int L(X, \tilde{Y}_P) dP$$

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

Appendix C. Proof of Theorem 2

Denote by \mathcal{F}_{SB} the feasibility set for Problem (9):

$$\mathcal{F}_{SB} = \left\{ Y \in B(\Sigma) : 0 \leq Y \leq X \text{ and } \int (X - Y) dv \geq R_0 \right\} \tag{A6}$$

Let $\mathcal{F}_{SB}^=$ be defined as:

$$\mathcal{F}_{SB}^= = \left\{ Y \in B(\Sigma) : 0 \leq Y \leq X \text{ and } \int (X - Y) dv = R_0 \right\} \tag{A7}$$

To rule out trivial situations, we will make the following assumption:

Assumption A1. $\mathcal{F}_{SB}^= \neq \emptyset$.

Assumption A1 then implies that $\mathcal{F}_{SB} \neq \emptyset$. By monotonicity of the Choquet integral, the non-emptiness of \mathcal{F}_{SB} implies that:

$$\int X dv \geq R_0$$

Appendix C.1. “Splitting”

Recall that by Lebesgue’s decomposition theorem ([37], Th. 10.61) there exists a unique pair (P_{ac}, P_s) of (nonnegative) finite measures on (S, Σ) , such that $P = P_{ac} + P_s$, $P_{ac} \ll Q$ and $P_s \perp Q$. That is, for all $B \in \Sigma$ with $Q(B) = 0$, we have $P_{ac}(B) = 0$, and there is some $A \in \Sigma$, such that $Q(S \setminus A) = P_s(A) = 0$. It then also follows that $P_{ac}(S \setminus A) = 0$ and $Q(A) = 1$. In the following, the Σ -measurable set A on which Q is concentrated is assumed to be fixed all throughout. Consider now the following two problems:

For a given $\beta \geq R_0$,

$$\begin{aligned} & \sup_{Y \in B(\Sigma)} \int_A u(W_0 - \Pi - X + Y) dP & (A8) \\ \text{s.t. } & 0 \leq Y \leq X, \\ & \int (X - Y) dT \circ Q = \beta \end{aligned}$$

and:

$$\begin{aligned} & \sup_{Y \in B(\Sigma)} \int_{S \setminus A} u(W_0 - \Pi - X + Y) dP & (A9) \\ \text{s.t. } & 0 \leq Y \mathbf{1}_{S \setminus A} \leq X \mathbf{1}_{S \setminus A}, \\ & \int_{S \setminus A} (X - Y) dT \circ Q = 0 \end{aligned}$$

Remark A1. Since the function u is continuous (Assumption 2), it is bounded on any closed and bounded subset of \mathbb{R} . Therefore, since the range of X is closed and bounded, the supremum of each of the above two problems is finite when their feasibility sets are non-empty. Now, the function X is feasible for Problem (A9), and so, Problem (A9) has a non-empty feasibility set.

Definition A5. For a given $\beta \geq R_0$, let $\Theta_{A,\beta}$ be the feasibility set of Problem (A8) with parameter β . That is,

$$\Theta_{A,\beta} := \left\{ Y \in B^+(\Sigma) : 0 \leq Y \leq X, \int (X - Y) dT \circ Q = \beta \right\}$$

Denote by Γ the collection of all β for which the feasibility set $\Theta_{A,\beta}$ is non-empty:

Definition A6. Let $\Gamma := \{ \beta \geq R_0 : \Theta_{A,\beta} \neq \emptyset \}$.

Lemma A4. $\Gamma \neq \emptyset$.

Proof. By Assumption A1, $R_0 \in \Gamma$. \square

Lemma A5. X is optimal for Problem (A9).

Proof. The feasibility of X for Problem (A9) is clear. To show optimality, let Y be any feasible solution for Problem (A9). Then, for each $s \in S \setminus A$, $Y(s) \leq X(s)$. Therefore, since u is increasing, we have $u(W_0 - \Pi - X(s) + Y(s)) \leq u(W_0 - \Pi - X(s) + X(s)) = u(W_0 - \Pi)$, for each $s \in S \setminus A$. Thus,

$$\int_{S \setminus A} u(W_0 - \Pi - X + Y) dP \leq \int_{S \setminus A} u(W_0 - \Pi - X + X) dP = u(W_0 - \Pi)P(S \setminus A)$$

\square

Remark A2. Since $Q(S \setminus A) = 0$ and $T(0) = 0$, it follows that $T \circ Q(S \setminus A) = 0$, and so, $\int \mathbf{1}_{S \setminus A} dT \circ Q = T \circ Q(S \setminus A) = 0$, by Proposition A1. Therefore, for any $Z \in B^+(\Sigma)$, it follows from the monotonicity and positive homogeneity of the Choquet integral (Proposition A1) that:

$$0 \leq \int_{S \setminus A} Z dT \circ Q = \int Z \mathbf{1}_{S \setminus A} dT \circ Q \leq \int \|Z\|_s \mathbf{1}_{S \setminus A} dT \circ Q = \|Z\|_s \int \mathbf{1}_{S \setminus A} dT \circ Q = 0$$

and so, $\int_{S \setminus A} Z dT \circ Q = 0$. Consequently, it follows from Proposition A1 that for any $Z \in B^+(\Sigma)$,

$$\int Z dT \circ Q \leq \int Z \mathbf{1}_A dT \circ Q = \int_A Z dT \circ Q$$

Now, consider the following problem:

$$\sup_{\beta \in \Gamma} \left\{ F_A^*(\beta) : F_A^*(\beta) \text{ is the supremum of Problem (A8), for a fixed } \beta \in \Gamma \right\} \tag{A10}$$

Lemma A6. If β^* is optimal for Problem (A10) and if Y_1^* is optimal for Problem (A8) with parameter β^* , then $Y^* := Y_1^* \mathbf{1}_A + X \mathbf{1}_{S \setminus A}$ is optimal for Problem (9).

Proof. By the feasibility of Y_1^* for Problem (A8) with parameter β^* , we have $0 \leq Y_1^* \leq X$ and $\int (X - Y_1^*) dT \circ P = \beta^*$. Therefore, $0 \leq Y^* \leq X$, and:

$$\begin{aligned} \int (X - Y^*) dT \circ Q &= \int [(X - Y_1^*) \mathbf{1}_A + (X - X) \mathbf{1}_{S \setminus A}] dT \circ Q \\ &= \int_A (X - Y_1^*) dT \circ Q \geq \int (X - Y_1^*) dT \circ Q = \beta^* \geq R_0 \end{aligned}$$

where the inequality $\int_A (X - Y_1^*) dT \circ Q \geq \int (X - Y_1^*) dT \circ Q$ follows from the same argument as in Remark A2. Hence, Y^* is feasible for Problem (9). To show the optimality of Y^* for Problem (9), let \bar{Y} be any other feasible function for Problem (9), and define α by $\alpha = \int (X - \bar{Y}) dT \circ Q$. Then, $\alpha \geq R_0$, and so, \bar{Y} is feasible for Problem (A8) with parameter α ; and α is feasible for Problem (A10). Hence:

$$F_A^*(\alpha) \geq \int_A u(W_0 - \Pi - X + \bar{Y}) dP$$

Now, since β^* is optimal for Problem (A10), it follows that $F_A^*(\beta^*) \geq F_A^*(\alpha)$. Moreover, \bar{Y} is feasible for Problem (A9) (since $0 \leq \bar{Y} \leq X$, and so, $\int_{S \setminus A} (X - \bar{Y}) dT \circ Q = 0$ by Remark A2). Thus,

$$\begin{aligned} F_A^*(\beta^*) + u(W_0 - \Pi)P(S \setminus A) &\geq F_A^*(\alpha) + u(W_0 - \Pi)P(S \setminus A) \\ &\geq \int_A u(W_0 - \Pi - X + \bar{Y}) dP + u(W_0 - \Pi)P(S \setminus A) \\ &\geq \int_A u(W_0 - \Pi - X + \bar{Y}) dP + \int_{S \setminus A} u(W_0 - \Pi - X + \bar{Y}) dP \\ &= \int u(W_0 - \Pi - X + \bar{Y}) dP \end{aligned}$$

However, $F_A^*(\beta^*) = \int_A u(W_0 - \Pi - X + Y_1^*) dP$. Therefore,

$$\int u(W_0 - \Pi - X + Y^*) dP = F_A^*(\beta^*) + u(W_0 - \Pi)P(S \setminus A) \geq \int u(W_0 - \Pi - X + \bar{Y}) dP.$$

Hence, Y^* is optimal for Problem (9). \square

Lemma A7. R_0 is optimal for Problem (A10).

Proof. The feasibility of R_0 for Problem (A10) follows from Assumption A1. The optimality of R_0 for Problem (A10) follows from the monotonicity of the utility function u appearing in Problem (A8) and of the Choquet integral, which imply that the function $\beta \mapsto F_A^*(\beta)$ is nonincreasing. \square

By Lemma A7, we can restrict ourselves to solving Problem (A8) with parameter $\beta = R_0$. That is, Lemmata A6 and A7 imply the following result.

Lemma A8. Consider the problem:

$$\begin{aligned} & \sup_{Y \in B(\Sigma)} \int_A u(W_0 - \Pi - X + Y) dP & (A11) \\ \text{s.t. } & 0 \leq Y \leq X, \\ & \int (X - Y) dT \circ Q = R_0 \end{aligned}$$

If Y_1^* is optimal for Problem (A11), then $Y^* := Y_1^* \mathbf{1}_A + X \mathbf{1}_{S \setminus A}$ is optimal for Problem (9).

Appendix C.2. Solving Problem (A11)

Recall that for all $Z \in B(\Sigma)$, $\int Z dP = \int_A Zh dQ + \int_{S \setminus A} Z dP$, where $h = dP_{ac}/dQ$ is the Radon–Nikodým derivative of P_{ac} with respect to Q . Moreover, by definition of the set $A \in \Sigma$, we have $Q(S \setminus A) = P_s(A) = 0$. Therefore, $\int_A Zh dQ = \int Zh dQ$, for each $Z \in B(\Sigma)$. Hence, we can rewrite Problem (A11) (recalling that $h = \phi \circ X$) as the following problem:

$$\begin{aligned} & \sup_{Y \in B(\Sigma)} \int u(W_0 - \Pi - X + Y) \phi(X) dQ & (A12) \\ \text{s.t. } & 0 \leq Y \leq X, \\ & \int (X - Y) dT \circ Q = R_0 \end{aligned}$$

We then obtain the following result immediately:

Lemma A9. If Y_1^* is optimal for Problem (A12), then it is optimal for Problem (A11), and hence, the function $Y^* := Y_1^* \mathbf{1}_A + X \mathbf{1}_{S \setminus A}$ is optimal for Problem (9).

Now, consider the following problem:

$$\begin{aligned} & \sup_{Z \in B(\Sigma)} \int u(W_0 - \Pi - Z) \phi(X) dQ & (A13) \\ \text{s.t. } & 0 \leq Z \leq X, \\ & \int Z dT \circ Q = R_0 = \int_0^{+\infty} T(Q(\{s \in S : Z(s) \geq t\})) dt \end{aligned}$$

Lemma A10. If Z^* is optimal for Problem (A13), then $Y^* := X - Z^*$ is optimal for Problem (A12). Consequently, the function $(X - Z^*) \mathbf{1}_A + X \mathbf{1}_{S \setminus A}$ is optimal for Problem (9).

Proof. Suppose that Z^* is optimal for Problem (A13), and define $Y^* := X - Z^*$. Then, $Y^* \in B(\Sigma)$. Moreover, since $0 \leq Z^* \leq X$, it follows that $0 \leq Y^* \leq X$. Now,

$$\int (X - Y^*) dT \circ Q = \int (X - (X - Z^*)) dT \circ Q = \int Z^* dT \circ Q = R_0$$

and so, Y^* is feasible for Problem (A12). To show optimality of Y^* for Problem (A12), suppose, by way of contradiction, that $\bar{Y} \neq Y^*$ is feasible for Problem (A12) and:

$$\int u(W_0 - \Pi - X + \bar{Y}) h dQ > \int u(W_0 - \Pi - X + Y^*) h dQ$$

that is, with $\bar{Z} := X - \bar{Y}$, we have:

$$\int u(W_0 - \Pi - \bar{Z}) h dQ > \int u(W_0 - \Pi - Z^*) h dQ$$

Now, since $0 \leq \bar{Y} \leq X$ and $\int (X - \bar{Y}) dT \circ Q = R_0$, \bar{Z} is feasible for Problem (A13), hence contradicting the optimality of Z^* for Problem (A13). Thus, $Y^* := X - Z^*$ is optimal for Problem (A12). The rest follows from Lemma A9. \square

Definition A7. If $Z_1, Z_2 \in B^+(\Sigma)$ are feasible for Problem (A13), we will say that Z_2 is an improvement of Z_1 (or is improving) when the following hold:

- (1) $\int u(W_0 - \Pi - Z_2) h dQ \geq \int u(W_0 - \Pi - Z_1) h dQ$; and,
- (2) $\int Z_2 dT \circ Q \geq \int Z_1 dT \circ Q$.

The next result shows that for any feasible claim for Problem (A13), there is a another feasible claim for Problem (A13), which is comonotonic with X and improving.

Lemma A11. If Z is feasible for Problem (A13) and if \tilde{Z} is the nondecreasing Q -rearrangement of Z with respect to X , then \tilde{Z} is feasible for Problem (A13), comonotonic with X and an improvement of Z .

Proof. Let Z be feasible for Problem (A13), and note that by Assumption 5, the map $\zeta(X, Z) := u(W_0 - \Pi - Z) \phi(X)$ is supermodular (see Example A1). Let \tilde{Z} denote the nondecreasing Q -rearrangement of Z with respect to X . Then, by Lemma A1 and by equimeasurability of Z and \tilde{Z} , the function $\tilde{\zeta}$ is feasible for Problem (A13). Furthermore, by Lemma A3 and by the supermodularity of $\zeta(X, Z)$, it follows that \tilde{Z} is an improvement of Z . \square

Appendix C.3. Quantile Reformulation

Let $Z \in B^+(\Sigma)$ be feasible for Problem (A13); let $F_Z(t) = Q(\{s \in S : Z(s) \leq t\})$ denote the distribution function of Z with respect to the probability measure Q ; and let $F_X(t) = Q(\{s \in S : X(s) \leq t\})$ denote the distribution function of X with respect to the probability measure Q . Let $F_Z^{-1}(t)$ be the left-continuous inverse of the distribution function F_Z (that is, the quantile function of Z), defined by:

$$F_Z^{-1}(t) = \inf \{z \in \mathbb{R}^+ : F_Z(z) \geq t\}, \forall t \in [0, 1]$$

Let \tilde{Z} denote the nondecreasing Q -rearrangement of Z with respect to X . Since $Z \in B^+(\Sigma)$, it can be written as $\psi \circ X$ for some nonnegative Borel-measurable and bounded map ψ on $X(S)$. Moreover, since $0 \leq Z \leq X$, ψ is a mapping of $[0, M]$ into $[0, M]$. Let $\zeta := Q \circ X^{-1}$ be the image measure of Q under X . By Assumption 4 (1), ζ is nonatomic. We can then define the

mapping $\tilde{\psi} : [0, M] \rightarrow [0, M]$ as in Appendix B (see Equation (A2) on p. 14) to be the nondecreasing ζ -rearrangement of ψ , that is,

$$\tilde{\psi}(t) := \inf \left\{ z \in \mathbb{R}^+ : \zeta(\{x \in [0, M] : \psi(x) \leq z\}) \geq \zeta([0, t]) \right\}$$

Then, as in Appendix B, $\tilde{Z} = \tilde{\psi} \circ X$. Therefore, for each $s_0 \in S$,

$$\tilde{Z}(s_0) = \tilde{\psi}(X(s_0)) = \inf \left\{ z \in \mathbb{R}^+ : \zeta(\{x \in [0, M] : \psi(x) \leq z\}) \geq \zeta([0, X(s_0)]) \right\}$$

However, for each $s_0 \in S$,

$$\zeta([0, X(s_0)]) = Q \circ X^{-1}([0, X(s_0)]) = F_X(X(s_0)) := F_X(X)(s_0)$$

Moreover,

$$\begin{aligned} \zeta(\{x \in [0, M] : \psi(x) \leq z\}) &= Q \circ X^{-1}(\{x \in [0, M] : \psi(x) \leq z\}) \\ &= Q(\{s \in S : \psi(X(s)) \leq z\}) = F_Z(z) \end{aligned}$$

Consequently, for each $s_0 \in S$,

$$\tilde{Z}(s_0) = \inf \left\{ z \in \mathbb{R}^+ : F_Z(z) \geq F_X(X)(s_0) \right\} = F_Z^{-1}(F_X(X)(s_0)) := F_Z^{-1}(F_X(X))(s_0)$$

That is,

$$\tilde{Z} = F_Z^{-1}(F_X(X)) \tag{A14}$$

where F_Z^{-1} is the left-continuous inverse of F_Z , as defined in Equation (10).

Hence, by Lemma A11 and Equation (A14), we can restrict ourselves to finding a solution to Problem (A13) of the form $F^{-1}(F_X(X))$, where F is the distribution function of a function $Z \in B^+(\Sigma)$, such that $0 \leq Z \leq X$ and $\int Z dT \circ Q = R_0$. Moreover, since X is a nondecreasing function of X and Q -equimeasurable with X , it follows from the Q -a.s. uniqueness of the equimeasurable nondecreasing Q -rearrangement (see Appendix B) that $X = F_X^{-1}(F_X(X))$, Q -a.s. (see also [44, Lemma A.21]). Thus, for any $Z \in B^+(\Sigma)$,

$$\begin{aligned} \int u(W_0 - \Pi - F_Z^{-1}(F_X(X))) \phi(F_X^{-1}(F_X(X))) dQ &= \int u(W_0 - \Pi - \tilde{Z}) \phi(X) dQ \\ &\geq \int u(W_0 - \Pi - Z) \phi(X) dQ \end{aligned}$$

where the inequality follows from the proof of Lemma A11. Moreover, since $\zeta = Q \circ X^{-1}$ is nonatomic (by Assumption 4), it follows that $F_X(X)$ has a uniform distribution over $(0, 1)$ ([44] Lemma A.21), that is, $Q(\{s \in S : F_X(X)(s) \leq t\}) = t$ for each $t \in (0, 1)$. Finally, letting $U := F_X(X)$, we have:

$$\begin{aligned} \int F^{-1}(U) dT \circ Q &= \int_0^{+\infty} T[Q(\{s \in S : F^{-1}(U)(s) \geq t\})] dt \\ &= \int_0^{+\infty} T[Q(\{s \in S : F^{-1}(U)(s) > t\})] dt \\ &= \int_0^{+\infty} T[1 - F(t)] dt \\ &= \int_0^1 T'(1 - t) F^{-1}(t) dt = \int T'(1 - U) F^{-1}(U) dQ \end{aligned}$$

where the third and last equalities above follow from the fact that U has a uniform distribution over $(0, 1)$ and where the second-to-last equality follows from a standard argument⁷.

Now, recall from Definition 5 that $\mathcal{A}Quant$ given in Equation (11) is the collection of all *admissible quantile functions*, that is the collection of all functions f of the form F^{-1} , where F is the distribution function of a function $Z \in B^+(\Sigma)$, such that $0 \leq Z \leq X$, and consider the following problem:

$$\begin{aligned} & \sup_f \int u(W_0 - \Pi - f(U)) \phi(F_X^{-1}(U)) dQ & (A15) \\ & \text{s.t. } f \in \mathcal{A}Quant, \\ & \int T'(1 - U) f(U) dQ = R_0 \end{aligned}$$

Lemma A12. *If f^* is optimal for Problem (A15), then the function $f^*(U)$ is optimal for Problem (A13), where $U := F_X(X)$. Moreover, $X - f^*(U)$ is optimal for Problem (A12). Consequently, the function $(X - f^*(U)) \mathbf{1}_A + X \mathbf{1}_{S \setminus A}$ is optimal for Problem (9).*

Proof. Suppose that $f^* \in \mathcal{A}Quant$ is optimal for Problem (A15), and let $Z^* \in B^+(\Sigma)$ be a corresponding function. That is, f^* is the quantile function of Z^* and $0 \leq Z^* \leq X$. Let $\tilde{Z}^* := f^*(U)$. Then, \tilde{Z}^* is the equimeasurable nondecreasing Q -rearrangement of Z^* with respect to X , and so, $0 \leq \tilde{Z}^* \leq X$ by Lemma A1. Moreover,

$$\begin{aligned} R_0 &= \int T'(1 - U) f^*(U) dQ = \int f^*(U) dT \circ Q \\ &= \int \tilde{Z}^* dT \circ Q = \int_0^{+\infty} T[Q(\{s \in S : \tilde{Z}^*(s) \geq t\})] dt \\ &= \int_0^{+\infty} T[Q(\{s \in S : Z^*(s) \geq t\})] dt = \int \tilde{Z}^* dT \circ Q \end{aligned}$$

where the second-to-last equality follows from the Q -equimeasurability of Z^* and \tilde{Z}^* . Therefore, $\tilde{Z}^* = f^*(U)$ is feasible for Problem (A13). To show optimality, let Z be any feasible solution for Problem (A13), and let F be the distribution function for Z . Then, by Lemma A11, the function $\tilde{Z} := F^{-1}(U)$ is feasible for Problem (A13), comonotonic with X and Pareto-improving. Moreover, \tilde{Z} has also F as a distribution function. To show the optimality of $\tilde{Z}^* = f^*(U)$ for Problem (A13), it remains to show that:

$$\int u(W_0 - \Pi - \tilde{Z}^*) \phi(X) dQ \geq \int u(W_0 - \Pi - \tilde{Z}) \phi(X) dQ$$

Now, let $f := F^{-1}$, so that $\tilde{Z} = f(U)$. Since \tilde{Z} is feasible for Problem (A13), we have:

$$\begin{aligned} R_0 &= \int \tilde{Z} dT \circ Q = \int F^{-1}(U) dT \circ Q \\ &= \int_0^1 T'(1 - t) F^{-1}(t) dt = \int T'(1 - U) f(U) dQ \end{aligned}$$

Hence, f is feasible for Problem (A15). Since f^* is optimal for Problem (A15), we have:

$$\int u(W_0 - \Pi - f^*(U)) \phi(F_X^{-1}(U)) dQ \geq \int u(W_0 - \Pi - f(U)) \phi(F_X^{-1}(U)) dQ$$

⁷ See, e.g. Denneberg [40], Proposition 1.4 on p. 8 and the discussion on pp. 61–62. See also [45], p. 418, [46], p. 210 and p. 213, or [30], p. 207.

Moreover, since $X = F_X^{-1}(U)$, Q -a.s., we have:

$$\int u(W_0 - \Pi - \tilde{Z}^*) \phi(X) dQ \geq \int u(W_0 - \Pi - \tilde{Z}) \phi(X) dQ$$

Therefore, $\tilde{Z}^* = f^*(U)$ is optimal for Problem (A13). Hence, by Lemma A10, $Y^* := X - \tilde{Z}^* = X - f^*(U)$ is optimal for Problem (A12). Finally, by Lemma A10, the function $(X - f^*(U)) \mathbf{1}_A + X \mathbf{1}_{S \setminus A}$ is optimal for Problem (9).

□

Lemma A12 completes the proof of Theorem 2. □

Appendix D. Proof of Corollary 1

Corollary A1. Under the assumptions of Corollary 1, an optimal solution Y^* for Problem (9) takes the following form:

$$Y^* = \left(X - \max \left[0, \min \left\{ F_X^{-1}(U), f_{\lambda^*}^*(U) \right\} \right] \right) \mathbf{1}_A + X \mathbf{1}_{S \setminus A}$$

where for each $t \in (0, 1) \setminus \{t : \phi \circ F_X^{-1}(t) = 0\}$,

$$f_{\lambda^*}^*(t) = W_0 - \Pi - (u')^{-1} \left(\frac{-\lambda^* T'(1-t)}{\phi(F_X^{-1}(t))} \right)$$

and λ^* is chosen so that:

$$\int_0^1 T'(1-t) \max \left[0, \min \left\{ F_X^{-1}(t), f_{\lambda^*}^*(t) \right\} \right] dt = R_0$$

Proof. Recall from Equation (11) that

$$\mathcal{A}Quant = \left\{ f \in Quant : 0 \leq f(z) \leq F_X^{-1}(z), \text{ for each } 0 < z < 1 \right\}$$

where $Quant = \left\{ f : (0, 1) \rightarrow \mathbb{R} \mid f \text{ is nondecreasing and left-continuous} \right\}$. Define the collection \mathcal{K} of functions on $(0, 1)$ as follows:

$$\mathcal{K} = \left\{ f : (0, 1) \rightarrow \mathbb{R} \mid 0 \leq f(z) \leq F_X^{-1}(z), \text{ for each } 0 < z < 1 \right\} \tag{A16}$$

Then, $\mathcal{A}Quant = Quant \cap \mathcal{K}$. Consider the following problem.

$$\begin{aligned} & \sup_f \int_0^1 u(W_0 - \Pi - f(t)) \phi(F_X^{-1}(t)) dt \\ & \text{s.t. } f \in \mathcal{A}Quant, \\ & \int_0^1 T'(1-t) f(t) dt = R_0 \end{aligned} \tag{A17}$$

Lemma A13. If $f^* \in \mathcal{A}Quant$ satisfies the following:

$$(1) \quad \int_0^1 T'(1-t) f^*(t) dt = R_0;$$

(2) There exists $\lambda \leq 0$, such that for all $t \in (0, 1) \setminus \{t : \phi \circ F_x^{-1}(t) = 0\}$,

$$f^*(t) = \arg \max_{0 \leq y \leq F_x^{-1}(t)} \left[u(W_0 - \Pi - y) \phi \left(F_x^{-1}(t) \right) - \lambda T'(1-t)y \right]$$

Then, f^* solves Problem (A17).

Proof. Suppose that $f^* \in \mathcal{A}Quant$ satisfies Conditions (1) and (2) above. Then, in particular, f^* is feasible for Problem (A17). To show the optimality of f^* for Problem (A17), let f by any other feasible solution for Problem (A17). Then, for all $t \in (0, 1) \setminus \{t : \phi \circ F_x^{-1}(t) = 0\}$,

$$\begin{aligned} u(W_0 - \Pi - f^*(t)) \phi \left(F_x^{-1}(t) \right) - \lambda T'(1-t)f^*(t) \\ \geq u(W_0 - \Pi - f(t)) \phi \left(F_x^{-1}(t) \right) - \lambda T'(1-t)f(t) \end{aligned}$$

That is, $\left[u(W_0 - \Pi - f^*(t)) - u(W_0 - \Pi - f(t)) \right] \phi \left(F_x^{-1}(t) \right) \geq \lambda T'(1-t) \left[f^*(t) - f(t) \right]$. Integrating yields $V(f^*) - V(f) \geq \lambda [R_0 - R_0] = 0$, that is $V(f^*) \geq V(f)$, as required. \square

Hence, in view of Lemma A13, in order to find a solution for Problem (A17), one can start by solving the problem:

$$\max_{0 \leq f_\lambda(t) \leq F_x^{-1}(t)} \left[u(W_0 - \Pi - f_\lambda(t)) \phi \left(F_x^{-1}(t) \right) - \lambda T'(1-t)f_\lambda(t) \right] \tag{A18}$$

for a given $\lambda \leq 0$ and a fixed $t \in (0, 1) \setminus \{t : \phi \circ F_x^{-1}(t) = 0\}$.

Consider first the following problem:

$$\max_{f_\lambda(t)} \left[u(W_0 - \Pi - f_\lambda(t)) \phi \left(F_x^{-1}(t) \right) - \lambda T'(1-t)f_\lambda(t) \right] \tag{A19}$$

for a given $\lambda \leq 0$ and a fixed $t \in (0, 1) \setminus \{t : \phi \circ F_x^{-1}(t) = 0\}$.

By concavity of the utility function u , in order to solve Problem (A19), it suffices to solve for the first-order condition:

$$-u'(W_0 - \Pi - f_\lambda^*(t)) \phi \left(F_x^{-1}(t) \right) - \lambda T'(1-t) = 0$$

which gives

$$f_\lambda^*(t) = W_0 - \Pi - (u')^{-1} \left(\frac{-\lambda T'(1-t)}{\phi \left(F_x^{-1}(t) \right)} \right) \tag{A20}$$

Then, the function $f_\lambda^*(t)$ solves Problem (A19), for a given $\lambda \leq 0$ and a fixed $t \in (0, 1) \setminus \{t : \phi \circ F_x^{-1}(t) = 0\}$.

By Assumption 7, the function $t \mapsto \frac{T'(1-t)}{\phi \left(F_x^{-1}(t) \right)}$ is nondecreasing. By Assumption 2, the function u is strictly concave and continuously differentiable. Hence, the function u' is both continuous and strictly decreasing. This then implies that $(u')^{-1}$ is continuous and strictly decreasing, by the inverse function theorem ([50] pp. 221–223). Therefore, the function $f_\lambda^*(t)$ in Equation (A20) is nondecreasing ($\lambda \leq 0$). Moreover, by Assumption 4, the function T' is continuous. Therefore, Assumption 6 then implies that $f_\lambda^*(t)$ is left-continuous.

Define the function f_λ^{**} by:

$$f_\lambda^{**}(t) = \max \left[0, \min \left\{ F_X^{-1}(t), f_\lambda^*(t) \right\} \right] \tag{A21}$$

Then, $f_\lambda^{**}(t) \in \mathcal{K}$. Moreover, since both F_X^{-1} and f_λ^* are nondecreasing and left-continuous functions, it follows that f_λ^{**} is nondecreasing and left-continuous. Consequently, $f_\lambda^{**}(t) \in \mathcal{AQuant}$. Finally, it is easily seen that $f_\lambda^{**}(t)$ solves Problem (A18) for the given λ . Now, if λ^* is chosen, so that $\int_0^1 T'(1-t) f_{\lambda^*}^{**}(t) dt = R_0$, then by Lemma A13, $f_{\lambda^*}^{**}$ is optimal for Problem (A17).

Hence, in view of Theorem 2 and Lemma A13, to conclude the proof of Corollary A1, it remains to show that there exists a $\lambda^* \leq 0$, such that $\int_0^1 T'(1-t) f_{\lambda^*}^{**}(t) dt = R_0$. This is given by Lemma A14 below.

Lemma A14. *Let ψ be the function of the parameter $\lambda \leq 0$ defined by $\psi(\lambda) := \int_0^1 T'(1-t) f_\lambda^{**}(t) dt$. Then, there exists a $\lambda^* \leq 0$, such that $\psi(\lambda^*) = R_0$.*

Proof. First note that ψ is a continuous and nonincreasing function of λ , where the continuity of ψ is a consequence of Lebesgue’s dominated convergence theorem ([37], Theorem 11.21). Indeed, since X is bounded and since F_X^{-1} is nondecreasing, it follows that for each $t \in [0, 1]$,

$$\min \left\{ F_X^{-1}(t), f_\lambda^*(t) \right\} \leq F_X^{-1}(t) \leq F_X^{-1}(1) \leq M = \|X\|_s < +\infty$$

Moreover, since T is concave and increasing, T' is nonincreasing and nonnegative, and so, for each $t \in [0, 1]$, $0 \leq T'(1-t) \leq T'(0)$. However, $T'(0) < +\infty$, by Assumption 4 (since T' is differentiable, it is continuous and, therefore, bounded on every closed and bounded subset of \mathbb{R}). Hence, for each $t \in [0, 1]$,

$$\min \left\{ F_X^{-1}(t), f_\lambda^*(t) \right\} T'(1-t) \leq F_X^{-1}(1) T'(0) \leq \|X\|_s T'(0) < +\infty$$

Furthermore, $\psi(0) = 0$ (by Assumption 2), and:

$$\begin{aligned} \lim_{\lambda \rightarrow -\infty} \psi(\lambda) &= \int_0^1 T'(1-t) \min \left\{ F_X^{-1}(t), W_0 - \Pi \right\} dt \\ &= \int_0^{F_X(W_0 - \Pi)} T'(1-t) F_X^{-1}(t) dt + (W_0 - \Pi) \int_{F_X(W_0 - \Pi)}^1 T'(1-t) dt \end{aligned}$$

However, by Assumption 3, we have $F_X(W_0 - \Pi) = 1$. This then implies that:

$$\lim_{\lambda \rightarrow -\infty} \psi(\lambda) = \int_0^1 T'(1-t) F_X^{-1}(t) dt = \int X dT \circ Q$$

Now, for any $Y \in B^+(\Sigma)$, which is feasible for Problem (A8), one has:

- (i) $0 \leq Y \leq X$; and,
- (ii) $\int (X - Y) dT \circ Q = R_0$

Hence, $0 \leq X - Y \leq X$, and so, by the monotonicity of the Choquet integral (Proposition A1), it follows that $R_0 = \int (X - Y) dT \circ Q \leq \int X dT \circ Q$. Consequently,

$$0 = \psi(0) \leq R_0 \leq \int X dT \circ Q = \lim_{\lambda \rightarrow -\infty} \psi(\lambda)$$

Hence, by the intermediate value theorem ([50], Theorem 4.23), there exists some $\lambda^* \leq 0$, such that $\psi(\lambda^*) = R_0$. \square

By Theorem 2 and Lemmas A13 and A14, this concludes the proof of Corollary A1. \square

Now, since the function $f_{\lambda}^{**}(t) = \max \left[0, \min \left\{ F_X^{-1}(t), f_{\lambda^*}^*(t) \right\} \right]$ is optimal for Problem (A17), it is optimal for Problem (A15). Therefore, by Lemma A12, the function:

$$f_{\lambda}^{**}(U) = \max \left[0, \min \left\{ F_X^{-1}(U), f_{\lambda^*}^*(U) \right\} \right]$$

is optimal for Problem (A13), and the function:

$$\left(X - \max \left[0, \min \left\{ F_X^{-1}(U), f_{\lambda^*}^*(U) \right\} \right] \right) \mathbf{1}_A + X \mathbf{1}_{S \setminus A}$$

is optimal for Problem (9). However, since $Q \circ X^{-1}$ is nonatomic (by Assumption 4), it follows that $U = F_X(X)$ has a uniform distribution over $(0, 1)$ ([44], Lemma A.21), and that $X = F_X^{-1}(F_X(X))$, Q -a.s. Therefore,

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

This Q -a.s. equality implies that:

$$\max \left[0, \min \left\{ X, W_0 - \Pi - (u')^{-1} \left(\frac{-\lambda^* T'(1-U)}{\phi(X)} \right) \right\} \right]$$

is also optimal for Problem (A13). Lemma A12 then implies that the function:

$$\left(X - \max \left[0, \min \left\{ X, W_0 - \Pi - (u')^{-1} \left(\frac{-\lambda^* T'(1-U)}{\phi(X)} \right) \right\} \right] \right) \mathbf{1}_A + X \mathbf{1}_{S \setminus A}$$

is optimal for Problem (9).

Finally, it can be easily verified that:

$$\begin{aligned} X - \max \left[0, \min \left\{ X, W_0 - \Pi - (u')^{-1} \left(\frac{-\lambda^* T'(1-U)}{\phi(X)} \right) \right\} \right] \\ = \min \left[X, \max \left(0, X - \left[W_0 - \Pi - (u')^{-1} \left(\frac{-\lambda^* T'(1-U)}{\phi(X)} \right) \right] \right) \right] \end{aligned}$$

The rest follows from Corollary A1, noting that since $X = F_X^{-1}(U)$, Q -a.s., one has:

$$\begin{aligned} R_0 &= \int_0^1 T'(1-t) \max \left[0, \min \left\{ F_X^{-1}(t), f_{\lambda^*}^*(t) \right\} \right] dt \\ &= \int T'(1-U) \max \left[0, \min \left\{ F_X^{-1}(U), f_{\lambda^*}^*(U) \right\} \right] dQ \\ &= \int T'(1-U) \max \left[0, \min \left\{ X, W_0 - \Pi - (u')^{-1} \left(\frac{-\lambda^* T'(1-U)}{\phi(X)} \right) \right\} \right] dQ \end{aligned}$$

This concludes the proof of Corollary 1. □

References

1. Arrow, K. *Essays in the Theory of Risk-Bearing*; Markham Publishing Company: Chicago, IL, USA, 1971.
2. De Finetti, B. La Préviation: Ses Lois Logiques, Ses Sources Subjectives. *Ann. l'Inst. Henri Poincaré* **1937**, *7*, 1–68.
3. Savage, L. *The Foundations of Statistics*, 2nd ed.; Dover Publications: New York, NY, USA, 1972.
4. Von Neumann, J.; Morgenstern, O. *Theory of Games and Economic Behavior*; Princeton University Press: Princeton, NJ, USA, 1944.
5. KNIGHT, F. *Risk, Uncertainty, and Profit*; Houghton Mifflin: Boston, MA, USA; New York, NY, USA, 1921.
6. Ellsberg, D. Risk, Ambiguity, and the Savage Axioms. *Q. J. Econ.* **1961**, *75*, 643–669.
7. Camerer, C. Individual Decision Making. In *Handbook of Experimental Economics*; Kagel, J.H., Roth, A.E., Eds.; Princeton University Press: Princeton, NJ, USA, 1995.
8. Schmeidler, D. Subjective Probability and Expected Utility without Additivity. *Econometrica* **1989**, *57*, 571–587.
9. Gilboa, I.; Schmeidler, D. Maxmin Expected Utility with a Non-Unique Prior. *J. Math. Econ.* **1989**, *18*, 141–153.
10. Ghirardato, P.; Maccheroni, F.; Marinacci, M. Differentiating Ambiguity and Ambiguity Attitude. *J. Econ. Theory* **2004**, *118*, 133–173.
11. Amarante, M. Foundations of Neo-Bayesian Statistics. *J. Econ. Theory* **2009**, *144*, 2146–2173.
12. Gilboa, I.; Marinacci, M. Ambiguity and the Bayesian Paradigm. In *Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress of the Econometric Society*; Acemoglu, D., Arellano, M., Dekel, E., Eds.; Cambridge University Press: New York, NY, USA, 2013.
13. Hogarth, R.; Kunreuther, H. Risk, Ambiguity, and Insurance. *J. Risk Uncertain.* **1989**, *2*, 5–35.
14. Amarante, M.; Ghossoub, M.; Phelps, E. Ambiguity on the Insurer's Side: The Demand for Insurance. *J. Math. Econ.* **2015**, *58*, 61–78.
15. Alary, D.; Gollier, C.; Treich, N. The Effect of Ambiguity Aversion on Insurance and Self-Protection. *Econ. J.* **2013**, *123*, 1188–1202.
16. Klibanoff, P.; Marinacci, M.; Mukerji, S. A Smooth Model of Decision Making under Ambiguity. *Econometrica* **2005**, *73*, 1849–1892.
17. Gollier, C. Optimal Insurance Design of Ambiguous Risks. *Economic Theory* **2014**, *57*, 555–576.
18. Jeleva, M. Background Risk, Demand for Insurance, and Choquet Expected Utility Preferences. *GENEVA Pap. Risk Insur.-Theory* **2000**, *25*, 7–28.
19. Young, V. Optimal Insurance under Wang's Premium Principle. *Insur. Math. Econ.* **1999**, *25*, 109–122.
20. Bernard, C.; He, X.; Yan, J.; Zhou, X. Optimal Insurance Design under Rank-Dependent Expected Utility. *Math. Financ.* **2015**, *25*, 154–186.
21. Quiggin, J. A Theory of Anticipated Utility. *J. Econ. Behav.* **1982**, *3*, 323–343.
22. Yaari, M. The Dual Theory of Choice under Risk. *Econometrica* **1987**, *55*, 95–115.
23. Doherty, N.; Eeckhoudt, L. Optimal Insurance without Expected Utility: The Dual Theory and the Linearity of Insurance Contracts. *J. Risk Uncertain.* **1995**, *10*, 157–179.
24. Karni, E. Optimal Insurance: A Nonexpected Utility Analysis. In *Contributions to Insurance Economics*; Dionne, G., Ed.; Kluwer Academic Publishers: Boston, MA, USA, 1992.
25. Machina, M. Non-Expected Utility and the Robustness of the Classical Insurance Paradigm. *GENEVA Pap. Risk Insur.-Theory* **1995**, *20*, 9–50.
26. Schlesinger, H. Insurance Demand without the Expected-Utility Paradigm. *J. Risk Insur.* **1997**, *64*, 19–39.
27. Carlier, G.; Dana, R.; Shahidi, N. Efficient Insurance Contracts under Epsilon-Contaminated Utilities. *GENEVA Pap. Risk Insur.-Theory* **2003**, *28*, 59–71.
28. Anwar, S.; Zheng, M. Competitive Insurance Market in the Presence of Ambiguity. *Insur. Math. Econ.* **2012**, *50*, 79–84.
29. Carlier, G.; Dana, R. Insurance Contracts with Deductibles and Upper Limits. Université Paris Dauphine: Paris, France, 2002; Preprint, Ceremade.
30. Carlier, G.; Dana, R. Core of Convex Distortions of a Probability. *J. Econ. Theory* **2003**, *113*, 199–222.

31. Carlier, G.; Dana, R. Two-persons Efficient Risk-sharing and Equilibria for Concave Law-invariant Utilities. *Econ. Theory* **2008**, *36*, 189–223.
32. Chateauneuf, A.; Dana, R.; Tallon, J. Optimal Risk-sharing Rules and Equilibria with Choquet-expected-utility. *J. Math. Econ.* **2000**, *34*, 191–214.
33. Balbás, A.; Balbás, B.; Balbás, R.; Heras, A. Optimal Reinsurance under Risk and Uncertainty. *Insur. Math. Econ.* **2015**, *60*, 61–74.
34. Ert, E.; Trautmann, S. Sampling Experience Reverses Preferences for Ambiguity. *J. Risk Uncertain.* **2014**, *49*, 31–42.
35. Kocher, M.; Lahno, A.; Trautmann, S. *Ambiguity Aversion is the Exception*; CESifo Group: Munich, Germany, 2015; Preprint, CESifo Working Paper Series No. 5261.
36. Trautmann, S.; van de Kuilen, G. Ambiguity Attitudes. In *The Wiley Blackwell Handbook of Judgment and Decision Making*; Keren, G., Wu, G., Eds.; Wiley-Blackwell: Oxford, UK, 2016.
37. Aliprantis, C.; Border, K. *Infinite Dimensional Analysis*, 3rd ed.; Springer-Verlag: Heidelberg, Germany, 2006.
38. Pichler, A. Insurance Pricing under Ambiguity. *Eur. Actuar. J.* **2014**, *4*, 335–364.
39. Ghossoub, M. Vigilant Measures of Risk and the Demand for Contingent Claims. *Insur. Math. Econ.* **2015**, *61*, 27–35.
40. Denneberg, D. *Non-Additive Measure and Integral*; Kluwer Academic Publishers: Dordrecht, Netherlands, 1994.
41. Marinacci, M.; Montrucchio, L. Introduction to the Mathematics of Ambiguity. In *Uncertainty in Economic Theory: Essays in Honor of David Schmeidlers 65th Birthday*; Gilboa, I., Ed.; Routledge: London, UK, 2004; pp. 46–107.
42. Schmeidler, D. Integral Representation without Additivity. *Proc. Am. Math. Soc.* **1986**, *97*, 255–261.
43. Cohn, D. *Measure Theory*; Birkhauser: Boston, MA, USA, 1980.
44. Föllmer, H.; Schied, A. *Stochastic Finance: An Introduction in Discrete Time*, 3rd ed.; Walter de Gruyter: Berlin, Germany, 2011.
45. Jin, H.; Zhou, X. Behavioral Portfolio Selection in Continuous Time. *Math. Financ.* **2008**, *18*, 385–426.
46. He, X.; Zhou, X. Portfolio Choice via Quantiles. *Math. Financ.* **2011**, *21*, 203–231.
47. Jin, H.; Zhou, X. Greed, Leverage, and Potential Losses: A Prospect Theory Perspective. *Math. Financ.* **2013**, *23*, 122–142.
48. Carlier, G.; Dana, R. Optimal Demand for Contingent Claims when Agents Have Law Invariant Utilities. *Math. Financ.* **2011**, *21*, 169–201.
49. Dudley, R. *Real Analysis and Probability*; Cambridge University Press: New York, NY, USA, 2002.
50. Rudin, W. *Principles of Mathematical Analysis*, 3rd ed.; McGraw-Hill Book Company: New York, NY, USA, 1976.
51. Ghossoub, M. Equimeasurable Rearrangements with Capacities. *Math. Oper. Res.* **2015**, *40*, 429–445.
52. Ghossoub, M. Contracting under Heterogeneous Beliefs. Ph.D. Thesis, Department of Statistics & Actuarial Science, University of Waterloo, Waterloo, ON, Canada, May 2011.

