OPTIMAL INSURANCE UNDER MAXMIN EXPECTED UTILITY

CORINA BIRGHILA
UNIVERSITY OF WATERLOO

TIM J. BOONEN
UNIVERSITY OF AMSTERDAM

MARIO GHOSSOUB
UNIVERSITY OF WATERLOO

THIS DRAFT: JUNE 16, 2023

Abstract. We examine a problem of demand for insurance indemnification, when the insured is sensitive to ambiguity and behaves according to the Maxmin-Expected Utility model of Gilboa and Schmeidler (1989), whereas the insurer is a (risk-averse or risk-neutral) Expected-Utility maximizer. We characterize optimal indemnity functions both with and without the customary *ex ante no-sabotage* requirement on feasible indemnities, and for both concave and linear utility functions for the two agents. This allows us to provide a unifying framework in which we examine the effects of the *no-sabotage* condition, marginal utility of wealth, belief heterogeneity, as well as ambiguity (multiplicity of priors) on the structure of optimal indemnity functions. In particular, we show how the singularity in beliefs leads to an optimal indemnity function that involves full insurance on an event to which the insurer assigns zero probability, while the decision maker assigns a positive probability. We examine several illustrative examples, and we provide numerical studies for the case of a Wasserstein and a Rényi ambiguity set.

Key words and phrases. Optimal Insurance, Ambiguity, Multiple Priors, Maxmin-Expected Utility, Heterogeneous Beliefs.

JEL Classification: C02, C61, D86, G22.

2010 Mathematics Subject Classification: 91B30, 91G99.

We thank the Editor, Associate Editor, and an anonymous reviewer for comments and suggestions that have improved this paper. We are grateful to Frank Riedel, Roger J.A. Laeven, and Michel Vellekoop, as well audiences at the 2020 Oberwolfach workshop "*New Challenges in the Interplay between Finance and Insurance*", the University of Amsterdam, and the 2021 SIAM Conference on Financial Mathematics and Engineering. Mario Ghossoub acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC Grant No. 2018-03961).

1. Introduction

A foundational problem in the theory of risk exchange is the problem of demand for insurance indemnification. Specifically, an insurance buyer, or decision maker (DM), faces a random insurable loss, against which she seeks coverage through the purchase of an insurance policy. The insurance pricing functional is assumed to be known by the DM, and to be given by the certainty equivalent of the insurer's utility. Although this is a classical problem, it has traditionally been confined to the accustomed framework of Expected-Utility Theory (EUT), going back to the pioneering work of Arrow (1963, 1971) and Mossin (1968). Arrow (1963) shows the optimality of deductible insurance (a zero indemnification below a fixed threshold of loss, and a linear indemnification above) in an EUT framework, when the DM is risk-averse, the insurer is risk-neutral, and the two parties have the same beliefs about the underlying loss probability distribution. We refer to Gollier (2013) and Schlesinger (2000) for surveys of the rather large literature on optimal insurance with EU preferences.

Ambiguity in Insurance Demand. The vast majority of this literature remains within the confines of the classical EUT. Yet, ever since the major challenges to the foundations of EUT that the work of Allais (1953) and Ellsberg (1961) has put forward, decision theory has been pulling away from parts of the axiomatic foundations of EUT, in favour of non-EU models that can rationalize behavior depicted by Allais (1953) and Ellsberg (1961), as well as other cognitive biases that are not captured by EUT. Arguably, one of the most important achievements of the modern theory of choice under uncertainty is the remarkable development spurred by the work of Ellsberg (1961), in the study of what came to be known as *ambiguity*, or *model uncertainty*. Two main approaches to the rationalization of attitudes toward ambiguity have been explored in the literature on axiomatic decision theory: the *non-additive prior* approach, and the *multiple additive priors* approach. These two approaches do intersect, but they are not equivalent. The first category is based on the seminal contributions of Yaari (1987) (Dual Theory, or DT), Quiggin (1982) (Rank-Dependent Expected-Utility, or RDEU), and Schmeidler (1989) (Choquet-Expected Utility, or CEU), which encompasses the previous two models. The second category was initiated by Gilboa and Schmeidler (1989) (Maxmin-Expected Utility, or MEU) and further refined by Ghirardato et al. (2004) (the α-maxmin model), Klibanoff et al. (2005) (the KMM model), and Amarante (2009) who provides a unifying framework. Additionally, it is important to note that while MEU and RDEU have a nonempty intersection (e.g., when the distortion function in RDEU is convex), they are vastly different models. In fact, few MEU models can be mapped to a RDEU model (see, e.g., Amarante (2014) for a detailed explanation of this point).

While the literature on non-EU preferences in risk-sharing or optimal insurance design problems is considerably thinner than the literature on risk-sharing with EU preferences, *behavioral* preferences, and ambiguity in particular play an increasing role in this literature. Yet, Machina (2013) points out that the robustness of standard optimal insurance results under situations of ambiguity is still very much an open question, despite a growing literature on the topic. For instance, Bernard et al. (2015) and Xu et al. (2018) study RDEU preferences of the DM and risk-neutral EU preferences of the insurer, and they derive optimal insurance indemnities. Ghossoub (2019b) extends the analysis to account for more general premium principles. Recently, Xu (2021) reconsiders the problem of optimal insurance under RDEU preferences for the DM and risk-neutral EU preferences of the insurer and provides a novel characterization of optimal indemnities using an ODE approach. Also within the first category of ambiguity representation as a non-additive prior, Jeleva (2000) considers the case of a DM who is a CEU-maximizer.

In the second category of ambiguity representation as a collection of additive priors and an aggregation rule, Alary et al. (2013) and Gollier (2014) consider the case of an ambiguity-averse DM, in the sense of KMM. However, they consider a finite state space and restrict the set of priors to have a given parametric form. More recently, Jiang et al. (2020) study a variant of KMM preferences applied to distortion risk measures with a finite set of priors. Under such preferences for the DM, and using an expected-value premium principle, the authors derive an implicit characterization of optimal indemnity functions.

Despite its appeal, for its ability to provide a separation of the effect of ambiguity aversion from that of risk aversion, as well as for its capability to define the notion of ambiguity neutrality, the KMM model, as a model of ambiguity with multiple priors, is arguably not as intuitive or popular as the MEU model of Gilboa and Schmeidler (1989). The MEU model gives rise to decision-making problems that can be embedded into to a larger class of *model uncertainty* problems, which lie at the core of the theory of *distributionally robust optimization* (DRO). In this framework, a decision-making problem is often modelled via a maxmin formulation: the agent is uncertain about the underlying model (prior), and therefore formulates an objective function using a collection of (additive) priors, also referred to as the *ambiguity set*. The agent then aims to maximize the objective under the worst-case model (e.g., Ben-Tal et al. (2009)). However, the intuitiveness and wide popularity of the MEU model notwithstanding, there has surprisingly been no study of optimal insurance contracting when the DM is an MEU-maximizer, to the best of our knowledge. This paper fills this void. Specifically, we extend the classical setup and results in two ways: (i) the DM is endowed with MEU preferences with a set $\mathcal{C}$ of priors; and (ii) the insurer is not necessarily risk-neutral (that is, the premium principle is not necessarily an expected-value premium principle). The main objective of this paper is to determine the shape of the optimal insurance indemnity when the DM is sensitive to ambiguity and behaves according to MEU.

This Paper's Contribution. In the literature on optimal insurance contracting, a popular assumption is the *no-sabotage condition*, typically imposed as an *ex ante* condition of feasibility of insurance indemnities. This condition stipulates that the insured (ceded) risk and the retained risk are comonotonic (they are both nondecreasing functions of the underlying loss). Under the no-sabotage condition, the DM has no incentive to under-report the underlying loss, nor does the DM have an incentive to create incremental losses. This condition is also sometimes referred to as *incentive compatibility*, or a condition that avoids *ex post* moral hazard; and it is further studied by Huberman et al. (1983) and Carlier and Dana (2003)[1]. In this paper, we characterize optimal insurance contracts under MEU, both with and without the no-sabotage condition. In doing so, this paper sheds light on the consequences of the no-sabotage assumption on the construction of optimal insurance indemnities, in the presence of belief heterogeneity as well as multiple priors for the DM. Furthermore, while the literature on optimal insurance with non-standard preferences assumes risk-neutrality of the insurer, we provide a more general treatment and allow for risk-aversion of the insurer not only of the insured. We do this both with and without the no-sabotage condition.

Our main results are the following. First, we examine in Section 3 the general case in which the insurer is a risk-averse EU-maximizer, and the DM is an MEU-maximizer with a concave utility function, displaying decreasing marginal utility of wealth. Following the vast majority of the literature, we provide an implicit characterization of optimal indemnity functions, both with and without the no-sabotage condition on feasible indemnities[2]. Optimal indemnity functions can be formulated as a solution to an ordinary differential equation, which can then be easily solved numerically in practice. The implicit characterization of optimal indemnities is then used to provide closed-form solutions when the relation between the DM's and insurer's beliefs is specified. For instance, when the ambiguity set of DM consists of all models that are absolutely continuous with respect to the insurer's belief such that the corresponding Radon-Nikodým derivative is an increasing function as in Furman and Zitikis (2008a, 2008b, 2009), we characterize the shape of optimal contracts in an explicit way (see Example 4.4). The main technique proposed to solve the problem consists of two main steps. First, the constrained optimization problem is reformulated as a minimax problem, for which Sion's Minimax Theorem can be applied to obtain the existence of the worst-case measure P^* in the ambiguity set $\mathcal{C}$. The Minimax Theorem is an important result and a standard tool that is frequently employed in the robust optimization and distributionally robust optimization literature (see, for example, Žáčková (1966), or Ben-Tal and Nemirovski (1998), or Shapiro and Kleywegt (2002)). In the framework of optimal insurance design, the Minimax Theorem is used in Cheung et al. (2019) to obtain the structure of the optimal insurance contract that minimizes a coherent risk measure and under the premium budget constraint. Here, the minimax reformulation of the original problem results from the Kusuoka representation of the coherent risk measures. In the presence of model uncertainty, Jiang et al. (2020) adopt the same result to derive an analytical form of the optimal insurance contract under distortion risk measures. In the context of MEU, the Minimax Theorem is further used in Xu et al. (2018) to prove the existence of Lagrange multiplier that binds the budget constraint. Similarly to Cheung et al. (2019), the structure of the worst-case $P^* \in \mathcal{C}$ is derived numerically, by specifying the structure of the ambiguity set. In the second step, once the worst-case measure P^* is obtained, the problem falls into the literature on belief heterogeneity in insurance contracting, to which we contribute significantly, as discussed below. The saddle point approach is also considered in the work of Birghila and Pflug (2019), where the insured's ambiguity set is a convex hull of a finite number of models that are within some ϵ-distance of a reference/baseline model. The distance between distributions is measured by the Wasserstein distance on the positive real line, with a distorted underlying metric. The saddle point, i.e., the optimal insurance contract and the worst-case distribution, is obtained using a numerical approach.

Additionally, by specifying the structure of the DM's ambiguity set $\mathcal{C}$, we are able to obtain explicitly the worst-case probability measure for the problem analyzed in Section 3. In particular, we examine the special case in which the DM's set of priors forms a neighborhood around the insurer's probability measure. In a general setting in which both participants are risk-averse, we define $\mathcal{C}$ to be a Rényi ambiguity set. In a discretized framework, we use a successive convex programming algorithm to solve the ordinary differential equation obtained in Section 3. We then assess the influence of the ambiguity set on the optimal value. In particular, we show numerically that a larger ambiguity set yields a lower certainty equivalent of final wealth for the DM, but increases the willingness-to-pay for insurance. Moreover, the impact of the no-sabotage condition on the feasible set of insurance indemnities is illustrated.

Second, as a special case of the above setting, we examine in Section 4 the situation in which the insurer is risk-neutral, and hence the premium principle is an expected-value premium principle, as is commonly assumed in the literature (e.g., Bernard et al. (2015), Xu et al. (2018), and Xu (2021)). In this case, we provide an explicit, closed-form characterization of optimal indemnity functions in the absence of the no-sabotage condition, and an implicit characterization in the presence of the no-sabotage condition. In particular, by doing so, we provide in both cases (with and without the no-sabotage condition) a crisp depiction of the effect of heterogeneity in beliefs between the two parties, showing how the singularity in beliefs leads to an optimal indemnity function that involves full insurance on an event to which the insurer assigns zero probability, but not the DM. This an important and intuitive feature of our optimal contracts in this case. Similarly to Section 3, we conclude this section with a numerical example. When the DM is risk-averse, $\mathcal{C}$ is a Wasserstein ambiguity set, and the insurer is risk-neutral, we are able to characterize the saddle point of the problem in Section 4. In this case, the optimal indemnity is a deductible contract, and the worst-case measure P^* dominates the insurer's probability measure in the sense of first-order stochastic dominance.

As an application of the results obtained in Section 4, one can also examine the situation in which both parties display constant marginal utility of wealth, that is, their utility functions are linear. In that case, it is straightforward to show that if the no-sabotage condition is imposed, then layer insurance is optimal. In the absence of the no-sabotage condition, the optimal indemnity makes use of a partition of the state space in three sets as in Theorem (3.4), providing no insurance for events in the first set, full insurance for events in the second set, and proportional insurance for events in the third set. Moreover, Artzner et al. (1999) and Delbaen (2002) show that the class of MEU preferences with linear utility is related to the class of coherent risk measures. Therefore, our analysis can be used to derive optimal insurance contracts when the DM is endowed with a general coherent risk measure.

Other Related Literature. Broadly speaking, this paper contributes to the literature focusing on incorporating behavioral models of decision-making into the literature on optimal insurance design. While our main focus is on ambiguity-sensitive preferences, it is important to note that other behavioral models of decision-making have been gaining popularity in the theory of insurance demand. For instance, Cheung et al. (2015) study disappointment theory (e.g., Bell (1985), Loomes and Sugden (1986), and Gul (1991)), while Chi and Zhuang (2020) study the effects of regret theory (e.g., Bell (1982) and Loomes and Sugden (1982)). Both settings accommodate for a deductible and partial insurance of losses above the deductible as an optimal indemnity function.

By explicitly incorporating model uncertainty into the problem formulation via a set of priors $\mathcal{C}$, the present paper also falls within the DRO framework. In this perspective, insurance contracts can be seen as saddle points of a DRO problem. The benefit of this technique is twofold. First, the worst-case approach ensures that the optimal decision is not sensitive to possible model misspecification. Second, in many situations, there exist tractable reformulations or algorithms to solve these distributionally robust models, even when the corresponding non-ambiguous problem (that is, when there is a unique prior) cannot be efficiently solved. The idea of incorporating multiple models in the decision-making process dates back to the fundamental work of Scarf (1958) in the inventory management applications. He considers a robust formulation of the newsvendor problem, where the optimal strategy is constructed over all possible demand functions with known mean and variance. This initial idea is further developed in the work of Ben-Tal et al. (2009) and Bertsimas and Sim (2004), among others. A key concept in DRO is the structure of the set of priors, known here as the ambiguity set. Clearly, the choice of the ambiguity set $\mathcal{C}$ influences the worst-case model, and thus the optimal decision, while it also facilitates a tractable reformulation and efficient algorithm implementation. The existing literature has focused so far on two types of ambiguity sets: those built using the moment-based approach (e.g., Delage and Ye (2010), Scarf (1958), and Zymler et al. (2013)), and those built using the statistical distance-based approach (such as the Kullback-Leibler divergence in Calafiore and El Ghaoui (2006), the L_1-ball in Thiele (2008), or the Wasserstein distance in Esfahani and Kuhn (2018)). Each such choice comes with useful structural properties, but also with shortcomings that need to be dealt with. Ultimately, it is the available set of observations and the type of application that would dictate a suitable choice of ambiguity set $\mathcal{C}$.

Furthermore, this paper also contributes to the literature on heterogeneity in beliefs between the DM and the insurer. When the DM does not perceive any ambiguity in the assessment of uncertainty, and thereby behaves as an EU-maximizer, the set $\mathcal{C}$ of priors is a singleton, that is, $\mathcal{C} = \{P\}$, for some probability measure P distinct from the insurer's probability measure Q, and potentially exhibiting some singularity with Q. Heterogeneity in beliefs has been studied recently in the context of optimal (re)insurance by Ghossoub (2019a), Boonen and Ghossoub (2019, 2021), Chi (2019), and Yu and Fang (2020). All of these studies focus on unambiguous subjective preferences on the side of the DM (that is, a unique subjective prior on the state space), but they differ in the formulation of the objective function that is optimized. As a special case in which the set of priors is a singleton, our results provide a unifying treatment of optimal insurance with belief heterogeneity, and extend the existing results in this literature in several ways, as we make no assumption on how the beliefs diverge, and we allow for risk-aversion of the insurer, which is not done in the literature. First, while the majority of the existing literature imposes some assumptions on the way the beliefs of the policyholder and the insurer (the measures P and Q, respectively) can diverge (e.g., a monotone likelihood ratio, a monotone hazard ratio, or some other assumption), we do not make any such assumption. We allow the measures to truly diverge in any way, and to exhibit singularity. Second, while the existing literature only considers the case of a risk-neutral insurer (a linear utility function v), we do consider the effect of risk-aversion of the insurer on the optimal indemnity. We also consider the case of risk-neutrality. Third, when it comes to the choice of the set of *ex ante* admissible indemnity functions, the existing literature either consider the case of no-sabotage indemnity functions (the set $\hat{\mathcal{I}}$ in eq. (2.3)) or the set of general nonnegative indemnity functions that do not exceed the loss (the set $\mathcal{I}$ in eq. (2.2)), but not both. We do consider both cases, thereby illustrating the impact of the no-sabotage assumption on the structure of optima.

The rest of the paper is organized as follows. Section 2 presents the setup of our problem together with the necessary background. In Section 3, we consider the case in which both the insurer and DM have concave utility functions, and the DM is an MEU-maximizer. We characterize optimal indemnity functions both in the presence and absence of the no-sabotage condition. The corresponding worst-case measures are obtained numerically, by imposing a specific structure of the ambiguity set $\mathcal{C}$. Section 4 considers the particular case of a risk-neutral insurer, and provides some illustrating examples. Section 5 concludes the paper. Some definitions and technical proofs are provided in Appendices A to C.

2. Setup and preliminaries

Let S be a nonempty collection of states of the world, and equip S with a σ-algebra $\mathcal{G}$ of events. A DM is facing an insurable state-contingent loss represented by a random variable X on the measurable space $(S, \mathcal{G})$, with values in the interval $[0, M]$, for some $M \in \mathbb{R}^+$. We denote by Σ the sub-σ-algebra $\sigma\{X\}$ of $\mathcal{G}$ on S generated by the random variable X.

Let $B(\Sigma)$ denote the vector space of all bounded, $\mathbb{R}$-valued, and Σ-measurable functions on (S, Σ), and let $B^+(\Sigma)$ be its positive cone. When endowed with the supnorm $\|.\|_{sup}$, $B(\Sigma)$ is a Banach space (e.g., Dunford and Schwartz (1958, IV.5.1)). By Doob's Measurability Theorem (e.g., Aliprantis and Border (2006, Theorem 4.41)), for any $Y \in B(\Sigma)$ there exists a bounded, Borel-measurable map $I : \mathbb{R} \to \mathbb{R}$ such that $Y = I \circ X$. Moreover, $Y \in B^+(\Sigma)$ if and only if the function I is nonnegative.

Definition 2.1. Two functions $Y_1, Y_2 \in B(\Sigma)$ are said to be comonotonic (resp., anti-comonotonic) if

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

For instance any $Y \in B(\Sigma)$ is comonotonic and anti-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 (resp., anti-comonotonic) with Y_1 if and only if the function I is nondecreasing (resp., nonincreasing).

Let $ba(\Sigma)$ denote the linear space of all bounded finitely additive set functions on Σ, endowed with the usual mixing operations. When endowed with the total variation norm $\|.\|_v$, $ba(\Sigma)$ is a Banach space. By a classical result (e.g., Dunford and Schwartz (1958, IV.5.1)), $(ba(\Sigma), \|.\|_v)$ is isometrically isomorphic to the norm-dual of $B(\Sigma)$, via the duality $<\phi, \lambda> = \int \phi \, d\lambda, \ \forall \lambda \in ba(\Sigma), \ \forall \phi \in B(\Sigma)$. Consequently, we can endow $ba(\Sigma)$ with the weak* topology $\sigma(ba(\Sigma), B(\Sigma))$.

Let $ca(\Sigma)$ denote the collection of all countably additive elements of $ba(\Sigma)$, and let $ca^+(\Sigma)$ denote its positive cone. Then $ca(\Sigma)$ is a $\|.\|_v$-closed linear subspace of $ba(\Sigma)$. Hence, $ca(\Sigma)$ is $\|.\|_v$-complete, i.e. $(ca(\Sigma), \|.\|_v)$ is a Banach space. Denote by

$$ca_1^+(\Sigma) := \Big\{\mu \in ca^+(\Sigma) : \mu(S) = 1\Big\}$$

the collection of probability measures on (S, Σ). We shall endow $ca_1^+(\Sigma)$ with the weak* topology inherited from $ba(\Sigma)$.

For any $Y \in B(\Sigma)$ and $P \in ca_1^+(\Sigma)$, let $F_{Y,P}(t) := P\big(\{s \in S : Y(s) \leqslant t\}\big)$ denote the cumulative distribution function (CDF) of Y with respect to the probability measure P, and let $F_{Y,P}^{-1}(t)$ denote the left-continuous inverse of the $F_{Y,P}$ (i.e., the quantile function of Y), defined by

$$F_{Y,P}^{-1}(t) := \inf\big\{z \in \mathbb{R} : F_{Y,P}(z) \geqslant t\big\}, \ \forall t \in [0,1].$$

2.1. The DM's and the Insurer's Preferences.

The DM can purchase insurance against the random loss X in a perfectly competitive insurance market, for a premium set by the insurer. In return for the premium payment, the DM is promised an indemnification against the realizations of X. An indemnity function is a random variable $Y = I(X)$ on (S, Σ), for some bounded, Borel-measurable map $I : X(S) \to \mathbb{R}$, which pays off the amount $I(X(s)) \in \mathbb{R}$ in state of world $s \in S$, corresponding to a realization $X(s)$ of X. That is, we can identify the set of indemnity functions with a subset of $B(\Sigma)$. For each indemnity function

$Y \in B(\Sigma)$, we define the corresponding retention function by $R := X - Y \in B(\Sigma)$. As the name suggests, R is the retained loss after insurance indemnification.

The DM has a preference relation over insurance indemnification functions (or over wealth profiles) that admits an MEU representation $V^{\text{MEU}} : B(\Sigma) \to \mathbb{R}$ as in Gilboa and Schmeidler (1989), of the form

$$V^{\text{MEU}}(Z) := \min_{\mu \in \mathcal{C}} \int u(Z)\, d\mu, \quad \forall\, Z \in B(\Sigma), \tag{2.1}$$

where $u : \mathbb{R} \to \mathbb{R}$ is a concave utility function, and $\mathcal{C}$ is a (unique) weak*-compact and convex subset of $ba_1^+(\Sigma)$. Moreover, we assume that the DM's preferences satisfy the Arrow-Villegas Monotone Continuity axiom as in Chateauneuf et al. (2005), so that $\mathcal{C} \subset ca_1^+(\Sigma)$, i.e., all priors are countably additive. Additionally, the DM's utility function u satisfies the following assumption.

Assumption 1. *The utility function $u : \mathbb{R} \longrightarrow \mathbb{R}$ is strictly increasing, concave and continuously differentiable.*

Let $W_0 \in \mathbb{R}^+$ be the DM's initial wealth. After purchasing insurance coverage for a premium $\Pi_0 > 0$, the DM's terminal wealth is a random variable $W \in B(\Sigma)$ given by

$$W := W_0 - X + Y - \Pi_0.$$

The insurer's preference over $B(\Sigma)$ admits an EU representation $V^{\text{Ins}} : B(\Sigma) \to \mathbb{R}$ of the form

$$V^{\text{Ins}}(Z) := \int v(Z)\, dQ, \quad \forall\, Z \in B(\Sigma),$$

for a utility function $v : \mathbb{R} \to \mathbb{R}$ satisfying Assumption 1 and a probability measure $Q \in ca_1^+(\Sigma)$.

The insurer has an initial wealth W_0^{Ins}, and faces an administration cost, often called an *indemnification cost*, associated with the handling of an indemnity payment. As customary in the literature (e.g., Bernard et al. (2015) and Xu et al. (2018)), we assume that for a given indemnity function $Y = I \circ X$, this indemnification cost is a proportional cost of the form ρY, for a given safety loading factor $\rho \geqslant 0$ specified exogenously and *a priori*. Hence, the insurer's terminal wealth is the random variable $W^{\text{Ins}} \in B(\Sigma)$ given by

$$W^{\text{Ins}} := W_0^{\text{Ins}} - (1+\rho)Y + \Pi_0.$$

2.2. **Admissible Indemnity Functions.** In Arrow (1963)'s original formulation of the optimal insurance problem under EUT, an *ex ante* condition of feasibility of indemnity schedules is the requirement that these be nonnegative and no larger than the realization of the loss in each state of the world. This is often referred to as the *indemnity principle*, and it translates into the requirement that an admissible set of indemnities be restricted to those $Y \in B(\Sigma)$ that satisfy $0 \leqslant Y \leqslant X$. We shall denote this set of indemnity functions by $\mathcal{I}$:

$$\mathcal{I} := \Big\{ Y = I \circ X \in B^+(\Sigma) : 0 \leqslant I(x) \leqslant x,\, \forall x \in [0, M] \Big\}. \tag{2.2}$$

A desirable property of optimal indemnities is that an indemnity function $Y = I \circ X$ and the corresponding retention function $R = X - Y$ be both nondecreasing functions of the loss X, that is both comonotonic with X (and hence Y and R are comonotonic). Indeed, if Y fails to be comonotonic with X, then the DM has an incentive to under-report the loss; whereas if R fails to be comonotonic with X, then the DM has an incentive to create additional damage. These situations of *ex post* moral hazard are not desirable, and one often seeks additional *ex ante* conditions that would rule out such behavior from the DM. In the setting of Arrow (1963), the optimal indemnity is a deductible contract of the form $Y = \max(X - d, 0)$, for some $d \in \mathbb{R}^+$. For such contracts, both the indemnity and retention functions are comonotonic with the loss, and optimal indemnities are *de facto* immune to the kind of *ex post* moral hazard described above. However, outside of EUT, optimality of deductible contracts does not always hold, and optimal indemnities might suffer from the aforementioned type of moral hazard, as in Bernard et al. (2015).

In order to rule out *ex post* moral hazard that might arise from a misreporting of the loss by the DM, an additional condition is often imposed *ex ante* on the set of feasible indemnity schedules (as in Xu et al. (2018)). Such a condition is called the *no-sabotage* condition, and it stipulates that admissible indemnity functions and the corresponding retention functions be comonotonic, hence resulting in the feasibility set $\hat{\mathcal{I}}$ given by:

$$\hat{\mathcal{I}} := \left\{ \hat{Y} \in \mathcal{I} : \hat{Y} \text{ and } \hat{R} = X - \hat{Y} \text{ are comonotonic} \right\}.$$

Since $Y \in \mathcal{I}$ is of the form $Y = I \circ X$, with $0 \leqslant I(x) \leqslant x$ for all $x \in [0, M]$, we can write $\hat{\mathcal{I}}$ as

$$\hat{\mathcal{I}} = \left\{ \hat{Y} = \hat{I} \circ X \in B^+(\Sigma) : \hat{I}(0) = 0,\, 0 \leqslant \hat{I}(x_1) - \hat{I}(x_2) \leqslant x_1 - x_2, \forall\, 0 \leqslant x_2 \leqslant x_1 \leqslant M \right\}. \tag{2.3}$$

The no-sabotage condition is also sometimes referred to as *incentive compatibility* by Xu et al. (2018), and it is further studied by Huberman et al. (1983) and Carlier and Dana (2003). The latter discuss various classes of *ex ante* admissible contracts, as well as their implications of optimal indemnities.

Remark 2.2. Let $C[0, M]$ denote the set of all continuous functions on $[0, M]$ (and hence bounded), equipped with the supnorm $\| \cdot \|_{sup}$. Note that $\hat{\mathcal{I}}$ is a uniformly bounded subset of $C[0, M]$ consisting of Lipschitz-continuous functions $[0, M] \to [0, M]$, with common Lipschitz constant $K = 1$. Therefore, $\hat{\mathcal{I}}$ is equicontinuous, and hence compact by the Arzelà-Ascoli Theorem (e.g., Dunford and Schwartz (1958, Theorem IV.6.7)).

In this paper, we will characterize optimal indemnity functions, both with and without the no-sabotage condition, in order to examine the impact of such an *ex ante* requirement on feasible indemnity schedules. This will first be done in the general setting of an MEU-maximizing DM with a concave utility and an EU-maximizing insurer with concave utility (Section 3), and then in a setting where the insurer is risk-neutral (hence uses an expected-value premium principle).

3. Optimal Indemnity Functions

In this section, we investigate the DM's problem of demand for insurance indemnification, when the DM is ambiguity-sensitive and has preferences admitting an MEU representation of the form given in eq. (2.1), whereas the insurer is a risk-averse EU-maximizer with a concave utility function v. We first examine in Section 3.1 the class $\mathcal{I}$ of indemnities that are nonnegative and cannot exceed the loss X (as defined in eq. (2.2)), and we provide in Theorem 3.2 a closed-form characterization of the optimal indemnity in this case. We then consider in Section 3.2 the class $\hat{\mathcal{I}}$ of indemnities that are such that both indemnity and retention functions are nondecreasing functions of the loss (as defined in eq. (2.3)). In that case, Theorem 3.4 provides an implicit characterization of the optimal indemnity function.

The DM chooses a premium π and an indemnity function Y to maximize her MEU preferences. Such a problem can be solved in two steps. In the first step, an optimal indemnification Y^* is determined, for a fixed premium π. In the second step, the optimal premium π^* is determined. The second step is a one-dimensional optimization problem, and this paper focuses on the first step. That is, we determine the optimal indemnity for a fixed premium $\pi = \Pi_0$, as is done in Bernard et al. (2015) and Xu et al. (2018).

Let $\mathcal{F}$ denote the set of admissible indemnity functions, which could be either the set $\mathcal{I}$ defined in eq. (2.2), or the set $\hat{\mathcal{I}} \subset \mathcal{I}$ defined in eq. (2.3). For a given insurance premium $\Pi_0 > 0$ and a compact and convex set $\mathcal{C}$ of probability measures, the optimal indemnity function is obtained as the solution of the problem

$$(P) \qquad \begin{cases} \sup\limits_{Y \in \mathcal{F}} \inf\limits_{P \in \mathcal{C}} & \mathbb{E}_P[u(W_0 - X + Y - \Pi_0)] \\ \text{s.t.} & \mathbb{E}_Q[v(W_0^{\text{Ins}} - (1+\rho)Y + \Pi_0)] \geqslant v(W_0^{\text{Ins}}), \end{cases}$$

where $\rho \geqslant 0$ is a given safety loading factor. The constraint in (P) is interpreted as the insurer's participation constraint.[3] Observe that $\mathbb{E}_P[u(W_0 - X + Y - \Pi_0)] \leqslant u(W_0 - \Pi_0)$, for all $P \in \mathcal{C}$ and all $I \in \mathcal{F}$, and thus Problem (P) is finite. If $\mathbb{E}_Q[v(W_0^{\text{Ins}} - (1+\rho)X + \Pi_0)] \geqslant v(W_0^{\text{Ins}})$, then we can eliminate the constraint in Problem (P), and the optimal indemnity is $Y^* = X$, Q-a.s. In the following, we assume $\mathbb{E}_Q[v(W_0^{\text{Ins}} - (1+\rho)X + \Pi_0)] < v(W_0^{\text{Ins}})$.

Remark 3.1. By the Lebesgue Decomposition Theorem, for any $P \in \mathcal{C}$ there are finite nonnegative countably additive measures P_{ac} and P_s on (S, Σ) such that $P = P_{ac} + P_s$, where $P_{ac} \ll Q$ and $P_s \perp Q$. Hence, for each $P \in \mathcal{C}$, there exists some $A_P \in \Sigma$ and $h_P : S \to [0, \infty)$ such that $Q(S \backslash A_P) = P_s(A_P) = 0$ and $h_P = dP_{ac}/dQ$. In particular, since h_P is Σ-measurable, there exists a nonnegative Borel measurable function $\xi_P : \mathbb{R}_+ \to \mathbb{R}_+$ such that $h_P = \xi_P \circ X$.

3.1. **Without the No-Sabotage Condition.** The solution (Y^*, P^*) to Problem (P), when $\mathcal{F} = \mathcal{I}$ is given in the following result.

Theorem 3.2. *Suppose that the utility functions u and v satisfy Assumption 1 and are, in addition, strictly concave with $\lim\limits_{x \to -\infty} u'(x) = \lim\limits_{x \to -\infty} v'(x) = +\infty$ and $\lim\limits_{x \to +\infty} u'(x) = \lim\limits_{x \to +\infty} v'(x) = 0$.*[4] *Let $\mathcal{F} = \mathcal{I}$ as defined in eq. (2.2) be the set of admissible indemnity functions. Then there exists $P^* \in \mathcal{C}$ such that $Y^* \in \mathcal{I}$ is optimal for Problem (P) if and only if is of the form:*

$$Y^* = \widetilde{Y}^* 1_{A \backslash A_{h^*}} + Y_{h^*} 1_{A_{h^*}} + X 1_{S \backslash A}, \tag{3.1}$$

where

(a) $A \in \Sigma$ is such that $P^* = P^*_{ac} + P^*_s$, with $P^*_s(A) = Q(S \backslash A) = 0$;

(b) $h^* : S \to [0, \infty)$ is such that $h^* = dP^*_{ac}/dQ$;

(c) $A_{h^*} := \{s \in A : h^*(s) = 0\}$;

(d) $\widetilde{Y}^*$ and Y_{h^*} are of the form:

Case 1. *If $\lambda^* > 0$, then $Y_{h^*} = 0$ and $\widetilde{Y}^* = \max[0, \min(X, Y_0^*)]$, where Y_0^* solves*

$$u'(W_0 - X(s) + Y(s) - \Pi_0) h^*(s) - \lambda^*(1+\rho) v'(W_0^{Ins} - (1+\rho)Y(s) + \Pi_0) = 0, \ \forall s \in A \backslash A_{h^*};$$

Case 2. *If $\lambda^* = 0$, then $\widetilde{Y}^* = X$ and Y_{h^*} solves*

$$\mathbb{E}_Q[v(W_0^{Ins} - (1+\rho)Y 1_{A_{h^*}} + \Pi_0)] = v(W_0^{Ins}) - \mathbb{E}_Q[v(W_0^{Ins} - (1+\rho)X 1_{A \backslash A_{h^*}} + \Pi_0)];$$

(e) $\lambda^* \in \mathbb{R}_+$ defined (d) is such that $\lambda^*\big(\mathbb{E}_Q[v(W_0^{Ins} - (1+\rho)Y^* + \Pi_0)] - v(W_0^{Ins})\big) = 0$.

Proof. Define the set $\mathcal{I}_0 := \big\{Y \in \mathcal{I} : \mathbb{E}_Q[v(W_0^{\text{Ins}} - (1+\rho)Y + \Pi_0)] \geqslant v(W_0^{\text{Ins}})\big\}$. Observe that for $Y_1, Y_2 \in \mathcal{I}_0$ and $\alpha \in (0,1)$, we have $\widetilde{Y} := \alpha Y_1 + (1-\alpha) Y_2 \in \mathcal{I}$ and

$$\mathbb{E}_Q[v(W_0^{\text{Ins}} - (1+\rho)\widetilde{Y} + \Pi_0)] \geqslant \alpha \mathbb{E}_Q[v(W_0^{\text{Ins}} - (1+\rho)Y_1 + \Pi_0)] + (1-\alpha)\mathbb{E}_Q[v(W_0^{\text{Ins}} - (1+\rho)Y_2 + \Pi_0)] \geqslant v(W_0^{\text{Ins}}),$$

and thus the set $\mathcal{I}_0$ is convex. The objective function $\mathbb{E}_P[u(W_0 - X + Y - \Pi_0)]$ of (P) is concave in $Y \in \mathcal{I}_0$ and continuous in Y with respect to supnorm $\| \cdot \|_{sup}$. Moreover, $\mathbb{E}_P[u(W_0 - X + Y - \Pi_0)]$ is linear in P

and continuous in $P \in \mathcal{C}$ in the weak* topology. The set $\mathcal{I}_0$ is convex and the set $\mathcal{C}$ is convex and weak*-compact. Therefore, Problem (P) satisfies the conditions of Sion's Minimax Theorem (see Appendix A.1), and hence there exists a saddle point $(Y^*, P^*) \in \mathcal{I}_0 \times \mathcal{C}$ such that

$$\begin{aligned}\sup_{Y \in \mathcal{I}_0} \inf_{P \in \mathcal{C}} \mathbb{E}_P[u(W_0 - X + Y - \Pi_0)] &= \sup_{Y \in \mathcal{I}_0} \min_{P \in \mathcal{C}} \mathbb{E}_P[u(W_0 - X + Y - \Pi_0)] \\ &= \min_{P \in \mathcal{C}} \sup_{Y \in \mathcal{I}_0} \mathbb{E}_P[u(W_0 - X + Y - \Pi_0)] \\ &= \mathbb{E}_{P^*}[u(W_0 - X + Y^* - \Pi_0)].\end{aligned}$$

For $P^* \in \mathcal{C}$, to characterize the optimal indemnity Y^*, we focus on the following inner problem:

$$\begin{aligned}&\sup_{Y \in \mathcal{I}} \quad \mathbb{E}_{P^*}[u(W_0 - X + Y - \Pi_0)] \\ &\text{s.t.} \quad \mathbb{E}_Q[v(W_0^{\text{Ins}} - (1+\rho)Y + \Pi_0)] \geqslant v(W_0^{\text{Ins}}).\end{aligned} \tag{3.2}$$

Problem (3.2) is a convex optimization problem, since the constraint can be equivalently written as $\mathbb{E}_Q[v_1(W_0^{\text{Ins}} - (1+\rho)Y + \Pi_0)] \leqslant v_1(W_0^{\text{Ins}})$, where $v_1 := -v$ is a convex utility function. For $P^* \in \mathcal{C}$, let $A := A_{P^*}$ and $h^* := h_{P^*}$ be as in Remark 3.1, and consider the following two problems:

$$\sup_{Y \in \mathcal{I}} \left\{ \int_{S \setminus A} u(W_0 - X + Y - \Pi_0) dP_s^* : 0 \leqslant Y 1_{S\setminus A} \leqslant X 1_{S\setminus A}, \int_{S\setminus A} v(W_0^{\text{Ins}} - (1+\rho)Y + \Pi_0)\, dQ = 0 \right\}. \tag{3.3}$$

$$\sup_{Y \in \mathcal{I}} \left\{ \int_A u(W_0 - X + Y - \Pi_0) h^*\, dQ : \int v(W_0^{\text{Ins}} - (1+\rho)Y + \Pi_0)\, dQ \geqslant v(W_0^{\text{Ins}}) \right\}. \tag{3.4}$$

Observe that $\overline{Y} := X$ is a feasible solution for Problem (3.3) and it holds that

$$\int_{S\setminus A} u(W_0 - X + \overline{Y} - \Pi_0) dP_s^* = u(W_0 - \Pi_0) P_s^*(S\setminus A) \geqslant \int_{S\setminus A} u(W_0 - X + Y - \Pi_0) dP_s^*,$$

for any feasible solution Y for Problem (3.3). Hence $\overline{Y} = X$ is optimal for (3.3).

Now, let $Y_1^* \in \mathcal{I}$ be an optimal solution for Problem (3.4). We claim that $Y^* := Y_1^* 1_A + X 1_{S\setminus A}$ is optimal for Problem (3.2). To see this, we remark that

$$\int_S v(W_0^{\text{Ins}} - (1+\rho)Y^* + \Pi_0)\, dQ = \int_A v(W_0^{\text{Ins}} - (1+\rho)Y_1^* + \Pi_0)\, dQ \geqslant v(W_0^{\text{Ins}}),$$

where the last inequality follows from the feasibility of Y_1^* for (3.4). Hence Y^* is feasible for Problem (3.2). The optimality of Y^* is then derived similarly to Ghossoub (2019a, Lemma C.6).

Next, we focus on the optimal indemnity Y_1^* that solves Problem (3.4). The associated Lagrange function is

$$\mathcal{L}(Y_1, \lambda) = \int_A \left[u(W_0 - X(s) + Y_1(s) - \Pi_0) h^*(s) + \lambda v(W_0^{\text{Ins}} - (1+\rho)Y_1(s) + \Pi_0) \right] dQ(s) - \lambda v(W_0^{\text{Ins}}),$$

where $\lambda \in \mathbb{R}_+$ is the Lagrange multiplier. As the domain $\mathcal{I}$ of Y_1 is convex, and $\mathcal{L}(Y_1, \lambda)$ is both concave and continuous in Y_1 with respect to supnorm $\| \cdot \|_{sup}$, as well as linear in λ, strong duality holds, i.e.,

$$\text{val}(\mathcal{L}) := \sup_{Y_1 \in \mathcal{I}} \inf_{\lambda \in \mathbb{R}_+} \mathcal{L}(Y_1, \lambda) = \inf_{\lambda \in \mathbb{R}_+} \sup_{Y_1 \in \mathcal{I}} \mathcal{L}(Y_1, \lambda),$$

where the optimal value val($\mathcal{L}$) of Problem (3.4) is finite, since (P) is finite. Moreover, by Sion's Minimax Theorem, (Y_1, λ) is a saddle point of (3.4).

For a fixed $\lambda \in \mathbb{R}_+$, a necessary and sufficient condition for $Y_1^* \in \mathcal{I}$ to be the optimal solution of Problem (3.4) is

$$\lim_{\theta \to 0^+} \mathcal{L}'((1-\theta)Y_1^* + \theta Y_1) \leqslant 0, \ \forall Y_1 \in \mathcal{I}. \tag{3.5}$$

By direct computation, (3.5) becomes

$$\int_A [u'(W_0 - X + Y_1^* - \Pi_0)h^* - \lambda(1+\rho)v'(W_0^{\text{Ins}} - (1+\rho)Y_1^* + \Pi_0)](Y_1 - Y_1^*)\, dQ \leqslant 0, \ \forall Y_1 \in \mathcal{I}. \tag{3.6}$$

Define the following sets, depending on Lagrange multiplier λ:

$$\begin{cases} A_\lambda^+ := \{s \in A : u'(W_0 - X(s) + Y_1^*(s) - \Pi_0)h^*(s) - \lambda(1+\rho)v'(W_0^{\text{Ins}} - (1+\rho)Y_1^*(s) + \Pi_0) > 0\}, \\ A_\lambda^0 := \{s \in A : u'(W_0 - X(s) + Y_1^*(s) - \Pi_0)h^*(s) - \lambda(1+\rho)v'(W_0^{\text{Ins}} - (1+\rho)Y_1^*(s) + \Pi_0) = 0\}, \\ A_\lambda^- := \{s \in A : u'(W_0 - X(s) + Y_1^*(s) - \Pi_0)h^*(s) - \lambda(1+\rho)v'(W_0^{\text{Ins}} - (1+\rho)Y_1^*(s) + \Pi_0) < 0\}. \end{cases}$$

First, observe that on A_λ^+ and A_λ^-, condition (3.6) holds for all $Y_1 \in \mathcal{I}$ only if

$$Y_1^* 1_{A_\lambda^+} = X 1_{A_\lambda^+} \text{ and } Y_1^* 1_{A_\lambda^-} = 0. \tag{3.7}$$

Next, define the set $A_{h^*} := \{s \in A : \ h^*(s) = 0\}$. To obtain the structure of Y_1^* in (3.4), we distinguish the following cases, depending on λ.

Case 3.2.1. If $\lambda > 0$, then $A_{h^*} \subseteq A_\lambda^-$ and thus $Y_1^* 1_{A_{h^*}} = 0$. On A_λ^0, Y_1^* satisfies the following condition:

$$u'(W_0 - X(s) + Y_1(s) - \Pi_0)h^*(s) - \lambda(1+\rho)v'(W_0^{\text{Ins}} - (1+\rho)Y_1(s) + \Pi_0) = 0. \tag{3.8}$$

Let Y_0^* be the solution of (3.8). In the view of (3.7), Y_1^* is thus $Y_1^* 1_{A \backslash A_{h^*}} = \max\left[0, \min\left(X, Y_0^*\right)\right] 1_{A \backslash A_{h^*}}$, which depends on the state of the world only through h^* and X.

Case 3.2.2. If $\lambda = 0$, then $A_{h^*} = A_\lambda^0$ and $u'(W_0 - X(s) + Y_1^*(s) - \Pi_0)h^*(s) > 0$, for all $s \in A \backslash A_{h^*}$. Thus $Y_1^* 1_A = Y 1_{A_{h^*}} + X 1_{A \backslash A_{h^*}}$, for any feasible $Y \in \mathcal{I}$.

The indemnity $Y_{1,\lambda}^* := Y_1^*$, depending on λ, is the optimal solution of Problem (3.4) if there exists some $\lambda^* \geqslant 0$ such that $\mathbb{E}_Q[v(W_0^{\text{Ins}} - (1+\rho)Y_{1,\lambda^*}^* + \Pi_0)] = v(W_0^{\text{Ins}})$. To see this, define the constant $\overline{\lambda} := \dfrac{\text{val}(\mathcal{L}) + \varepsilon}{v(W_0^{\text{Ins}} + \Pi_0) - v(W_0^{\text{Ins}})} \in \mathbb{R}_+$, for some large $\varepsilon > \left|u(W_0 - M - \Pi_0)\right|$. Then for any $\lambda > \overline{\lambda}$, we obtain

$$\begin{aligned} \sup_{Y_{1,\lambda} \in \mathcal{I}} \mathcal{L}(Y_{1,\lambda}, \lambda) \geqslant \mathcal{L}(0, \lambda) &= \int_A u(W_0 - X - \Pi_0)h^* dQ + \lambda\left(\int_A v(W_0^{\text{Ins}} + \Pi_0)dQ - v(W_0^{\text{Ins}})\right) \\ &\geqslant u(W_0 - M - \Pi_0)P_{ac}^*(A) + \overline{\lambda}\big(v(W_0^{\text{Ins}} + \Pi_0) - v(W_0^{\text{Ins}})\big) \\ &= u(W_0 - M - \Pi_0)P_{ac}^*(A) + \text{val}(\mathcal{L}) + \varepsilon > \text{val}(\mathcal{L}), \end{aligned}$$

where the second inequality follows from the monotonicity of the utility u. Therefore, for all $\lambda > \overline{\lambda}$, $\sup_{Y_{1,\lambda} \in \mathcal{I}} \mathcal{L}(Y_{1,\lambda}, \lambda) > \text{val}(\mathcal{L})$ and thus,

$$\inf_{\lambda \geqslant \overline{\lambda}} \sup_{Y_{1,\lambda} \in \mathcal{I}} \mathcal{L}(Y_{1,\lambda}, \lambda) > \text{val}(\mathcal{L}) = \inf_{\lambda \geqslant 0} \sup_{Y_{1,\lambda} \in \mathcal{I}} \mathcal{L}(Y_{1,\lambda}, \lambda).$$

Hence the feasible set of λ reduces to the compact interval $[0, \overline{\lambda}]$.

Now, for $\lambda \in [0, \overline{\lambda}]$, let $\text{val}(\mathcal{L}; \lambda) := \sup_{Y \in \mathcal{I}} \mathcal{L}(Y, \lambda)$ be the optimal value as a function of λ. We claim that $\text{val}(\mathcal{L}; \lambda)$ is convex in λ: let $\theta \in (0,1)$ and $\lambda_1, \lambda_2 \in [0, \overline{\lambda}]$ and consider the following:

$$\begin{aligned} \text{val}(\mathcal{L}; \theta\lambda_1 + (1-\theta)\lambda_2) &= \sup_{Y \in \mathcal{I}} \mathcal{L}(Y, \theta\lambda_1 + (1-\theta)\lambda_2) = \sup_{Y \in \mathcal{I}} \theta\mathcal{L}(Y, \lambda_1) + (1-\theta)\mathcal{L}(Y, \lambda_2) \\ &\leqslant \theta \sup_{Y \in \mathcal{I}} \mathcal{L}(Y, \lambda_1) + (1-\theta) \sup_{Y \in \mathcal{I}} \mathcal{L}(Y, \lambda_2), \end{aligned}$$

where for the second inequality we use the fact that $\mathcal{L}(Y,\lambda)$ is linear in λ, for any given Y. Thus we obtain:

$$\sup_{Y_{1,\lambda}\in\mathcal{I}}\min_{\lambda\in[0,\overline{\lambda}]}\mathcal{L}(Y_{1,\lambda},\lambda)=\sup_{Y_{1,\lambda}\in\mathcal{I}}\mathcal{L}(Y_{1,\lambda^*},\lambda^*)=\mathcal{L}(Y^*_{1,\lambda^*},\lambda^*)=\inf_{\lambda\in[0,\overline{\lambda}]}\operatorname{val}(\mathcal{L};\lambda)=\operatorname{val}(\mathcal{L}),$$

where $\lambda^*(\mathbb{E}_Q[v(W_0^{\text{Ins}}-(1+\rho)Y^*_{1,\lambda^*}+\Pi_0)]-v(W_0^{\text{Ins}}))=0$.

Therefore, Y_1^* is an optimal solution to (3.1) if and only if Y_1^* is of the form $Y_1^*=\widetilde{Y}^*\mathbf{1}_{A\setminus A_{h^*}}+Y_{h^*}\mathbf{1}_{A_{h^*}}$, where $\widetilde{Y}^*$ and Y_{h^*} are defined in Case (3.2.1) and (3.2.2), respectively. To see this, let $Y_1\in\mathcal{I}$ be a feasible solution for (3.1) and consider the following:

(3.9)

$$\begin{aligned}\mathcal{L}(Y_1^*,\lambda^*)-\mathcal{L}(Y_1,\lambda^*)&=\int_A[u(W_0-X+Y_1^*-\Pi_0)h^*+\lambda^*v(W_0^{Ins}-(1+\rho)Y_1^*+\Pi_0)]dQ\\&\quad-\int_A[u(W_0-X+Y_1-\Pi_0)h^*+\lambda^*v(W_0^{Ins}-(1+\rho)Y_1+\Pi_0)]dQ\\&=\int_A[u(W_0-X+Y_1^*-\Pi_0)-u(W_0-X+Y_1-\Pi_0)]h^*dQ\\&\quad-\lambda^*\int_A[v(W_0^{Ins}-(1+\rho)Y_1^*+\Pi_0)-v(W_0^{Ins}-(1+\rho)Y_1+\Pi_0)]dQ\\&\geqslant\int_A\big(u'(W_0-X+Y_1^*-\Pi_0)h^*-\lambda^*v'(W_0^{Ins}-(1+\rho)Y_1^*+\Pi_0)\big)(Y_1^*-Y_1)dQ\\&=\int_{A^+_{\lambda^*}}\big(u'(W_0-\Pi_0)h^*-\lambda^*v'(W_0^{Ins}-(1+\rho)X+\Pi_0)\big)(X-Y_1)dQ\\&\quad+\int_{A^-_{\lambda^*}}\big(u'(W_0-X-\Pi_0)h^*-\lambda^*v'(W_0^{Ins}+\Pi_0)\big)(-Y_1)dQ\geqslant 0,\end{aligned}$$

where the third inequality uses the first-order Taylor approximation for a concave function. □

Note that the optimal indemnity Y^* in eq. (3.1) provides full insurance over the event $S\setminus A$, whenever the DM's worst-case measure P^* is such that $P^*(S\setminus A)\neq 0$. Note also that on the event A, Y^* satisfies an ordinary differential equation for which an analytical expression is difficult to provide in general. However, for particular choices of $\mathcal{C}$, we can obtain numerically the structure of Y^*, as well as P^* (see Example 3.6).

3.2. **With the No-Sabotage Condition.** Next we analyze the case when the set $\mathcal{F}$ of Problem (P) is restricted to the set of indemnities satisfying the no-sabotage condition. In this case the feasibility set becomes:

$$\hat{\mathcal{I}}_0:=\Big\{\hat{Y}\in\hat{\mathcal{I}}:\ \mathbb{E}_Q[v(W_0^{\text{Ins}}-(1+\rho)\hat{Y}+\Pi_0)]\geqslant v(W_0^{\text{Ins}})\Big\}.\tag{3.10}$$

Remark 3.3. Since $\hat{\mathcal{I}}$ is a compact subset of the space $(C[0,M],\|\cdot\|_{sup})$ (see Remark 2.2), and $\hat{\mathcal{I}}_0$ is a closed subset of $\hat{\mathcal{I}}$, it follows that $\hat{\mathcal{I}}_0$ is compact.

Theorem 3.4. *Suppose that the utility functions u and v satisfy Assumption 1. Let $\mathcal{F}=\hat{\mathcal{I}}$ as defined in eq. (2.3) be the set of admissible indemnity functions. Then there exists $P^*\in\mathcal{C}$ such that $\hat{Y}^*\in\hat{\mathcal{I}}$ is optimal solution of Problem (P) if and only if it is Q-a.s. of the form*

$$\hat{Y}^*=\hat{Y}_1^*1_A+X1_{S\setminus A},$$

where

(a) $A \in \Sigma$ *is such that* $P^* = P^*_{ac} + P^*_s$, *with* $P^*_s(A) = Q(S \backslash A) = 0$;

(b) $h^* : S \to [0, \infty)$ *is such that* $h^* = dP^*_{ac}/dQ$;

(c) $\xi^* : \mathbb{R}_+ \to \mathbb{R}_+$ *is a Borel measurable function such that* $h^* = \xi^* \circ X$;

(d) $\hat{Y}^*_1 = \hat{I}^* \circ X$, *where* $\hat{I}^*(x) = \int_0^x (\hat{I}^*)'(t)\, dt$, $\forall x \in [0, M]$ *and*

$$(\hat{I}^*)'(t) = \begin{cases} 0, & \text{if } \int_{[t,M] \cap X(A)} (u'(W_0 - x + \hat{I}^*(x) - \Pi_0)\xi^*(x) - \lambda^*(1+\rho)v'(W_0^{Ins} - (1+\rho)\hat{I}^*(x) + \Pi_0))\, dF_{X,Q}(x) < 0, \\ \kappa(t), & \text{if } \int_{[t,M] \cap X(A)} (u'(W_0 - x + \hat{I}^*(x) - \Pi_0)\xi^*(x) - \lambda^*(1+\rho)v'(W_0^{Ins} - (1+\rho)\hat{I}^*(x) + \Pi_0))\, dF_{X,Q}(x) = 0, \\ 1, & \text{if } \int_{[t,M] \cap X(A)} (u'(W_0 - x + \hat{I}^*(x) - \Pi_0)\xi^*(x) - \lambda^*(1+\rho)v'(W_0^{Ins} - (1+\rho)\hat{I}^*(x) + \Pi_0))\, dF_{X,Q}(x) > 0, \end{cases}$$

for some Lebesgue measurable and $[0,1]$*-valued function* κ;

(e) $\lambda^* \in \mathbb{R}_+$ *is such that* $\lambda^*\big(\mathbb{E}_Q[v(W_0^{Ins} - (1+\rho)\hat{Y}^* + \Pi_0)] - v(W_0^{Ins})\big) = 0$.

Proof. Similarly to Theorem 3.2, there exists a saddle point $(\hat{Y}^*, P^*) \in \hat{\mathcal{I}}_0 \times \mathcal{C}$ such that

$$\begin{aligned} \sup_{\hat{Y} \in \hat{\mathcal{I}}_0} \inf_{P \in \mathcal{C}} \mathbb{E}_P[u(W_0 - X + \hat{Y} - \Pi_0)] &= \min_{P \in \mathcal{C}} \max_{\hat{Y} \in \hat{\mathcal{I}}_0} \mathbb{E}_P[u(W_0 - X + \hat{Y} - \Pi_0)] \\ &= \mathbb{E}_{P^*}[u(W_0 - X + \hat{Y}^* - \Pi_0)] \end{aligned} \tag{3.11}$$

where $\hat{\mathcal{I}}_0$ given in eq. (3.10) is compact (see Remark 3.3). For $P^* \in \mathcal{C}$, the inner optimization problem in (3.11) becomes:

$$\begin{aligned} \sup_{\hat{Y} \in \hat{\mathcal{I}}} \quad & \int_A u(W_0 - X + \hat{Y} - \Pi_0) h^*\, dQ + \int_{S \backslash A} u(W_0 - X + \hat{Y} - \Pi_0)\, dP^*_s \\ \text{s.t.} \quad & \int_A v(W_0^{\text{Ins}} - (1+\rho)\hat{Y} + \Pi_0)\, dQ \geqslant v(W_0^{\text{Ins}}), \end{aligned}$$

where $A := A_{P^*}$ and $h^* := h_{P^*}$, depending on P^*, are defined in Remark 3.1. Similar to Theorem 3.2, the optimal indemnity function $\hat{Y}^*$ can be obtained as $\hat{Y}^* = \hat{Y}^*_1 1_A + X 1_{S \backslash A}$, where $\hat{Y}^*_1$ solves Problem (3.12) below.

$$\sup_{\hat{Y}_1 \in \hat{\mathcal{I}}} \left\{ \int u(W_0 - X + \hat{Y}_1 - \Pi_0) h^*\, dQ \; : \; \int v(W_0^{\text{Ins}} - (1+\rho)\hat{Y}_1 + \Pi_0)\, dQ \geqslant v(W_0^{\text{Ins}}) \right\}. \tag{3.12}$$

The Lagrange function of Problem (3.12) is

$$\mathcal{L}(\hat{Y}_1, \lambda) = \int_A \big[u(W_0 - X + \hat{Y}_1 - \Pi_0) h^* + \lambda v(W_0^{\text{Ins}} - (1+\rho)\hat{Y}_1 + \Pi_0)\big]\, dQ - \lambda v(W_0^{\text{Ins}}),$$

where $\lambda \in \mathbb{R}_+$ is the Lagrange multiplier. The domain $\hat{\mathcal{I}}$ of $\hat{Y}_1$ is convex, and $\mathcal{L}(\hat{Y}_1, \lambda)$ is concave in $\hat{Y}_1$, continuous in Y_1 with respect to supnorm $\|\cdot\|_{sup}$, and linear in λ. Thus the strong duality holds. Therefore, for fixed $\lambda \in \mathbb{R}_+$, a necessary and sufficient condition for $\hat{Y}^*_1 \in \mathcal{I}$ to be the optimal solution of Problem (3.12) is

$$\lim_{\theta \to 0^+} \mathcal{L}'((1-\theta)\hat{Y}^*_1 + \theta \hat{Y}_1) \leqslant 0, \; \forall\, \hat{Y}_1 \in \hat{\mathcal{I}}. \tag{3.13}$$

Since $Q(S\backslash A) = 0$, we can extend the domain of the integral above in (3.13) over $[0, M]$. By direct computation, (3.13) becomes: for all $\hat{Y}_1 = \hat{I}(X) \in \hat{\mathcal{I}}$,

(3.14)
$$\int_0^M \left[u'(W_0 - t + \hat{I}^*(t) - \Pi_0)\xi^*(t) - \lambda(1+\rho)v'(W_0^{\text{Ins}} - (1+\rho)\hat{I}^*(t) + \Pi_0)\right]\left(\hat{I}(t) - \hat{I}^*(t)\right) dQ{\circ}X^{-1}(t) \leqslant 0.$$

As any $\hat{Y}_1 = \hat{I}(X) \in \hat{\mathcal{I}}$ is absolutely continuous, it is almost everywhere differentiable on $[0, M]$, and hence (3.14) becomes

$$\begin{aligned} 0 &\geqslant \int_0^M \int_0^t (u'(W_0 - t - \hat{I}^*(t) - \Pi_0)\xi^*(t) - \lambda(1+\rho)v'(W_0^{\text{Ins}} - (1+\rho)\hat{I}^*(t) + \Pi_0))\times \\ &\qquad\qquad \times (\hat{I}'(x) - (\hat{I}^*)'(x))\, dx\, dQ \circ X^{-1}(t) \\ &= \int_0^M \int_x^M (u'(W_0 - t - \hat{I}^*(t) - \Pi_0)\xi^*(t) - \lambda(1+\rho)v'(W_0^{\text{Ins}} - (1+\rho)\hat{I}^*(t) + \Pi_0))\, dQ \circ X^{-1}(t)\times \\ &\qquad\qquad \times (\hat{I}'(x) - (\hat{I}^*)'(x))\, dx, \end{aligned}$$

for all $\hat{Y}_1 = \hat{I}(X) \in \hat{\mathcal{I}}$; hence $\hat{Y}_1^* = \hat{I}^*(X)$ is of the form:

$$(\hat{I}^*)'(x) = \begin{cases} 0, & \text{if } \displaystyle\int_{[x,M]\cap X(A)} \tau(t)\, dQ \circ X^{-1}(t) < 0, \\ \kappa(x), & \text{if } \displaystyle\int_{[x,M]\cap X(A)} \tau(t)\, dQ \circ X^{-1}(t) = 0, \\ 1, & \text{if } \displaystyle\int_{[x,M]\cap X(A)} \tau(t)\, dQ \circ X^{-1}(t) > 0, \end{cases}$$

where $\tau(x) := u'(W_0 - x + \hat{I}^*(x) - \Pi_0)\xi^*(x) - \lambda^*(1+\rho)v'(W_0^{\text{Ins}} - (1+\rho)\hat{I}^*(x) + \Pi_0)$ and κ is some Lebesgue measurable and $[0,1]$-valued function. The existence of the Lagrange multiplier $\lambda^* \in \mathbb{R}_+$ that guarantees the existence of the solution $\hat{Y}_1^*$ follows similar to Theorem 3.2.

Thus $\hat{Y}_1^* = \hat{I}^*(X) \in \hat{\mathcal{I}}$ is an optimal solution of (3.12) if and only if $\hat{Y}_1^*$ is of the form (d). To see this, let $\hat{Y}_1 \in \hat{\mathcal{I}}$ be a feasible solution of (3.12) and, similar to (3.9) in Theorem (3.2), we have:

$$\begin{aligned} \mathcal{L}(\hat{Y}_1^*, \lambda^*) - \mathcal{L}(\hat{Y}_1, \lambda^*) &\geqslant \int_0^M \int_x^M \tau(t) dQ \circ X^{-1}(t)((\hat{I}_1^*)'(x) - \hat{I}_1'(x))dx \\ &= \int_{X(A^+_{\lambda^*})} \int_{[x,M]\cap X(A)} \tau(t) dQ \circ X^{-1}(t)((\hat{I}_1^*)'(x) - \hat{I}_1'(x))dx \\ &\quad + \int_{X(A^0_{\lambda^*})} \int_{[x,M]\cap X(A)} \tau(t) dQ \circ X^{-1}(t)((\hat{I}_1^*)'(x) - \hat{I}_1'(x))dx \\ &\quad + \int_{X(A^-_{\lambda^*})} \int_{[x,M]\cap X(A)} \tau(t) dQ \circ X^{-1}(t)((\hat{I}_1^*)'(x) - \hat{I}_1'(x))dx \\ &= \int_{X(A^+_{\lambda^*})} \int_{[x,M]\cap X(A)} \tau(t) dQ \circ X^{-1}(t)(1 - \hat{I}_1'(x))dx \\ &\quad + \int_{X(A^-_{\lambda^*})} \int_{[x,M]\cap X(A)} \tau(t) dQ \circ X^{-1}(t)(-\hat{I}_1'(x))dx \geqslant 0, \end{aligned}$$

where $X(A^+_{\lambda^*}) = \{x \in [0,M] \cap X(A) : \int_{[x,M]\cap X(A^+_{\lambda^*})} \tau(t) dQ \circ X^{-1}(t) > 0\}$ and $X(A^0_{\lambda^*})$ and $X(A^-_{\lambda^*})$ are defined similarly. □

Theorem 3.4 above provides a general characterization of the optimal solution of Problem (P), when the set of admissible indemnity functions is given by $\hat{\mathcal{I}}$. The exact structure of the optimal indemnity $\hat{Y}^*$

may be difficult to interpret, due to its implicit form. However, under the (more common) assumption of risk-neutrality of the insurer, closed-form solutions for $\hat{Y}^*$ can be obtained, as seen in the Section 4.

3.3. **Numerical Example.** This section presents a numerical example that illustrates the structure of the optimal indemnity $\hat{Y}^*$, as well as the worst-case distribution P^*, obtained in Section 3, when the ambiguity set $\mathcal{C}$ is constructed as a specific neighbourhood around a *reference/ baseline distribution*. Throughout this analysis, we assume that the underlying space S is a Polish space, equipped with its Borel sigma-algebra.

As before, X is a nonnegative random variable representing the insurable loss, whose true distribution may be unknown. The insurer's belief $Q \in ca_1^+(\Sigma)$ regarding the loss X can be the empirical distribution, derived from experts' opinion or estimated using standard statistical tools. The DM's ambiguity regarding the realizations of X is described by a δ-neighbourhood around Q defined as:

$$\mathcal{C}_\delta := \{P \in ca_1^+(\Sigma) : \mathrm{d}(P, Q) \leqslant \delta\}, \tag{3.15}$$

where $\mathrm{d} : ca_1^+(\Sigma) \times ca_1^+(\Sigma) \to \mathbb{R}_+$ is some *discrepancy measure* between probability measures P and Q, and $\delta > 0$ is a *tolerance level/ambiguity radius*. The mapping d satisfies $\mathrm{d}(P, Q) = 0$ if and only if $P = Q$. It is worth mentioning that the worst-case distribution P^* depends not only on the choice of d, but also on the ambiguity radius δ. In general, the size of $\mathcal{C}_\delta$ is connected to the amount of observations available: if δ is close to zero, the impact of ambiguity is negligible; while large values of δ indicate high levels of ambiguity. The question of how to optimally choose the ambiguity radius is an ongoing stream of research in robust optimization. One possible approach is to interpret δ as the degree of ambiguity about the reference model and thus argues that this choice depends on the risk preferences of market participants (e.g., Breuer and Csiszár (2016); Wozabal (2012)). In Example 3.6, we follow this approach and solve Problem (P) for different levels of ambiguity. This allows us to analyze the impact of ambiguity on the optimal indemnity $\hat{Y}^*$ and the worst-case distribution P^*.

The following observation characterizes the change in the DM's expected utility, as a function of the ambiguity radius δ. This dynamic is later illustrated in Figure 3 in Example 3.6.

Remark 3.5. For a fixed premium $\Pi_0 > 0$, let $\hat{\mathcal{I}}_0$ (defined in eq. (3.10)) be the feasible set of indemnities in Theorem 3.4. Moreover, for a discrepancy measure d and some ambiguity radii $\delta_1 \leqslant \delta_2$, let $\mathcal{C}_{\delta_1}$ and $\mathcal{C}_{\delta_2}$ be the corresponding ambiguity sets, as defined in eq. (3.15). Let $(\hat{Y}_1^*, P_1^*)$ and $(\hat{Y}_2^*, P_2^*)$ be the saddle points of Problems (P), for $\mathcal{C}_{\delta_1}$ and $\mathcal{C}_{\delta_2}$, respectively. It holds that

$$\mathbb{E}_{P_2^*}[u(W_0 - X + \hat{Y}_2^* + \Pi_0)] \leqslant \mathbb{E}_{P_1^*}[u(W_0 - X + \hat{Y}_2^* + \Pi_0)] \leqslant \mathbb{E}_{P_1^*}[u(W_0 - X + \hat{Y}_1^* + \Pi_0)],$$

where the first inequality follows from $\mathcal{C}_{\delta_1} \subseteq \mathcal{C}_{\delta_2}$, as $\delta_1 \leqslant \delta_2$. Hence, for increasing values of δ, the optimal DM's expected utility decreases.

Example 3.6 (Rényi ambiguity set)**.** For this example we focus on Problem (P), when the admissible set of indemnities is $\mathcal{F} = \hat{\mathcal{I}}$ as defined in eq. (2.3). Let DM's ambiguity set $\mathcal{C}_\delta$ be given by

$$\mathcal{C}_\delta^{D_\alpha} := \{P \ll Q : D_\alpha(P\|Q) \leqslant \delta\},$$

where D_α is the Rényi divergence of order α between P and Q, i.e.,

$$D_\alpha(P\|Q) := \frac{1}{\alpha - 1} \log \mathbb{E}_Q\left[\left(\frac{dP}{dQ}\right)^\alpha\right].$$

We observe that for every $\alpha > 1$, $D_\alpha(P\|Q) = 0$ if and only if $P = Q$. When $\alpha \to 1$, D_α is the well-known Kullback-Leibler divergence. Moreover, since S is a Polish space, for any ambiguity radius $\delta \in [0, \infty)$ and degree $\alpha \geqslant 1$ the set $\mathcal{C}_\delta^{D_\alpha}$ is a convex and compact in the topology of weak convergence (e.g., Van Erven and Harremos (2014, Theorem 20)). For more on the properties of the divergence D_α, we refer to Rényi (1961) and Liese and Vajda (1987).

To illustrate our results, we follow the existing literature and consider a discretely distributed loss X. For a sample of size n, we assume without loss of generality that $x_1 \leqslant \cdots \leqslant x_n$, and we denote this n-sample by $\mathbf{x} = [x_1, \ldots, x_n]^\top$. For our example, a random sample $\mathbf{x}$ of size $n = 100$ is drawn from a truncated exponential distribution with mean parameter $\mu = 20$, and with an upper bound $M = W_0 - \Pi_0$. Moreover, the insurer's belief $\widehat{Q}$ is the empirical distribution of the sample $\mathbf{x}$. Let $\widehat{\mathbf{q}} = [\widehat{q}_1, \ldots, \widehat{q}_n]^\top$ be the insurer's probability mass function (pmf), where $\widehat{q}_i := \widehat{Q}(X = x_i)$, $\widehat{q}_i \geqslant 0$, $i = 1, \ldots, n$, $\mathbf{1}^\top \widehat{\mathbf{q}} = 1$.

Let $\hat{\mathbf{y}} = [\hat{y}_1, \ldots, \hat{y}_n]^\top \in \mathbb{R}^n_+$ be the indemnification function corresponding to the loss $\mathbf{x}$. Following the approach in Asimit et al. (2017), the feasibility constraints $0 \leqslant \hat{y}_i \leqslant x_i$ and $0 \leqslant \hat{y}_i - \hat{y}_{i-1} \leqslant x_i - x_{i-1}$, for $i = 1, \ldots, n$, are represented by $\mathbf{0} \leqslant \hat{\mathbf{y}} \leqslant \mathbf{x}$ and $\mathbf{0} \leqslant \mathbf{A}_{n-1}\hat{\mathbf{y}} \leqslant \mathbf{A}_{n-1}\mathbf{x}$, where for $i = 1, \ldots, n-1$, the matrix $\mathbf{A}_i \in \mathbb{R}^{(n-1)\times n}$ is defined as follows:

$$\mathbf{A}_i := \begin{pmatrix} -1 & 1 & 0 & 0 & \ldots & & \ldots & & \ldots & 0 \\ 0 & -1 & 1 & 0 & \ldots & & \ldots & & \ldots & 0 \\ & \ddots & & \ddots & & \ddots & & \ddots & & \ddots \\ \ldots & & \ldots & & \ldots & 0 & -1 & 1 & 0 & \ldots \\ \ldots & & \ldots & & \ldots & 0 & 0 & 0 & 0 & \ldots \\ & \ddots & & \ddots & & \ddots & & \ddots & & \ddots \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i\text{-th row.} \\ \\ \\ \end{matrix} \tag{3.16}$$

Moreover, $\mathbf{p} = [p_1, \ldots, p_n]^\top \in [0,1]^n$ belongs to $\mathcal{C}_\delta^{D_\alpha}$ if it satisfies the following conditions:

(i) is a pmf: $\mathbf{1}^\top \mathbf{p} = 1$;
(ii) is absolutely continuous with respect to $\widehat{\mathbf{q}}$: if $\exists\, i \in \{1, \ldots, n\}$ such that $\widehat{q}_i = 0$, then $p_i = 0$.
(iii) lies in a Rényi ambiguity set around $\widehat{\mathbf{q}}$:
$$\mathbf{p}^\alpha \cdot \widehat{\mathbf{q}}^{\,1-\alpha} = \sum_{i=1}^n p_i^\alpha\, \widehat{q}_i^{\,1-\alpha} \leqslant \overline{\delta},$$
where $\overline{\delta} := \exp(\delta(\alpha - 1))$.

To simplify the notation, let $D := \{\mathbf{p} \in [0,1]^n : p_i = 0 \text{ if } \widehat{q}_i = 0, i = 1, \ldots, n\}$. With the above representations for the variables $\hat{\mathbf{y}}$ and $\mathbf{p}$, Problem (P) can be formulated as follows:

$$(\mathrm{P}^n) \qquad \begin{cases} \max\limits_{\hat{\mathbf{y}} \in \mathbb{R}^n_+} \ \min\limits_{\mathbf{p} \in D} & \displaystyle\sum_{i=1}^n u(W_0 - x_i + \hat{y}_i - \Pi_0) p_i \\ \qquad \text{s.t.} & \mathbf{0} \leqslant \mathbf{A}_{n-1}\hat{\mathbf{y}} \leqslant \mathbf{A}_{n-1}\mathbf{x}, \\ & \mathbf{0} \leqslant \hat{\mathbf{y}} \leqslant \mathbf{x}, \\ & \displaystyle\sum_{i=1}^n -v(W_0^{\mathrm{Ins}} - (1+\theta)\hat{y}_i + \Pi_0)\widehat{q}_i \leqslant -v(W_0^{\mathrm{Ins}}), \\ & \mathbf{p}^\alpha \cdot \widehat{\mathbf{q}}^{\,1-\alpha} \leqslant \overline{\delta}, \\ & \mathbf{1}^\top \mathbf{p} = 1. \end{cases}$$

Observe that Problem (P^n) is a convex optimization problem, as the objective function is concave in $\hat{y}_i$ and linear in p_i, for $i = 1, \ldots, n$, while the constraints are convex in $\hat{y}_i$ and p_i, for any $\alpha > 1$. Problem (P^n) is solved via successive convex programming (SCP – see Pflug and Picher (2014)). The idea is to approximate the infinite dimensional ambiguity set $\mathcal{C}_\delta^{D_\alpha}$ by a finitely generated set $\mathcal{P}^{(m)} := \{\widehat{\mathbf{q}}, \mathbf{p}^{(1)}, \ldots, \mathbf{p}^{(m)}\}$, obtained iteratively from solving the inner problem in (P^n). The algorithm starts with $m = 0$, $\mathcal{P}^{(m)} := \mathcal{P}^{(0)} = \{\widehat{\mathbf{q}}\}$, and solves the outer problem:

$$(\mathrm{P}^n_{\text{outer}}) \qquad \begin{cases} \max\limits_{\hat{\mathbf{y}}\in\mathbb{R}^n_+} \min\limits_{\mathbf{p}\in\mathcal{P}^{(m)}} & \sum\limits_{i=1}^{n} u(W_0 - x_i + \hat{y}_i - \Pi_0)p_i \\ \text{s.t.} & \mathbf{0} \leqslant \mathbf{A}_{n-1}\hat{\mathbf{y}} \leqslant \mathbf{A}_{n-1}\mathbf{x}, \\ & \mathbf{0} \leqslant \hat{\mathbf{y}} \leqslant \mathbf{x}, \\ & \sum\limits_{i=1}^{n} -v(W_0^{\text{Ins}} - (1+\theta)\hat{y}_i + \Pi_0)\widehat{q}_i \leqslant -v(W_0^{\text{Ins}}). \end{cases}$$

The solution $\hat{\mathbf{y}}^{(m)} := \hat{\mathbf{y}}^{(1)}$ acts as input for the inner problem:

$$(\mathrm{P}^n_{\text{inner}}) \qquad \begin{cases} \min\limits_{\mathbf{p}\in D} & \sum\limits_{i=1}^{n} u(W_0 - x_i + \hat{y}_i - \Pi_0)p_i \\ \text{s.t.} & \mathbf{p}^{\alpha} \cdot \widehat{\mathbf{q}}^{1-\alpha} \leqslant \overline{\delta}, \\ & \mathbf{1}^{\top}\mathbf{p} = 1. \end{cases}$$

The new $\mathbf{p}^{(m+1)} := \mathbf{p}^{(1)}$ is added to the discrete set, i.e., $\mathcal{P}^{(m+1)} = \mathcal{P}^{(m)} \cup \{\mathbf{p}^{(m+1)}\}$, $m = m+1$, and the outer Problem ($\mathrm{P}^n_{\text{outer}}$) is solved using the updated $\mathcal{P}^{(m)}$. The algorithm stops when no new model is found. The convergence of the algorithm is proven in Pflug and Picher (2014). For completeness, a sketch of the proof of this result in our setting is presented in Appendix C.

To obtain an explicit solution, suppose that the DM's initial wealth is $W_0 = 200$, the insurance premium is $\Pi_0 = 10$, the safety loading is $\rho = 0.2$, and the DM's utility is given by $u(x) = x^{1/3}$. The insurers' initial wealth is $W_0^{\text{Ins}} = 400$, and the utility is $v(x) = x^{1/5}$. For the ambiguity set $\mathcal{C}_{\delta}^{D_\alpha}$, we choose the ambiguity radius $\delta = 0.7$ and the order of Rényi divergence $\alpha = 3$. Figure 1 shows the optimal indemnity $\hat{\mathbf{y}}^*$ (left) and the worst-case distribution $F_{X,P*}$, corresponding to $\mathbf{p}^*$ (right).

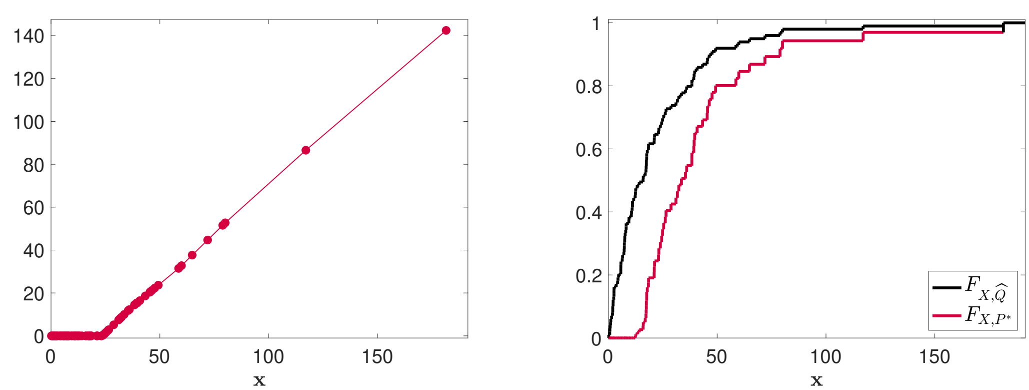

FIGURE 1. Left: the optimal indemnity $\hat{\mathbf{y}}^*$ as function of $\mathbf{x}$. Right: the DM's optimal distribution $F_{X,P*}$ (red) compared to insurers' belief $F_{X,\widehat{Q}}$ (black).

We also solve Problem (P^n) for the same ambiguity set $\mathcal{C}_{\delta}^{D_\alpha}$, when the feasibility set is $\mathcal{F} = \mathcal{I}$ as defined in eq. (2.2). This implies that the constraint $\mathbf{0} \leqslant \mathbf{A}_{n-1}\hat{\mathbf{y}} \leqslant \mathbf{A}_{n-1}\mathbf{x}$ is removed from the optimization Problem (P^n). Figure 2 illustrates the difference between the optimal indemnities corresponding to $\mathcal{I}$ and $\hat{\mathcal{I}}$.

We next investigate the decrease in optimal expected utility, when the ambiguity set increases. Certainty equivalence is used to quantify the impact of ambiguity radii on the optimal value of Problem (P^n). For

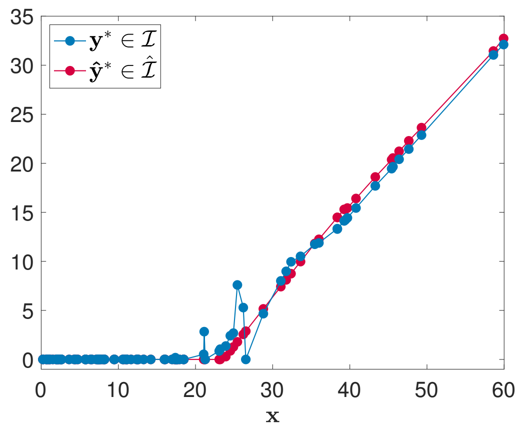

FIGURE 2. Optimal indemnities $\mathbf{y}^*$ and $\hat{\mathbf{y}}^*$ for Problem (P^n) when the feasibility sets are $\mathcal{I}$ and $\hat{\mathcal{I}}$, respectively.

each δ, let $(\hat{\mathbf{y}}^*, \mathbf{p}^*)$ be an optimal solution of (P^n) and define the certainty equivalents CE_1 and CE_2 as follows:

$$\begin{cases} \inf_{P \in \mathcal{C}_\delta^{D_\alpha}} \mathbb{E}_P[u(W_0 - \mathbf{x} + \mathrm{CE}_1(\delta))] = \sup_{\hat{\mathbf{y}} \in \hat{\mathcal{I}}} \inf_{P \in \mathcal{C}_\delta^{D_\alpha}} \mathbb{E}_P[u(W_0 - \mathbf{x} + \hat{\mathbf{y}} - \Pi_0)], \\ u(\mathrm{CE}_2(\delta)) = \sup_{\hat{\mathbf{y}} \in \hat{\mathcal{I}}} \inf_{P \in \mathcal{C}_\delta^{D_\alpha}} \mathbb{E}_P[u(W_0 - \mathbf{x} + \hat{\mathbf{y}} - \Pi_0)], \end{cases}$$

where P is the probability measure corresponding to $\mathbf{p}$. The constant CE_1 quantifies the marginal benefit of the optimal insurance contract, which we interpret as the willingness-to-pay for insurance. Moreover, CE_2 measures the certainty equivalent of DM's final wealth position. Figure 3 displays the changes in certainty equivalents for increased values of ambiguity radius. The left figure shows that a larger ambiguity radius yields a higher marginal benefit of the optimal insurance contract. This implies that the DM has a higher willingness-to-pay for the optimal insurance contract if the ambiguity set gets larger. On the other hand, the certainty equivalent of the final wealth position decreases when the ambiguity set gets larger because the DM is more ambiguity-averse (Figure 3 (right)).

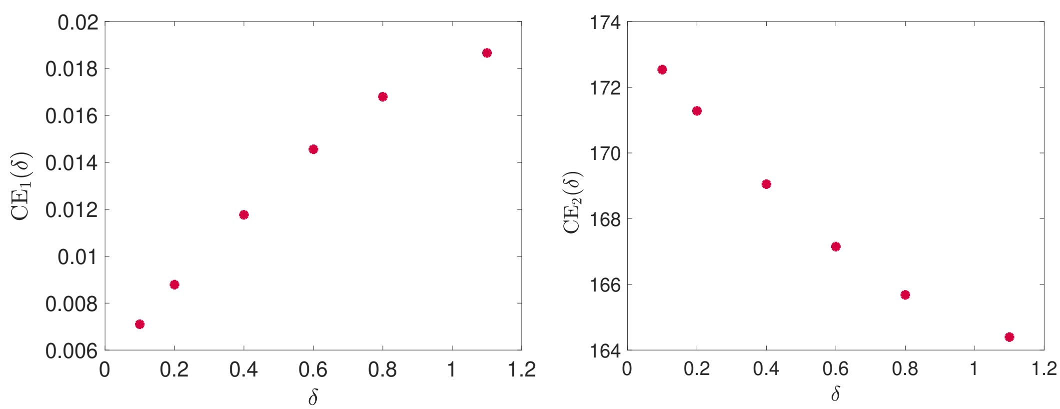

FIGURE 3. Left: certainty equivalent CE_1 as function of the ambiguity radius δ. Right: certainty equivalent CE_2 as function of the ambiguity radius δ.

Figure 4 (left) provides a closer look at the optimal indemnities $\hat{\mathbf{y}}^*$, when the ambiguity set $\mathcal{C}_{\delta}^{D_\alpha}$ becomes wider. Figure 4 (right) shows the worst-case distribution F_{X,P^*} for several values of δ. For increasing values of the ambiguity radius, it can be observed that each F_{X,P^*} dominates all the previous distributions in the first stochastic order.

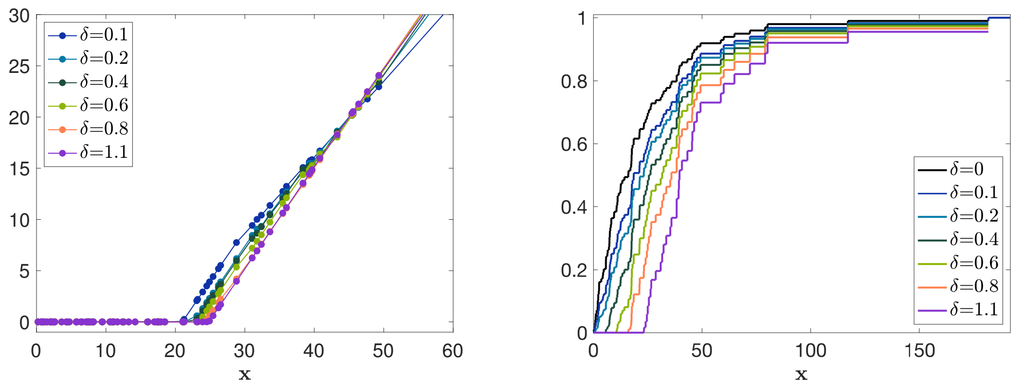

FIGURE 4. Left: optimal indemnities $\hat{\mathbf{y}}^* \in \hat{\mathcal{I}}$ of Problem (P^n) for several values of δ. Right: the corresponding worst-case probability distributions F_{X,P^*} for several values of δ.

4. The case of a Risk-Neutral Insurer

In this section, we examine the case of a risk-neutral insurer, that is, when the utility function v is linear, and we characterize the optimal indemnity both without and with the no-sabotage condition. In the former case, we obtain a closed-form characterization of the optimal indemnity (Proposition 4.1), whereas in the latter case, the optimal indemnity is determined implicitly (Proposition 4.3). The results are illustrated in Example 4.4 for a specific ambiguity set $\mathcal{C}$, and closed-form solutions are obtained. We conclude the section with a concrete example, to illustrate the structure of the optimal indemnity function, when the DM's ambiguity set is a Wasserstein ball centered around the insurer's belief Q. By specifying the ambiguity set $\mathcal{C}$, the optimal measures P^* in Proposition 4.1 and Proposition 4.3 are obtained numerically.

Specifically, we study Problem (P) under the assumption of risk neutrality of insurer:

$$(P_1) \qquad \begin{cases} \sup\limits_{Y\in\mathcal{F}} \inf\limits_{P\in\mathcal{C}} & \mathbb{E}_P[u(W_0 - X + Y - \Pi_0)] \\ \text{s.t.} & (1+\rho)\mathbb{E}_Q[Y] \leqslant \Pi_0. \end{cases}$$

If $(1+\rho)\mathbb{E}_Q[X] \leqslant \Pi_0$, then we can eliminate the constraint in Problem (P_1), as the insurance premium is large. In this case, the optimal indemnity is $Y^* = X$, Q-a.s. In the following, we assume that $(1+\rho)\,\mathbb{E}_Q[X] > \Pi_0$.

4.1. Without the No-Sabotage Condition.

Proposition 4.1. *Suppose that the utility function u satisfies Assumption 1 and is, in addition, strictly concave and such that $\lim\limits_{x\to-\infty} u'(x) = +\infty$ and $\lim\limits_{x\to+\infty} u'(x) = 0$. Let $\mathcal{F} = \mathcal{I}$ as defined in eq. (2.2) be the set of admissible indemnity functions. Then there exists $P^* \in \mathcal{C}$ such that an optimal solution $Y^* \in \mathcal{I}$ of Problem (P_1) is of the form:*

$$Y^* = (X - R^*)1_{A\setminus A_{h^*}} + (X - R_{h^*})1_{A_{h^*}} + X1_{S\setminus A}, \tag{4.1}$$

where

(a) $A \in \Sigma$ *is such that* $P^* = P^*_{ac} + P^*_s$*, with* $P^*_s(A) = Q(S \backslash A) = 0$*;*

(b) $h^* : S \to [0, \infty)$ *is such that* $h^* = dP^*_{ac}/dQ$*;*

(c) $A_{h^*} := \{s \in A : h^*(s) = 0\}$*;*

(d) R^* *and* R_{h^*} *are given by:*

Case 1. *If* $(1+\rho)\mathbb{E}_Q[X 1_{S \backslash A_{h^*}}] > \Pi_0$*, then* $R^* = \max\left[0, \min\left(X, W_0 - \Pi_0 - (u')^{-1}\left(\frac{\lambda^*}{h^*}\right)\right)\right] 1_{A \backslash A_{h^*}}$ *and* $R_{h^*} = X 1_{A_{h^*}}$*, Q-a.s., where* $\lambda^* \in \mathbb{R}_+$ *is such that* $(1+\rho)\mathbb{E}_Q[Y^*] = \Pi_0$*;*

Case 2. *If* $(1+\rho)\mathbb{E}_Q[X 1_{S \backslash A_{h^*}}] \leqslant \Pi_0$*, then* $R^* = 0$*, Q-a.s. and* $R_{h^*} = c\, X 1_{A_{h^*}}$*, where* $c \in (0, 1]$ *is defined as* $c := \dfrac{\mathbb{E}_Q[X] - (1+\rho)^{-1}\Pi_0}{\mathbb{E}_Q[X 1_{A_{h^*}}]}$.

The above result is a special case of Theorem 3.2.

Remark 4.2. In the setting of Proposition 4.1, Case (2), an important special case is when $h^* \equiv 0$, i.e., $P^* \perp Q$, where P^* is the worst-case measure that attains the infimum in (P_1). Thus the optimal indemnity function is $Y^* = (X - R_{h^*})1_A + X 1_{S \backslash A}$, where $R_{h^*} \in \mathcal{I}$ satisfies $\int_A R_{h^*}(s)\, dQ(s) = \widetilde{\Pi}_0$. A possible choice for R^*_h is shown in Proposition 4.1.

4.2. **With the No-Sabotage Condition.** The following result characterizes the optimal indemnity Y^*, when the no-sabotage condition is enforced.

Proposition 4.3. *Suppose that the utility function* u *satisfies Assumption 1 and let* $\mathcal{F} = \hat{\mathcal{I}}$ *as defined in eq. (2.3) be the set of admissible indemnity functions. Then there exists* $P^* \in \mathcal{C}$ *such that the optimal solution of Problem* (P_1) *is* $\hat{Y}^* \in \hat{\mathcal{I}}$*, Q-a.s. and is of the form*

$$\hat{Y}^* = (X - \hat{R}^*)1_A + X 1_{S \backslash A},$$

where

(a) $A \in \Sigma$ *is such that* $P^* = P^*_{ac} + P^*_s$*, with* $P^*_s(A) = Q(S \backslash A) = 0$*;*

(b) $h^* : S \to [0, \infty)$ *is such that* $h^* = dP^*_{ac}/dQ$*;*

(c) $\xi^* : \mathbb{R}_+ \to \mathbb{R}_+$ *is a Borel measurable function such that* $h^* = \xi^* \circ X$*;*

(d) $\hat{R}^* = \hat{r}^* \circ X$*, where* $\hat{r}^*(x) = \int_0^x (\hat{r}^*)'(t)\, dt$*,* $\forall x \in [0, M]$ *and*

$$(\hat{r}^*)'(t) = \begin{cases} 0, & \text{if } \int_{[t,M] \cap X(A)} (u'(W_0 - \hat{r}^*(x) - \Pi_0)\xi^*(x) - \lambda^*)\, dQ \circ X^{-1}(x) > 0, \\ \kappa(t), & \text{if } \int_{[t,M] \cap X(A)} (u'(W_0 - \hat{r}^*(x) - \Pi_0)\xi^*(x) - \lambda^*)\, dQ \circ X^{-1}(x) = 0, \\ 1, & \text{if } \int_{[t,M] \cap X(A)} (u'(W_0 - \hat{r}^*(x) - \Pi_0)\xi^*(x) - \lambda^*)\, dQ \circ X^{-1}(x) < 0, \end{cases}$$

for some Lebesgue measurable and $[0, 1]$*-valued function* κ*;*

(e) $\lambda^* \in \mathbb{R}_+$ *is such that* $(1+\rho)\mathbb{E}_Q[\hat{Y}^*] = \Pi_0$.

Proposition 4.3 is a particular case of Theorem 3.4 when the insurer's utility v is linear.

4.3. **Example.** The following example analyzes the structure of the optimal Y^* and $\hat{Y}^*$ in Propositions (4.1) and (4.3), respectively, when all probability measures P in $\mathcal{C}$ are absolutely continuous with respect to Q, with a particular structure of the Radon-Nikodým derivatives. This is done both with and without the no-sabotage condition. Specifically, we assume that each $P \in \mathcal{C}$ is such that $P \ll Q$ with

$$\frac{dP}{dQ} = \frac{w(X)}{\int w(X)dQ}, \tag{4.2}$$

for some nonnegative and increasing weight function w satisfying $\int w(X)dQ > 0$. Such measure transformations have a long tradition in insurance pricing, dating back to the Esscher transform (e.g., Bühlmann (1980)), in which the function w takes the form $w(x) = e^{bx}$, for a given $b \in (0, +\infty)$. More generally, Furman and Zitikis (2008a, 2008b, 2009) discuss the general class of weighted premium principles where pricing is done via measure transformations as in eq. (4.2).

Suppose that the utility function u satisfies Assumption 1, and assume that insurer's probability measure Q has a continuous CDF over $[0, M]$.

Example 4.4. Let the DM's ambiguity set $\mathcal{C}$ be defined as follows:

$$\mathcal{C}_{\mathcal{W}} := \left\{ P \in ca_1^+(\Sigma) : \frac{dP}{dQ} = \frac{w(X)}{\int w(X)dQ},\ w \in \mathcal{W} \right\},$$

where $\mathcal{W} \subset L^1\left(\mathbb{R}, \mathcal{B}(\mathbb{R}), Q \circ X^{-1}\right)$ is a collection of nonnegative increasing weight functions, such that $\int w(X)dQ > 0$ for all $w \in \mathcal{W}$. Appendix B provides conditions under which the set $\mathcal{C}_{\mathcal{W}}$ is convex and weak*-compact.

First we analyze the case when the feasible set of indemnities is $\mathcal{F} = \mathcal{I}$, as defined in eq. (2.2). By definition of $\mathcal{C}_{\mathcal{W}}$, any optimal P^* is absolutely continuous with respect to Q. Moreover, by monotonicity of $h^* = dP^*/dQ = \xi^*(X)$, there exists some $a \geqslant 0$ such that $\xi^*(x) = 0$, for $x \in [0, a]$ and $\xi^*(x) > 0$, for $x > a$, i.e., the set A_{h^*} in Proposition 4.1 is precisely $A_{h^*} = X^{-1}([0, a])$.

If $(1+\rho)\ \mathbb{E}_Q[X 1_{S \setminus A_{h^*}}] = (1+\rho) \displaystyle\int_a^M x\, dQ \circ X^{-1}(x) > \Pi_0$, the optimal indemnity $Y^* = I^*(X)$ in (4.1) is such that

$$I^*(x) = \max\left[0, \min\left(x, x - W_0 + \Pi_0 + (u')^{-1}\left(\frac{\lambda^*}{\xi^*}\right)\right)\right] 1_{[a,M]},$$

where $\lambda^* \in \mathbb{R}_+$ is such that $(1+\rho)\mathbb{E}_Q[Y^*] = \Pi_0$.

If $(1+\rho)\ \mathbb{E}_Q[X 1_{S \setminus A_{h^*}}] = (1+\rho) \displaystyle\int_a^M x\, dQ \circ X^{-1}(x) \leqslant \Pi_0$, then (4.1) becomes

$$I^*(x) = \begin{cases} (1-c)x, & \text{if } x \leqslant a, \\ x, & \text{if } x > a, \end{cases}$$

where the constant $c \in (0, 1]$ is chosen as in Proposition 4.1.

Next, let $\mathcal{F} = \hat{\mathcal{I}}$, as defined in eq. (2.3). Following the setting of Proposition 4.3, the utility function u need not to be strictly concave, but only concave. According to Proposition 4.3, the optimal retention can be equivalently written as

$$(\hat{r}^*)'(t) = \begin{cases} 0, & \text{if } \int_t^M (\lambda^* - u'(W_0 - \hat{r}^*(x) - \Pi_0)\xi^*(x))\, dF_{X,Q}(x) < 0, \\ \kappa(t), & \text{if } \int_t^M (\lambda^* - u'(W_0 - \hat{r}^*(x) - \Pi_0)\xi^*(x))\, dF_{X,Q}(x) = 0, \\ 1, & \text{if } \int_t^M (\lambda^* - u'(W_0 - \hat{r}^*(x) - \Pi_0)\xi^*(x))\, dF_{X,Q}(x) > 0. \end{cases}$$

Observe that the function $\varphi : [0, M] \to \mathbb{R}$, $\varphi(x) := -u'(W_0 - \hat{r}^*(x) - \Pi_0)\xi^*(x)$ is a continuous, decreasing function. We distinguish the following cases:

Case 1. If $-\lambda^* < \varphi(M)$, then $\int_t^M (\varphi(x) + \lambda^*)\, dF_{X,Q}(x) > 0$, for all $t \in [0, M]$, and thus $(\hat{r}^*)' \equiv 1$.

Case 2. If $-\lambda^* > \varphi(0)$, then $\int_t^M (\varphi(x) + \lambda^*)\, dF_{X,Q}(x) < 0$, for all $t \in [0, M]$, and thus $(\hat{r}^*)' \equiv 0$.

Case 3. If $-\lambda^* \in [\varphi(M), \varphi(0)]$, then there exists some $d \in (0, M)$ such that $\varphi(x) + \lambda^* \geqslant 0$, for all $x \leqslant d$ and $\varphi(x) + \lambda^* < 0$, for all $x > d$. This implies that for all $t > d$, $\int_t^M (\varphi(x) + \lambda^*)\, dF_{X,Q}(x) < 0$, and thus $(\hat{r}^*)'(t) = 0$, for $t \in (d, M]$. Moreover, for $t_1 < t_2 \leqslant d$, it holds that

$$\int_{t_1}^{d} (\varphi(x) + \lambda^*)\, dF_{X,Q}(x) + \int_{d}^{M} (\varphi(x) + \lambda^*)\, dF_{X,Q}(x) \geqslant \int_{t_2}^{d} (\varphi(x) + \lambda^*)\, dF_{X,Q}(x) + \int_{d}^{M} (\varphi(x) + \lambda^*)\, dF_{X,Q}(x).$$

Therefore, there exists some $d^* \geqslant 0$ such that for all $t \geqslant d^*$,

$$\int_{t}^{M} (\varphi(x) + \lambda^*)\, dF_{X,Q}(x) < 0.$$

Hence, $(\hat{r}^*)'(t) = 1$ for all $t < d^*$, and $(\hat{r}^*)'(t) = 0$, for all $t > d^*$. In this case, $\hat{r}^*(t) = \min(t, d^*)$ and thus $\hat{I}^*(t) = \max(t - d^*, 0)$.

4.4. **Numerical Example.** In this section, we examine the structure of the saddle point $(\hat{Y}^*, P^*)$ in the setting of Problem (P_1), when the insurer is risk-neutral, the admissible set of indemnities is $\mathcal{F} = \hat{\mathcal{I}}$ as defined in eq. (2.3) and the DM's ambiguity set $\mathcal{C}$ is a Wasserstein ball around Q, the insurer's belief. Similar to Section 3.3, we assume S is a Polish space and X is a nonnegative random variable, with unknown true distribution. However, compared to the approach in Example 3.6 of selecting the ambiguity radius δ, we follow this time a data-driven approach to obtain δ. It consists of estimating δ either by evaluating the discrepancy between the empirical model $\widehat{P}_n$ and the calibrated model, or using measure concentration inequalities to target a certain confidence level $\beta \in (0, 1)$, i.e., $\mathbb{P}(\mathrm{d}(P^*, \widehat{P}_n) \leqslant \delta) \geqslant 1 - \beta$ (see Esfahani and Kuhn (2018, Theorem 3.4 and the discussion afterwards), Blanchet et al. (2020, Section 5.1)). We investigate the former method in Example 4.5, when the ambiguity set is constructed using the Wasserstein metric.

Example 4.5 (Wasserstein ambiguity set)**.** In this example, the DM's ambiguity about the realizations of X is characterized by the ambiguity set $\mathcal{C}_\delta$ given by

$$\mathcal{C}_\delta^{\mathcal{W}_1} := \{P \in ca_1^+(\Sigma) : \mathcal{W}_1\left(P \circ X^{-1}, Q \circ X^{-1}\right) \leqslant \delta\},$$

where $\mathcal{W}_1$ is the Wasserstein distance on $\mathbb{R}$, with the L_1-norm being the underlying metric (e.g., Vallender (1974)):

$$\mathcal{W}_1\left(P \circ X^{-1}, Q \circ X^{-1}\right) := \int_{\mathbb{R}} \left|F_{X,P}(x) - F_{X,Q}(x)\right| dx = \int_0^1 \left|F_{X,P}^{-1}(t) - F_{X,Q}^{-1}(t)\right| dt.$$

The Wasserstein distance is a metric satisfying $\mathcal{W}_1\left(P \circ X^{-1}, Q \circ X^{-1}\right) = 0$ if and only if $P \circ X^{-1} = Q \circ X^{-1}$ (e.g., Villani (2008, Ex 6.3 p. 94). This induces a metric on $\mathcal{C}_\delta^{\mathcal{W}_1}$, such that $\mathcal{C}_\delta^{\mathcal{W}_1}$ is convex and weak*-compact. See Villani (2008) for further properties of the Wasserstein distance. By a slight abuse of notation, we write $\mathcal{W}_1(P, Q)$ instead of $\mathcal{W}_1\left(P \circ X^{-1}, Q \circ X^{-1}\right)$.

Let Q be the insurer's belief regarding the loss X. In this example, we assume that $F_{X,Q}$ is a truncated Generalized Pareto distribution with an upper bound M, with shape and scale parameters 0.3 and 5, respectively. Next, we simulate from the distribution $F_{X,Q}$, and obtain the empirical distribution. We then construct a piecewise linear approximation $F_{X,\widehat{Q}}$ of this empirical distribution, with given knots $\{x_1, \ldots, x_n\}$, where the partition $0 = x_1 < \cdots < x_n = M$ is chosen arbitrarily, but kept fixed all throughout. That is, $F_{X,\widehat{Q}}$ is given by a system $\left\{[x_i, x_{i+1}], [F_{X,\widehat{Q}}(x_i), F_{X,\widehat{Q}}(x_{i+1})]\right\}_{i=1}^{n-1}$. Note that by construction, $F_{X,\widehat{Q}}(x_n) = 1$. The corresponding density $\widehat{\mathbf{q}} = [\widehat{q}_1, \ldots, \widehat{q}_{n-1}]^\top$ is piecewise constant on each interval

$[x_i, x_{i+1}]$, for $i = 1, \ldots, n-1$. More precisely, $\widehat{q}_i$ is the slope of the line passing through the points $(x_i, F_{X,\widehat{Q}}(x_i))$ and $(x_{i+1}, F_{X,\widehat{Q}}(x_{i+1}))$, i.e., $\widehat{q}_i = \dfrac{F_{X,\widehat{Q}}(x_{i+1}) - F_{X,\widehat{Q}}(x_i)}{x_{i+1} - x_i}$, for $i = 1, \ldots, n-1$.

This representation of $\widehat{Q}$ allows us to compute the Wasserstein distance between $F_{X,P}$ and $F_{X,\widehat{Q}}$, and thus to characterize the alternative distributions via the system $\{F_{X,P}(x_1), \ldots, F_{X,P}(x_n)\}$, for the same segments $[x_i, x_{i+1}]$, $i = 1, \ldots, n-1$. For two such distributions, the Wasserstein distance is the sum of the areas of the trapezoids with corners $\{F_{X,\widehat{Q}}(x_i), F_{X,\widehat{Q}}(x_{i+1}), F_{X,P}(x_i), F_{X,P}(x_{i+1})\}$ formed by $F_{X,P}$ and $F_{X,\widehat{Q}}$ (see the shaded area in Figure 5), i.e.,

$$\mathcal{W}_1(P, \widehat{Q}) = \frac{1}{2} \sum_{i=1}^{n-1} (x_{i+1} - x_i)\, \phi\Big(F_{X,P}(x_i) - F_{X,\widehat{Q}}(x_i), F_{X,P}(x_{i+1}) - F_{X,\widehat{Q}}(x_{i+1})\Big), \tag{4.3}$$

where the function $\phi : [-1,1]^2 \to \mathbb{R}_+$ defined below is convex in each component (e.g., Pflug et al. (2017)):

$$\phi(a, b) = \begin{cases} |a| + |b|, & \text{if } ab \geqslant 0, \\ \dfrac{a^2 + b^2}{|a| + |b|}, & \text{otherwise.} \end{cases}$$

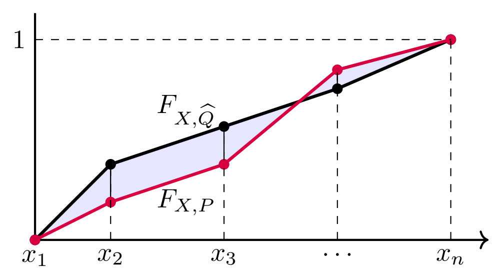

FIGURE 5. Computation of Wasserstein distance between piecewise linear probability distributions $F_{X,\widehat{Q}}$ and $F_{X,P}$.

The alternative measure P is represented by an $(n-1)$-dimensional vector $\mathbf{p} = [p_1, \ldots, p_{n-1}]^\top$, where $p_i \in [0,1]$ is the constant forming the piecewise constant density of $F_{X,P}$. More precisely, the alternative CDF $F_{X,P}$ will be linear on each interval $[x_i, x_{i+1}]$, and will differ from $F_{X,\widehat{Q}}$ only in the cumulative probabilities $F_{X,P}(x_i)$. Thus, p_i will be the slope of the line passing through the points $(x_i, F_{X,P}(x_i))$ and $(x_{i+1}, F_{X,P}(x_{i+1}))$. The representation of $F_{X,P}$ is shown in Figure 5. Therefore, the variable $\mathbf{p}$ must satisfy $\mathbf{p}^\top \mathbf{A}_{n-1}\mathbf{x} = 1$, where $\mathbf{A}_{n-1} \in \mathbb{R}^{(n-1)\times n}$ is defined in eq. (3.16). Using the matrix $\mathbf{A}_i$, for $i = 1, \ldots, n-1$, $F_{X,P}$ can also be represented via $F_{X,P}(x_i) = \mathbf{p}^\top \mathbf{A}_i \mathbf{x}$.

Next, to identify the optimal $\hat{Y}^*$, we follow the equivalent formulation of Problem (P_1) and describe the decision variable in terms of retention function $\hat{R}$. First, note that since the feasible set of indemnity functions is assumed to be $\mathcal{F} = \hat{\mathcal{I}}$, it follows that $\hat{R} = \hat{r} \circ X$ where $\hat{r}$ is a 1-Lipschitz, and hence continuous function. We further assume that $\hat{r}$ is linear between the segments $[x_i, x_{i+1}]$, and thus piecewise linear and of the form $\hat{r}(x) = a_i x + b_i$, for $x_i \leqslant x \leqslant x_{i+1}$, $i = 1, \ldots, n-1$. Since $\hat{r} \in \hat{\mathcal{I}}$, it follows that $a_i \in [0,1]$ and $b_i \in [-x_i, 0]$, for $i = 1, \ldots, n-1$.

The error introduced by solving (P_1) in terms of $\widehat{Q}$ instead of Q can be used to estimate the ambiguity radius δ. The estimator $\delta = \delta_n$ depends on the number of piecewise linear segments, and thus it is informed by the data. In particular, we propose to approximate δ_n as

$$\delta_n := \big|\mathbb{E}_{\widehat{Q}}[X] - \mathbb{E}_Q[X]\big| \leqslant \mathcal{W}_1(\widehat{Q}, Q).$$

The inequality becomes an equality if $F_{X,\widehat{Q}}$ dominates $F_{X,Q}$ in the first stochastic order.

With the above representations for $\mathbf{p}$ and $\hat{\mathbf{r}}$, Problem (P_1) is then approximated by the following problem:

$$(\mathrm{P}_1^n) \qquad F\begin{cases} \max\limits_{\substack{\mathbf{a}\in[0,1]^n,\\ \mathbf{b}\in[-\mathbf{x},0]}} \min\limits_{\mathbf{p}\in[0,1]^{n-1}} \displaystyle\sum_{i=1}^{n-1}\int_{x_i}^{x_{i+1}} u(W_0 - x + a_i x + b_i - \Pi_0)dx\, p_i \\ \text{s.t.} \quad 0 \leqslant a_i x_i + b_i \leqslant x_i,\ i = 1,\dots,n-1, \\ \quad a_i x_{i+1} + b_i = a_{i+1}x_{i+1} + b_{i+1},\ i = 1,\dots,n-2, \\ \quad \displaystyle\sum_{i=1}^{n-1}\int_{x_i}^{x_{i+1}} (x - a_i x + b_i)dx\, q_i \leqslant \frac{\Pi_0}{1+\rho}, \\ \quad \mathcal{W}_1(P,\widehat{Q}) \leqslant \delta_n,\ F_{X,P}(x_i) = \mathbf{p}^\top \mathbf{A}_i \mathbf{x},\ i = 1,\dots,n, \\ \quad \mathbf{p}^\top \mathbf{A}_{n-1}\mathbf{x} = 1, \end{cases}$$

where $\mathcal{W}_1(P,\widehat{Q})$ is computed as in eq. (4.3) and $\mathbf{a} = [a_1,\dots,a_{n-1}]^\top$ and $\mathbf{b} = [b_1,\dots,b_{n-1}]^\top$. The first two constraints in (P_1^n) specify that the retention function $\hat{r}$ is continuous and linear between the segments $[x_i, x_{i+1}]$, $i = 1,\dots,n-1$. The objective function in (P_1^n) is concave in $\mathbf{a}$ and $\mathbf{b}$ and linear in $\mathbf{p}$, while the constraints are convex in $\mathbf{p}$ and linear in $\mathbf{a}$ and $\mathbf{b}$. Similar to Example 3.6, Problem (P_1^n) is solved in a step-wise manner, by splitting the initial problem into an inner and outer problem.

For the implementation, we resume the input for (P_1^n): the DM's initial wealth is $W_0 = 250$ and the utility is $u(x) = (1-\exp(-\gamma x))/\gamma$, for $\gamma = 0.03$, while the premium $\Pi_0 = 4$ and the safety loading is $\rho = 0.2$. For $n = 200$, we simulate from the distribution $F_{X,Q}$, and construct the piecewise linear approximation $F_{X,\widehat{Q}}$ of the empirical distribution on the partition $0 = x_1 < x_2 < \cdots < x_n = M = W_0 - \Pi_0 = 246$. The CDF $F_{X,\widehat{Q}}$ will play the role of the baseline distribution in Problem (P_1^n). Finally, the ambiguity radius δ_n is estimated to be approximately 0.3. Figure 6 shows one of the saddle points of Problem (P_1^n): the optimal $\hat{\mathbf{y}}^*$ is piecewise-linear, while the corresponding $F_{X,P*}$ dominates $F_{X,\widehat{Q}}$ in the first stochastic order.

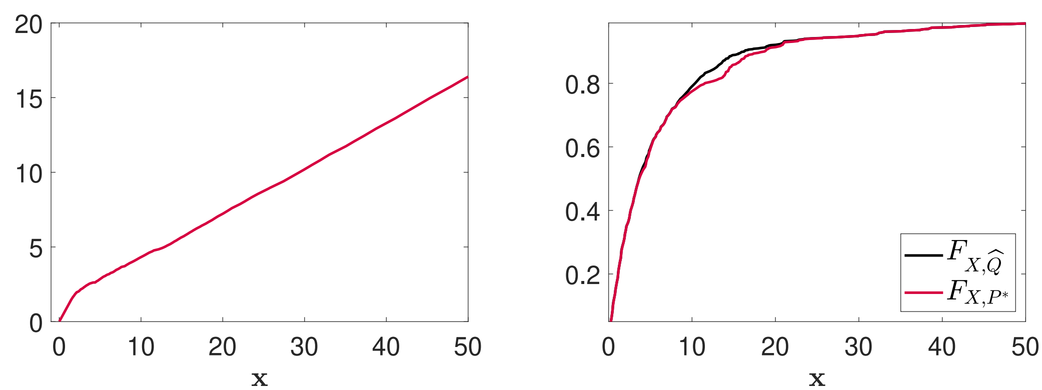

FIGURE 6. Left: the optimal indemnity $\hat{\mathbf{y}}^* = \mathbf{x} - \hat{\mathbf{r}}^*$ as function of $\mathbf{x}$. Right: the DM's optimal distribution $F_{X,P*}$ (red) compared to insurers' belief $F_{X,\widehat{Q}}$ (black). Here, for sake of presentation, we only display values of $\mathbf{x}$ below 50.

Next, we study a problem related to Problem (P_1^n), in which retention function $\mathbf{r}$ is required only to be bounded by $\mathbf{x}$, i.e. $\mathbf{r} \in \mathcal{I}$, where $\mathcal{F} = \mathcal{I}$ as defined in eq. (2.2). It implies that $\mathbf{a} \in \mathbb{R}^n$ in (P_1^n). In Figure 7 (left) we display the difference between the optimal retention functions of Problem (P_1^n), for the sets $\mathcal{I}$ and $\hat{\mathcal{I}}$, respectively, and we also display a zoomed-in perspective for small values of $\mathbf{x}$. Figure 7 (right) provides the corresponding indemnities $\mathbf{y}^*$ and $\hat{\mathbf{y}}^*$, respectively. In the absence of the no-sabotage condition (the blue lines in Figure 7), the indemnity $\mathbf{y}^*$ can be decreasing with respect to the loss $\mathbf{x}$ on some parts of its domain (see the blue line in Figure 7 (left)).

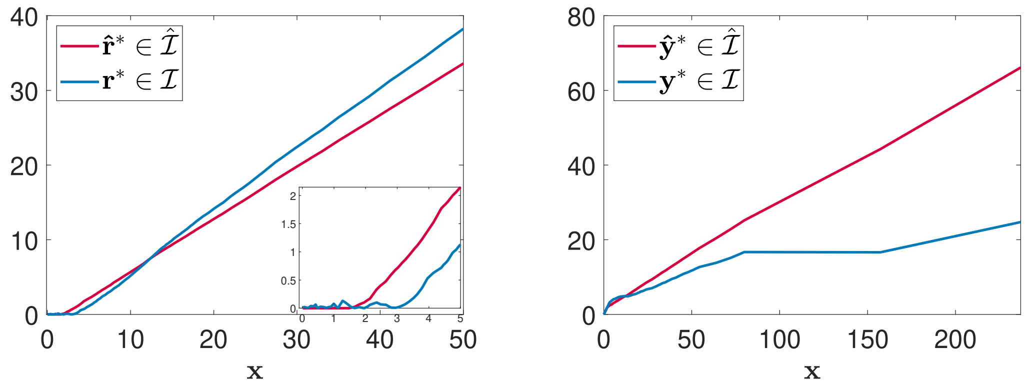

FIGURE 7. Left: optimal retention functions $\mathbf{r}^*$ and $\hat{\mathbf{r}}^*$ for Problem (P_1^n) when the feasibility sets are $\mathcal{I}$ and $\hat{\mathcal{I}}$, respectively, with a zoomed-in perspective for small values of the underlying loss $\mathbf{x}$. Right: the corresponding indemnities $\mathbf{y}^* = \mathbf{x} - \mathbf{r}^*$ and $\hat{\mathbf{y}}^* = \mathbf{x} - \hat{\mathbf{r}}^*$.

5. Conclusion

The impact of ambiguity on insurance markets in general, and insurance contracting in particular, is by now well-documented. One of the most popular and intuitive ways to model sensitivity of preferences to ambiguity is the Maxmin-Expected Utility (MEU) model of Gilboa and Schmeidler (1989). Nonetheless, to the best of our knowledge, none of the theoretical studies of risk sharing in insurance markets in the presence of ambiguity have examined the case in which the decision maker (DM) is an MEU-maximizer. This paper fills this void. Specifically, we extend the classical setup and results in two ways: (i) the DM is endowed with MEU preferences; and (ii) the insurer is an Expected-Utility-maximizer who is not necessarily risk-neutral (that is, the premium principle is not necessarily an expected-value premium principle). The main objective of this paper is then to determine the shape of the optimal insurance indemnity in that case.

We characterize optimal indemnity functions both with and without the customary *ex ante no-sabotage* requirement on feasible indemnities, and for both concave and linear utility functions for the two agents. The no-sabotage condition is shown to play a key role in determining the shape of optimal indemnity functions. An equally important factor in characterizing optimal indemnities is the singularity in beliefs between the two agents. We subsequently examine several illustrative examples, and we provide numerical studies for the case of a Wasserstein and a Rényi ambiguity set. Specifically, we provide a successive convex programming algorithm to compute optimal insurance indemnities in a discretized framework. The Wasserstein and Rényi distances are two popular metrics to construct probability ambiguity sets. We show in numerical examples that a larger ambiguity set yields a lower certainty equivalent of final wealth, but increases the willingness-to-pay for insurance. As a by-product of our analysis, we provide a comprehensive and unifying treatment of optimal insurance design under subjective expected-utility theory in the presence of belief heterogeneity, thereby extending many result in the related literature.

An interesting direction for future research would be to give the policyholder the possibility of partially hedging her loss exposure, through some hedging instrument that would act as an uninsurable background risk. Specifically, the policyholder initially faces an insurable loss random variable X and is able to partially hedge her loss exposure through another random variable Z. For a given hedging investment decision, the latter can be interpreted as a background risk, which might be correlated with X. Optimal insurance design in the presence of a background risk has been examined in expected-utility theory by Dana and Scarsini (2007) (without the no-sabotage condition) and Chi and Wei (2020) (with the no-sabotage condition).

Appendix A. A Useful Tool

A.1. Sion's Minimax Theorem.

Theorem A.1 (Sion (1958))**.** *Let $\mathbb{X}$ and $\mathbb{Y}$ be convex, compact spaces, and f a function on $\mathbb{X} \times \mathbb{Y}$. If $x \mapsto f(x, y)$ is quasi-convex and lower semi-continuous, for all $y \in \mathbb{Y}$ and $y \mapsto f(x, y)$ is quasi-concave and upper semi-continuous, for all $x \in \mathbb{X}$, then*

$$\min_{x \in \mathbb{X}} \max_{y \in \mathbb{Y}} f(x, y) = \max_{y \in \mathbb{Y}} \min_{x \in \mathbb{X}} f(x, y).$$

If the assumption about the compactness of $\mathbb{X}$ is dropped, then

$$\inf_{x \in \mathbb{X}} \max_{y \in \mathbb{Y}} f(x, y) = \max_{y \in \mathbb{Y}} \inf_{x \in \mathbb{X}} f(x, y).$$

If the assumption about the compactness of $\mathbb{Y}$ is dropped, then

$$\min_{x \in \mathbb{X}} \sup_{y \in \mathbb{Y}} f(x, y) = \sup_{y \in \mathbb{Y}} \min_{x \in \mathbb{X}} f(x, y).$$

Appendix B. Convexity and Compactness of $\mathcal{C}_{\mathcal{W}}$ in Example 4.4

Lemma B.1. *For a fixed $Q \in ca_1^+(\Sigma)$, let $\mathcal{C}_{\mathcal{W}}$ be the set defined as follows:*

$$\mathcal{C}_{\mathcal{W}} := \left\{ P \in ca_1^+(\Sigma) : \frac{dP}{dQ} = \frac{w(X)}{\int w(X) dQ},\ w \in \mathcal{W} \right\}, \tag{B.1}$$

where $\mathcal{W} \subset L^1\left(\mathbb{R}, \mathcal{B}(\mathbb{R}), Q \circ X^{-1}\right)$ is a collection of nonnegative increasing weight functions, such that $\int w(X) dQ > 0$, for all $w \in \mathcal{W}$. Then the following hold:

(i) If $\mathcal{W}$ is a convex cone, then $\mathcal{C}_{\mathcal{W}}$ is convex.

(ii) If $\mathcal{C}_{\mathcal{W}}$ is uniformly absolutely continuous with respect to some $\mu \in ca^+(\Sigma)$, then $\mathcal{C}_{\mathcal{W}}$ is weak-compact.*

Proof. (i) is easy to verify. To show (ii), first note that $\mathcal{C}_{\mathcal{W}}$ is norm-bounded. Since $\mathcal{C}_{\mathcal{W}}$ is also uniformly absolutely continuous with respect to $\mu \in ca^+(\Sigma)$, it follows from Dunford and Schwartz (1958, Theorem IV.9.2) that $\mathcal{C}_{\mathcal{W}}$ is weakly sequentially compact, and hence weak*-compact, by Maccheroni and Marinacci (2001, Theorem 1). □

Remark B.2. In Lemma B.1, if $\mathcal{C}_{\mathcal{W}}$ is countable, that is, is of the form

$$\left\{ P_n \in ca_1^+(\Sigma) : n \in \mathbb{N},\ \frac{dP_n}{dQ} = \frac{w(X)}{\int w(X) dQ},\ w \in \mathcal{W} \right\},$$

and if $\lim_{n \to +\infty} P_n(A)$ exists for each $A \in \Sigma$, then the requirement of uniform absolute continuity of $\mathcal{C}_{\mathcal{W}}$ is superfluous by the Vitali-Hahn-Saks Theorem (Dunford & Schwartz, 1958, Theorem III.7.2).

Proposition B.3. *If $\mathcal{W}$ is order bounded in the Banach lattice $L^1\left(\mathbb{R}, \mathcal{B}(\mathbb{R}), Q \circ X^{-1}\right)$, with a constant upper bound and a nonnegative lower bound having nonzero L^1-norm, then $\mathcal{C}_{\mathcal{W}}$ is uniformly absolutely continuous with respect to Q.*

Proof. Suppose that $\mathcal{W}$ is order bounded in $L^1\left(\mathbb{R}, \mathcal{B}(\mathbb{R}), Q \circ X^{-1}\right)$, with a constant upper bound and a nonnegative lower bound having nonzero L^1-norm. Then there exists $f \in L^1_+\left(\mathbb{R}, \mathcal{B}(\mathbb{R}), Q \circ X^{-1}\right)$ and $M \in \mathbb{R}^+$, such that $M < +\infty$, $\|f\|_1 = \int f \, dQ \circ X^{-1} > 0$, and $f \leqslant w \leqslant M$, for each $w \in \mathcal{W}$. Consequently, for each $P \in \mathcal{C}_{\mathcal{W}}$,

$$\frac{dP}{dQ} \leqslant \frac{M}{\|f\|_1} < +\infty.$$

Hence, for each $P \in \mathcal{C}_{\mathcal{W}}$ and each $A \in \Sigma$,

$$P(A) = \int_A dP \leqslant \frac{M}{\|f\|_1} Q(A).$$

Consequently, for each $\varepsilon > 0$, letting $\delta := \frac{\|f\|_1}{M} \varepsilon > 0$, it follows that for each $A \in \Sigma$ and each $P \in \mathcal{C}_{\mathcal{W}}$,

$$Q(A) < \delta \Longrightarrow P(A) < \frac{M}{\|f\|_1} \delta = \varepsilon.$$

Hence, $\mathcal{C}_{\mathcal{W}}$ is uniformly absolutely continuous with respect to Q. □

Appendix C. Convergence of the Algorithm in Examples 3.6 and 4.5

The convergence of the SCP algorithm in Examples 3.6 and 4.5 is proven in Pflug and Picher (2014, Proposition B.6).

Proposition C.1. *Every cluster point $\hat{Y}^*$ of the iteration:*

$$\hat{Y}^{(m+1)} \in \underset{\hat{Y} \in \hat{\mathcal{I}}_0}{\arg\max} \min_{P^{(i)} \in \mathcal{P}^{(m)}} u(\hat{Y}, P^{(i)}), \tag{C.1}$$

$$P^{(m+1)} \in \underset{P \in \mathcal{C}_\delta}{\arg\min}\, u(\hat{Y}^{(m+1)}, P) \tag{C.2}$$

is a solution of Problem (P).

ENDNOTES

[1]We refer to Carlier and Dana (2003) for a discussion of various notions of *ex ante* admissible contracts.

[2]For instance, under the no-sabotage condition, feasible indemnities are Lipschitz-continuous and hence absolutely continuous. Optimal indemnities are then characterized in implicit form through their derivative, that is, the so-called "marginal indemnification function" (MIF), as in Assa (2015). The vast majority of the literature on optimal insurance with the no-sabotage condition uses the MIF semi-implicit characterization of the optimal solution (see Xu et al. (2018) or Zhuang et al. (2016), for example).

[3]All of this paper's results can be derived for any participation constraint of the form $\mathbb{E}_Q[v(W_0^{\text{Ins}} - (1+\rho)Y + \Pi_0)] \geqslant k$, with $k \leqslant v(W_0^{\text{Ins}} + \Pi_0)$. To maintain a direct economic interpretation, we choose throughout the paper $k = v(W_0^{\text{Ins}})$, i.e., the insurer's reservation utility.

[4]The limit conditions on u and v are the customary Inada (1963) conditions, often encountered in the literature.

References

Alary, D., Gollier, C., & Treich, N. (2013). The effect of ambiguity aversion on insurance and self-protection. *Economic Journal*, *123*(573), 1188-1202.

Aliprantis, C., & Border, K. (2006). *Infinite dimensional analysis - 3rd edition.* Springer-Verlag.

Allais, M. (1953). Le Comportement de l'Homme Rationnel Devant le Risque: Critique des Axiomes et Postulats de l'École Américaine. *Econometrica*, *21*(4), 503-546.

Amarante, M. (2009). Foundations of Neo-Bayesian statistics. *Journal of Economic Theory*, *144*(5), 2146-2173.

Amarante, M. (2014). A characterization of exact non-atomic market games. *Journal of Mathematical Economics*, *54*(1), 59-62.

Arrow, K. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*, *53*(5), 941-973.

Arrow, K. (1971). *Essays in the theory of risk-bearing.* Chicago: Markham Publishing Company.

Artzner, P., Delbaen, F., Eber, J., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, *9*(3), 203-228.

Asimit, A., Bignozzi, V., Cheung, K., Hu, J., & Kim, E.-S. (2017). Robust and Pareto optimality of insurance contracts. *European Journal of Operational Research*, *262*(2), 720-732.

Assa, H. (2015). On optimal reinsurance policy with distortion risk measures and premiums. *Insurance: Mathematics and Economics*, *61*, 70–75.

Bell, D. (1982). Regret in decision making under uncertainty. *Operations Research*, *30*(5), 961-981.

Bell, D. (1985). Disappointment in decision making under uncertainty. *Operations Research*, *33*(1), 1-27.

Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). *Robust optimization.* Princeton University Press, Princeton, New Jersey.

Ben-Tal, A., & Nemirovski, A. (1998). Robust convex optimization. *Mathematics of Operations Research*, *23*(4), 769–805.

Bernard, C., He, X., Yan, J., & Zhou, X. (2015). Optimal insurance design under rank-dependent expected utility. *Mathematical Finance*, *25*(1), 154-186.

Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research*, *52*(1), 35-53.

Birghila, C., & Pflug, G. (2019). Optimal XL-insurance under Wasserstein-type ambiguity. *Insurance: Mathematics and Economics*, *88*, 30-43.

Blanchet, J., He, F., & Murthy, K. (2020). On distributionally robust extreme value analysis. *Extremes*, 1-31.

Boonen, T., & Ghossoub, M. (2019). On the existence of a representative reinsurer under heterogeneous beliefs. *Insurance: Mathematics and Economics*, *88*, 209-225.

Boonen, T., & Ghossoub, M. (2021). Optimal reinsurance with multiple reinsurers: Distortion risk measures, distortion premium principles, and heterogeneous beliefs. *Insurance: Mathematics and Economics*, *101*(1), 23-37.

Breuer, T., & Csiszár, I. (2016). Measuring distribution model risk. *Mathematical Finance*, *26*(2), 395-411.

Bühlmann, H. (1980). An economic premium principle. *ASTIN Bulletin*, *11*(1), 52-60.

Calafiore, G., & El Ghaoui, L. (2006). On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory and Applications*, *130*(1), 1-22.

Carlier, G., & Dana, R. (2003). Pareto efficient insurance contracts when the insurer's cost function is discontinuous. *Economic Theory*, *21*(4), 871-893.

Chateauneuf, A., Maccheroni, F., Marinacci, M., & Tallon, J. (2005). Monotone continuous multiple priors. *Economic Theory*, *26*(4), 973-982.

Cheung, K., Chong, W., & Lo, A. (2019). Budget-constrained optimal reinsurance design under coherent risk measures. *Scandinavian Actuarial Journal*, *2019*(9), 729–751.

Cheung, K., Chong, W., & Yam, S. (2015). The optimal insurance under disappointment theories. *Insurance: Mathematics and Economics*, *64*, 77–90.

Chi, Y. (2019). On the optimality of a straight deductible under belief heterogeneity. *ASTIN Bulletin*, *49*(1), 243-262.

Chi, Y., & Wei, W. (2020). Optimal insurance with background risk: An analysis of general dependence structures. *Finance and Stochastics*, *24*(4), 903–937.

Chi, Y., & Zhuang, S. (2020). *Regret-based optimal insurance design.* (Mimeo)

Dana, R., & Scarsini, M. (2007). Optimal risk sharing with background risk. *Journal of Economic Theory*, *133*(1), 152-176.

Delage, E., & Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, *58*(3), 595-612.

Delbaen, F. (2002). Coherent risk measures on general probability spaces. *Advances in finance and stochastics. Essays in Honour of Dieter Sondermann*, 1-37.

Dunford, N., & Schwartz, J. (1958). *Linear operators, part 1: General theory.* Wiley-Interscience.

Ellsberg, D. (1961). Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics*, *75*(4), 643-669.

Esfahani, P. M., & Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, *171*(1-2), 115-166.

Furman, E., & Zitikis, R. (2008a). Weighted premium calculation principles. *Insurance: Mathematics and Economics*, *42*(1), 459-465.

Furman, E., & Zitikis, R. (2008b). Weighted risk capital allocations. *Insurance: Mathematics and Economics*, *43*(2), 263-269.

Furman, E., & Zitikis, R. (2009). Weighted pricing functionals with applications to insurance. *North American Actuarial Journal*, *13*(4), 483-496.

Ghirardato, P., Maccheroni, F., & Marinacci, M. (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, *118*(2), 133-173.

Ghossoub, M. (2019a). Budget-constrained optimal insurance with belief heterogeneity. *Insurance: Mathematics and Economics*, *89*, 79-91.

Ghossoub, M. (2019b). Optimal insurance under rank-dependent expected utility. *Insurance: Mathematics and Economics*, *87*, 51-66.

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*, *18*(2), 141-153.

Gollier, C. (2013). The economics of optimal insurance design. In *G. dionne (ed.), Handbook of Insurance – 2^{nd} ed. .* Springer.

Gollier, C. (2014). Optimal insurance design of ambiguous risks. *Economic Theory*, *57*(3), 555-576.

Gul, F. (1991). A theory of disappointment aversion. *Econometrica*, *59*(3), 667-686.

Huberman, G., Mayers, D., & Smith Jr., C. (1983). Optimal insurance policy indemnity schedules. *Bell Journal of Economics*, *14*(2), 415-426.

Inada, K. (1963). On a two-sector model of economic growth: Comments and a generalization. *The Review of Economic Studies*, *30*(2), 119-127.

Jeleva, M. (2000). Background risk, demand for insurance, and Choquet expected utility preferences. *Geneva Papers on Risk and Insurance Theory*, *25*(1), 7-28.

Jiang, W., Escobar-Anel, M., & Ren, J. (2020). Optimal insurance contracts under distortion risk measures with ambiguity aversion. *ASTIN Bulletin*, *50*(2), 619–646.

Klibanoff, P., Marinacci, M., & Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, *73*(6), 1849-1892.

Liese, F., & Vajda, I. (1987). *Convex statistical distances.* Teubner, Leipzig, Germany.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, 805-824.

Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies*, *53*(2), 271-282.

Maccheroni, F., & Marinacci, M. (2001). A Heine-Borel Theorem for $ba(\Sigma)$. *RISEC*, *48*(3), 353-362.

Machina, M. (2013). Non-expected utility and the robustness of the classical insurance paradigm. In *G. dionne (ed.), Handbook of Insurance Economics, Second Ed.* Springer, New York.

Mossin, J. (1968). Aspects of rational insurance purchasing. *Journal of Political Economy*, *76*(4), 553-568.

Pflug, G., & Picher, A. (2014). *Multistage stochastic optimization.* Springer.

Pflug, G., Timonina-Farkas, A., & Hochrainer-Stigler, S. (2017). Incorporating model uncertainty into optimal insurance contract design. *Insurance: Mathematics and Economics*, *73*, 68-74.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, *3*(4), 323-343.
Rényi, A. (1961). On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics.* University of California Press, Oakland, California.
Scarf, H. (1958). A min-max solution of an inventory problem. *Studies in The Mathematical Theory of Inventory and Production*.
Schlesinger, H. (2000). The theory of insurance demand. In *G. dionne (ed.), Handbook of Insurance.* Kluwer Academic publishers.
Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, *57*(3), 571-587.
Shapiro, A., & Kleywegt, A. (2002). Minimax analysis of stochastic problems. *Optimization Methods and Software*, *17*(3), 523–542.
Sion, M. (1958). On general minimax theorems. *Pacific Journal of Mathematics*, *8*(1), 171–176.
Thiele, A. (2008). Robust stochastic programming with uncertain probabilities. *IMA Journal of Management Mathematics*, *19*(3), 289-321.
Vallender, S. (1974). Calculation of the Wasserstein distance between probability distributions on the line. *Theory of Probability & Its Applications*, *18*(4), 784-786.
Van Erven, T., & Harremos, P. (2014). Rényi divergence and Kullback-Leibler divergence. *IEEE Transactions on Information Theory*, *60*(7), 3797-3820.
Villani, C. (2008). *Optimal transport: Old and new.* Springer Science & Business Media.
Wozabal, D. (2012). A framework for optimization under ambiguity. *Annals of Operations Research*, *193*(1), 21-47.
Xu, Z. (2021). *Pareto optimal moral-hazard-free insurance contracts in behavioral finance framework.* (Mimeo)
Xu, Z., Zhou, X., & Zhuang, S. (2018). Optimal insurance under rank-dependent utility and incentive compatibility. *Mathematical Finance*, *21*(4), 503-34.
Yaari, M. (1987). The dual theory of choice under risk. *Econometrica*, *55*(1), 95-115.
Yu, H., & Fang, Y. (2020). Optimal insurance under heterogeneous belief. *Journal of Financial Risk Management*, *9*, 179-189.
Žáčková, J. (1966). On minimax solutions of stochastic linear programming problems. *Časopis pro pěstování matematiky*, *91*(4), 423–430.
Zhuang, S., Weng, C., Tan, K. S., & Assa, H. (2016). Marginal indemnification function formulation for optimal reinsurance. *Insurance: Mathematics and Economics*, *67*, 65–76.
Zymler, S., Kuhn, D., & Rustem, B. (2013). Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, *137*(1-2), 167-198.

Corina Birghila: University of Waterloo – Department of Statistics and Actuarial Science – 200 University Ave. W. – Waterloo, ON, N2L 3G1 – Canada

Email address: corina.birghila@uwaterloo.ca

Tim J. Boonen: Amsterdam School of Economics – University of Amsterdam – Roetersstraat 11, 1018 WB – Amsterdam – The Netherlands

Email address: t.j.boonen@uva.nl

Mario Ghossoub: University of Waterloo – Department of Statistics and Actuarial Science – 200 University Ave. W. – Waterloo, ON, N2L 3G1 – Canada

Email address: mario.ghossoub@uwaterloo.ca