Revisiting optimal insurance design under smooth ambiguity aversion*

Yichun Chi† Richard Peter‡ Wei Wei§

August 10, 2022

Abstract

We analyze optimal insurance design for a risk- and ambiguity-averse policyholder who is uncertain about the distribution of losses and faces linear transaction costs. We use smooth ambiguity preferences, a flexible ambiguity structure, and focus on indemnity schedules that satisfy the principle of indemnity and the no-sabotage condition for incentive compatibility. We characterize optimal insurance contracts and find that the marginal indemnity is either zero or one except at critical points. We then provide a condition for a straight deductible to be optimal and show that this condition is satisfied under various stochastic ordering assumptions on the priors. We discuss specific ambiguity structures, some of which give rise to indemnities with multiple layers. We also derive comparative statics. Greater ambiguity aversion always raises insurance demand whereas greater ambiguity has indeterminate effects. For policyholders with relative ambiguity prudence between zero and two, greater ambiguity raises insurance demand.

Keywords: Insurance design · risk sharing · ambiguity · deductible · no-sabotage condition · comparative statics

JEL-Classification: D81 · D86 · G22

---

*Chi's work was supported by grants from the National Natural Science Foundation of China (Grant No. 11971505) and 111 Project of China (No. B17050). Wei acknowledges the financial support by the Research and Creative Activities Support grant (Grant No. AAC2253) from the University of Wisconsin-Milwaukee.

†Central University of Finance and Economics, China Institute for Actuarial Science, Beijing 102206, China. E-Mail: yichun@cufe.edu.cn.

‡University of Iowa, Department of Finance, Iowa City, IA 52240, USA. E-Mail: richard-peter@uiowa.edu.

§Corresponding author. University of Illinois at Urbana-Champaign, Department of Mathematics, Urbana, IL 61801, USA. E-Mail: weiw@illinois.edu.

1 Introduction

Risk sharing is critically important for the functioning of our modern economies. It is thus no surprise that risk sharing has received substantial attention in the field of development economics (e.g., Townsend, 1994; Fafchamps and Lund, 2003; Fafchamps and Gubert, 2007). In the absence of transaction costs, the mutuality principle prescribes that a risk allocation is only efficient if it completely washes out those risks that are fully diversifiable (Borch, 1962). Undiversifiable risk is shared across economic agents according to their degree of risk tolerance (Wilson, 1968; Rubinstein, 1974; Constantinides, 1982). If risk-neutral agents are present in the economy, they fully insure risk-averse agents to eliminate costly risk premiums. Insurance markets facilitate the mutualization of risks. However, the transfer of risk to an insurer does involve significant transaction costs. According to data from the National Association of Insurance Commissioners between 1990 and 2015, operating expenses account for one third of insurance premiums charged by US insurance companies.

When insurance involves such large deadweight losses, a complete transfer of risk is no longer optimal. Instead the policyholder prefers to retain some risk via clauses such as coinsurance, deductibles, or limits to economize on the insurance premium. This raises the question how to optimally design the insurance contract. The famous contributions by Arrow (1963, 1965, 1971, 1974) provide the cornerstone result in this literature. When the premium is proportional to the actuarial value of the contract, the optimal indemnity schedule takes the form of a straight deductible. The policyholder retains losses below the deductible. For losses above the deductible, she only pays the deductible and the insurer covers the remainder of the loss. Intuitively, the insurance indemnity is most valuable in states with high marginal utility. For a concave utility function, marginal utility is high when losses are large. A straight deductible prioritizes the indemnification of large losses over small losses.

In this paper, we revisit the question of optimal insurance design when the policyholder is uncertain about the loss distribution. We incorporate ambiguity aversion with the help of Klibanoff et al.'s (2005) smooth ambiguity model. We use a flexible ambiguity structure that is not limited to a finite set of priors. We restrict the set of admissible indemnity schedules by imposing the so-called no-sabotage condition. It stipulates that the indemnity and the retained loss are both increasing in loss severity so that the policyholder and the insurer both bear more of the loss the larger its realization. Such indemnity schedules are incentive-compatible in the sense that they do not encourage upward or downward manipulation of the loss. Our approach complements Gollier's (2014) who restricts the analysis to a finite set of priors and does not impose the no-sabotage condition. While the difference between the two settings may appear slight, the results can be remarkably different.

Our contribution to the literature is fourfold. First, we characterize optimal indemnity schedules and show that the marginal indemnity is either zero or one except at critical points. Second, we provide a condition for the optimal indemnity schedule to be a straight deductible, thus extending Arrow's celebrated result to the case of smooth ambiguity aversion. We then identify stochastic ordering assumptions on the priors, under which this condition is satisfied. Third, we discuss specific ambiguity structures to isolate the effect of the no-sabotage condition on the shape of the optimal indemnity schedule. Fourth, we conduct comparative statics. Greater ambiguity aversion always raises insurance demand in our model whereas greater ambiguity has indeterminate effects. For ambiguity-prudent policyholders with relative ambiguity prudence less than two, the intuitive result prevails and greater ambiguity raises insurance demand. These results are consistent with the comparative statics of smooth ambiguity aversion in portfolio choice (Gollier, 2011; Huang and Tzeng, 2018), self-insurance and self-protection (Snow, 2011; Alary et al., 2013), nonperformance risk (Peter and Ying, 2020), and precautionary saving (Peter, 2019) but do not arise in Gollier's (2014) model.

The problem of optimal insurance design has been extended in many directions. Raviv (1979), Huberman et al. (1983) and Young (1999) consider other premium principles and nonlinear transaction costs. Due to a dominance result by Gollier and Schlesinger (1996), a straight deductible is optimal in any decision-theoretic framework that respects second-order stochastic dominance (see Zilcha and Chew, 1990; Karni, 1992; Machina, 1995). Other researchers have studied the effect of background risk,[1] belief heterogeneity,[2] and probability distortions.[3] Gollier (2013) summarizes the literature on optimal insurance design under expected utility and Ghossoub (2019b) the literature under non-expected utility.

Camerer and Weber (1992) define ambiguity as "uncertainty about probability, created by missing information that is relevant and could be known." In the context of optimal insurance design, a policyholder typically faces uncertainty about the loss distribution. Even with the most comprehensive data, it is impossible to narrow down the loss distribution perfectly and any parameter estimate comes with nontrivial confidence intervals. Against this background, a realistic approach to optimal insurance design should take uncertainty into account. Ellsberg's (1961) famous thought experiment reveals that people are sensitive to ambiguity. Ambiguity aversion has been documented in laboratory experiments (e.g., Einhorn and Hogarth, 1986; Chow and Sarin, 2001), market settings with educated individuals (e.g., Sarin and Weber, 1993), and surveys of business owners and managers (e.g., Viscusi and Chesson, 1999; Chesson and Viscusi, 2003). While not universal (Kocher et al., 2018; Baillon and Emirmahmutoglu, 2018), recent survey evidence of US households confirms the role of ambiguity aversion for financial decision-making in the field (Dimmock et al., 2016)

Popular models of decision-making under ambiguity include Choquet expected utility (Schmeidler, 1989), the maxmin expected utility model (Gilboa and Schmeidler, 1989), the α-maxmin expected utility model (Ghirardato et al., 2004), and smooth ambiguity aversion

[1] See Mayers and Smith Jr (1983), Eeckhoudt and Kimball (1992), Gollier (1996), Dana and Scarsini (2007), Chi and Wei (2018) and Chi and Tan (2021).

[2] See Marshall (1992), Ghossoub (2017), Boonen and Ghossoub (2019) and Chi and Wei (2020).

[3] See Bernard et al. (2015), Xu et al. (2019) and Ghossoub (2019b).

(Klibanoff et al., 2005). From a conceptual standpoint, the main advantage of the smooth model is that it disentangles tastes and beliefs. This property allows us to derive clean comparative statics by varying the degree of ambiguity aversion while keeping the level of uncertainty fixed, or varying the level of uncertainty while keeping the degree of ambiguity aversion fixed. Based on the exchange between Epstein (2010) and Klibanoff et al. (2012), Cubitt et al. (2020) construct an experimental test that discriminates between different classes of decision-making models under ambiguity. They find greater support for the smooth ambiguity model than for the maxmin or α-maxmin models, with the relative support being stronger for subjects who are classified as ambiguity-averse.[4]

We are not the first to study optimal insurance design under ambiguity. Carlier et al. (2003) analyze Pareto-optimal insurance contracts under Choquet expected utility with epsilon-contaminated priors. In their model, optimal insurance contracts necessarily satisfy the no-sabotage condition. Unlike the smooth model, Choquet expected utility does not separate tastes and beliefs. Alary et al. (2013) show the optimality of a straight deductible in the smooth ambiguity model under very restrictive conditions. They require that only the loss probability is subject to uncertainty and that policyholders know the loss severity distribution perfectly. Gollier (2014) provides a more general analysis of optimal insurance design in the smooth model. He derives a condition under which a disappearing deductible is optimal and highlights the possibility that an increase in the policyholder's degree of ambiguity aversion may have the counterintuitive effect to reduce optimal insurance demand.

We take a fresh look at optimal insurance design under smooth ambiguity aversion. Our setting yields a general characterization of optimal indemnity schedules, which is to date lacking in the literature. A straight deductible is optimal under a broader set of circumstances in our model. Arrow's famous result is thus often "robust" to uncertainty. Unlike previous studies, we find that indemnities with multiple layers can arise. Finally, the comparative statics of greater ambiguity aversion and greater ambiguity yield intuitive results in our setting that are consistent with the literature. Compared to Gollier's (2014) approach, the no-sabotage condition often leads to qualitatively different results. Whether insurers are willing to offer contracts that violate the no-sabotage condition is an empirical issue. If not, the no-sabotage condition only rules out indemnity schedules that are not observed in practice anyway.

The paper proceeds as follows. Section 2 outlines the model. Section 3 presents properties of optimal indemnity schedules including our main characterizing theorem. Section 4 provides a condition for a straight deductible to be optimal and discusses stochastic ordering assumptions under which this condition is satisfied. Section 5 analyzes several specific ambiguity structures including left-against-right ambiguity, one-against-all ambiguity, and two-state ambiguity. Section 6 derives comparative statics with respect to the policyholder's degree of ambiguity aversion and the level of ambiguity. A final section concludes.

[4] Cubitt et al. (2020) summarize the evidence on distinguishing between different ambiguity models empirically and the conceptual challenges that arise in this literature, see their Section 2.3.

2 The model

A policyholder is endowed with initial wealth w and faces a loss of $X \geq 0$. The random variable X is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and the policyholder is uncertain about the loss distribution. We let the distribution of X be parameterized by Θ, which is also a random variable. Our ambiguity structure is thus quite flexible and includes Gollier's (2014) setting with a finite number of priors as a special case. To see this, let random variable Θ take n possible values, $\theta_1, \ldots, \theta_n$, and define

$$F_i(x) = \mathbb{P}(X \leq x | \Theta = \theta_i) \qquad \text{for } x \in \mathbb{R} \text{ and } i = 1, \ldots, n.$$

F_i denotes the cumulative distribution function of loss X conditional on $\Theta = \theta_i$. If we set $q_i = \mathbb{P}(\Theta = \theta_i)$, we can interpret $(q_1, \ldots, q_n)$ as the policyholder's second-order belief, where q_i denotes the probability that F_i is the true loss distribution.

In our analysis, we do not restrict Θ to have a finite number of outcomes. For example, the policyholder may have a particular family of loss distributions in mind and estimate a location parameter from past observations. In this case, we can let Θ take values from an estimated confidence interval and let the distribution of Θ under $\mathbb{P}$ be the distribution of the estimator. Our ambiguity structure thus generalizes the ambiguity structure in Gollier (2014). Denote by $\mathcal{M}_\theta$ the essential supremum of X conditional on $\Theta = \theta$. Clearly, $\mathcal{M} = \sup_\theta \mathcal{M}_\theta$ is then the essential supremum of random variable X, and we assume $\mathcal{M} < \infty$.

The policyholder can purchase insurance to mitigate the risk of loss. An insurance contract is a pair (I, π), where I is the indemnity schedule and π is the insurance premium. The indemnity schedule specifies the amount paid by the insurer to the policyholder. This amount is $I(x)$ conditional on a loss of $X = x$. The policyholder's retained loss is given by $R_I(x) = x - I(x)$. In actuarial terms, the functions I and R_I are usually called the policyholder's ceded and retained loss functions. One can think of $I(X)$ as the amount of risk ceded to the insurer and of $R_I(X)$ as the amount of risk retained by the policyholder. We make the following assumptions about the indemnity schedule and the retained loss function:

A1: Principle of indemnity: For all x, we have $0 \leq I(x) \leq x$;

A2: No-sabotage condition: I and R_I are increasing in x.

Arrow (1963) introduces $I(x) \geq 0$ for all x and Raviv (1979) assumes the full principle of indemnity. It is widely accepted in the insurance economics literature because it rules out short-selling of insurance and overinsurance.[5] The no-sabotage condition states that both the policyholder and the insurer bear more of the insurable loss the larger its realization. Huberman et al. (1983) justify increasing indemnity schedules with the policyholder's ability

[5] Gollier (1987) and Ghossoub (2019a) analyze optimal insurance design without the principle of indemnity.

to misrepresent downward the magnitude of loss. They also show that indemnity schedules with an increasing retained loss help rule out counterfactual insurance contracts. If the retained loss were decreasing, the policyholder would have an incentive to inflate the size of the claim or create further damage, leading to an ex-post moral hazard problem. Chiappori et al. (2006) focus on indemnity schedules for which the retained loss is increasing and state that "this property is satisfied empirically" and "relies on compelling theoretical arguments." More recently, Xu et al. (2019) and Chi and Wei (2020) use the no-sabotage condition for optimal insurance design in the context of rank-dependent utility and belief heterogeneity.

Let $\mathfrak{C}$ denote the set of admissible indemnity schedules satisfying A1 and A2. Note that $I \in \mathfrak{C}$ if, equivalently, $I(0) = 0$, I is increasing and one-Lipschitz continuous,

$$0 \leq I(x) - I(y) \leq x - y \qquad \text{for all } 0 \leq y \leq x.$$

It follows that any admissible indemnity schedule $I \in \mathfrak{C}$ is differentiable with $0 \leq I'(x) \leq 1$ almost everywhere.

The insurer operates on a competitive market, is risk- and ambiguity-neutral, and faces linear transaction costs. For each dollar of indemnity payment, the insurer incurs a cost of $(1+\tau)$ dollars. The premium π for indemnity schedule I is thus given by

$$\pi(I(X)) = (1+\tau)\mathbb{E}[I(X)].$$

We can interpret τ as a safety loading coefficient. From the policyholder's perspective, the insurance premium is actuarially fair when $\tau = 0$ and actuarially unfair when $\tau > 0$.

The policyholder's final wealth when choosing insurance contract (I, π) is given by

$$W_I = w - X + I(X) - (1+\tau)\mathbb{E}[I(X)].$$

Following Klibanoff et al. (2005) and Neilson (2010), we characterize the policyholder's risk and ambiguity preferences with functions u and ϕ. In this setting, u is a von Neumann-Morgenstern (vNM) utility function of final wealth that is strictly increasing, $u' > 0$, and strictly concave, $u'' < 0$. The ambiguity function ϕ is assumed strictly increasing, $\phi' > 0$. Its curvature characterizes the policyholder's attitude towards ambiguity. If ϕ is linear, $\phi'' = 0$, the policyholder is ambiguity-neutral and the model collapses to expected utility. If ϕ is strictly concave, $\phi'' < 0$, the policyholder is ambiguity-averse, and her ex-ante welfare is lowered by the presence of ambiguity. On a competitive insurance market, the insurer offers a contract that maximizes the policyholder's ex-ante welfare and thus solves

$$\max_{I \in \mathfrak{C}} J(I) = \mathbb{E}[\phi(\mathbb{E}[u(W_I)|\Theta])]. \tag{1}$$

When the policyholder is ambiguity-neutral, we obtain Arrow's (1971) classical model as a special case. For later reference, we will briefly revisit its solution. Consider the problem

$$\max_{I \in \mathfrak{C}} \mathbb{E}[u(W_I)], \tag{2}$$

which has the same solution as Problem (1) when ϕ is linear. We define by $S_X(x) = \mathbb{P}(X > x)$ the survival function of the loss X. Let

$$x_\tau = \inf \left\{ x \geq 0 : S_X(x) \leq \frac{1}{1+\tau} \right\} \tag{3}$$

be the smallest loss amount so that the survival function is below the inverse of the insurer's gross premium rate. Loss amount x_τ is a critical level. The probability of losses larger than x_τ is less than $1/(1+\tau)$ whereas the probability of losses smaller than x_τ is at least $\tau/(1+\tau)$. The threshold x_τ is increasing in τ, and when $\tau = 0$, we obtain $x_\tau = 0$. For $D \in [0, \mathcal{M})$, we consider the straight deductible contract $I_D(x) = \max(0, x - D)$ with associated premium $\pi_D = (1+\tau)\mathbb{E}[\max(0, X - D)]$. In the actuarial literature, this shape of the indemnity schedule is often referred to as stop-loss insurance. Chi and Wei (2018) introduce the following function:

$$V^0(D) = \frac{u'(w - D - \pi_D)}{\mathbb{E}[u'(w - \min(X, D) - \pi_D)]}.$$

It relates the policyholder's marginal utility at the deductible D to her expected marginal utility over the whole range of possible loss realizations. Chi and Wei (2018) show that $V^0(D)$ is increasing in D over $[x_\tau, \mathcal{M})$ and obtain the following result.

Proposition 1. *Suppose the policyholder is ambiguity-neutral (i.e., $\phi'' = 0$). Then, the optimal insurance contract is a straight deductible, $I_{D_0^*}(x)$ for all $x \geq 0$. The optimal deductible D_0^* is given as follows:*

$$D_0^* = \max \left(\sup \left\{ D \in [x_\tau, \mathcal{M}) : V^0(D) \leq 1 + \tau \right\}, x_\tau \right). \tag{4}$$

Proposition 1 is Arrow's (1965, 1971) famous result that a straight deductible is the optimal indemnity schedule. It also arises as a special case of Gollier's (2014) model when the policyholder is ambiguity-neutral, see his Proposition 1, and continues to hold under our generalized ambiguity structure. In addition, we provide the optimal level of the deductible in Eq. (4) explicitly. Set $\sup \emptyset = -\infty$ by convention. Then, the optimal deductible is at least x_τ as defined in Eq. (3). It coincides with x_τ if $V^0(D)$ exceeds $(1+\tau)$ for all deductible levels D larger than x_τ. Otherwise, the optimal deductible is the largest value so that $V^0(D)$ is just below the insurer's gross premium rate $(1+\tau)$. Intuitively, insurance is most valuable when final wealth is low because then marginal utility is high. This is the case when losses are large. A straight deductible prioritizes the indemnification of large losses over small losses, and is thus an effective way to maximize the policyholder's expected utility.

3 Properties of optimal indemnity schedules

In this section, we first show the existence and uniqueness of a solution to Problem (1) for a policyholder who is ambiguity-averse. We then characterize optimal indemnity schedules and use our characterization to verify Mossin's (1968) Theorem.

Lemma 1.

(i) A solution to Problem (1) *exists.*

(ii) The solution to Problem (1) *is unique almost surely if one of the following two conditions is satisfied:*

(a) The loading is strictly positive, $\tau > 0$.

(b) A loss of zero belongs to the support of X*, that is,* $\mathbb{P}(X < \epsilon) > 0$ *for all* $\epsilon > 0$.

Then, if I_1 *and* I_2 *are both solutions to Problem* (1)*, we have* $\mathbb{P}(I_1(X) = I_2(X)) = 1$.

The proof is similar to the proof of Lemma 2.1 in Chi and Wei (2020) and is therefore omitted. The idea behind result (i) is to define a metric on $\mathfrak{C}$ under which $\mathfrak{C}$ is compact, and then apply the Arzelà-Ascoli Theorem. Result (ii) is obtained by establishing the optimality of convex combinations of two optimal solutions I_1 and I_2, which allows us to show that the associated final wealth prospects coincide almost surely under $\mathbb{P}$. In practice, conditions (a) and (b) are not restrictive because insurance contracts are rarely actuarially fair unless governmental subsidies are in place, and a loss of zero is usually an outcome with positive probability. So in most cases, conditions (a) and (b) will both be satisfied although each one of them alone is sufficient to ensure the uniqueness of the solution.

To provide a necessary and sufficient condition for an indemnity schedule to be optimal, we need to make some technical assumptions. Let $W_0 = w - X - (1+\tau)\mathbb{E}[X]$ be the policyholder's final wealth if she bears the entire loss and the cost of full insurance. Throughout the paper, we assume that the following two assumptions hold:

A3: Boundedness: $\mathbb{E}[u'(W_0)] < \infty$ and $\sup_\theta \phi'(\mathbb{E}[u(W_0)|\Theta = \theta]) < \infty$;

A4: Dominated convergence: $\mathbb{E}[Xu'(W_0)] < \infty$.

For any admissible indemnity schedule $I \in \mathfrak{C}$, we then have

$$W_0 = w - X - (1+\tau)\mathbb{E}[X] \leq w - X + I(X) - (1+\tau)\mathbb{E}[I(X)] = W_I.$$

This inequality follows from the principle of indemnity, which implies $\mathbb{E}[I(X)] \leq \mathbb{E}[X]$. As a result, $u(W_0) \leq u(W_I)$ and $u'(W_I) \leq u'(W_0)$. Assumption A3 then ensures that the policyholder's expected marginal welfare does not explode,

$$\mathbb{E}[\phi'(\mathbb{E}[u(W_I)|\Theta])u'(W_I)] \leq \sup_\theta \phi'(\mathbb{E}[u(W_0)|\Theta = \theta]) \cdot \mathbb{E}[u'(W_0)] < \infty.$$

This guarantees that the following auxiliary function is well-defined:

$$V_I(x) \quad = \quad \frac{\mathbb{E}[\phi'(\mathbb{E}[u(W_I)|\Theta])u'(W_I)|X > x]}{\mathbb{E}[\phi'(\mathbb{E}[u(W_I)|\Theta])u'(W_I)]}, \qquad \text{for } x \in [0, \mathcal{M}). \tag{5}$$

Auxiliary function $V_I(x)$ plays a central role in our analysis. It compares the policyholder's expected marginal welfare for losses in excess of x to her overall expected marginal welfare under indemnity schedule $I \in \mathfrak{C}$. Assumption A4 is a technical condition that allows us to exchange the operations of expectation and differentiation due to Lebesgue's dominated convergence theorem. The following result characterizes the optimal indemnity schedule.

Theorem 1. *An indemnity schedule $I^* \in \mathfrak{C}$ is a solution to Problem* (1) *if and only if it satisfies the following condition almost everywhere:*

$$I^{*\prime}(x) = \begin{cases} 1, & \text{for } V_{I^*}(x) > 1 + \tau, \\ 0, & \text{for } V_{I^*}(x) < 1 + \tau. \end{cases} \tag{6}$$

Appendix A.1 provides the proof. Theorem 1 gives a general characterization of the optimal solution to Problem (1). Specifically, the marginal indemnity $I^{*\prime}(x)$ is either 0 or 1, with some irregularities at the critical point(s) where $V_{I^*}(x) = 1 + \tau$. This means that, at the margin, when the loss increases by a dollar, the additional loss is either fully covered if $I^{*\prime}(x) = 1$, or not covered at all if $I^{*\prime}(x) = 0$. While Theorem 1 does not explicitly solve the insurance design problem, it is useful in verifying the optimality of an indemnity schedule and establishing qualitative properties of optimal solutions. Theorem 1 also yields a strategy to improve suboptimal indemnity schedules, which we show in Appendix C. Furthermore, it implies that the policyholder optimally retains small losses in the following sense.

Corollary 1. *An optimal indemnity schedule I^* satisfies $I^*(x) = 0$ for all $x < x_\tau$, where x_τ is defined in Eq.* (3)*.*

Proof. This follows directly from Theorem 1. For $x < x_\tau$, we obtain

$$V_{I^*}(x) \leq \frac{1}{\mathbb{P}(X > x)} < 1 + \tau.$$

□

In other words, the optimal insurance contract generally has a deductible if $x_\tau > 0$. An exception arises if $\tau = 0$ because then the insurance contract is actuarially fair and $x_\tau = 0$. In this case, full insurance is optimal. As it turns out, $\tau = 0$ is also necessary for the optimality of full insurance as summarized in our next result.

Proposition 2. *Full insurance is an optimal solution to Problem* (1) *if and only if the contract is actuarially fair ($\tau = 0$).*

Proof. Denote the indemnity schedule for full insurance by $I_f(x) = x$. With full insurance, the policyholder's final wealth is constant, $W_{I_f} = w - (1+\tau)\mathbb{E}[X]$. As a result, auxiliary function V_{I_f} is also constant, $V_{I_f}(x) = 1$ for all $x \in [0, \mathcal{M})$. Thus, Theorem 1 implies that $I_f(x)$ is an optimal solution if and only if $1 = V_{I_f}(x) \geq 1+\tau$, which is equivalent to $\tau = 0$. □

Proposition 2 is often referred to as Mossin's (1968) Theorem in insurance economics (see Schlesinger, 2013). Gollier (2014) shows it in his Proposition 2 and we extend it to our generalized ambiguity structure. It is not obvious that Mossin's Theorem remains valid under ambiguity. Dow and da Costa Werlang (1992) obtain a no-trade result, and Mukerji and Tallon (2001) find that financial markets are incomplete under ambiguity. As explained by Lang (2017), the preferences in the last two papers exhibit first-order ambiguity aversion whereas Klibanoff et al.'s (2005) smooth model has second-order ambiguity aversion. For this reason, complete risk transfer can only be optimal when the premium is actuarially fair.

So far we have seen that a straight deductible is optimal under ambiguity neutrality (Proposition 1) and that full insurance is optimal if and only if the premium is actuarially fair (Proposition 2). Gollier (2014) also finds these results. More realistically, policyholders are ambiguity-averse, $\phi'' < 0$, and the premium is actuarially unfair, $\tau > 0$. Then, full insurance is no longer optimal, which raises the question of optimal insurance design. As we will show in the next section, Theorem 1 allows us to derive a condition for the optimality of a straight deductible, which is satisfied under various stochastic ordering assumptions on the priors. It is in this case that our results often differ from those in Gollier (2014).

4 Optimality of straight deductible contracts

Recall that $I_D(x)$ denotes the indemnity schedule with a straight deductible $D \geq 0$. Indemnity schedule $I_{\mathcal{M}}$ represents the no-insurance strategy because $I_{\mathcal{M}}(X) = 0$ almost surely. The following result holds.

Theorem 2. *If $V_{I_D}(x)$ is increasing in x on $[x_\tau, \mathcal{M})$ for any $D \geq x_\tau$, then the solution to Problem* (1) *is a straight deductible with deductible level*

$$D^* = \inf \{ D \geq x_\tau : V_{I_D}(D) \geq 1+\tau \} . \tag{7}$$

Appendix A.2 gives the proof. When the auxiliary function $V_{I_D}(x)$ is increasing in x on $[x_\tau, \mathcal{M})$ for any deductible level $D \geq x_\tau$, we can apply Theorem 1 to indemnity schedule I_{D^*} to verify its optimality. We emphasize that Eq. (7) allows us to find the optimal deductible with the help of auxiliary function $V_{I_D}(x)$ evaluated at $x = D$. One might wonder how restrictive it is to require $V_{I_D}(x)$ to be increasing in $x \in [x_\tau, \mathcal{M})$ for all $D \geq x_\tau$. To answer this question and make Theorem 2 more applicable, we now introduce different stochastic ordering assumptions.

Definition 1. Let Y and Z be two random variables.

(i) Y is smaller than Z in the first-order stochastic dominance (FSD) sense, denoted as $Y \leq_{fsd} Z$, if $\mathbb{P}(Z \leq x) \leq \mathbb{P}(Y \leq x)$ for all x.

(ii) Y is smaller than Z in the hazard rate order, denoted as $Y \leq_{hr} Z$, if $\mathbb{P}(Z > x)/\mathbb{P}(Y > x)$ is increasing in x.

(iii) Y is smaller than Z in the likelihood ratio order, denoted as $Y \leq_{lr} Z$, if, for all measurable sets A and B with $A \leq B$, we have $\mathbb{P}(Y \in A) \cdot \mathbb{P}(Z \in B) \geq \mathbb{P}(Y \in B) \cdot \mathbb{P}(Z \in A)$.

(iv) Y is less risky than Z in the sense of Rothschild and Stiglitz (1970), denoted as $Y \leq_{RS} Z$, if $\mathbb{E}[Y] = \mathbb{E}[Z]$ and $\int_{-\infty}^{x} \mathbb{P}(Y \leq t)\,\mathrm{d}t \leq \int_{-\infty}^{x} \mathbb{P}(Z \leq t)\,\mathrm{d}t$ for all x.

First-order stochastic dominance is often referred to as the usual stochastic order in actuarial science. If Z dominates Y by FSD, then $\mathbb{E}[v(Z)] \geq \mathbb{E}[v(Y)]$ for all increasing functions v such that the expectations exist. Intuitively, Z is more likely to take on large values than Y, which is appreciated when higher values are a good thing. The hazard rate is defined as the intensity of failure, and the hazard rate order then ranks random variables according to this intensity. Hazard rates feature prominently in survival analysis. The likelihood ratio order is used extensively in the moral hazard literature. Rogerson (1985a,b) shows that the first-order approach to the principal-agent problem is valid when the monotone likelihood ratio property holds and when the distribution function of outcomes is concave in the agent's effort level. The likelihood ratio property also features prominently in the theory of monotone comparative statics under uncertainty (Athey, 2002). The order in Rothschild and Stiglitz (1970) is a variability order because if Y is less risky than Z, it has a lower variance, $\mathrm{Var}[Y] \leq \mathrm{Var}[Z]$, provided the variances exist. It is referred to as the convex order in actuarial science because $\mathbb{E}[v(Z)] \geq \mathbb{E}[v(Y)]$ for all convex functions v, given the expectations exist.

Shaked and Shanthikumar (2007) provide a systematic overview of properties of these stochastic orders in their Chapter 1. Specifically, they show the following:

$$Y \leq_{lr} Z \qquad \Longrightarrow \qquad Y \leq_{hr} Z \qquad \Longrightarrow \qquad Y \leq_{fsd} Z.$$

Ordering random variables by the likelihood ratio order is more restrictive than ordering them by the hazard rate order which, in turn, is more restrictive than ordering them by first-order stochastic dominance. All of these orders are partial in the sense that any two random variables may or may not be ordered in a particular way. Based on Definition 1, we can now define the notion of stochastic increasingness.

Definition 2. Random variable X is increasing in random variable Θ in the FSD sense, denoted as $X \uparrow_{fsd} \Theta$, if $[X|\Theta = \theta_1]$ is smaller than $[X|\Theta = \theta_2]$ in the FSD sense for all θ_1 and θ_2 such that $\theta_1 \leq \theta_2$.

As Θ takes on larger values, random variable X becomes larger stochastically. Stochastic increasingness in the hazard rate order and in the likelihood ratio order are defined analogously. We can likewise define stochastic decreasingness by assuming that higher realizations of Θ decrease random variable X stochastically. Jindapon and Neilson (2007) use a stochastic ordering assumption for higher-order risk to characterize comparative Arrow-Pratt and comparative Ross risk aversion (see also Liu and Neilson, 2019). Crainich et al. (2016) use a similar set-up to analyze how risk changes affect effort provision.

The following result uses stochastic ordering assumptions to ensure that auxiliary function $V_{I_D}(x)$ satisfies the criterion in Theorem 2.

Proposition 3. *The solution to Problem* (1) *is a straight deductible with the deductible level specified in Eq.* (7) *if one of the following conditions holds:*

(i) The ratio $\mathbb{P}(X > x|\Theta = \theta)/\mathbb{P}(X > x_\tau|\Theta = \theta)$ is independent of θ for all $x \geq x_\tau$;

(ii) Random variables X and Θ are increasing or decreasing in each other in the FSD sense;

(iii) Random variable X is increasing or decreasing in Θ in the hazard rate order.

Appendix A.3 states the proof. Condition (i) implies so-called tail independence between X and Θ. When given $X > x_\tau$, the distribution of X conditional on Θ is independent of Θ. Condition (i) is satisfied when the ambiguity is concentrated on losses below x_τ. Gollier (2014) considers this special case in his Proposition 5 and concludes that a straight deductible is optimal if the degree of ambiguity aversion is small enough. Result (i) extends this conclusion to our more flexible ambiguity structure and removes the restriction on the degree of ambiguity aversion by virtue of the no-sabotage condition. A straight deductible is thus the optimal insurance contract in a larger set of circumstances under our approach.

Condition (ii) requires that X and Θ are mutually ordered in the FSD sense. If X is increasing in Θ in the FSD order, then losses become larger in the FSD sense as θ increases. As such random variable Θ orders the priors from better to worse because stochastically smaller losses are preferred over stochastically larger losses. Condition (ii) also requires Θ to be increasing in X. Conditional on observing a large loss, the probability that the underlying realization of Θ is large is greater than conditional on observing a small loss. Assuming Θ to be increasing in X justifies the inference that large losses are more likely to originate from worse priors. As discussed in Cai and Wei (2012), monotonicity in the FSD sense is not symmetric, and we thus need to assume X to be monotonic in Θ and vice versa.

Gollier (2014) only studies the case in which Θ ranks the priors in the FSD sense in his Proposition 9 for the special case that ambiguity is concentrated on losses below x_τ. Our result (ii) shows that letting X and Θ be mutually ordered in the FSD sense ensures the optimality of a straight deductible regardless of whether ambiguity is concentrated on losses below x_τ or not. Yet again, this expands the set of circumstances under which a straight deductible is optimal. Our condition (iii) shows that ordering priors by the hazard rate order also leads to the optimality of a straight deductible. Monotonicity in the hazard rate order is implied by monotonicity in the liklihood ratio order, which proves the following result.

Corollary 2. *The solution to Problem* (1) *is a straight deductible with deductible level specified in Eq.* (7) *if random variable X is increasing or decreasing in Θ in the likelihood ratio order.*

Corollary 2 follows from Proposition 3(*iii*) because ranking random variables by the likelihood ratio order implies their ranking by the hazard rate order. It also follows from Proposition 3(*ii*) because increasingness in the likelihood ratio order is symmetric, see Appendix B.4, and implies increasingness in the FSD order. Gollier (2014) considers the likelihood ratio order in his Proposition 8 only for the special case of two priors, that is, $\mathbb{P}(\Theta \in \{\theta_1, \theta_2\}) = 1$. He finds a so-called "disappearing deductible" to be optimal because in his result the retained loss above the deductible is nonincreasing in x. Corollary 2 removes the restrictive assumption of only two priors. Furthermore, the no-sabotage condition rules out disappearing deductibles and we then find a straight deductible to be optimal. Proposition 3(*iii*) shows that the likelihood ratio order is unnecessarily constraining though because it suffices to rank priors by the hazard rate order to obtain the same result. Yet again, our approach extends the optimality of a straight deductible considerably by virtue of the no-sabotage condition.

We discussed after Proposition 3 that condition (*i*) holds when ambiguity is concentrated on losses below x_τ. We now show that Theorem 1 also applies to the dual scenario in which ambiguity is concentrated on losses above D_0^*, where D_0^* is the optimal deductible for an ambiguity-neutral policyholder as defined in Eq. (4).

Proposition 4. *Let ambiguity be concentrated on losses above D_0^* so that $\mathbb{P}(X \leq x|\Theta = \theta)$ is independent of θ for any $x \leq D_0^*$. In this case, the solution to Problem* (1) *is a straight deductible with deductible level D_0^*. Ambiguity has no effect on the optimal insurance contract.*

Appendix A.4 gives the proof. Gollier's (2014) Proposition 3 reaches the same conclusion. Our Proposition 4 extends this result to our more flexible ambiguity structure. Whether ambiguity is concentrated on losses below x_τ or above D_0^*, or is not concentrated at all is an empirical issue. In practice, the answer to this question is likely to depend on the type of loss exposure. In the situation of parameter uncertainty mentioned in the introduction, ambiguity is not concentrated on any particular portion of the loss distribution. In this case, conditions (*ii*) and (*iii*) establish the optimality of a straight deductible whereas Gollier's (2014) model remains silent about the shape of the optimal indemnity schedule. Our Eq. (7) even specifies the optimal deductible level, which will, in general, be different from D_0^*. Before we discuss these comparative statics, we will investigate some specific ambiguity structures to isolate the effect of the no-sabotage condition on optimal insurance design and highlight additional differences between our results and those in Gollier (2014).

5 Some specific ambiguity structures

5.1 Preliminaries

Consider the following three-piece setting:

$$\mathbb{P}(X \le x|\Theta = \theta) = (1-p)\mathbb{P}(\hat{X} \le x) + p(1-\theta)\mathbb{P}(X_1 \le x) + p\theta\mathbb{P}(X_2 \le x) \quad \text{for } x \ge 0, \tag{8}$$

where we let $p \in [0,1]$, $\theta \in [0,1]$, and $\hat{X}, X_1$ and X_2 be nonnegative random variables. The ambiguity structure in Eq. (8) says intuitively that the policyholder is certain about the loss distribution on one set of values but uncertain about the loss distribution on another set of values. If $p = 0$, the policyholder is certain that the loss distribution is that of $\hat{X}$. As soon as $p > 0$, the value of θ matters and the policyholder faces uncertainty over the true loss distribution. In the special case of $p = 1$, the policyholder believes that the loss distribution can be any mixture of the distributions of X_1 and X_2 depending on θ. For example, if $\theta = 0$ and $\theta = 1$ each have a 50% chance of occurring, the policyholder thinks that the true loss distribution is equally likely to be that of X_1 or that of X_2.

Eq. (8) includes many ambiguity structures as special cases including left-against-right ambiguity, one-against-all ambiguity, and two-state ambiguity. We discuss them in the sequel.

5.2 Left-against-right ambiguity

We obtain the special case of left-against-right ambiguity by ordering the random variables X_1, $\hat{X}$, and X_2 from left to right. One possibility to order them is to require

$$ess\sup X_1 \le ess\inf \hat{X} \qquad \text{and} \qquad ess\sup \hat{X} \le ess\inf X_2, \tag{9}$$

where $ess\sup X$ and $ess\inf X$ denote the essential supremum and the essential infimum of random variable X. Intuitively, under the left-against-right ambiguity structure the policyholder is uncertain about the loss distribution on the two tails but certain about the loss distribution for losses of intermediate size. We obtain the following result.

Lemma 2. *Under assumptions* (8) *and* (9)*, X is increasing in Θ in the likelihood ratio order.*

Appendix A.5 gives the proof. If the positions of X_1 and X_2 are switched in condition (9), leading to $ess\sup X_2 \le ess\inf \hat{X}$ and $ess\sup \hat{X} \le ess\inf X_1$, then X is decreasing in Θ in the likelihood ratio order.

Ordering assumption (9) is restrictive. The proof of Lemma 2 reveals that, for X to be increasing in Θ in the likelihood ratio order, all we need is that $X_1 \le_{lr} \hat{X} \le_{lr} X_2$. Condition (9) is sufficient but not necessary for ranking X_1, $\hat{X}$ and X_2 by the likelihood ratio order. In fact, if we impose the constraint that X_1, $\hat{X}$ and X_2 have disjoint supports, then $X_1 \le_{lr} \hat{X} \le_{lr} X_2$ also implies condition (9) but requiring disjoint supports is itself restrictive.

One might thus wonder how X and Θ are ordered if we relax the ordering assumption on X_1, $\hat{X}$ and X_2. The next lemma answers this question.

Lemma 3. *Under assumption* (8)*, if* $X_1 \leq_{hr} \hat{X} \leq_{hr} X_2$*, then* X *is increasing in* Θ *in the hazard rate order.*

Appendix A.6 states the proof. Weakening the ordering assumption on X_1, $\hat{X}$ and X_2 also weakens the ordering of X in Θ. Stochastic increasingness in the hazard rate order is, however, strong enough to establish the optimality of a straight deductible. We summarize this result in the next proposition.

Proposition 5. *Under assumption* (8) *with* $X_1 \leq_{hr} \hat{X} \leq_{hr} X_2$*, the solution to Problem* (1) *is a straight deductible with the deductible level specified in Eq.* (7)*.*

Proposition 5 follows directly from Proposition 3(*iii*) because X is increasing in Θ in the hazard rate order when X_1, $\hat{X}$ and X_2 are ordered in the hazard rate order (Lemma 3). As a corollary, a straight deductible with the deductible level specified in Eq. (7) is also optimal when X_1, $\hat{X}$ and X_2 are ordered in the likelihood ratio order or when they satisfy ordering assumption (9). We can either conclude this from Proposition 5 because the likelihood ratio order is a special case of the hazard rate order, or from Lemma 2 and Corollary 2.

Gollier (2014) does not discuss left-against-right ambiguity. We add this ambiguity structure to the list for two reasons. First, it illustrates the usefulness of Theorem 2 and Proposition 3. Second, it corroborates our main point that a straight deductible is the solution to the optimal insurance design problem under smooth ambiguity aversion in a large set of circumstances. Left-against-right ambiguity is one of them.

5.3 One-against-all ambiguity

We obtain the one-against-all ambiguity structure by setting $p = 1$ and $X_1 = x_1$ with $x_1 \geq 0$ in Eq. (8). Specifically, the distribution of the loss X conditional on $\Theta = \theta$ is given by

$$\mathbb{P}(X \leq x|\Theta = \theta) = (1-\theta)\mathbf{1}_{\{x_1 \leq x\}} + \theta\mathbb{P}(X_2 \leq x) \qquad \text{for } x \geq 0, \tag{10}$$

where $\mathbf{1}_A$ denotes the indicator function of event A. Gollier (2014) considers one-against-all ambiguity and refers to x_1 as the ambiguous state. Intuitively, Eq. (10) means that the probability of state x_1 is ambiguous, and this ambiguity is compensated by all other states. The distribution of X_2 does not depend on θ, and therefore all other states exhibit a constant level of ambiguity. Gollier (2014) finds that *"[t]he optimal deductible applied to the ambiguous loss is smaller than the deductible applied to all other losses."* This rules out a straight deductible in his model. In our model, the optimal contract may be of the straight deductible type even though other shapes are possible as we show in the next proposition.

Proposition 6. *Under the one-against-all ambiguity structure, the solution to Problem* (1) *is a multi-layer contract that takes one of the following two forms:*

(i) $I_l(x; D_l, U) = \min(\max(x - D_l, 0), U) + \max(x - x_1, 0)$ *with* $0 \leq D_l \leq D_l + U \leq x_1$.

(ii) $I_r(x; D_r, U) = \min(\max(x - x_1 + U, 0), U) + \max(x - D_r, 0)$ *with* $0 \leq U \leq x_1 \leq D_r$.

Appendix A.7 shows this result. To interpret the indemnity schedules in Proposition 6, we rewrite them to see how losses of different size are indemnified. For case (i), we have

$$I_l(x; D_l, U) = \begin{cases} 0, & \text{for } x \leq D_l, \\ x - D_l, & \text{for } D_l < x \leq U + D_l, \\ U, & \text{for } U + D_l < x \leq x_1, \\ x - (x_1 - U), & \text{for } x > x_1. \end{cases}$$

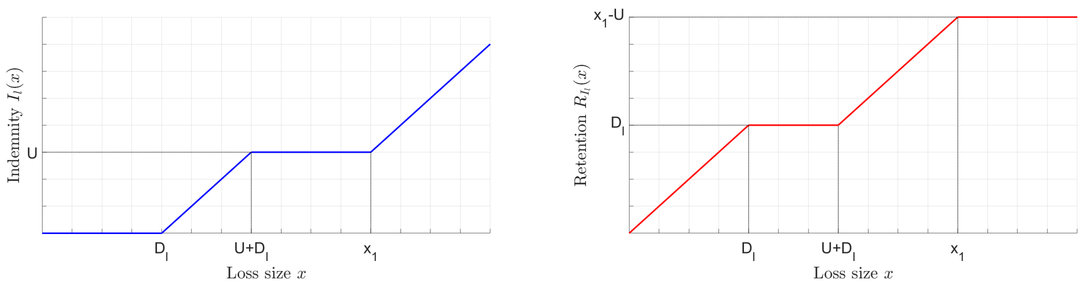
Indemnity schedule I_l provides U dollars of coverage for losses below x_1 subject to a deductible of D_l. For losses above x_1, it provides full insurance above a deductible of $x_1 - U$. Malamud et al. (2016) refer to these segments of the loss distribution as tranches and show how they arise endogenously in an insurance market with multiple providers. As stated in Proposition 6(i), we have $D_l \leq x_1 - U$ so the deductible increases as we go from the loss layer below x_1 to the loss layer above x_1. For case (ii), we have

$$I_r(x; D_r, U) = \begin{cases} 0, & \text{for } x \leq x_1 - U, \\ x - (x_1 - U), & \text{for } x_1 - U < x \leq x_1, \\ U, & \text{for } x_1 < x \leq D_r, \\ x - (D_r - U), & \text{for } x > D_r. \end{cases}$$

Indemnity schedule I_r provides U dollars of coverage for losses below D_r subject to a deductible of $x_1 - U$. For losses above D_r, it provides full insurance above a deductible of $D_r - U$. We have $x_1 \leq D_r$, as stated in Proposition 6(ii), so the deductible increases from $x_1 - U$ to $D_r - U$ as we go from the loss layer below D_r to the loss layer above D_r. In Gollier's (2014) model, losses below and above x_1 are subject to a larger deductible than a loss of x_1. The deductible thus decreases when going from losses below x_1 to a loss of x_1, and then increases when going from a loss of x_1 to losses above x_1. This non-monotonic behavior can lead to violations of the no-sabotage condition in his model, and is thus ruled out in our model.

Figure 1 illustrates indemnity schedule I_l in panel (a) and the associated retained loss function R_{I_l} in panel (b). If we wanted to illustrate I_r and R_{I_r} instead, all we need to do is to change the labels on the x-axis and on the y-axis. We observe that the principle of indemnity is satisfied because $I_l(x)$ is always between zero and x, and that the no-sabotage condition holds because both I_l and R_{I_l} are increasing. Furthermore, we can directly observe Theorem 1 because $I_l'(x)$ is either zero or one except at the critical points D_l, $U + D_l$ and x_1. At those points, the indemnity schedule has kinks and is not differentiable. As a consequence, $R_{I_l}'(x)$ is also either zero or one except at the critical points.

In some special cases, the optimal insurance contract can be simplified to a straight deductible. We provide some sufficient conditions in the following result.



(a) Indemnity schedule $I_l(x; D_l, U)$

(b) Retained loss function $R_{I_l}(x)$

Figure 1: Illustration of indemnity schedule $I_l(x; D_l, U)$ and its associated retained loss function $R_{I_l}(x)$ from Proposition 6(i). To illustrate $I_r(x; D_r, U)$, replace D_l by $x_1 - U$, $U + D_l$ by x_1, and x_1 by D_r on the x-axis. To illustrate R_{I_r}, make the same replacements on the x-axis and replace D_l by U and $x_1 - U$ by $D_r - U$ on the y-axis.

Proposition 7. *Under the one-against-all ambiguity structure, the solution to Problem* (1) *is a straight deductible with the deductible level specified in Eq.* (7) *if one of the following conditions holds:*

(i) $x_1 \leq x_\tau$.

(ii) $x_1 \leq ess \inf X_2$.

(iii) $x_1 \geq ess \sup X_2$.

Appendix A.8 provides the proof. The point is that Gollier (2014) finds a deductible that is V-shaped in the loss and jumps down to a lower level when the loss is x_1. We restrict the set of admissible indemnity schedules by imposing the no-sabotage condition, which rules out indemnity schedules like the one found by Gollier (2014). Instead, we find an indemnity with multiple layers to be optimal, which sometimes collapses to a straight deductible. So while Gollier (2014) shows that a straight deductible is never optimal under the one-against-all ambiguity structure in his model, it may well be optimal in our model.

5.4 Two-state ambiguity

Consider $X_1 = x_1$ and $X_2 = x_2$ in Eq. (8) with $0 \leq x_1 < x_2$. In this case, the distribution of the loss X conditional on $\Theta = \theta$ is given by

$$\mathbb{P}(X \leq x|\Theta = \theta) = (1-p)\mathbb{P}(\hat{X} \leq x) + p(1-\theta)\mathbf{1}_{\{x_1 \leq x\}} + p\theta\mathbf{1}_{\{x_2 \leq x\}} \qquad \text{for } x \geq 0,$$

where we let the distribution of $\hat{X}$ be independent of θ and assume $\mathbb{P}(\hat{X} \in \{x_1, x_2\}) = 0$. There is a $(1-p)$ chance that the loss distribution is that of $\hat{X}$ and a p chance that the loss is either x_1 and x_2. Conditional on the loss being either x_1 or x_2, the probability of a loss of x_1 is $(1-\theta)$ and that of a loss of x_2 is θ. So the two states x_1 and x_2 are ambiguous because their probability of occurrence depends on the value of θ.

The following special case shows that a straight deductible can be optimal under the two-state ambiguity structure.

Corollary 3. *Under the two-state ambiguity structure with $x_1 = 0$ and $x_2 = ess \sup \hat{X}$, the solution to Problem* (1) *is a straight deductible with the deductible level specified in Eq.* (7).

Proof. With $x_1 = 0$ and $x_2 = ess \sup \hat{X}$, the two-state ambiguity structure is a special case of left-against-right ambiguity and condition (9) holds so that Proposition 5 applies. □

In general, however, the solution will not be a straight deductible but will be of the multi-layer form again. This is our next result.

Proposition 8. *Under the two-state ambiguity structure, the solution to Problem* (1) *is a multi-layer contract that takes the following form:*

$$I^*(x) = \min(\max(x - D_l, 0), U_l) + \min(\max(x - D_m, 0), U_m - U_l) + \max(x - D_r, 0). \quad (11)$$

The parameters U_l, U_m, D_l, D_m and D_r are nonnegative and satisfy

$$U_l \leq U_m, \; D_l \leq x_1 \leq D_m \leq x_2 \leq D_r, \; D_l + U_l \leq x_1, \; D_m + U_m - U_l \leq x_2. \quad (12)$$

Appendix A.9 provides the proof. Indemnity schedule I^* has three layers. It provides up to U_l dollars of coverage for losses below D_m subject to a deductible of D_l. For losses between D_m and D_r, indemnity schedule I^* provides U_l dollars from the first layer and up to $U_m - U_l$ in additional coverage for the portion of the loss in the second layer, $x - D_m$. For losses between $D_m + U_m - U_l$ and D_r, the policyholder thus receives $U_l + (U_m - U_l) = U_m$ dollars in indemnification. The third layer starts at D_r and the policyholder receives U_m dollars from the first two layers and the entire portion of the loss that falls into the third layer, $x - D_r$.

In Gollier's (2014) model, losses other than x_1 and x_2 are subject to the same deductible, a loss of x_1 is fully covered, and a loss of x_2 is subject to a smaller deductible than all other losses, see his Proposition 6. So the deductible starts at a positive level, jumps down to zero at x_1, is at the same positive level between x_1 and x_2, jumps down to a smaller positive level at x_2, and then returns to the initial positive level for losses in excess of x_2. The no-sabotage condition rules out indemnity schedules like the one found by Gollier (2014). As in the case of one-against-all ambiguity, we obtain an indemnity of the multi-layer form and the additional ambiguous state adds another layer to the indemnity schedule.

In our model, this argument can be extended to multi-state ambiguity structures with more than two ambiguous states, with each extra state potentially adding another layer. Despite this complexity, we draw the same general conclusion as in the case of one-against-all ambiguity. A straight deductible is never optimal in Gollier's (2014) model but arises as a special case in our analysis, see Corollary 3. The specific ambiguity structures considered in this section substantiate our conclusion that the no-sabotage condition leads to the optimality of a straight deductible in a broader set of circumstances.

6 Comparative statics

This section provides comparative statics. One of the main advantages of Klibanoff et al.'s (2005) smooth ambiguity model is that it achieves a clean separation of tastes and beliefs. We can vary the policyholder's degree of ambiguity aversion while keeping the level of ambiguity fixed, or we can vary the level of ambiguity while keeping the policyholder's degree of ambiguity aversion fixed. Gollier (2014) only looks at the effect of ambiguity aversion at the extensive margin in one specific case. He compares the optimal deductible of a subjective expected utility maximizer (i.e., $\phi'' = 0$) to the optimal deductible of an ambiguity-averse policyholder (i.e., $\phi'' < 0$) when ambiguity is concentrated on losses below x_τ. In this case and under some additional stochastic ordering assumptions on the priors, ambiguity aversion has the counterintuitive effect that it lowers the optimal demand for insurance because the optimal deductible increases so that the policyholder retains more risk.

We first define an increase in the degree of ambiguity aversion. In the field and in laboratory experiments researchers typically observe that the degree of ambiguity aversion varies across people (e.g., Dimmock et al., 2015; Berger and Bosetti, 2020). The smooth ambiguity model allows us to accommodate this observation, and we use the notion of comparative ambiguity aversion from Klibanoff et al. (2005).

Definition 3. A policyholder is *more ambiguity-averse* than another policyholder if they share the same vNM utility function, hold the same beliefs, and if the ambiguity function of the first one, ϕ_1, is more concave than the ambiguity function of the second one, ϕ_2, in the sense of Arrow-Pratt, that is, if

$$-\frac{\phi_1''(z)}{\phi_1'(z)} \geq -\frac{\phi_2''(z)}{\phi_2'(z)} \qquad \text{for any } z \text{ in the domain of } \phi_1 \text{ and } \phi_2.$$

We can then establish the following result.

Proposition 9. *Let X be increasing in Θ by first-order stochastic dominance ($X \uparrow_{fsd} \Theta$) and let Θ be increasing in X in the hazard rate order ($\Theta \uparrow_{hr} X$). Then, an increase in the policyholder's degree of ambiguity aversion lowers the optimal deductible.*

Appendix A.10 provides the proof. The conditions $X \uparrow_{fsd} \Theta$ and $\Theta \uparrow_{hr} X$ are easy to satisfy. They hold if the loss is increasing in Θ in the likelihood ratio order because increasingness in the likelihood ratio order is symmetric, see Appendix B.4, and implies increasingness in the hazard rate order. Section 5.2 shows that the left-against-right ambiguity structure is a special case in which $X \uparrow_{lr} \Theta$ holds. Another example is the one-against-all ambiguity structure with $x_1 = 0$. In this case, ambiguity is concentrated on the probability of the no-loss state and the policyholder is certain about the loss distribution conditional on a loss occurring. Alary et al. (2013) consider this knife-edge case and demonstrate the optimality of a straight deductible (their Proposition 6) and that an increase in the policyholder's degree of ambiguity aversion lowers the optimal deductible (their Proposition 7).

With the help of the no-sabotage condition, Proposition 9 establishes the intuitive result in much more generality. While we do impose stochastic ordering assumptions on the loss X and random variable Θ, we do not require the loss probability to be the only source of ambiguity as Alary et al. (2013) do. Schlesinger (1981) shows that an increase in risk aversion lowers the optimal deductible and thus increases insurance demand under expected utility. In our setting, ambiguity aversion reinforces risk aversion and can thus explain a higher demand for low deductibles compared to expected utility. Conventional models of risk aversion fail to explain the strong preference for low deductibles that researchers have documented in the field (Cohen and Einav, 2007; Sydnor, 2010) and in laboratory experiments (Shapira and Venezia, 2008; Jaspersen et al., 2022b), see also Jaspersen et al. (2022a). It is plausible that people are uncertain about the loss distribution at the time they purchase insurance, and that this ambiguity not only concerns the probability of loss but also its severity. Ambiguity aversion may thus contribute to the well-documented propensity to overinsure modest risks.

In a next step, we answer the dual question of the effect of greater ambiguity. Jewitt and Mukerji (2017) define the relation "more ambiguous." Their definition is technical and hard to operationalize. To simplify the analysis, we use the three-piece ambiguity structure in Eq. (8). With this specification, ambiguity arises from the uncertainty over the mixture weight $\theta \in [0, 1]$. The key simplification lies in the fact that, for any function g, the conditional expectation $\mathbb{E}[g(X)|\Theta]$ is then a linear function of Θ. We pose the following definition.

Definition 4. A policyholder *perceives greater ambiguity* if his second-order belief becomes riskier in the sense of Rothschild and Stiglitz (1970), that is, if random variable Θ_1 changes to Θ_2 with $\Theta_2 \geq_{RS} \Theta_1$.

Snow (2010) introduces this definition of greater ambiguity to study the value of information, Snow (2011) uses it in the context of self-insurance and self-protection, Hoy et al. (2014) apply it to genetic testing decisions under ambiguity, and Peter (2019) utilizes it for the comparative statics of precautionary saving. Under the three-piece ambiguity structure specified in Eq. (8), this definition is consistent with Jewitt and Mukerji's (2017) notion of "more ambiguous (I)" because such a change in beliefs does not affect ambiguity-neutral policyholders but makes every ambiguity-averse policyholder worse off. To formulate our last result, we introduce an intensity measure of the policyholder's prudence in ambiguity.

Definition 5. A policyholder with ambiguity function ϕ is *ambiguity-prudent* if $\phi''' \geq 0$. His *relative ambiguity prudence* is given by $\mathcal{P}(z) = -z\phi'''(z)/\phi''(z)$.

The notion of ambiguity prudence mimics the concept of risk prudence in utility theory. Kimball (1990) coins the term "prudence" for a positive third derivative of the utility function and shows that it is necessary and sufficient for a precautionary savings motive in the discounted expected utility model. Baillon (2017) introduces prudence with respect to ambiguity and shows its equivalence to $\phi''' \geq 0$ in the smooth ambiguity model. Ambiguity prudence matters for the survival of ambiguity-averse agents (Guerdjikova and Sciubba, 2015) and affects precautionary saving (Berger, 2014) and prevention decision under ambiguity (Berger, 2016). Evidence from the laboratory supports that the majority of individual decisions are consistent with ambiguity prudence (Baillon et al., 2018).

We are now in a position to formulate our last result.

Proposition 10. *Consider the three-piece ambiguity structure stated in Eq.* (8)*, assume that* $X_1 \leq_{hr} \hat{X} \leq_{hr} X_2$*, and let the policyholder be ambiguity-prudent with relative ambiguity prudence below two. Then, greater ambiguity lowers the optimal deductible.*

Appendix A.11 provides the proof. Under the stated conditions, we obtain the intuitive result that an increase in the level of ambiguity raises insurance demand. The comparative statics of ambiguity are different from the comparative statics of risk. For an increase in risk, Eeckhoudt et al. (1991) show that the effects depend on the portion of the loss distribution that becomes riskier. If only losses below the deductible are affected and the policyholder is prudent, the optimal deductible increases resulting in lower insurance demand. If only losses above the deductible are affected and the policyholder has decreasing absolute risk aversion, the optimal deductible decreases.

Proposition 10 requires an assumption about the policyholder's intensity of ambiguity prudence. This assumption also arises in the context of portfolio choice under ambiguity (Huang and Tzeng, 2018), precautionary saving (Peter, 2019), and insurance demand under nonperformance risk (Peter and Ying, 2020). The intuition behind this threshold condition comes from the presence of two conflicting effects. Greater ambiguity has a positive substitution effect on insurance demand because the more comprehensively the policyholder is insured, the less she is affected by risk and the associated uncertainty over the loss distribution. At the same time, a policyholder who is ambiguity prudent has a precautionary motive to increase expected final wealth and thereby raise expected utility. One way to increase final wealth is to buy less insurance, which reduces the insurance premium. For those losses that are not too large, the premium savings dominate the reduced indemnification.

One way to mute this precautionary motive is by assuming $\phi''' = 0$, which seems unrealistic. Alternatively, we can allow for ambiguity prudence but restrict its intensity. Berger and Bosetti (2020) conduct an experimental test of ambiguity attitudes using model uncertainty. They fit the ϕ-function of the smooth model and find that a power form fits better than an expo-power form. Relative ambiguity aversion is 0.53 in their study with a standard error of 0.0261. This estimate implies relative ambiguity prudence of 1.53.[6] As such, the results in Berger and Bosetti (2020) support our assumption of $\mathcal{P} \leq 2$ for most, if not all subjects in their data. Admittedly more research is needed on the intensity of ambiguity prudence, and it would be helpful to have some evidence from the field.

[6] If ϕ takes the power form, say $\phi(z) = (z^{1-\gamma} - 1)/(1-\gamma)$ for $\gamma > 0$, then relative ambiguity aversion is given by γ and relative ambiguity prudence by $1 + \gamma$. An estimate of $\gamma = 0.53$ thus implies relative ambiguity prudence of 1.53, which is less than 2.

7 Conclusion

In this paper, we revisited the problem of optimal insurance design for ambiguity-averse policyholders under Klibanoff et al.'s (2005) smooth ambiguity model. Unlike previous literature, we used a flexible ambiguity structure that is not restricted to a finite set of priors and imposed both the principle of indemnity and the no-sabotage condition on the set of admissible indemnity schedules. We characterized the optimal indemnity schedule by showing that its slope is either zero or one except at critical points. We also derived a sufficient condition for a straight deductible to be optimal, thus extending Arrow's cornerstone result to smooth ambiguity aversion. This condition is satisfied when ambiguity does not affect large losses (above x_τ), when the loss X and random variable Θ are increasing or decreasing in each other in the FSD sense, or when the loss X is monotonic in Θ in the hazard rate order. Our characterization is substantially more general than previous literature and shows that a straight deductible is optimal in a broad set of circumstances.

We discussed several specific ambiguity structures to isolate the effect of the no-sabotage condition. A straight deductible is optimal under left-against-right ambiguity and may be optimal under one-against-all and two-state ambiguity. In the last two cases, the optimal contract is of the multi-layer form, and the number of ambiguous states determines how many layers are possible. Our model yields intuitive comparative statics. Previous research either presupposes that ambiguity is absent conditional on the loss (Proposition 7 in Alary et al., 2013) or finds counterintuitive effects (Proposition 9 in Gollier, 2014). Under some simple stochastic ordering assumptions, greater ambiguity aversion always lowers the optimal deductible and hence raises insurance demand. Ambiguity aversion can thus help explain the overinsurance puzzle of modest risks (Sydnor, 2010). We also studied the comparative statics of greater ambiguity and find that it raises insurance demand for ambiguity-prudent policyholders with relative ambiguity prudence less than two, which seems empirically plausible.

Gollier (2014) was the first to study optimal insurance design under smooth ambiguity aversion and we only come second. We trade off a smaller contract space by imposing the no-sabotage condition for a more flexible ambiguity structure that includes, for example, the case of parameter uncertainty. While ambiguity and ambiguity aversion certainly affect the optimal level of risk transfer, Arrow's (1963) famous result that a straight deductible is optimal holds broadly in our setting by virtue of the no-sabotage condition.

References

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

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

Arrow, K. J. (1965). Aspects of the theory of risk-bearing. *Helsinki: Yrjö Jahnsson Foundation*.

Arrow, K. J. (1971). The theory of risk aversion. *Essays in the theory of risk-bearing*, pages 90–120.

Arrow, K. J. (1974). Optimal insurance and generalized deductibles. *Scandinavian Actuarial Journal*, 1974(1): 1–42.

Athey, S. (2002). Monotone comparative statics under uncertainty. *The Quarterly Journal of Economics*, 117(1): 187–223.

Baillon, A. (2017). Prudence with respect to ambiguity. *The Economic Journal*, 127(604): 1731–1755.

Baillon, A. and Emirmahmutoglu, A. (2018). Zooming in on ambiguity attitudes. *International Economic Review*, 59(4): 2107–2131.

Baillon, A., Schlesinger, H., and van de Kuilen, G. (2018). Measuring higher order ambiguity preferences. *Experimental Economics*, 21(2): 233–256.

Berger, L. (2014). Precautionary saving and the notion of ambiguity prudence. *Economics Letters*, 123(2): 248–251.

Berger, L. (2016). The impact of ambiguity and prudence on prevention decisions. *Theory and Decision*, 80(3): 389–409.

Berger, L. and Bosetti, V. (2020). Characterizing ambiguity attitudes using model uncertainty. *Journal of Economic Behavior & Organization*, 180: 621–637.

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

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

Borch, K. (1962). Equilibrium in a reinsurance market. *Econometrica*, 30(3): 424–444.

Cai, J. and Wei, W. (2012). On the invariant properties of notions of positive dependence and copulas under increasing transformations. *Insurance: Mathematics and Economics*, 50(1): 43–49.

Camerer, C. and Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, 5(4): 325–370.

Carlier, G., Dana, R. A., and Shahidi, N. (2003). Efficient insurance contracts under epsilon-contaminated utilities. *The Geneva Papers on Risk and Insurance Theory*, 28(1): 59–71.

Chesson, H. and Viscusi, W. (2003). Commonalities in time and ambiguity aversion for long-term risks. *Theory and Decision*, 54(1): 57–71.

Chi, Y. and Tan, K. S. (2021). Optimal incentive-compatible insurance with background risk. *ASTIN Bulletin*, 51(2): 661–688.

Chi, Y. and Wei, W. (2018). Optimum insurance contracts with background risk and higher-order risk attitudes. *ASTIN Bulletin*, 48(3): 1025–1047.

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

Chiappori, P.-A., Jullien, B., Salanié, B., and Salanie, F. (2006). Asymmetric information in insurance: General testable implications. *The RAND Journal of Economics*, 37(4): 783–798.

Chow, C. and Sarin, R. (2001). Comparative ignorance and the Ellsberg paradox. *Journal of Risk and Uncertainty*, 22(2): 129–139.

Cohen, A. and Einav, L. (2007). Estimating risk preferences from deductible choice. *American Economic Review*, 97(3): 745–788.

Constantinides, G. M. (1982). Intertemporal asset pricing with heterogeneous consumers and without demand aggregation. *Journal of Business*, 55(2): 253–267.

Crainich, D., Eeckhoudt, L., and Menegatti, M. (2016). Changing risks and optimal effort. *Journal of Economic Behavior & Organization*, 125: 97–106.

Cubitt, R., Van De Kuilen, G., and Mukerji, S. (2020). Discriminating between models of ambiguity attitude: A qualitative test. *Journal of the European Economic Association*, 18(2): 708–749.

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

Dimmock, S., Kouwenberg, R., Mitchell, O., and Peijnenburg, K. (2016). Ambiguity aversion and household portfolio choice puzzles: Empirical evidence. *Journal of Financial Economics*, 119(3): 559–577.

Dimmock, S. G., Kouwenberg, R., Mitchell, O. S., and Peijnenburg, K. (2015). Estimating ambiguity preferences and perceptions in multiple prior models: Evidence from the field. *Journal of Risk and Uncertainty*, 51(3): 219–244.

Dow, J. and da Costa Werlang, S. R. (1992). Uncertainty aversion, risk aversion, and the optimal choice of portfolio. *Econometrica*, 60(1): 197–204.

Eeckhoudt, L., Gollier, C., and Schlesinger, H. (1991). Increases in risk and deductible insurance. *Journal of Economic Theory*, 55(2): 435–440.

Eeckhoudt, L. and Kimball, M. (1992). Background risk, prudence, and the demand for insurance. In *Contributions to Insurance Economics*, pages 239–254. Springer.

Einhorn, H. and Hogarth, R. (1986). Decision making under ambiguity. *Journal of Business*, 59(4): 225–250.

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

Epstein, L. G. (2010). A paradox for the "smooth ambiguity" model of preference. *Econometrica*, 78(6): 2085–2099.

Fafchamps, M. and Gubert, F. (2007). The formation of risk sharing networks. *Journal of Development Economics*, 83(2): 326–350.

Fafchamps, M. and Lund, S. (2003). Risk-sharing networks in rural Philippines. *Journal of Development Economics*, 71(2): 261–287.

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

Ghossoub, M. (2017). Arrow's theorem of the deductible with heterogeneous beliefs. *North American Actuarial Journal*, 21(1): 15–35.

Ghossoub, M. (2019a). Budget-constrained optimal insurance without the nonnegativity constraint on indemnities. *Insurance: Mathematics and Economics*, 84: 22–39.

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

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

Gollier, C. (1987). The design of optimal insurance contracts without the nonnegativity constraint on claims. *Journal of Risk and Insurance*, 54(2): 314–324.

Gollier, C. (1996). Optimum insurance of approximate losses. *Journal of Risk and Insurance*, 63(3): 369–380.

Gollier, C. (2011). Portfolio choices and asset prices: The comparative statics of ambiguity aversion. *The Review of Economic Studies*, 78(4): 1329–1344.

Gollier, C. (2013). The economics of optimal insurance design. In Dionne, G., editor, *Handbook of Insurance*, chapter 4, pages 107–122. Springer.

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

Gollier, C. and Schlesinger, H. (1996). Arrow's theorem on the optimality of deductibles: a stochastic dominance approach. *Economic Theory*, 7(2): 359–363.

Guerdjikova, A. and Sciubba, E. (2015). Survival with ambiguity. *Journal of Economic Theory*, 155: 50–94.

Hoy, M., Peter, R., and Richter, A. (2014). Take-up for genetic tests and ambiguity. *Journal of Risk and Uncertainty*, 48(2): 111–133.

Huang, Y.-C. and Tzeng, L. Y. (2018). A mean-preserving increase in ambiguity and portfolio choices. *Journal of Risk and Insurance*, 85(4): 993–1012.

Huberman, G., Mayers, D., and Smith Jr, C. W. (1983). Optimal insurance policy indemnity schedules. *The Bell Journal of Economics*, 14(2): 415–426.

Jaspersen, J. G., Peter, R., and Ragin, M. A. (2022a). Probability weighting and insurance demand in a unified framework. *The Geneva Risk and Insurance Review (forthcoming)*.

Jaspersen, J. G., Ragin, M. A., and Sydnor, J. R. (2022b). Predicting insurance demand from risk attitudes. *Journal of Risk and Insurance*, 89(1): 63–96.

Jewitt, I. and Mukerji, S. (2017). Ordering ambiguous acts. *Journal of Economic Theory*, 171: 213–267.

Jindapon, P. and Neilson, W. S. (2007). Higher-order generalizations of Arrow–Pratt and Ross risk aversion: A comparative statics approach. *Journal of Economic Theory*, 136(1): 719–728.

Karni, E. (1992). Optimal insurance: A nonexpected utility analysis. *Contributions to Insurance Economics*, pages 217–238.

Kimball, M. S. (1990). Precautionary saving in the small and in the large. *Econometrica*, 58(1): 53–73.

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

Klibanoff, P., Marinacci, M., and Mukerji, S. (2012). On the smooth ambiguity model: A reply. *Econometrica*, 80(3): 1303–1321.

Kocher, M., Lahno, A., and Trautmann, S. (2018). Ambiguity aversion is not universal. *European Economic Review*, 101: 268–283.

Lang, M. (2017). First-order and second-order ambiguity aversion. *Management Science*, 63(4): 1254–1269.

Liu, L. and Neilson, W. S. (2019). Alternative approaches to comparative nth-degree risk aversion. *Management Science*, 65(8): 3824–3834.

Machina, M. J. (1995). Non-expected utility and the robustness of the classical insurance paradigm. *The Geneva Papers on Risk and Insurance Theory*, 20(1): 9–50.

Malamud, S., Rui, H., and Whinston, A. (2016). Optimal reinsurance with multiple tranches. *Journal of Mathematical Economics*, 65: 71–82.

Marshall, J. M. (1992). Optimum insurance with deviant beliefs. In *Contributions to Insurance Economics*, pages 255–274. Springer.

Mayers, D. and Smith Jr, C. W. (1983). The interdependence of individual portfolio decisions and the demand for insurance. *Journal of Political Economy*, 91(2): 304–311.

Mossin, J. (1968). Aspects of rational insurance purchasing. *Journal of Political Economy*, 46(4): 553–568.

Mukerji, S. and Tallon, J.-M. (2001). Ambiguity aversion and incompleteness of financial markets. *The Review of Economic Studies*, 68(4): 883–904.

Neilson, W. S. (2010). A simplified axiomatic approach to ambiguity aversion. *Journal of Risk and Uncertainty*, 41(2): 113–124.

Ohlin, J. (1969). On a class of measures of dispersion with application to optimal reinsurance. *ASTIN Bulletin*, 5(2): 249–266.

Peter, R. (2019). Revisiting precautionary saving under ambiguity. *Economics Letters*, 174: 123–127.

Peter, R. and Ying, J. (2020). Do you trust your insurer? Ambiguity about contract nonperformance and optimal insurance demand. *Journal of Economic Behavior & Organization*, 180: 938–954.

Raviv, A. (1979). The design of an optimal insurance policy. *American Economic Review*, 69(1): 84–96.

Rogerson, W. P. (1985a). The first-order approach to principal-agent problems. *Econometrica*, 53(6): 1357–1367.

Rogerson, W. P. (1985b). Repeated moral hazard. *Econometrica*, 53(1): 69–76.

Rothschild, M. and Stiglitz, J. E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory*, 2(3): 225–243.

Rubinstein, M. (1974). An aggregation theorem for securities markets. *Journal of Financial Economics*, 1(3): 225–244.

Sarin, R. and Weber, M. (1993). Effects of ambiguity in market experiments. *Management Science*, 39(5): 602–615.

Schlesinger, H. (1981). The optimal level of deductibility in insurance contracts. *Journal of Risk and Insurance*, (3): 465–481.

Schlesinger, H. (2013). The theory of insurance demand. In Dionne, G., editor, *Handbook of Insurance*, chapter 7, pages 167–184. Springer, 2 edition.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, 57(3): 571–587.

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic orders*. Springer Science & Business Media.

Shanthikumar, J. G. and Yao, D. D. (1991). Bivariate characterization of some stochastic order relations. *Advances in Applied Probability*, 23(3): 642–659.

Shapira, Z. and Venezia, I. (2008). On the preference for full-coverage policies: Why do people buy too much insurance? *Journal of Economic Psychology*, 29(5): 747–761.

Snow, A. (2010). Ambiguity and the value of information. *Journal of Risk and Uncertainty*, 40(2): 133–145.

Snow, A. (2011). Ambiguity aversion and the propensities for self-insurance and self-protection. *Journal of Risk and Uncertainty*, 42(1): 27–43.

Sydnor, J. (2010). (Over)insuring modest risks. *American Economic Journal: Applied Economics*, 2(4): 177–99.

Townsend, R. M. (1994). Risk and insurance in village India. *Econometrica*, 62(3): 539–591.

Viscusi, W. and Chesson, H. (1999). Hopes and fears: The conflicting effects of risk ambiguity. *Theory and Decision*, 47(2): 157–184.

Wilson, R. (1968). The theory of syndicates. *Econometrica*, 36(1): 119–132.

Xu, Z. Q., Zhou, X. Y., and Zhuang, S. C. (2019). Optimal insurance under rank-dependent utility and incentive compatibility. *Mathematical Finance*, 29(2): 659–692.

Young, V. R. (1999). Optimal insurance under Wang's premium principle. *Insurance: Mathematics and Economics*, 25(2): 109–122.

Zilcha, I. and Chew, S. H. (1990). Invariance of the efficient sets when the expected utility hypothesis is relaxed. *Journal of Economic Behavior & Organization*, 13(1): 125–131.

A Proofs

A.1 Proof of Theorem 1

"Necessity." Let I^* be an optimal solution to Problem (1). For any $I \in \mathfrak{C}$ and $p \in [0,1]$, denote by $\tilde{I}_p(x) = pI^*(x) + (1-p)I(x)$ the convex combination of I^* and I for $x \geq 0$. Then, $\tilde{I}_p \in \mathfrak{C}$ because I^* and I are in $\mathfrak{C}$. Furthermore, we obtain

$$\begin{aligned}\frac{\partial}{\partial p}\mathbb{E}[u(W_{\tilde{I}_p})|\Theta]\Big|_{p=1} &= \mathbb{E}[u'(W_{I^*})(I^*(X) - I(X) - (1+\tau)(\mathbb{E}[I^*(X)] - \mathbb{E}[I(X)]))|\Theta] \\ &= \int_0^\infty \mathbb{E}[u'(W_{I^*})(\mathbf{1}_{\{X>x\}} - (1+\tau)S_X(x))|\Theta](I^{*\prime}(x) - I'(x))\,\mathrm{d}x,\end{aligned}$$

where $\mathbf{1}_A$ denotes the indicator function of event A. The first equality holds because Assumption A4 allows us to exchange the operations of expectation and differentiation. The second equality follows from the fundamental theorem of calculus because

$$I(x) = I(x) - I(0) = \int_0^x I'(y)\,\mathrm{d}y = \int_0^\infty I'(y)\mathbf{1}_{\{x>y\}}\,\mathrm{d}y,$$

and likewise for I^*. The optimality of I^* implies

$$\frac{\partial}{\partial p}\mathbb{E}[\phi(\mathbb{E}[u(W_{\tilde{I}_p})|\Theta])]\Big|_{p=1} \geq 0,$$

which is equivalent to

$$\begin{aligned}0 &\leq \mathbb{E}\left[\phi'(\mathbb{E}[u(W_{I^*})|\Theta])\frac{\partial}{\partial p}\mathbb{E}[u(W_{\tilde{I}_p})|\Theta]\Big|_{p=1}\right] \\ &= \int_0^\infty \mathbb{E}[\phi'(\mathbb{E}[u(W_{I^*})|\Theta])\mathbb{E}[u'(W_{I^*})(\mathbf{1}_{\{X>x\}} - (1+\tau)S_X(x))|\Theta]](I^{*\prime}(x) - I(x))\,\mathrm{d}x \\ &= \int_0^\infty \mathbb{E}[\phi'(\mathbb{E}[u(W_{I^*})|\Theta])u'(W_{I^*})(\mathbf{1}_{\{X>x\}} - (1+\tau)S_X(x))](I^{*\prime}(x) - I(x))\,\mathrm{d}x \\ &= \mathbb{E}[\phi'(\mathbb{E}[u(W_{I^*})|\Theta])u'(W_{I^*})]\int_0^{\mathcal{M}} S_X(x)(V_{I^*}(x) - (1+\tau))(I^{*\prime}(x) - I'(x))\,\mathrm{d}x.\end{aligned}$$

The first equality holds by substituting $\frac{\partial}{\partial p}\mathbb{E}[u(W_{\tilde{I}_p})|\Theta]\Big|_{p=1}$ from above and by exchanging the operations of expectation and integration, which is justified by Assumption A4. The second equality holds by the law of total expectation. The third equality uses the definition of auxiliary function $V_{I^*}(x)$ in Eq. (5). Since the above inequality holds for any $I \in \mathfrak{C}$, property (6) in Theorem 1 must be satisfied.

"Sufficiency."Assume that I^* satisfies property (6). Since both ϕ and u are increasing and concave, then, for any $I \in \mathfrak{C}$, it holds that

$$\mathbb{E}[\phi(\mathbb{E}[u(W_{I^*})|\Theta])] - \mathbb{E}[\phi(\mathbb{E}[u(W_I)|\Theta])]$$

$$
\begin{aligned}
&\geq \mathbb{E}[\phi'(\mathbb{E}[u(W_{I^*})|\Theta])(\mathbb{E}[u(W_{I^*})|\Theta] - \mathbb{E}[u(W_I)|\Theta])] \\
&\geq \mathbb{E}[\phi'(\mathbb{E}[u(W_{I^*})|\Theta])\mathbb{E}[u'(W_{I^*})(W_{I^*} - W_I)|\Theta]] \\
&= \int_0^\infty \mathbb{E}[\phi'(\mathbb{E}[u(W_{I^*})|\Theta])u'(W_{I^*})(\mathbf{1}_{\{X>x\}} - (1+\tau)S_X(x))](I^{*\prime}(x) - I'(x))\,\mathrm{d}x.
\end{aligned}
$$

Consequently, if I^* satisfies (6), then $\mathbb{E}[\phi(\mathbb{E}[u(W_{I^*})|\Theta])] \geq \mathbb{E}[\phi(\mathbb{E}[u(W_I)|\Theta])]$ for all $I \in \mathfrak{C}$, which makes I^* an optimal solution to Problem (1).

A.2 Proof of Theorem 2

We proceed in two steps and first show the result in case of $D^* = \infty$ and then for $D^* < \infty$. If $D^* = \infty$, the set $\{D \geq x_\tau : V_{I_D}(D) \geq 1+\tau\}$ is empty and hence $V_{I_D}(D) < 1+\tau$ for all $D \in [x_\tau, \mathcal{M})$. We assume $V_{I_D}(x)$ to be increasing in x on $[x_\tau, \mathcal{M})$ for a given $D \geq x_\tau$. Therefore, it holds that

$$
V_{I_D}(x) \leq V_{I_D}(D) < 1+\tau \qquad \text{for any } D \geq x.
$$

Consequently, we have

$$
V_{I_\mathcal{M}}(x) = \lim_{D\to\mathcal{M}} V_{I_D}(x) \leq 1+\tau \qquad \text{for any } x \in [x_\tau, \mathcal{M}).
$$

For $x \in [0, x_\tau)$, we have $S_X(x) > 1/(1+\tau)$ and therefore

$$
\begin{aligned}
V_{I_\mathcal{M}}(x) &= \frac{\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_\mathcal{M}})|\Theta])u'(W_{I_\mathcal{M}})|X>x]}{\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_\mathcal{M}})|\Theta])u'(W_{I_\mathcal{M}})]} \\
&= \frac{1}{S_X(x)} \cdot \frac{\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_\mathcal{M}})|\Theta])u'(W_{I_\mathcal{M}})] \cdot \mathbf{1}_{\{X>x\}}}{\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_\mathcal{M}})|\Theta])u'(W_{I_\mathcal{M}})]} < 1+\tau.
\end{aligned}
$$

As a result, $V_{I_\mathcal{M}} \leq 1+\tau$ holds for all $x \in [0, \mathcal{M})$, and $I_\mathcal{M}$ is thus the solution to Problem (1) according to Theorem 1. Indemnity schedule $I_\mathcal{M}$ represents the no-insurance strategy, or equivalently, the straight deductible contract with deductible level of $D = \infty$. This proves the desired conclusion in case of $D^* = \infty$.

Now assume $D^* \in [x_\tau, \mathcal{M})$. We then have $V_{I_D}(D) < 1+\tau$ for any $D \in [x_\tau, D^*)$, and thus $\lim_{D\uparrow D^*} V_{I_D}(D) \leq 1+\tau$. Furthermore, we have $V_{I_{D^*}}(D^*) \geq 1+\tau$ due to the right continuity of $V_{I_D}(D)$ in D, see Lemma 4(iv) in Appendix B.1. Consider the function $V_{I_{D^*}}(x)$, which is assumed to be increasing in $x \in [x_\tau, \mathcal{M})$. We then have

$$
V_{I_{D^*}}(x) \geq V_{I_{D^*}}(D^*) \geq 1+\tau \qquad \text{for any } x \geq D^*,
$$

and

$$
V_{I_{D^*}}(x) = \lim_{\substack{D\uparrow D^* \\ D>x}} V_{I_D}(x) \leq \lim_{\substack{D\uparrow D^* \\ D>x}} V_{I_D}(D) \leq 1+\tau \qquad \text{for any } x \in [x_\tau, D^*).
$$

The equality is due to the continuity of $V_{I_D}(x)$ in D and the first inequality is due to the increasingness of $V_{I_D}(x)$ in x.

We are now in a position to apply Theorem 1. We have shown that $V_{I_{D^*}}(x) > 1+\tau$ can only occur for $x \geq D^*$. But for $x \in (D^*, \mathcal{M})$, we obtain $I'_{D^*}(x) = 1$. Furthermore, $V_{I_{D^*}}(x) < 1+\tau$ can only occur for $x \leq D^*$. But for $x \in [0, D^*)$, we have $I'_{D^*}(x) = 0$. Property (6) is thus satisfied, which makes the straight deductible contract with indemnity schedule I_{D^*} the optimal solution to Problem (1)

A.3 Proof of Proposition 3

According to Theorem 2, it suffices to show that $V_{I_D}(x)$ is increasing in $x \in [x_\tau, \mathcal{M})$ for any $D \geq x_\tau$ under each of the conditions (*i*) to (*iii*).

Assume that condition (*i*) holds and let $I \in \mathfrak{C}$. Then, for any $x \geq x_\tau$, we obtain

$$\frac{\mathbb{E}[u'(W_I)\mathbf{1}_{\{X>x\}}|\Theta]}{\mathbb{P}(X > x_\tau|\Theta)} = \frac{\mathbb{E}[u'(W_I)\mathbf{1}_{\{X>x\}}]}{\mathbb{P}(X > x_\tau)} \tag{13}$$

because $\mathbb{P}(X > x|\Theta = \theta)/\mathbb{P}(X > x_\tau|\Theta = \theta)$ is independent of θ. Rewrite auxiliary function $V_I(x)$ as follows:

$$V_I(x) \quad = \quad \frac{\mathbb{E}[\phi'(\mathbb{E}[u(W_I)|\Theta])u'(W_I)|X > x]}{\mathbb{E}[\phi'(\mathbb{E}[u(W_I)|\Theta])u'(W_I)]} = \frac{\mathbb{E}[\phi'(\mathbb{E}[u(W_I)|\Theta]) \cdot \mathbb{E}[u'(W_I)\mathbf{1}_{\{X>x\}}|\Theta]]}{\mathbb{P}(X > x) \cdot \mathbb{E}[\phi'(\mathbb{E}[u(W_I)|\Theta])u'(W_I)]}.$$

Substituting Eq. (13) into the expression for $V_I(x)$ yields

$$V_I(x) \quad = \quad \frac{\mathbb{E}[\phi'(\mathbb{E}[u(W_I)|\Theta]) \cdot \mathbb{P}(X > x_\tau|\Theta)]}{\mathbb{P}(X > x_\tau) \cdot \mathbb{E}[\phi'(\mathbb{E}[u(W_I)|\Theta])u'(W_I)]} \cdot \frac{\mathbb{E}[u'(W_I)\mathbf{1}_{\{X>x\}}]}{\mathbb{P}(X > x)}.$$

The first factor does not depend on x. The second factor simplifies to $\mathbb{E}[u'(W_I)|X > x)]$, which is an increasing function of x. Indeed, the larger the realization of X, the smaller the value of W_I because the policyholder's retained loss function is increasing due to the no-sabotage condition A2. Marginal utility is decreasing due to risk aversion, which then makes $u'(W_I)$ increasing in X. Therefore, $V_I(x)$ is increasing in $x \in [x_\tau, \mathcal{M})$ for any $I \in \mathfrak{C}$. It follows that $V_{I_D}(x)$ is increasing in $x \in [x_\tau, \mathcal{M})$ for any $D \geq x_\tau$, and Theorem 2 applies.

Assume that condition (*ii*) holds. Then, $W_{I_D} = w - X + \max(0, X-D) - (1+\tau)\mathbb{E}[I_D(X)]$ is a decreasing function of X because I_D satisfies the no-sabotage condition. When $X \uparrow_{fsd} \Theta$, we have that $\mathbb{E}[u(W_{I_D})|\Theta]$ is a decreasing function of Θ because utility function u is increasing. The ambiguity function ϕ is concave due to ambiguity aversion, which makes ϕ' decreasing. Therefore, $\phi'(\mathbb{E}[u(W_{I_D})|\Theta])$ is an increasing function of Θ. When coupled with $\Theta \uparrow_{fsd} X$, this yields that $\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_D})|\Theta])|X]$ is an increasing function of X. Risk aversion is represented by a concave utility function so that u' is decreasing. This makes $u'(W_{I_D})$ an increasing function of X. Hence, assuming $X \uparrow_{fsd} \Theta$ and $\Theta \uparrow_{fsd} X$ ensures that

$$V_{I_D}(x) = \frac{\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_D})|\Theta])u'(W_{I_D})|X > x]}{\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_D})|\Theta])u'(W_{I_D})]} = \frac{\mathbb{E}[\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_D})|\Theta])|X]u'(W_{I_D})|X > x]}{\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_D})|\Theta])u'(W_{I_D})]}$$

is increasing in x, and Theorem 2 applies. When $X \downarrow_{fsd} \Theta$ instead, then $\mathbb{E}[u(W_{I_D})|\Theta]$ is an increasing function of Θ and $\phi'(\mathbb{E}[u(W_{I_D})|\Theta])$ is a decreasing function of Θ. Coupled with $\Theta \downarrow_{fsd} X$, we have that $\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_D})|\Theta])|X]$ is an increasing function of X, and the rest of the argument is identical.

Assume that condition (iii) holds and let $X \uparrow_{hr} \Theta$. Increasingness in the hazard rate order implies increasingness in the FSD order, $X \uparrow_{fsd} \Theta$. Then we know from the proof of (ii) that $u'(W_{I_D})$ is an increasing function of X, that $\mathbb{E}[u(W_{I_D})|\Theta]$ is a decreasing function of Θ, and that $\phi'(\mathbb{E}[u(W_{I_D})|\Theta])$ is an increasing function of Θ. Both $u'(W_{I_D})$ and $\phi'(\mathbb{E}[u(W_{I_D})|\Theta])$ are positive. Since $X \uparrow_{hr} \Theta$, Lemma 5(i) in Appendix B.2 implies that $[\Theta|X > x]$ increases in x in the likelihood ratio order, and thus also in the FSD order. Lemma 6 in Appendix B.3 then implies that $V_{I_D}(x)$ is increasing in x, which completes the proof.

A.4 Proof of Proposition 4

We prove that $I'_{D_0^*}$ satisfies criterion (6) in Theorem 1 under the assumptions made. The criterion involves auxiliary function $V_{I_{D_0^*}}(x)$. When ambiguity is concentrated on losses above D_0^*, the distribution of $W_{I_{D_0^*}} = w - \min(X, D_0^*) - (1+\tau)\mathbb{E}[I_{D_0^*}(X)]$ conditional on Θ is independent of Θ. Therefore, $\mathbb{E}[u(W_{I_{D_0^*}})|\Theta]$ is a deterministic quantity. We then obtain

$$\begin{aligned} V_{I_{D_0^*}}(x) &= \frac{\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_{D_0^*}})|\Theta])u'(W_{I_{D_0^*}})|X > x]}{\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_{D_0^*}})|\Theta])u'(W_{I_{D_0^*}})]} \\ &= \frac{\phi'(\mathbb{E}[u(W_{I_{D_0^*}})]) \cdot \mathbb{E}[u'(W_{I_{D_0^*}})|X > x]}{\phi'(\mathbb{E}[u(W_{I_{D_0^*}})]) \cdot \mathbb{E}[u'(W_{I_{D_0^*}})]} = \frac{\mathbb{E}[u'(W_{I_{D_0^*}})|X > x]}{\mathbb{E}[u'(W_{I_{D_0^*}})]}. \end{aligned}$$

Since $I_{D_0^*}$ is a solution to Problem (2), which is the special case of Problem (1) for $\phi(x) = x$, it follows that

$$I'_{D_0^*}(x) = \begin{cases} 1, & \text{for } V_{I_{D_0^*}}(x) > 1+\tau; \\ 0, & \text{for } V_{I_{D_0^*}}(x) < 1+\tau. \end{cases}$$

Therefore, Theorem 1 establishes that $I_{D_0^*}$ is already the solution to Problem (1) for any increasing and concave ϕ-function.

A.5 Proof of Lemma 2

To prove $X \uparrow_{lr} \Theta$, we need to show that $[X|\Theta = \theta_1] \leq_{lr} [X|\Theta = \theta_2]$ for any $\theta_1 \leq \theta_2$. Let Y_1 and Y_2 be independent random variables with the same distribution as $[X|\Theta = \theta_1]$ and

$[X|\Theta = \theta_2]$. According to Shanthikumar and Yao's (1991) bivariate functional characterization of the likelihood ratio order in their Theorem 2.3, we obtain $Y_1 \leq_{lr} Y_2$ if and only if $\mathbb{E}[g(Y_1, Y_2)] \geq \mathbb{E}[g(Y_2, Y_1)]$ for any arrangement-increasing function $g(x, y)$, where a function is called arrangement-increasing if $g(x, y) \geq g(y, x)$ for any $x \leq y$.

Let $(\hat{X}^{\perp}, X_1^{\perp}, X_2^{\perp})$ be an independent copy of $(\hat{X}, X_1, X_2)$. Due to Eq. (8), we can represent Y_1 as a mixture of $\hat{X}, X_1$ and X_2 with probability weights $(1-p)$, $p(1-\theta_1)$ and $p\theta_1$, and Y_2 as a mixture of $\hat{X}^{\perp}, X_1^{\perp}$ and $X_2^{\perp}$ with probability weights $(1-p)$, $p(1-\theta_2)$ and $p\theta_2$. By direct computation, we obtain

$$\begin{aligned}
\mathbb{E}[g(Y_1, Y_2)] \quad = \quad & (1-p)^2\mathbb{E}[g(\hat{X}, \hat{X}^{\perp})] + (1-p)p(1-\theta_2)\mathbb{E}[g(\hat{X}, X_1^{\perp})] + (1-p)p\theta_2\mathbb{E}[g(\hat{X}, X_2^{\perp})] \\
& +(1-p)p(1-\theta_1)\mathbb{E}[g(X_1, \hat{X}^{\perp})] + p^2(1-\theta_1)(1-\theta_2)\mathbb{E}[g(X_1, X_1^{\perp})] \\
& +p^2(1-\theta_1)\theta_2\mathbb{E}[g(X_1, X_2^{\perp})] + (1-p)p\theta_1\mathbb{E}[g(X_2, \hat{X}^{\perp})] \\
& +p^2(1-\theta_2)\theta_1\mathbb{E}[g(X_2, X_1^{\perp})] + p^2\theta_1\theta_2\mathbb{E}[g(X_2, X_2^{\perp})],
\end{aligned}$$

and similarly,

$$\begin{aligned}
\mathbb{E}[g(Y_2, Y_1)] \quad = \quad & (1-p)^2\mathbb{E}[g(\hat{X}^{\perp}, \hat{X})] + (1-p)p(1-\theta_1)\mathbb{E}[g(\hat{X}^{\perp}, X_1)] + (1-p)p\theta_1\mathbb{E}[g(\hat{X}^{\perp}, X_2)] \\
& +(1-p)p(1-\theta_2)\mathbb{E}[g(X_1^{\perp}, \hat{X})] + p^2(1-\theta_2)(1-\theta_1)\mathbb{E}[g(X_1^{\perp}, X_1)] \\
& +p^2(1-\theta_2)\theta_1\mathbb{E}[g(X_1^{\perp}, X_2)] + (1-p)p\theta_2\mathbb{E}[g(X_2^{\perp}, \hat{X})] \\
& +p^2(1-\theta_1)\theta_2\mathbb{E}[g(X_2^{\perp}, X_1)] + p^2\theta_2\theta_1\mathbb{E}[g(X_2^{\perp}, X_2)].
\end{aligned}$$

We have $\mathbb{E}[g(\hat{X}, \hat{X}^{\perp})] = \mathbb{E}[g(\hat{X}^{\perp}, \hat{X})]$ because $\hat{X}^{\perp}$ is an independent copy of $\hat{X}$. Furthermore, $ess \sup X_1 \leq ess \inf \hat{X}$ implies $X_1 \leq_{lr} \hat{X}$ so that

$$\mathbb{E}[g(X_1, \hat{X}^{\perp})] - \mathbb{E}[g(\hat{X}^{\perp}, X_1)] = \mathbb{E}[g(X_1^{\perp}, \hat{X})] - \mathbb{E}[g(\hat{X}, X_1^{\perp})] \geq 0.$$

The equality holds because $X_1^{\perp}$ and $\hat{X}^{\perp}$ are independent copies of X_1 and $\hat{X}$. The inequality follows from $X_1^{\perp} \leq_{lr} \hat{X}$ and Shanthikumar and Yao's (1991) Theorem 2.3. Since $\theta_1 \leq \theta_2$, we have $(1-p)p(1-\theta_2) \leq (1-p)p(1-\theta_1)$, and thus

$$\begin{aligned}
& (1-p)p(1-\theta_2)\mathbb{E}[g(\hat{X}, X_1^{\perp})] + (1-p)p(1-\theta_1)\mathbb{E}[g(X_1, \hat{X}^{\perp})] \\
\geq \quad & (1-p)p(1-\theta_1)\mathbb{E}[g(\hat{X}^{\perp}, X_1)] + (1-p)p(1-\theta_2)\mathbb{E}[g(X_1^{\perp}, \hat{X})].
\end{aligned}$$

Similarly we find

$$\begin{aligned}
& (1-p)p\theta_2\mathbb{E}[g(\hat{X}, X_2^{\perp})] + (1-p)p\theta_1\mathbb{E}[g(X_2, \hat{X}^{\perp})] \\
\geq \quad & (1-p)p\theta_1\mathbb{E}[g(\hat{X}^{\perp}, X_2)] + (1-p)p\theta_2\mathbb{E}[g(X_2^{\perp}, \hat{X})]
\end{aligned}$$

because $ess\sup X \leq ess\inf X_2$ implies $X \leq_{lr} X_2$, and

$$\begin{aligned} & p^2(1-\theta_1)\theta_2\mathbb{E}[g(X_1, X_2^{\perp})] + p^2(1-\theta_2)\theta_1\mathbb{E}[g(X_2, X_1^{\perp})] \\ \geq \ & p^2(1-\theta_2)\theta_1\mathbb{E}[g(X_1^{\perp}, X_2)] + p^2(1-\theta_1)\theta_2\mathbb{E}[g(X_2^{\perp}, X_1)]. \end{aligned}$$

because $ess\sup X_1 \leq ess\inf X_2$ implies $X_1 \leq_{lr} X_2$. Finally, $\mathbb{E}[g(X_i, X_i^{\perp})] = \mathbb{E}[g(X_i^{\perp}, X_i)]$ for $i = 1, 2$ because $X_i^{\perp}$ is an independent copy of X_i. Combining inequalities accordingly shows that $\mathbb{E}[g(Y_1, Y_2)] \geq \mathbb{E}[g(Y_2, Y_1)]$ for any arrangement-increasing function $g(x, y)$, and therefore $Y_1 \leq_{lr} Y_2$ as desired.

A.6 Proof of Lemma 3

Denote the survival functions of X_1, $\hat{X}$ and X_2 by S_1, $\hat{S}$ and S_2, respectively. The conditional survival function of X in Eq. (8) given $\Theta = \theta$ is then given by

$$S(x|\theta) = (1-p)\hat{S}(x) + pS_1(x) + p\theta(S_2(x) - S_1(x)).$$

According to Definition 1(ii), we need to show that $S(x|\theta_2)/S(x|\theta_1)$ is increasing in x for any $\theta_1 \leq \theta_2$. Introduce

$$m(x) = \frac{S_2(x) - S_1(x)}{(1-p)\hat{S}(x) + pS_1(x)}$$

so that $S(x|\theta) = \left((1-p)\hat{S}(x) + pS_1(x)\right)(1 + p\theta m(x))$. We then obtain

$$\frac{S(x|\theta_2)}{S(x|\theta_1)} = \frac{1 + p\theta_2 m(x)}{1 + p\theta_1 m(x)} = \frac{\theta_2}{\theta_1} - \frac{\theta_2/\theta_1 - 1}{1 + p\theta_1 m(x)}. \tag{14}$$

Now $X_1 \leq_{hr} \hat{X} \leq_{hr} X_2$ implies that both $\hat{S}(x)/S_1(x)$ and $S_2(x)/\hat{S}(x)$ are increasing in x. This implies that $m(x)$ is also increasing in x, which we can see by rewriting it as follows:

$$m(x) = \frac{S_2(x)/\hat{S}(x)}{(1-p) + pS_1(x)/\hat{S}(x)} - \frac{1}{(1-p)\hat{S}(x)/S_1(x) + p}.$$

It then follows from Eq. (14) that $S(x|\theta_2)/S(x|\theta_1)$ is increasing in x since $\theta_2 \geq \theta_1$. This demonstrates $X \uparrow_{hr} \Theta$.

A.7 Proof of Proposition 6

For any given indemnity $I \in \mathfrak{C}$, the policyholder retains $R = x_1 - I(x_1)$ when a loss of x_1 occurs. Indemnity schedules in $\mathfrak{C}$ satisfy the principle of indemnity so the amount of retention is nonnegative, $R \geq 0$. Let $I_R(x) = \max(x - R, 0) = \max(x - x_1 + I(x_1), 0)$ be the straight deductible that uses this amount of retention as the deductible level. Assume that $\mathbb{E}[I(X_2)] \geq \mathbb{E}[I_R(X_2)]$; then, there exists a $D_l \in [0, R]$ such that $\mathbb{E}[I(X_2)] = \mathbb{E}[I_l(X_2; D_l, I(x_1))]$, with I_l as defined in Proposition 6(i). For brevity, we write $I_l(x)$ as shorthand for $I_l(x; D_l, I(x_1))$.

There is some $x_0 \in [D_l, D_l + I(x_1)]$ so that indemnity schedule $I_l(x)$ up-crosses indemnity schedule $I(x)$ at x_0, that is, $I_l(x) \geq I(x)$ for any $x > x_0$ and $I_l(x) \leq I(x)$ for any $x < x_0$. In particular, we have $I_l(x_1) = I(x_1)$. According to Ohlin's (1969) Lemma 3, we then obtain

$$I_l(X_2) \geq_{RS} I(X_2) \qquad \text{and} \qquad X_2 - I_l(X_2) \leq_{RS} X_2 - I(X_2).$$

Under the one-against-all ambiguity structure (10), $\mathbb{E}[I(X_2)] = \mathbb{E}[I_l(X_2; D_l, I(x_1))]$ ensures

$$\mathbb{E}[I(X)] = I(x_1)\mathbb{E}[1-\Theta] + \mathbb{E}[\Theta]\mathbb{E}[I(X_2)] = I_l(x_1)\mathbb{E}[1-\Theta] + \mathbb{E}[\Theta]\mathbb{E}[I_l(X_2)] = \mathbb{E}[I_l(X)].$$

Hence, the premiums for I and I_l are equal. Coupled with $X_2 - I_l(X_2)$ being less risky than $X_2 - I(X_2)$ in the sense of Rothschild and Stiglitz (1970), this implies that, for any increasing and concave utility function u, we have

$$\mathbb{E}[u(w - X_2 + I_l(X_2) - (1+\tau)\mathbb{E}[I_l(X)])] \geq \mathbb{E}[u(w - X_2 + I(X_2) - (1+\tau)\mathbb{E}[I(X)])].$$

Consequently, we find

$$\begin{aligned}
\mathbb{E}[u(W_{I_l})|\Theta] \quad &= \quad (1-\Theta)u(w - x_1 + I_l(x_1) - (1+\tau)\mathbb{E}[I_l(X)]) \\
&\quad\quad +\Theta\mathbb{E}[u(w - X_2 + I_l(X_2) - (1+\tau)\mathbb{E}[I_l(X)])] \\
&\geq_{a.s.} \quad (1-\Theta)u(w - x_1 + I(x_1) - (1+\tau)\mathbb{E}[I(X)]) \\
&\quad\quad +\Theta\mathbb{E}[u(w - X_2 + I(X_2) - (1+\tau)\mathbb{E}[I(X)])] = \mathbb{E}[u(W_I)|\Theta],
\end{aligned}$$

which implies

$$J(I_l) = \mathbb{E}[\phi(\mathbb{E}[u(W_{I_l})|\Theta])] \geq \mathbb{E}[\phi(\mathbb{E}[u(W_I)|\Theta])] = J(f)$$

because ϕ is increasing. As a result, indemnity schedule I is dominated by indemnity schedule I_l, which is of the form stated in Proposition 6(i).

If $\mathbb{E}[I(X_2)] \geq \mathbb{E}[I_R(X_2)]$ does not hold, we must have $\mathbb{E}[I(X_2)] < \mathbb{E}[I_R(X_2)]$ instead. In this case, we can show with a similar argument that indemnity schedule I is dominated by an indemnity schedule of the form stated in Proposition 6(ii). Indemnity I was arbitrarily chosen so for any indemnity schedule in $\mathfrak{C}$, we can find an indemnity schedule either of the form in Proposition 6(i) or (ii) that increases the policyholder's ex-ante welfare. The optimal solution to Problem (1) under the one-against-all ambiguity structure must therefore be either of the form $I_l(x; D_l, U)$ or of the form $I_r(x; D_r, U)$.

A.8 Proof of Proposition 7

For (i), observe that $\mathbb{P}(X > x|\Theta = \theta)/\mathbb{P}(X > x_\tau|\Theta = \theta)$ is independent of θ for any $x > x_\tau$ in case of $x_1 \leq x_\tau$. It then follows from Proposition 3(i) that a straight deductible is optimal with the deductible level specified in Eq. (7).

To show (ii), note that the one-against-all ambiguity structure (10) reduces to a special case of the left-against-right ambiguity structure when $x_1 \le ess\inf X_2$. The optimality of a straight deductible with deductible level specified in Eq. (7) then follows from Proposition 5.

Similarly, if $x_1 \ge ess\sup X_2$, the one-against-all ambiguity structure (10) reduces to a special case of the left-against-right ambiguity structure with the positions of x_1 and X_2 switched. The optimality of a straight deductible with deductible level specified in Eq. (7) then follows from the discussion after Lemma 2 and Proposition 5.

A.9 Proof of Proposition 8

For any given indemnity $I \in \mathfrak{C}$, set $U_l = I(x_1)$ and $U_m = I(x_2)$. We can then find D_l, D_m and D_r satisfying (12) such that

$$\begin{aligned}\mathbb{E}\left[I(\hat{X})\cdot\mathbf{1}_{\{0\le\hat{X}<x_1\}}\right] &= \mathbb{E}\left[I^*(\hat{X})\cdot\mathbf{1}_{\{0\le\hat{X}<x_1\}}\right],\\ \mathbb{E}\left[I(\hat{X})\cdot\mathbf{1}_{\{x_1\le\hat{X}<x_2\}}\right] &= \mathbb{E}\left[I^*(\hat{X})\cdot\mathbf{1}_{\{x_1\le\hat{X}<x_2\}}\right],\\ \mathbb{E}\left[I(\hat{X})\cdot\mathbf{1}_{\{\hat{X}\ge x_2\}}\right] &= \mathbb{E}\left[I^*(\hat{X})\cdot\mathbf{1}_{\{\hat{X}\ge x_2\}}\right].\end{aligned}$$

Due to the crossing positions between $I(x)$ and $I^*(x)$, we conclude that

$$I(\hat{X}) \le_{RS} I^*(\hat{X}) \qquad \text{and} \qquad \hat{X} - I^*(\hat{X}) \le_{RS} \hat{X} - I(\hat{X}).$$

By construction we obtain that $I^*(x_1) = I(x_1)$ and $I^*(x_2) = I(x_2)$. Using a similar argument as in the proof of Proposition 6 in Appendix A.7, we conclude that $J(I) \le J(I^*)$. The optimal indemnity schedule must therefore be of the form specified in Eq. (11).

A.10 Proof of Proposition 9

Let policyholder 1 be more ambiguity-averse than policyholder 2. For $I \in \mathfrak{C}$, auxiliary function $V_I(x)$ defined in (5) depends on the policyholder's ambiguity function. To make this clear, we write $V_I^{\phi_1}(x)$ and $V_I^{\phi_2}(x)$ to distinguish between the two policyholders. We have that $\Theta \uparrow_{hr} X$ implies $\Theta \uparrow_{fsd} X$. Together with $X \uparrow_{fsd} \Theta$, we then know from Proposition 3(ii) that a straight deductible is optimal for each policyholder. Let D_1^* and D_2^* denote their deductible levels, which are determined by

$$D_1^* = \inf\left\{D \ge x_\tau : V_{I_D}^{\phi_1}(D) \ge 1+\tau\right\} \qquad \text{and} \qquad D_2^* = \inf\left\{D \ge x_\tau : V_{I_D}^{\phi_2}(D) \ge 1+\tau\right\}.$$

If $D_2^* = \infty$, then $D_1^* \le D_2^*$ follows immediately. Otherwise, if $D_2^* < \infty$, we can show $D_1^* \le D_2^*$ by showing $V_{I_{D_2^*}}^{\phi_1}(D_2^*) \ge V_{I_{D_2^*}}^{\phi_2}(D_2^*)$.

Let $\xi = \mathbb{E}[-u(W_{I_{D_2^*}})|\Theta]$. Then, ξ is an increasing transformation of Θ since $X \uparrow_{fsd} \Theta$. According to Lemma 5(ii) in Appendix B.2, $\Theta \uparrow_{hr} X$ then implies $\xi \uparrow_{hr} X$. In other words, $[\xi|X = x_1] \le_{hr} [\xi|X = x_2]$ for any $x_1 \le x_2$. If ϕ_1 is more concave than ϕ_2 in the sense of Arrow-

Pratt, then $\phi_1'(-z)/\phi_2'(-z)$ is increasing in z.[7] Both $\phi_1'(-z)$ and $\phi_2'(-z)$ are nonnegative and increasing in z. Lemma 5(iii) in Appendix B.2 then implies that

$$\frac{\mathbb{E}[\phi_1'(-\xi)|X=x_1]}{\mathbb{E}[\phi_2'(-\xi)|X=x_1]} \leq \frac{\mathbb{E}[\phi_1'(-\xi)|X=x_2]}{\mathbb{E}[\phi_2'(-\xi)|X=x_2]} \tag{15}$$

for any $x_1 \leq x_2$. The ratio $\mathbb{E}[\phi_1'(-\xi)|X=x]/\mathbb{E}[\phi_2'(-\xi)|X=x]$ is thus increasing in x.

Denote by

$$h_i(x) = \mathbb{E}[\phi_i'(\mathbb{E}[u(W_{I_{D_2^*}})|\Theta])u'(W_{I_{D_2^*}})|X=x], \qquad \text{for } i=1,2,$$

the expected marginal welfare of policyholders 1 and 2 under the straight deductible D_2^* conditional on a loss of x. On $\{X=x\}$, the quantity $u(W_{I_{D_2^*}})$ is deterministic so that

$$\frac{h_1(x)}{h_2(x)} = \frac{\mathbb{E}[\phi_1'(-\xi)u'(W_{I_{D_2^*}})|X=x]}{\mathbb{E}[\phi_2'(-\xi)u'(W_{I_{D_2^*}})|X=x]} = \frac{\mathbb{E}[\phi_1'(-\xi)|X=x]}{\mathbb{E}[\phi_2'(-\xi)|X=x]}.$$

Therefore, the ratio $h_1(x)/h_2(x)$ is increasing in x, see (15) above. According to Lemma A.3 in Chi and Wei (2018), we then obtain

$$\frac{\mathbb{E}[h_1(X)\cdot \mathbf{1}_{\{X>D_2^*\}}]}{\mathbb{E}[h_1(X)\cdot \mathbf{1}_{\{X\leq D_2^*\}}]} \geq \frac{\mathbb{E}[h_2(X)\cdot \mathbf{1}_{\{X>D_2^*\}}]}{\mathbb{E}[h_2(X)\cdot \mathbf{1}_{\{X\leq D_2^*\}}]},$$

which in turn implies

$$\frac{\mathbb{E}[h_1(X)\cdot \mathbf{1}_{\{X>D_2^*\}}]}{\mathbb{E}[h_1(X)]} \geq \frac{\mathbb{E}[h_2(X)\cdot \mathbf{1}_{\{X>D_2^*\}}]}{\mathbb{E}[h_2(X)]}.$$

Using the definition of h_1 and h_2, this demonstrates

$$\begin{aligned} V_{I_{D_2^*}}^{\phi_1}(D_2^*) &= \frac{\mathbb{E}\left[\phi_1'(\mathbb{E}[u(W_{I_{D_2^*}})|\Theta])u'(W_{I_{D_2^*}})\middle|X>D_2^*\right]}{\mathbb{E}\left[\phi_1'(\mathbb{E}[u(W_{I_{D_2^*}})|\Theta])u'(W_{I_{D_2^*}})\right]} \\ &\geq \frac{\mathbb{E}\left[\phi_2'(\mathbb{E}[u(W_{I_{D_2^*}})|\Theta])u'(W_{I_{D_2^*}})\middle|X>D_2^*\right]}{\mathbb{E}\left[\phi_2'(\mathbb{E}[u(W_{I_{D_2^*}})|\Theta])u'(W_{I_{D_2^*}})\right]} = V_{I_{D_2^*}}^{\phi_2}(D_2^*), \end{aligned}$$

Therefore, $D_1^* \leq D_2^*$ as desired.

[7] By direct computation, we find

$$\frac{\mathrm{d}}{\mathrm{d}z}\frac{\phi_1'(-z)}{\phi_2'(-z)} = \frac{\phi_1'(-z)}{\phi_2'(-z)}\cdot\left[-\frac{\phi_1''(-z)}{\phi_1'(-z)} - \left(-\frac{\phi_2''(-z)}{\phi_2'(-z)}\right)\right] \geq 0.$$

A.11 Proof of Proposition 10

According to Proposition 5, the optimal insurance contract is a straight deductible for both second-order beliefs Θ_1 and Θ_2. Let X_{Θ_1} and X_{Θ_2} denote the loss random variables under the two second-order beliefs. As stated in Definition 1(iv), $\Theta_2 \geq_{RS} \Theta_1$ entails $\mathbb{E}[\Theta_1] = \mathbb{E}[\Theta_2]$. Therefore, X_{Θ_1} and X_{Θ_2} have the same unconditional distributions because it only depends on $\mathbb{E}[\Theta]$ under the three-piece ambiguity structure (8). Let

$$V_{I_D}^{\Theta_i}(D) = \frac{\mathbb{E}\left[\phi'(\mathbb{E}[u(W_{I_D}^{\Theta_i})\big|\Theta_i])u'(W_{I_D}^{\Theta_i})|X_{\Theta_i} > D\right]}{\mathbb{E}\left[\phi'(\mathbb{E}[u(W_{I_D}^{\Theta_i})|\Theta_i])u'(W_{I_D}^{\Theta_i})\right]} \qquad \text{for } i \in \{1, 2\},$$

denote auxiliary function (5) under Θ_1 and Θ_2. As in Appendix A.10, it suffices to show that $V_{I_D}^{\Theta_1}(D) \leq V_{I_D}^{\Theta_2}(D)$. When $X_{\Theta_i} > D$, then $u'(W_{I_D}^{\Theta_i}) = u'(w - D - (1+\tau)\mathbb{E}[I_D(X_{\Theta_i})])$, which is a deterministic quantity because $\mathbb{E}[I_D(X_{\Theta_1})] = \mathbb{E}[I_D(X_{\Theta_2})]$. Therefore, $V_{I_D}^{\Theta_1}(D) \leq V_{I_D}^{\Theta_2}(D)$ is equivalent to

$$\frac{\mathbb{E}\left[\phi'(\mathbb{E}[u(W_{I_D}^{\Theta_1})|\Theta_1])\big|X_{\Theta_1} > D\right]}{\mathbb{E}\left[\phi'(\mathbb{E}[u(W_{I_D}^{\Theta_1})|\Theta_1])u'(W_{I_D}^{\Theta_1})\right]} \leq \frac{\mathbb{E}\left[\phi'(\mathbb{E}[u(W_{I_D}^{\Theta_2})|\Theta_2])\big|X_{\Theta_2} > D\right]}{\mathbb{E}\left[\phi'(\mathbb{E}[u(W_{I_D}^{\Theta_2})|\Theta_2])u'(W_{I_D}^{\Theta_2})\right]}. \tag{16}$$

To show this, we first derive the posterior distribution of Θ_i conditional on $X_{\Theta_i} > D$. For ease of exposition, we assume that Θ_1 and Θ_2 have probability density functions $\pi_1(\theta)$ and $\pi_2(\theta)$. The posterior density function of Θ_i conditional on $X_{\Theta_i} > D$ is then given by

$$\pi_{i|D} = \frac{\pi_i(\theta) \cdot \mathbb{P}(X_{\Theta_i} > D|\Theta_i = \theta)}{\mathbb{P}(X_{\Theta_i} > D)} = \frac{\pi_i(\theta) \cdot (m_D\theta + b_D)}{\mathbb{P}(X_{\Theta_i} > D)},$$

where $m_D = p(\mathbb{P}(X_2 > D) - \mathbb{P}(X_1 > D))$ and $b_D = (1-p)\mathbb{P}(\hat{X} > D) + p\mathbb{P}(X_1 > D)$.

Under the three-piece ambiguity structure (8), $\mathbb{E}[u(W_{I_D}^{\Theta_i})|\Theta_i]$ and $\mathbb{E}[u'(W_{I_D}^{\Theta_i})|\Theta_i]$ are both linear functions in Θ_i. Denote

$$\mathbb{E}[u(W_{I_D}^{\Theta_i})|\Theta_i] = m_u\Theta_i + b_u \qquad \text{and} \qquad \mathbb{E}[u'(W_{I_D}^{\Theta_i})|\Theta_i] = m_{u'}\Theta_i + b_{u'}$$

with

$$m_u = p(\mathbb{E}[u(W_{I_D}(X_2))] - \mathbb{E}[u(W_{I_D}(X_1))]), \qquad b_u = (1-p)\mathbb{E}[u(W_{I_D}(\hat{X}))] + p\mathbb{E}[u(W_{I_D}(X_1))],$$
$$m_{u'} = p(\mathbb{E}[u'(W_{I_D}(X_2))] - \mathbb{E}[u'(W_{I_D}(X_1))]), \quad b_{u'} = (1-p)\mathbb{E}[u'(W_{I_D}(\hat{X}))] + p\mathbb{E}[u'(W_{I_D}(X_1))].$$

We use the notation $W_{I_D}(X_1)$ as shorthand for $w - X_1 + I_D(X_1) - (1+\tau)\mathbb{E}[I_D(X)]$, and likewise for $\hat{X}$ and X_2. We assume $X_1 \leq_{hr} \hat{X} \leq_{hr} X_2$ in Proposition 10, which implies $X_1 \leq_{fsd} X_2$. Final wealth $W_{I_D}(x)$ is decreasing in x, u is increasing, and u' is decreasing. We then have $m_u \leq 0$ and $m_{u'} \geq 0$ from $X_1 \leq_{fsd} X_2$.

With the help of these notations, we can now derive expressions for the quantities in inequality (16). For the numerators, we obtain

$$\begin{aligned}&\mathbb{E}\left[\phi'(\mathbb{E}[u(W_{I_D}^{\Theta_i})|\Theta_i])\big|X_{\Theta_i} > D\right] = \mathbb{E}[\phi'(m_u\Theta_i + b_u)|X_{\Theta_i} > D]\\ = &\int_0^1 \phi'(m_u\theta + b_u)\cdot \pi_{i|D}(\theta)\,\mathrm{d}\theta = \int_0^1 \phi'(m_u\theta + b_u)\cdot \frac{\pi_i(\theta)(m_D\theta + b_D)}{\mathbb{P}(X_{\Theta_i} > D)}\,\mathrm{d}\theta\\ = &\frac{1}{\mathbb{P}(X_{\Theta_i} > D)}\cdot \mathbb{E}[\phi'(m_u\Theta_i + b_u)(m_D\Theta_i + b_D)].\end{aligned}$$

For the denominators, we apply the double expectation theorem to obtain

$$\begin{aligned}&\mathbb{E}\left[\phi'(\mathbb{E}[u(W_{I_D}^{\Theta_i})|\Theta_i])u'(W_{I_D}^{\Theta_i})\right] = \mathbb{E}\left[\phi'(m_u\Theta_i + b_u)u'(W_{I_D}^{\Theta_i})\right]\\ = &\mathbb{E}\left[\phi'(m_u\Theta_i + b_u)\mathbb{E}[u'(W_{I_D}^{\Theta_i})|\Theta_i]\right] = \mathbb{E}\left[\phi'(m_u\Theta_i + b_u)(m_{u'}\Theta_i + b_{u'})\right].\end{aligned}$$

Combining the expressions for the numerators and the denominators, condition (16) is equivalent to

$$\frac{\mathbb{E}[\phi'(m_u\Theta_1 + b_u)(m_D\Theta_1 + b_D)]}{\mathbb{E}\left[\phi'(m_u\Theta_1 + b_u)(m_{u'}\Theta_1 + b_{u'})\right]} \geq \frac{\mathbb{E}[\phi'(m_u\Theta_2 + b_u)(m_D\Theta_2 + b_D)]}{\mathbb{E}\left[\phi'(m_u\Theta_2 + b_u)(m_{u'}\Theta_2 + b_{u'})\right]}.$$

Denote $g(x) = \phi'(m_ux+b_u)(m_Dx+b_D)$ and $h(x) = \phi'(m_ux+b_u)(m_{u'}x+b_{u'})$. We can assume without loss of generality that Θ_1 and Θ_2 are independent, in which case condition (16) is further equivalent to

$$\mathbb{E}[g(\Theta_1)h(\Theta_2)] \leq \mathbb{E}[g(\Theta_2)h(\Theta_1)].$$

Since $\Theta_1 \leq_{RS} \Theta_2$, we can use the bivariate characterization of the convex order in Theorem 3.A.6 in Shaked and Shanthikumar (2007). A sufficient condition for the above inequality to hold is that $g(x)h(y) - g(y)h(x)$ is convex in x for all y. By direct computation, we obtain

$$g(x)h(y) - g(y)h(x) = (m_Db_{u'} - m_{u'}b_D)\phi'(m_uy + b_u)\phi'(m_ux + b_u)(x - y).$$

We complete the proof by showing that $\gamma(x) = \phi'(m_ux + b_u)(x - y)$ is convex in x for all y and that $m_Db_{u'} - m_{u'}b_D \geq 0$.

We obtain

$$\begin{aligned}\gamma''(x) &= m_u^2\phi'''(m_ux + b_u)(x - y) + 2m_u\phi'''(m_ux + b_u)\\ &= \underbrace{m_u}_{\leq 0}\underbrace{\phi''(m_ux + b_u)}_{\leq 0}\cdot\underbrace{[2 - \mathcal{P}_\phi(m_ux + b_u)]}_{\geq 0} - \underbrace{m_u}_{\leq 0}\underbrace{(m_uy + b_u)}_{\geq 0}\underbrace{\phi'''(m_ux + b_u)}_{\geq 0}.\end{aligned}$$

By definition, $m_uy + b_u = \mathbb{E}[u(W_{I_D}(X))|\Theta = y] \geq 0$ so that $\gamma''(x) \geq 0$ under our assumptions. To show that $m_Db_{u'} - m_{u'}b_D \geq 0$, let $\widetilde{X}_i$ follow a mixture distribution of $\hat{X}$ and X_i with weights $(1-p)$ and p, that is, $\mathbb{P}(\widetilde{X}_i \le x) = (1-p)\mathbb{P}(\hat{X} \le x) + p\mathbb{P}(X_i \le x)$ for $i = 1, 2$. According to Theorem 1.B.8 in Shaked and Shanthikumar (2007), the hazard rate order is preserved under mixture. Since $X_1 \le_{hr} \hat{X} \le_{hr} X_2$, we then have $\widetilde{X}_1 \le_{hr} \widetilde{X}_2$.

We express the coefficients m_D, $b_{u'}$, $m_{u'}$, and b_D in terms of $\widetilde{X}_1$ and $\widetilde{X}_2$ as follows:

$$\begin{aligned} m_D &= \mathbb{P}(\widetilde{X}_2 > D) - \mathbb{P}(\widetilde{X}_1 > D), & b_D &= \mathbb{P}(\widetilde{X}_1 > D), \\ m_{u'} &= \mathbb{E}[u'(W_{I_D}(\widetilde{X}_2))] - \mathbb{E}[u'(W_{I_D}(\widetilde{X}_1))], & b_{u'} &= \mathbb{E}[u'(W_{I_D}(\widetilde{X}_1))]. \end{aligned}$$

The desired inequality $m_D b_{u'} - m_{u'} b_D \ge 0$ is then equivalent to

$$\mathbb{E}[u'(W_{I_D}(\widetilde{X}_1))] \cdot \mathbb{P}(\widetilde{X}_2 > D) \quad \ge \quad \mathbb{E}[u'(W_{I_D}(\widetilde{X}_2))] \cdot \mathbb{P}(\widetilde{X}_1 > D). \tag{17}$$

For $i = 1, 2$, we rewrite

$$\begin{aligned} \mathbb{E}[u'(W_{I_D}(\widetilde{X}_i))]) &= \mathbb{E}\left[u'(W_{I_D}(\widetilde{X}_i)) \cdot \mathbf{1}_{\{\widetilde{X}_i \le D\}}\right] + \mathbb{E}\left[u'(W_{I_D}(\widetilde{X}_i)) \cdot \mathbf{1}_{\{\widetilde{X}_i > D\}}\right] \\ &= \mathbb{E}\left[u'(W_{I_D}(\widetilde{X}_i)) \cdot \mathbf{1}_{\{\widetilde{X}_i \le D\}}\right] + u'(W_{I_D}(D))\mathbb{P}(\widetilde{X}_i > D). \end{aligned}$$

Letting $\widetilde{X}_1$ and $\widetilde{X}_2$ be independent, which we can assume without loss of generality, condition (17) is further equivalent to

$$\mathbb{E}\left[u'(W_{I_D}(\widetilde{X}_1)) \cdot \mathbf{1}_{\{\widetilde{X}_1 \le D\}}\right] \cdot \mathbb{P}(\widetilde{X}_2 > D) \ge \mathbb{E}\left[u'(W_{I_D}(\widetilde{X}_2)) \cdot \mathbf{1}_{\{\widetilde{X}_2 \le D\}}\right] \cdot \mathbb{P}(\widetilde{X}_1 > D),$$

which is in turn equivalent to

$$\mathbb{E}\left[u'(W_{I_D}(\widetilde{X}_1)) \cdot \mathbf{1}_{\{\widetilde{X}_1 \le D, \widetilde{X}_2 > D\}}\right] \ge \mathbb{E}\left[u'(W_{I_D}(\widetilde{X}_2)) \cdot \mathbf{1}_{\{\widetilde{X}_2 \le D, \widetilde{X}_1 > D\}}\right].$$

Recalling that $\widetilde{X}_1 \le_{hr} \widetilde{X}_2$, we can use the bivariate characterization of the hazard rate order in Theorem 1.B.10 of Shaked and Shanthikumar (2007). The inequality above holds if

$$u'(W_{I_D}(x)) \cdot \mathbf{1}_{\{x \le D, y > D\}} - u'(W_{I_D}(y)) \cdot \mathbf{1}_{\{y \le D, x > D\}}$$

increases in y on $\{y \ge x\}$. This is the case because $\mathbf{1}_{\{y \le D, x > D\}} = 0$ when $y \ge x$ and $\mathbf{1}_{\{x \le D, y > D\}}$ increases in y.

Together with the convexity of $\gamma(x)$, this proves that $g(x)h(y) - g(y)h(x)$ is convex in x for all y, which completes the proof.

B Auxiliary results

B.1 Properties of function $V_{I_D}(x)$

For $I \in \mathfrak{C}$, the function $V_I(x)$ is defined in Eq. (5). For $D \geq 0$, indemnity schedule I_D denotes the straight deductible $I_D(x) = \max(0, x - D)$, and function $V_{I_D}(x)$ is defined accordingly. The following lemma summarizes useful properties of $V_{I_D}(x)$ that we use in various proofs.

Lemma 4. *The function $V_{I_D}(x)$ has the following properties.*

(i) $V_{I_D}(x)$ is right continuous in x with left limit for any fixed D.

(ii) $V_{I_D}(x)$ is continuous in D for any fixed x.

Furthermore, if $V_{I_D}(x)$ is increasing in x on $[x_\tau, \mathcal{M})$ for any $D \geq x_\tau$, the following holds.

(iii) For $x_0 \in [x_\tau, \mathcal{M})$, the double limits $\lim_{(D,x)\to(x_0,x_0-)} V_{I_D}(x)$ and $\lim_{(D,x)\to(x_0,x_0+)} V_{I_D}(x)$ both exist.[8]

(iv) As a univariate function, $V_{I_D}(D)$ is right continuous in $D \in [x_\tau, \mathcal{M})$ with left limit.

Proof. Result (i) follows from the dominated convergence theorem because $\mathbb{E}\left[g(X)\mathbf{1}_{\{X>x\}}\right]$ is right continuous in x for any function g such that $\mathbb{E}\left[|g(X)|\right] < \infty$. Therefore,

$$V_{I_D}(x) = \frac{\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_D})|\Theta])u'(W_{I_D})|X > x]}{\mathbb{E}[\phi'(\mathbb{E}[u(W_{I_D})|\Theta])u'(W_{I_D})]} = \frac{\mathbb{E}\left[\phi'(\mathbb{E}[u(W_{I_D})|\Theta])u'(W_{I_D}) \cdot \mathbf{1}_{\{X>x\}}\right]}{S_X(x) \cdot \mathbb{E}[\phi'(\mathbb{E}[u(W_{I_D})|\Theta])u'(W_{I_D})]}$$

is right continuous in x. Left limits exist due to the monotone convergence theorem.

Result (ii) is also due to the dominated convergence theorem.

To show (iii), let $D \geq x_\tau$ and select a sequence $\{x_n\}_{n=1}^\infty \subset [x_\tau, x_0)$ such that $x_n \uparrow x_0$. Since $V_{I_D}(x)$ is increasing in x, $V_{I_D}(x_n)$ monotonically converges to $V_{I_D}(x_0-)$. Both $\{V_{I_D}(x_n)\}_{n=1}^\infty$ and $V_{I_D}(x_0-)$ are continuous functions of D. According to Dini's theorem, $V_{I_D}(x_n)$ then converges uniformly to $V_{I_D}(x_0-)$ over $D \in [x_\tau, \mathcal{M})$. Therefore, for any $\varepsilon > 0$, there exists $N \geq 1$ such that

$$|V_{I_D}(x_n) - V_{I_D}(x_0-)| < \varepsilon \qquad \text{for any } n \geq N \text{ and } D \in [x_\tau, \mathcal{M}).$$

Then, for any $x \in [x_N, x_0)$, it holds that

$$|V_{I_D}(x) - V_{I_D}(x_0-)| \leq |V_{I_D}(x_N) - V_{I_D}(x_0-)| < \varepsilon \qquad \text{for any } D \in [x_\tau, \mathcal{M}).$$

This implies the uniform convergence of $V_{I_D}(x)$ to $V_{I_D}(x_0-)$ over $D \in [x_\tau, \mathcal{M})$ as x approaches x_0 from the left. The double limit $\lim_{(D,x)\to(x_0,x_0-)} V_{I_D}(x)$ then exists due to the

[8] The notation $(D, x) \to (x_0, x_0-)$ (resp., $(D, x) \to (x_0, x_0+)$) means that (D, x) approaches (x_0, x_0) with the constraint $x < x_0$ (resp., $x > x_0$).

Moore-Osgood theorem. A similar argument proves the existence of the other double limit $\lim_{(D,x)\to(x_0,x_0+)} V_{I_D}(x)$.

To show result (iv), note that the existence of $\lim_{(D,x)\to(x_0,x_0+)} V_{I_D}(x)$ for any $x_0 \in [x_\tau, \mathcal{M})$ implies that $V_{I_D}(x)$ reaches the same limit as (D,x) approaches (x_0,x_0) from different directions in the half plane $\{(D,x) : x > x_0\}$. Therefore,

$$\lim_{D\downarrow x_0} V_{I_D}(D) = \lim_{(D,x)\to(x_0,x_0)} V_{I_D}(x) = \lim_{D\downarrow x}\lim_{x\downarrow x_0} V_{I_D}(x) = \lim_{D\downarrow x_0} V_{I_D}(x_0) = V_{I_{x_0}}(x_0),$$

where the last two equalities are due to the right continuity of $V_{I_D}(x)$ in x and the continuity of $V_{I_D}(x_0)$ in D, respectively. This proves the right continuity of $V_{I_D}(D)$ in D. The left limit of $V_{I_D}(D)$ in D exists due to

$$\lim_{D\uparrow x_0} V_{I_D}(D) = \lim_{(D,x)\to(x_0,x_0-)} V_{I_D}(x) = \lim_{x\uparrow x_0}\lim_{D\uparrow x_0} V_{I_D}(x) = \lim_{x\uparrow x_0} V_{I_{x_0}}(x).$$

The left limit of $V_{I_{x_0}}(x)$ in x exists because of result (i). □

B.2 Three useful properties of the hazard rate order

Property (i) is used in the proof of Proposition 3(iii) in Appendix A.3, properties (ii) and (iii) are used in the proof of Proposition 9 in Appendix A.10.

Lemma 5. *The hazard rate order possesses the following properties.*

(i) If $X \uparrow_{hr} \Theta$, then $[\Theta|X > x]$ increases in x in the likelihood ratio order.

(ii) If $X \leq_{hr} Y$, then $g(X) \leq_{hr} g(Y)$ for any increasing function g.

(iii) If $X \leq_{hr} Y$, then $\frac{\mathbb{E}[f(X)]}{\mathbb{E}[g(X)]} \leq \frac{\mathbb{E}[f(Y)]}{\mathbb{E}[g(Y)]}$ for any pair of functions f and g such that $g(x)$ is nonnegative and increasing, and $\frac{f(x)}{g(x)}$ is increasing in x.

Proof. To show (i), assume for ease of exposition that Θ has a probability density function $\pi(\theta)$. We denote the conditional density function of Θ given $X > x$ by $\pi(\theta|X > x)$. We then obtain $\mathbb{P}(X > x|\Theta = \theta) = \pi(\theta|X > x)\mathbb{P}(X > x)/\pi(\theta)$ from Bayes' theorem. Now $X \uparrow_{hr} \Theta$ implies that $\mathbb{P}(X > x|\Theta = \theta_2)/\mathbb{P}(X > x|\Theta = \theta_1)$ increases in x for any $\theta_1 \leq \theta_2$. Therefore, for any $x_1 \leq x_2$ and $\theta_1 \leq \theta_2$ we obtain the following equivalences:

$$\frac{\mathbb{P}(X > x_2|\Theta = \theta_2)}{\mathbb{P}(X > x_2|\Theta = \theta_1)} \geq \frac{\mathbb{P}(X > x_1|\Theta = \theta_2)}{\mathbb{P}(X > x_1|\Theta = \theta_1)}$$

$$\iff \frac{\pi(\theta_2|X > x_2)\mathbb{P}(X > x_2)/\pi(\theta_2)}{\pi(\theta_1|X > x_2)\mathbb{P}(X > x_2)/\pi(\theta_1)} \geq \frac{\pi(\theta_2|X > x_1)\mathbb{P}(X > x_1)/\pi(\theta_2)}{\pi(\theta_1|X > x_1)\mathbb{P}(X > x_1)/\pi(\theta_1)}$$

$$\iff \frac{\pi(\theta_2|X > x_2)}{\pi(\theta_1|X > x_2)} \geq \frac{\pi(\theta_2|X > x_1)}{\pi(\theta_1|X > x_1)} \quad \iff \quad \frac{\pi(\theta_2|X > x_2)}{\pi(\theta_2|X > x_1)} \geq \frac{\pi(\theta_1|X > x_2)}{\pi(\theta_1|X > x_1)}.$$

The last inequality states that $\pi(\theta|X > x_2)/\pi(\theta|X > x_1)$ increases in θ for any $x_1 \leq x_2$, that is, that $[\Theta|X > x]$ increases in x in the sense of the likelihood ratio order.

Property (ii) is taken from Theorem 1.B.2 of Shaked and Shanthikumar (2007).

For (iii), let $y_2 \geq y_1 \geq x$ so that

$$\begin{aligned}&(f(y_2)g(x) - f(x)g(y_2)) - (f(y_1)g(x) - f(x)g(y_1)) \\ = \;& g(x)g(y_2)\left(\frac{f(y_2)}{g(y_2)} - \frac{f(y_1)}{g(y_1)}\right) + g(x)(g(y_2) - g(y_1))\left(\frac{f(y_1)}{g(y_1)} - \frac{f(x)}{g(x)}\right) \geq 0.\end{aligned}$$

Without loss of generality, let X and Y be independent and let $(X^\perp, Y^\perp)$ be an independent copy of (X, Y). According to Theorem 1.B.10 of Shaked and Shanthikumar (2007), we obtain

$$\mathbb{E}[f(Y)]\mathbb{E}[g(X)] = \mathbb{E}[f(Y^\perp)g(X^\perp)] \geq \mathbb{E}[f(X^\perp)g(Y^\perp)] = \mathbb{E}[g(Y)]\mathbb{E}[f(X)],$$

which implies the desired inequality. □

B.3 Two properties of the conditional expectation

The following two properties of the conditional expectation are used in the proof of Proposition 3(iii) in Appendix A.3.

Lemma 6. *Let X and Y be two random variables.*

(i) *$\mathbb{E}[X|Y > y]$ increases in y if and only if $\mathbb{E}[X|Y > y] \geq \mathbb{E}[X|Y = y]$ for any y.*

(ii) *If $[X|Y > y]$ increases in y in the FSD sense, then $\mathbb{E}[g(X)h(Y)|Y > y]$ is increasing in y for any nonnegative and increasing functions g and h.*

Proof. To show (i), we assume for simplicity that Y has a probability density function $f_Y(y)$. In this case,

$$\frac{\mathrm{d}}{\mathrm{d}y}\mathbb{E}[X|Y > y] = \frac{f_Y(y)}{\mathbb{P}(Y > y)} \cdot \left(\mathbb{E}[X|Y > y] - \mathbb{E}[X|Y = y]\right),$$

and the assertion follows immediately.

For (ii), if g is an increasing function and $[X|Y > y]$ increases in y in the FSD sense, then $\mathbb{E}[g(X)|Y > y]$ increases in y. According to result (i), this is equivalent to $\mathbb{E}[g(X)|Y \geq y] \geq \mathbb{E}[g(X)|Y = y]$ for any y. We obtain

$$\mathbb{E}[g(X)h(Y)|Y > y] \geq \mathbb{E}[g(X)h(y)|Y > y] \geq \mathbb{E}[g(X)h(y)|Y = y] = \mathbb{E}[g(X)h(Y)|Y = y]$$

for any y. The first inequality holds because h is increasing and g is nonnegative, and the second inequality holds because h is nonnegative. According to result (i), $\mathbb{E}[g(X)h(Y)|Y > y]$ is then increasing in y as claimed. □

B.4 Symmetry of the likelihood ratio order

Corollary 2 follows easily from Proposition 3(*iii*) because $X \uparrow_{lr} \Theta$ implies $X \uparrow_{hr} \Theta$, see Shaked and Shanthikumar (2007). It also follows from Proposition 3(*ii*) because $X \uparrow_{lr} \Theta$ implies $X \uparrow_{fsd} \Theta$ and because the likelihood ratio order is symmetric. To see the symmetry, assume that (X, Θ) has a positive joint density function to simplify the exposition. The general case can be shown in a similar way but is notationally more cumbersome. Denote by $\pi(\theta)$ the probability density function (p.d.f.) of Θ, by $f(x)$ the p.d.f. of X, and by $f(x|\theta)$ the conditional p.d.f. of X given $\Theta = \theta$. Then, the joint p.d.f. is $f(x,\theta) = f(x|\theta)\pi(\theta)$, and the conditional p.d.f. of Θ given $X = x$ is $\pi(\theta|x) = f(x|\theta)\pi(\theta)/f(x)$ by Bayes' theorem. Now $X \uparrow_{lr} \Theta$ implies that, for any $\theta_2 \geq \theta_1$, the likelihood ratio $f(x|\theta_2)/f(x|\theta_1)$ is increasing in x. Then, for any $x_2 \geq x_1$, we obtain the following chain of implications:

$$
\begin{aligned}
& \frac{f(x_2|\theta_2)}{f(x_2|\theta_1)} \geq \frac{f(x_1|\theta_2)}{f(x_1|\theta_1)} \quad \Longrightarrow \quad \frac{f(x_2|\theta_2)}{f(x_1|\theta_2)} \geq \frac{f(x_2|\theta_1)}{f(x_1|\theta_1)} \\
\Longrightarrow \quad & \frac{f(x_2,\theta_2)/\pi(\theta_2)}{f(x_1,\theta_2)/\pi(\theta_2)} \geq \frac{f(x_2,\theta_1)/\pi(\theta_1)}{f(x_1,\theta_1)/\pi(\theta_1)} \quad \Longrightarrow \quad \frac{f(x_2,\theta_2)/f(x_2)}{f(x_1,\theta_2)/f(x_1)} \geq \frac{f(x_2,\theta_1)/f(x_2)}{f(x_1,\theta_1)/f(x_1)} \\
\Longrightarrow \quad & \frac{\pi(\theta_2|x_2)}{\pi(\theta_2|x_1)} \geq \frac{\pi(\theta_1|x_2)}{\pi(\theta_1|x_1)}.
\end{aligned}
$$

Therefore, for any $x_2 \geq x_1$, the likelihood ratio $\pi(\theta|x_2)/\pi(\theta|x_1)$ is increasing in θ, which is the definition of $\Theta \uparrow_{lr} X$. The fact that $X \downarrow_{lr} \Theta$ implies $\Theta \downarrow_{lr} X$ can be shown similarly.

C Improving a suboptimal indemnity schedule with Theorem 1

The approach for improving a suboptimal indemnity schedule is similar to the idea in Chi and Wei (2020). Let $I \in \mathfrak{C}$ be an indemnity schedule that does not solve Problem (1). Let

$$
\mathcal{A} = \{x : I'(x) > 0, V_I(x) < 1+\tau\} \qquad \text{and} \qquad \mathcal{B} = \{x : I'(x) < 1, V_I(x) > 1+\tau\}
$$

be the violation areas where I does not satisfy property (6). If I were optimal, both $\mathcal{A}$ and $\mathcal{B}$ would be empty. Indemnity schedule I satisfies condition (6) on the complement of $\mathcal{A} \cup \mathcal{B}$. To improve indemnity schedule I for the policyholder, we can increase the marginal indemnity at points in $\mathcal{B}$ and reduce the marginal indemnity at points in $\mathcal{A}$. To accomplish this, we introduce the following new indemnity schedule:

$$
\hat{I}_p(x) = I(x) + p \int_0^x \left\{(1 - I'(t))\mathbf{1}_{\mathcal{B}} - I'(t)\mathbf{1}_{\mathcal{A}}\right\} \mathrm{d}t, \qquad \text{for } p \in [0,1].
$$

Clearly, $\hat{I}_p \in \mathfrak{C}$ for all $p \in [0,1]$, and $\hat{I}_p = I$ for $p = 0$. The policyholder can be made better off by moving from indemnity schedule I into the direction of the new indemnity schedule.

Proposition 11. *There exists a* $p^* \in (0,1]$ *such that* $\mathbb{E}[\phi(\mathbb{E}[u(W_I)|\Theta])] < \mathbb{E}[\phi(\mathbb{E}[u(W_{\hat{I}_{p^*}})|\Theta])]$.

Proof. Consider the policyholder's welfare under indemnity schedule $\hat{I}_p$, that is, $J(\hat{I}_p) = \mathbb{E}[\phi(\mathbb{E}[u(W_{\hat{I}_p})|\Theta])]$. Take the derivative of $J(\hat{I}_p)$ with respect to p and evaluate it at $p = 0$:

$$\frac{\partial}{\partial p} J(\hat{I}_p)\bigg|_{p=0} = \mathbb{E}[\phi(\mathbb{E}[u(W_I)|\Theta])u'(W_I)]$$
$$\cdot \left(\int_{\mathcal{A}} ((1+\tau) - V_I(x))I'(x)S_X(x)\,\mathrm{d}x + \int_{\mathcal{B}} (V_I(x) - (1+\tau))(1 - I'(x))S_X(x)\,\mathrm{d}x \right).$$

This is obtained along the lines of Appendix A.1. We then find that $\frac{\partial}{\partial p} J(\hat{I}_p)\Big|_{p=0} \geq 0$ because ϕ and u are strictly increasing and because the large round bracket is nonnegative. For $x \in \mathcal{A}$ we have $((1+\tau) - V_I(x))I'(x) > 0$, and for $x \in \mathcal{B}$ we have $(V_I(x) - (1+\tau))(1 - I'(x)) > 0$ from the definition of the sets $\mathcal{A}$ and $\mathcal{B}$. Furthermore, $\frac{\partial}{\partial p} J(\hat{I}_p)\Big|_{p=0} = 0$ if and only if the Lebesgue measure of $(\mathcal{A} \cup \mathcal{B}) \cap [0, \mathcal{M})$ is zero. If this were the case, I would already be an optimal solution to Problem (1) because I satisfies property (6) in Theorem 1 on the complement of $\mathcal{A} \cup \mathcal{B}$. As we assumed I to be suboptimal, the Lebesgue measure of $(\mathcal{A} \cup \mathcal{B}) \cap [0, \mathcal{M})$ must be strictly positive so that $\frac{\partial}{\partial p} J(\hat{I}_p)\Big|_{p=0} > 0$. In this case, a marginal increase of p above zero raises the policyholder's welfare and we can find a $p^* \in (0,1]$ such that $J(W_{\hat{I}_{p^*}}) > J(W_I)$. $\square$

It follows from the concavity of u and ϕ that the policyholder's welfare with indemnity schedule $\hat{I}_p$ is a concave function in p. In principal, it is thus possible to find the maximizer $p^* = \arg\max_{p \in [0,1]} J(\hat{I}_p)$, which yields the best improvement of I under the scheme above. Of course, $\hat{I}_{p^*}$ may not necessarily be the optimal indemnity schedule.