Risk Sharing in Monopolistic Insurance Markets: Hidden Types and Bowley-Optimal Pricing

by

Benxuan Shi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Actuarial Science and Quantitative Finance

Waterloo, Ontario, Canada, 2025

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Bin Zou Associate Professor, Dept. of Mathematics University of Connecticut
Supervisors:	Mario Ghossoub Associate Professor, Dept. of Statistics and Actuarial Science University of Waterloo
	Bin Li Associate Professor, Dept. of Statistics and Actuarial Science University of Waterloo
Internal Members:	Jun Cai Professor, Dept. of Statistics and Actuarial Science University of Waterloo
	Fangda Liu Associate Professor, Dept. of Statistics and Actuarial Science University of Waterloo
Internal-External Member:	Anqi Li Assistant Professor, Dept. of Economics University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis investigates optimal insurance design in a monopolistic market under both complete and incomplete information. The motivation stems from real-world insurance settings where insurers may either fully understand policyholders' risk characteristics or must infer them from observable behavior. We develop theoretical frameworks to model both environments and derive analytical characterizations of optimal contracts under varying assumptions.

In the first part, we analyze a classical full-information setting in which the insurer knows both the policyholder's loss distribution and risk preferences. The policyholder is assumed to be risk-averse, modeled by an increasing and concave utility function, while the insurer is risk-neutral. In a monopolistic market, the insurer, as the sole contract provider, holds significant influence over both the structure and pricing of insurance contracts. We study the impact of contract forms—such as deductibles and coinsurance—on the insurer's optimal pricing strategy, which we express through a *loading function* drawn from a class of increasing and convex functions. A central concept introduced in this framework is the *Bowley solution*, which captures the sequential nature of decision-making between the insurer and the policyholder. We relate this framework to foundational literature, particularly Chan and Gerber (1985). Our analysis shows that linear loading functions (yielding expected-value premiums) are optimal under coinsurance, while piecewise linear functions (aligned with stop-loss premiums) are optimal under deductible contracts.

The second part retains the full-information assumption but departs from traditional convex pricing rules. Instead, we introduce ambiguity in risk assessment by distorting the probability measure using a distortion function, reflecting subjective or behavioral risk perceptions. Symmetrically, the policyholder evaluates contracts using a distortion risk measure rather than expected utility. We retain the Bowley sequential structure but relax restrictions on the contract form, assuming only that indemnity schedules are uniformly Lipschitz continuous—an assumption that helps address moral hazard. Under this generalized framework, we find that full insurance becomes optimal when the policyholder is strictly risk-averse. If the policyholder evaluates risk using Value-at-Risk (VaR), the optimal contract becomes a policy limit contract with a sharp pricing distortion aligned with the VaR confidence level. For policyholders with inverse-S-shaped distortion functions (common in behavioral models), the optimal contract takes a deductible form, and the insurer's distortion partially mirrors the policyholder's up to a key threshold. These results offer insight into how non-linear transformations of risk perception shape contract design.

In the third part, we consider an incomplete information setting in which the insurer cannot observe a policyholder's risk attitude. We model heterogeneity using Yaari's dual

utility theory, parameterizing preferences via a continuum of distortion functions indexed by a type parameter θ. This setup introduces adverse selection: policyholders may misreport their type to secure better terms. To address this, the insurer must design a menu of contracts—each pairing a specific indemnity schedule and premium—to ensure *individual rationality* (voluntary participation) and *incentive compatibility* (truthful type revelation).

We formulate the insurer's profit maximization problem subject to these constraints and apply tools from mechanism design and contract theory to characterize the optimal solution. Under suitable assumptions, we find that the optimal menu consists of layered contracts with desirable properties: the most risk-averse types receive full insurance (a property known as efficiency at the top), and both coverage and pricing increase with the degree of risk aversion. The least risk-averse type is indifferent between participating and opting out, while the insurer extracts strictly positive profit from more risk-averse individuals. We also examine how the optimal menu is affected by the introduction of a fixed participation cost. In this case, the insurer chooses to withdraw part of the menu, excluding contracts targeted at the least risk-averse individuals. Additionally, we study an alternative objective in which the insurer designs an *incentive-efficient* menu—one that incorporates policyholder welfare alongside profit. We show that the layered structure remains optimal in this setting and provide a detailed characterization of the associated properties of the incentive-efficient contract menu.

Overall, this thesis contributes to the theoretical foundations of insurance economics in monopolistic markets and provides insights into the design and pricing of insurance contracts under both complete and asymmetric information.

Acknowledgements

I would like to express my heartfelt gratitude to Professors Mario Ghossoub and Bin Li for their invaluable support. Their patience, guidance, and encouragement have been instrumental throughout my Ph.D. journey.

I also extend my sincere thanks to Professors Jun Cai, Anqi Li, Fangda Liu, and Bin Zou for serving on my thesis committee. Their time, insightful comments, and constructive feedback have greatly contributed to the improvement of this thesis.

I am especially grateful to Professor Zhimin Zhang for his consistent support and encouragement over the years.

Finally, I would like to thank my parents, my cousin, and my friend Ran Liu for their unwavering love and support.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Background and Motivation

Risk-sharing is a fundamental problem in the insurance industry, enabling individuals and organizations to protect themselves against financial losses stemming from uncertain events. In competitive markets, insurers and policyholders engage in dynamic exchanges, with multiple providers offering a variety of policy options. However, when the market becomes monopolistic—dominated by a single insurer—the dynamics of risk-sharing change significantly. The absence of competition grants the monopolistic insurer considerable power over premium pricing and policy design often resulting in inefficiencies and suboptimal risk allocation.

This thesis investigates risk-sharing in monopolistic insurance markets, with particular emphasis on hidden types and Bowley-optimal pricing. Hidden types arise from asymmetric information, where the insurer cannot fully observe the policyholder's risk characteristics, complicating the task of setting fair and efficient premiums. Bowley-optimal pricing offers a framework for determining premiums that balance the insurer's financial risk with the policyholder's willingness to pay. By analyzing the pricing mechanism and asymmetric information, we aim to shed light on how monopolistic insurers can structure contracts to achieve efficient and equitable risk-sharing outcomes.

In classical actuarial science, insurance models often focus on characterizing the preferences of policyholders under uncertainty and selecting contracts that maximize the insurer's profit in monopolistic settings. A crucial aspect of this analysis involves comparing policyholders' preferences in the presence and absence of insurance coverage, referred to as *individually rationality* behavior. Such preferences are typically described using the expected utility of the policyholder or a chosen risk measure. In full-information models, this framework provides a foundation for analyzing optimal insurance design, aiming to strike a balance between adequate risk transfer for the policyholder and financial sustainability for the insurer. In contrast, under asymmetric information, the presence of hidden types leads to adverse selection. Capturing the true nature of this information asymmetry and designing *incentive-compatible* contracts becomes essential. Moreover, issues such as moral hazard play a critical role in shaping the structure and effectiveness of optimal contracts.

A critical challenge arises from moral hazard—the phenomenon where the policyholder may alter their behavior after obtaining insurance coverage, thereby increasing the insurer's exposure to risk. To address moral hazard, one tractable approach is to impose regularity conditions on the insurance contract to control incentives. For instance, consider a setting in which the insurer cannot observe the policyholder's realized loss. In this context, a 1-Lipschitz continuity condition on the indemnity function can act as a mathematical safeguard to promote behavioral stability and prevent manipulation.

Specifically, suppose the policyholder faces a random loss X, and the insurance contract specifies a reimbursement function $I(X)$. The 1-Lipschitz condition of the form

$$|I(x_1) - I(x_2)| \leq |x_1 - x_2|$$

for all x_1, x_2, also satisfies a 1-Lipschitz condition, which helps prevent underreporting or manipulation of behavior that would reduce apparent losses, as discussed in Huberman et al. (1983). This assumption will be adopted throughout the thesis to ensure that the policyholder truthfully reflects the realized loss and does not strategically alter their behavior after entering the contract.

In addition to moral hazard, this thesis also explores the challenges of premium pricing, policy design, and asymmetric information in a monopolistic insurance market. I begin by examining how each of these problems can be addressed under full information, and then extend the analysis to the more realistic setting of information asymmetry within the monopoly framework.

1.2 Bowley Optimality

Consider a monopolistic insurance market where the risk characteristics of policyholders are fully observable to the insurer. We assume that the policyholder faces a bounded random loss X and makes decisions based on expected utility maximization. The utility function u is assumed to be increasing and concave, reflecting risk-averse preferences. On the other hand, the insurer is assumed to be risk-neutral, meaning that they are indifferent to the risk associated with providing coverage and care only about profitability.

The insurance contract is represented by a function $I(X)$, where I s selected from the set of 1-Lipschitz continuous functions. The policyholder pays a premium $\pi \in \mathbb{R}_+$ to the insurance company in exchange for the coverage specified by $I(X)$. For a given premium principle π, the indemnity structure $I^*(X)$ is chosen to maximize the policyholder's welfare, which is given by

$$\max_{I \in \mathcal{I}_L} \mathbb{E}[u(w_0 - X + I(X) - \pi)], \tag{1.1}$$

where w_0 represents the initial wealth of the policyholder. If the utility function u is strictly concave, then the policyholder exhibits strict risk aversion. The solution to such a problem has been well established in the literature. If the premium π depends solely on the actuarial value $\mathbb{E}[I(X)]$, the result in Arrow (1963) demonstrates that the deductible contract, characterized by $I^*(x) = (x - d)_+$ for some $d \geq 0$ provides an optimal solution to the problem defined in (1.1). For more detailed discussions on risk sharing under the expected utility framework, see Young (1999) and Chi and Zhuang (2020). For alternative approaches, such as the rank-dependent utility framework—which applies distortions to both terminal wealth (or losses) and the probability measure—refer to Bernard et al. (2015), Xu et al. (2019), and Ghossoub (2019). Under the distortion risk measure framework, relevant analyses can be found in Cai and Chi (2020) and Boonen and Ghossoub (2021).

In a competitive insurance market, the optimal contract design typically takes into account the welfare of the policyholders, as insurers must compete to attract customers. In such markets, insurers are constrained by competition, which forces them to offer contracts that balance their own profitability with the policyholder's needs. In contrast, in a monopolistic insurance market, where a single insurer dominates, the insurer has significantly more power over contract design. With no competition, the monopolistic insurer can influence the structure of the insurance contract, including the type of coverage offered, the coverage amount, and the premium charged. This power allows the insurer to potentially set terms that maximize their own profit, but this raises the question of whether such a cost function—or premium structure—would be optimal for the insurer in a monopolistic market.

The key issue in this context is whether a premium function of the expected-value form would indeed be the optimal choice for the monopolistic insurer. While this approach might seem reasonable for the insurer to cover the cost of providing insurance, it also leaves room for the insurer to leverage their monopoly power to increase premiums beyond what would be optimal in a monpolistic setting. Thus, the monopolist's ability to set premiums and determine the structure of the insurance contract calls for careful examination. It is important to analyze whether this approach leads to an efficient allocation of risk and whether it is in the best interest of both the insurer and the policyholder. The paper RAVIV (1979) discusses another formulation, which the author refers to as a *suboptimal* solution. This alternative problem arises when the insurer is assumed to be risk-averse rather than risk-neutral. Specifically, the optimal indemnity is slected to maximize their own expected utility rather than focusing solely on cost recovery. The insurer's problem is formulated as

$$\max_{I \in \mathcal{I}_L} \mathbb{E}[v(W_0 + \pi - I(X))], \tag{1.2}$$

where W_0 represents the insurer's initial wealth, and v is a concave, increasing utility function that captures the insurer's risk-averse preferences.

Importantly, this optimization is solved under the constraint that the premium π is a function of the actuarial value $\mathbb{E}[I(X)]$, ensuring that the premium is tied to the expected indemnity payment in a structured way. It is observed that under this formulation, the optimal insurance contract takes the form of a policy limit contract, given by $I^*(x) = (x, c)$, for some threshold $c \geq 0$. In this structure, indemnity payments are capped at a fixed amount c, regardless of the magnitude of the loss.

From the above discussion, we observe a notable shift in perspective in the design of insurance contracts: rather than designing the insurance to maximize the policyholder's welfare (as is typically done in competitive markets), the contract is structured to maximize the insurer's expected utility. As a result, this shift leads to a policy that limits large payouts and stabilizes the insurer's retained losses. In this sense, the "suboptimal" label refers to the fact that the solution is not necessarily optimal from the policyholder's standpoint, but rather from the insurer's risk management perspective.

Another way to characterize the solution that reflects the insurer's bargaining power is through the concept known as the *Bowley solution*. In the context of insurance design, the Bowley solution captures a sequential decision-making process: the policyholder first decides whether to accept the offered contract based on their own utility maximization given the proposed premium, and then the insurer selects the insurance structure and corresponding premium to maximize their own objective, anticipating the policyholder's response. The seminal work CHAN and GERBER (1985) mathematically formalizes this setting in the context of optimal insurance design. In this work, the objective is to solve the following two problems:

1. The policyholder's problem: Suppose that the random variable P is a pricing density, satisfying $P > 0$ and $\mathbb{E}[P] = 1$. For a given P, the policyholder chooses the indemnity function I that solves

$$\max_{I \in \mathcal{I}_L} \mathbb{E}[u(w_0 - X + I(X) - \pi)], \tag{1.3}$$

where the premium is given by $\pi = \mathbb{E}[P(I(X))]$. Denote the optimal solution by I_P^*.

2. The insurer's problem: Having determined the mapping: $P \mapsto I_P^*$, the insurer chooses P to maximize

$$\max_{P} \mathbb{E}[v(W_0 + \pi_P - I_P^*(X))], \tag{1.4}$$

where $\pi_P = \mathbb{E}[P(I_P^*(X))]$. Denote the optimal solution by P^*.

The Bowley optimal solution is a pair $(I_{P^*}^*, P^*)$, where P^* maximizes the insurer's expected utility, subject to the constraint that the policyholder optimally responds by selecting $I_{P^*}^*$, which maximizes their own expected utility given P^*. Results concerning the Bowley solution have been extensively studied in recent literature. For continuous-time insurance optimization settings, see Chen and Shen (2018), Li and Young (2021), and Cao et al. (2022). For one-period insurance markets, related analyses appear in Cheung et al. (2019), Chi et al. (2020), and Boonen and Jiang (2024). Boonen and Ghossoub (2023) compares the Bowley solution with Pareto-optimal contracts when distortion risk measures are employed and the pricing mechanism follows that of Chan and Gerber (1985). The studies Boonen et al. (2021), Cao et al. (2023) and Ghossoub and Zhu (2024) further extend the monopoly framework by incorporating multiple insurance buyers. Beyond the monopolistic setting, the sequential game framework has also been explored in competitive markets where multiple insurers offer contracts, as discussed in Zhu et al. (2023).

Building on a similar framework, we study the design of an optimal insurance contract that constitutes a Bowley solution to a two-step optimization problem. We assume that the insurer is risk-neutral. Unlike the classical setting where premiums are determined by a pricing density, we consider a different pricing structure: the premium is given by the expectation of a convex function of the indemnity. Specifically,

$$\pi(I(X)) = \mathbb{E}[g(I(X))], \tag{1.5}$$

where $g : [0, \infty) \to [0, \infty)$ is an increasing and convex function. We refer to this pricing structure as the *convex-loaded premium principle*. Such premium principles are economically appealing, since the premium loading increases with an additional unit of indemnification, and the increment in loading will not decrease when the insurance company cedes more losses. For a given loading factor $\theta \geq 0$, common examples of such premium principles include, for instance, the expected-value premium principle, corresponding to $g(x) = (1+\theta)x$; the quadratic premium principle, corresponding to $g(x) = x + \theta x^2$; and the stop-loss premium principle, corresponding to $g(x) = x + \theta(x-s)_+$, for some $s > 0$.

In Chapter 2, we examine the Bowley-optimal pair when the feasible indemnity functions are restricted to proportional forms and deductible forms. We show that, for proportional contracts, the expected-value premium principle is the optimal convex-loaded premium principle. For deductible contracts, the optimal convex-loaded premium principle is again an expected-value type, but with a piecewise constant loading function. We refer to this structure as the stop-loss premium principle.

Beyond considering the general premium formulation given by (1.5), we also examine an alternative pricing method introduced in Wang (1996), where the premium is calculated using a distortion function—commonly known as the *distortion premium principle*. In this setting, we assume that policyholders evaluate risk by minimizing a distortion risk measure, while the insurer remains a risk-neutral expected-profit maximizer.

For the convex-loaded premium principle, the premium is determined by the expectation of a convex transformation of the indemnity function. Specifically, for each realization of the ceded loss $I(x)$, the policyholder is charged $g(I(x))$ instead of $I(x)$ itself, leading to the following premium calculation:

$$\pi(I(X)) = \int g(I(X))\, \mathrm{d}\mathbb{P},$$

where g is an increasing and convex function. In the distortion-based pricing method, rather than altering the realized indemnity $I(x)$, at each loss level, the probability distribution of the loss itself is modified through a distortion function. The premium for the contract $I(\cdot)$ is then given by

$$\pi(I(X)) = \int I(X)\, \mathrm{d}g \circ \mathbb{P}.$$

where g is a distortion function, that is, a non-decreasing function satisfying $g(0) = 0$, $g(1) = 1$. The objective can be formulated as the following two-step problem:

1. The policyholder's problem: Suppose that g is a pricing distortion function satisfying $g(0) = 0$ and $g(1) = 1$. Given g, the policyholder chooses the indemnity function I that solves

$$\max_{I \in \mathcal{I}_L} \rho(X - I(X) + \pi_g(I)),$$

where ρ is a distortion risk measure defined by $\rho(Z) = \int_0^\infty T(\overline{F}_Z(t))\mathrm{d}t$ for any non-negative random variable Z with T being a distortion function. The premium $\pi_g(I)$ is given by $\pi_g(I) = \int I(X)\,\mathrm{d}g \circ \mathbb{P}$. Denote the optimal solution by I_g.

2. The insurer's problem: Having determined the mapping $g \mapsto I_g$, the insurer chooses I_g and g to maximize its expected profit:

$$\max_{I_g, g} \pi_g(I_g) - \mathbb{E}[I_g(X)],$$

where $\pi_g(I_g) = \int I_g(X)\,\mathrm{d}g \circ \mathbb{P}$. Denote the optimal solution by $(I^*_{g^*}, g^*)$.

The Bowley optimal solution is a pair $(I^*_{g^*}, g^*)$. This setting has been previously discussed in the literature, for instance in Cheung et al. (2019). However, their analysis restricts the policyholder to be strictly risk-averse, corresponding to a concave distortion function T, or to the special case of a Value-at-Risk minimizer. In contrast, our framework extends the setting to a more general class of policyholders, without imposing any restrictions on the distortion function T.

In Chapter 3, we first analyze the policyholder's problem and show that, for a given pricing distortion g, a layered insurance contract is optimal. The structure of the layers depends on the relationship between the pricing distortion g and the distortion function T that reflects the policyholder's risk attitude. Then, we study the insurer's problem and show that, for different types of policyholders—characterized by their varying attitudes toward risk—the optimal insurance contract and the optimal pricing distortion vary accordingly. In particular, the optimal insurance contract depends on both the risk distribution and the policyholder's risk attitude, while the optimal pricing distortion depends heavily on the distortion function T. For other forms of premium principles beyond the convex-loaded expected value principle and distortion-based methods, see, for example, Deprez and Gerber (1985), Kaluszka (2001), and Chi and Tan (2013).

1.3 Hidden Risk Types

Chapters 2 and 3 contribute to the literature by examining the issue of uninsurable sources of risk, as initially discussed in Arrow (1963). In this market context, we address a different dimension of the problem faced by insurers: information asymmetry. Specifically, we focus on situations where the risk characteristics of policyholders are not fully observable by the insurer. The first paper examines the problem of asymmetric information in insurance markets is Rothschild and Stiglitz (1976), which considers a competitive setting. In this scenario, the equilibrium outcome maximizes the policyholder's welfare while ensuring zero profit for the insurer, reflecting the pressure of competition. The authors demonstrate that the inability to fully observe risk types can significantly affect the structure and existence of equilibrium in the market. The subsequent work by Stiglitz (1977) considers the problem of asymmetric information in a monopolistic insurance market. Their findings show that under such circumstances, the information about the risk distribution and the risk attitude of policyholders significantly influence the equilibrium market structure and the design of equilibrium contracts.

An important finding in Stiglitz (1977) is that, in a monopolistic setting, both the distribution of risk types and the risk preferences of policyholders influence the market equilibrium. When information about the risk distribution is unobservable, the resulting equilibrium structure is analyzed in Chade and Schlee (2012), Chade and Schlee (2020), and Gershkov et al. (2023). If the insurer lacks information about the policyholders' risk preferences or attitudes, the equilibrium is characterized in Landsberger and Meilijson (1994) and Boonen and Zhang (2021). When both the risk distribution and the risk preferences are hidden, relevant results can be found in Landsberger and Meilijson (1999), Chade et al. (2022), and Cheung et al. (2025).

Our objective is to design a set of insurance contracts — a *menu of contracts* — such that each type of policyholder selects the contract intended for them, while simultaneously maximizing the insurer's profit. Such a menu is constructed to extract the true information of the policyholders, even though they have incentives to misreport their characteristics to maximize their own benefit. Suppose that a policyholder faces a bounded random loss $X \in [0, M]$ and that heterogeneity in risk attitudes is captured by a continuous random variable $\theta \in \Theta$, with density $f(\theta)$. In this framework, the insurer designs different contracts for different types of policyholders, where the menu of contracts is denoted by $(I_\theta, p_\theta)_{\theta \in \Theta}$. Here, I_θ represents the insurance coverage provided to a type-θ policyholder, and p_θ is the corresponding price. We adopt separate notation for the premium because, in this setting, the premium is a nonlinear function of the insurance coverage I_θ. This menu must satisfy two requirements:

- Individual Rationality (IR): Each policyholder must benefit from accepting the contract designed for them, compared to having no insurance.
- Incentive Compatibility (IC): Each policyholder should prefer the contract intended for their own type over contracts designed for any other type.

When a menu of contracts satisfies the individual rationality and incentive compatibility conditions described above, it constitutes a feasible solution to the insurer's problem.

Among this set of feasible solutions, we identify the one that maximizes the insurer's expected profit—formulated similarly to (1.2)—which leads to the following optimization problem:

$$\max_{(I_\theta, p_\theta)_{\theta \in \Theta}} \int_\Theta \left(p_\theta - \mathbb{E}[I_\theta(X)]\right) f(\theta)\, \mathrm{d}\theta,$$

where $\mathbb{E}[I_\theta(X)]$ represents the insurer's expected payout to a type-θ policyholder.

In Chapter 4, we fully characterize the menu of contracts that satisfies both the individual rationality and incentive compatibility conditions. We show that in this setting, the endogenous variable p_θ plays a crucial role in ensuring these conditions are met. Under mild assumptions, we also identify the menu that maximizes the insurer's expected profit. The result reveals that the optimal insurance coverage takes a layered structure, with the specific configuration determined by the risk distribution, the distribution of types θ, and the way Yaari's utility varies with θ. We further investigate the role of friction costs in shaping the equilibrium structure, building significantly on Chade and Schlee (2020), which models such frictions through fixed costs associated with offering insurance. Finally, we study Pareto-optimal menus of contracts by incorporating the welfare of both the insurer and the policyholders into the objective function.

The rest of this thesis is organized as follows:

- Chapter 2 introduces the optimal insurance policy and pricing strategy in the context of Bowley optimality. The feasible indemnity functions considered are of deductible and proportional types, both of which satisfy the aforementioned 1-Lipschitz condition. Premiums are determined by taking the expectation of an increasing and convex function of the indemnity.

- Chapter 3 formulates the Bowley-optimal insurance strategy when the policyholder evaluates outcomes using a risk measure instead of expected utility. The admissible indemnity functions remain 1-Lipschitz continuous, and the premium is computed using a distorted probability measure.

- Chapter 4 focuses on the problem of adverse selection. In this setting, the insurer faces uncertainty regarding the policyholder's risk attitude. The optimal contract is designed to screen these hidden types and reveal true preferences through the structure of the insurance offering.

- Chapter 5 concludes the thesis by summarizing the main findings and discussing potential directions for future research.

Chapter 2

Bowley-Optimal Convex-Loaded Premium Principles

2.1 Introduction

In a monopolistic insurance market, we examine the design of insurance products by taking into account the interests of both the insurer and the policyholder. This setup corresponds to the *Stackelberg equilibrium*, also known as the *Bowley solution*. In this sequential game, the insurer and the policyholder interact in stages: given a fixed premium principle, we first solve the policyholder's demand problem to identify the set of indemnity functions that are optimal from the policyholder's perspective. Among the premium principles that support these optimal indemnities, we then select the one that maximizes the insurer's expected profit. An early discussion of such game-theoretic interactions in insurance markets appears in Chan and Gerber (1985), where the Bowley solution is studied under the assumption that the policyholder maximizes expected exponential utility, and the premium principle is determined through a pricing density.

Building on the pioneering work of Chan and Gerber (1985), this sequential game framework is extended to a continuous-time insurance optimization setting in Chen and Shen (2018), where a complete solution is obtained under the assumption that the policyholder has exponential utility. A later study, Li and Young (2021), investigates the Bowley solution by modeling the policyholder's preferences using a mean-variance functional form of terminal wealth, and derives the optimal parameter of the mean-variance premium principle accordingly. Cao et al. (2022) also examine this sequential game in a continuous-time framework by fixing the form of the mean-variance premium principle and introducing a penalty term to account for ambiguity about the loss distribution, without requiring the uniformly Lipschitz continuity condition. For one-period insurance policies, the Bowley solution is discussed in Cheung et al. (2019), where the policyholder makes decisions based on a distortion risk measure. Chi et al. (2020) investigate a similar setting to that of Chan and Gerber (1985), assuming that the indemnity function is subject to constraints on the first two moments. The aforementioned works assume a single type of policyholder in the market and no asymmetric information between the insurer and the policyholder. In contrast, Boonen et al. (2021) examine insurance policies that sequentially solve the policyholder's and insurer's problems under the assumption that there are two distinct types of policyholders, who differ in their risk appetite. Boonen and Ghossoub (2023) investigates a sequential insurance market setting and examines the relationship between the Bowley solution and Pareto-optimal contracts when distortion risk measures are employed. In their framework, the insurance contract is priced using a density function, in the spirit of Chan and Gerber (1985). Ghossoub and Zhu (2024) extend this framework by incorporating multiple insurance buyers into the analysis. Beyond the monopolistic insurance market, the sequential game framework has also been explored in settings where multiple insurers compete by offering contracts. In these markets, price competition among insurers gives rise to a different game structure, which is analyzed using the concept of *Subgame Perfect Nash Equilibrium*, as in Zhu et al. (2023).

This chapter investigates the behavior of the insurance policyholder under a sequential decision-making process in a monopolistic market. Specifically, this chapter characterizes the premium principles that are Bowley-optimal when the indemnity functions follow two of the most popular and practically relevant indemnity schedules: (i) the deductible indemnities of the form $I(x) = (x - d)_+$, for some $d \geq 0$; and (ii) the proportional indemnities of the form $I(x) = \alpha\, x$, for some $\alpha \in [0, 1]$. Both of these classes of functions belong to the class of 1-Lipschitz functions and hence satisfy the no-sabotage condition. The Bowley optima within each of these two classes of indemnity functions, using the class of convex-loaded premium principles, and assuming that the policyholder is a risk-averse EU-maximizer, while the insurer is a risk-neutral expected-profit maximizer. Regarding the optimality problem with general indemnity functions, we present a detailed discussion in the Appendix. Our major findings are summarized as follows. First, we find that the *expected-value premium principle* is Bowley optimal for the class of proportional indemnity functions. Second, the expected-value premium principle is also Bowley optimal for the class of deductible indemnity functions if the loss has decreasing mean residual life (DMRL). Third, the *stop-loss premium principle*, with the change point being the largest mean excess of loss level, is Bowley optimal for deductible indemnities under a mild condition. Methodologically, we propose a novel *dual approach* to study the Bowley optimality problem. More specifically, for a fixed indemnity function, we first obtain from the policyholder's demand problem the set of premium principles under which that indemnity function is optimal for the policyholder. Among these premium principles, we select one that maximizes the insurer's expected profit. Lastly, varying the indemnity functions, we verify the equivalence between the dual problem and the original problem.

This chapter is organized as follows. Section 2.2 sets out the model and formulates the problem. Section 2.3 examines the Bowley optimal convex-loaded premium principles when the set of acceptable indemnities is the set of proportional indemnity functions, and Section 2.4 provides a similar analysis in the case of deductible indemnity functions. Section 2.5 concludes. The Appendix provides an extension of our analysis to the general class of 1-Lipschitz indemnity functions, as well as a study of the special case of binary loss random variables.

2.2 Model Setup and Problem Formulation

2.2.1 The Market

Let X be a continuous random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, representing a random loss against which a policyholder is seeking insurance coverage. We assume that $X(\Omega) = [0, 1]$, with a probability density function $f_X(x) > 0$, for all $x \in [0, 1]$.

The market provides the policyholder the opportunity to purchase insurance coverage in return for a premium payment. Let I denote the indemnity function, that is, the portion of the loss covered by the insurer. In particular, $I(X) = X$ represents full insurance coverage, and $I(X) = 0$ represents no insurance. We assume given a set of *ex ante* admissible indemnity functions $\mathcal{I}_L$, which could be, for instance, the customary set of 1-Lipschitz indemnity schedules that satisfy the so-called no-sabotage condition. Specifically,

$$\mathcal{I}_L = \left\{ I : X(\Omega) \to X(\Omega) : I\left(0\right) = 0,\ 0 \leq I(x) - I(y) \leq x - y \text{ for all } x > y \geq 0 \right\}.$$

To issue an indemnity of $I \in \mathcal{I}_L$, the insurer charges a premium of amount $\mathbb{E}[g(I(X))]$, where $g : [0, \infty) \to [0, \infty)$ represents a *loading function*. In this paper, we consider the class of *convex-loaded premium principle*, where g is an increasing and convex function. Convex-loaded premium principles are economically appealing. For two indemnity functions $I_1, I_2 \in \mathcal{I}_L$ such that $I_2(x) \geq I_1(x)$ for all $x \in [a, b]$ over some $[a, b] \subset [0, 1]$, the convexity of g implies that

$$g'\left(I_2(x)\right) \geq g'\left(I_1(x)\right),\ \forall\, x \in [a, b].$$

That is, the marginal premium loading increases with the size of the covered loss. In particular, if $I_2(x) \geq I_1(x)$ for all $x \in [0,1]$, then $\mathbb{E}[g'(I_2(X))] \geq \mathbb{E}[g'(I_1(X))]$.

Common examples of such premium principles include the expected-value premium principle, the quadratic premium principle, and the stop-loss premium principle. Further, the admissible set of convex-loaded premium principles is given by the following set:

$$\mathcal{A}_0 = \left\{ g : X(\Omega) \to [0,\infty) \;\middle|\; g \text{ is increasing and convex, } g(0) = 0,\ g(x) \geq x, \right.$$
$$\left. g(x) \not\equiv x,\ \mathbb{E}[g(X)] < +\infty \right\}. \tag{2.1}$$

First, $g(0) = 0$ implies that a zero premium is charged for no insurance coverage. Second, $g(x) \geq x$ means that the insurer charges a non-negative risk loading. Third, the condition $g(x) \not\equiv x$ is to exclude the degenerate case with zero risk loading. Indeed, it is well known by Mossin's Theorem (e.g., Mossin (1968)) that a risk-averse, EU-maximizing policyholder will purchase full insurance if there is no risk loading.

We assume that the policyholder is a risk-averse EU-maximizer with initial wealth w_0 and a strictly increasing and strictly concave utility function $u : (-\infty, w_0) \to \mathbb{R}$. Hence, for a given indemnity function $I \in \mathcal{I}_L$ and a premium principle $g \in \mathcal{A}_0$, the policyholder's expected utility of end-of-period wealth is given by

$$\mathbb{E}\left[u\left(w_0 - X + I(X) - \mathbb{E}[g(I(X))]\right)\right]. \tag{2.2}$$

We also make the customary assumption that the insurer is a risk-neutral expected-profit maximizer. Consequently, the insurer's expected utility of end-of-period profit is given by

$$\mathbb{E}[g(I(X)) - I(X)].$$

Clearly, the policyholder's demand for insurance coverage depends on the given premium principle, which in turn affects the insurer's profit.

2.2.2 Bowley Optima

For a given $I \in \mathcal{I}_L$ and $g \in \mathcal{A}_0$, consider the following quantities:

$$U^{Pol}(I,g) = \mathbb{E}\left[u\left(w_0 - X + I(X) - \mathbb{E}[g(I(X))]\right)\right];$$

$$U^{In}(I,g) = \mathbb{E}[g(I(X))] - \mathbb{E}[I(X)].$$

Our aim is to characterize Stackelberg equilibria and examine their *ex post* efficiency. These notions are defined below.

Definition 2.2.1. *Let $\mathcal{I}_L$ be a given collection of admissible indemnity functions.*

1. *A market mechanism is a pair $(I,g) \in \mathcal{I}_L \times \mathcal{A}_0$.*

2. *A market mechanism $(I,g) \in \mathcal{I}_L \times \mathcal{A}_0$ is Individually Rational (IR) if it is preferred by both agents to the status quo:*

$$U^{Pol}(I,g) \geq U^{Pol}(0,\cdot) = \mathbb{E}[u(w_0 - X)] \quad \text{and} \quad U^{In}(I,g) \geq U^{In}(0,\cdot) = 0.$$

   Let $\mathcal{IR} \subset \mathcal{I}_L \times \mathcal{A}_0$ denote the set of all IR mechanisms.

3. *A market mechanism $(I^*,g^*) \in \mathcal{I}_L \times \mathcal{A}_0$ is Bowley Optimal (BO), or a Stackelberg Equilibrium, if:*

   (a) $I^* \in \underset{I \in \mathcal{I}_L}{\arg\max}\, U^{Pol}(I,g^*)$;

   (b) $U^{In}(I^*,g^*) \geq U^{In}(I,g)$, for all (I,g) such that $I \in \underset{\widetilde{I} \in \mathcal{I}_L}{\arg\max}\, U^{Pol}(\widetilde{I},g)$.

   Let $\mathcal{BO} \subset \mathcal{I}_L \times \mathcal{A}_0$ denote the set of all BO mechanisms.

Henceforth, we will use the following terminology:

- For a given $g \in \mathcal{A}_0$, the *policyholder's problem* is that of finding an indemnity that maximizes their expected utility of end-of-period wealth:

$$\max_I U^{Pol}(I,g). \tag{2.3}$$

  We denote by I_g an optimal indemnity function corresponding to the premium principle g.

- Given the optimal indemnity I_g above, the *insurer's problem* is that of finding an optimal premium principle:

$$\max_g U^{In}(I_g,g). \tag{2.4}$$

Furthermore, for a fixed $g \in \mathcal{A}_0$, letting

$$\mathcal{I}_g = \arg\max \ U^{Pol}(I, g),$$

it follows that $(I^*, g^*) \in \mathcal{BO}$ if and only if (I^*, g^*) is optimal for the following problem:

$$\max_{g \in \mathcal{A}_0,\, I_g \in \mathcal{I}_g} U^{In}(I_g, g). \tag{2.5}$$

It is worth noting that in order to solve problem (2.5), we propose a novel *dual approach*. Specifically, for a fixed indemnity function I, we first obtain from the policyholder's problem (2.3) the set of premium principles under which it is optimal for the policyholder to choose the indemnity function I, i.e.,

$$\mathcal{A}(I) = \left\{ g \in \mathcal{A}_0 : I \in \arg\max_{\widetilde{I} \in \mathcal{I}_L} U^{Pol}(\widetilde{I}, g) \right\}.$$

Then, under the same indemnity function I, we find a premium principle g^* that maximizes the insurer's expected profit, i.e.,

$$g^* \in \arg\max_{g \in \mathcal{A}(I)} U^{In}(I, g).$$

The following theorem verifies the equivalence between the original problem and the dual problem:

$$\max_{g \in \mathcal{A}_0} \max_{I \in \mathcal{I}_g} U^{In}(I, g) \quad \text{and} \quad \max_{I \in \mathcal{I}_L} \max_{g \in \mathcal{A}(I)} U^{In}(I, g).$$

Theorem 2.2.1. *We have*

$$\max_{g \in \mathcal{A}_0} \max_{I \in \mathcal{I}_g} U^{In}(I, g) = \max_{I \in \mathcal{I}_L} \max_{g \in \mathcal{A}(I)} U^{In}(I, g).$$

Proof. Let (I_{g^*}, g^*) denote an optimal solution to $\max_{g \in \mathcal{A}_0} \max_{I \in \mathcal{I}_g} U^{In}(I, g)$, where $I_{g^*} \in \mathcal{I}_{g^*}$. Then

$$\max_{g \in \mathcal{A}_0} \max_{I \in \mathcal{I}_g} U^{In}(I, g) = U^{In}(I_{g^*}, g^*) \leq \max_{g \in \mathcal{A}(I_{g^*})} U^{In}(I_{g^*}, g) \leq \max_{I \in \mathcal{I}} \max_{g \in \mathcal{A}(I)} U^{In}(I, g). \tag{2.6}$$

On the other hand, assume that $\max_{I \in \mathcal{I}_L} \max_{g \in \mathcal{A}(I)} U^{In}(I, g) = U^{In}(I^*, g_{I^*})$, for some $I^* \in \mathcal{I}_L$ and $g_{I^*} \in \mathcal{A}(I^*)$. Then it follows that

$$\max_{I \in \mathcal{I}_L} \max_{g \in \mathcal{A}(I)} U^{In}(I, g) = U^{In}(I^*, g_{I^*}) \leq \max_{I \in \mathcal{I}_{g_{I^*}}} U^{In}(I, g_{I^*}) \leq \max_{g \in \mathcal{A}_0} \max_{I \in \mathcal{I}_g} U^{In}(I, g). \tag{2.7}$$

Combining (2.6) and (2.7), we obtain

$$\max_{g\in\mathcal{A}_0}\max_{I\in\mathcal{I}_g} U^{In}(I,g)=\max_{I\in\mathcal{I}_L}\max_{g\in\mathcal{A}(I)} U^{In}(I,g),$$

which concludes the proof. □

We elaborate below on the utilization of the dual approach. The original problem

$$\max_{g\in\mathcal{A}_0}\max_{I\in\mathcal{I}_g} U^{In}(I,g)$$

suggests solving for the optimal indemnity function, for a given loading function g, as the first step; and, then determining the optimal loading function as a second step. However, the first step presents a significant challenge, as the optimal indemnity function I_g is inherently *implicit*, due to the generality of g. This is evident to see from our Theorems 2.3.1 and 2.4.1. Even after we restrict the indemnity structure to proportional or deductible contracts, for a general loading function g, the corresponding optimal proportional rate α_g^* and the optimal deductible level d_g^* can only be implicitly determined as solutions to some equations. This complexity makes solving the subsequent insurer's problem nearly impossible.

The dual problem reverses the order and aims to solve for the optimal loading function g for a given indemnity I as the first step. This approach is not only more feasible, as demonstrated in Theorems 2.3.2 and 2.4.2, where optimal loading functions are *explicitly* solved for proportional or deductible contracts, but is also more natural because the primary objective of this paper is to investigate the *optimality of loading functions*. The subsequent insurer's problem is a straightforward optimization with respect to some parameters (namely, the co-insurance level or the deductible level). It is worth noting that, even if future work may extend indemnities to more general forms, Theorems 2.3.2 and 2.4.2 remain valuable as they provide the optimal loading functions for proportional and deductible contracts.

2.3 Proportional Indemnity Functions

In this section, we study the policyholder's problem (2.3) and the insurer's problem (2.4), when the indemnities are restricted to be of the proportional insurance type, i.e., of the form $I(X)=\alpha X$, for some $\alpha\in[0,1]$. For a given such α and a given $g\in\mathcal{A}_0$, we denote the proportional market mechanism (I_α,g), with $I_\alpha(X)=\alpha\,X$, by (α,g).

2.3.1 The Policyholder's Problem

For a given premium principle $g \in \mathcal{A}_0$, the policyholder's problem (2.3) becomes

$$\max_{\alpha \in [0,1]} U^{Pol}(\alpha, g) := \max_{\alpha \in [0,1]} \mathbb{E}\left[u\left(w_0 - X + \alpha X - \mathbb{E}[g(\alpha X)]\right)\right], \tag{2.8}$$

where $U^{Pol}(\alpha, g)$ denotes policyholder's expected utility under the proportional market mechanism (α, g). For a given premium principle $g \in \mathcal{A}_0$, we define the set of optimal proportional indemnities for Problem (2.8) by

$$A_g = \text{argmax}_{\alpha \in [0,1]} U^{Pol}(\alpha, g). \tag{2.9}$$

The following theorem characterizes optimal proportional indemnity functions for the policyholder problem (2.8), for a given premium principle $g \in \mathcal{A}_0$.

Theorem 2.3.1. *For a given premium principle $g \in \mathcal{A}_0$, any optimal proportional indemnity function for Problem (2.8) satisfies $\alpha_g^* \in [0, 1)$ and solves the following equation:*

$$\mathbb{E}\left[g'\left(\alpha_g^* X\right) X\right] = \frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha_g^* X - \mathbb{E}[g(\alpha_g^* X)]\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha_g^* X - \mathbb{E}[g(\alpha_g^* X)]\right)\right]}. \tag{2.10}$$

Furthermore, a necessary and sufficient condition for policyholders to participate, i.e., for $\alpha_g^* \neq 0$ (hence $\alpha_g^* \in (0, 1)$), is

$$g'(0) < \frac{\mathbb{E}\left[u'\left(w_0 - X\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X\right)\right] \mathbb{E}\left[X\right]}. \tag{2.11}$$

Proof. Differentiating $U^{Pol}(\alpha, g) = \mathbb{E}\left[u\left(w_0 - X + \alpha X - \mathbb{E}[g(\alpha X)]\right)\right]$ yields

$$\frac{\partial U^{Pol}(\alpha, g)}{\partial \alpha} = \mathbb{E}\left[u'\left(w_0 - X + \alpha X - \mathbb{E}[g(\alpha X)]\right)\left(X - \mathbb{E}[g'(\alpha X)\, X]\right)\right].$$

Hence,

$$\begin{aligned} \left.\frac{\partial U^{Pol}(\alpha, g)}{\partial \alpha}\right|_{\alpha=1} &= \mathbb{E}\left[u'\left(w_0 - \mathbb{E}[g(X)]\right)\left(X - \mathbb{E}[g'(X)\, X]\right)\right] \\ &= u'\left(w_0 - \mathbb{E}[g(X)]\right)\left(\mathbb{E}[X] - \mathbb{E}[g'(X)\, X]\right) < 0, \end{aligned}$$

where the last step is due to $g \in \mathcal{A}_0$. This implies that the policyholder's expected utility decreases when α approaches to 1, and thus $\alpha_g^* < 1$. In other words, full insurance is not optimal.

Moreover, we have

$$\left.\frac{\partial U^{Pol}(\alpha, g)}{\partial \alpha}\right|_{\alpha=0} = \mathbb{E}\left[u'\left(w_0 - X\right)\left(X - \mathbb{E}[g'(0)\,X]\right)\right]$$
$$= \mathbb{E}\left[u'\left(w_0 - X\right)X\right] - g'(0)\mathbb{E}\left[u'\left(w_0 - X\right)\right]\mathbb{E}\left[X\right]. \qquad (2.12)$$

Since $U^{Pol}(\alpha, g)$ is a composition of the increasing and concave utility function $u(\cdot)$ and the function $(\alpha-1)X - \mathbb{E}[g(\alpha X)]$, that is concave in α, we know that $U^{Pol}(\alpha, g)$ is concave in α. If $g'(0) < \frac{\mathbb{E}[u'(w_0-X)X]}{\mathbb{E}[u'(w_0-X)]\mathbb{E}[X]}$, we have from (2.12) that $\left.\frac{\partial U^{Pol}(\alpha,g)}{\partial \alpha}\right|_{\alpha=0} > 0$, which implies that $\alpha_g^* \neq 0$, that is, it is optimal for policyholders to buy some insurance. Otherwise, if $g'(0) \geq \frac{\mathbb{E}[u'(w_0-X)X]}{\mathbb{E}[u'(w_0-X)]\mathbb{E}[X]}$, from again (2.12) and the concavity of $U^{Pol}(\cdot, g)$, we deduce that $\frac{\partial U^{Pol}(\alpha,g)}{\partial \alpha} \leq 0$ for all $\alpha \in [0, 1]$ and it is optimal for policyholders to not participate in the insurance contract, i.e., $\alpha_g^* = 0$. □

We note the following implications of Theorem 2.3.1. First, for any premium principle $g \in \mathcal{A}_0$, full insurance is never optimal ($\alpha_g^* \neq 1$). This is intuitive, as all admissible premium principles in $\mathcal{A}_0$ charge additional risk loading on top of the expected loss (as $g(x) \geq x$ and $g(x) \neq x$). Second, the optimal proportion α_g^* is characterized by the first-order condition (2.10). It is possible that optimal proportions may not unique, in the absence of further conditions on admissible premium principles. Third, (2.11) is the *policyholder's participation constraint*, meaning that they will not buy any insurance if this condition is not met. When condition (2.11) fails to hold, the convexity of g implies that

$$g'(x) - 1 \geq g'(0) - 1 \geq \frac{\mathbb{E}\left[u'\left(w_0 - X\right)X\right]}{\mathbb{E}\left[u'\left(w_0 - X\right)\right]\mathbb{E}\left[X\right]} - 1 \geq 0,$$

for any $x \in [0, 1]$, where the last step follows from

$$\mathbb{E}\left[u'\left(w_0 - X\right)X\right] - \mathbb{E}\left[u'\left(w_0 - X\right)\right]\mathbb{E}\left[X\right] = \mathrm{Cov}\left(u'\left(w_0 - X\right), X\right) > 0.$$

This means that policyholders will not participate if the risk loading factor $g'(x) - 1$ is too high.

2.3.2 The Insurer's Problem

By Theorem 2.3.1, it is only meaningful to consider premium principles satisfying the policyholder's participation constraint (2.11). As such, we further restrict admissible premium principles to the set

$$\mathcal{A}_1 = \left\{ g \in \mathcal{A}_0 : g'(0) < \frac{\mathbb{E}\left[u'\left(w_0 - X\right)X\right]}{\mathbb{E}\left[u'\left(w_0 - X\right)\right]\mathbb{E}\left[X\right]} \right\}.$$

According to the first-order condition (2.10), for any given $\alpha \in (0,1)$, we define the set

$$\mathcal{A}_1(\alpha) = \left\{ g \in \mathcal{A}_1 : \mathbb{E}\left[g'\left(\alpha X\right) X\right] = \frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - \mathbb{E}[g(\alpha X)]\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - \mathbb{E}[g(\alpha X)]\right)\right]} \right\}.$$

Intuitively, $\mathcal{A}_1(\alpha)$ is the set of premium principles under which α is the optimal proportion. Strictly speaking, $\mathcal{A}_1(\alpha)$ can be *larger* than the set of premium principles corresponding to the optimal proportion α, because we do not necessarily have a unique optimal solution from the first-order condition (2.10). However, by Theorem 2.3.1, we still have

$$\mathcal{A}_1 = \bigcup_{\alpha \in (0,1)} \mathcal{A}_1(\alpha).$$

In other words, any $g \in \mathcal{A}_1$ must belong to at least one set $\mathcal{A}_1(\alpha)$, for some $\alpha \in (0,1)$.

For a fixed $\alpha \in (0,1)$, we then consider what we call the *α-proportional indemnity problem* given by

$$\max_{g \in \mathcal{A}_1(\alpha)} U^{In}(\alpha, g) := \max_{g \in \mathcal{A}_1(\alpha)} \mathbb{E}\left[-\alpha X + \mathbb{E}[g(\alpha X)]\right], \tag{2.13}$$

where $U^{In}(\alpha, g)$ denotes the insurer's expected profit under the proportional market mechanism (α, g). Problem (2.13) is that of finding the premium principles in $\mathcal{A}_1(\alpha)$ that maximize the insurer's expected profit.

Theorem 2.3.2. *For a given proportion level $\alpha \in (0,1)$, an optimal premium principle for Problem (2.13) is given by*

$$g_\alpha(x) = \left(\frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right] \mathbb{E}[X]} \right) x, \tag{2.14}$$

where

$$c_\alpha = \sup \left\{ c \geq 0 : \frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c\right)\right]} - \frac{c}{\alpha} \geq 0 \right\} \tag{2.15}$$

is such that $c_\alpha \in (0, \infty)$. Furthermore, $U^{In}(\alpha, g_\alpha) > 0$ for any fixed $\alpha \in (0,1)$.

Proof. Fix $\alpha \in (0,1)$. We first show that

$$\max_{g \in \mathcal{A}_1(\alpha)} \mathbb{E}[g(\alpha X)] \leq c_\alpha.$$

For any $g \in \mathcal{A}_1(\alpha)$, by $g(0) = 0$ and the convexity of g, we have

$$\mathbb{E}[g(\alpha X)] = \int_0^1 g(\alpha x) f_X(x) \mathrm{d}x = \int_0^1 \int_0^x \alpha g'(\alpha y) \mathrm{d}y f_X(x) \mathrm{d}x$$

$$
\begin{aligned}
&\leq \int_0^1 \alpha g'(\alpha x) x f_X(x) \mathrm{d}x \\
&= \alpha \mathbb{E}[g'(\alpha X) X] \\
&= \alpha \frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - \mathbb{E}[g(\alpha X)]\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - \mathbb{E}[g(\alpha X)]\right)\right]},
\end{aligned}
$$

where the last step is by the definition of $\mathcal{A}_1(\alpha)$. As we enlarge the admissible set, it follows that

$$
\begin{aligned}
\max_{g \in \mathcal{A}_1(\alpha)} \mathbb{E}[g(\alpha X)] &\leq \left\{ \begin{array}{c} \max_g \mathbb{E}[g(\alpha X)], \\ \text{s.t. } \mathbb{E}[g(\alpha X)] \leq \alpha \frac{\mathbb{E}[u'(w_0 - X + \alpha X - \mathbb{E}[g(\alpha X)])X]}{\mathbb{E}[u'(w_0 - X + \alpha X - \mathbb{E}[g(\alpha X)])]}, \end{array} \right. \\
&\leq \sup \left\{ c \geq 0 : \alpha \frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c\right)\right]} - c \geq 0 \right\} \\
&:= c_\alpha.
\end{aligned}
$$

Further, since

$$
\frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X\right)\right]} > 0,
$$

and

$$
\limsup_{c \to \infty} \frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c\right)\right]} - \frac{c}{\alpha} \leq \limsup_{c \to \infty} \left(1 - \frac{c}{\alpha}\right) = -\infty,
$$

we deduce that $c_\alpha \in (0, \infty)$, and c_α must satisfy the condition under which the above inequality becomes an equality, i.e.,

$$
\alpha \frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right]} = c_\alpha. \tag{2.16}
$$

Next show that g_α defined in (2.14) achieves the optimality of Problem (2.13), i.e., it satisfies

$$
\mathbb{E}[g_\alpha(\alpha X)] = c_\alpha,
$$

and $g_\alpha \in \mathcal{A}_1(\alpha)$. Indeed, it is seen that

$$
\begin{aligned}
\mathbb{E}[g_\alpha(\alpha X)] &= \alpha \left(\frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right] \mathbb{E}[X]} \right) \mathbb{E}[X] \\
&= \alpha \frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right]} \\
&= c_\alpha,
\end{aligned}
$$

where the last step is by (2.16). Moreover,

$$\begin{aligned}\mathbb{E}\left[g'_\alpha(\alpha X) X\right] &= \frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right]\mathbb{E}[X]}\mathbb{E}[X] \\ &= \frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right]},\end{aligned}$$

which implies that $g_\alpha \in \mathcal{A}_1(\alpha)$. Note that g_α satisfies the policyholder's participation constraint, i.e., $g_\alpha \in \mathcal{A}_1$, that is an immediate corollary of the followed Proposition 2.3.1.

Lastly, we show that insurer has a positive expected profit under the premium principal g_α. Indeed,

$$\begin{aligned}U^{In}(\alpha, g_\alpha) &= \mathbb{E}[-\alpha X + \mathbb{E}[g_\alpha(\alpha X)]] \\ &= -\alpha\mathbb{E}[X] + \alpha\left(\frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right]\mathbb{E}[X]}\right)\mathbb{E}[X] \\ &= \alpha\frac{\mathrm{Cov}(u'\left(w_0 - X + \alpha X - c_\alpha\right), X)}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right]} \\ &> 0,\end{aligned}$$

where the last inequality is due to both $u'\left(w_0 - x + \alpha x - c_\alpha\right)$ and x are strictly increasing in x. □

Theorem 2.3.2 shows that the *expected-value premium principle* with loading

$$\frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right]\mathbb{E}[X]} - 1$$

is optimal for Problem (2.13). The insurer is better off by offering the mechanism (α, g_α), as the expected profit $U^{In}(\alpha, g_\alpha)$ is positive.

Recall that the set of optimal proportions for Problem (2.8), for a given premium principle g, is denoted by A_g, see (2.9). The following proposition shows that, under the premium principle g_α, α is the *unique* optimal proportion level for the policyholder.

Proposition 2.3.1. *For any $\alpha \in (0,1)$, we have $A_{g_\alpha} = \{\alpha\}$.*

Proof. Fix $\alpha \in (0,1)$. For an arbitrary $\widetilde{\alpha} \in [0,1]$, we have

$$U^{Pol}(\widetilde{\alpha}, g_\alpha) = \mathbb{E}\left[u\left(w_0 - X + \widetilde{\alpha} X - \mathbb{E}[g_\alpha(\widetilde{\alpha} X)]\right)\right]$$

$$= \mathbb{E}\left[u\left(w_0 + (\widetilde{\alpha}-1)X - \widetilde{\alpha}\frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right]}\right)\right].$$

Since the utility function $u(\cdot)$ is strictly increasing and concave and

$$\left(\widetilde{\alpha}-1\right)X - \widetilde{\alpha}\left(\frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - \pi_\alpha\right)X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - \pi_\alpha\right)\right]}\right)$$

is linear in $\widetilde{\alpha}$, we deduce that $U^{Pol}(\widetilde{\alpha}, g_\alpha)$ is strictly concave in $\widetilde{\alpha}$. We then show that the maximum of $U^{Pol}(\widetilde{\alpha}, g_\alpha)$ is reached at $\widetilde{\alpha} = \alpha$ via the first-order condition. Indeed,

$$\begin{aligned}
&\left.\frac{\partial U^{Pol}(\widetilde{\alpha}, g_\alpha)}{\partial \widetilde{\alpha}}\right|_{\widetilde{\alpha}=\alpha} \\
&= \mathbb{E}\left[u'\left(w_0 + (\alpha-1)X - \alpha\frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right]}\right)\left(X - \frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right]}\right)\right] \\
&= \mathbb{E}\left[u'\left(w_0 + (\alpha-1)X - c_\alpha\right)\left(X - \frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right]}\right)\right] \\
&= \mathbb{E}\left[u'\left(w_0 + (\alpha-1)X - c_\alpha\right)X\right] - \mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right]\frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - c_\alpha\right)\right]} \\
&= 0,
\end{aligned}$$

where we used the definition of c_α in the second equality. This completes the proof. □

The following theorem verifies the equivalence between the original problem $\max\limits_{g\in\mathcal{A}_1}\max\limits_{\alpha\in A_g} U^{In}(\alpha, g)$ and the dual problem $\max\limits_{\alpha\in(0,1)}\max\limits_{g\in\mathcal{A}_1(\alpha)} U^{In}(\alpha, g)$. The proof is similar as Theorem 2.2.1.

Theorem 2.3.3. *For proportional indemnities,*

$$\max_{g\in\mathcal{A}_1}\max_{\alpha\in A_g} U^{In}(\alpha, g) = \max_{\alpha\in(0,1)}\max_{g\in\mathcal{A}_1(\alpha)} U^{In}(\alpha, g) = \max_{\alpha\in(0,1)} U^{In}(\alpha, g_\alpha). \tag{2.17}$$

Proof. First, by Theorem 2.3.2, it is known that for any $\alpha \in (0, 1)$,

$$\max_{g\in\mathcal{A}_1(\alpha)} U^{In}(\alpha, g) = U^{In}(\alpha, g_\alpha).$$

As such, the second equality in (2.17) is immediate.

Next we prove the first equality. Let (α_{g^*}, g^*) be an optimal solution to $\max\limits_{g\in\mathcal{A}_1}\max\limits_{\alpha\in A_g} U^{In}(\alpha, g)$, where $\alpha_{g^*} \in A_{g^*}$. From Theorem 2.3.1, we know that $\alpha_{g^*} \in (0, 1)$, and this mechanism

(α_{g^*}, g^*) must satisfy the first-order condition (2.10), hence implying that $g^* \in \mathcal{A}_1(\alpha_{g^*})$. Thus,

$$\max_{g\in\mathcal{A}_1}\max_{\alpha\in A_g} U^{In}(\alpha, g) = U^{In}\left(\alpha_{g^*}, g^*\right) \leq \max_{g\in\mathcal{A}_1(\alpha_{g^*})} U^{In}\left(\alpha_{g^*}, g\right) \leq \max_{\alpha\in(0,1)}\max_{g\in\mathcal{A}_1(\alpha)} U^{In}\left(\alpha, g\right). \tag{2.18}$$

On the other hand, assume that $\max_{\alpha\in(0,1)} U^{In}\left(\alpha, g_\alpha\right) = U^{In}\left(\alpha^*, g_{\alpha^*}\right)$. Then

$$\max_{\alpha\in(0,1)}\max_{g\in\mathcal{A}_1(\alpha)} U^{In}\left(\alpha, g\right) = U^{In}\left(\alpha^*, g_{\alpha^*}\right) = \max_{\alpha\in A_{g_{\alpha^*}}} U^{In}\left(\alpha, g_{\alpha^*}\right) \leq \max_{g\in\mathcal{A}_1}\max_{\alpha\in A_g} U^{In}\left(\alpha, g\right), \tag{2.19}$$

where we used Proposition 2.3.1 in the second equality. Combining (2.18) and (2.19), we obtain that

$$\max_{g\in\mathcal{A}_1}\max_{\alpha\in A_g} U^{In}\left(\alpha, g\right) = \max_{\alpha\in(0,1)}\max_{g\in\mathcal{A}_1(\alpha)} U^{In}\left(\alpha, g\right),$$

which concludes the proof. □

The first equality in (2.17) is a kind of "*dual representation*" of the insurer's value function. The inside α-proportional indemnity problem $\max_{g\in\mathcal{A}_1(\alpha)} U^{In}\left(\alpha, g\right)$ has been solved by Theorem 2.3.2, and the outside problem is a trivial optimization with respect to the parameter $\alpha \in (0, 1)$. We conclude from Theorem 2.3.3 that the *expected-value premium principle* is a Bowley optimal premium principle for proportional insurance contracts.

Example 2.3.1. *Suppose that X follows a truncated exponential distribution supported on $[0, 1]$, with survival function given by*

$$\overline{F}_X(x) = \frac{e^{-\lambda x} - e^{-\lambda}}{1 - e^{-\lambda}}, \quad x \in [0, 1],$$

where $\lambda > 0$. Assume the exponential utility $u(x) = \frac{1-e^{-\gamma x}}{\gamma}$ with risk aversion parameter $\gamma > 0$. Figure 2.1 plots the insurer's expected profit $U^{In}(\alpha, g_\alpha)$, as a function of $\alpha \in (0, 1)$, where we set $\lambda = 7$ and $\gamma = 9$. It is seen that the highest profit can be reached at $\alpha = 0.366$.

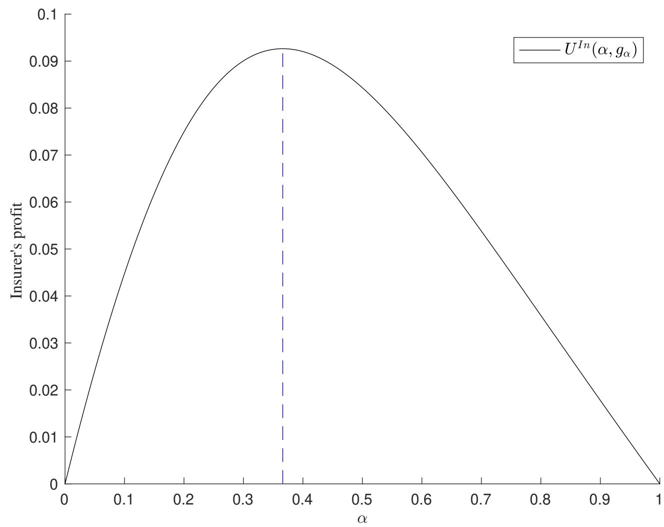


Figure 2.1: Insurer's expected profit under the expected-value premium principle

At the end of this section, we conclude that the derived mechanism (α^*, g_{α^*}) is Bowley optimal and individually rational.

Theorem 2.3.4. *Consider the proportional mechanism* (α^*, g_{α^*})*, where* $\alpha^* \in \arg\max_{\alpha\in(0,1)} U^{In}(\alpha, g_\alpha)$*. We have* $(\alpha^*, g_{\alpha^*}) \in \mathcal{BO} \cap \mathcal{IR}$*.*

Proof. It follows from Proposition 2.3.1 that

$$U^{Pol}(\alpha^*, g_{\alpha^*}) = \max_{\alpha\in[0,1]} U^{Pol}(\alpha, g_{\alpha^*}).$$

By Theorem 2.3.3, we have

$$\max_{g\in\mathcal{A}_1,\, \alpha\in A_g} U^{In}(\alpha, g) = \max_{\alpha\in(0,1)} \max_{g\in\mathcal{A}_1(\alpha)} U^{In}(\alpha, g) = \max_{\alpha\in(0,1)} U^{In}(\alpha, g_\alpha) = U^{In}(\alpha^*, g_{\alpha^*}),$$

where the last equality is due to $\alpha^* \in \arg\max_{\alpha\in(0,1)} U^{In}(\alpha, g_\alpha)$. Therefore, $(\alpha^*, g_{\alpha^*}) \in \mathcal{BO}$.

On the other hand, since $g_{\alpha^*}(x) \geq x$, we have

$$U^{In}(\alpha^*, g_{\alpha^*}) = \mathbb{E}[g_{\alpha^*}(\alpha^* X)] - \mathbb{E}[\alpha^* X] \geq \mathbb{E}[\alpha^* X] - \mathbb{E}[\alpha^* X] = 0 = U^{In}(0, \cdot).$$

Moreover, we know from Proposition 2.3.1 that $A_{g_{\alpha^*}} = \{\alpha^*\}$ implying that the proportional level α^* is the policyholder's optimal choice for Problem (2.8) under the premium principle g_{α^*}. Thus,

$$\begin{aligned} U^{Pol}(\alpha^*, g_{\alpha^*}) &= \mathbb{E}\left[u(w_0 - X + \alpha^* X - \mathbb{E}[g_{\alpha^*}(\alpha^* X)]\right] \\ &\geq \mathbb{E}\left[u(w_0 - X + 0 - \mathbb{E}[g_{\alpha^*}(0)]\right] = \mathbb{E}\left[u(w_0 - X)\right] = U^{Pol}(0, \cdot), \end{aligned}$$

and so $(\alpha^*, g_{\alpha^*}) \in \mathcal{IR}$. □

2.4 Deductible Indemnity Functions

In this section, we study the policyholder's problem (2.3) and the insurer's problem (2.4), when the indemnities are restricted to be of the deductible insurance type, i.e., of the form $I(X) = (X - d)_+$ for some $d \in [0, 1]$. For a given such d and a given $g \in \mathcal{A}_0$, we denote the deductible market mechanism (I_d, g), with $I_d(X) = (X - d)_+$, by (d, g).

2.4.1 The Policyholder's Problem

For a given premium principle $g \in \mathcal{A}_0$, under deductible contracts $I(X) = (X - d)_+$, the policyholder's objective function (2.3) is given by

$$\max_{d \in [0,1]} U^{Pol}(d, g) := \max_{d \in [0,1]} \mathbb{E}\left[u\left(w_0 - X + (X - d)_+ - \mathbb{E}[g((X - d)_+)]\right)\right], \tag{2.20}$$

where $U^{Pol}(d, g)$ denotes the policyholder's expected utility under the deductible market mechanism (d, g). We define the set of optimal deductible levels for a given $g \in \mathcal{A}_0$ by

$$D_g = \text{argmax}_{d \in [0,1]} U^{Pol}(d, g). \tag{2.21}$$

The following result provides a characterization of the solution to the policyholder's problem (2.20).

Theorem 2.4.1. *For a given premium principle $g \in \mathcal{A}_0$, any optimal deductible level d_g^* for Problem (2.20) is such that $d_g^* \in (0, 1]$, and it satisfies the equation*

$$\mathbb{E}\left[g'(X - d_g^*) \mid X > d_g^*\right] = \frac{u'\left(w_0 - d_g^* - \mathbb{E}[g((X - d_g^*)_+)]\right)}{\mathbb{E}\left[u'\left(w_0 - X + (X - d_g^*)_+ - \mathbb{E}[g((X - d_g^*)_+)]\right)\right]}. \tag{2.22}$$

Moreover, a sufficient condition for the policyholder to participate, i.e., for $d_g^* \neq 1$ (that is, $d_g^* \in (0,1)$), is given by

$$g'(0) < \frac{u'(w_0 - 1)}{\mathbb{E}\left[u'(w_0 - X)\right]}. \tag{2.23}$$

Proof. By (2.20), the policyholder's expected utility can be rewritten as

$$\begin{aligned} U^{Pol}(d,g) &= \mathbb{E}\left[u\left(w_0 - X + (X-d)_+ - \mathbb{E}[g((X-d)_+)]\right)\right] \\ &= \int_0^d u\left(w_0 - x - \mathbb{E}[g((X-d)_+)]\right) f_X(x)\mathrm{d}x + u\left(w_0 - d - \mathbb{E}[g((X-d)_+)]\right) \overline{F}_X(d), \end{aligned}$$

where $\overline{F}_X(d) := \mathbb{P}(X > d)$. Differentiating $U^{Pol}(d,g)$ yields

$$\begin{aligned} \frac{\partial U^{Pol}(d,g)}{\partial d} &= u\left(w_0 - d - \mathbb{E}[g((X-d)_+)]\right) f_X(d) \\ &- \int_0^d u'\left(w_0 - x - \mathbb{E}[g((X-d)_+)]\right) f_X(x)\mathrm{d}x \frac{\partial \mathbb{E}[g((X-d)_+)]}{\partial d} \\ &- u'\left(w_0 - d - \mathbb{E}[g((X-d)_+)]\right) \overline{F}_X(d) \left(1 + \frac{\partial \mathbb{E}[g((X-d)_+)]}{\partial d}\right) \\ &- u\left(w_0 - d - \mathbb{E}[g((X-d)_+)]\right) f_X(d) \\ &= -u'\left(w_0 - d - \mathbb{E}[g((X-d)_+)]\right) \overline{F}_X(d) \\ &+ \int_d^1 g'(x-d) f_X(x)\mathrm{d}x \cdot \mathbb{E}\left[u'\left(w_0 - X + (X-d)_+ - \mathbb{E}[g((X-d)_+)]\right)\right]. \end{aligned}$$

It is seen that the first-order condition $\frac{\partial U^{Pol}(d,g)}{\partial d} = 0$ implies

$$\frac{u'\left(w_0 - d - \mathbb{E}[g((X-d)_+)]\right)}{\mathbb{E}\left[u'\left(w_0 - X + (X-d)_+ - \mathbb{E}[g((X-d)_+)]\right)\right]} = \frac{\int_d^1 g'(x-d) f_X(x)\mathrm{d}x}{\overline{F}_X(d)} = \mathbb{E}[g'(X-d) | X > d].$$

Moreover,

$$\left.\frac{\partial U^{Pol}(d,g)}{\partial d}\right|_{d=0} = -u'(w_0 - \mathbb{E}[g(X)])\left(1 - \mathbb{E}\left[g'(X)\right]\right) > 0,$$

and $\left.\frac{\partial U^{Pol}(d,g)}{\partial d}\right|_{d=1} = 0$. To further determine the monotonicity of $U^{Pol}(d,g)$ at $d=1$, we take the second-order derivative and obtain

$$\left.\frac{\partial^2 U^{Pol}(d,g)}{\partial d^2}\right|_{d=1} = f_X(1)\left(u'(w_0 - 1) - g'(0)\mathbb{E}\left[u'(w_0 - X)\right]\right) > 0,$$

where the inequality is due to condition (2.23). From $\left.\frac{\partial U^{Pol}(d,g)}{\partial d}\right|_{d=1} = 0$ and $\left.\frac{\partial^2 U^{Pol}(d,g)}{\partial d^2}\right|_{d=1} > 0$, we deduce that $d_g^* \neq 1$. □

Note that the policyholder's participation constraint $g'(0) < \frac{u'(w_0-1)}{\mathbb{E}[u'(w_0-X)]}$ given in (2.23) is similar to (2.11) for proportional insurance contracts. The main difference is that (2.23) is only a sufficient but not necessary condition to ensure the policyholder's participation in *deductible contracts*, while (2.11) is both a necessary and sufficient condition for the policyholder's participation in *proportional contracts*. Moreover, it is also easily seen that the upper bound $\frac{u'(w_0-1)}{\mathbb{E}[u'(w_0-X)]}$ is greater than 1, indicating positive risk loading.

2.4.2 The Insurer's Problem

By Theorem 2.4.1, we further restrict the set of admissible premium principles to those that satisfy the participation constraint (2.23), and we hence consider the set to

$$\mathcal{A}_2 = \left\{ g \in \mathcal{A}_0 : g'(0) < \frac{u'(w_0 - 1)}{\mathbb{E}\left[u'(w_0 - X)\right]} \right\}.$$

We then look within this set for a premium principle that maximizes the insurer's expected profit.

Following the first-order condition (2.22), for a given $d \in (0, 1)$, we define the following set

$$\mathcal{A}_2(d) = \left\{ g \in \mathcal{A}_2 : \mathbb{E}\left[g'(X - d) \mid X > d\right] = \frac{u'\left(w_0 - d - \mathbb{E}[g((X - d)_+)]\right)}{\mathbb{E}\left[u'\left(w_0 - X + (X - d)_+ - \mathbb{E}[g((X - d)_+)]\right)\right]} \right\},$$

which includes all premium principles under which d is the optimal deductible level. Similarly, we have

$$\mathcal{A}_2 = \bigcup_{d \in (0,1)} \mathcal{A}_2(d).$$

For a fixed deductible level $d \in (0, 1)$, we then consider what we call the *d-deductible contract problem* given by

$$\max_{g \in \mathcal{A}_2(d)} U^{In}(d, g) := \max_{g \in \mathcal{A}_2(d)} \mathbb{E}\left[-(X - d)_+\right] + \mathbb{E}[g((X - d)_+)], \tag{2.24}$$

where $U^{In}(d, g)$ denotes insurer's expected profit under the deductible mechanism (d, g). For ease of notation, we denote the *mean excess of loss* function of a loss X by

$$M_X(x) = \mathbb{E}[X - x | X > x].$$

Theorem 2.4.2. *For a given deductible level $d \in (0,1)$, an optimal premium loading function for Problem (2.24) is given by*

$$g_d(x) = \begin{cases} x, & 0 \leq x \leq z_d^*, \\ (1+K_d)x - K_d z_d^*, & z_d^* < x \leq 1-d, \end{cases} \tag{2.25}$$

where

$$z_d^* = \text{argmax}_{z \in [0,1-d)} M_X(d+z),$$

and

$$K_d = \left(\frac{u'(w_0 - d - c_d)}{\mathbb{E}\left[u'(w_0 - X + (X-d)_+ - c_d)\right]} - 1 \right) \frac{\overline{F}_X(d)}{\overline{F}_X(d+z_d^*)},$$

with

$$c_d = \sup\left\{ c \geq 0 : \frac{u'(w_0 - d - c)\,\overline{F}_X(d) M_X(d+z_d^*)}{\mathbb{E}\left[u'(w_0 - X + (X-d)_+ - c)\right]} - \overline{F}_X(d)(M_X(d+z_d^*) - M_X(d)) - c \geq 0 \right\}$$

is such that $c_d \in (0,\infty)$. Furthermore, $U^{In}(d, g_d) > 0$ for any $d \in (0,1)$.

Proof. To maximize the insurer's expected profit, it suffices to consider the part

$$\begin{aligned} \mathbb{E}[g((X-d)_+)] &= -\int_d^1 g(x-d)\mathrm{d}\overline{F}_X(x) \\ &= \int_d^1 \overline{F}_X(x) g'(x-d)\mathrm{d}x \\ &= \int_0^{1-d} g'(y)\overline{F}_X(y+d)\mathrm{d}y. \end{aligned}$$

To optimize g, we separate it into two parts $g'(0)$ and $g'(y) - g'(0) := h(y)$, and obtain

$$\mathbb{E}[g((X-d)_+)] = \int_0^{1-d} h(y)\overline{F}_X(y+d)\mathrm{d}y + g'(0)\int_0^{1-d} \overline{F}_X(y+d)\mathrm{d}y. \tag{2.26}$$

The first-order condition in $\mathcal{A}_2(d)$ can be rewritten as

$$\begin{aligned} \frac{u'(w_0 - d - \mathbb{E}[g((X-d)_+)])}{\mathbb{E}\left[u'(w_0 - X + (X-d)_+ - \mathbb{E}[g((X-d)_+)])\right]} &= \mathbb{E}[g'(X-d)|X>d] \\ &= \frac{\int_d^1 g'(x-d) f_X(x)\mathrm{d}x}{\overline{F}_X(d)} \end{aligned}$$

$$= \frac{\int_0^{1-d} h(y) f_X(y+d)\mathrm{d}y}{\overline{F}_X(d)} + g'(0). \quad (2.27)$$

Next we aim to find an upper bound of $\mathbb{E}[g((X-d)_+)]$ for $g \in \mathcal{A}_2(d)$. For any $h \in \mathcal{A}_h := \{h(y) = g'(y) - g'(0) : g \in \mathcal{A}_2\}$, we construct an increasing sequence of step functions $\{h_n\}_{n\in\mathbb{N}}$ to approximate h as follows. Let $\{\mathcal{P}_n\}_{n\in\mathbb{N}}$ be an increasing sequence of partitions of $[0, 1-d]$, that is $\mathcal{P}_n = \{z_{0,n}, z_{1,n}, \ldots, z_{n,n}\}$ such that $z_{0,n} = 0$, $z_{n,n} = 1-d$, $\mathcal{P}_n \subset \mathcal{P}_{n+1}$, and $\lim_{n\to\infty} \max_{k=1,\ldots,n} (z_{k,n} - z_{k-1,n}) = 0$. Define

$$h_n(y) := \sum_{k=1}^{n} a_k \, \mathbb{I}_{(z_{k-1,n}, z_{k,n})}(y),$$

where $a_k := h(z_{k-1,n})$, for $k = 1, \ldots, n$. By this construction, we have $h_n(y) \uparrow h(y)$ as $n \to \infty$, a.e.. Let $b_k := \int_{z_{k,n}}^{1-d} f_X(y+d)\mathrm{d}y$. Note that $a_1 = 0$ and $b_n = 0$. It follows that

$$\begin{aligned}
\int_0^{1-d} h_n(y) f_X(y+d)\mathrm{d}y &= \sum_{k=2}^{n} a_k \int_{z_{k-1,n}}^{z_{k,n}} f_X(y+d)\mathrm{d}y = \sum_{k=2}^{n} a_k(b_{k-1} - b_k) = \sum_{k=2}^{n} a_k b_{k-1} - \sum_{k=2}^{n-1} a_k b_k \\
&= \sum_{k=2}^{n} a_k b_{k-1} - \sum_{k=3}^{n} a_{k-1} b_{k-1} = \sum_{k=2}^{n} a_k b_{k-1} - \sum_{k=2}^{n} a_{k-1} b_{k-1} \\
&= \sum_{k=2}^{n} (a_k - a_{k-1}) \overline{F}_X(d + z_{k-1,n}), \qquad (2.28)
\end{aligned}$$

where we used $a_1 = 0$ in the first and the second-to-last equality, and $b_n = 0$ in the third equality. Now, since $h_n(y) \le h(y)$, a.e., (2.28) and the constraint in (2.27) give

$$\begin{aligned}
&\sum_{k=2}^{n} (a_k - a_{k-1}) \, \overline{F}_X(d + z_{k-1,n}) \\
&\quad \le \overline{F}_X(d) \left(\frac{u'\left(w_0 - d - \mathbb{E}[g((X-d)_+)]\right)}{\mathbb{E}\left[u'\left(w_0 - X + (X-d)_+ - \mathbb{E}[g((X-d)_+)]\right)\right]} - g'(0) \right).
\end{aligned} \qquad (2.29)$$

Let $c_k := \int_{z_{k,n}}^{1-d} \overline{F}_X(y+d)\,\mathrm{d}y$, with $c_n = 0$. Similarly to (2.28), we have

$$\begin{aligned}
&\int_0^{1-d} h_n(y) \overline{F}_X(y+d)\mathrm{d}y \\
&= \sum_{k=2}^{n} a_k \int_{z_{k-1,n}}^{z_{k,n}} \overline{F}_X(y+d)\mathrm{d}y = \sum_{k=2}^{n} a_k (c_{k-1} - c_k) = \sum_{k=2}^{n} (a_k - a_{k-1}) c_{k-1}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=2}^{n}(a_k - a_{k-1}) \int_{z_{k-1,n}}^{1-d} \overline{F}_X(y+d)\mathrm{d}y \\
&= \sum_{k=2}^{n}(a_k - a_{k-1})\overline{F}_X(d+z_{k-1,n})M_X(d+z_{k-1,n}) \\
&\leq \overline{F}_X(d)\left(\frac{u'\left(w_0 - d - \mathbb{E}[g((X-d)_+)]\right)}{\mathbb{E}\left[u'\left(w_0 - X + (X-d)_+ - \mathbb{E}[g((X-d)_+)]\right)\right]} - g'(0)\right) M_X(d+z_d^*),
\end{aligned}
$$

where

$$
z_d^* := \underset{z\in[0,1-d)}{\arg\max} M_X(d+z),
$$

and the last step is by (2.29). The monotone convergence theorem yields that, for any $h \in \mathcal{A}_h$,

$$
\begin{aligned}
&\int_0^{1-d} h(y)\,\overline{F}_X(y+d)\mathrm{d}y \\
&\leq \overline{F}_X(d)\left(\frac{u'\left(w_0 - d - \mathbb{E}[g((X-d)_+)]\right)}{\mathbb{E}\left[u'\left(w_0 - X + (X-d)_+ - \mathbb{E}[g((X-d)_+)]\right)\right]} - g'(0)\right) M_X(d+z_d^*).
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\mathbb{E}[g((X-d)_+)] = \int_0^{1-d} h(y)\overline{F}_X(y+d)\mathrm{d}y + g'(0)\int_0^{1-d}\overline{F}_X(y+d)\mathrm{d}y \\
&\leq \overline{F}_X(d)\left(\frac{u'\left(w_0 - d - \mathbb{E}[g((X-d)_+)]\right)}{\mathbb{E}\left[u'\left(w_0 - X + (X-d)_+ - \mathbb{E}[g((X-d)_+)]\right)\right]} - g'(0)\right) M_X(d+z_d^*) \\
&+ g'(0)\int_0^{1-d}\overline{F}_X(y+d)\mathrm{d}y \\
&= \overline{F}_X(d)\left(\frac{u'\left(w_0 - d - \mathbb{E}[g((X-d)_+)]\right)}{\mathbb{E}\left[u'\left(w_0 - X + (X-d)_+ - \mathbb{E}[g((X-d)_+)]\right)\right]} - g'(0)\right) M_X(d+z_d^*) \\
&+ \overline{F}_X(d)g'(0)M_X(d) \\
&\leq \frac{u'\left(w_0 - d - \mathbb{E}[g((X-d)_+)]\right)\overline{F}_X(d)M_X(d+z_d^*)}{\mathbb{E}\left[u'\left(w_0 - X + (X-d)_+ - \mathbb{E}[g((X-d)_+)]\right)\right]} \\
&- \overline{F}_X(d)(M_X(d+z_d^*) - M_X(d)), \qquad (2.30)
\end{aligned}
$$

where the last step is by $g'(0) \geq 1$ and $M_X(d+z_d^*) \geq M_X(d)$. It follows that

$$
\max_{g\in\mathcal{A}_2(d)} \mathbb{E}[g((X-d)_+)]
$$

$$
\leq \left\{ \begin{array}{c} \max_{g \in \mathcal{A}_2} \mathbb{E}[g((X-d)_+)] \\ \text{s.t. } \mathbb{E}[g((X-d)_+)] \leq \frac{u'(w_0 - d - \mathbb{E}[g((X-d)_+)])\overline{F}_X(d) M_X(d+z_d^*)}{\mathbb{E}[u'(w_0 - X + (X-d)_+ - \mathbb{E}[g((X-d)_+)])]} - \overline{F}_X(d)(M_X(d+z_d^*) - M_X(d)) \end{array} \right.
$$

$$
\leq \sup \left\{ c \geq 0 : c \leq \frac{u'\left(w_0 - d - c\right)\overline{F}_X(d) M_X(d+z_d^*)}{\mathbb{E}\left[u'\left(w_0 - X + (X-d)_+ - c\right)\right]} - \overline{F}_X(d)(M_X(d+z_d^*) - M_X(d)) \right\}
$$

$$
:= c_d. \tag{2.31}
$$

Since

$$
\begin{aligned}
&\frac{u'\left(w_0 - d\right)\overline{F}_X(d) M_X(d+z_d^*)}{\mathbb{E}\left[u'\left(w_0 - X + (X-d)_+\right)\right]} - \overline{F}_X(d)(M_X(d+z_d^*) - M_X(d)) \\
&\geq \overline{F}_X(d) M_X(d+z_d^*) - \overline{F}_X(d)(M_X(d+z_d^*) - M_X(d)), \\
&> 0,
\end{aligned}
$$

and

$$
\frac{u'\left(w_0 - d - c\right)\overline{F}_X(d)}{\mathbb{E}\left[u'\left(w_0 - X + (X-d)_+ - c\right)\right]} = \frac{u'\left(w_0 - d - c\right)\overline{F}_X(d)}{\int_0^d u'(w_0 - x - c) f_X(x)\mathrm{d}x + u'\left(w_0 - d - c\right)\overline{F}_X(d)} < 1,
$$

we deduce that $c_d \in (0, \infty)$.

Then we verify that the upper bound c_d in (2.31) can be attained by the loading function g_d given in (2.25). By (2.25) and (2.26),

$$
\begin{aligned}
&\mathbb{E}[g_d((X-d)_+)] = \int_0^{1-d} (g_d'(y) - g_d'(0))\overline{F}_X(y+d)\mathrm{d}y + g_d'(0)\int_0^{1-d} \overline{F}_X(y+d)\mathrm{d}y \\
&= (K_d + 1)\int_{z_d^*}^{1-d} \overline{F}_X(y+d)\mathrm{d}y + \int_0^{z_d^*} \overline{F}_X(y+d)\mathrm{d}y \\
&= \left(\left(\frac{u'\left(w_0 - d - c_d\right)}{\mathbb{E}\left[u'\left(w_0 - X + (X-d)_+ - c_d\right)\right]} - 1 \right) \frac{\overline{F}_X(d)}{\overline{F}_X(d+z_d^*)} + 1 \right) \int_{z_d^*}^{1-d} \overline{F}_X(y+d)\mathrm{d}y \\
&+ \int_0^{z_d^*} \overline{F}_X(y+d)\mathrm{d}y \\
&= \left(\frac{u'\left(w_0 - d - c_d\right)}{\mathbb{E}\left[u'\left(w_0 - X + (X-d)_+ - c_d\right)\right]} - 1 \right) \overline{F}_X(d) \frac{\int_{z_d^*}^{1-d} \overline{F}_X(y+d)\mathrm{d}y}{\overline{F}_X(d+z_d^*)} + \int_0^{1-d} \overline{F}_X(y+d)\mathrm{d}y \\
&= \frac{u'\left(w_0 - d - c_d\right)\overline{F}_X(d) M_X(d+z_d^*)}{\mathbb{E}\left[u'\left(w_0 - X + (X-d)_+ - c_d\right)\right]} - \overline{F}_X(d) M_X(d+z_d^*) + \overline{F}_X(d) M_X(d) \\
&= c_d.
\end{aligned}
$$

Moreover, it is straightforward to verify that g_d satisfies the first-order condition (2.27). Indeed, if $z_d^* > 0$, we have

$$\begin{aligned}\frac{\int_0^{1-d}\left(g_d'(y)-g_d'(0)\right)f_X(y+d)\mathrm{d}y}{\overline{F}_X(d)}+g_d'(0)&=K_d\frac{\int_{z_d^*}^{1-d}f_X(y+d)\mathrm{d}y}{\overline{F}_X(d)}+1\\&=K_d\frac{\overline{F}_X(d+z_d^*)}{\overline{F}_X(d)}+1\\&=\frac{u'\left(w_0-d-c_d\right)}{\mathbb{E}\left[u'\left(w_0-X+(X-d)_+-c_d\right)\right]},\end{aligned}$$

where the last step is by the definition of K_d. If $z_d^* = 0$, we have

$$\begin{aligned}\frac{\int_0^{1-d}\left(g_d'(y)-g_d'(0)\right)f_X(y+d)\mathrm{d}y}{\overline{F}_X(d)}+g_d'(0)&=1+K_d\\&=\frac{u'\left(w_0-d-c_d\right)}{\mathbb{E}\left[u'\left(w_0-X+(X-d)_+-c_d\right)\right]}.\end{aligned}$$

Further, it is seen that g_d satisfies the policyholder's participation constraint, i.e., $g_d \in \mathcal{A}_2$ if $z_d^* > 0$. When $z_d^* = 0$, the policyholder's participation constraint can be implied by the followed Proposition 2.4.1. This confirms that $g_d \in \mathcal{A}_2$.

Lastly, we show that the insurer's expected profit is positive under g_d. By (2.26),

$$\begin{aligned}U^{In}(d,g_d)&=\mathbb{E}[g_d((X-d)_+)]-\mathbb{E}[(X-d)_+]\\&=\int_0^{1-d}\left(g_d'(y)-g_d'(0)\right)\overline{F}_X(y+d)\mathrm{d}y+\left(g_d'(0)-1\right)\int_0^{1-d}\overline{F}_X(y+d)\mathrm{d}y\\&=K_d\int_{z_d^*}^{1-d}\overline{F}_X(y+d)\mathrm{d}y>0,\end{aligned}$$

where the last step is by $K_d > 0$, which can be deduced from

$$u'\left(w_0-d-c_d\right)>\mathbb{E}\left[u'\left(w_0-X+(X-d)_+-c_d\right)\right],$$

as u is strictly concave, $d \in (0, 1)$, and the density of X is positive. □

The most important implication of Theorem 2.4.2 is that a *stop-loss premium principle* g_d is optimal for the insurer, when the class of acceptable indemnity functions is the collection of deductible indemnities with deductible level d. It is seen from (2.25) that the loading is 0 for losses smaller than z_d^*, and it increases to K_d for larger losses, where the *change point* z_d^* is the largest *mean excess of loss* level, by noting that $X - d$ is the loss covered by the insurer. This is contrary to proportional insurance indemnities, where the expected-value premium principle is optimal (see Theorem 2.3.2). Additionally, Theorem 2.4.2 also shows that the insurer is better off offering the mechanism $(d,\, g_d)$.

Remark 1. *The function g_d under consideration is piecewise linear and hence differentiable almost everywhere, except possibly at the turning point z_d^*. In our analysis, the derivative is interpreted as the right-hand derivative, unless otherwise specified. Moreover, all integrals and comparisons involving the derivative are taken to hold almost everywhere, excluding the point of non-differentiability, which form a null set.*

Remark 2. *We provide some intuition for the proof of Theorem 2.4.2. By (2.26) and fixing $g'(0)$, we know that maximizing the insurer's premium $\mathbb{E}[g((X-d)_+)]$ is equivalent to maximize $\int_0^{1-d} h(z)\overline{F}_X(d+z)\mathrm{d}z$. Further, by (2.27) and assuming that the insurer's premium $\mathbb{E}[g((X-d)_+)]$ attains some maximum value, one essentially needs to consider the following problem*

$$\begin{cases} \max\limits_{h\in\mathcal{A}_h} \displaystyle\int_0^{1-d} h(z)\overline{F}_X(z+d)\mathrm{d}z, \\[2ex] s.t. \displaystyle\int_0^{1-d} h(z)f_X(z+d)\mathrm{d}z = c, \end{cases}$$

where c is a constant. Informally, we assume the differentiability of h. Then the objective function can be rewritten as

$$\begin{aligned} \max_h \int_0^{1-d} h(z)\,\overline{F}_X(z+d)\,\mathrm{d}z &= \max_h \int_0^{1-d} h'(z)\left(\int_{z+d}^1 \overline{F}_X(y)\mathrm{d}y\right)\mathrm{d}z \\ &= \max_h \int_0^{1-d} h'(z)\overline{F}_X(z+d)M_X(z+d)\mathrm{d}z, \end{aligned}$$

and the constraint can be rewritten as

$$c = \int_0^{1-d} h(z)\,f_X(z+d)\,\mathrm{d}z = -\int_0^{1-d} h(z)\,\mathrm{d}\overline{F}_X(z+d) = \int_0^{1-d} h'(z)\,\overline{F}_X(z+d)\,\mathrm{d}z.$$

As such, it is optimal to choose $h'(z)\,\overline{F}_X(z+d)$ (and then $h'(z)$) as a Dirac function with mass point at the maximizer of $M_X(z+d)$, that is denoted by z_d^*.

The following numerical example is to demonstrate the optimality of the stop-loss premium principle g_d.

Example 2.4.1. *Let X be a truncated Pareto distribution supported on $[0,1]$ with survival function*

$$\overline{F}_X(x) = \frac{(1+x)^{-\beta} - 2^{-\beta}}{1 - 2^{-\beta}}, \quad x \in [0,1],$$

where $\beta > 2$. Assume the exponential utility $u(x) = \frac{1-e^{-\gamma x}}{\gamma}$ with risk aversion parameter $\gamma > 0$. For each fixed $d \in (0,1)$, the following figure plots the insurer's expected profit $U^{In}(d, g_d)$ under the stop-loss premium principle g_d. We compare it with the expected-value premium principle

$$g_\theta(x) = (1+\theta)x,$$

where $\theta > 0$ is such that $g_\theta \in \mathcal{A}_2(d)$, the same admissible set as g_d. The corresponding insurer's expected profit is denoted by $U^{In}(d, g_\theta)$. From Figure 2.2, it is clear that insurer's expected profit is improved by adopting the stop-loss premium principle. We set $\beta = 11$ and $\gamma = 6$.

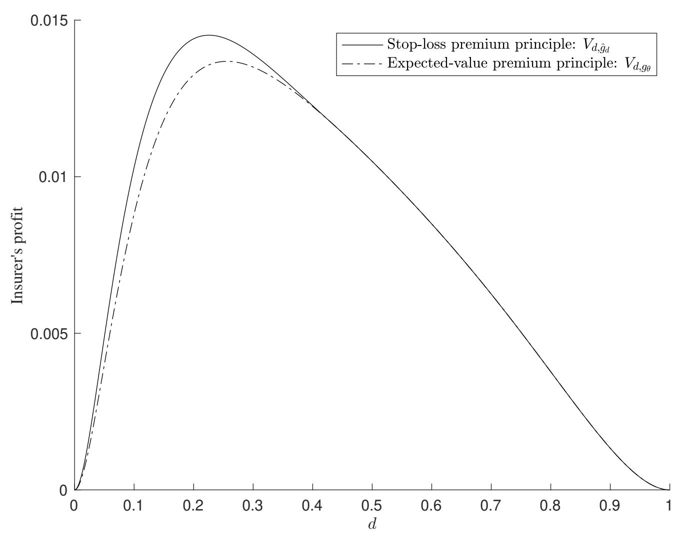


Figure 2.2: Insurer's expected profit under the stop-loss premium principle vs the expected-value premium principle

Recall from (2.21) that D_g denotes the set of optimal deductible levels for the policyholder under a premium principle g. The following proposition shows that, under the stop-loss premium principle g_d obtained in Theorem 2.4.2, the unique optimal deductible level for the policyholder is indeed d, provided that the mean excess of loss function $M_X(x)$

is *decreasing*, that is, when X has decreasing mean residual life (DMRL), which is often viewed as suggesting a *light-tailed distribution*. Examples of distributions with DMRL include the uniform distribution, the truncated exponential distribution, the truncated normal, etc. For DMRL losses, the stop-loss premium principle reduces to the expected-value premium principle. Moreover, the following proposition shows that, under the premium principle g_d, d is the *unique* optimal deductible level for the policyholder.

Proposition 2.4.1. *For a fixed $d \in (0,1)$, suppose that*

$$z_d^* := \text{argmax}_{z \in [0,1-d)} M_X(d+z) = 0,$$

and note that this holds for all losses with DMRL. Then we have

$$g_d(x) = \frac{u'(w_0 - d - c_d)}{\mathbb{E}\left[u'(w_0 - X + (X-d)_+ - c_d)\right]} x, \quad \forall x \in (0, 1-d],$$

with

$$c_d = \sup\left\{c \geq 0 : \frac{u'(w_0 - d - c)\overline{F}_X(d) M_X(d)}{\mathbb{E}\left[u'(w_0 - X + (X-d)_+ - c)\right]} - c \geq 0\right\}.$$

Further,

$$D_{g_d} = \{d\}.$$

Proof. Since $z_d^* = 0$, by letting $a_d := \frac{u'(w_0-d-c_d)}{\mathbb{E}[u'(w_0-X+(X-d)_+-c_d)]}$, we have

$$\begin{aligned}U^{Pol}(\widetilde{d}, g_d) &= \mathbb{E}\left[u\left(w_0 - X + (X-\widetilde{d})_+ - a_d \mathbb{E}[(X-\widetilde{d})_+]\right)\right] \\ &= \int_0^{\widetilde{d}} u\left(w_0 - x - a_d\mathbb{E}[(X-\widetilde{d})_+]\right) f_X(x)dx + u\left(w_0 - \widetilde{d} - a_d\mathbb{E}[(X-\widetilde{d})_+]\right)\overline{F}(\widetilde{d}).\end{aligned}$$

By $\frac{\partial \mathbb{E}[(X-\widetilde{d})_+]}{\partial \widetilde{d}} = \frac{\partial \int_{\widetilde{d}}^1 \overline{F}(x)dx}{\partial \widetilde{d}} = -\overline{F}(\widetilde{d})$, it follows that

$$\begin{aligned}\frac{\partial U^{Pol}(\widetilde{d}, g_d)}{\partial \widetilde{d}} &= u\left(w_0 - \widetilde{d} - a_d\mathbb{E}[(X-\widetilde{d})_+]\right) f_X(\widetilde{d}) \\ &\quad + \int_0^{\widetilde{d}} u'\left(w_0 - x - a_d\mathbb{E}[(X-\widetilde{d})_+]\right) a_d\overline{F}(\widetilde{d}) f_X(x)\mathrm{d}x \\ &\quad - u\left(w_0 - \widetilde{d} - a_d\mathbb{E}[(X-\widetilde{d})_+]\right) f_X(\widetilde{d}) \\ &\quad + u'\left(w_0 - \widetilde{d} - a_d\mathbb{E}[(X-\widetilde{d})_+]\right)\left(-1 + a_d\overline{F}(\widetilde{d})\right)\overline{F}(\widetilde{d})\end{aligned}$$

$$= \overline{F}(\widetilde{d}) l_d(\widetilde{d}), \tag{2.32}$$

where

$$\begin{aligned} l_d(\widetilde{d}) :=& \int_0^{\widetilde{d}} u'\left(w_0 - x - a_d\mathbb{E}[(X-\widetilde{d})_+]\right) a_d f_X(x)\mathrm{d}x \\ &+ u'\left(w_0 - \widetilde{d} - a_d\mathbb{E}[(X-\widetilde{d})_+]\right)\left(-1 + a_d\overline{F}(\widetilde{d})\right) \\ =& -u'\left(w_0 - \widetilde{d} - a_d\mathbb{E}[(X-\widetilde{d})_+]\right) \\ &+ a_d\mathbb{E}\left[u'\left(w_0 - X + (X-\widetilde{d})_+ - a_d\mathbb{E}[(X-\widetilde{d})_+]\right)\right]. \end{aligned}$$

By $c_d = a_d\mathbb{E}[(X-\widetilde{d})_+]$, it is clear that

$$l_d(d) = -u'\left(w_0 - \widetilde{d} - a_d\mathbb{E}[(X-\widetilde{d})_+]\right) + a_d\mathbb{E}\left[u'\left(w_0 - X + (X-\widetilde{d})_+ - c_d\right)\right] = 0.$$

Further,

$$\begin{aligned} \frac{\partial l_d(\widetilde{d})}{\partial \widetilde{d}} =& u'\left(w_0 - \widetilde{d} - a_d\mathbb{E}[(X-\widetilde{d})_+]\right) a_d f_X(\widetilde{d}) \\ &+ \int_0^{\widetilde{d}} u''\left(w_0 - x - a_d\mathbb{E}[(X-\widetilde{d})_+]\right) a_d^2 f_X(x)\overline{F}(\widetilde{d})\mathrm{d}x \\ &+ u''\left(w_0 - \widetilde{d} - a_d\mathbb{E}[(X-\widetilde{d})_+]\right)\left(-1 + a_d\overline{F}(\widetilde{d})\right)^2 \\ &- u'\left(w_0 - \widetilde{d} - a_d\mathbb{E}[(X-\widetilde{d})_+]\right) a_d f_X(\widetilde{d}) \\ =& \int_0^{\widetilde{d}} u''\left(w_0 - x - a_d\mathbb{E}[(X-\widetilde{d})_+]\right) a_d^2 f_X(x)\overline{F}(\widetilde{d})\mathrm{d}x \\ &+ u''\left(w_0 - \widetilde{d} - a_d\mathbb{E}[(X-\widetilde{d})_+]\right)\left(-1 + a_d\overline{F}(\widetilde{d})\right)^2 \\ <& \ 0. \end{aligned}$$

Thus, $l_d(\widetilde{d})$ is decreasing in $\widetilde{d}$ and has a unique zero at $\widetilde{d} = d$. Consequently, by (2.32), we conclude that $U^{Pol}(\widetilde{d}, g_d)$ is increasing for $\widetilde{d} < d$, decreasing for $\widetilde{d} > d$, and reaches a global maximum at $\widetilde{d} = d$. □

Similarly to Theorem 2.3.3, we have the following results regarding the relationship between the original problem $\max_{g\in\mathcal{A}_2}\max_{d\in D_g} U^{In}(d,g)$ and the dual problem $\max_{d\in(0,1)}\max_{g\in\mathcal{A}_2(d)} U^{In}(d,g)$.

Theorem 2.4.3. *For deductible contracts, we have*

$$\max_{g\in\mathcal{A}_2}\max_{d\in D_g} U^{In}(d,g) \le \max_{d\in(0,1)}\max_{g\in\mathcal{A}_2(d)} U^{In}(d,g) = \max_{d\in(0,1)} U^{In}(d,g_d). \tag{2.33}$$

Moreover, the equality holds in (2.33) if there exists $d^*\in \operatorname{argmax}_{d\in(0,1)} U^{In}(d,g_d)$ such that

$$d^* \in D_{g_{d^*}}. \tag{2.34}$$

In particular, (2.34) holds for all losses with DMRL.

Proof. By Theorem 2.4.2, it is known that for any $d\in(0,1)$,

$$\max_{g\in\mathcal{A}_2(d)} U^{In}(d,g) = U^{In}(d,g_d),$$

which implies the second equality in (2.33). Next we prove the inequality in (2.33). Denote by (d_{g^*}, g^*) the optimal solution to $\max\limits_{g\in\mathcal{A}_2}\max\limits_{d\in D_g} U^{In}(d,g)$, where $d_{g^*}\in D_{g^*}$. From Theorem 2.4.1, we know that $d_{g^*}\in(0,1)$, and thus the mechanism (d_{g^*}, g^*) must satisfy the first-order condition (2.22), implying that $g^*\in\mathcal{A}_2(d_{g^*})$. Thus,

$$\max_{g\in\mathcal{A}_2}\max_{d\in D_g} U^{In}(d,g) = U^{In}\left(d_{g^*}, g^*\right) \le \max_{g\in\mathcal{A}_2(d_{g^*})} U^{In}\left(d_{g^*}, g\right) \le \max_{d\in(0,1)}\max_{g\in\mathcal{A}_2(d)} U^{In}\left(d,g\right).$$

It remains to show that

$$\max_{d\in(0,1)} U^{In}(d,g_d) \le \max_{g\in\mathcal{A}_2}\max_{d\in D_g} U^{In}(d,g)$$

under condition (2.34). Suppose that $\max\limits_{d\in(0,1)} U^{In}(d,g_d) = U^{In}(d^*, g_{d^*})$ for some $d^*\in(0,1)$. Then

$$\max_{d\in(0,1)} U^{In}(d,g_d) = U^{In}(d^*,g_{d^*}) = \max_{d\in D_{g_{d^*}}} U^{In}(d,g_{d^*}) \le \max_{g\in\mathcal{A}_2,\, d\in D_g} U^{In}(d,g),$$

which concludes the proof. Lastly, note that condition (2.34) holds for all losses have DMRL by Proposition 2.4.1. □

We have the following results regarding the Bowley optimality and individual rationality of our derived deductible mechanism (d^*, g_{d^*}). We omit the proof, as it follows similar steps to those in Theorem 2.3.4.

Theorem 2.4.4. *Consider the deductible mechanism* (d^*, g_{d^*})*, where* $d^* \in \underset{d\in(0,1)}{\arg\max}\, U^{In}(d, g_d)$*. Suppose that*

$$d^* \in D_{g_{d^*}},$$

and note that this holds for all losses with DMRL. Then $(d^*, g_{d^*}) \in \mathcal{BO} \cap \mathcal{IR}$*.*

It is worth noting that Theorem 2.4.4 assures that if the loss distribution has DMRL, this is a *sufficient* condition to guarantee the Bowley optimality of the derived mechanism (d^*, g_{d^*}), where d^* is the deductible level that maximizes the insurer's expected profit. However, this does not imply that the Bowley optimality cannot be achieved with non-DMRL distributions.

In the following example, we revisit the truncated Pareto distribution, a non-DMRL distribution, to demonstrate this point. We show that our derived mechanism (d^*, g_{d^*}), which maximizes the insurer's expected profit, can still achieve Bowley optimality for the policyholder by verifying that

$$d^* \in D_{g_{d^*}} = \text{argmax}_{d\in[0,1]}\, U^{Pol}(d, g_{d^*}).$$

Therefore, DMRL is *sufficient but not necessary* for the Bowley optimality of (d^*, g_{d^*}).

Example 2.4.2. *Suppose that* X *follows a truncated Pareto distribution supported on the interval* $[0, 1]$*, with survival function*

$$\overline{F}_X(x) = \frac{(1+x)^{-\beta} - 2^{-\beta}}{1 - 2^{-\beta}}, \quad x \in [0, 1],$$

where $\beta > 2$*. One can verify that* X *has non-DMRL. Suppose that* $u(x) = \frac{1-e^{-\gamma x}}{\gamma}$*, where we set* $\gamma = 1$*.*

In Figure 2.3, we examine three values of β*:* $\beta = 15$*,* $\beta = 40$*, and* $\beta = 65$*. The left three panels plot the function* $d \mapsto U^{In}(d, g_d)$*, showing that the deductible level that maximizes the insurer's expected profit that is*

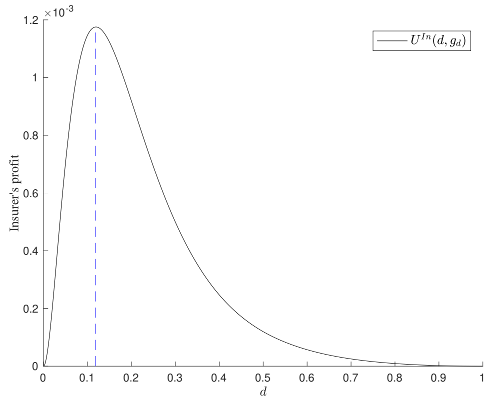
$$d^* = \text{argmax}_{d\in[0,1]}\, U^{In}(d, g_d)$$

is achieved at 0.119*,* 0.042*, and* 0.025 *respectively, for each value of* β*. The right three panels justify that, given the premium principle* g_{d^*}*, the optimal deductible level for the policyholder coincides with* d^**, i.e.,*

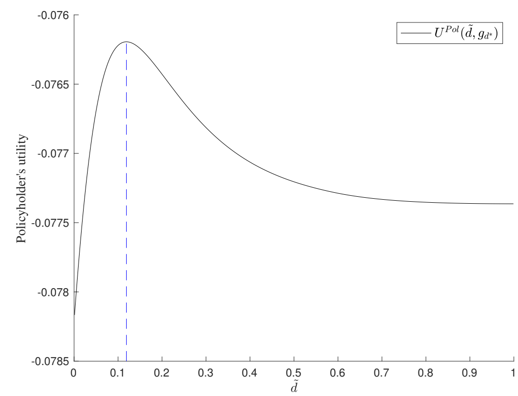
$$D_{g_{d^*}} = \{d^*\}.$$

This implies the Bowley optimality of (d^*, g_{d^*}) *by Theorem 2.4.4.*

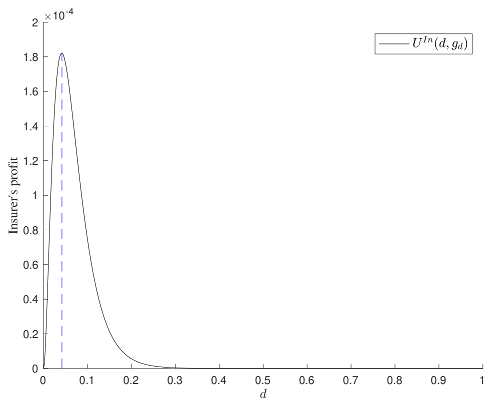
In fact, we have verified the Bowley optimality for $\beta = 3, 4, \ldots, 100$*. We only present three representative values of* β *to present in Figure 2.3.*



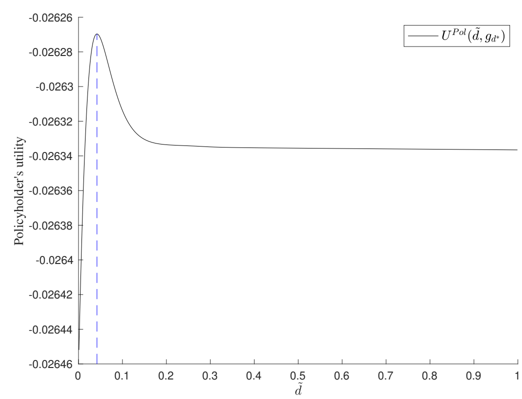
(a) $d^* = 0.119$ when $\beta = 15$.



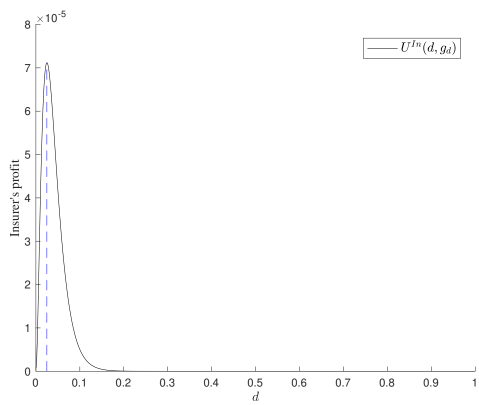
(b) $D_{g_{d^*}} = \{0.119\}$ when $\beta = 15$.



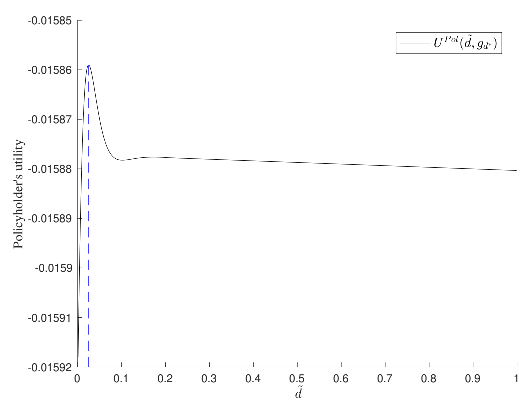
(c) $d^* = 0.042$ when $\beta = 40$.



(d) $D_{g_{d^*}} = \{0.042\}$ when $\beta = 40$.



(e) $d^* = 0.025$ when $\beta = 65$.



(f) $D_{g_{d^*}} = \{0.025\}$ when $\beta = 65$.

Figure 2.3: Truncated Pareto distribution

Theorem 2.4.4 and Example 2.4.2 imply that a broader sufficient condition—beyond DMRL losses—to ensure the Bowley optimality and individual rationality of the mechanism (d^*, g_{d^*}) is that

$$d^* \in D_{g_{d^*}}. \tag{2.35}$$

Even if this condition is not satisfied, Theorem 2.4.3 implies that the mechanism (d^*, g_{d^*}) still yields the highest profit for the insurer, albeit at the expense of reducing the policyholder's welfare. As the leader in the Stackelberg game, the insurer retains the option to implement this mechanism, provided the policyholder's individual rationality condition is satisfied. Alternatively, when condition (2.35) is not satisfied, the insurer may also consider adopting the *proportional mechanism* (α^*, g_{α^*}) derived in Section 2.3. This is intuitive, as condition 2.35 fails only for non-DMRL losses, which have *heavier tails*. The proportional mechanism thus offers a *more balanced* way for both parties to share the risk.

The following example offers an alternative comparison between the two forms of insurance contracts. It computes the insurer's profit under both proportional and deductible contracts, using the Bowley-optimal loading function in each case. The results suggest that the insurer's optimal choice between a proportional and a deductible contract depends on the policyholder's risk aversion and the distribution of the insured loss.

Example 2.4.3. *Suppose the policyholder is an expected utility maximizer with exponential utility given by*

$$u(x) = \frac{1 - e^{-\gamma x}}{\gamma},$$

where $\gamma > 0$ is the risk aversion parameter. We compare the insurer's profit under proportional and deductible contracts for different loss distributions.

In the left panel of Figure 2.4, the loss variable X follows a truncated exponential distribution with survival function

$$\overline{F}_X(x) = \frac{e^{-\lambda x} - e^{-\lambda}}{1 - e^{-\lambda}}, \quad x \in [0, 1],$$

where $\lambda > 0$. In the right panel, X follows a truncated Pareto distribution supported on $[0, 1]$, with survival function

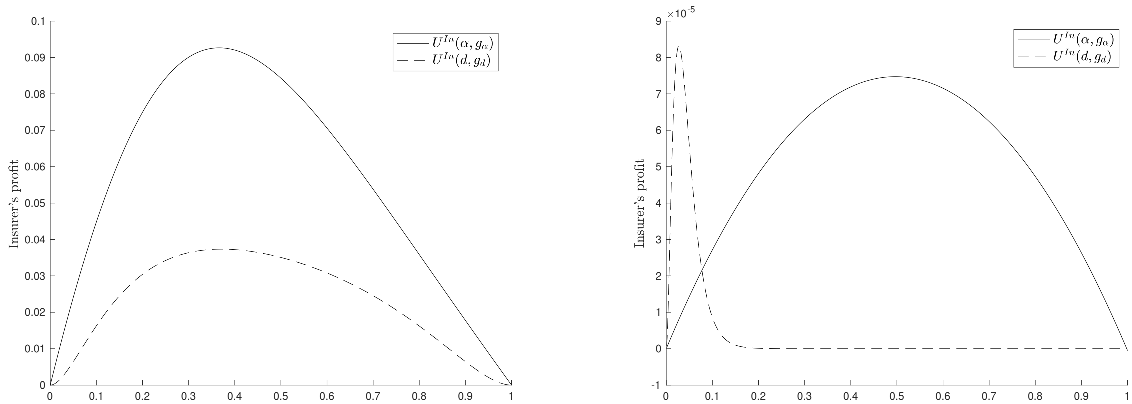
$$\overline{F}_X(x) = \frac{(1 + x)^{-\beta} - 2^{-\beta}}{1 - 2^{-\beta}}, \quad x \in [0, 1],$$

where $\beta > 2$.

The results show that when X follows the truncated exponential distribution and the risk aversion parameter is $\gamma = 9$, the insurer's profit is maximized by a proportional contract, which outperforms the deductible contract. In this case, the insurer should adopt the proportional Bowley-optimal mechanism (α^*, g_{α^*}).

On the other hand, when $\gamma = 1$ and the loss follows a truncated Pareto distribution with $\beta = 60$, the deductible contract (d^*, g_{d^*}) yields a higher profit than the proportional contract at the optimal point.

In summary, whether a proportional or deductible contract is optimal for the insurer depends critically on the policyholder's risk aversion γ and the distributional properties of the loss X.



(a) $\gamma = 9$ and $\lambda = 7$

(b) $\gamma = 1$ and $\beta = 60$

Figure 2.4: Comparison of the insurer's profit under different contract forms

2.5 Conclusion

This chapter contributes to the literature on Stackelberg equilibria (Bowley optima) in monopolistic centralized sequential-move insurance markets in several ways. We restrict the set of market pricing mechanisms to those premium principles that are expectations of some increasing and convex function of the indemnity function. Such premium principles are economically appealing, since the marginal premium increases with ceded loss function. We assume that the policyholder is a risk-averse EU-maximizer, while the insurer is a risk-neutral expected-profit maximizer.

We then study Bowley optimality of such premium principles, when the indemnity functions are chosen from the two most popular and practically relevant classes of indemnity schedules: the deductible indemnities and the proportional indemnities, both of which satisfy the no-sabotage condition. To the best of our knowledge, this is the first work that studies the optimal structure of premium principles for given classes of indemnity functions. We find that the expected-value premium principle is Bowley optimal for proportional indemnities, while the stop-loss premium principle is Bowley optimal for deductible indemnities under some mild conditions. Methodologically, we introduce a novel dual approach to study the Bowley optimality problem. In the appendix, we include some initial attempts to explore Bowley's optimal principle for general indemnities and a special case involving binary losses.

Appendix A: An Extension to General 1-Lipschitz Indemnities

A.1 General Losses

This Appendix discusses the Bowley-optimal convex-loaded premium principles with general 1-Lipschitz indemnity functions. In this context, the policyholder's problem with general indemnities (all $I \in \mathcal{I}_L$) is

$$\max_{I \in \mathcal{I}_L} \mathbb{E}\left[u\left(w_0 - X + I(X) - \mathbb{E}[g(I(X))]\right)\right]. \tag{A.1}$$

The following theorem provides a characterization of the solution to problem (A.1).

Theorem A.1. *For a given convex-loaded premium principle $g \in \mathcal{A}_0$, an indemnity function I_g is optimal to (A.1) if and only if $I_g \in \mathcal{I}$ and I_g satisfies*

$$(I_g)'(x) = \begin{cases} 0, & L(x) < 0, \\ \phi_g(x) \in [0,1], & L(x) = 0, \\ 1, & L(x) > 0, \end{cases} \tag{A.2}$$

where

$$L(x) = \frac{\int_x^1 u'\left(w_0 - y + I_g(y) - \mathbb{E}[g(I_g(X))]\right) \mathrm{d}F_X(y)}{\mathbb{E}\left[u'\left(w_0 - X + I_g(X) - \mathbb{E}[g(I_g(X))]\right)\right]} - \int_x^1 g'\left(I_g(y)\right) \mathrm{d}F_X(y).$$

Proof. Suppose that $\hat{I}$ solves problem (A.1). Let $I \in \mathcal{I}_L$ be arbitrary and fixed. For any $\varepsilon \in (0,1)$, we know that $\widetilde{I} = (1-\varepsilon)\hat{I} + \varepsilon I \in \mathcal{I}_L$. By the optimality of $\hat{I}$, we have that

$$\begin{aligned}
0 &\geq \mathbb{E}[u(w_0 - X + \widetilde{I}(X) - \mathbb{E}[g(\widetilde{I}(X))])] - \mathbb{E}[u(w_0 - X + \hat{I}(X) - \mathbb{E}[g(\hat{I}(X))])] \\
&= \int_0^1 \Big[u(w_0 - x + \widetilde{I}(x) - \mathbb{E}[g(\widetilde{I}(X))]) - u(w_0 - x + \hat{I}(x) - \mathbb{E}[g(\hat{I}(X))]) \Big] \mathrm{d}F_X(x) \\
&\geq \int_0^1 u'(w_0 - x + \widetilde{I}(x) - \mathbb{E}[g(\widetilde{I}(X))]) \Big(\widetilde{I}(x) - \hat{I}(x) - \mathbb{E}[g(\widetilde{I}(X))] + \mathbb{E}[g(\hat{I}(X))] \Big) \mathrm{d}F_X(x) \\
&= \int_0^1 u'(w_0 - x + \widetilde{I}(x) - \mathbb{E}[g(\widetilde{I}(X))]) \Big(\varepsilon (I(x) - \hat{I}(x)) - \mathbb{E}[g(\widetilde{I}(X))] + \mathbb{E}[g(\hat{I}(X))] \Big) \mathrm{d}F_X(x),
\end{aligned}$$

where the third step is by the concavity of u. It follows that

$$\begin{aligned}
&\varepsilon \int_0^1 (I'(x) - \hat{I}'(x)) \left(\int_x^1 u'(w_0 - y + \widetilde{I}(y) - \mathbb{E}[g(\widetilde{I}(X))])) \mathrm{d}F_X(y) \right) \mathrm{d}x \\
&= \varepsilon \int_0^1 u'(w_0 - x + \widetilde{I}(x) - \mathbb{E}[g(\widetilde{I}(X))])(I(x) - \hat{I}(x)) \mathrm{d}F_X(x) \\
&\leq \Big(\mathbb{E}[g(\widetilde{I}(X))] - \mathbb{E}[g(\hat{I}(X))] \Big) \int_0^1 u'(w_0 - x + \widetilde{I}(x) - \mathbb{E}[g(\widetilde{I}(X))]) \mathrm{d}F_X(x) \\
&\leq \int_0^1 g'(\widetilde{I}(x))(\widetilde{I}(x) - \hat{I}(x)) \mathrm{d}F_X(x) \mathbb{E}\left[u'\left(w_0 - X + \widetilde{I}(X) - \mathbb{E}[g(\widetilde{I}(X))] \right) \right] \\
&= \varepsilon \int_0^1 g'(\widetilde{I}(x))(I(x) - \hat{I}(x)) \mathrm{d}F_X(x) \mathbb{E}\left[u'\left(w_0 - X + \widetilde{I}(X) - \mathbb{E}[g(\widetilde{I}(X))] \right) \right] \\
&= \varepsilon \int_0^1 (I'(x) - \hat{I}'(x)) \left(\int_x^1 g'(\widetilde{I}(y)) \mathrm{d}F_X(y) \right) \mathrm{d}x \mathbb{E}\left[u'\left(w_0 - X + \widetilde{I}(X) - \mathbb{E}[g(\widetilde{I}(X))] \right) \right],
\end{aligned}$$

where the third step is by the convexity of g. Dividing both sides by ε and letting $\varepsilon \to 0$ yields

$$\begin{aligned}
&\int_0^1 (I'(x) - \hat{I}'(x)) \left(\int_x^1 u'(w_0 - y + \hat{I}(y) - \mathbb{E}[g(\hat{I}(X))]) \mathrm{d}F_X(y) \right) \mathrm{d}x \\
&\leq \int_0^1 (I'(x) - \hat{I}'(x)) \left(\int_x^1 g'(\hat{I}(y)) \mathrm{d}F_X(y) \right) \mathrm{d}x \mathbb{E}\left[u'\left(w_0 - X + \hat{I}(X) - \mathbb{E}[g(\hat{I}(X))] \right) \right],
\end{aligned}$$

or equivalently,

$$\int_0^1 \Big(I'(x) - \hat{I}'(x) \Big) \Phi(x) \mathrm{d}x \leq 0,$$

where

$$\Phi(x) := \int_x^1 u'(w_0 - y + \hat{I}(y) - \mathbb{E}[g(\hat{I}(X))])\mathrm{d}F_X(y)$$
$$- \int_x^1 g'(\hat{I}(y))\mathrm{d}F_X(y)\mathbb{E}\left[u'\left(w_0 - X + \hat{I}(X) - \mathbb{E}[g(\hat{I}(X))]\right)\right],$$

This implies that $\hat{I}$ maximizes

$$\max_{I \in \mathcal{I}_L} \int_0^1 I'(x)\Phi(x)\mathrm{d}x.$$

Therefore, we deduce that $I_g \in \mathcal{I}_g$ only if

$$\left(I_g\right)'(x) = \begin{cases} 0, & L(x) < 0, \\ \phi_g(x) \in [0,1], & L(x) = 0, \\ 1, & L(x) > 0, \end{cases} \tag{A.3}$$

where

$$L(x) := \frac{\int_x^1 u'\left(w_0 - y + I_g(y) - \mathbb{E}[g(I_g(X))]\right)\mathrm{d}F_X(y)}{\mathbb{E}\left[u'\left(w_0 - X + I_g(X) - \mathbb{E}[g(I_g(X))]\right)\right]} - \int_x^1 g'(I_g(y))\mathrm{d}F_X(y)$$

On the other hand, suppose that I_g satisfies (A.3), for any $I \in \mathcal{I}_L$, we have

$$\mathbb{E}\left[u\left(w_0 - X + I_g(X) - \mathbb{E}[g(I_g(X))]\right)\right] - \mathbb{E}\left[u\left(w_0 - X + I(X) - \mathbb{E}[g(I(X))]\right)\right]$$
$$= \int_0^1 u\left(w_0 - x + I_g(x) - \mathbb{E}\left[g\left(I_g(X)\right)\right]\right)\mathrm{d}F_X(x) - \int_0^1 u\left(w_0 - x + I(x) - \mathbb{E}\left[g\left(I(X)\right)\right]\right)\mathrm{d}F_X(x)$$
$$\geq \int_0^1 u'\left(w_0 - x + I_g(x) - \mathbb{E}\left[g\left(I_g(X)\right)\right]\right)\left(I_g(x) - I(x) - \mathbb{E}\left[g\left(I_g(X)\right)\right] + \mathbb{E}\left[g\left(I(X)\right)\right]\right)\mathrm{d}F_X(x)$$
$$= \int_0^1 \left(I_g'(x) - I'(x)\right)\int_x^1 u'\left(w_0 - y + I_g(y) - \mathbb{E}\left[g\left(I_g(X)\right)\right]\right)\mathrm{d}F_X(y)\mathrm{d}x$$
$$+ \int_0^1 \left(g(I(x)) - g(I_g(x))\right)\mathrm{d}F_X(x)\int_0^1 u'\left(w_0 - x + I_g(x) - \mathbb{E}\left[g\left(I_g(X)\right)\right]\right)\mathrm{d}F_X(x)$$
$$\geq \int_0^1 \left(I_g'(x) - I'(x)\right)\int_x^1 u'\left(w_0 - y + I_g(y) - \mathbb{E}\left[g\left(I_g(X)\right)\right]\right)\mathrm{d}F_X(y)\mathrm{d}x$$
$$+ \int_0^1 g'(I_g(x))\left(I(x) - I_g(x)\right)\mathrm{d}F_X(x)\int_0^1 u'\left(w_0 - x + I_g(x) - \mathbb{E}\left[g\left(I_g(X)\right)\right]\right)\mathrm{d}F_X(x)$$

$$
\begin{aligned}
&= \int_0^1 \left(I_g'(x) - I'(x)\right) \int_x^1 u'\left(w_0 - y + I_g(y) - \mathbb{E}\left[g\left(I_g(X)\right)\right]\right) \mathrm{d}F_X(y)\mathrm{d}x \\
&+ \int_0^1 \left(I'(x) - I_g'(x)\right) \int_x^1 g'(I_g(y))\mathrm{d}F_X(y)\mathrm{d}x \int_0^1 u'\left(w_0 - x + I_g(x) - \mathbb{E}\left[g\left(I_g(X)\right)\right]\right) \mathrm{d}F_X(x) \\
&\geq 0,
\end{aligned}
$$

where we use the concavity of u in the second step, the convexity of g in the fourth step, and that I_g satisfies (A.3) in the last step. This completes the proof. □

Remark 3. *In recent work by* CAO *et al. (2024), our convex-loaded premium principle is adopted, and the study focuses on the optimal indemnity function under the exponential utility* $u(x) = -e^{-\gamma x}$*. Note that the authors assume that the indemnity functions satisfy* $0 \leq I(x) \leq x$*, without imposing the 1-Lipschitz condition. They find that the optimal indemnity function takes the form of a deductible contract, i.e.,*

$$
I_g(x) = \begin{cases} 0, & \text{if} \quad x < d_g, \\ \in (0, x), & \text{if} \quad x \geq d_g, \end{cases}
$$

for some deductible level $d_g > 0$*, and the 1-Lipschitz continuity of this solution is confirmed. It is important to note that, as* $L(0) = 1 - \mathbb{E}\left[g'\left(I_g(X)\right)\right] < 0$*, our findings indicate that the optimal indemnity function under a general utility, given in(3.11), still includes a deductible level.*

The subsequent insurer's problem is to find an optimal loading function g that maximizes the expected profit, i.e.,

$$
\max_{g \in \mathcal{A}_0} \mathbb{E}[g(I_g(X)) - I_g(X)]. \tag{A.4}
$$

Substituting (3.11) into (A.4) gives

$$
\begin{aligned}
\mathbb{E}\left[g\left(I_g(X)\right)\right] - \mathbb{E}\left[I_g(X)\right] &= \int_0^1 \left(g'\left(I_g(x)\right) - 1\right) I_g'(x)\overline{F}_X(x)\mathrm{d}x \\
&= \int_{\mathcal{C}_g} \left(g'\left(I_g(x)\right) - 1\right) \overline{F}_X(x)\mathrm{d}x + \int_{\mathcal{D}_g} \left(g'\left(I_g(x)\right) - 1\right) \phi_g(x)\overline{F}_X(x)\mathrm{d}x,
\end{aligned}
$$

where $\mathcal{C}_g = \{x : L(x) > 0\}$ and $\mathcal{D}_g = \{x : L(x) = 0\}$. It is evident that the loading function g is rather implicitly embedded in both the objective function and the integration regions. This poses fundamental challenges for further analytical studies. In the next subsection, we consider a special case with binary losses which can be solved explicitly.

A.2 The Special Case of Binary Losses

In this subsection, we consider a special case where the loss X is binary, taking the value $l \in (0,1]$ with probability p, and the value 0 with probability $1-p$. Then

$$F_X(x) = \begin{cases} 1-p, & 0 \le x < l, \\ 1, & l \le x \le 1. \end{cases} \tag{A.5}$$

Theorem A.2. *For a given premium principle $g \in \mathcal{A}_0$, if X follows the binary distribution (A.5), then any optimal indemnity function I_g for Problem (A.1) is such that $I_g(l)$ is a solution to the following equation in y:*

$$g'(y) = \frac{u'\left(w_0 - l + y - pg(y)\right)}{pu'\left(w_0 - l + y - pg(y)\right) + (1-p)u'\left(w_0 - pg(y)\right)}. \tag{A.6}$$

Furthermore, a necessary and sufficient condition for policyholders to participate, i.e., $I_g(l) \neq 0$, is

$$g'(0) < \frac{u'\left(w_0 - l\right)}{pu'\left(w_0 - l\right) + (1-p)u'\left(w_0\right)}.$$

Proof. Given a given premium principle $g \in \mathcal{A}_0$, the policyholder's objective function in (A.1) is to maximize

$$U^{Pol}(y,g) = pu\left(w_0 - l + y - pg(y)\right) + (1-p)u\left(w_0 - pg(y)\right),$$

where $y = I(l)$. It follows that

$$\frac{\partial U^{Pol}(y,g)}{\partial y} = pu'\left(w_0 - l + y - pg(y)\right)\left(1 - pg'(y)\right) - p(1-p)u'\left(w_0 - pg(y)\right)g'(y).$$

Note that full insurance is never optimal since

$$\begin{aligned} \left.\frac{\partial U^{Pol}(y,g)}{\partial y}\right|_{y=l} &= pu'\left(w_0 - pg(l)\right)\left(1 - pg'(l)\right) - p(1-p)u'\left(w_0 - pg(l)\right)g'(l) \\ &= pu'\left(w_0 - pg(l)\right)\left(1 - g'(l)\right) < 0. \end{aligned}$$

Moreover, since u is increasing and concave, and $g \in \mathcal{A}_0$, it follows that

$$\frac{\partial^2 U^{Pol}(y,g)}{\partial y^2} = pu''\left(w_0 - l + y - pg(y)\right)\left(1 - pg'(y)\right)^2 - p^2u'\left(w_0 - l + y - pg(y)\right)g''(y)$$

$$+ p^2(1-p)u''\left(w_0 - pg(y)\right)\left(g'(y)\right)^2 - p(1-p)u'\left(w_0 - pg(y)\right)g''(y)$$
$$< 0.$$

The global concavity of $U^{Pol}(y, g)$ in y implies that zero indemnification is not optimal if and only if

$$\left.\frac{\partial U^{Pol}(y,g)}{\partial y}\right|_{y=0} = pu'\left(w_0 - l\right)\left(1 - pg'(0)\right) - p(1-p)u'\left(w_0\right)g'(0) > 0,$$

or equivalently, $g'(0) < \frac{u'(w_0 - l)}{pu'(w_0-l)+(1-p)u'(w_0)}$. □

Remark 4. *Alternatively, we can obtain Theorem A.2 from Theorem A.1 by setting $F_X(x)$ to the form given in (A.5). In this case, the function $L(x)$ in Theorem A.1 reduces to a constant,*

$$L(x) = \frac{pu'\left(w_0 - l + I_g(l) - pg(I_g(l))\right)}{pu'\left(w_0 - l + I_g(l) - pg(I_g(l))\right) + (1-p)u'\left(w_0 - pg(I_g(l))\right)} - pg'(I_g(l)),$$

for $0 \le x < l$. One can deduce that $L(x) \le 0$. Indeed, suppose by way of contradiction that $L(x) > 0$. Then by Theorem A.1, we have $I_g(l) = l$, and thus

$$L(x) = \frac{pu'\left(w_0 - pg(l)\right)}{pu'\left(w_0 - pg(l)\right) + (1-p)u'\left(w_0 - pg(l)\right)} - pg'(l) = p(1 - g'(l)) \le 0,$$

which leads to a contradiction.

If $L(x) = 0$, then the optimal insurance coverage $y = I_g(l)$ solves the following equation in y,

$$\frac{u'\left(w_0 - l + y - pg(y)\right)}{pu'\left(w_0 - l + y - pg(y)\right) + (1-p)u'\left(w_0 - pg(y)\right)} = g'(y),$$

that is the first-order condition (A.6). If $L(x) < 0$, we deduce that zero indemnification is optimal, i.e., $I_g(l) = 0$, which implies that

$$L(x) = \frac{pu'\left(w_0 - l\right)}{pu'\left(w_0 - l\right) + (1-p)u'\left(w_0\right)} - pg'(0) < 0.$$

Therefore, zero indemnification is optimal if and only if

$$g'(0) \ge \frac{u'\left(w_0 - l\right)}{pu'\left(w_0 - l\right) + (1-p)u'\left(w_0\right)}.$$

Otherwise, the policyholder is better off participating, and the optimal contract $y = I_g(l)$ solves equation (A.6).

Note that, by the monotonicity and concavity of u, the upper bound of $g'(0)$ in Theorem A.2 satisfies

$$\frac{u'(w_0-l)}{pu'(w_0-l)+(1-p)u'(w_0)} > \frac{u'(w_0-l)}{pu'(w_0-l)+(1-p)u'(w_0-l)} = 1.$$

This means that the subset of premium principles within set $\mathcal{A}_0$ that policyholders are willing to participate is nonempty.

We then examine the insurer's dual problem. Similar to the analysis with proportional or deductible indemnities, we define the admissible premium principles to be the set

$$\mathcal{A}_3 = \left\{ g \in \mathcal{A}_0 : g'(0) < \frac{u'\left(w_0-l\right)}{pu'\left(w_0-l\right)+(1-p)u'\left(w_0\right)} \right\}.$$

According to the first-order condition (A.6), for any given $y \in (0,l)$, we define the set

$$\mathcal{A}_3(y) = \left\{ g \in \mathcal{A}_3 : g'(y) = \frac{u'\left(w_0-l+y-pg(y)\right)}{pu'\left(w_0-l+y-pg(y)\right)+(1-p)u'\left(w_0-pg(y)\right)} \right\}.$$

For a given coverage $y \in (0,l)$ and a premium principle g, the insurer's expected profit is given by

$$U^{In}(y,g) = p\left(g(y)-y\right).$$

Then the insurer's dual problem is given by

$$\max_{y\in(0,l)} \max_{g\in\mathcal{A}_3(y)} p\left(g(y)-y\right).$$

The following theorem shows that *expected-value premium principle* is optimal for a given coverage $y \in (0,l)$.

Theorem A.3. *For any given coverage $y \in (0,l)$, an optimal premium principle to*

$$\max_{g\in\mathcal{A}_3(y)} p\left(g(y)-y\right) \tag{A.7}$$

is an expected-value premium principle given by

$$g_y(x) = \frac{u'\left(w_0-l+y-c_y\right)}{pu'\left(w_0-l+y-c_y\right)+(1-p)u'\left(w_0-c_y\right)}x, \tag{A.8}$$

where

$$c_y = \sup\left\{ c \geq 0 : \frac{u'\left(w_0-l+y-c\right)y}{pu'\left(w_0-l+y-c\right)+(1-p)u'\left(w_0-c\right)} - \frac{c}{p} \geq 0 \right\}$$

is such that $c_y \in (0,\infty)$. Furthermore, $U^{In}(y,g_y) > 0$ for any $y \in (0,l)$.

Proof. For any $g \in \mathcal{A}_3(y)$, by $g(0) = 0$ and the convexity of g, we have

$$\frac{pg(y)}{p} = \int_0^y g'(x)\mathrm{d}x \le yg'(y) = \frac{u'\left(w_0 - l + y - pg(y)\right) y}{pu'\left(w_0 - l + y - pg(y)\right) + (1-p)u'\left(w_0 - pg(y)\right)}.$$

It follows that

$$\max_{g \in \mathcal{A}_3(y)} pg(y) \le \sup\left\{c \ge 0 : \frac{u'\left(w_0 - l + y - c\right) y}{pu'\left(w_0 - l + y - c\right) + (1-p)u'\left(w_0 - c\right)} - \frac{c}{p} \ge 0\right\} := c_y. \tag{A.9}$$

Further, since

$$\frac{u'\left(w_0 - l + y\right) y}{pu'\left(w_0 - l + y\right) + (1-p)u'\left(w_0\right)} > 0,$$

and

$$\begin{aligned}
&\limsup_{c \to \infty}\left(\frac{u'\left(w_0 - l + y - c\right) y}{pu'\left(w_0 - l + y - c\right) + (1-p)u'\left(w_0 - c\right)} - \frac{c}{p}\right) \\
&= \frac{y}{p}\limsup_{c \to \infty}\left(\frac{pu'\left(w_0 - l + y - c\right)}{pu'\left(w_0 - l + y - c\right) + (1-p)u'\left(w_0 - c\right)} - \frac{c}{y}\right) \\
&\le \frac{y}{p}\limsup_{c \to \infty}\left(1 - \frac{c}{y}\right) \\
&= -\infty,
\end{aligned}$$

we deduce that that $c_y \in (0, \infty)$.

Next we show that $g_y \in \mathcal{A}_3(y)$ and it achieves the optimality of Problem (A.7). By the definition of g_y and c_y, we have that

$$g_y(y) = \frac{u'\left(w_0 - l + y - c_y\right)}{pu'\left(w_0 - l + y - c_y\right) + (1-p)u'\left(w_0 - c_y\right)} y = \frac{c_y}{p}.$$

This together with (A.9) implies that g_y indeed achieves the optimality of Problem (A.7). We then verify $g_y \in \mathcal{A}_3(y)$. By $c_y = pg_y(y)$, we have

$$\begin{aligned}
g_y'(y) &= \frac{u'\left(w_0 - l + y - c_y\right)}{pu'\left(w_0 - l + y - c_y\right) + (1-p)u'\left(w_0 - c_y\right)} \\
&= \frac{u'\left(w_0 - l + y - pg_y(y)\right)}{pu'\left(w_0 - l + y - pg_y(y)\right) + (1-p)u'\left(w_0 - pg_y(y)\right)},
\end{aligned}$$

implying that g_y satisfies the first-order condition of $\mathcal{A}_3(y)$. Moreover, since the policyholder's expected utility under g_y is given by

$$U^{Pol}(x, g_y) := pu\left(w_0 - l + x - pg_y(x)\right) + (1-p)u\left(w_0 - pg_y(x)\right),$$

it is straightforward to verify that $\left.\frac{\partial U^{Pol}(x,g_y)}{\partial x}\right|_{x=y} = 0$ and $\frac{\partial^2 U^{Pol}(x,g_y)}{\partial x^2} < 0$ for any $x \in [0,1]$, implying that

$$0 < \left.\frac{\partial U^{Pol}(x,g_y)}{\partial x}\right|_{x=0} = pu'\left(w_0 - l\right)\left(1 - pg_y'(0)\right) + (1-p)u'\left(w_0\right)\left(-pg_y'(0)\right),$$

which is equivalent to $g_y'(0) < \frac{u'(w_0 - l)}{pu'(w_0-l)+(1-p)u'(w_0)}$. Thus, $g_y \in \mathcal{A}_3(y)$ and it achieves the optimality of Problem (A.7).

Lastly, we show that $U^{In}(y, g_y) > 0$ for any $y \in (0, l)$. Indeed,

$$\begin{aligned}
U^{In}(y, g_y) &= p\left(g_y(y) - y\right) \\
&= p\left(\frac{u'\left(w_0 - l + y - c_y\right)y}{pu'\left(w_0 - l + y - c_y\right) + (1-p)u'\left(w_0 - c_y\right)} - y\right) \\
&> p\left(\frac{u'\left(w_0 - l + y - c_y\right)y}{pu'\left(w_0 - l + y - c_y\right) + (1-p)u'\left(w_0 - l + y - c_y\right)} - y\right) \\
&= 0,
\end{aligned}$$

where the inequality follows from $y < l$. □

Finally, we present a theorem that summarizes the equivalence between the original problem and the dual problem, and confirms the Bowley optimality of the expected-value premium principle for binary losses. The proof is omitted, as it follows the same lines as Theorems 2.3.3 and 2.3.4.

Theorem A.4. *For binary losses,*

$$\max_{g \in \mathcal{A}_3} U^{In}(I_g, g) = \max_{y \in (0,l)} U^{In}(y, g_y). \tag{A.10}$$

Moreover, the mechanism $(y^*, g_{y^*}) \in \mathcal{BO} \cap \mathcal{IR}$*, where* $y^* \in \arg\max_{y \in (0,l)} U^{In}(y, g_y)$.

Chapter 3

Bowley-Optimal Distortion Premium Principles

3.1 Introduction

Risk measures are fundamental tools in the design and evaluation of insurance contracts, as they quantify the potential losses faced by agents and guide optimal risk-sharing decisions. Among various classes of risk measures, distortion risk measures stand out due to their flexibility in capturing different risk attitudes through the application of distortion functions to the loss distribution. This class encompasses widely used measures such as Value-at-Risk (VaR) and Tail Value-at-Risk (TVaR), and has been widely adopted in both risk evaluation and premium calculation.

In the context of optimal (re)insurance design, risk measures have been directly incorporated into the optimization criteria. For example, Cai and Tan (2007) introduce VaR and TVaR into the insurance optimization problem under the assumption that the indemnity function is of deductible form and that the premium is calculated using the expected value of the indemnity function with a positive loading factor—commonly referred to as the expected-value premium principle. Cai et al. (2008) relax the restriction on the form of the indemnity function and study more general contract structures. They show that deductible, proportional, and mixed contracts can be optimal for minimizing VaR or TVaR, still under the expected-value premium principle. Extending beyond this setting, Assa (2015) consider a more general framework in which both the decision-maker's preferences and the premium calculation are modeled using distortion risk measures—a class that includes VaR and TVaR as special cases. In this setting, the distortion premium reflects the cost of transferring risk, while the agent's perception of risk is represented by a separate distortion function.

The result in Assa (2015) not only addresses the individual optimization problem of either the policyholder or the insurer but also provides valuable insights into the solution of a social planner problem, where both parties are involved and the objective is to minimize the aggregate risk, measured as the sum of two distortion risk measures—one from each party. In another study, Cai et al. (2016) examine a setting in which both the insurer and the policyholder are VaR minimizers. They propose solving the problem by considering a convex combination of the two parties' risk measures, while the cost of transferring risk is determined using the expected-value premium principle. This framework is further extended in Cai et al. (2017), where both parties are assumed to be TVaR minimizers, allowing for a broader exploration of optimal contract structures under distortion risk measures.

In the aforementioned individual or two-party optimization problems, the decision-making process is simultaneous: there is no designated leader or follower, and both parties make their choices at the same time. The solution in these settings corresponds to the optimal insurance coverage that minimizes one or two party's respective risk measure. An alternative and equally important direction in the study of insurance market optimization—particularly in monopolistic markets—involves sequential decision-making. This setup is typically modeled as a Stackelberg game, or more specifically, a Bowley solution. The application of this game-theoretic structure to insurance was first introduced in Chan and Gerber (1985). The concrete formulation of this sequential optimization process can be found in equations (1.3) and (1.4). The Stackelberg structure reflects the asymmetric power dynamics inherent in a monopolistic market. In this setting, the insurer—acting as the market leader—has greater control over the design and pricing of insurance contracts due to the absence of competition. Consequently, the insurer can not only determine the form of the insurance coverage but also choose the pricing rule, thereby shaping both the quantity and cost of insurance in a way that maximizes its own objective while anticipating the response of the policyholder.

The sequential game framework introduced in Chan and Gerber (1985) has since been extended to various settings. In continuous-time insurance optimization models, it has been developed further in Chen and Shen (2018), Li and Young (2021), and Cao et al. (2022). For one-period insurance contracts, related extensions can be found in Cheung et al. (2019), Chi et al. (2020), Boonen et al. (2021), Zhu et al. (2023), and Ghossoub and Zhu (2024).

In this chapter, we characterize the Bowley solution in a monopolistic insurance market with a single policyholder. For related studies in non-monopolistic settings, see Zhu et al. (2023); for models with multiple policyholders or heterogeneous types, refer to Boonen et al. (2021) and Ghossoub and Zhu (2024), respectively. We assume that the policyholder evaluates insurance contracts using a distortion risk measure, following the approach in Assa (2015) and Cheung et al. (2019). The admissible class of premium principles is given by distortion premium principles—also known as Wang's premium principle—originally introduced in Wang (1996) and further developed in Wang et al. (1997). These works demonstrate that distortion premiums satisfy a set of axioms that reflect essential properties of market-consistent insurance pricing. The insurer is assumed to be risk-neutral and seeks to maximize expected profit. To avoid moral hazard, we restrict the class of admissible indemnity functions to be 1-Lipschitz continuous. The modeling framework is most closely related to that of Cheung et al. (2019). However, in contrast to their study—which assumes that the policyholder is either strictly risk-averse or a VaR minimizer—we impose no specific restriction on the curvature of the policyholder's distortion function. Our analysis employs the quantile formulation and proceeds in two steps. First, for a fixed pricing distortion function, we characterize the optimal indemnity function that minimizes the policyholder's perceived risk. Second, using the resulting optimal indemnity, we solve the insurer's problem of selecting the pricing distortion function that maximizes expected profit.

We show that the optimal indemnity function is of layer type, with its structure driven by the interplay between the policyholder's and insurer's distortion functions. In particular, for strictly risk-averse individuals (e.g., represented by concave distortion functions), the optimal contract provides full insurance. For VaR minimizers, the optimal contract involves a coverage limit. For individuals with inverse-S shaped distortion functions (e.g., reflecting loss aversion or probability weighting), the optimal contract takes the form of a deductible. These results extend the work of Cheung et al. (2019) by allowing for a broader class of policyholders, including those whose preferences are characterized by general distortion risk measures—capturing not only traditional risk aversion but also behavioral distortions. Moreover, if we restrict our analysis to the first step and fix the premium principle—for instance, by adopting the expected-value premium principle—our model reduces to the individual optimization problem studied in Assa (2015) and related literature.

The chapter is organized as follows. Section 3.2 introduces the model and formulates the problem. Section 3.3 examines the optimal insurance indemnity functions for a given pricing rule. Section 3.4 analyzes the insurer's optimal pricing distortion. Finally, Section 3.5 concludes the chapter.

3.2 Model Setup and Problem Formulation

3.2.1 The Market

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let X be a random variable representing a potential risk with distribution function F_X. A positive realization of X is seen as a loss. In this section, we assume that the policyholder makes decisions based on a distortion risk measure. The insurer is risk-neutral and the objective is to maximize the expected profit. The policyholder's risk attitude is characterized by a distortion function, defined as follows.

Definition 3.2.1. *A mapping $T : [0,1] \to [0,1]$ is a distortion function if $T(0) = 0$, $T(1) = 1$, and T is non-decreasing and differentiable on $[0,1]$. In addition, it is a*

1. *convex distortion if $T'(t)$ is non-decreasing on $[0,1]$.*
2. *concave distortion if $T'(t)$ is non-increasing on $[0,1]$.*
3. *Inverse-S shaped distortion (ISSD) if it is twice-differentiable on $(0,1)$, and there exists a $t_0 \in (0,1)$ such that $T'(t)$ is strictly deceasing on $(0,t_0)$ and strictly increasing on $(t_0,1)$. Moreover, $\lim_{t\downarrow 0} T'(t) > 1$, and $\lim_{t\uparrow 1} T'(t) > 1$.*

If a policyholder faces a random risk X, then the policyholder's risk preference under a distortion risk measure is defined by the following Choquet integral:

$$\rho(X) := \int X \,\mathrm{d}T \circ \mathbb{P} := \int_0^{+\infty} T(\mathbb{P}(X \geq x))\,\mathrm{d}x + \int_{-\infty}^{0} [T(\mathbb{P}(X \geq x)) - 1]\,\mathrm{d}x. \tag{3.1}$$

where T is a distortion function. If the distortion function T is concave, then the policyholder exhibits strict risk aversion, placing greater weight on unfavorable outcomes. Conversely, if T is convex, the policyholder is risk seeking, emphasizing favorable outcomes. If T is an ISSD, this indicates that the policyholder places heightened attention on extreme events, particularly in the tails of the loss distribution.

Suppose that the policyholder faces a bounded loss $X \in [0, M]$, and this risk is to be shared between the policyholder and the insurance company. The portion of the loss covered by the insurer is denoted by $I(X)$, where I is selected from the following class of 1-Lipschitz continuous indemnity functions:

$$\mathcal{I}_L := \Big\{ I : [0,M] \to [0,M] \;\Big|\; I(0) = 0,\ 0 \leq I(x_1) - I(x_2) \leq x_1 - x_2, \forall\, x_2 \leq x_1 \in [0,M] \Big\}.$$

The remaining part of the loss retained by the policyholder is $R(X) := X - I(X)$. The cost of transferring the loss $I(X)$ to the insurer is calculated using a distortion premium principle, which evaluates the expected loss under a distorted probability measure induced by a distortion function g. The resulting premium charged by the insurer is given by

$$\pi_g\left(I(X)\right) = \int I(X)\, dg \circ \mathbb{P} \tag{3.2}$$

Denote an insurance contract with coverage function I and premium calculated under the distortion function g by (g, I). After purchasing this contract, the policyholder's end-of-period total loss becomes

$$X - I(X) + \pi_g(I(X)),$$

where $\pi_g(I(X))$ is the premium paid to the insurer. The policyholder evaluates this net loss using a distortion risk measure induced by a preference distortion function T. The associated perceived risk of the insurance contract is then given by the Choquet integral:

$$V^{Pol}(g, I) := \rho\left(X - I(X) + \pi_g\left(I(X)\right)\right). \tag{3.3}$$

Issuing the contract (g, I) to the policyholder ensures profitability for the risk-neutral insurer, whose profit is the premium received minus the indemnity paid. The expected profit is given by:

$$V^{In}(g, I) = \pi_g\left(I(X)\right) - \mathbb{E}\left[I(X)\right]. \tag{3.4}$$

In this chapter, we aim to design the insurance contract (g, I) that achieves *Bowley Optimality* for both parties. That is, a contract which maximizes the insurer's profit while ensuring that the policyholder minimizes their risk measure, given the charged premium. Mathematically, this is defined as follows.

Definition 3.2.2. *A market mechanism (g^*, I^*) is Bowley Optimal (BO), or a Stackelberg Equilibrium, if:*

1. $I^* \in \underset{\overline{I} \in \mathcal{I}_L}{\arg\min}\, V^{Pol}(g^*, \overline{I})$*;*
2. $V^{In}(g^*, I^*,) \geq V^{In}(g, I)$*, for all (g, I) such that* $I \in \underset{\overline{I} \in \mathcal{I}_L}{\arg\min}\, V^{Pol}(g, \overline{I})$*.*

To achieve this goal, we adopt a quantile-based approach. The details of this method are presented in the next section.

3.2.2 The Quantile Approach

We assume that the random loss X satisfies the following assumption.

Assumption 3.2.1. *The random loss X has a strictly increasing distribution function F_X.*

By Föllmer and Schied (2011, Lemma A.21), Assumption 3.2.1 guarantees that the random variable X can be represented as $X = F_X^{-1}(U)$, where U is uniformly distributed on $(0,1)$ and F_X^{-1} is the quantile function of X. When interpreting X as a loss, and applying the definition of the Choquet integral in (3.1), the distortion risk measure of X can be equivalently expressed in its quantile form as:

$$\begin{aligned}\rho(X) &= \int X \,\mathrm{d}T \circ \mathbb{P} = \int_0^{+\infty} T\left(1 - F_X(x)\right)\,\mathrm{d}x \\ &= \int_0^1 T(1-t)\,\mathrm{d}F_X^{-1}(t) = \int_0^1 \left(F_X^{-1}\right)'(t) T(1-t)\,\mathrm{d}t. \end{aligned} \tag{3.5}$$

We refer to the representation in (3.5) as the *quantile form* of the distortion risk measure of X. Using this formulation, we now rewrite the policyholder's risk measure in (3.3) and the insurer's expected profit in (3.4) in a similar quantile-based form.

For every $I \in \mathcal{I}_L$, the indemnity $I(X)$ and the retention $X - I(X)$ are comonotone. By the comonotonic additivity of the quantile function, it follows that:

$$F_X^{-1} = F_{I(X)+X-I(X)}^{-1} = F_{I(X)}^{-1} + F_{X-I(X)}^{-1},$$

Thus, the quantile function of the retention is given by

$$q(t) := F_{R(X)}^{-1}(t) = F_X^{-1}(t) - F_{I(X)}^{-1}(t),$$

for $t \in [0,1]$. The function I belongs to $\mathcal{I}_L$ if and only if the corresponding quantile function q lies in the class:

$$\mathcal{Q}_L = \left\{ q : (0,1) \to \mathbb{R} \,\middle|\, q(0) = 0,\ 0 \le q'(t) \le \left(F_X^{-1}\right)'(t) \right\}.$$

Using this representation, the distortion premium $\Pi_g\left(I(X)\right)$ can be expressed in terms of the quantile function q as:

$$\pi_g\left(I(X)\right) = \int_0^1 \left(F_{I(X)}^{-1}\right)'(t)\, g(1-t)\,\mathrm{d}t = \int_0^1 \left(\left(F_X^{-1}\right)'(t) - q'(t)\right) g(1-t)\,\mathrm{d}t := \pi_g(q). \tag{3.6}$$

Similarly, the policyholder's welfare given in (3.3) can be rewritten as:

$$\begin{aligned} V^{Pol}(g, I) &= \int (X - I(X) + \pi_g(I(X)))\, \mathrm{d}T \circ \mathbb{P} \\ &= \pi_g(I(X)) + \int R(X)\, \mathrm{d}T \circ \mathbb{P} \\ &= \pi_g(q) + \int_0^1 q'(t)\, T(1-t)\mathrm{d}t := V^{Pol}(g, q), \end{aligned} \tag{3.7}$$

where the second equality uses the fact that $\pi_g(I(X))$ is constant with respect to the probability measure.

3.3 The Policyholder's Problem

In this section, we focus on the policyholder's problem, which is to determine an optimal insurance strategy under a given premium structure:

$$\min_{I \in \mathcal{I}_L} V^{Pol}(g, I). \tag{3.8}$$

Using the representation in (3.7), this problem reduces to finding the optimal quantile retention function q_g for a given pricing distortion g, by solving:

$$\min_{q \in \mathcal{Q}_L} V^{Pol}(g, q). \tag{3.9}$$

The following lemma delineates the optimal retention quantile when examining the solution through the distortion function g.

Lemma 3.3.1. *For a given g, a quantile q_g is optimal for problem (3.9) if and only if*

$$(q_g)'(t) = \begin{cases} 0, & g(1-t) < T(1-t), \\ \phi_g(t) \in \left[0, \left(F_X^{-1}\right)'(t)\right], & g(1-t) = T(1-t), \\ \left(F_X^{-1}\right)'(t), & g(1-t) > T(1-t). \end{cases} \tag{3.10}$$

Proof. Substituting the expression for the premium (3.6) into the welfare function (3.7), we obtain:

$$V^{Pol}(g, q) = \pi_g(q) + \int_0^1 q'(t)\, T(1-t)\mathrm{d}t$$

$$= \int_0^1 \left(\left(F_X^{-1}\right)'(t) - q'(t) \right) g(1-t)\,\mathrm{d}t + \int_0^1 q'(t)\,T(1-t)\mathrm{d}t$$
$$= \int_0^1 \left(F_X^{-1}\right)'(t)\,g(1-t)\,\mathrm{d}t + \int_0^1 q'(t)\,\left(T(1-t) - g(1-t)\right)\mathrm{d}t.$$

The first term in the final line is independent of q, so minimizing the risk measure $V^{Pol}(g,q)$ over $q \in \mathcal{Q}_L$ is equivalent to solving:

$$\min_{q \in \mathcal{Q}_L} \int_0^1 q'(t)\,(T(1-t) - g(1-t))\,\mathrm{d}t.$$

By the Marginal Indemnity Function Approach given in Assa (2015), the optimal quantile function q_g must satisfy:

$$(q_g)'(t) = \begin{cases} 0, & g(1-t) < T(1-t), \\ \in \left[0, \left(F_X^{-1}\right)'(t)\right], & g(1-t) = T(1-t), \\ \left(F_X^{-1}\right)'(t), & g(1-t) > T(1-t). \end{cases} \tag{3.11}$$

□

This rule characterizes the structure of the Bowley-optimal indemnity for a given pricing distortion g: it prescribes full retention (no indemnity) when the pricing weight g is lower than the policyholder's valuation T, full coverage when the pricing weight exceeds the policyholder's valuation, and arbitrary choice of indemnity in the case of equality.

The following lemma establishes the equivalence between the solution obtained via the quantile approach and the solution to the original policyholder's optimization problem.

Lemma 3.3.2. *For a given g, the quantile function $q_g(t)$ is optimal for problem (3.9) if and only if the indemnity function $I_g(t) = t - q_g(F_X(t))$ is optimal for problem (3.8).*

Proof. Suppose that $q_g(t)$ is optimal for problem (3.9). Let I be any other feasible solution to problem (3.8), and $q(t)$ its retention quantile function. We have

$$V^{Pol}(g, I) = \int (X - I(X) + \pi_g(I(X)))\,\mathrm{d}T \circ \mathbb{P}$$
$$= \pi_g(q) + \int_0^1 q'(t)\,T(1-t)\mathrm{d}t$$

$$
\begin{aligned}
&\geq \pi_g(q_g) + \int_0^1 (q_g)'(t)\, T(1-t)\mathrm{d}t \\
&= \pi_g\left(I_g(X)\right) + \int_0^1 \left(F_X^{-1}(t) - I_g\left(F_X^{-1}(t)\right)\right)' T(1-t)\mathrm{d}t \\
&= \pi_g\left(I_g(X)\right) + \int_0^1 \left(F_X^{-1}\right)'(t)\, T(1-t)\,\mathrm{d}t - \int_0^1 \left(F^{-1}_{I_g(X)}\right)'(t)\, T(1-t)\mathrm{d}t \\
&= \int \left(X - I_g(X) + \pi_g\left(I_g(X)\right)\right) \mathrm{d}T \circ \mathbb{P} = V^{Pol}\left(g, I_g\right),
\end{aligned}
$$

where the third equality is by the fact that $q_g(t) = q_g\left(F_X\left(F_X^{-1}(t)\right)\right) = F_X^{-1}(t) - I_g\left(F_X^{-1}(t)\right)$, and

$$
\begin{aligned}
\pi_g(q_g) &= \int_0^1 \left(\left(F_X^{-1}\right)(t) - q_g(t)\right)' g(1-t)\mathrm{d}t = \int_0^1 \left(I_g\left(F_X^{-1}(t)\right)\right)' g(1-t)\mathrm{d}t \\
&= \int_0^1 \left(F^{-1}_{I_g(X)}\right)'(t)\, g(1-t)\mathrm{d}t = \pi_g\left(I_g(X)\right).
\end{aligned}
$$

This implies that $I_g(t)$ minimizes value function (3.3).

On the other hand, suppose that $I_g(t) = t - q_g\left(F_X(t)\right)$ is optimal for problem (3.8). If $q(t)$ is a feasible solution to problem (3.9), and $I(t) = t - q\left(F_X(t)\right)$, we have

$$
\begin{aligned}
V^{Pol}(g, q) &= \pi_g(q) + \int_0^1 q'(t)\, T(1-t)\mathrm{d}t \\
&= \int \left(X - I(X) + \pi_g\left(I(X)\right)\right) \mathrm{d}T \circ \mathbb{P} \\
&\geq \int \left(X - I_g(X) + \pi_g\left(I_g(X)\right)\right) \mathrm{d}T \circ \mathbb{P} \\
&= \pi_g(q_g) + \int_0^1 (q_g)'(t)\, T(1-t)\mathrm{d}t = V^{Pol}\left(g, q_g\right),
\end{aligned}
$$

which indicates the optimality of q_g for problem (3.9). This completes the proof. □

This result characterizes the optimal indemnity function I_g corresponding to a given premium distortion function g. The indemnity function is determined via the associated quantile function q_g. However, the solution is not unique, and this non-uniqueness stems from the potential non-uniqueness of q_g as defined in (3.10). Specifically, if $g(1-t) = T(1-t)$, then any choice of $\phi_g \in [0, F_X^{-1}(t)]$ yields the same risk exposure level for the policyholder and thus equally optimizes their objective function. Nevertheless, these different choices can lead to varying profits for the insurer. A more detailed discussion of this issue is provided in the proof of the theorem concerning the insurer's problem. In what follows, we analyze the optimal indemnity function and the corresponding pricing distortion from the insurer's perspective.

3.4 The Insurer's Problem

In this section, we will delve into the second-step optimization problem of the Stackelberg equilibrium. This involves determining the optimal premium distortion function given that the policyholder's welfare is maximized. Specifically, the objective is to identify an optimal contract $(g^*, I^*_{g^*})$ that solves the following problem,

$$\max_{g,I_g} V^{In}(g,I) = \max_{g,I_g} \pi_g\left(I_g\right) - \mathbb{E}\left[I_g(X)\right]. \tag{3.12}$$

Building on the previous section, we will explore the optimal solution using the quantile approach. This entails examining the optimal distortion premium by representing profit as a function of the retention quantile:

$$\max_{g,q_g} V^{In}(g,q_g) = \max_{g,q_g} \pi_g\left(q_g\right) - \int_0^1 \left(F_X^{-1}(t) - q_g(t)\right) \mathrm{d}t. \tag{3.13}$$

The following result elaborates on the solution.

Theorem 3.4.1. *The contract* $\left(g^*, q^*_{g^*}\right)$ *is optimal for problem* (3.13) *if and only if*

$$g^*(t) = 1 - \widetilde{g}^*(1-t),$$

where $\widetilde{g}^*(t)$ *is defined as*

$$\widetilde{g}^*(t) = \begin{cases} \widetilde{T}(t), & \text{if } \widetilde{T}(t) < t, \\ \in \left[\sup\left\{z < t : \widetilde{g}^*(z) < z\right\}, \widetilde{T}(t)\right], & \text{if } \widetilde{T}(t) \geq t, \end{cases} \tag{3.14}$$

with $\widetilde{T}(t) := 1 - T(1-t)$. *Moreover, the optimal quantile function* $q^*_{g^*}$ *satisfies*

$$\left(q^*_{g^*}\right)'(t) = \begin{cases} 0, & \text{if } \widetilde{T}(t) < t, \\ \phi_{g^*}(t) \in \left[0, \left(F_X^{-1}\right)'(t)\right], & \text{if } \widetilde{g}^*(t) = \widetilde{T}(t) = t, \\ \left(F_X^{-1}\right)'(t), & \text{if } \widetilde{g}^*(t) < \widetilde{T}(t) = t \text{ or } \widetilde{T}(t) > t. \end{cases} \tag{3.15}$$

Proof. By the expression of $\Pi_g\left(q_g\right)$ in (3.6), we have

$$\begin{aligned}
V^{In}\left(g, q_g\right) &= \int_0^1 \left(\left(F_X^{-1}\right)'(t) - \left(q_g\right)'(t)\right) g(1-t)\,\mathrm{d}t - \int_0^1 \left(F_X^{-1}(t) - q_g(t)\right)\,\mathrm{d}t \\
&= \int_0^1 \left(\left(F_X^{-1}\right)'(t) - \left(q_g\right)'(t)\right) g(1-t)\,\mathrm{d}t - \int_0^1 \left(\left(F_X^{-1}\right)'(t) - \left(q_g\right)'(t)\right)(1-t)\,\mathrm{d}t \\
&= \int_0^1 \left(\left(F_X^{-1}\right)'(t) - \left(q_g\right)'(t)\right)\left(t-1+g(1-t)\right)\mathrm{d}t. \qquad (3.16)
\end{aligned}$$

Let $h(t) = t - 1 + g(1-t)$, and define the sets: $\mathcal{A}_1\left(h\right) = \left\{t : h(t) < t - \widetilde{T}(t)\right\}$, $\mathcal{A}_2\left(h\right) = \left\{t : h(t) = t - \widetilde{T}(t)\right\}$, and $\mathcal{A}_3\left(h\right) = \left\{t : h(t) > t - \widetilde{T}(t)\right\}$, where $\widetilde{T}(t) = 1 - T(1-t)$ represents the conjugate of T. The quantile q_g in (3.10) can be rewritten as

$$\left(q_h\right)'(t) = \begin{cases} 0, & t \in \mathcal{A}_1\left(h\right), \\ \phi_h(t) \in \left[0, \left(F_X^{-1}\right)'(t)\right], & t \in \mathcal{A}_2\left(h\right), \\ \left(F_X^{-1}\right)'(t), & t \in \mathcal{A}_3\left(h\right). \end{cases} \qquad (3.17)$$

The form of q_h indicates that for a given h (i.e., for a given g), the marginal quantile for retention is 0 when $t \in \mathcal{A}_1\left(h\right)$, and it is $\left(F_X^{-1}\right)'(t)$ when $t \in \mathcal{A}_3\left(h\right)$. When $t \in \mathcal{A}_2\left(h\right)$, the optimal retention allows for some flexibility, as long as $q_h \in \mathcal{Q}_L$. The arbitrary choice of $\phi_h(t)$ does not affect the policyholder's risk exposure level, but it does impact the insurer's profit. To maximize the insurer's profit, we must take a further step to determine the value of ϕ_h. We achieve this by analyzing the profit over a finer partition $\{\mathcal{B}_i\}_{i=1}^3$ when $t \in \mathcal{A}_2\left(h\right)$, where $\mathcal{B}_1 = \{t : \widetilde{T}(t) < t\}$, $\mathcal{B}_2 = \{t : \widetilde{T}(t) = t\}$, and $\mathcal{B}_3 = \{t : \widetilde{T}(t) > t\}$. The profit (3.16) can then be expressed as,

$$\begin{aligned}
V^{In}\left(h, q_h\right) &:= \int_0^1 \left(F_X^{-1}\right)'(t)\, h(t)\, \mathbb{1}_{\mathcal{A}_1(h)}(t)\,\mathrm{d}t + \int_0^1 \left(\left(F_X^{-1}\right)'(t) - \phi_h(t)\right) h(t)\, \mathbb{1}_{\mathcal{A}_2(h)}(t)\,\mathrm{d}t \\
&= \int_0^1 \left(F_X^{-1}\right)'(t)\, h(t)\, \mathbb{1}_{\mathcal{A}_1(h)}(t)\,\mathrm{d}t \\
&+ \int_0^1 \left(\left(F_X^{-1}\right)'(t) - \phi_h(t)\right) h(t)\, \mathbb{1}_{\mathcal{A}_2(h)}(t) \left(\mathbb{1}_{\mathcal{B}_1}(t) + \mathbb{1}_{\mathcal{B}_2}(t) + \mathbb{1}_{\mathcal{B}_3}(t)\right)\mathrm{d}t. \qquad (3.18)
\end{aligned}$$

On $\mathcal{A}_2\left(h\right) \cap \mathcal{B}_1$, we have that $h(t) = t - \widetilde{T}(t) > 0$. Thus, the insurer's profit (3.18) is decreasing in $\phi_h(t)$, and the optimal choice is to set $\phi_h(t) = 0$ to maximize profit. On

$\mathcal{A}_2(h) \cap \mathcal{B}_2$, we have $h(t) = t - \widetilde{T}(t) \equiv 0$, so the profit contribution is always zero regardless of the value of $\phi_h(t)$. Therefore, $\phi_h(t)$ an take any value in the interval $\left[0, \left(F_X^{-1}\right)'(t)\right]$ without affecting the insurer's profit. On $\mathcal{A}_2(h) \cap \mathcal{B}_3$, we get $h(t) = t - \widetilde{T}(t) < 0$. In this case, the insurer's profit increases with $\phi_h(t)$, so it is optimal to set $\phi_h(t) = \left(F_X^{-1}\right)'(t)$, on $\mathcal{A}_2(h) \cap \mathcal{B}_3$. Therefore, for a given h, the derivative of the optimal retention quantile is

$$\left(q_h^*\right)'(t) = \begin{cases} 0, & t \in \mathcal{A}_1(h) \cup \left(\mathcal{A}_2(h) \cap \mathcal{B}_1\right), \\ \phi_h(t) \in \left[0, \left(F_X^{-1}\right)'(t)\right], & t \in \mathcal{A}_2(h) \cap \mathcal{B}_2, \\ \left(F_X^{-1}\right)'(t), & t \in \mathcal{A}_3(h) \cup \left(\mathcal{A}_2(h) \cap \mathcal{B}_3\right). \end{cases} \tag{3.19}$$

As $\mathcal{A}_1(h)$ and $\mathcal{A}_2(h)$ are disjoint, the insurer's profit can be expressed as

$$\begin{aligned} V^{In}\left(h, q_h^*\right) &= \int_0^1 \left(F_X^{-1}\right)'(t)\, h(t) \left(\mathbb{1}_{\mathcal{A}_1(h)}(t) + \mathbb{1}_{\mathcal{A}_2(h) \cap \mathcal{B}_1}(t)\right) \mathrm{d}t \\ &= \int_0^1 \left(F_X^{-1}\right)'(t)\, h(t) \left(\mathbb{1}_{(\mathcal{A}_1(h) \cup \mathcal{A}_2(h)) \cap \mathcal{B}_1}(t) + \mathbb{1}_{\mathcal{A}_1(h) \cap (\mathcal{B}_2 \cup \mathcal{B}_3)}(t)\right) \mathrm{d}t. \end{aligned}$$

Hence, the insurer's problem reduces to solving the following optimization:

$$\begin{aligned} &\max_h V^{In}\left(h, q_h^*\right) \\ &= \max_h \left\{ \int_0^1 \left(F_X^{-1}\right)'(t)\, h(t) \left(\mathbb{1}_{\{h(t) \le t - \widetilde{T}(t)\}} \mathbb{1}_{\{\widetilde{T}(t) < t\}} + \mathbb{1}_{\{h(t) < t - \widetilde{T}(t)\}} \mathbb{1}_{\{\widetilde{T}(t) \ge t\}}\right) \mathrm{d}t \right\}. \end{aligned} \tag{3.20}$$

We solve problem (3.20) by maximizing the integrand pointwise. For each fixed $t_0 \in (0,1)$, define the auxiliary function

$$m(y; t_0) := y \left(\mathbb{1}_{\{y \le t_0 - \widetilde{T}(t_0)\}} \mathbb{1}_{\{\widetilde{T}(t_0) < t_0\}} + \mathbb{1}_{\{y < t_0 - \widetilde{T}(t_0)\}} \mathbb{1}_{\{\widetilde{T}(t_0) \ge t_0\}}\right),$$

which represents the pointwise contribution of $h(t_0) = y$ to the integral in (3.20). The goal is to find the value of y to maximizes $m(y; t_0)$ for each $t_0 \in (0,1)$. We observe that $m(y; t_0)$ is a piecewise linear function in y for fixed t_0. Moreover, let

$$y_{t_0} \in \arg\max_y \; m(y; t_0),$$

then the maximum value of $m(y; t_0)$ is given by

$$m(y_{t_0}; t_0) = \begin{cases} t_0 - \widetilde{T}(t_0), & \widetilde{T}(t_0) < t_0, \\ 0, & \widetilde{T}(t_0) \ge t_0, \end{cases}$$

and any maximizer y_{t_0} satisfies

$$y_{t_0} \begin{cases} = t_0 - \widetilde{T}(t_0), & \widetilde{T}(t_0) < t_0, \\ \geq t_0 - \widetilde{T}(t_0), & \widetilde{T}(t_0) \geq t_0, \end{cases} \tag{3.21}$$

for any $t_0 \in [0,1]$.

We now proceed to establish the characterization of the optimal solution: A function h^* maximizes the insurer's profit in problem (3.20) if and only if it satisfies, for every t_0,

$$h^*(t_0) = y_{t_0}$$

with y_{t_0} as given in (3.21).

We first study the necessary condition. Assume that h^* is a solution to problem (3.20). If $\widetilde{T}(t_0) < t_0$, we show that $h^*(t_0) = t_0 - \widetilde{T}(t_0)$ for any $t_0 \in (0,1)$. Suppose, by way of contradiction, that $h^*(t_0) \neq t_0 - \widetilde{T}(t_0)$. As $\widetilde{T}(t_0) < t_0$, we can find an arbitrary small $\epsilon > 0$ such that $\widetilde{T}(t) < t$ when $t \in (t_0 - \epsilon, t_0 + \epsilon)$. If $h^*(t_0) < t_0 - \widetilde{T}(t_0)$, we have that $m(h^*(t_0); t_0) = h^*(t_0)$. Then there exists a function $\widetilde{h}$ such that $\widetilde{h}(t) = y_t > h^*(t)$ when $t \in (t_0 - \epsilon, t_0 + \epsilon)$, and $\widetilde{h}(t) = h^*(t)$ when $t \notin (t_0 - \epsilon, t_0 + \epsilon)$, where y_t is given in (3.21). Then, we have that

$$\begin{aligned} V^{In}\left(\widetilde{h}, q^*_{\widetilde{h}}\right) - V^{In}\left(h^*, q^*_{h^*}\right) &= \int_{t_0-\epsilon}^{t_0+\epsilon} \left(F_X^{-1}\right)'(t)\ \left(m(y_t;t) - m\left(h^*(t);t\right)\right) \mathrm{d}t \\ &= \int_{t_0-\epsilon}^{t_0+\epsilon} \left(F_X^{-1}\right)'(t)\ \left(y_t - h^*(t)\right) \mathrm{d}t > 0, \end{aligned}$$

since $y_t > h^*(t)$ when $t \in (t_0 - \epsilon, t_0 + \epsilon)$. This contradicts to the fact that h^* is optimal for problem (3.20). Similarly, if $h^*(t_0) > t_0 - \widetilde{T}(t_0)$, we have that $m\left(h^*(t_0); t_0\right) = 0$. With the same ϵ, there exists a function $\widetilde{h}$ such that $\widetilde{h}(t) = y_t < h^*(t)$ when $t \in (t_0 - \epsilon, t_0 + \epsilon)$. Meanwhile, $\widetilde{h}(t) = h^*(t)$ when $t \notin (t_0 - \epsilon, t_0 + \epsilon)$. It is seen that

$$\begin{aligned} V^{In}\left(\widetilde{h}, q^*_{\widetilde{h}}\right) - V^{In}\left(h^*, q^*_{h^*}\right) &= \int_{t_0-\epsilon}^{t_0+\epsilon} \left(F_X^{-1}\right)'(t)\ \left(m(y_t;t) - m\left(h^*(t);t\right)\right) \mathrm{d}t \\ &= \int_{t_0-\epsilon}^{t_0+\epsilon} \left(F_X^{-1}\right)'(t)\ \left(y_t - 0\right) \mathrm{d}t > 0, \end{aligned}$$

which is a contradiction. Therefore, we have that $h^*(t_0) = t_0 - \widetilde{T}(t_0)$ for any $t_0 \in (0,1)$ when $\widetilde{T}(t_0) < t_0$.

If $\widetilde{T}(t_0) \geq t_0$, we show that $h^*(t_0) \geq t_0 - \widetilde{T}(t_0)$. Suppose, by way of contradiction, that $h^*(t_0) < t_0 - \widetilde{T}(t_0)$. Then $m\left(h^*(t_0); t_0\right) = h^*(t_0) < t_0 - \widetilde{T}(t_0) \leq 0$. Since $\widetilde{T}(t_0) \geq t_0$, then there is an arbitrary small $\epsilon > 0$ such that $\widetilde{T}(t) \geq t$ on the interval $(t_0 - \epsilon, t_0]$ or $[t_0, t_0 + \epsilon)$. Assume that $\widetilde{T}(t) \geq t$ on $(t_0 - \epsilon, t_0]$. There exists a function $\widetilde{h}$ such that $\widetilde{h}(t) = y_t > h^*(t)$ when $t \in (t_0 - \epsilon, t_0]$, and $\widetilde{h}(t) = h^*(t)$ when $t \notin (t_0 - \epsilon, t_0]$. Therefore,

$$\begin{aligned} V^{In}\left(\widetilde{h}, q^*_{\widetilde{h}}\right) - V^{In}\left(h^*, q^*_{h^*}\right) &= \int_{t_0-\epsilon}^{t_0} \left(F_X^{-1}\right)'(t)\ \left(m(y_t; t) - m\left(h^*(t); t\right)\right) \mathrm{d}t \\ &= \int_{t_0-\epsilon}^{t_0} \left(F_X^{-1}\right)'(t)\ \left(0 - m\left(h^*(t); t\right)\right) \mathrm{d}t > 0, \end{aligned}$$

since $m\left(h^*(t); t\right) < 0$ when $t \in (t_0 - \epsilon, t_0]$. It is a contradiction. Hence, $h^*(t_0) \geq t_0 - \widetilde{T}(t_0)$. We can derive a similar result if $\widetilde{T}(t) \geq t$ on $[t_0, t_0 + \epsilon)$. In short, we have

$$h^*(t_0) = y_{t_0} \begin{cases} = t_0 - \widetilde{T}(t_0), & \widetilde{T}(t_0) < t_0, \\ \geq t_0 - \widetilde{T}(t_0), & \widetilde{T}(t_0) \geq t_0, \end{cases} \tag{3.22}$$

for any $t_0 \in [0, 1]$.

Then, we check the sufficient condition. Assume that h^* is defined as in (3.22). Then, for any other feasible solution h, we compare the insurer's profit under h^* and h. From the structure of (3.20), we have

$$\begin{aligned} &V^{In}\left(h^*, q^*_{h^*}\right) - V^{In}\left(h, q^*_{h}\right) = \int_0^1 \left(F_X^{-1}\right)'(t)\,(t - \widetilde{T}(t))\mathbb{1}_{\{\widetilde{T}(t)<t\}}\mathrm{d}t \\ &- \int_0^1 \left(F_X^{-1}\right)'(t)\, h(t) \Big(\mathbb{1}_{\{h(t)\leq t-\widetilde{T}(t)\}}\mathbb{1}_{\{\widetilde{T}(t)<t\}}(t) + \mathbb{1}_{\{h(t)<t-\widetilde{T}(t)\}}\mathbb{1}_{\{\widetilde{T}(t)=t\}}(t) \\ &+ \mathbb{1}_{\{h(t)<t-\widetilde{T}(t)\}}\mathbb{1}_{\{\widetilde{T}(t)>t\}}(t)\Big)\mathrm{d}t \\ &= \int_0^1 \left(F_X^{-1}\right)'(t)\ \left(t - \widetilde{T}(t) - h(t)\mathbb{1}_{\{h(t)\leq t-\widetilde{T}(t))\}}\right)\mathbb{1}_{\{\widetilde{T}(t)<t\}}\mathrm{d}t \\ &- \int_0^1 \left(F_X^{-1}\right)'(t)\, h(t)\left(\mathbb{1}_{\{h(t)<t-\widetilde{T}(t)\}}\mathbb{1}_{\{\widetilde{T}(t)=t\}} + \mathbb{1}_{\{h(t)<t-\widetilde{T}(t)\}}\mathbb{1}_{\{\widetilde{T}(t)>t\}}\right)\mathrm{d}t, \end{aligned} \tag{3.23}$$

We analyze the terms separately. For the first integral in (3.23), we know that

$$t - \widetilde{T}(t) - h(t)\mathbb{1}_{\{h(t)\leq t-\widetilde{T}(t)\}} = \begin{cases} t - \widetilde{T}(t), & h(t) > t - \widetilde{T}(t) \\ t - \widetilde{T}(t) - h(t), & h(t) \leq t - \widetilde{T}(t) \end{cases}$$

which is always nonnegative. For the second integral, if $\widetilde{T}(t) = t$, then

$$-h(t)\mathbb{1}_{\{h(t)<t-\widetilde{T}(t)\}} = -h(t)\mathbb{1}_{\{h(t)<0\}} \geq 0.$$

If $\widetilde{T}(t) > t$, then $t - \widetilde{T}(t) < 0$, and thus

$$-h(t)\mathbb{1}_{\{h(t)<t-\widetilde{T}(t)\}} \geq 0.$$

Thus, both terms in (3.23) are nonnegative for all $t \in [0,1]$ which means the integrand is nonnegative pointwise. Therefore, the difference in profit is nonnegative:

$$V^{In}\left(h^*, q^*_{h^*}\right) \geq V^{In}\left(h, q^*_h\right).$$

Hence, h^* maximizes the insurer's profit functional. This completes the sufficiency proof.

Additionally, when h^* is given as in (3.22), we can characterize the structure of the critical sets. In particular, we observe:

$$\mathcal{A}_1\left(h^*\right) \cup \left(\mathcal{A}_2\left(h^*\right) \cap \mathcal{B}_1\right) = \emptyset \cup \{t : h^*(t) = t - \widetilde{T}(t) > 0\} = \{t : t - \widetilde{T}(t) > 0\},$$

$$\mathcal{A}_2\left(h^*\right) \cap \mathcal{B}_2 = \{t : h^*(t) = t - \widetilde{T}(t) = 0\},$$

and

$$\begin{aligned}\left(\mathcal{A}_2\left(h^*\right) \cap \mathcal{B}_3\right) \cup \mathcal{A}_3\left(h^*\right) &= \{t : h^*(t) = t - \widetilde{T}(t) < 0\} \cup \{t : h^*(t) > t - \widetilde{T}(t)\} \\ &= \{t : h^*(t) > t - \widetilde{T}(t) = 0\} \cup \{t : \widetilde{T}(t) > t\}\end{aligned}$$

Using these classifications, the derivative of the optimal indemnity-induced premium function $q^*_{h^*}$ given in (3.19) becomes

$$\left(q^*_{h^*}\right)'(t) = \begin{cases} 0, & \widetilde{T}(t) < t, \\ \phi_h(t) \in \left[0, \left(F_X^{-1}\right)'(t)\right], & t \in \{z : h^*(z) = z - \widetilde{T}(z) = 0\} \\ \left(F_X^{-1}\right)'(t) & t \in \{z : h^*(z) > z - \widetilde{T}(z) = 0\} \cup \{z : \widetilde{T}(z) > z\}. \end{cases}$$

By the relationship $\widetilde{g}^*(t) = t - h^*(t)$, the the optimal h^* leads to the optimal premium distortion function $\widetilde{g}^*$, given by

$$\widetilde{g}^*(t) = \begin{cases} \widetilde{T}(t), & \widetilde{T}(t) < t, \\ \in \left[0, \widetilde{T}(t)\right], & \widetilde{T}(t) \geq t. \end{cases} \tag{3.24}$$

The formula in (3.24) provides a pointwise characterization of the optimal premium structure. Since $\widetilde{g}^*$ is itself a distortion function, it must satisfy the properties outlined in Definition 4.2.1, namely $\widetilde{g}^*(0) = 0$, $\widetilde{g}^*(1) = 1$, and it is increasing and differentiable. Thus, in the case $\widetilde{T}(t) \geq t$, these constraints further refine the admissible values of $\left[\sup\left\{z < t : \widetilde{g}^*(z)\right\}, \widetilde{T}(t)\right]$. Therefore,

$$\widetilde{g}^*(t) = \begin{cases} \widetilde{T}(t), & \widetilde{T}(t) < t, \\ \in \left[\sup\left\{z < t : \widetilde{g}^*(z)\right\}, \widetilde{T}(t)\right], & \widetilde{T}(t) \geq t. \end{cases} \tag{3.25}$$

The optimal retention quantile q^* as a function of the distortion premium function g^* can be derived by

$$\left(q^*_{g^*}\right)'(t) = \begin{cases} 0, & \widetilde{T}(t) < t, \\ \phi_{g^*}(t) \in \left[0, \left(F_X^{-1}\right)'(t)\right], & t \in \{z : \widetilde{g}^*(z) = \widetilde{T}(z) = z\}, \\ \left(F_X^{-1}\right)'(t), & t \in \{z : \widetilde{g}^*(z) < \widetilde{T}(z) = z\} \cup \{z : \widetilde{T}(z) > z\}. \end{cases}$$

This completes the proof. □

The above theorem characterizes the Bowley-optimal distortion premium structure under the assumption that policyholders evaluate contracts using a distortion risk measure. It shows that the Bowley-optimal solution is not unique: any premium distortion function satisfying the condition in (3.14), together with the optimal indemnity choice defined by (3.15), yields the highest possible profit for the insurer. When $\widetilde{T}(t) < t$, we have $\widetilde{g}^*(t) = \widetilde{T}(t)$ and $\left(q^*_{g^*}\right)'(t) = 0$, which implies

$$g^*(t) = T(t).$$

In this case, any additional loss is fully transferred to the insurer, meaning the policyholder receives full insurance coverage. This situation corresponds to higher risk aversion on the part of the policyholder, and the optimal pricing distortion exactly reflects the policyholder's preferences. When $\widetilde{T}(t) > t$ (or equivalently $T(t) < t$), some flexibility is allowed in the choice of g^*, and we have $\left(q^*_{g^*}\right)'(t) = \left(F_X^{-1}\right)'(t)$. Any distortion function satisfying

$$\widetilde{g}^*(t) \in \left[\sup\left\{z < t : \widetilde{g}^*(z) < z\right\}, \widetilde{T}(t)\right]$$

is admissible, and this interval ensures the overall distortion function $\widetilde{g}^*$ (or equivalently g^*) remains increasing to satisfy feasibility. In this case, incremental losses are retained by the policyholder. In the borderline case where $\widetilde{T}(t) = t$, any distortion function that preserves monotonicity and any quantile that is uniformly continuous will result in the same insurer profit.

Furthermore, the maximum profit attainable by the insurer, consistent with the definition of Bowley optimality in Definition 3.2.2, is given by:

$$\begin{aligned} V^{In}\left(g^*, q^*_{g^*}\right) &= \int_0^1 \left(F_X^{-1}\right)'(t)\,(t - 1 + T(1-t))\; \mathbb{1}_{\{1-t<T(1-t)\}}\,\mathrm{d}t \\ &= \int_0^1 \left(F_X^{-1}\right)'(1-t)\,(T(t) - t)\; \mathbb{1}_{\{T(t)>t\}}\,\mathrm{d}t. \end{aligned} \tag{3.26}$$

It is observed that the Bowley-optimal profit given in (3.26) depends entirely on the policyholder's risk characteristics. When two policyholders share the same risk attitude, i.e., they use the same distortion function T, but differ in their underlying risk distributions, the one with a riskier distribution in the sense of the first stochastic order (i.e., a larger loss distribution $F_X(x)$ for any $x \in [0, M]$) generates a higher profit for the insurer. Conversely, for fixed risk distributions, a more risk-averse policyholder—reflected by a more concave distortion function T—also yields greater profit for the insurer.

The following result provides a characterization of the Bowley-optimal insurance indemnity function when the solution is expressed in terms of the quantile function.

Lemma 3.4.1. *The contract* $\left(g^*, q^*_{g^*}\right)$ *is optimal to problem* (3.13) *if and only if* $(g^*, I^*_{g^*})$ *is optimal to problem* (3.12), *where* $I^*_{g^*}(t) = t - q^*_{g^*}(F_X(t))$.

Proof. This is similar to the proof of Lemma 3.3.2. □

The result in Theorem 3.4.1 shows that the Bowley-optimal insurance contract $(g^*, q^*_{g^*})$, consists of an optimal pricing distortion g^* given in (3.14), and the corresponding optimal retention quantile $q^*_{g^*}$, given in (3.15). An important implication of this result is that the optimal indemnity, or equivalently, the optimal retention quantile, depends on both the risk distribution F_X and the policyholder's risk attitude T. In contrast, the optimal distortion function g^* depends only on the policyholder's risk attitude T. This underscores that a policyholder's perception and response to risk play a pivotal role in determining the structure of the optimal premium principle—regardless of the underlying loss distribution. Therefore, incorporating a detailed discussion of the policyholder's risk attitude is crucial.

In what follows, we present several examples that illustrate how the optimal insurance contract evolves under different risk attitudes, as captured by various forms of the distortion function T. These examples highlight the flexibility and generality of our approach, demonstrating its ability to accommodate a broader range of risk preferences compared to existing frameworks in the literature.

3.4.1 Extension of the Work in Cheung et al. (2019)

The cases of policyholders who exhibit strict risk aversion or who act as Value-at-Risk minimizers are thoroughly analyzed in Cheung et al. (2019). Our results in this paper fully encompass and extend those findings. Specifically, the framework we develop not only recovers the optimal contracts identified in Cheung et al. (2019) for these two types of policyholders but also generalizes the analysis to a broader class of distortion-based risk attitudes.

Example 3.4.1 (Concave distortion)**.** *Suppose the policyholder's distortion function $T(t)$ is concave on $[0,1]$. Then the conjugate distortion function $\widetilde{T}(t) = 1 - T(1-t)$ is convex. In particular, we have $\widetilde{T}(t) \le t$ for all $t \in (0,1)$. As a result, a Bowley-optimal conjugate pricing distortion function is given by*

$$\widetilde{g}^*(t) = \widetilde{T}(t) \quad \textit{for all } t \in (0,1),$$

and hence the optimal distortion function for pricing is

$$g^*(t) = T(t) \quad \textit{for all } t \in [0,1].$$

An optimal retention function is flat:

$$\left(q^*_{g^*}\right)'(t) \equiv 0,$$

which implies that the optimal indemnity function satisfies

$$I^*_{g^*}(x) = x, \quad \textit{for } x \in [0, M].$$

That is, the insurer offers full insurance under the Bowley-optimal contract.

When a concave distortion function is used to define the distortion risk measure—reflecting that the policyholder is strictly risk-averse—our result shows that full insurance is optimal in the sense of Bowley optimality for such individuals. We observe that, in this case, the policyholder is indifferent to participating. Indeed,

$$\rho(X - I^*_{g^*}(X) + \pi_{g^*}(I^*_{g^*}(X))) = \pi_{g^*}(X) = \int X \, \mathrm{d}g^* \circ \mathbb{P} = \int X \, \mathrm{d}T \circ \mathbb{P} = \rho(X).$$

On the other hand, the insurer's profit in (3.26) becomes

$$V^{In}\left(g^*, q^*_{g^*}\right) = \int_0^1 \left(F_X^{-1}\right)'(1-t)\,(T(t) - t)\, \mathrm{d}t.$$

That is, the insurer's profit increases as the policyholder becomes more risk-averse — that is, as the function T becomes more concave.

Example 3.4.2 (Value-at-Risk)**.** *Suppose the policyholder minimizes a Value-at-Risk measure at confidence level* $\alpha \in (0, 1)$*. This corresponds to the indicator distortion function:*

$$T(t) = \mathbb{1}_{(1-\alpha,1]}(t).$$

Then the conjugate distortion is given by $\widetilde{T}(t) = \mathbb{1}_{[\alpha,1]}(t)$*. Hence, the Bowley-optimal conjugate distortion for pricing becomes:*

$$\widetilde{g}^*(t) = \widetilde{T}(t) = 0, \quad \textit{for } t < \alpha,$$

and using the definition $g^*(t) = 1 - \widetilde{g}^*(1-t)$*, we get*

$$g^*(t) = 1, \quad \textit{for } t > 1 - \alpha.$$

This implies that an optimal marginal retention function satisfies

$$\left(q^*_{g^*}\right)'(t) = \begin{cases} 0, & t < \alpha. \\ \left(F_X^{-1}\right)'(t), & t \geq \alpha \end{cases}$$

Therefore, the insurer provides insurance according to

$$\widetilde{I}^*_{g^*}(x) = \begin{cases} x, & x < F_X^{-1}(\alpha), \\ F_X^{-1}(\alpha), & x \geq F_X^{-1}(\alpha). \end{cases}$$

This result illustrates that, under Value-at-Risk minimization, the Bowley-optimal contract takes the form of a coverage limit contract, providing full insurance on the lower tail of the loss distribution (i.e., for losses below the $\alpha-$quantile) while leaving the upper tail uninsured.

The results presented in Examples 3.4.1 and 3.4.2 are also established in Cheung et al. (2019), where the policyholder is assumed to evaluate risk using either a concave distortion function or by minimizing Value-at-Risk. In both cases, the resulting premium structure coincides with that derived in our setting. In particular, when the policyholder's preference is modeled by a concave distortion function, the outcome in Example 3.4.1 aligns with Cheung et al. (2019, Theorem 3.1, case $\gamma = 0$). Similarly, the result in Example 3.4.2, where the distortion corresponds to a Value-at-Risk measure, is consistent with Cheung et al. (2019, Theorem 3.5, case $\gamma = 0$).

The following example provides another solution to the Bowley optimization problem in a setting where the distortion function does not fall into the special cases commonly studied in the literature. In particular, we assume that the policyholder places greater weight on the tail of the risk distribution, reflecting heightened concern for extreme losses.

Example 3.4.3 (ISSD). *Suppose the policyholder evaluates risk using a distortion risk measure with the following distortion function:*

$$T(t) = \frac{t^\theta}{(t^\theta + (1-t)^\theta)^{\frac{1}{\theta}}},$$

where $\theta = 0.55$*. It is observed that,* $\widetilde{T}(t) < t$ *if* $t > t_0 = 0.6931$*, and* $\widetilde{T}(t) \geq t$ *if* $t \leq t_0$*. Therefore, the conjugate of the optimal pricing distortion satisfies*

$$\widetilde{g}^*(t) = \widetilde{T}(t), \quad \textit{for } t > t_0.$$

The quantile of the marginal retention, given the premium is determined by g^**,*

$$\left(q^*_{g^*}\right)'(t) = \begin{cases} \left(F_X^{-1}\right)'(t), & t \leq t_0 \\ 0, & t > t_0. \end{cases}$$

By Lemma 3.3.2, the optimal indemnity function is

$$I^*_{g^*}(t) = t - q^*_{g^*}(F_X(t)) = \begin{cases} 0, & x \leq F_X^{-1}(t_0). \\ x - F_X^{-1}(t_0), & x > F_X^{-1}(t_0). \end{cases}$$

This example demonstrates that when the policyholder is more concerned with extreme losses, the deductible contract becomes Bowley-optimal. The deductible level in this case depends on both the risk distribution F_X and the policyholder's attitude toward risk, as reflected by the turning point of ISSD.

3.5 Conclusion

This chapter focuses on Stackelberg equilibria (also known as Bowley optima) in monopolistic, centralized, sequential-move insurance markets, where policyholders exhibit varying risk preferences and purchase different types of insurance contracts, as introduced in Chapter 2. Specifically, we depart from the assumption in Chapter 2 that premiums are based on the expected value of the indemnity. Instead, we assume that premiums are determined using a Choquet integral, which distorts probabilities to assign different weights across the distribution of risk. To model policyholder preferences, we adopt the distortion risk measure, under which each policyholder evaluates risk by applying a personalized distortion function that assigns different weights to outcomes. This framework is consistent with the setting considered in Cheung et al. (2019). In contrast to Chapter 2, where the indemnity function is restricted to specific forms—such as deductible or coinsurance contracts—we relax this constraint by allowing a broader class of indemnity functions. This is made possible by the flexibility of the distortion risk measure, which serves as the core tool for modeling decision-making in this framework.

To achieve the Stackelberg equilibrium in our setting, we employ the Quantile Approach. This method requires that the underlying risk variable has a strictly increasing cumulative distribution function. The equilibrium contract is derived through a two-step optimization procedure. In the first step, we determine the optimal indemnity function for a given pricing distortion. We show that a layered insurance contract emerges as optimal. Specifically, if the distortion function used for pricing is lower than that used to model the policyholder's risk attitude, then any marginal increase in loss will be fully covered by the insurer. In contrast, if the pricing distortion is higher than the policyholder's distortion function, then any marginal increase in loss will be entirely retained by the policyholder, implying that no insurance coverage is provided in that region.

In the second step, we determine the optimal pricing distortion, given that the policyholder selects the optimal insurance contract identified in the first step. We find that the structure of the Bowley-optimal premium is highly sensitive to the policyholder's risk preferences. For instance, if the policyholder is strictly risk averse—represented by a concave distortion function in the distortion risk measure—the optimal pricing distortion coincides with the policyholder's own distortion function. In contrast, if the policyholder evaluates risk using Value-at-Risk, the Bowley-optimal pricing distortion becomes an indicator function that equals one above a certain threshold, relating to the chosen VaR confidence level. These two special cases are also discussed in Cheung et al. (2019). However, our framework is more general. We can accommodate a broader class of policyholder preferences, as long as they are represented by a distortion risk measure. This includes, for example, inverse-S shaped distortion functions, which reflect greater sensitivity to extreme outcomes and are consistent with the empirical behavior of individuals under risk.

Chapter 4

Hidden Risk Attitudes

4.1 Introduction

In this chapter, we examine a monopolistic insurance market under hidden information. The issue of imperfect information in insurance markets was first addressed in Rothschild and Stiglitz (1976), where the authors demonstrated that the presence or absence of perfect information significantly influences market equilibrium. In particular, under perfect information in a competitive market with multiple insurers and zero-profit conditions, full insurance is typically provided to policyholders. However, when information is imperfect, a separating equilibrium may arise—meaning that individuals with different risk types are offered different insurance contracts tailored to their characteristics. This type of equilibrium structure under imperfect information is also studied in alternative insurance settings by Young and Browne (2000), where policyholders are assumed to make decisions based on the Dual Utility (DU) theory of Yaari (1987). The subsequent work Stiglitz (1977) examines imperfect information in a monopolistic insurance market, where a single insurer operates without price competition and is able to earn strictly positive profits. In such a market, if information is perfect, the insurer offers full insurance and extracts the entire surplus, leaving the policyholder indifferent between participating and not participating. Under imperfect information, however, the equilibrium differs from the full-information case and typically takes the form of a separating equilibrium.

Building on the findings of Stiglitz (1977), a large body of literature has explored how imperfect information affects monopolistic insurance markets. In both Rothschild and Stiglitz (1976) and Stiglitz (1977), the analysis is restricted to two types of insurance policyholders. The work of Chade and Schlee (2012) extends this framework to a continuum of policyholder types. Many key properties observed in the two-type case are preserved in the continuum setting. For instance, full insurance is provided to the highest-risk type, while the lowest-risk type receives a contract that leaves them indifferent between participating and not participating. Moreover, the extension to a continuum of types offers greater flexibility for analyzing the structure of equilibrium contracts. Chade and Schlee (2012) shows that, under mild conditions and assuming a smooth distribution of types, the premium function exhibits a backward-S shape with respect to the amount of coverage. In a related study, Chade and Schlee (2020) investigates how the structure of costs affects the equilibrium contract. Specifically, when a provision cost is introduced—rather than modeling costs as the expected indemnity—the market may exhibit a pooling equilibrium, and in some cases, there may be no gains from trade at all. Gershkov et al. (2023) presents a significant extension in the study of asymmetric information in monopolistic insurance markets. The authors move beyond the binary-distributed risk assumption by allowing the loss random variable to follow a continuous distribution. Furthermore, they analyze this framework under the assumption that policyholders are Dual Utility (DU) maximizers. A separating equilibrium is derived under certain conditions, and commonly used indemnity structures—such as deductible and policy limit contracts—are shown to emerge as optimal under specific circumstances.

Another key insight from Stiglitz (1977) is that in a monopoly setting, not only the distribution of risk types but also the risk preferences of policyholders influence the market equilibrium. In contrast, in a competitive market, only the distribution of risk types affects the equilibrium outcome. Motivated by this observation, Landsberger and Meilijson (1994) study how imperfect information about risk attitudes affects the equilibrium when policyholders are expected-utility maximizers sharing the same risk distribution. Boonen and Zhang (2021) examine this effect under the framework of distortion risk measures.

In this work, we extend these results to a monopolistic market with a continuum of policyholder types. Specifically, the market consists of a risk-neutral, profit-maximizing insurer and policyholders with Yaari's dual utility (DU) preferences. Each policyholder faces an insurable loss with a known continuous distribution, but the insurer cannot observe the policyholder's risk attitude. Instead, the risk attitude—captured by a distortion function—is drawn from a continuum of types. We characterize the optimal (profit-maximizing), incentive-compatible, and individually rational menu of insurance contracts. Under suitable regularity conditions, we show that the optimal menu can be characterized via the marginal loss retention for each type, resulting in a layered deductible contract. Each layer in this structure is determined by the policyholder's risk type. Notably, such layered indemnity structures are frequently observed in real-world insurance markets.

In addition, and following Chade and Schlee (2020), we examine the effect of insurance provision costs on equilibria in our monopoly market with DU policyholder. Specifically, we assume that in addition to the expected insurance indemnification payment, the insurer incurs "friction costs", such as claims processing costs, administrative expenses, actuarial loading, etc. Moreover, we characterize incentive-efficient menus of contracts in the context of an arbitrary type space. We show that individually rational and incentive compatible contracts that are Pareto optimal can be achieved by maximizing a social welfare function that accounts for hidden types. While it is difficult to solve such a problem in the general case, we are able to provide a crisp characterization of solutions under a few assumptions.

The rest of this chapter is structured as follows. Section 4.2 presents our market model and formulates the problem. Section 4.3 characterizes the set of individually rational and incentive compatible menus of contracts. Section 4.4 characterizes optimal menus and separating equilibria, and examines properties thereof. In Section 4.5 we study the effects of friction costs on the structure of equilibria, and in Section 4.6 we examine Pareto-optimal menus of contracts. Finally, Section 4.7 concludes.

4.2 The Insurance Market

4.2.1 Setup

We consider a one-period monopoly insurance market in which an policyholder is subject to an insurable loss and seeks insurance coverage from a monopolistic insurer, in exchange for a premium payment. We assume that the loss random variable X is an element of $B(\Sigma)$, the space of bounded and Σ-measurable real-valued functions on a given probability space $(S, \Sigma, \mathbb{P})$, with range $[0, M]$, for some $M < +\infty$, and cumulative distribution function $F_X(x) := \mathbb{P}\left(X \leq x\right)$.

The monopolistic insurer is assumed to be risk-neutral, who aims to maximize profit. The policyholder behaves according to the Dual Utility (DU) framework of Yaari (1987), in which risk aversion is entirely captured by the policyholder's probability weighting function, also referred to as a distortion function.

Definition 4.2.1. *A function $g : [0, 1] \to [0, 1]$ is called a distortion function if it satisfies the following conditions:*

1. $g(0) = 0$ *and* $g(1) = 1$*;*
2. g *is increasing on* $[0, 1]$.

Let W be a bounded random variable. The dual utility of W is the Choquet integral of W with respect to the distorted probability measure $g \circ \mathbb{P}$, given by

$$\begin{aligned} DU(W) = \int W \, \mathrm{d}g \circ \mathbb{P} &:= \int_0^{+\infty} g(1 - \mathbb{P}(W \leq x)) \, \mathrm{d}x + \int_{-\infty}^{0} [g(1 - \mathbb{P}(W \leq x)) - 1] \, \mathrm{d}x \\ &= - \int_{\mathbb{R}} x \, \mathrm{d}g \, (1 - F_W(x)) \end{aligned} \tag{4.1}$$

By basic properties of the Choquet integral (e.g., Denneberg (1994)), DU is monotone, positively homogeneous, and translation-invariant. In particular,

$$DU(W + c) = DU(W) + c,$$

for any constant $c \in \mathbb{R}$.

The policyholder can purchase insurance coverage $I(X)$ against X, satisfying $0 \leq I(X) \leq X$, from the insurer at a premium p. The policyholder's resulting end-of-period wealth is given by

$$-p - X + I(X) = -p - R(X),$$

where $R(X) = X - I(X) \geq 0$ is the part of the loss that is not covered by the insurer, and hence retained by the policyholder. We refer to $R(X)$ as the retained loss. The resulting utility of the policyholder is then given by

$$\begin{aligned} DU(-p - R(X)) &= -p + DU(-R(X)) \\ &= -p + \int_{-\infty}^{0} [g(1 - \mathbb{P}(-R(X) \leq x)) - 1] \, \mathrm{d}x \\ &= -p + \int_{-\infty}^{0} [g(\mathbb{P}(R(X) \leq -x)) - 1] \, \mathrm{d}x \\ &= -p - \int_0^{+\infty} [1 - g(\mathbb{P}(R(X) \leq x))] \, \mathrm{d}x. \\ &= -p - \int_0^{M} [1 - g(\mathbb{P}(R(X) \leq x))] \, \mathrm{d}x. \end{aligned}$$

We use the convention that $U(X, 0)$ denotes the utility of wealth in the absence of insurance. That is,

$$U(X, 0) = DU(-X) = - \int_0^M [1 - g(F_X(x))] \, \mathrm{d}x.$$

By a classical result of Quiggin (1993) and Chateauneuf and Cohen (1994), the functional DU in (4.1) is weakly risk averse if and only if $g(s) \leq s$, for all $s \in [0,1]$. See Ghossoub and He (2021) for more about risk aversion in Rank-Dependent Utility Theory and Dual Utility Theory. We make this assumption all throughout.

Assumption 4.2.1 (Risk Aversion)**.** *The distortion function g is differentiable and weakly risk averse, that is, $g(s) \leq s$ for all $s \in [0,1]$.*

Following Quiggin (1993) and Ghossoub and He (2021), the quantity

$$\Delta(W) := \mathbb{E}[W] - DU(W)$$

is called the *risk premium* associated with the random variable W, and weak risk aversion was shown to be equivalent to a nonnegative risk premium. Hence, an implication of weak risk aversion is that

$$DU(X) \leq DU(\mathbb{E}[X]) = \mathbb{E}[X],$$

and we therefore obtain the following immediate result.

Proposition 4.2.1. *If g satisfies Assumption 4.2.1, then $U(X,0) \leq -\mathbb{E}[X]$.*

The insurer is a risk-neutral expected-utility maximizer. After receiving the premium payment p from the policyholder in exchange for a promised indemnity payment of $I(X)$, the end-of-period wealth of the insurer is given by

$$p - I(X) = p - (X - R(X)),$$

resulting in an expected utility of

$$p - \mathbb{E}[X - R(X)],$$

which is also the insurer's expected profit from the contractual agreement.

4.2.2 Hidden Risk Attitudes

In our setting, the monopolistic insurer is able to observe the distribution of the loss random variable X, but not the risk attitude of the policyholder. Specifically, we assume that the policyholder's type θ is drawn from a continuum of types $\Theta = \left[\underline{\theta}, \overline{\theta}\right] \subset \mathbb{R}_+$, and it determines the distortion function g_θ used in the policyholder's utility function

$$DU_\theta(\cdot) = \int \cdot \, \mathrm{d}\, g_\theta \circ \mathbb{P}.$$

We assume that θ has a density function $f(\theta)$ with a corresponding cumulative distribution function $F(\theta)$ over Θ, and survival function $\overline{F}(\theta) = 1 - F(\theta)$. Hence, the utility of wealth for a policyholder of type $\theta \in \Theta$ in the absence of insurance is given by

$$U_\theta(X, 0) := DU_\theta(-X) = -\int_0^M \left[1 - g_\theta\left(F_X\left(x\right)\right)\right] \mathrm{d}x. \tag{4.2}$$

Furthermore, we assume that the type space Θ is ordered, with larger values of θ indicating a more risk-averse attitude. That is, by Ghossoub and He (2021), if $\theta_1 \leq \theta_2$, then $g_{\theta_1}(s) \geq g_{\theta_2}(s)$, for all $s \in [0,1]$. Specifically, we make the following additional assumption about the collection $\{g_\theta\}_{\theta \in \Theta}$.

Assumption 4.2.2. *The family $\{g_\theta\}_{\theta\in\Theta}$ of distortion functions satisfies the following:*

1. **Continuity in Types:** *$\{g_\theta\}_{\theta\in\Theta}$ is uniformly Lipschitz continuous in θ, with common Lipschitz coefficient $c < +\infty$.*
2. **Ordered Type Space:** *$\frac{\partial g_\theta(s)}{\partial \theta} \leq 0$ for $s \in (0, 1)$.*

Examples that satisfy such conditions include the quadratic distortion function $g_\theta(s) = \theta s^2 + (1-\theta)s$ for $\theta \in [0,1]$, and the class $g_\theta(s) = \frac{s}{1+\theta(1-s)}$ for $\theta \in [0,1]$, for instance.

To mitigate the potential loss, a policyholder of type θ will purchase an insurance contract at a cost of p_θ. The retained loss is $R_\theta(X) = X - I_\theta(X)$, where I_θ is the indemnity schedule offered by the insurer. The policyholder's resulting end-of-period utility of wealth is given by

$$U_\theta\left(R_\theta, p_\theta\right) = -p_\theta - \int_0^M \left[1 - g_\theta\left(\mathbb{P}\left[R_\theta(X) \leq x\right]\right)\right] \mathrm{d}x. \tag{4.3}$$

It is customary in the literature to restrict the set of feasible indemnity schedules offered in the market to those functions $I : [0, M] \to [0, M]$ that satisfy $0 \leq I(X) \leq X$, and are such that both $I(X)$ and $R(X) = X - I(X)$ are non-decreasing functions of X. This restriction, often referred to as the *no-sabotage condition* (e.g., Carlier and Dana (2003, 2005)) is meant to prevent *ex post* moral hazard that might result from the policyholder's misreporting of the true vaue of the realized loss X (e.g., Huberman et al. (1983)).

Note that any indemnity function $I : [0, M] \to [0, M]$ that satisfies the *no-sabotage* condition is a non-decreasing and 1-Lipschitz function (e.g., (Denneberg, 1994, Proposition 4.5)), and consequently so is the associated retention function R. Henceforth, we make the assumption that the market only offers indemnity functions that satisfy the *no-sabotage* condition. Hence, the set of *ex ante* admissible indemnity functions is given by

$$\mathcal{I}_L := \Big\{ I : [0, M] \to [0, M] \;\Big|\; I(0) = 0,\ 0 \leq I(x_1) - I(x_2) \leq x_1 - x_2, \forall\, 0 \leq x_2 \leq x_1 \leq M \Big\}. \tag{4.4}$$

Remark 5. *Let $C[0, M]$ denote the set of all continuous functions on $[0, M]$ (and thus bounded), equipped with the supnorm $\|\cdot\|_{sup}$. Note that $\mathcal{I}_L$ is a uniformly bounded subset of $C[0, M]$ consisting of Lipschitz-continuous functions $[0, M] \to [0, M]$, with common Lipschitz constant $K = 1$. Therefore, $\mathcal{I}_L$ is equicontinuous, and hence compact by the Arzelà-Ascoli Theorem (Dunford and Schwartz, 1958, Theorem IV.6.7).*

Since Lipschitz-continuous functions are absolutely continuous, we can equivalently restrict the set of feasible retention functions to the set

$$\mathcal{R}_L := \Big\{ R : [0, M] \to [0, M] \;\Big|\; R(0) = 0, 0 \leq \frac{\partial R(x)}{\partial x} \leq 1, \forall\, x \in [0, M], \text{ a.e.} \Big\}. \tag{4.5}$$

Hence, a contract consists of a feasible indemnity function (equivalently, a feasible retention function) and an associated premium.

Definition 4.2.2. *A contract is a pair $(R, p) \in \mathcal{R}_L \times \mathbb{R}_+$.*

4.2.3 Problem Formulation

For the risk-neutral insurer, the profit from providing a contract (R_θ, p_θ) to a policyholder of type θ is

$$\pi\left(R_\theta, p_\theta\right) := p_\theta - \mathbb{E}\left[X - R_\theta(X)\right] = p_\theta - \mathbb{E}[X] + \int_0^M \left[1 - F_X\left(x\right)\right] \frac{\partial R_\theta(x)}{\partial x} \mathrm{d}x, \tag{4.6}$$

and the expected profit resulting from offering a menu of contracts $(R_\theta, p_\theta)_{\theta \in \Theta}$ is

$$V\left((R_\theta, p_\theta)_{\theta\in\Theta}\right) := \int_\Theta \pi\left(R_\theta, p_\theta\right) \mathrm{d}F(\theta). \tag{4.7}$$

To incentivize the policyholder to participate in the market, we require *individual rationality*. Additionally, to ensure that policyholder reveals their true type θ and selects contracts that match their preferences, the menu of insurance contracts must also satisfy *incentive compatibility*.

Definition 4.2.3. *A menu of contracts* $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{R}_L^\Theta \times \mathbb{R}_+^\Theta$ *is said to be:*

1. *Individually rational (IR) if* $U_\theta(R_\theta, p_\theta) \geq U_\theta(X, 0)$*, for all* $\theta \in \Theta$*.*
2. *Incentive compatible (IC) if* $U_\theta(R_\theta, p_\theta) \geq U_\theta(R_{\theta'}, p_{\theta'})$*, for all* $\theta, \theta' \in \Theta$*.*

Let $\mathcal{IR}$ and $\mathcal{IC}$ denote the set of all IR and IC menus, and note that both sets are convex, by linearity of U_θ in (4.3). Consequently $\mathcal{IR} \cap \mathcal{IC}$ is a convex set.

As the monopolist in the insurance market, the insurer aims to offer a menu that maximizes expected profit. Our goal is therefore to identify the set of all profit-maximizing IR and IC menus of contracts $(R_\theta^*, p_\theta^*)_{\theta\in\Theta}$. That is, we wish to solve the following problem.

Problem 1.

$$(R_\theta^*, p_\theta^*)_{\theta\in\Theta} \in \underset{(R_\theta,p_\theta)_{\theta\in\Theta}\in\mathcal{IR}\cap\mathcal{IC}}{\arg\max} V\left((R_\theta, p_\theta)_{\theta\in\Theta}\right).$$

4.3 Characterization of IR and IC Menus

In this section, we present preliminary results about individually rational and incentive compatible menus. These results provide an intuitive understanding of how the contract influences the insurer's expected profit and what the optimal menu could look like.

We first note an immediate implication of the definition of individual rationality.

Proposition 4.3.1. *A menu of contracts* $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IR}$ *if and only if for each* $\theta \in \Theta$*,* $R_\theta \in \mathcal{R}_L$ *and*

$$p_\theta \leq \int_0^M [1 - g_\theta(F_X(x))] \left(1 - \frac{\partial R_\theta(x)}{\partial x}\right) \mathrm{d}x. \tag{4.8}$$

Proof. Consider a contract $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{R}_L^\Theta \times \mathbb{R}_+^\Theta$. If $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IR}$, then for each $\theta \in \Theta$,

$$\begin{aligned} -\int_0^M [1 - g_\theta(F_X(x))] \,\mathrm{d}x = U_\theta(X, 0) &\leq U_\theta(R_\theta, p_\theta) \\ &= -p_\theta - \int_0^M [1 - g_\theta(F_X(x))] \frac{\partial R_\theta(x)}{\partial x} \mathrm{d}x, \end{aligned}$$

implying that

$$p_\theta \leq \int_0^M [1 - g_\theta(F_X(x))] \left(1 - \frac{\partial R_\theta(x)}{\partial x}\right) \mathrm{d}x.$$

The converse follows similarly. ∎

The next result provides necessary conditions for incentive compatibility. It offers a clear characterization of the premium structure and the insurer's profit, when the policyholder truthfully reveals their information.

Proposition 4.3.2. *If $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IC}$, then the following two properties hold:*

1. *For any $\theta \in \Theta$, the premium p_θ is of the form*

$$p_\theta = p_{\underline{\theta}} + \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X\left(x\right)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\mathrm{d}x \; - \; \int_{\underline{\theta}}^{\theta}\int_0^M \frac{\partial g_s\left(F_X\left(x\right)\right)}{\partial s}\frac{\partial R_s(x)}{\partial x}\mathrm{d}x\mathrm{d}s - \int_0^M \left[1 - g_\theta\left(F_X\left(x\right)\right)\right] \frac{\partial R_\theta(x)}{\partial x}\mathrm{d}x, \tag{4.9}$$

   where $p_{\underline{\theta}} \in \mathbb{R}_+$ is arbitrary.

2. *The insurer's expected profit is given by*

$$\begin{aligned} V\left((R_\theta, p_\theta)_{\theta\in\Theta}\right) = p_{\underline{\theta}} + \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X\left(x\right)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\mathrm{d}x \\ - \mathbb{E}[X] - \int_{\underline{\theta}}^{\overline{\theta}} \left(\int_0^M J_\theta(x)\frac{\partial R_\theta(x)}{\partial x}\mathrm{d}x\right) f(\theta)\mathrm{d}\theta, \end{aligned} \tag{4.10}$$

   where

$$J_\theta(x) := F_X(x) - g_\theta\left(F_X\left(x\right)\right) + \frac{\overline{F}(\theta)}{f(\theta)}\frac{\partial g_\theta\left(F_X\left(x\right)\right)}{\partial \theta}. \tag{4.11}$$

Proof. The utility for policyholder of type θ with contract (R_θ, p_θ) in (4.3) can be written as

$$\begin{aligned} U_\theta\left(R_\theta, p_\theta\right) &= -p_\theta - \int_0^{R_\theta(M)} 1 - g_\theta\left(\mathbb{P}\left[R_\theta(X) \le x\right]\right)\mathrm{d}x \\ &= -p_\theta - \int_0^{R_\theta(M)} 1 - g_\theta\left(F_X\left(R_\theta^{-1}(x)\right)\right)\mathrm{d}x \\ &= -p_\theta - \int_0^{M} 1 - g_\theta\left(F_X\left(x\right)\right)\mathrm{d}R_\theta(x) \\ &= -p_\theta - \int_0^{M} \left[1 - g_\theta\left(F_X\left(x\right)\right)\right]\frac{\partial R_\theta(x)}{\partial x}\mathrm{d}x, \end{aligned} \tag{4.12}$$

where $R_\theta^{-1}(x) = \sup\{z : R_\theta(z) \leq x\}$. Similarly, for the contract $(R_{\theta'}, p_{\theta'})$, we have,

$$U_\theta\left(R_{\theta'}, p_{\theta'}\right) = -p_{\theta'} - \int_0^M \left[1 - g_\theta\left(F_X\left(x\right)\right)\right] \frac{\partial R_{\theta'}(x)}{\partial x} \mathrm{d}x.$$

Note that $U_\theta\left(R_{\theta'}, p_{\theta'}\right)$ is Lipschitz continuous in θ, and hence absolutely continuous in θ, since

$$\left|\frac{\partial U_\theta\left(R_{\theta'}, p_{\theta'}\right)}{\partial \theta}\right| \leq \int_0^M \left|\frac{\partial g_\theta\left(F_X\left(x\right)\right)}{\partial \theta} \frac{\partial R_{\theta'}(x)}{\partial x}\right| \mathrm{d}x \leq \int_0^M \left|\frac{\partial g_\theta\left(F_X\left(x\right)\right)}{\partial \theta}\right| \mathrm{d}x \leq c\, M < +\infty,$$

where the second inequality holds since $R_\theta \in \mathcal{R}_L$, and the third inequality holds by Assumption 4.2.2(1).

Then by the envelope theorem (e.g., Milgrom and Segal (2002, Theorem 2)), if $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IC}$, then for any $\theta \in \Theta$,

$$\begin{aligned} U_\theta\left(R_\theta, p_\theta\right) &= U_{\underline{\theta}}\left(R_{\underline{\theta}}, p_{\underline{\theta}}\right) + \int_{\underline{\theta}}^{\theta} \left.\frac{\partial U_{s'}(R_s, p_s)}{\partial s'}\right|_{s'=s} \mathrm{d}s \\ &= -p_{\underline{\theta}} - \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X\left(x\right)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x} \mathrm{d}x + \int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X\left(x\right)\right)}{\partial s} \frac{\partial R_s(x)}{\partial x} \mathrm{d}x\, \mathrm{d}s. \end{aligned} \tag{4.13}$$

Equating (4.12) and (4.13), it follows that

$$\begin{aligned} p_\theta = p_{\underline{\theta}} &+ \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X\left(x\right)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x} \mathrm{d}x - \int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X\left(x\right)\right)}{\partial s} \frac{\partial R_s(x)}{\partial x} \mathrm{d}x \mathrm{d}s \\ &- \int_0^M \left[1 - g_\theta\left(F_X\left(x\right)\right)\right] \frac{\partial R_\theta(x)}{\partial x} \mathrm{d}x. \end{aligned}$$

Substituting the above expression for the premium p_θ into (4.6), the insurer's expected profit given in (4.7) becomes

$$\begin{aligned} &V\left((R_\theta, p_\theta)_{\theta\in\Theta}\right) = \int_{\underline{\theta}}^{\overline{\theta}} \pi\left(R_\theta, p_\theta\right) f(\theta) \mathrm{d}\theta \\ &= p_{\underline{\theta}} + \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X\left(x\right)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x} \mathrm{d}x - \mathbb{E}[X] \end{aligned}$$

$$
\begin{aligned}
&- \int_{\underline{\theta}}^{\overline{\theta}} \left(\int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s} \frac{\partial R_s(x)}{\partial x} \mathrm{d}x \mathrm{d}s + \int_0^M \left[F_X(x) - g_\theta\left(F_X(x)\right)\right] \frac{\partial R_\theta(x)}{\partial x} \mathrm{d}x \right) f(\theta) \mathrm{d}\theta \\
&= p_{\underline{\theta}} + \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x} \mathrm{d}x - \mathbb{E}[X] - \int_{\underline{\theta}}^{\overline{\theta}} \left(\int_0^M \left[F_X(x) - g_\theta\left(F_X(x)\right)\right] \frac{\partial R_\theta(x)}{\partial x} \mathrm{d}x \right) f(\theta) \mathrm{d}\theta \\
&+ \left(\int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s} \frac{\partial R_s(x)}{\partial x} \mathrm{d}x \mathrm{d}s \overline{F}(\theta) \right) \Bigg|_{\theta=\underline{\theta}}^{\theta=\overline{\theta}} - \int_{\underline{\theta}}^{\overline{\theta}} \overline{F}(\theta) \int_0^M \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta} \frac{\partial R_\theta(x)}{\partial x} \mathrm{d}x \mathrm{d}\theta \\
&= p_{\underline{\theta}} + \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x} \mathrm{d}x - \mathbb{E}[X] \\
&- \int_{\underline{\theta}}^{\overline{\theta}} \left(\int_0^M \frac{\partial R_\theta(x)}{\partial x} \left(F_X(x) - g_\theta\left(F_X(x)\right) + \frac{\overline{F}(\theta)}{f(\theta)} \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta} \right) \mathrm{d}x \right) f(\theta) \mathrm{d}\theta \\
&= p_{\underline{\theta}} + \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x} \mathrm{d}x - \mathbb{E}[X] - \int_{\underline{\theta}}^{\overline{\theta}} \left(\int_0^M \frac{\partial R_\theta(x)}{\partial x} J_\theta(x) \mathrm{d}x \right) f(\theta) \mathrm{d}\theta,
\end{aligned}
$$

where $J_\theta(x) := F_X(x) - g_\theta\left(F_X(x)\right) + \frac{\overline{F}(\theta)}{f(\theta)} \frac{\partial g_\theta(F_X(x))}{\partial \theta}$. □

An important implication of (4.9) is that if $(R_\theta, p_\theta)_{\theta \in \Theta} \in \mathcal{IC}$, then for a given type $\theta \in \Theta$, the premium p_θ is fully determined by the choice of $p_{\underline{\theta}}$, the premium payment of the lowest type (the least risk averse), and the contracts R_s, for all types $s \leq \theta$. Such reasonable type-dependent nonlinear pricing schedules are also examined in Stiglitz (1977) and subsequently applied to other screening problems, as seen in Chade and Schlee (2012), Boonen and Zhang (2021), and Gershkov et al. (2023), for instance. The proof of Proposition 4.3.2 reveals that the representation of p_θ given in (4.9) depends solely on the incentive compatibility condition. It conveys the policyholders' true information—reflected through their choice of contract and corresponding utility—to the insurer. The insurer then reassesses their potential expected profit from such transactions based on this information. Further, the profit expression in (4.10) reveals that, for any given θ and x, the profit is a monotonic function of the marginal retention function $\frac{\partial R_\theta(x)}{\partial x}$, and the optimal value can then be identified by analyzing the sign of $J_\theta(x)$. A detailed analysis of this will be provided in the next section.

When policyholders exhibit different attitudes toward risk, their demand for insurance varies accordingly. A highly risk-averse individual is more inclined to purchase a contract that covers a larger portion of potential random losses, thereby retaining less risk than a less risk-averse individual. For example, consider two types of individuals, θ_1 and θ_2, where $\theta_1 < \theta_2$. If only one type of contract, $(R_{\theta_2}, p_{\theta_2})$, which provides high coverage at a high premium, is offered, individuals of type θ_2 may choose to purchase it. However, individuals of type θ_1 may be more comfortable with the risk and unwilling to pay a high premium for extensive coverage. Conversely, if only $(R_{\theta_1}, p_{\theta_1})$, which provides low coverage at a low premium, is available, type θ_1 individuals will participate, but type θ_2 individuals—who are more risk-averse and willing to pay more for greater coverage—would be underserved. In this case, offering separate contracts tailored to each individual type improves overall welfare. If type θ_1 individuals receive $(R_{\theta_1}, p_{\theta_1})$ and type θ_2 individuals receive $(R_{\theta_2}, p_{\theta_2})$, this design ensures that each policyholder benefits. When the coverage level of a menu of contracts $(R_\theta, p_\theta)_{\theta\in\Theta}$ follows such a monotonic structure for any level of risk, meaning that more risk-averse individuals receive greater coverage, thereby retaining less loss, we say that the $\{R_\theta\}_{\theta\in\Theta}$ is *submodular*. This concept is formally defined as follows.

Definition 4.3.1. *A collection $\{R_\theta\}_{\theta\in\Theta}$ of retention functions is submodular if $\frac{\partial R_\theta(x)}{\partial x}$ is non-increasing in θ for all $x \in [0, M]$.*

Combined with Proposition 4.3.2, the submodularity of the retention function allows us to derive a necessary and sufficient condition for a menu to be IC, as shown in the following result.

Proposition 4.3.3. *If $\{R_\theta\}_{\theta\in\Theta}$ is submodular, then the menu $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IC}$ if and only if $\{p_\theta\}_{\theta\in\Theta}$ satisfies (4.9).*

Proof. The "only if" part follows directly from Proposition 4.3.2. For the other direction, note first that for any $\theta < \theta'$, we have

$$
\begin{aligned}
&U_\theta\left(R_{\theta'}, p_{\theta'}\right) = -p_{\theta'} - \int_0^M \left[1 - g_\theta\left(F_X(x)\right)\right] \frac{\partial R_{\theta'}(x)}{\partial x}\mathrm{d}x \\
&= -p_{\underline{\theta}} - \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\mathrm{d}x + \int_{\underline{\theta}}^{\theta'} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s}\frac{\partial R_s(x)}{\partial x}\mathrm{d}x\mathrm{d}s \\
&+ \int_0^M \left[g_\theta\left(F_X(x)\right) - g_{\theta'}\left(F_X(x)\right)\right] \frac{\partial R_{\theta'}(x)}{\partial x}\mathrm{d}x \\
&= -p_{\underline{\theta}} - \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\mathrm{d}x + \int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s}\frac{\partial R_s(x)}{\partial x}\mathrm{d}x\mathrm{d}s \\
&+ \int_\theta^{\theta'} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s}\frac{\partial R_s(x)}{\partial x}\mathrm{d}x\mathrm{d}s - \int_0^M \int_\theta^{\theta'} \frac{\partial g_s\left(F_X(x)\right)}{\partial s}\mathrm{d}s\frac{\partial R_{\theta'}(x)}{\partial x}\mathrm{d}x \\
&= U_\theta\left(R_\theta, p_\theta\right) + \int_\theta^{\theta'} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s}\frac{\partial R_s(x)}{\partial x}\mathrm{d}x\mathrm{d}s - \int_\theta^{\theta'} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s}\frac{\partial R_{\theta'}(x)}{\partial x}\mathrm{d}s\mathrm{d}x \\
&\leq U_\theta\left(R_\theta, p_\theta\right),
\end{aligned}
$$

where the second equality follows from the expression for $t_{\theta'}$ in (4.9), while the fourth equality is based on the expression for $U_\theta\left(R_\theta, p_\theta\right)$ in (4.13), when p_θ satisfies (4.9). Lastly, the inequality follows from Assumption 4.2.2 and the fact that $\frac{\partial R_\theta(x)}{\partial x} \leq \frac{\partial R_{\theta'}(x)}{\partial x}$ for any $\theta < \theta'$ given that R is submodular. Similarly, we can show that $U_{\theta'}\left(R_\theta, p_\theta\right) \leq U_{\theta'}\left(R_{\theta'}, p_{\theta'}\right).$ Thus, the menu is IC. □

The necessary part of this corollary is obvious from Proposition 4.3.2. The sufficient part holds when R is monotone with the risk type. This corollary indicates that an IC menu is equivalent to any insurance menu with monotonic insurance coverage and premium determined by (4.9). This aligns with the condition for a submodular retention function to be an IC menu when the hidden information is the loss distribution ratrher than the policyholder's level of risk aversion. This is characterized in Gershkov et al. (2023, Proposition 1(ii)). Notably, in Gershkov et al. (2023), submodular retention implies offering greater insurance coverage to individuals characterized by a stochastically larger loss distribution.

In addition to the IC condition, the menus of contracts must be designed to ensure that policyholders are willing to participate. Thus, we also require that the menus of contracts be IR. The following result characterizes a menu that satisfies both the IC and IR conditions.

Proposition 4.3.4. *If $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IC}$, then $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IR}$ if and only if $(R_{\underline{\theta}}, p_{\underline{\theta}}) \in \mathcal{IR}$.*

Proof. Consider a menu $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IC}$. By Proposition 4.3.2, we have that

$$\frac{\partial U_\theta(X,0)}{\partial\theta} = \int_0^M \frac{\partial g_\theta\left(F_X\left(x\right)\right)}{\partial\theta}\mathrm{d}x \leq \int_0^M \frac{\partial g_\theta\left(F_X\left(x\right)\right)}{\partial\theta}\frac{\partial R_\theta(x)}{\partial x}\mathrm{d}x = \frac{\partial U_\theta(R_\theta, p_\theta)}{\partial\theta},$$

where the inequality follows from Assumption 4.2.2, and the fact that $R_\theta \in \mathcal{R}_L$. Assume that for the lowest type $\underline{\theta}$, the contract is IR. Then, $U_{\underline{\theta}}(R_{\underline{\theta}}, p_{\underline{\theta}}) \geq U_{\underline{\theta}}(X, 0)$. For any other $\theta \in \Theta$, we have that

$$U_\theta(X,0) = U_{\underline{\theta}}(X,0) + \int_{\underline{\theta}}^{\theta} \frac{\partial U_s(X,0)}{\partial s}\mathrm{d}s \leq U_{\underline{\theta}}(R_{\underline{\theta}}, p_{\underline{\theta}}) + \int_{\underline{\theta}}^{\theta} \frac{\partial U_s(R_s, p_s)}{\partial s}\mathrm{d}s = U_\theta(R_\theta, p_\theta),$$

which implies that $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IR}$. □

This result indicates that if the menu is IC, it is also IC, provided that the contract offered to the lowest type satisfies the IR condition. This significantly simplifies the analysis by reducing the need to separately verify IR constraints for all types. This observation aligns with the setting in which the private information lies in the loss distribution rather than in the level of risk aversion, as demonstrated in Lemma 1 of Gershkov et al. (2023). Based on Proposition 4.3.4, we can determine an interval of values for $p_{\underline{\theta}}$ that ensures that an IC menu is IR. This is given below.

Corollary 4.3.1. *If $\{R_\theta\}_{\theta\in\Theta}$ is submodular, then $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IR}\cap\mathcal{IC}$ if and only if $\{p_\theta\}_{\theta\in\Theta}$ satisfies (4.9), with*

$$p_{\underline{\theta}} \leq \int_0^M \left[1 - g_{\underline{\theta}}(F_X(x))\right]\left(1 - \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\right) \mathrm{d}x.$$

Proof. First, suppose that $\{R_\theta\}_{\theta\in\Theta}$ is submodular and that $\{p_\theta\}_{\theta\in\Theta}$ satisfies (4.9). By Proposition 4.3.3, it follows that the menu $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IC}$. Additionally, if

$$p_{\underline{\theta}} \leq \int_0^M \left[1 - g_{\underline{\theta}}(F_X(x))\right]\left(1 - \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\right) \mathrm{d}x,$$

then we have

$$\begin{aligned}
U_{\underline{\theta}}(R_{\underline{\theta}}, p_{\underline{\theta}}) &= -p_{\underline{\theta}} - \int_0^M \left[1 - g_{\underline{\theta}}(F_X(x))\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x} \mathrm{d}x \\
&\geq -\int_0^M \left[1 - g_{\underline{\theta}}(F_X(x))\right]\left(1 - \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\right) \mathrm{d}x - \int_0^M \left[1 - g_{\underline{\theta}}(F_X(x))\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x} \mathrm{d}x \\
&= -\int_0^M \left[1 - g_{\underline{\theta}}(F_X(x))\right] \mathrm{d}x = U_{\underline{\theta}}(X, 0).
\end{aligned}$$

Hence, $(R_{\underline{\theta}}, p_{\underline{\theta}}) \in \mathcal{IR}$. By Proposition 4.3.4, it follows that $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IR}$.

Conversely, suppose that $\{R_\theta\}_{\theta\in\Theta}$ is submodular and that $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IR}\cap\mathcal{IC}$. By Proposition 4.3.3, $\{p_\theta\}_{\theta\in\Theta}$ satisfies (4.9). Furthermore, since $(R_{\underline{\theta}}, p_{\underline{\theta}}) \in \mathcal{IR}$, we obtain $U_{\underline{\theta}}(R_{\underline{\theta}}, p_{\underline{\theta}}) \geq U_{\underline{\theta}}(X, 0)$, which implies that

$$p_{\underline{\theta}} \leq \int_0^M \left[1 - g_{\underline{\theta}}(F_X(x))\right]\left(1 - \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\right) \mathrm{d}x.$$

□

This result follows directly from Proposition 4.3.3 and Proposition 4.3.4. It implies that whenever we have an IC menu with monotone insurance coverage, it is enough to verify the premium of the lowest risk type to determine whether the menu is individually rational.

In the next section, we will use the results given above in order to characterize the profit-maximizing, IR, and IC menus.

4.4 Solution and Properties of the Optimal Menu

4.4.1 Equilibrium Contracts with Full Information

As argued by STIGLITZ (1977), in a monopoly market with a risk-neutral insurer and a risk-averse EU-maximizing policyholder, and under perfect information, equilibria consist of full-insurance contracts at premia that make the policyholder indifferent between purchasing insurance and foregoing coverage. That is, all the policyholder surplus involved in the reduction of risk is extracted by the monopoly, as is classically understood. In this section we show that this insight still holds when the policyholder is Yaari DU-maximizer.

To that end, consider the benchmark case of full information in the market, whereby the risk aversion of the policyholder is observable by the insurer. This is tantamount to assuming that the type space is a singleton of the form $\Theta = \{\theta_0\}$. In this case, the insurer no longer needs to offer menus of contracts, but rather a single contract to the policyholder, which is *de facto* incentive compatible. The monopolist's problem is therefore to design a contract that is profit-maximizing and individually rational:

Problem 2.

$$\sup_{(R,p)\in\mathcal{R}_L\times\mathbb{R}_+} \Big\{\pi(R,p) : U_{\theta_0}(R,p) \geq U_{\theta_0}(X,0)\Big\}.$$

Noting that for all $R \in \mathcal{R}_L$,

$$\int_0^M \left[1 - g_{\theta_0}\left(\mathbb{P}\left[R(X) \leq x\right]\right)\right] \mathrm{d}x = \int_0^M \left[1 - g_{\theta_0}\left(F_X(x)\right)\right] R'(x)\, \mathrm{d}x,$$

and using (4.3) and (4.6), Problem 2 can be rewritten as:

$$\sup_{(R,p)\in\mathcal{R}_L\times\mathbb{R}_+} \left\{p - \mathbb{E}[X] + \int_0^M \left[1 - F_X(x)\right] R'(x)\, \mathrm{d}x \,\middle|\, p \leq \int_0^M \left[1 - g_{\theta_0}\left(F_X(x)\right)\right] \left[1 - R'(x)\right] \mathrm{d}x\right\}. \tag{4.14}$$

The following result shows that the constraint in Problem (4.14) is binding at an optimum.

Lemma 4.4.1. *If $(R^*, p^*) \in \mathcal{R}_L \times \mathbb{R}_+$ is optimal for Problem (4.14), then*

$$p^* = \int_0^M \left[1 - g_{\theta_0}\left(F_X(x)\right)\right] \left[1 - (R^*)'(x)\right] \mathrm{d}x. \tag{4.15}$$

Proof. Suppose that $(R^*, p^*) \in \mathcal{R}_L \times \mathbb{R}_+$ is optimal for Problem (4.14), but that

$$0 \leq p^* < \int_0^{+\infty} [1 - g_{\theta_0}(F_X(x))] \, [1 - (R^*)'(x)] \, \mathrm{d}x := \widetilde{p},$$

Then, for the contract $(R^*, \widetilde{p})$, we have

$$\begin{aligned}
U_{\theta_0}(R^*, \widetilde{p}) &= -\widetilde{p} - \int_0^{+\infty} [1 - g_{\theta_0}(F_X(x))] \, (R^*)'(x) \, \mathrm{d}x \\
&= -\int_0^{+\infty} [1 - g_{\theta_0}(F_X(x))] \, [1 - (R^*)'(x)] \, \mathrm{d}x - \int_0^{+\infty} [1 - g_{\theta_0}(F_X(x))] \, (R^*)'(x) \, \mathrm{d}l \\
&= -\int_0^{+\infty} [1 - g_{\theta_0}(F_X(x))] \, \mathrm{d}x \\
&= U_{\theta_0}(X, 0),
\end{aligned}$$

so that the contract $(R^*, \widetilde{p})$ is feasible for Problem (4.14). Moreover,

$$\begin{aligned}
\pi(R^*, \widetilde{p}) &= \widetilde{p} - \mathbb{E}[X] + \int_0^M [1 - F_X(x)] \, (R^*)'(x) \, \mathrm{d}x \\
&> p^* - \mathbb{E}[X] + \int_0^M [1 - F_X(x)] \, (R^*)'(x) \, \mathrm{d}x \\
&= \pi(R^*, p^*),
\end{aligned}$$

contradicting the optimality of (R^*, p^*) for Problem (4.14). Hence,

$$p^* = \int_0^{+\infty} [1 - g_{\theta_0}(F_X(x))] \, [1 - (R^*)'(x)] \, \mathrm{d}x.$$

□

The following result shows that in the case of perfect information in our monopoly market, full insurance is optimal, at a premium that makes the policyholder indifferent between insurance and no insurance.

Proposition 4.4.1. *The optimal solution to Problem 2 is given by the contract*

$$(R^*_{\theta_0}, p^*_{\theta_0}) = \left(0, \int_0^M [1 - g_{\theta_0}(F_X(x))] \, \mathrm{d}x\right).$$

Moreover,

$$U_{\theta_0}(R^*_{\theta_0}, p^*_{\theta_0}) = U_{\theta_0}(X, 0).$$

Proof. By Lemma 4.4.1, a contract (R^*, p^*) is optimal for Problem (4.14) if and only if R^* is optimal for

$$\sup_{R\in\mathcal{R}_L}\left\{\aleph_{\theta_0}(R) := \int_0^M \left[1-g_{\theta_0}\left(F_X(x)\right)\right]\left[1-R'(x)\right]\mathrm{d}x - \mathbb{E}[X] + \int_0^M \left[1-F_X\left(x\right)\right] R'(x)\,\mathrm{d}x\right\}, \tag{4.16}$$

and

$$p^* := \int_0^{+\infty} \left[1-g_{\theta_0}\left(F_X(x)\right)\right]\left[1-R'(x)\right]\mathrm{d}x = \int_0^M \left[1-g_{\theta_0}\left(F_X(x)\right)\right]\left[1-R'(x)\right]\mathrm{d}x.$$

Now, for each $R \in \mathcal{R}_L$,

$$\begin{aligned}
\aleph_{\theta_0}(R) &= -\mathbb{E}[X] + \int_0^M \left[1-g_{\theta_0}\left(F_X(x)\right)\right]\left[1-R'(x)\right]\mathrm{d}x + \int_0^M \left[1-F_X\left(x\right)\right] R'(x)\,\mathrm{d}x \\
&= -\mathbb{E}[X] + \int_0^M \left[1-g_{\theta_0}\left(F_X(x)\right)\right]\mathrm{d}x - \int_0^M \left[F_X(x)-g_{\theta_0}\left(F_X(x)\right)\right] R'(x)\,\mathrm{d}x \\
&= -\int_0^M \left[1-F_X(x)\right]\mathrm{d}x + \int_0^M \left[1-g_{\theta_0}\left(F_X(x)\right)\right]\mathrm{d}x - \int_0^M \left[F_X(x)-g_{\theta_0}\left(F_X(x)\right)\right] R'(x)\,\mathrm{d}x \\
&= \int_0^M \left[F_X(x)-g_{\theta_0}\left(F_X(x)\right)\right]\mathrm{d}x - \int_0^M \left[F_X(x)-g_{\theta_0}\left(F_X(x)\right)\right] R'(x)\,\mathrm{d}x \\
&= \int_0^M \left[F_X(x)-g_{\theta_0}\left(F_X(x)\right)\right]\left[1-R'(x)\right]\mathrm{d}x.
\end{aligned}$$

Since $F_X(x) - g_{\theta_0}\left(F_X(x)\right) \geq 0$, for all $x \in [0, M]$ (by Assumption 4.2.1), and since $R' \in [0, 1]$ for all $R \in \mathcal{R}_L$, it follows that the $(R^*)' \equiv 0$ is optimal for Problem 4.16. Hence, the contract $\left(0, \int_0^M \left[1-g_{\theta_0}\left(F_X(x)\right)\right]\mathrm{d}x\right) = (0, -U_{\theta_0}(X, 0))$ is optimal for Problem (4.14). Moreover, by (4.3),

$$U_{\theta_0}\left(0, -U_{\theta_0}(X, 0)\right) = U_{\theta_0}(X, 0),$$

meaning that the policyholder is indifferent between purchasing insurance and not doing so.

□

The optimal contract entails zero retention, that is, full insurance, at a premium of

$$p^*_{\theta_0} = \int_0^M \left[1-g_{\theta_0}\left(F_X\left(x\right)\right)\right]\mathrm{d}x = -U_{\theta_0}(X, 0) = -DU(-X),$$

which leads to all the policyholder surplus being absorbed by the monopoly.

It should be noted that the equilibrium contract discussed in this section differs from the equilibrium contracts presented in Chapters 2 and 3, where the analysis is conducted within a Stackelberg game framework. In this section, the optimal contract is designed to maximize the insurer's profit while satisfying both IR and IC conditions. In this setting, the policyholder can only choose among contract types offered by the insurer but does not influence the structure of the insurance contract itself. The structure is determined entirely from the insurer's perspective, with the goal of maximizing profit. When there is only a single type of policyholder in the market, the optimal contract is the one that maximizes the insurer's profit subject to the policyholder's IR constraint. In contrast, the optimal contracts defined in Definitions 2.2.1 and 3.2.2 are derived by first considering the policyholder's welfare. These contracts aim to maximize the policyholder's utility, and thus are inherently IR. Among such contracts, those that also yield higher profits for the insurer can serve as equilibrium outcomes in the BO market mechanism.

4.4.2 Equilibrium Contracts with Hidden Risk Attitudes

We now extend the benchmark case of a market with full information to a setting with imperfect information, in which the policyholder's risk aversion is unobservable to the insurer. Theorem 4.4.1 characterizes the profit-maximizing menu under imperfect information, ensuring incentive compatibility and individual rationality, with the solution expressed in terms of marginal loss retention.

Theorem 4.4.1. *Suppose that $J_\theta(x)$ given in (4.11) is non-decreasing in θ, for all x. An optimal solution $(R^*_\theta, p^*_\theta)_{\theta\in\Theta}$ for Problem 1 is characterized by the following:*

1. *For each $\theta \in \Theta$, the retention function R^*_θ satisfies*

$$\frac{\partial R^*_\theta(x)}{\partial l} = \begin{cases} 0, & J_\theta(x) > 0, \\ \in [0,1], & J_\theta(x) = 0, \\ 1, & J_\theta(x) < 0. \end{cases} \tag{4.17}$$

2. *For each $\theta \in \Theta$, the corresponding premium p^*_θ is given by:*

$$p^*_\theta := \int_0^M [1 - g_{\underline{\theta}}(F_X(x))]\,\mathrm{d}x - \int_{\underline{\theta}}^{\theta}\int_0^M \frac{\partial g_s(F_X(x))}{\partial s}\frac{\partial R^*_s(x)}{\partial x}\mathrm{d}x\mathrm{d}s - \int_0^M [1 - g_\theta(F_X(x))]\frac{\partial R^*_\theta(x)}{\partial x}\,\mathrm{d}x. \tag{4.18}$$

Moreover, the collection $\{R^*_\theta\}_{\theta\in\Theta}$ of optimal retention functions is submodular.

Proof. First note that the profit in (4.10) is an increasing function of $p_{\underline{\theta}}$. Therefore, at the optimum, $p_{\underline{\theta}}$ must take its largest value, provided the IR condition is still satisfied. By Proposition 4.3.1, we conclude that $p^*_{\underline{\theta}} = \int_0^M [1 - g_{\underline{\theta}}(F_X(x))] \left(1 - \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\right) \mathrm{d}x$, where $R_{\underline{\theta}}$ is the retention function for the lowest type policyholder.

For $R_\theta \in \mathcal{R}_L$, we know that $R_\theta(0) = 0$, and $\frac{\partial R_\theta(x)}{\partial x} \in [0, 1]$ for any $\theta \in \Theta$. An optimal solution for Problem 1 exists because of the continuity of the profit functional and the compactness of the set of retention functions (Remark 5).

With $p_{\underline{\theta}} := \int_0^M [1 - g_{\underline{\theta}}(F_X(x))] \left(1 - \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\right) \mathrm{d}x$, the insurer's expected profit in (4.10) becomes

$$\int_0^M [1 - g_{\underline{\theta}}(F_X(x))] \,\mathrm{d}x - \mathbb{E}[X] - \int_{\underline{\theta}}^{\bar{\theta}} \left(\int_0^M J_\theta(x) \frac{\partial R_\theta(x)}{\partial x} \mathrm{d}x \right) f(\theta) \mathrm{d}\theta. \tag{4.19}$$

We maximize (4.19) pointwise. Specifically, for a fixed $\theta \in \Theta$, we look for

$$R^*_\theta \in \underset{R_\theta \in \mathcal{R}_L}{\arg\max} \; - \int_0^M J_\theta(x) \frac{\partial R_\theta(x)}{\partial x} \mathrm{d}x.$$

The maximum is achieved when the retention function takes the form given in (4.17). For any $\theta < \theta'$, since $J_\theta(x)$ is increasing in θ for all x, it follows that $J_\theta(x) \leq J_{\theta'}(x)$. Thus, the pointwise maximization solution satisfies

$$\frac{\partial R^*_{\theta'}(x)}{\partial x} \leq \frac{\partial R^*_\theta(x)}{\partial x}$$

for all l, meaning $R^*_\theta(x)$ is submodular. Therefore, by Proposition 4.3.3, it is IC. Additionally, the optimal contract for the lowest risk type $(R^*_{\underline{\theta}}, p^*_{\underline{\theta}})$ satisfies Corollary 4.3.1, which confirms that the menu is IR. □

The above results asserts that, under the assumption that, for each $x \in [0, M]$, the function $\theta \mapsto J_\theta(x)$ is non-decreasing, the optimal menu of contracts consists of a collection of layered indemnity schedules, and hence layered corresponding retention functions. Specifically, the marginal retention at a loss of value $x \in [0, M]$ is 0 when $J_\theta(x) > 0$, corresponding to full insurance. When $J_\theta(x) < 0$, the marginal retention is 1, corresponding to no insurance for that level of loss. When $J_\theta(x) = 0$, the optimal retention allows for some flexibility, as long as feasibility is maintained. Additionally, at an optimum, the premium determined by a non-linear function of the insurance coverage as described in (4.18). In other words, a profit-maximizing, IR, and IC menu of contracts leads to a separating equilibrium. We comment further below on the monotonicity of $J_\theta(x)$ in θ.

Additionally, the optimal menu has an intuitive property. Since the collection $\{R^*_\theta\}_{\theta\in\Theta}$ of optimal retention functions is submodular, it follows at any loss level $x \in [0, M]$, the marginal loss retention of a lower risk type always exceeds that of a higher risk type. A less risk-policyholder retains more marginal loss than a higher risk-averse policyholder at an optimum.

While in competitive insurance markets only differences in risk distribution affect equilibria and the nature of separating contracts, in a monopoly both risk and risk attitude have an impact on equilibrium contracts, as observed by Stiglitz (1977). The latter showed if the loss distribution is known by the monopolist insurer, but the policyholder's level of risk aversion is hidden from the insurer, then equilibrium contracts can have different forms, and they can be separating or pooling equilibria. We show below that this is not the case in our setting, and that separating equilibria always outperform pooling equilibria.

Proposition 4.4.2. *Suppose that $J_\theta(x)$ given in (4.11) is non-decreasing in θ for all x. The separating equilibrium described in Theorem 4.4.1 is more valuable to the insurer than a pooling equilibrium contract.*

Proof. Let (R_p, p_p) denote the single contract offered to all policyholders, where R_p represents the retention and t_p the premium. For an policyholder of type θ, the utility under this contract is given by

$$U_\theta (R_p, p_p) = -p_p - \int_0^M [1 - g_\theta (F_X (x))] \frac{\partial R_p(x)}{\partial x} \, \mathrm{d}x.$$

With a single contract, one only needs to verify that the IR condition holds. Specifically, for each $\theta \in \Theta$, the contract (R_p, p_p) must satisfy $U_\theta (R_p, p_p) \geq U_\theta (X, 0)$, or equivalently,

$$\int_0^M [1 - g_\theta (F_X (x))] \left(1 - \frac{\partial R_p(x)}{\partial x}\right) \mathrm{d}x \geq p_p,$$

for any $\theta \in \Theta$. This condition holds for all $\theta \in \Theta$ if and only if it holds for $\underline{\theta}$, as g_θ is decreasing in θ. Furthermore, since the profit is increasing in p_p, the optimal premium p_p for any given R_p satisfies:

$$\int_0^M [1 - g_{\underline{\theta}} (F_X (x))] \left(1 - \frac{\partial R_p(x)}{\partial x}\right) \mathrm{d}x = p_p.$$

Thus, the insurer's expected profit with the above premium is given by

$$\begin{aligned}
V\left(R_p, p_p\right) &= \int_{\theta}\left(p_p - \mathbb{E}[X] + \int_0^M \left(1 - F_X\left(x\right)\right)\frac{\partial R_p(x)}{\partial x}\,\mathrm{d}x\right)\mathrm{d}F(\theta) \\
&= \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X\left(x\right)\right)\right]\left(1 - \frac{\partial R_p(x)}{\partial x}\right)\mathrm{d}x - \int_0^M \left[1 - F_X\left(x\right)\right]\left(1 - \frac{\partial R_p(x)}{\partial x}\right)\mathrm{d}x \\
&= \int_0^M \left[F_X\left(x\right) - g_{\underline{\theta}}\left(F_X\left(x\right)\right)\right]\left(1 - \frac{\partial R_p(x)}{\partial x}\right)\mathrm{d}x.
\end{aligned}$$

Hence, the profit decreases as the marginal retention function increases. Thus, at the optimum, the retention function satisfies $R_p^* = 0$. Consequently, the maximum profit with a single contract is,

$$V\left(R_p^*, p_p^*\right) = \int_0^M \left[F_X\left(x\right) - g_{\underline{\theta}}\left(F_X\left(x\right)\right)\right]\mathrm{d}x,$$

and

$$\begin{aligned}
V\left(R_p^*, p_p^*\right) &\leq \int_0^M \left[F_X(x) - g_{\underline{\theta}}\left(F_X\left(x\right)\right)\right]\mathrm{d}x - \int_{\underline{\theta}}^{\overline{\theta}}\left(\int_0^M J_\theta(x)\mathbb{1}_{\{J_\theta(x)<0\}}\mathrm{d}x\right) f(\theta)\,\mathrm{d}\theta \\
&= -U_{\underline{\theta}}\left(X, 0\right) - \mathbb{E}[X] - \int_{\underline{\theta}}^{\overline{\theta}}\left(\int_0^M J_\theta(x)\mathbb{1}_{\{J_\theta(x)<0\}}\mathrm{d}x\right) f(\theta)\,\mathrm{d}\theta,
\end{aligned}$$

which is the expected profit obtained from the solution given in Theorem 4.4.1. Thus, for policyholders with varying risk attitudes, designing a menu of contracts to achieve a separating equilibrium is always more advantageous for the insurer than providing a uniform contract that leads to a pooling equilibrium. □

The above result demonstrates that in a monopolistic market, selecting contracts that satisfy the IR and IC conditions, as outlined in Theorem 4.4.1, always outperforms offering a single contract. This result extends a key finding of Boonen and Zhang (2021). More properties of the optimal menu will be examined in Section 4.4.4.

4.4.3 The Function $\theta \mapsto J_\theta(x)$

Two Components

Next, we provide an analysis of how the virtual value function

$$J_\theta(x) = F_X(x) - g_\theta\left(F_X\left(x\right)\right) + \frac{\overline{F}(\theta)}{f(\theta)}\frac{\partial g_\theta\left(F_X\left(x\right)\right)}{\partial \theta}$$

affects the insurer's profit, and we provide conditions along with interpretations for the monotonicity of the map $\theta \mapsto J_\theta(x)$ to hold.

First, note that $J_\theta(x)$ can be decomposed into two components. The first component is

$$F_X(x) - g_\theta\left(F_X(x)\right),$$

which is nonnegative by Assumption 4.2.1 (weak risk aversion), and non-decreasing in θ by Assumption 4.2.2 (ordered type space). This quantity can be seen as a proxy for the level of risk aversion of type θ at the point $F_X(x) \in [0,1]$, since weak risk aversion of type θ is equivalent to $g_\theta(s) \leq s$, for $s \in [0,1]$.

The second component of $J_\theta(x)$ is

$$\frac{\overline{F}(\theta)}{f(\theta)} \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta},$$

which can be interpreted as the information rent. It serves as a cost since $\frac{\partial g_\theta(F_X(x))}{\partial \theta}$ is non-positive, by Assumption 4.2.2. This arises because policyholders have private knowledge about their risk attitudes, which the insurer does not observe directly. The insurer must account for this private information when designing contracts to mitigate adverse selection. If this component is increasing in θ, then marginal changes in coverage or retention result in a smaller reduction in the insurer's profit for higher values of θ. This implies that as risk aversion increases, it becomes less costly to design incentive compatible contracts for more risk averse individuals.

Together, these two components determine how the insurer's profit changes for any marginal increase in coverage or marginal decrease in retention.

Monotonicity

Next, we provide some intuitive sufficient conditions for the monotonicity of the map $\theta \mapsto J_\theta(x)$, for a given $x \in [0, M]$. First, as mentioned above, the first component is non-decreasing in θ, by assumption. For the second component, differentiating with respect to θ yields

$$\begin{aligned}\frac{\partial}{\partial \theta}\left(\frac{\overline{F}(\theta)}{f(\theta)} \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta}\right) &= \left(\frac{\overline{F}(\theta)}{f(\theta)}\right)' \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta} + \frac{\overline{F}(\theta)}{f(\theta)} \frac{\partial^2 g_\theta\left(F_X(x)\right)}{\partial \theta^2} \\ &= -\frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta}\left(-\frac{\overline{F}(\theta)}{f(\theta)}\right)' + \frac{\overline{F}(\theta)}{f(\theta)} \frac{\partial^2 g_\theta\left(F_X(x)\right)}{\partial \theta^2}. \end{aligned} \tag{4.20}$$

This quantity is non-negative if the following two conditions hold jointly:

1. The hazard rate $f/\overline{F}$ is non-decreasing over Θ;
2. For each $s \in (0,1)$, the function $\theta \mapsto g_\theta(s)$ is convex.

The first condition is imposed on the distribution F of types over the type space Θ. A non-decreasing hazard rate for the distribution F ensures that the first term in (4.20) is non-negative, since $\frac{\partial g_\theta(F_X(x))}{\partial \theta} \leq 0$, by assumption.

The second condition implies that the second term in (4.20) is non-negative. Together with Assumption 4.2.2, this condition implies that the utility specified in (4.2) is non-increasing and convex with respect to the risk aversion parameter θ. In other words, the difference in the utility of wealth between risk types diminishes as the type increases. This observation is consistent with commonly used utility functions, such as exponential utility functions or power utility functions, for instance.

Many distortion functions satisfy the above two conditions. For example: $g_\theta(s) = \theta s^2 + (1-\theta)s$, where $\theta \in [0,1]$; $g_\theta(s) = \frac{s}{1+\theta(1-s)}$, where $\theta \in [0,1]$; and the family of exponential distortions: $g_\theta(s) = \frac{e^{-\theta(1-s)} - e^{-\theta}}{1-e^{-\theta}}$, where $\theta \in [0,1]$.

If the above two conditions hold, then $J_\theta(x)$ is a monotonic function of θ for any $x \in [0, M]$. A similar monotonicity condition is also introduced when the hidden information pertains to the risk distribution, as discussed in Gershkov et al. (2023). In their study, a similar function to J appears. To ensure monotonicity of this function, the authors assume that the type distribution has a non-decreasing hazard rate. Additionally, they impose the following other conditions:

1. The distortion function in Yaari's dual utility framework must be convex. This ensures that the utility function reflects aversion to mean-preserving spreads (strong risk aversion).
2. The collection of loss distributions indexed by types must exhibit stochastic concavity with respect to the type parameter. This implies that as the risk type increases, the risk distributions satisfy a concave ordering.
3. For the highest risk type $\overline{\theta}$, with cumulative distribution function $H_{\overline{\theta}}$, it is assumed that $g(H_{\overline{\theta}}(0)) \geq 1$. This condition corresponds to the requirement that the probability of an accident for the highest risk type must meet a minimum threshold.

4.4.4 Properties of Optimal Menus

The following proposition elaborates on several properties of the optimal menu.

Proposition 4.4.3. *Suppose that $J_\theta(x)$ given in (4.11) is non-decreasing in θ, for all x. The optimal menu $(R^*_\theta, p^*_\theta)_{\theta\in\Theta}$ in Theorem 4.4.1 satisfies the following properties.*

1. *The retention function R^*_θ decreases with θ, and the premium p^*_θ increases with θ.*
2. *For the most risk averse policyholder, $R_{\overline{\theta}}(x) = 0$ for all $x \in [0, M]$.*
3. *For the least risk averse policyholder, $U_{\underline{\theta}}(R^*_{\underline{\theta}}, p^*_{\underline{\theta}}) = U_{\underline{\theta}}(X, 0)$.*
4. *The policyholder's utility $U_\theta\left(R^*_\theta, p^*_\theta\right)$ decreases with θ. Moreover, it is convex in θ if g_θ is convex in θ.*
5. *The insurer's expected profit from a type-θ contract, that is $\pi\left(R^*_\theta, p^*_\theta\right)$, increases with θ.*
6. *The insurer's overall expected profit is positive: $V\left((R^*_\theta, p^*_\theta)_{\theta\in\Theta}\right) > 0$.*

Proof. 1. Following from Theorem 4.4.1, R^*_θ is decreasing in θ due to submodularity. Substituting the optimal retention into the premium expression in (4.9) and taking the derivative yields

$$\frac{\partial p^*_\theta}{\partial \theta} = -\int_0^M \frac{\partial g_\theta\left(F_X(x)\right)}{\partial s}\frac{\partial R^*_\theta(x)}{\partial x}\mathrm{d}x + \int_0^M \frac{\partial g_\theta\left(F_X(x)\right)}{\partial s}\frac{\partial R^*_\theta(x)}{\partial x}\mathrm{d}x$$
$$\int_0^M \left(1 - g_\theta\left(F_X(x)\right)\right)\frac{\partial}{\partial \theta}\frac{\partial R^*_\theta(x)}{\partial x}\mathrm{d}x \geq 0,$$

since $\frac{\partial}{\partial\theta}\frac{\partial R^*_\theta(x)}{\partial x} \leq 0$.

2. This follows from the fact that $J_{\overline{\theta}}(x) \leq 0$ for all x.
3. This follows from the expression for $p^*_{\underline{\theta}}$ in (4.18).
4. From the expression of the policyholder's utility for an IC menu given in (4.13),

$$\frac{\partial U_\theta\left(R^*_\theta, p^*_\theta\right)}{\partial \theta} = \int_0^M \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta}\frac{\partial R^*_\theta(x)}{\partial x}\mathrm{d}x \leq 0,$$

since g_θ satisfies Assumption 4.2.2. In addition, the second derivative is given by

$$\frac{\partial^2 U_\theta\left(R^*_\theta, p^*_\theta\right)}{\partial \theta^2} = \int_0^M \frac{\partial}{\partial\theta}\left(\frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta}\right)\cdot\frac{\partial R^*_\theta(x)}{\partial x}\mathrm{d}x$$
$$+\int_0^M \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta}\cdot\frac{\partial}{\partial\theta}\left(\frac{\partial R^*_\theta(x)}{\partial x}\right)\mathrm{d}x \geq 0,$$

if $\frac{\partial g_\theta(s)}{\partial \theta}$ is increasing. Therefore, the utility of wealth at an optimum is decreasing and convex in θ.

5. The profit in (4.6) for the contract (R_θ^*, p_θ^*) is

$$
\begin{aligned}
&\pi\left(R_\theta^*, p_\theta^*\right) = p_\theta^* - \mathbb{E}[X] + \int_0^M \left[1 - F_X(x)\right] \frac{\partial R_\theta^*(x)}{\partial x} \mathrm{d}x \\
&= \int_0^M \left[1 - g_{\underline{\theta}}(F_X(x))\right] \mathrm{d}l - \int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s} \frac{\partial R_s^*(x)}{\partial x} \mathrm{d}x\mathrm{d}s - \int_0^M \left[1 - g_\theta\left(F_X(x)\right)\right] \frac{\partial R_\theta^*(x)}{\partial x} \mathrm{d}x \\
&- \mathbb{E}[X] + \int_0^M \left[1 - F_X(x)\right] \frac{\partial R_\theta^*(x)}{\partial x} \mathrm{d}x \\
&= \int_0^M \left[F_X(x) - g_{\underline{\theta}}\left(F_X(x)\right)\right] \mathrm{d}x - \int_0^M \left[F_X(x) - g_\theta\left(F_X(x)\right)\right] \frac{\partial R_\theta^*(x)}{\partial x} \mathrm{d}x \\
&- \int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s} \frac{\partial R_s^*(x)}{\partial x} \mathrm{d}x\mathrm{d}s. \qquad (4.21)
\end{aligned}
$$

Taking the derivative with respect to θ yields

$$
\begin{aligned}
\frac{\partial \pi\left(R_\theta^*, p_\theta^*\right)}{\partial \theta} &= -\int_0^M \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta} \frac{\partial R_\theta^*(x)}{\partial x} \mathrm{d}x + \int_0^M \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta} \frac{\partial R_\theta^*(x)}{\partial x} \mathrm{d}x \\
&- \int_0^M \left[F_X(x) - g_\theta\left(F_X(x)\right)\right] \frac{\partial}{\partial \theta} \frac{\partial R_\theta^*(x)}{\partial x} \mathrm{d}x \\
&= -\int_0^M \left[F_X(x) - g_\theta\left(F_X(x)\right)\right] \frac{\partial}{\partial \theta} \frac{\partial R_\theta^*(x)}{\partial x} \mathrm{d}x \geq 0,
\end{aligned}
$$

for any $\theta \in \Theta$.

6. At an optimum, the expected profit in (4.10) is

$$
\begin{aligned}
V\left((R_\theta^*, p_\theta^*)_{\theta \in \Theta}\right) &= \int_0^M \left[F_X(x) - g_{\underline{\theta}}\left(F_X(x)\right)\right] \mathrm{d}x - \int_{\underline{\theta}}^{\overline{\theta}} \left(\int_0^M \frac{\partial R_\theta^*(x)}{\partial x} J_\theta(x) \mathrm{d}x\right) f(\theta) \mathrm{d}\theta \\
&= \int_0^M \left[F_X(x) - g_{\underline{\theta}}\left(F_X(x)\right)\right] \mathrm{d}x - \int_{\underline{\theta}}^{\overline{\theta}} \left(\int_0^M J_\theta(x) \mathbb{1}_{\{J_\theta(x) < 0\}} \mathrm{d}x\right) f(\theta) \mathrm{d}\theta \\
&\geq \int_0^M \left[F_X(x) - g_{\underline{\theta}}\left(F_X(x)\right)\right] \mathrm{d}x = -\mathbb{E}[X] - U_{\underline{\theta}}(X, 0) > 0,
\end{aligned}
$$

where the last inequality follows from Proposition 4.2.1(1). This inequality is strict since $\underline{\theta}$ is risk averse. Thus, the insurer earns a positive profit.

□

This proposition outlines key properties of the profit-maximizing menu. The first result shows that insurance coverage and premia are monotone in the level of risk aversion. A more risk averse policyholder is offered a contract with greater coverage but at a higher price. This finding aligns with GERSHKOV et al. (2023, Theorem 1) and CHADE and SCHLEE (2012), in a setting with the hidden information is the loss distribution rather than the risk attitude of the policyholder. However, the key distinction is that in CHADE and SCHLEE (2012), the monotonicity of coverage and premia is imposed by assumption. Indeed, their approach relies on the assumption that the policyholder's utility satisfies the single-crossing property, which implies that the second derivative of the utility with respect to the retention and the premium is positive, ensuring the monotonicity of coverage and premia with risk types, as the authors show. In contrast, our results establish this monotonicity as a property of optimal menus. Rather than imposing conditions directly on the policyholder's utility, we impose a monotonicity condition on the virtual value function $J_\theta(x)$, as is the case in GERSHKOV et al. (2023).

The second result demonstrates that the most risk-averse policyholder will be offered full insurance coverage. This phenomenon, referred to as *efficiency at the top* in CHADE and SCHLEE (2012), also holds when the hidden information pertains to the loss distribution rather than the policyholder's risk attitude, as shown in GERSHKOV et al. (2023, Theorem 1).

The third result shows that the monopoly will absorb all surplus from the least risk-averse policyholder, and the contract offered to this policyholder leaves them indifferent between participation and non-participation. This property is also observed in CHADE and SCHLEE (2012, Theorem 1(iv)) and is particularly evident in BOONEN and ZHANG (2021), where the insurance market consists of only two types of policyholders with hidden risk attitude.

The fourth result elicits an interesting monotonicity property of the policyholder's utility of wealth at an optimum, as a function of their level of risk aversion, i.e., their risk type. As discussed earlier, if g_θ is convex in θ, then the utility of wealth of a type-θ policyholder in the absence of insurance (i.e., $U_\theta(X, 0)$) is decreasing and convex in θ. After trading, the policyholder's utility becomes $U_\theta(R_\theta, p_\theta)$, which decreasing and convex in θ, but attains a higher value than $U_\theta(X, 0)$, with equality holding for the type $\underline{\theta}$ policyholder. This echoes the finding of GERSHKOV et al. (2023, Theorem 1). In CHADE and SCHLEE (2012, Lemma 4), it is shown that the utility is a decreasing function of the risk type, but the convexity is not discussed.

The fifth result demonstrates that policyholders with a higher level of risk aversion, who are willing to pay more for a product with a higher coverage, are generally more valuable to the insurer, in that they induce a higher expected profit for the insurer. This contradicts the findings in scenarios with a hidden loss distribution, such as in Chade and Schlee (2012) and Gershkov et al. (2023), where the most profitable policyholders are always of intermediate, rather than extreme, types.

The final result asserts that the insurer is expected to earn a positive profit, making it profitable for the monopoly to design contracts even in the presence of asymmetric information. This finding aligns with established results on optimal insurance under adverse selection in a monopoly market, as in Chade and Schlee (2012, Theorem 1(vi)) and Gershkov et al. (2023, Lemma 2).

Remark 6. *It is worth noting that in both Gershkov et al. (2023, Proposition 2) and Chade and Schlee (2012, Theorem 2 (ii)), where the loss distribution is hidden from the insurer, the authors show that with an optimal menu that fully sorts policyholders, the insurer's expected profit increases as policyholders become more risk averse. However, in our setting, even with the optimal menu described in Theorem 4.4.1, the insurer's profit does not necessarily increase when the loss distribution shifts under first-order stochastic dominance, whether increasing or decreasing. Indeed, the profit with* (R^*_θ, p^*_θ) *can be rewritten as*

$$\begin{aligned}
\pi(R^*_\theta, p^*_\theta) &= \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \mathrm{d}x - \int_0^M \left(\int_{\underline{\theta}}^{\theta} \frac{\partial R^*_s(x)}{\partial x} \mathrm{d}g_s\left(F_X(x)\right) \right) \mathrm{d}x \\
&\quad - \int_0^M \left[1 - g_\theta\left(F_X(x)\right)\right] \frac{\partial R^*_\theta(x)}{\partial x} \mathrm{d}x - \int_0^M \left[1 - F_X(x)\right] \left(1 - \frac{\partial R^*_\theta(x)}{\partial x}\right) \mathrm{d}x \\
&= \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \mathrm{d}x - \int_0^M \left[1 - g_\theta\left(F_X(x)\right)\right] \frac{\partial R^*_\theta(x)}{\partial x} \mathrm{d}x - \int_0^M \left[1 - F_X(x)\right] \left(1 - \frac{\partial R^*_\theta(x)}{\partial x}\right) \mathrm{d}x \\
&\quad - \int_0^M \left(\frac{\partial R^*_\theta(x)}{\partial x} g_\theta\left(F_X(x)\right) - \frac{\partial R^*_{\underline{\theta}}(x)}{\partial x} g_{\underline{\theta}}\left(F_X(x)\right) - \int_{\underline{\theta}}^{\theta} g_s\left(F_X(x)\right) \frac{\partial}{\partial s} \frac{\partial R^*_s(x)}{\partial x} \mathrm{d}s \right) \mathrm{d}x \\
&= - \int_0^M \left[1 - F_X(x)\right] \left(1 - \frac{\partial R^*_\theta(x)}{\partial x}\right) \mathrm{d}x + \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \left(1 - \frac{\partial R^*_{\underline{\theta}}(x)}{\partial x}\right) \mathrm{d}x \\
&\quad + \int_0^M \left(\frac{\partial R^*_{\underline{\theta}}(x)}{\partial x} - \frac{\partial R^*_\theta(x)}{\partial x} \right) \mathrm{d}x + \int_0^M \int_{\underline{\theta}}^{\theta} g_s\left(F_X(x)\right) \frac{\partial}{\partial s} \frac{\partial R^*_s(x)}{\partial x} \mathrm{d}s\mathrm{d}x.
\end{aligned}$$

The first term, $\int_0^M \left[1 - F_X(x)\right] \left(1 - \frac{\partial R^*_\theta(x)}{\partial x}\right) \mathrm{d}x$, *represents the expected coverage of the type-*θ *policyholder. This term increases as* F_X *increases in the sense of first-order stochastic dominance, which subsequently leads to a decrease in the insurer's profit. The remaining three items represent the premium charged to the policyholders, and the result shows that the endogenous premium* p_θ *increases when* F_X *becomes more risky, leading to an increase* *in the insurer's profit. Combining these two effects, one cannot determine the direction of change in the insurer's total profit for this type of policyholder, nor can one conclusively determine the resulting expected profit.*

4.4.5 Special Cases

Another implication of Theorem 4.4.1 is that when $J_\theta(x)$ is monotone in θ, the contract menu consists of layered options, with any additional marginal loss either fully borne by the policyholder or entirely transferred to the insurer. We now examine two special cases where the layered contracts are among the most popular types of insurance contracts: the deductible contract menu and the coverage limit contract.

Theorem 4.4.2. *Suppose that $J_\theta(x)$, as defined in (4.11), is non-decreasing with respect to θ, for all x.*

1. *If for each $\theta < \overline{\theta}$ there exists a unique $x_\theta \in [0, M]$ such that $J_\theta(x) \leq 0$ for $x \leq x_\theta$ and $J_\theta(x) \geq 0$ for $x \geq x_\theta$, then the optimal contract is of the deductible type, with the menu given by $(\max(X, x_\theta), p_\theta)_{\theta\in\Theta}$.*
2. *If for each $\theta < \overline{\theta}$ there exists a unique $x_\theta \in [0, M]$ such that $J_\theta(x) \geq 0$ for $x \leq x_\theta$ and $J_\theta(x) \leq 0$ for $x \geq x_\theta$, then the optimal contract is of coverage limit type, with the menu given by $((X - x_\theta)_+, p_\theta)_{\theta\in\Theta}$.*

Proof. By Theorem 4.4.1, for any $\theta < \overline{\theta}$, we have that an optimal solution satisfies $\frac{\partial R_\theta^*(x)}{\partial x} = 0$ for $J_\theta(x) \geq 0$, and $\frac{\partial R_\theta^*(x)}{\partial x} = 1$ for $J_\theta(x) < 0$. When there exists a unique $x_\theta \in [0, M]$ such that $J_\theta(x) \leq 0$ for $x \leq x_\theta$ and $J_\theta(x) \geq 0$ for $x \geq x_\theta$, we obtain

$$R_\theta^*(x) = \begin{cases} x, & \textit{if} \quad x \leq x_\theta, \\ x_\theta, & \textit{if} \quad x > x_\theta. \end{cases}$$

In contract, if for each $\theta < \overline{\theta}$, there exists a unique $x_\theta \in [0, M]$ such that $J_\theta(x) \geq 0$ for $x \leq x_\theta$ and $J_\theta(x) \leq 0$ for $x \geq x_\theta$, we obtain

$$R_\theta^*(x) = \begin{cases} 0, & \textit{if} \quad x \leq x_\theta, \\ x - x_\theta, & \textit{if} \quad x > x_\theta. \end{cases}$$

□

The above results imply that if $J_\theta(x)$ crosses the zero line from below for all θ, then a deductible menu will be offered to the policyholders. Furthermore, since Theorem 4.4.1 establishes that the solution is submodular, the deductible level x_θ decreases as θ increases. Conversely, if $J_\theta(x)$ crosses the zero line from above for all θ, then a coverage limit contract is provided, and the limit x_θ increases with θ. The following are two examples that illustrate the optimality of deductible contracts and coverage limit contracts.

Example 4.4.1 (Optimality of a deductible menu)**.** *Suppose that each g_θ belongs to the family of inverse S-shaped distortion (ISSD) functions introduced in* Xu *and* Zhou *(2013):*

$$g_\theta(s) = \begin{cases} (1-\theta)s - (1-\theta)s^2, & \text{if } 0 \le s \le \frac{1}{2}, \\ (3+\theta)s^2 - (3+\theta)s + 1, & \text{if } \frac{1}{2} \le s \le 1. \end{cases}$$

Suppose that θ follows a uniform distribution on $[0,1]$. Then

$$\begin{aligned} J_\theta(x) &= F_X(x) - g_\theta\left(F_X(x)\right) + (1-\theta)\frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta} \\ &= \begin{cases} \theta F_X(x) + (1-\theta)F_X^2(x) + (1-\theta)\left(F_X^2(x) - F_X(x)\right), & \text{if } 0 \le F_X(x) \le \frac{1}{2}, \\ (4+\theta)F_X(x) - (3+\theta)F_X^2(x) - 1 + (1-\theta)\left(F_X^2(x) - F_X(x)\right), & \text{if } \frac{1}{2} \le F_X(x) \le 1. \end{cases} \\ &= \begin{cases} F_X(x)\left((2-2\theta)F_X(x) + (2\theta-1)\right), & \text{if } 0 \le F_X(x) \le \frac{1}{2}, \\ (-2-2\theta)\,F_X^2(x) + (3+2\theta)\,F_X(x) - 1, & \text{if } \frac{1}{2} \le F_X(x) \le 1. \end{cases} \end{aligned}$$

For a given $\theta \in \left[0, \frac{1}{2}\right)$, it is evident that $J_\theta(0) = J_\theta(M) = 0$. When $l \in \left(0, F_X^{-1}\left(\frac{1}{2}\right)\right]$, $(2-2\theta)\,F_X(x) + (2\theta - 1)$ is an increasing function of x for a given θ, and it crosses zero exactly once. Denote the root by x_θ. For $x \in \left(F_X^{-1}\left(\frac{1}{2}\right), M\right)$, the function $(-2-2\theta)\,F_X^2(x) + (3+2\theta)\,F_X(x) - 1$ is a downward-opening quadratic function of $F_X(x)$. This quadratic function has two roots, $F_X^{-1}\left(\frac{1}{2\theta+2}\right)$ and $F_X^{-1}(1)$, but neither lies within the interval. Thus, there is no root in this range. Consequently, $J_\theta(x) \le 0$ when $x < x_\theta$, and $J_\theta(x) \ge 0$ when $x \ge x_\theta$. By Theorem 4.4.2, the optimal contract is a deductible, with the deductible level given by

$$x_\theta = F_X^{-1}\left(\frac{1-2\theta}{2-2\theta}\right).$$

If $\theta \ge \frac{1}{2}$, $J_\theta(x)$ remains non-negative for all x, implying that $x_\theta = 0$. In this case, full insurance is optimal.

The above example examines the scenario where the policyholder of type θ uses a ISSD function g_θ, and is therefore mostly concerned with the tails of the loss distribution, that is relatively small and relatively large losses. The results show that for policyholders who are less risk averse ($\theta < \frac{1}{2}$), partial insurance will be provided in the form of a deductible contract. For policyholders who are more risk averse ($\theta \geq \frac{1}{2}$), full insurance is provided under each contract, leading to a form of pooling at the top (pooling with full insurance).

Example 4.4.2 (Optimality of a coverage limit menu)**.** *Suppose that* $g_\theta(s) = \frac{s}{1+\theta(1-s)}$*, where* $\theta \in [0,1]$*. Suppose further that* θ *follows a non-decreasing hazard rate distribution. We obtain that*

$$\begin{aligned} J_\theta(x) &= F_X(x) - g_\theta\left(F_X(x)\right) + \frac{\overline{F}(\theta)}{f(\theta)} \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta} \\ &= \frac{\theta F_X(x)\left(1 - F_X(x)\right)}{1 + \theta\left(1 - F_X(x)\right)} - \frac{\overline{F}(\theta)}{f(\theta)} \frac{F_X(x)\left(1 - F_X(x)\right)}{\left(1 + \theta\left(1 - F_X(x)\right)\right)^2} \\ &= \left(\theta - \frac{\overline{F}(\theta)}{f(\theta)} \frac{1}{1 + \theta\left(1 - F_X(x)\right)}\right) \frac{F_X(x) - F_X^2(x)}{1 + \theta\left(1 - F_X(x)\right)}. \end{aligned}$$

For a fixed θ*,* $J_\theta(x)$ *always crosses zero from above since the function* $\frac{F_X(x)-F_X^2(x)}{1+\theta(1-F_X(x))}$ *remains non-negative on* $[0,M]$*, and* $\theta - \frac{\overline{F}(\theta)}{f(\theta)}\frac{1}{1+\theta(1-F_X(x))}$ *is a decreasing function of* x*. This implies that the optimal menu is a coverage limit contract by Theorem 4.4.2. If* $\theta < \frac{1}{1+\theta}\frac{\overline{F}(\theta)}{f(\theta)}$*, then* $J_\theta(x) \leq 0$*, and full retention (no coverage) is optimal, with* $x_\theta = 0$*. If* $\frac{1}{1+\theta}\frac{\overline{F}(\theta)}{f(\theta)} \leq \theta < \frac{\overline{F}(\theta)}{f(\theta)}$*, then the coverage limit is given by*

$$x_\theta = F_X^{-1}\left(1 + \frac{1}{\theta} - \frac{\overline{F}(\theta)}{\theta^2 f(\theta)}\right).$$

If $\theta > \frac{\overline{F}(\theta)}{f(\theta)}$*,* $J_\theta(x) \geq 0$ *and thus* $x_\theta = M$*, or full insurance is optimal.*

If the policyholder places greater weight on higher levels of loss, as illustrated in Example 4.4.2 with a convex distortion, the coverage limit contract becomes optimal. The results indicate that no insurance is provided to the least risk-averse policyholder, full insurance is provided to the most risk-averse policyholder, and partial insurance is offered to those in between. This example illustrates that pooling occurs at both the bottom (pooling with no insurance) and the top when the policyholder's risk attitude is hidden information. In this case, coverage denial may occur, and typically, only a few contracts are offered, most of which providing full insurance.

These two types of contracts are also analyzed in Gershkov et al. (2023, Theorem 3). The authors argue that, for any given expected indemnification, a deductible contract is the most preferable option for a highly risk-averse policyholder, whereas a coverage limit contract is favored when it must satisfy the IC condition. In Gershkov et al. (2023, Example 2), where hidden information arises from the probability of a loss, the deductible menu is selected, with pooling observed both at the bottom and at the top. In Gershkov et al. (2023, Example 3), where hidden information pertains to the loss size, the coverage limit menu emerges as the optimal choice.

To conclude this section, we have shown that in a monopoly market without hidden risk attitudes, full insurance maximizes the insurer's profit, with the premium set to bind the policyholder's welfare constraint. However, when the policyholder's risk attitude is hidden from the insurer, the profit-maximizing, IR, and IC menu of contracts consists of a set of layered contracts. This menu yields a strictly positive profit for the monopoly, and, under these contracts, less risk-averse policyholders bear a larger portion of the incurred loss at a relatively lower price, while more risk-averse policyholders retain a smaller portion of the loss and pay a higher premium.

4.5 The Effect of Friction Costs on Equilibria

Thus far, the monopoly market that we consider did not account for any friction costs incurred by the insurer. The cost of insurance was simply the expected value of the indemnity payment, and the revenue to the insurer was the insurance premium collected from the policyholder. As a result, the insurer's profit was the premium payment less the expected indemnity payment.

In reality, insurers often incur additional "friction costs", such as claims processing costs, administrative expenses, actuarial loading, etc. Moreover, those costs can have a significant impact on the shape of contracts offered. This was shown, for instance, by Raviv (1979), Huberman et al. (1983), and Carlier and Dana (2003), in a setting of perfect information.

4.5.1 Equilibria with Costs and Hidden Risk Attitudes

In this section, we extend our model to account for such costs. Specifically, we consider the case of a fixed cost that occurs only when insurance is provided. We use the same cost function as the one proposed in Chade and Schlee (2020), which incorporates a fixed friction cost in addition to the indemnification cost. This friction cost arises whenever insurance is provided, i.e., $R_\theta(x) \not\equiv x$. Let the friction cost be denoted by $k > 0$. Notably, this cost does not alter any fundamental properties of the IR or IC menu; rather, it only impacts the insurer's profit to some extent. Consequently, Proposition 4.3.2(1), Proposition 4.3.3, Proposition 4.3.4, and Corollary 4.3.1 still hold. The insurer's profit from offering a contract (R_θ, p_θ) becomes

$$\widetilde{\pi}\left(R_\theta, p_\theta\right) = \pi\left(R_\theta, p_\theta\right) - k \cdot \mathbb{1}_{\{R_\theta(x) \not\equiv x\}}, \tag{4.22}$$

where π is defined in (4.6) and p_θ satisfies (4.9). Therefore, the expected profit for the insurer is given by

$$\widetilde{V}\left(\left(R_\theta, p_\theta\right)_{\theta\in\Theta}\right) = \int_\Theta \widetilde{\pi}\left(R_\theta, p_\theta\right) \mathrm{d}F(\theta), \tag{4.23}$$

and the profit-maximizing, IR, and IC menus are obtained by solving the following problem.

Problem 3.

$$(R^*_\theta, p^*_\theta)_{\theta\in\Theta} \in \underset{(R_\theta,p_\theta)_{\theta\in\Theta}\in\mathcal{IR}\cap\mathcal{IC}}{\arg\max} \widetilde{V}\left(\left(R_\theta, p_\theta\right)_{\theta\in\Theta}\right).$$

We solve this problem using a similar approach to that presented in Theorem 4.4.1. First, we derive the expression for the insurer's profit with an IC menu, similar to Proposition 4.3.2(2). The result is summarized below.

Proposition 4.5.1. *If $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IC}$, then the insurer's expected profit, as defined in (4.23), is given by*

$$\widetilde{V}\left(\left(R_\theta, p_\theta\right)_{\theta\in\Theta}\right) = V\left(\left(R_\theta, p_\theta\right)_{\theta\in\Theta}\right) - k \cdot \mathbb{P}\left(\theta \notin \Theta_R\right), \tag{4.24}$$

where $\Theta_R := \{\theta \in \Theta \mid R_\theta(x) = x \text{ for all } x \in [0, M]\}$, and $V\left(\left(R_\theta, p_\theta\right)_{\theta\in\Theta}\right)$ is defined in (4.10).

Proof. For any IR contract, if $R_\theta(x) = x$ for all $x \in [0, M]$, we know that $p_\theta = 0$, and thus $\pi(R_\theta, p_\theta) = 0$. Let Θ_R denote the set of θ values for which $R_\theta(x) = x$ for all $x \in [0, M]$, i.e., $\Theta_R = \{\theta \in \Theta \mid R_\theta(x) = x \text{ for } x \in [0, M]\}$. By the definition of $\widetilde{\pi}$ in (4.22), we have

$$\begin{aligned}\widetilde{V}\left(\left(R_\theta, p_\theta\right)_{\theta\in\Theta}\right) &= \int_{\Theta/\Theta_R} \left(\pi(R_\theta, p_\theta) - k\right) \mathrm{d}F(\theta) \\ &= \int_\Theta \left(\pi(R_\theta, p_\theta) - k\right) \mathrm{d}F(\theta) - \int_{\Theta_R} \left(\pi(R_\theta, p_\theta) - k\right) \mathrm{d}F(\theta)\end{aligned}$$

$$
\begin{aligned}
&= V\left((R_\theta, p_\theta)_{\theta\in\Theta}\right) - k - \int_{\Theta_R} (0-k)\,\mathrm{d}F(\theta) \\
&= V\left((R_\theta, p_\theta)_{\theta\in\Theta}\right) - k + k\cdot \int_{\theta} \mathbb{1}_{\{\theta\in\Theta_R\}}\mathrm{d}F(\theta) \\
&= V\left((R_\theta, p_\theta)_{\theta\in\Theta}\right) - k\cdot \mathbb{P}\left(\theta\notin\Theta_R\right).
\end{aligned}
$$

□

Subsequently, replacing the expected cost in Problem 3 by the expression obtained in (4.24), we find the optimal insurance structure by pointwise maximization. The result is stated in the following theorem.

Theorem 4.5.1. *Suppose that $J_\theta(x)$ given in (4.11) is non-decreasing in θ, for all x. Then,*

1. *If $k > \int_0^M \left[F_X(x) - g_{\overline{\theta}}\left(F_X(x)\right)\right]\mathrm{d}x$, then coverage is denied.*
2. *Otherwise, Problem 3 has a solution $(R_\theta^*, p_\theta^*)_{\theta\in\Theta} \in \mathcal{IR}\cap\mathcal{IC}$, where $R_\theta^*(x)\equiv x$ when $\theta<\theta^*$, and (R_θ^*, p_θ^*) is the same as the solution given in Theorem 4.4.1 when $\theta\geq\theta^*$, where,*
$$
\theta^* := \max\left\{\theta\in\Theta \;\middle|\; \int_0^M J_\theta(x)\mathbb{1}_{\{J_\theta(x)>0\}}\mathrm{d}x \leq k\right\}.
$$
Moreover, this menu is more profitable than a pooling equilibrium contract.

Proof. By Proposition 4.5.1(1), the menu that maximizes the insurer profit when cost is included in (4.23) also solves the following problem

$$
\max_{R\in\mathcal{R}_L} \quad \int_0^M \left[1-g_{\underline{\theta}}(F_X(x))\right]\mathrm{d}x - \mathbb{E}[X] - \int_{\underline{\theta}}^{\overline{\theta}} \left(\int_0^M J_\theta(x)\frac{\partial R_\theta(x)}{\partial x}\mathrm{d}x\right) f(\theta)\mathrm{d}\theta - k + k\cdot\mathbb{P}\left(\theta\in\Theta_R\right). \tag{4.25}
$$

We solve problem (4.25) pointwise. Specifically, for a fixed $\theta\in\Theta$, we look for R_θ^* such that

$$
R_\theta^* \in \underset{R\in\mathcal{R}_L}{\arg\max} \quad -\left(\int_0^M J_\theta(x)\frac{\partial R_\theta(x)}{\partial x}\mathrm{d}x\right) + k\cdot\mathbb{1}_{\{\theta\in\Theta_R\}},
$$

or

$$
R_\theta^* \in \underset{R\in\mathcal{R}_L}{\arg\max} \quad -\left(\int_0^M J_\theta(x)\frac{\partial R_\theta(x)}{\partial x}\mathrm{d}x\right) + k\cdot\mathbb{1}_{\{R_\theta(x)\equiv x\}}.
$$

If $\mathbb{1}_{\{R_\theta(x)\equiv x\}}=0$, at the optimum, we know that $\frac{\partial R^*_\theta(x)}{\partial x}=1$ when $J_\theta(x)<0$, and $\frac{\partial R^*_\theta(x)}{\partial x}=0$ when $J_\theta(x)>0$. Hence the total profit is $-\int_0^M J_\theta(x)\mathbb{1}_{\{J_\theta(x)<0\}}\mathrm{d}x$. If $\mathbb{1}_{\{R_\theta(x)\equiv x\}}=1$, the profit is $-\int_0^M J_\theta(x)\mathrm{d}x+k$.

If $-\int_0^M J_\theta(x)\mathbb{1}_{\{J_\theta(x)<0\}}\mathrm{d}x \geq -\int_0^M J_\theta(x)\mathrm{d}x+k$, or equivalently, $k\leq\int_0^M J_\theta(x)\mathbb{1}_{\{J_\theta(x)>0\}}\mathrm{d}x$, then it remains profitable to provide some insurance even with a fixed cost. However, if $k>\int_0^M J_\theta(x)\mathbb{1}_{\{J_\theta(x)>0\}}\mathrm{d}x$, the fixed cost exceeds the benefits collected, resulting in no insurance being provided. In this case, the optimal retention function is given by $R^*_\theta(x)=x$, for all x.

Additionally, the integral $\int_0^M J_\theta(x)\mathbb{1}_{\{J_\theta(x)>0\}}\mathrm{d}x$ is an increasing function of θ. This is because, for $\theta_1<\theta_2$, we know that $J_{\theta_1}(x)\leq J_{\theta_2}(x)$ for any $x\in[0,M]$, by assumption. Thus, $\mathbb{1}_{\left\{J_{\theta_1}(x)>0\right\}}\leq\mathbb{1}_{\left\{J_{\theta_2}(x)>0\right\}}$. Then

$$\int_0^M J_{\theta_1}(x)\mathbb{1}_{\left\{J_{\theta_1}(x)>0\right\}}\mathrm{d}x\leq\int_0^M J_{\theta_2}(x)\mathbb{1}_{\left\{J_{\theta_1}(x)>0\right\}}\mathrm{d}x\leq\int_0^M J_{\theta_2}(x)\mathbb{1}_{\left\{J_{\theta_2}(x)>0\right\}}\mathrm{d}x.$$

Thus:

1. If $\int_0^M J_{\overline{\theta}}(x)\mathbb{1}_{\left\{J_{\overline{\theta}}(x)>0\right\}}\mathrm{d}x=\int_0^M F_X(x)-g_{\overline{\theta}}\left(F_X(x)\right)\mathrm{d}x<k$, it is unprofitable for the insurer to trade with the highest-risk type, and thus it becomes unprofitable for all risk types. Consequently, the insurer chooses not to offer any contracts.

2. Otherwise, let $\theta^*=\max\left\{\theta\in\Theta\mid\int_0^M J_\theta(x)\mathbb{1}_{\{J_\theta(x)>0\}}\mathrm{d}x\leq k\right\}$. For any $\theta\leq\theta^*$, we have $R^*_\theta(x)=x$ for $x\in[0,M]$. When $\theta>\theta^*$, R^*_θ takes the form given in (4.17), which is submodular as shown by Theorem 4.4.1. Therefore, the solution to (4.25) is also submodular and thus satisfies the IC condition. The IR condition follows similarly to the proof of Theorem 4.4.1, and is omitted here.

Next, we compare the separating equilibrium contract with the pooling equilibrium. Let (R_p,p_p) denote the single contract offered to all policyholders. By Proposition 4.4.2, the maximum profit with a single contract is

$$\widetilde{V}\left(R^*_p,p^*_p\right)=\int_0^M\left[F_X(x)-g_{\underline{\theta}}\left(F_X(x)\right)\right]\mathrm{d}x-k.$$

For the separating menu of contracts, the insurer's profit at an optimum is

$$\widetilde{V}\left((R^*_\theta,p^*_\theta)_{\theta\in\Theta}\right)=\int_0^M\left[F_X(x)-g_{\underline{\theta}}\left(F_X(x)\right)\right]\mathrm{d}x-\int_{\underline{\theta}}^{\overline{\theta}}\left(\int_0^M J_\theta(x)\mathbb{1}_{\{J_\theta(x)<0\}}\mathrm{d}x\right)f(\theta)\,\mathrm{d}\theta-k\overline{F}\left(\theta^*\right).$$

Therefore,

$$\widetilde{V}\left((R_\theta^*, p_\theta^*)_{\theta\in\Theta}\right) - \widetilde{V}\left(R_p^*, p_p^*\right) = -\int_{\underline{\theta}}^{\overline{\theta}} \left(\int_0^M J_\theta(x)\mathbb{1}_{\{J_\theta(x)<0\}}\mathrm{d}x\right) f(\theta)\,\mathrm{d}\theta + kF\left(\theta^*\right) \geq 0, \tag{4.26}$$

which demonstrates the optimality of separating insurance. □

Note that the cost threshold $\int_0^M \left[F_X(x) - g_{\overline{\theta}}\left(F_X(x)\right)\right]\mathrm{d}x$ is the risk premium associated with the wealth of the type $\overline{\theta}$ policyholder (the most risk-averse type) in the absence of insurance. Indeed,

$$\begin{aligned}\int_0^M \left[F_X(x) - g_{\overline{\theta}}\left(F_X(x)\right)\right]\mathrm{d}x &= -\int_0^M \left[1 - F_X(x)\right]\mathrm{d}x + \int_0^M \left[1 - g_{\overline{\theta}}(F_X(x))\right]\mathrm{d}x \\ &= \mathbb{E}[-X] - U_{\overline{\theta}}(X,0) = \mathbb{E}[-X] - DU_{\overline{\theta}}(-X) \\ &= \Delta_{\overline{\theta}}(-X)\ (>0).\end{aligned}$$

This risk premium is non-decreasing with θ, by Assumption 4.2.2: for $\theta_1, \theta_2 \in \Theta$ such that $\theta_1 \leq \theta_2$, $g_{\theta_1}(s) \geq g_{\theta_2}(s)$, for all $s \in [0,1]$, i.e., θ_2 is more risk averse than θ_1. Consequently,

$$\int_0^M \left[F_X(x) - g_\theta\left(F_X(x)\right)\right]\mathrm{d}x \leq \int_0^M \left[F_X(x) - g_{\overline{\theta}}\left(F_X(x)\right)\right]\mathrm{d}x, \quad \forall\,\theta \in \Theta. \tag{4.27}$$

In addition, $\int_0^M \left(F_X(x) - g_{\overline{\theta}}\left(F_X(x)\right)\right)\mathrm{d}x$ is the expected profit of the insurer when there is no hidden information (with $\Theta = \{\overline{\theta}\}$) and no friction costs. Indeed, by Proposition 4.4.1, the optimal contract in that case is

$$(R_{\overline{\theta}}^*, p_{\overline{\theta}}^*) = \left(0, \int_0^M \left[1 - g_{\overline{\theta}}\left(F_X(x)\right)\right]\mathrm{d}x\right),$$

which leads to an expected profit of

$$\pi(R_{\overline{\theta}}^*, p_{\overline{\theta}}^*) = \int_0^M \left[1 - g_{\overline{\theta}}\left(F_X(x)\right)\right]\mathrm{d}x - \mathbb{E}[X] + \int_0^M \left[1 - F_X(x)\right]\,(R_{\overline{\theta}}^*)'(x)\,\mathrm{d}x$$

$$
\begin{aligned}
&= \int_0^M \left[1 - g_{\overline{\theta}}\left(F_X(x)\right)\right] \mathrm{d}x - \int_0^M \left[1 - F_X(x)\right] \, \mathrm{d}x \\
&= \int_0^M \left[F_X(x) - g_{\overline{\theta}}\left(F_X(x)\right)\right] \mathrm{d}x.
\end{aligned}
$$

Theorem 4.5.1 states that if the fixed cost k exceeds the profit that the insurer could have obtained in a market with perfect information and no friction cost, and in which the policyholder is of the most risk-averse type $\overline{\theta}$, then the insurer anticipates losses even on the most risk-averse policyholders (and hence on all other types θ, by (4.27)). In that case, there will be no gains from trade, and the insurer is not willing to offer any contract. A similar conclusion is reached in Chade and Schlee (2020, Theorem 1), which states that if there are no gains from trade at a given belief (the distribution of loss probabilities), then there are no gains from trade at any belief that likelihood-ratio dominates it (i.e., is more risky in the sense of the likelihood ratio).

On the other hand, when the fixed cost k is less than this threshold

$$
\int_0^M \left[F_X(x) - g_{\overline{\theta}}\left(F_X(x)\right)\right] \mathrm{d}x,
$$

Theorem 4.5.1 shows that the introduction of the fixed cost does not alter the equilibrium derived in Theorem 4.4.1, where no friction costs were present, with the exception that certain contracts are excluded. Indeed, for risk types $\theta < \theta^*$, the equilibrium indemnification provides for no insurance coverage (full loss retention). This excluded part of the no-cost equilibrium of Theorem 4.4.1 can be interpreted as a pooling at no-coverage. Note that this exclusion occurs from below, meaning that the contracts removed are those designed for the least risk-averse policyholders, who are in fact less valuable for the insurer. A similar effect is discussed in Chade and Schlee (2020, Theorem 3 and Corollary 1), which demonstrates that when a fixed cost is introduced, only pooling contracts and no-coverage contracts remain. However, contrary to Chade and Schlee (2020), Theorem 4.5.1 suggest that when trade occurs, separating equilibrium contracts always outperform pooling equilibrium contracts, despite the inclusion of fixed costs.

4.5.2 Equilibria with Costs and Full Information

If the insurer is able to observe the risk type of the policyholder, say $\Theta = \{\theta_0\}$, then the introduction of the fixed cost $k > 0$ has important implications on the full-information optimal contract derived in Proposition 4.4.1.

Corollary 4.5.1. *If $\theta \equiv \theta_0$, then an optimal contract is given by the following:*

1. *If $k > \int_0^M [F_X(x) - g_{\theta_0}(F_X(x))]\, \mathrm{d}x$, then coverage is denied. That is,*

$$(R^*_{\theta_0}, p^*_{\theta_0}) = (X, 0).$$

2. *If $k \le \int_0^M [F_X(x) - g_{\theta_0}(F_X(x))]\, \mathrm{d}x$, then full insurance is optimal, at a premium that makes the policyholder indifferent between insurance and no-insurance. That is,*

$$(R^*_{\theta_0}, p^*_{\theta_0}) = \left(0, \int_0^M [1 - g_{\theta_0}(F_X(x))]\, \mathrm{d}x\right).$$

Corollary 4.5.1 suggests that if the fixed cost k exceeds the profit that the insurer could have obtained in the absence of friction costs, then the insurer anticipates losses and will not provide any coverage. However, when the fixed cost is lower than the threshold $\int_0^M [F_X(x) - g_{\theta_0}(F_X(x))]\, \mathrm{d}x$, the optimal coverage is similar to the no-cost full-information optimum of Proposition 4.4.1.

4.6 Pareto-Optimal Menus of Contracts

The concept of incentive-compatible Pareto-optimality (often referred to as incentive efficiency) in the context of adverse selection has been examined by Prescott and Townsend (1984), Crocker and Snow (1985), d'Aspremont et al. (1990), Jerez (2003), and Bisin and Gottardi (2006), for instance, but with finitely many types. To define incentive-efficient menus of contracts in a context of an arbitrary (unaccountably infinite) set of types Θ, we use the approach to perfect competition in large economies introduced by Aumann (1964), and in particular the notion of a continuum of policyholders, or small traders in the language of Aumann (1964) and Shitovitz (1973).

4.6.1 Incentive-Efficient Menus

Equip $\Theta = [\underline{\theta}, \overline{\theta}]$ with the Borel sigma-algebra $\Sigma_{\mathcal{B}}(\Theta)$ and Lebesgue measure $\mathcal{L}$. A probability measure η_1 on $(\Theta, \Sigma_{\mathcal{B}}(\Theta))$ is said to be absolutely continuous with respect to a probability measure η_2 on $(\Theta, \Sigma_{\mathcal{B}}(\Theta))$, if $\eta_2(B) = 0 \Longrightarrow \eta_1(B) = 0$, for each $B \in \Sigma_{\mathcal{B}}(\Theta)$. If η_1 is absolutely continuous with respect to η_2 and η_2 is absolutely continuous with respect to η_1, the two measures are said to be equivalent, and we write $\eta_1 \sim \eta_2$.

Let μ be a probability measure on $(\Theta, \Sigma_{\mathcal{B}}(\Theta))$ that represents the distribution of policyholder types in the market. Since $\mu(\Theta) = 1$ by normalization, one can think of the quantity $\mu(B)$ as representing the proportion of types in the set $B \in \Sigma_{\mathcal{B}}(\Theta)$ in the market. For each $\theta \in \Theta$, let $F(\theta) := \mu([\underline{\theta}, \theta])$ define the cumulative distribution function over types, and let $\overline{F}(\theta) := \mu\left((\theta, \overline{\theta}]\right)$ define the decumulative distribution function over types. We assume that μ is absolutely continuous with respect to $\mathcal{L}$, with Radon–Nikodym derivative given by the probability density function f of the distribution function F over types.

Definition 4.6.1. *A menu $(R^*_\theta, p^*_\theta)_{\theta\in\Theta} \in \mathcal{IR} \cap \mathcal{IC}$ is said to be incentive efficient, or incentive-Pareto-optimal (IPO), if there does not exist any menu $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IR} \cap \mathcal{IC}$ such that the following two conditions hold:*

1. *$U_\theta(R_\theta, p_\theta) \geq U_\theta(R^*_\theta, p^*_\theta)$, for all $\theta \in \Theta$, and $\int_\Theta \pi(R_\theta, p_\theta)\, \mathrm{d}F(\theta) \geq \int_\Theta \pi(R^*_\theta, p^*_\theta)\, \mathrm{d}F(\theta)$, with at least one strict inequality.*
2. *If $\int_\Theta \pi(R_\theta, p_\theta)\, \mathrm{d}F(\theta) = \int_\Theta \pi(R^*_\theta, p^*_\theta)\, \mathrm{d}F(\theta)$, then the set of types for which the inequalities are strict is non-null. That is,*

$$\mu\bigg(\Big\{\theta \in \Theta : U_\theta(R_\theta, p_\theta) > U_\theta(R^*_\theta, p^*_\theta)\Big\}\bigg) > 0.$$

We denote by $\mathcal{IPO} \subseteq \mathcal{IR} \cap \mathcal{IC}$ the set of all IPO menus.

Equivalently, one can drop condition (2) from the above definition, with the understanding that statements made for all θ are to be taken in the μ-a.s. sense, as is common in the literature on markets with a continuum of traders (e.g., the convention used in Aumann (1966) and the subsequent literature).

Proposition 4.6.1. *If there exists a probability measure η on $(\Theta, \Sigma_{\mathcal{B}}(\Theta))$ such that μ is absolutely continuous with respect to η, and some $\alpha \in (0, 1)$, such that $(R^*_{\eta,\alpha,\theta}, p^*_{\eta,\alpha,\theta})_{\theta\in\Theta}$ is optimal for the problem*

$$\sup_{(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IR}\cap\mathcal{IC}} \left[\alpha \int_\Theta U_\theta(R_\theta, p_\theta)\, \mathrm{d}\eta + (1-\alpha) \int_\Theta \pi(R_\theta, p_\theta)\, \mathrm{d}\mu\right], \tag{4.28}$$

then $(R^*_{\eta,\alpha,\theta}, p^*_{\eta,\alpha,\theta})_{\theta\in\Theta} \in \mathcal{IPO}$.

Proof. Suppose that there exists a probability measure η on $(\Theta, \Sigma_{\mathcal{B}}(\Theta))$ such that μ is absolutely continuous with respect to η, and some $\alpha \in (0,1)$, such that $(R^*_\theta, p^*_\theta)_{\theta\in\Theta}$ is optimal for (4.28), but that $(R^*_\theta, p^*_\theta)_{\theta\in\Theta} \notin \mathcal{IPO}$. Then there exists $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IR} \cap \mathcal{IC}$ such that

$$U_\theta\left(R_\theta, p_\theta\right) \geq U_\theta\left(R^*_\theta, p^*_\theta\right), \ \forall\, \theta \in \Theta, \quad \text{and} \quad \int_\Theta \pi\left(R_\theta, p_\theta\right) \mathrm{d}\mu \geq \int_\Theta \pi\left(R^*_\theta, p^*_\theta\right) \mathrm{d}\mu,$$

with at least one strict inequality. Moreover, if $\int_\Theta \pi\left(R_\theta, p_\theta\right) \mathrm{d}\mu = \int_\Theta \pi\left(R^*_\theta, p^*_\theta\right) \mathrm{d}\mu$, then

$$\mu\Bigg(\Big\{\theta \in \Theta : U_\theta\left(R_\theta, p_\theta\right) > U_\theta\left(R^*_\theta, p^*_\theta\right)\Big\}\Bigg) > 0,$$

and hence

$$\eta\Bigg(\Big\{\theta \in \Theta : U_\theta\left(R_\theta, p_\theta\right) > U_\theta\left(R^*_\theta, p^*_\theta\right)\Big\}\Bigg) > 0.$$

Therefore,

$$\alpha \int_\Theta U_\theta\left(R_\theta, p_\theta\right)\, \mathrm{d}\eta + (1-\alpha) \int_\Theta \pi\left(R_\theta, p_\theta\right)\, \mathrm{d}\mu > \alpha \int_\Theta U_\theta\left(R^*_\theta, p^*_\theta\right)\, \mathrm{d}\eta + (1-\alpha) \int_\Theta \pi\left(R^*_\theta, p^*_\theta\right)\, \mathrm{d}\mu,$$

contradicting the optimality of $(R^*_\theta, p^*_\theta)_{\theta\in\Theta}$ for (4.28). Hence, $(R^*_\theta, p^*_\theta)_{\theta\in\Theta} \in \mathcal{IPO}$. □

4.6.2 Characterizing Incentive-Efficient Menus

By Proposition 4.6.1, we focus on maximizing the social welfare function

$$W_{\eta,\alpha}\left((R_\theta, p_\theta)_{\theta\in\Theta}\right) := \alpha \int_\Theta U_\theta\left(R_\theta, p_\theta\right)\, \mathrm{d}\eta + (1-\alpha) \int_\Theta \pi\left(R_\theta, p_\theta\right)\, \mathrm{d}\mu, \tag{4.29}$$

for a given probability measure η on $(\Theta, \Sigma_{\mathcal{B}}(\Theta))$, and a given weight $\alpha \in (0,1)$:

Problem 4.

$$\sup_{(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IR}\cap\mathcal{IC}} W_{\eta,\alpha}\left((R_\theta, p_\theta)_{\theta\in\Theta}\right). \tag{4.30}$$

Solutions to Problem 4 are difficult to characterize in general. Hereafter, we make some assumptions on η that allow us to provide a crisp characterization of optima. First, let F_η and $\overline{F}_\eta$ denote, respectively, the cumulative distribution function and survival function over types induced by η, that is, $F_\eta(\theta) := \eta([\underline{\theta}, \theta])$ and $\overline{F}_\eta(\theta) := \eta((\theta, \overline{\theta}])$, for all $\theta \in \Theta$. We will assume that η and μ are equivalent, which ensures the existence of a probability density function f_η for F_η, and that μ is smaller than η in the hazard rate order:

Assumption 4.6.1. *The measure η satisfies the following:*

1. $\eta \sim \mu$.
2. $\frac{f(\theta)}{\overline{F}(\theta)} \geq \frac{f_\eta(\theta)}{\overline{F}_\eta(\theta)}$, *for all* $\theta \in \Theta$.

Remark 7. *Assumption 4.6.1(2) implies that* $\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}$ *is non-decreasing in* $\theta \in \Theta$. *Indeed,*

$$\left(\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}\right)' = \frac{-f_\eta(\theta)\overline{F}(\theta) + \overline{F}_\eta(\theta) f(\theta)}{\overline{F}(\theta)^2} = \frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)} \left(\frac{f(\theta)}{\overline{F}(\theta)} - \frac{f_\eta(\theta)}{\overline{F}_\eta(\theta)}\right) \geq 0, \quad \forall \theta \in \Theta.$$

Moreover, since $\left.\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}\right|_{\theta=\underline{\theta}} = 1$, *and*

$$\left.\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}\right|_{\theta=\overline{\theta}} = \lim_{\theta \to \overline{\theta}} \frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)} = \lim_{\theta \to \overline{\theta}} \frac{-f_\eta(\theta)}{-f(\theta)} = \frac{f_\eta(\overline{\theta})}{f(\overline{\theta})},$$

where the second equality is by L'Hospital's rule, it follows that $\frac{f_\eta(\overline{\theta})}{f(\overline{\theta})} \geq 1$, *or* $f_\eta(\overline{\theta}) \geq f(\overline{\theta})$, *implying that* $\frac{f(\overline{\theta})}{f(\overline{\theta}) + f_\eta(\overline{\theta})} \leq \frac{1}{2}$.

Under Assumption 4.6.1, the following result provides a characterization of solutions of Problem 4.

Theorem 4.6.1. *Under Assumption 4.6.1, an optimal solution to Problem 4 is characterized as follows:*

1. *If* $\alpha \in \left(0, \frac{f(\overline{\theta})}{f(\overline{\theta}) + f_\eta(\overline{\theta})}\right)$ *and if the function* $J_{\eta,\alpha,\theta}(x)$ *given in (4.32) is non-decreasing in* θ *for all* l, *then there exists a solution* $\left(R^*_{\eta,\alpha,\theta}, p^*_{\eta,\alpha,\theta}\right)_{\theta \in \Theta}$. *For each* $\theta \in \Theta$, *the optimal retention function is such that the marginal retention satisfies the following:*

$$\frac{\partial R^*_{\eta,\alpha,\theta}(x)}{\partial x} = \begin{cases} 0, & J_{\eta,\alpha,\theta}(x) > 0, \\ \in [0,1], & J_{\eta,\alpha,\theta}(x) = 0, \\ 1, & J_{\eta,\alpha,\theta}(x) < 0, \end{cases} \tag{4.31}$$

and the premia $\left\{p^*_{\eta,\alpha,\theta}\right\}_{\theta \in \Theta}$ *satisfy (4.18). The function* $J_{\eta,\alpha,\theta}(x)$ *is defined as*

$$J_{\eta,\alpha,\theta}(x) := (1-\alpha)\left[F_X(x) - g_\theta\left(F_X(x)\right)\right] + \frac{\overline{F}(\theta)}{f(\theta)}\left(1 - \alpha - \alpha\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}\right)\frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta}. \tag{4.32}$$

2. *If* $\alpha \in \left[\frac{f(\overline{\theta})}{f(\overline{\theta})+f_\eta(\overline{\theta})}, \frac{1}{2}\right]$ *and if the function* $J_{\eta,\alpha,\theta}(x)$ *given in* (4.32) *is non-decreasing in* θ *for all* x *when* $\theta < \theta_\alpha$*, then* $R^*_{\eta,\alpha,\theta}$ *follows the form given in* (4.31) *for* $\theta < \theta_\alpha$*, and* $R^*_{\eta,\alpha,\theta} = 0$ *for* $\theta \geq \theta_\alpha$*, where* θ_α *is determined by the equation* $\frac{\overline{F}_\eta(\theta_\alpha)}{\overline{F}(\theta_\alpha)} = \frac{1-\alpha}{\alpha}$*. The premia* $\left\{p^*_{\eta,\alpha,\theta}\right\}_{\theta\in\Theta}$ *satisfy* (4.18).

3. *If* $\alpha \in \left(\frac{1}{2}, 1\right)$*, then* $R^*_{\eta,\alpha,\theta} = 0$ *and* $p^*_{\eta,\alpha,\theta} = 0$*, for all* $\theta \in \Theta$*.*

Moreover, the collection $\left\{R^*_{\eta,\alpha,\theta}\right\}_{\theta\in\Theta}$ *of optimal retention functions is submodular, for a given* η *and* α*.*

Proof. Form (4.12) and (4.6), it follows that the social welfare function can be rewritten as

$$\begin{aligned}
W_{\eta,\alpha}\left((R_\theta, p_\theta)_{\theta\in\Theta}\right) &= \alpha \int_{\underline{\theta}}^{\overline{\theta}} U_\theta\left(R_\theta, p_\theta\right) \mathrm{d}F_\eta(\theta) + (1-\alpha) \int_{\underline{\theta}}^{\overline{\theta}} \pi\left(R_\theta, p_\theta\right) \mathrm{d}F(\theta) \\
&= \alpha \int_{\underline{\theta}}^{\overline{\theta}} \left(- p_\theta - \int_0^M \left[1 - g_\theta(F_X(x))\right] \frac{\partial R_\theta(x)}{\partial x} \mathrm{d}x \right) \mathrm{d}F_\eta(\theta) \\
&\qquad + (1-\alpha) \int_{\underline{\theta}}^{\overline{\theta}} \left(p_\theta - \mathbb{E}[X] + \int_0^M \left[1 - F_X(x)\right] \frac{\partial R_\theta(x)}{\partial x} \mathrm{d}x \right) \mathrm{d}F(\theta) \\
&= \int_{\underline{\theta}}^{\overline{\theta}} \int_0^M \left((1-\alpha)[1 - F_X(x)] - \alpha \frac{\mathrm{d}F_\eta(\theta)}{\mathrm{d}F(\theta)} [1 - g_\theta(F_X(x))] \right) \frac{\partial R_\theta(x)}{\partial x} \mathrm{d}x \, \mathrm{d}F(\theta) \\
&\qquad + \int_{\underline{\theta}}^{\overline{\theta}} \left(1 - \alpha - \alpha \frac{\mathrm{d}F_\eta(\theta)}{\mathrm{d}F(\theta)} \right) p_\theta \, \mathrm{d}F(\theta) - (1-\alpha)\mathbb{E}[X].
\end{aligned}$$

Now, for every $(R_\theta, p_\theta)_{\theta\in\Theta} \in \mathcal{IC}$, it follows from Proposition 4.3.2 that the premium satisfies (4.9). Substituting this premium structure into the social welfare function yields

$$\begin{aligned}
W_{\eta,\alpha}\left((R_\theta, p_\theta)_{\theta\in\Theta}\right) &= \int_{\underline{\theta}}^{\overline{\theta}} \int_0^M \left((1-\alpha)[1 - F_X(x)] - \alpha \frac{\mathrm{d}F_\eta(\theta)}{\mathrm{d}F(\theta)} [1 - g_\theta(F_X(x))] \right) \frac{\partial R_\theta(x)}{\partial x} \mathrm{d}x \, \mathrm{d}F(\theta) \\
&+ \int_{\underline{\theta}}^{\overline{\theta}} \left(1 - \alpha - \alpha \frac{\mathrm{d}F_\eta(\theta)}{\mathrm{d}F(\theta)} \right) \left(p_{\underline{\theta}} + \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x} \mathrm{d}x \right. \\
&- \left. \int_0^M \left[1 - g_\theta\left(F_X(x)\right)\right] \frac{\partial R_\theta(x)}{\partial x} \mathrm{d}x \right) \mathrm{d}F(\theta) \\
&- \int_{\underline{\theta}}^{\overline{\theta}} \left(1 - \alpha - \alpha \frac{\mathrm{d}F_\eta(\theta)}{\mathrm{d}F(\theta)} \right) \left(\int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s} \frac{\partial R_s(x)}{\partial x} \mathrm{d}x \mathrm{d}s \right) \mathrm{d}F(\theta) - (1-\alpha)\mathbb{E}[X]
\end{aligned}$$

$$
\begin{aligned}
&= \left(p_{\underline{\theta}} + \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\mathrm{d}x\right) \int_{\underline{\theta}}^{\overline{\theta}} \left(1 - \alpha - \alpha \frac{\mathrm{d}F_\eta(\theta)}{\mathrm{d}F(\theta)}\right) \mathrm{d}F(\theta) - (1-\alpha)\mathbb{E}[X] \\
&- (1-\alpha) \int_{\underline{\theta}}^{\overline{\theta}} \int_0^M \left[F_X(x) - g_\theta\left(F_X(x)\right)\right] \frac{\partial R_\theta(x)}{\partial x}\mathrm{d}x\,\mathrm{d}F(\theta) - (1-\alpha)\cdot \\
&\quad \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s} \frac{\partial R_s(x)}{\partial x}\mathrm{d}x\mathrm{d}s\,\mathrm{d}F(\theta) + \alpha \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s} \frac{\partial R_s(x)}{\partial x}\mathrm{d}x\mathrm{d}s\,\mathrm{d}F_\eta(\theta) \\
&= (1-2\alpha)\left(p_{\underline{\theta}} + \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\mathrm{d}x\right) - (1-\alpha)\mathbb{E}[X] - (1-\alpha)\cdot \\
&\quad \left(\int_{\underline{\theta}}^{\overline{\theta}} \int_0^M \left[F_X(x) - g_\theta\left(F_X(x)\right)\right] \frac{\partial R_\theta(x)}{\partial x}\mathrm{d}x\,\mathrm{d}F(\theta) + \int_{\underline{\theta}}^{\overline{\theta}} \int_0^M \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta} \frac{\partial R_\theta(x)}{\partial x}\mathrm{d}x\,\overline{F}(\theta)\mathrm{d}\theta\right) \\
&+ (1-\alpha)\left(\int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s} \frac{\partial R_s(x)}{\partial x}\mathrm{d}x\mathrm{d}s\right) \overline{F}(\theta)\Big|_{\theta=\underline{\theta}}^{\theta=\overline{\theta}} - \alpha\cdot \\
&\quad \left(\int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s} \frac{\partial R_s(x)}{\partial x}\mathrm{d}x\mathrm{d}s\right) \overline{F}_\eta(\theta)\Big|_{\theta=\underline{\theta}}^{\theta=\overline{\theta}} + \alpha \int_{\underline{\theta}}^{\overline{\theta}} \int_0^M \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta} \frac{\partial R_\theta(x)}{\partial x}\mathrm{d}x\,\overline{F}_\eta(\theta)\mathrm{d}\theta \\
&= (1-2\alpha)\left(p_{\underline{\theta}} + \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\mathrm{d}x\right) - (1-\alpha)\mathbb{E}[X] \\
&- \int_{\underline{\theta}}^{\overline{\theta}} \int_0^M J_{\eta,\alpha,\theta}(x) \frac{\partial R_\theta(x)}{\partial x}\mathrm{d}x f(\theta)\,\mathrm{d}\theta, \qquad (4.33)
\end{aligned}
$$

where

$$
\begin{aligned}
J_{\eta,\alpha,\theta}(x) &:= (1-\alpha)\left[F_X(x) - g_\theta\left(F_X(x)\right)\right] + \frac{(1-\alpha)\overline{F}(\theta) - \alpha\overline{F}_\eta(\theta)}{f(\theta)} \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta} \\
&= (1-\alpha)\left[F_X(x) - g_\theta\left(F_X(x)\right)\right] + \frac{\overline{F}(\theta)}{f(\theta)}\left(1 - \alpha - \alpha\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}\right) \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta}. \qquad (4.34)
\end{aligned}
$$

1. If $\alpha \leq \frac{1}{2}$, the social welfare function (4.33) is a non-decreasing function of $p_{\underline{\theta}}$. Therefore, at the optimum, $p_{\underline{\theta}}$ must take its largest value, provided the IR condition is still satisfied. By Proposition 4.3.1, we conclude that $p^*_{\underline{\theta}} = \int_0^M \left[1 - g_{\underline{\theta}}(F_X(x))\right]\left(1 - \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\right)\mathrm{d}x$, where $R_{\underline{\theta}}$ is the retention function for the lowest type policyholder. With $p_{\underline{\theta}} := \int_0^M \left[1 - g_{\underline{\theta}}(F_X(x))\right]\left(1 - \frac{\partial R_{\underline{\theta}}(x)}{\partial x}\right)\mathrm{d}x$, the social welfare function simplifies to:

$$
(1-2\alpha)\int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right]\mathrm{d}x - (1-\alpha)\mathbb{E}[X] - \int_{\underline{\theta}}^{\overline{\theta}} \left(\int_0^M J_{\eta,\alpha,\theta}(x)\frac{\partial R_\theta(x)}{\partial x}\mathrm{d}x\right) f(\theta)\mathrm{d}\theta.
$$

We maximize this function pointwise, following the approach in Theorem 4.4.1. First, we analyze the function $J_{\eta,\alpha,\theta}(x)$ in (4.34). The first term, $(1-\alpha)\left[F_X(x) - g_\theta\left(F_X(x)\right)\right] \geq 0$ by Assumption 4.2.1. For the second term, it is evident that $\frac{\overline{F}(\theta)}{f(\theta)} \frac{\partial g_\theta(F_X(x))}{\partial \theta} \leq 0$, since g_θ satisfies Assumption 4.2.2. For the remaining part, we observe that $\left(1-\alpha-\alpha\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}\right)\Big|_{\theta=\underline{\theta}} = 1-2\alpha$. Furthermore, at $\theta=\overline{\theta}$, using Remark 7, we obtain

$$\left(1-\alpha-\alpha\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}\right)\Bigg|_{\theta=\overline{\theta}} = 1-\alpha-\alpha\frac{f_\eta(\overline{\theta})}{f(\overline{\theta})}.$$

Since $\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}$ is non-decreasing in θ, it follows that the function

$$\theta \mapsto 1-\alpha-\alpha\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}$$

is non-increasing. Therefore, if $1-\alpha-\alpha\frac{f_\eta(\overline{\theta})}{f(\overline{\theta})} > 0$, or equivalently, $\alpha < \frac{f(\overline{\theta})}{f_\eta(\overline{\theta})+f(\overline{\theta})}$, then

$$\frac{\overline{F}(\theta)}{f(\theta)}\left(1-\alpha-\alpha\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}\right)\frac{\partial g_\theta\left(F_X(x)\right)}{\partial\theta} \leq 0.$$

For a fixed $\theta \in \Theta$, the optimal retention solves the following maximization problem:

$$R^*_{\eta,\alpha,\theta} \in \underset{R_{\eta,\alpha,\theta}\in\mathcal{R}_L}{\arg\max} \; -\int_0^M J_{\eta,\alpha,\theta}(x)\frac{\partial R_{\eta,\alpha,\theta}(x)}{\partial x}\mathrm{d}x.$$

Hence, the optimal retention function satisfies (4.31). When α and η are specified, for any $\theta < \theta'$, and if $J_{\eta,\alpha,\theta}(x)$ is non-decreasing in θ for all x, it follows that $J_{\eta,\alpha,\theta}(x) \leq J_{\eta,\alpha,\theta'}(x)$. Thus, the pointwise maximization solution satisfies

$$\frac{\partial R^*_{\eta,\alpha,\theta'}(x)}{\partial x} \leq \frac{\partial R^*_{\eta,\alpha,\theta}(x)}{\partial x}$$

for all x, implying that $R^*_{\eta,\alpha,\theta}(x)$ is submodular. Therefore, by Proposition 4.3.3, it satisfies the IC condition. Additionally, the optimal contract for the lowest risk type $(R^*_{\eta,\alpha,\underline{\theta}}, p^*_{\eta,\alpha,\underline{\theta}})$ satisfies Corollary 4.3.1, ensuring that the menu is IR.

2. If $1-2\alpha \geq 0$ and $1-\alpha-\alpha\frac{f_\eta(\bar{\theta})}{f(\bar{\theta})} \leq 0$, or equivalently, $\alpha \in \left[\frac{f(\bar{\theta})}{f_\eta(\bar{\theta})+f(\bar{\theta})}, \frac{1}{2}\right]$, then there exists $\theta_\alpha \in \Theta$ such that $1-\alpha-\alpha\frac{\overline{F}_\eta(\theta_0)}{\overline{F}(\theta_0)} = 0$. Additionally, $1-\alpha-\alpha\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)} \geq 0$ when $\theta < \theta_\alpha$, and $1-\alpha-\alpha\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)} \leq 0$ when $\theta > \theta_\alpha$. Therefore, if the function $J_{\eta,\alpha,\theta}(x)$ is non-decreasing in θ when $\theta < \theta_\alpha$, the optimal retention function $R^*_{\eta,\alpha,\theta}$ follows the form given in (4.31). When $\theta \geq \theta_\alpha$, since $J_{\eta,\alpha,\theta}(x) \geq 0$ for all l, it follows that $R^*_{\eta,\alpha,\theta} = 0$.

3. If $\alpha > \frac{1}{2}$, the social welfare function (4.33) is a non-increasing function of $p_{\eta,\alpha,\underline{\theta}}$. Therefore, at the optimum, $p_{\eta,\alpha,\underline{\theta}}$ must take its smallest value while still satisfying the IR condition. By Proposition 4.3.1, we obtain $p^*_{\eta,\alpha,\underline{\theta}} = 0$. Substituting this into the social welfare function (4.33), we obtain

$$\begin{aligned}
&(1-2\alpha)\int_0^M \left[1-g_{\underline{\theta}}\left(F_X(x)\right)\right]\frac{\partial R_{\eta,\alpha,\underline{\theta}}(x)}{\partial x}\mathrm{d}x - (1-\alpha)\mathbb{E}[X] \\
&-\int_{\underline{\theta}}^{\overline{\theta}}\left(\int_0^M J_{\theta,\alpha,\eta}(x)\frac{\partial R_{\eta,\alpha,\theta}(x)}{\partial x}\mathrm{d}x\right)f(\theta)\mathrm{d}\theta \\
&=\int_{\underline{\theta}}^{\overline{\theta}}\int_0^M \left((1-2\alpha)\left[1-g_{\underline{\theta}}\left(F_X(x)\right)\right]\frac{\partial R_{\eta,\alpha,\underline{\theta}}(x)}{\partial x} - J_{\eta,\alpha,\theta}(x)\frac{\partial R_{\eta,\alpha,\theta}(x)}{\partial x}\right)\mathrm{d}x f(\theta)\mathrm{d}\theta \\
&-(1-\alpha)\mathbb{E}[X].
\end{aligned}$$

In this case, $1-\alpha-\alpha\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)} < 0$, for all θ. Consequently, $J_{\eta,\alpha,\theta}(x) \geq 0$, for all θ. For $\theta = \underline{\theta}$, we define

$$\widetilde{J}_{\eta,\alpha,\underline{\theta}}(x) := -(1-2\alpha)\left[1-g_{\underline{\theta}}\left(F_X(x)\right)\right] + J_{\eta,\alpha,\underline{\theta}}(x) \geq J_{\eta,\alpha,\underline{\theta}}(x) \geq 0.$$

The optimal retention function for the lowest risk type satisfies

$$R^*_{\eta,\alpha,\underline{\theta}} \in \underset{R_{\eta,\alpha,\underline{\theta}}\in\mathcal{R}_L}{\arg\max}\ -\int_0^M \widetilde{J}_{\eta,\alpha,\underline{\theta}}(x)\frac{\partial R_{\eta,\alpha,\underline{\theta}}(x)}{\partial x}\mathrm{d}x.$$

For all other risk types θ, the optimal retention function satisfies

$$R^*_{\eta,\alpha,\theta} \in \underset{R_{\eta,\alpha,\theta}\in\mathcal{R}_L}{\arg\max}\ -\int_0^M J_{\eta,\alpha,\theta}(x)\frac{\partial R_{\eta,\alpha,\theta}(x)}{\partial x}\mathrm{d}x,$$

The pointwise maximization implies that $\frac{\partial R^*_{\eta,\alpha,\theta}(x)}{\partial x} = 0$ for all x, and thus $R^*_{\eta,\alpha,\theta} \equiv 0$ for any $\theta \in \Theta$.

□

This result characterizes incentive-efficient contracts as the solutions to a social welfare maximization problem, where the policyholder's welfare is weighted by α, and η can be seen as the social planner's prior over unknown types. The structure of incentive-efficient contracts depends on the value of α. When $\alpha \in \left(0, \frac{f(\overline{\theta})}{f(\overline{\theta})+f_\eta(\overline{\theta})}\right)$, a separating equilibrium emerges, and the optimal retention function follows the submodular layered structure described in (4.31), ensuring that higher-risk types receive more coverage, similarly to Theorem 4.4.1. When $\alpha \in \left[\frac{f(\overline{\theta})}{f(\overline{\theta})+f_\eta(\overline{\theta})}, \frac{1}{2}\right]$, the equilibrium remains separating, but with a pooling segment at the top: full insurance is offered to all policyholders with type $\theta \geq \theta_\alpha$. Moreover, θ_α decreases with α since θ_α is the solution to $\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)} = \frac{1-\alpha}{\alpha}$, where the fraction $\frac{1-\alpha}{\alpha}$ decreases with α, while the ratio $\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}$ increases with θ. This implies that as α increases, more policyholder types become pooled at the top. When $\alpha > \frac{1}{2}$, a pooling equilibrium arises. That is, if a sufficiently high weight is placed on policyholder welfare, the previously separating equilibrium transitions into a pooling equilibrium, where full insurance is provided to all policyholders. This result is at odds with the findings in the base cases discussed in Sections 4.4 and 4.5, where a separating structure is always preferred over a pooling structure in terms of optimal contract design.

To achieve the separating layered equilibrium when $\alpha \leq \frac{1}{2}$, we impose a monotonicity condition on the function $J_{\eta,\alpha,\theta}$, analogous to the condition on J_θ in Theorem 4.4.1. These two functions are closely related, as expressed in the following relationship:

$$J_{\eta,\alpha,\theta}(x) = (1-\alpha)\, J_\theta(x) \;-\; \alpha \left(\frac{\overline{F}(\theta)}{f(\theta)}\right) \left(\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}\right) \frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta}. \tag{4.35}$$

In particular, when $\alpha = 0$, $J_{\eta,\alpha,\theta}(x)$ reduces to $J_\theta(x)$, recovering the standard case without explicit policyholder welfare considerations. The optimal retention function structure in (4.31) resembles the solution in (4.17), with the key distinction that the layer formation now depends on $J_{\eta,\alpha,\theta}(x)$ rather than $J_\theta(x)$. When this function is non-decreasing in θ for a given x, the existence of an optimal retention function is ensured. Furthermore, the condition guaranteeing the monotonicity of $J_\theta(x)$ in θ, as discussed in Section 4.4.3, is also sufficient to ensure the monotonicity of $J_{\eta,\alpha,\theta}(x)$ in θ. Specifically,

$$\begin{aligned}\frac{\partial J_{\eta,\alpha,\theta}(x)}{\partial \theta} = -\frac{\partial g_\theta\left(F_X(x)\right)}{\partial \theta} &\left((1-\alpha)\theta - \frac{\overline{F}(\theta)}{f(\theta)}\left(1-\alpha-\alpha\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}\right)\right)' \\ &+ \frac{\overline{F}(\theta)}{f(\theta)}\left(1-\alpha-\alpha\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}\right)\frac{\partial^2 g_\theta\left(F_X(x)\right)}{\partial \theta^2}.\end{aligned}$$

A non-decreasing hazard rate $f/\overline{F}$ under μ, the convexity of $\theta \mapsto g_\theta(s)$ for each $s \in (0,1)$, and Assumption 4.6.1 together ensure that $J_{\eta,\alpha,\theta}(x)$ is non-decreasing in θ for all x when $\alpha \in \left(0, \frac{f(\overline{\theta})}{f(\overline{\theta})+f_\eta(\overline{\theta})}\right)$, and $J_{\eta,\alpha,\theta}(x)$ is non-decreasing in θ when $\theta < \theta_\alpha$ for all x when $\alpha \in \left[\frac{f(\overline{\theta})}{f(\overline{\theta})+f_\eta(\overline{\theta})}, \frac{1}{2}\right]$. These results follow from the fact that $\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)}$ increases with θ and, in these cases, $1-\alpha-\alpha\frac{\overline{F}_\eta(\theta)}{\overline{F}(\theta)} \geq 0$.

The following proposition outlines key properties of the IPO menu established in Theorem 4.6.1.

Proposition 4.6.2. *Under the same assumption as Theorem 4.6.1, the following hold:*

1. *For each $\theta \in \Theta$, the marginal function $\frac{\partial R^*_{\eta,\alpha,\theta}(x)}{\partial x}$ decreases with α for all x. Consequently, the optimal retention function $R^*_{\eta,\alpha,\theta}(x)$ decreases with α.*
2. *For each $\theta \in \Theta$, $U_\theta(R^*_{\eta,\alpha,\theta}, p^*_{\eta,\alpha,\theta})$ increases with α.*
3. *The insurer's total profit $V\big((R^*_{\eta,\alpha,\theta}, p^*_{\eta,\alpha,\theta})_{\theta\in\Theta}\big)$ decreases with α. Moreover,*
$$V\big((R^*_{\eta,\alpha,\theta}, p^*_{\eta,\alpha,\theta})_{\theta\in\Theta}\big) > 0,$$
for $\alpha \in \left(0, \frac{1}{2}\right]$.

Proof. 1. By (4.35), if $J_{\eta,\alpha,\theta}(x) < 0$ then $J_\theta(x) < 0$, since $\frac{\partial g_\theta(F_X(x))}{\partial \theta} \leq 0$. Now, suppose there exist two weight parameters, α_1 and α_2 such that $\alpha_1 < \alpha_2$. Suppose, by way of contradiction, that
$$\frac{\partial R^*_{\eta,\alpha_1,\theta}(x)}{\partial x} < \frac{\partial R^*_{\eta,\alpha_2,\theta}(x)}{\partial x}$$
for some $x \in [0, M]$. Since the right-hand side of the inequality is at its maximum, it follows that $\frac{\partial R^*_{\eta,\alpha_2,\theta}(x)}{\partial x} = 1$. This implies that $J_{\eta,\alpha_2,\theta}(x) < 0$, which in turn leads to $J_\theta(x) < 0$. Given that $\alpha_1 < \alpha_2$, we obtain
$$(1-\alpha_1)J_\theta(x) < (1-\alpha_2)J_\theta(x),$$
which further implies that
$$J_{\eta,\alpha_1,\theta}(x) < J_{\eta,\alpha_2,\theta}(x) < 0.$$
Consequently, we must also have $\frac{\partial R^*_{\eta,\alpha_1,\theta}(x)}{\partial x} = 1$, contradicting our initial assumption. Therefore, we conclude that $\frac{\partial R^*_{\eta,\alpha_1,\theta}(x)}{\partial x} \geq \frac{\partial R^*_{\eta,\alpha_2,\theta}(x)}{\partial x}$ for all x.

2. The utility function for a type-θ policyholder under the contract $(R^*_{\eta,\alpha,\theta}, p^*_{\eta,\alpha,\theta})$ is given by

$$\begin{aligned}
U_\theta(R^*_{\eta,\alpha,\theta}, p^*_{\eta,\alpha,\theta}) &= -p^*_{\eta,\alpha,\theta} - \int_0^M \left[1 - g_\theta\left(F_X(x)\right)\right] \frac{\partial R^*_{\eta,\alpha,\theta}(x)}{\partial x} \mathrm{d}x \\
&= -p^*_{\eta,\alpha,\underline{\theta}} - \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \frac{\partial R^*_{\eta,\alpha,\underline{\theta}}(x)}{\partial x} \mathrm{d}x + \int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s} \frac{\partial R^*_{\eta,\alpha,s}(x)}{\partial x} \mathrm{d}x\mathrm{d}s \\
&= -\int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \mathrm{d}x + \int_{\underline{\theta}}^{\theta} \int_0^M \frac{\partial g_s\left(F_X(x)\right)}{\partial s} \frac{\partial R^*_{\eta,\alpha,s}(x)}{\partial x} \mathrm{d}x\mathrm{d}s. \qquad (4.36)
\end{aligned}$$

If $\alpha_1 < \alpha_2$, then

$$U_\theta(R^*_{\eta,\alpha_1,\theta}, p^*_{\eta,\alpha_1,\theta}) \le U_\theta(R^*_{\eta,\alpha_2,\theta}, p^*_{\eta,\alpha_2,\theta})$$

since $\frac{\partial g_s(F_X(x))}{\partial s} \le 0$, and, by Proposition 4.6.2(1), $\frac{\partial R^*_{\eta,\alpha_1,\theta}(x)}{\partial x} \ge \frac{\partial R^*_{\eta,\alpha_2,\theta}(x)}{\partial x}$. When $\alpha \ge \frac{1}{2}$, the utility function satisfies $U_\theta(R^*_{\eta,\alpha,\theta}, p^*_{\eta,\alpha,\theta}) = 0$, for every $\theta \in \Theta$.

3. For the insurer, when $\alpha \le \frac{1}{2}$, the premium schedule satisfies (4.9), and the total profit is given by

$$\begin{aligned}
&V\big((R^*_{\eta,\alpha,\theta}, p^*_{\eta,\alpha,\theta})_{\theta\in\Theta}\big) = p^*_{\eta,\alpha,\underline{\theta}} \\
&+ \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \frac{\partial R^*_{\eta,\alpha,\underline{\theta}}(x)}{\partial x} \mathrm{d}x - \mathbb{E}[X] - \int_{\underline{\theta}}^{\overline{\theta}} \left(\int_0^M J_\theta(x) \frac{\partial R^*_{\eta,\alpha,\theta}(x)}{\partial x} \mathrm{d}x \right) f(\theta)\mathrm{d}\theta \\
&= \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \mathrm{d}x - \mathbb{E}[X] - \int_{\underline{\theta}}^{\overline{\theta}} \left(\int_0^M J_\theta(x) \frac{\partial R^*_{\eta,\alpha,\theta}(x)}{\partial x} \mathrm{d}x \right) f(\theta)\mathrm{d}\theta. \qquad (4.37)
\end{aligned}$$

When $\alpha_1 < \alpha_2$, we have $\frac{\partial R^*_{\eta,\alpha_1,\theta}(x)}{\partial x} \ge \frac{\partial R^*_{\eta,\alpha_2,\theta}(x)}{\partial x}$ by Proposition 4.6.2(1). Since, at the optimum, $\frac{\partial R^*_{\eta,\alpha,\theta}(x)}{\partial x} = \{0, 1\}$, suppose that for some x,

$$1 = \frac{\partial R^*_{\eta,\alpha_1,\theta}(x)}{\partial x} > \frac{\partial R^*_{\eta,\alpha_2,\theta}(x)}{\partial x} = 0.$$

Then, we have $J_{\eta,\alpha_1,\theta}(x) < 0$, which, by (4.35), implies that $J_\theta(x) < 0$. Thus, it follows that

$$-\int_0^M \frac{\partial R^*_{\eta,\alpha_1,\theta}(x)}{\partial x} J_\theta(x)\mathrm{d}x \ge -\int_0^M \frac{\partial R^*_{\eta,\alpha_2,\theta}(x)}{\partial x} J_\theta(x)\mathrm{d}x.$$

Integrating over θ in Θ, we obtain

$$V\big((R^*_{\eta,\alpha_1,\theta}, p^*_{\eta,\alpha_1,\theta})_{\theta\in\Theta}\big) \ge V\big((R^*_{\eta,\alpha_2,\theta}, p^*_{\eta,\alpha_2,\theta})_{\theta\in\Theta}\big).$$

Additionally, from (4.37), we derive the bound

$$\begin{aligned} V\big((R^*_{\eta,\alpha,\theta}, p^*_{\eta,\alpha,\theta})_{\theta\in\Theta}\big) &\geq \int_0^M \left[1 - g_{\underline{\theta}}\left(F_X(x)\right)\right] \mathrm{d}x - \mathbb{E}[X] \\ &= \int_0^M \left[F_X(x) - g_{\underline{\theta}}\left(F_X(x)\right)\right] \mathrm{d}x > 0. \end{aligned}$$

When $\alpha > \frac{1}{2}$, the insurer's total profit is given by

$$\begin{aligned} V\big((R^*_{\eta,\alpha,\theta}, p^*_{\eta,\alpha,\theta})_{\theta\in\Theta}\big) &= \int_{\underline{\theta}}^{\overline{\theta}} \pi\left(R^*_{\eta,\alpha,\theta}, p^*_{\eta,\alpha,\theta}\right) f(\theta)\mathrm{d}\theta \\ &= \int_{\underline{\theta}}^{\overline{\theta}} \left(p^*_{\eta,\alpha,\theta} - \mathbb{E}\left[X - R^*_{\eta,\alpha,\theta}(X)\right]\right) f(\theta)\mathrm{d}\theta = -\mathbb{E}[X]. \end{aligned}$$

□

The first result shows that the IPO menu provides greater insurance coverage for each policyholder type, with the retained losses always decreasing as more weight is placed on policyholder welfare. Second, the incentive efficient menu, which differs in coverage and pricing structure from the profit-maximizing contract, benefits all policyholder types, with utility levels increasing monotonically with α. In contrast, the insurer's total profit decreases with α. The derived contract menu may become unprofitable for the insurer when policyholder welfare is heavily prioritized, particularly when $\alpha > \frac{1}{2}$. However, for $\alpha \leq \frac{1}{2}$, the insurer's profit remains strictly positive under this IPO menu.

4.6.3 Pareto Optima with Full Information

If the insurer is able to observe the risk type of the policyholder, say $\Theta = \{\theta_0\}$, then there is no adverse selection or screening problem. In this case, it is easy to verify that a contract $(R, p) \in \mathcal{R}_L \times \mathbb{R}_+$ is Pareto optimal if and only if there is some $\alpha \in [0, 1]$ such that (R, p) is optimal for the following problem:

Problem 5.

$$\sup_{(R,p)\in\mathcal{R}_L\times\mathbb{R}_+} \Big\{\alpha\, U_{\theta_0}(R,p) + (1-\alpha)\,\pi(R,p) : U_{\theta_0}(R,p) \geq U_{\theta_0}(X,0)\Big\}.$$

The following result shows that under complete information, full insurance is provided to the policyholder, as in Proposition 4.4.1, while the premium depends on the weight α.

Corollary 4.6.1. *If $\Theta = \{\theta_0\}$, an optimal solution to Problem 5 is given by the following:*

1. *If $\alpha \leq \frac{1}{2}$, then*
$$(R^*_{\theta_0}, p^*_{\theta_0}) = \left(0, \int_0^M \left[1 - g_{\theta_0}\left(F_X(x)\right)\right] \mathrm{d}x\right).$$

2. *If $\alpha > \frac{1}{2}$, then*
$$(R^*_{\theta_0}, p^*_{\theta_0}) = (0, 0).$$

Proof. If $\Theta = \{\theta_0\}$, the social welfare function is given by:

$$\begin{aligned}
&\alpha\, U_{\theta_0}(R, p) + (1-\alpha)\,\pi(R, p) \\
&= \alpha \left(-p_{\theta_0} - \int_0^M \left[1 - g_{\theta_0}\left(F_X(x)\right)\right] \frac{\partial R_{\theta_0}(x)}{\partial x} \mathrm{d}x\right) + (1-\alpha)\cdot \\
&\quad \left(p_{\theta_0} - \mathbb{E}[X] + \int_0^M \left[1 - F_X(x)\right] \frac{\partial R_{\theta_0}(x)}{\partial x} \mathrm{d}x\right) \\
&= (1-2\alpha) p_{\theta_0} + \int_0^M \left((1-\alpha)\left[1 - F_X(x)\right] - \alpha\left[1 - g_{\theta_0}\left(F_X(x)\right)\right]\right) \frac{\partial R_{\theta_0}(x)}{\partial x} \mathrm{d}x \\
&- (1-\alpha)\mathbb{E}[X].
\end{aligned} \tag{4.38}$$

1. If $\alpha \leq \frac{1}{2}$, we see that the function in (4.38) is a non-decreasing function of p_{θ_0}. By Proposition 4.3.1, we conclude that at the optimum, when $R_{\theta_0} \in \mathcal{R}_L$, the premium is
$$p^*_{\theta_0} = \int_0^M \left[1 - g_\theta(F_X(x))\right] \left(1 - \frac{\partial R_{\theta_0}(x)}{\partial x}\right) \mathrm{d}x.$$
Thus, the social welfare function becomes:
$$\begin{aligned}
&(1-2\alpha) \int_0^M \left[1 - g_\theta(F_X(x))\right] \left(1 - \frac{\partial R_{\theta_0}(x)}{\partial x}\right) \mathrm{d}x \\
&+ \int_0^M \left((1-\alpha)\left[1 - F_X(x)\right] - \alpha\left[1 - g_{\theta_0}\left(F_X(x)\right)\right]\right) \frac{\partial R_{\theta_0}(x)}{\partial x} \mathrm{d}x - (1-\alpha)\mathbb{E}[X] \\
&= (1-2\alpha) \int_0^M \left[1 - g_\theta(F_X(x))\right] \mathrm{d}x - (1-\alpha)\mathbb{E}[X] \\
&+ (1-\alpha) \int_0^M \left(\left[1 - F_X(x)\right] - \left[1 - g_{\theta_0}\left(F_X(x)\right)\right]\right) \frac{\partial R_{\theta_0}(x)}{\partial x} \mathrm{d}x
\end{aligned}$$

$$= (1-2\alpha)\int_0^M [1-g_\theta(F_X(x))]\,\mathrm{d}x + (1-\alpha)\int_0^M [g_{\theta_0}(F_X(x)) - F_X(x)]\frac{\partial R_{\theta_0}(x)}{\partial x}\mathrm{d}x$$
$$-(1-\alpha)\mathbb{E}[X].$$

The maximizer of this equation is $\frac{\partial R^*_{\theta_0}(x)}{\partial x} \equiv 0$, since $g_{\theta_0}(F_X(x)) - F_X(x) \leq 0$ for all x.

2. If $\alpha > \frac{1}{2}$, the function in (4.38) becomes a decreasing function of p_{θ_0}. Thus, at the optimum, the premium is $p^*_{\theta_0} = 0$. The social welfare function then becomes:

$$\int_0^M ((1-\alpha)[1-F_X(x)] - \alpha[1-g_{\theta_0}(F_X(x))])\frac{\partial R_{\theta_0}(x)}{\partial x}\mathrm{d}x - (1-\alpha)\mathbb{E}[X].$$

Since $1 - F_X(x) \leq 1 - g_{\theta_0}(F_X(x))$ for all x, and $1-\alpha < \alpha$, it follows that

$$(1-\alpha)[1-F_X(x)] \leq \alpha[1-g_{\theta_0}(F_X(x))].$$

Thus, the maximizer of this equation is again $\frac{\partial R^*_{\theta_0}(x)}{\partial x} \equiv 0$.

□

This result states that when α is small, the insurer sets the premium in a way to extract all of the policyholder surplus, yielding a strictly positive profit of $\Delta_{\theta_0}(-X)$. However, when policyholder welfare is prioritized (i.e., α is large), the premium is reduced, and the insurer charges a zero premium.

4.7 Conclusion

This chapter examines a monopolistic insurance market in which policyholders have private information about their risk attitudes, while the insurer observes only the overall risk distribution. This hidden information is modeled using the Dual Utility model of Yaari (1987), and equilibrium contracts are designed to maximize the insurer's profit while ensuring incentive compatibility—that is, revealing policyholders' private information. Our model extends the classical setting of Stiglitz (1977) by incorporating a continuum of types and assuming that losses are continuously distributed. Under certain conditions, a separating equilibrium is optimal and can be fully characterized. The continuum framework allows for a detailed analysis of key properties of the optimal contract, including policyholder welfare and insurer profit.

We characterize the optimal (profit-maximizing) incentive-compatible and individually rational menus of insurance contracts in terms of the marginal loss retention per type of policyholder. Optima consist of layered deductible contracts, where each such layered structure is determined by the risk type of the policyholder, proxied by their probability weighting function. Such layered indemnity structures are widely observed in practice. The key properties of our separating equilibria are that (i) insurance coverage and premia are monotone in the level of risk aversion; (ii) the most risk-averse policyholder receives full insurance (*efficiency at the top*); (iii) the monopoly will absorb all surplus from the least-risk averse policyholder, and the contract offered to this policyholder leaves them indifferent between participation and non-participation; (iv) at an optimum, the policyholder's utility of wealth is a decreasing function of their risk type; and (v) policyholders with a higher level of risk aversion, who are willing to pay more for a product with a higher coverage, are generally more valuable to the insurer, in that they induce a higher expected profit for the insurer.

In addition, we examine the effect of fixed insurance provision costs on equilibria. We show that the optimal menu is the same as in the absence of such fixed costs, with the exception that only part of the menu is excluded. The excluded contracts are those designed for policyholders with relatively lower risk aversion, who are less valuable to the insurer. Moreover, if the fixed cost is prohibitively high, exceeding the profit that the insurer could have obtained in a market with perfect information and no friction cost, and in which the policyholder is of the most risk-averse type, then the insurer anticipates losses even on the most risk-averse policyholders (and hence on all other types). In that case, there will be no gains from trade, and the insurer is not willing to offer any contract. However, when trade occurs, separating equilibrium contracts always outperform pooling equilibrium contracts.

Finally, we characterize incentive-efficient menus of contracts in the context of an arbitrary type space. We show that individually rational and incentive compatible contracts that are Pareto optimal can be achieved by maximizing a social welfare function that accounts for hidden types. While it is difficult to solve such a problem in the general case, we are able to provide a crisp characterization of solutions under a few assumptions.

Chapter 5

Future Work

This thesis examines optimal insurance design in monopolistic markets under both complete and incomplete information. We develop a unified framework that captures a wide range of policyholder behaviors and insurer strategies, driven by both classical expected utility theory and more flexible non-expected utility models such as Yaari's dual theory and distortion risk measures.

In the full-information setting, we first analyze optimal pricing through convex loading functions when the insurer fully observes the policyholder's risk preferences. We identify the conditions under which traditional contract forms—such as coinsurance and deductible contracts—emerge as optimal, and characterize their associated pricing rules. We then extend this model to account for behavioral distortions in probability perceptions, offering a broader class of pricing and contract design principles that align with both academic theory and empirical observations. This analysis highlights the central role of the Bowley solution in sequential insurance markets and demonstrates how insurer and policyholder distortions jointly shape contract structures.

In the incomplete information setting, we turn to mechanism design to address the challenge of adverse selection. Modeling policyholder heterogeneity through a continuum of distortion functions, we derive the optimal menu of contracts subject to incentive compatibility and individual rationality constraints. Our results reveal a layered structure of optimal contracts, where full insurance is offered to the most risk-averse types, and coverage and premiums increase with risk aversion. Extensions including fixed costs and incentive-efficient(Pareto-optimal) contract menus further deepen our understanding of practical trade-offs faced by insurers in markets with asymmetric information.

Looking ahead, several promising directions remain open for future research.

5.1 Extending to General Loading Functions

Recall from Chapter 2 that we initially considered a set of admissible loading functions given by:

$$\mathcal{A}_0 = \Bigg\{ g : X(\Omega) \to [0,\infty) \;\Bigg|\; g \text{ is increasing and convex, } g(0)=0,\ g(x)\geq x, \\ g(x) \not\equiv x,\ \mathbb{E}[g(X)] < +\infty \Bigg\}.$$

We now consider a broader class of loading functions:

$$\mathcal{F} = \Bigg\{ g : X(\Omega) \to [0,\infty) \;\Bigg|\; g(0)=0,\ |g'(x+) - g'(x-)| \leq M, \mathbb{E}[X] < \mathbb{E}[g(X)] < +\infty \Bigg\}$$

for some constant $M > 0$. This extension allows for more general pricing mechanisms that may include kinks or non-convexity.

To optimize g, we decompose it using:

$$g'(x) = g'(\alpha) + h(x),$$

and obtain

$$\mathbb{E}[g(\alpha X)] = \alpha \int_0^1 g'(\alpha x)\overline{F}_X(x)\mathrm{d}x = \alpha \int_0^1 h(\alpha x)\overline{F}_X(x)\mathrm{d}x + \alpha g'(\alpha)\mathbb{E}[X]. \tag{5.1}$$

For the policyholder's problem, the first-order condition for the optimal proportion α under a proportional contract—originally derived in equation (2.10)—still applies in this generalized setting. It becomes:

$$\begin{aligned} \frac{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - \mathbb{E}[g(\alpha X)]\right) X\right]}{\mathbb{E}\left[u'\left(w_0 - X + \alpha X - \mathbb{E}[g(\alpha X)]\right)\right]} &= \mathbb{E}\left[g'\left(\alpha X\right) X\right] \\ &= g'(\alpha)\mathbb{E}[X] + \mathbb{E}\left[h\left(\alpha X\right) X\right], \end{aligned} \tag{5.2}$$

For the insurer's problem, our goal is to determine the optimal pricing function $g \in \mathcal{F}$ that maximizes the insurer's expected profit, as defined in equation (2.13). Adopting a similar variational approach as in Theorem 2.3.2, we analyze the structure of g by decomposing it into its derivative components. This allows us to derive integral expressions for the insurer's profit and characterize the necessary conditions that an optimal pricing function must satisfy. Specifically, we approximate the derivative g' via a sequence of step functions and analyze the impact of marginal changes in g on expected profit under the constraint that $g \in \mathcal{F}$. The resulting inequalities yield useful bounds and ultimately guide the design of the profit-maximizing pricing rule within this broader admissible class. A complete solution to this generalized problem remains an open question and is left for future work.

5.2 Extension to the Two-Dimensional Hidden Information Case

Based on the progress made in addressing the asymmetric information problem in monopolistic markets, as discussed in Chapter 4, an interesting direction for future research is the extension of the current continuum-type model to a multidimensional hidden information framework within a monopolistic market. This extension would involve incorporating both hidden risk distributions and hidden risk attitudes.

The analysis of discrete-type policyholders, particularly in the two-type case, has been explored in the literature. For example, Landsberger and Meilijson (1994) focuses solely on hidden risk attitudes, while Landsberger and Meilijson (1999) incorporates both hidden risk distributions and attitudes. Their findings indicate that when policyholders differ only in their risk attitudes, the more risk-averse group is offered full insurance, while the less risk-averse type receives a two-valued partial insurance contract. When policyholders also differ in their risk distributions, the policyholder with the lower certainty equivalent is offered full insurance, while the other type receives partial insurance. The level of coverage in this case depends on the ratio of the two risk distribution functions.

An open question is whether these conclusions extend to the continuum-type setting. In the one-dimensional screening problem, transitioning from discrete to continuous types uncovers interesting structural properties of the contract menu, such as the convexity of policyholder welfare and the monotonicity of insurer profitability. However, when policyholders differ not only in risk attitudes but also in risk distributions, it remains unclear whether these key properties still hold. Moreover, it is important to examine whether the separating equilibrium continues to be more profitable and whether the equilibrium contract depends on the relative ordering of policyholders' risk distributions, as discussed in Landsberger and Meilijson (1999).

A very important contribution in this direction can be found in the recent work of Cheung et al. (2025). In this paper, the authors consider a setting where both the risk distribution and the risk attitude of policyholders are hidden information. policyholders are assumed to evaluate risk using VaR, while the premium is calculated using a distortion premium principle. The heterogeneity in private information is captured through a vector (α, k), where α reflects the policyholder's confidence level (VaR quantile), and k represents uncertainty in the risk distribution L_k. The paper derives optimal insurance contracts that are incentive-compatible, individually rational, and profit-maximizing, under specific contract structures—namely deductible, proportional, and combinations thereof. This paper serves as a valuable reference for future research in this direction.

References

ARROW, K. J. (1963). Uncertainty and the Welfare Economics of Medical Care. *American Economic Review*, 53(5):941–973.

ASSA, H. (2015). On Optimal Reinsurance Policy with Distortion Risk Measures and Premiums. *Insurance: Mathematics and Economics*, 61:70–75.

AUMANN, R. (1964). Markets with a Continuum of Traders. *Econometrica*, 32(1/2):39–50.

AUMANN, R. (1966). Existence of Competitive Equilibria in Markets with a Continuum of Traders. *Econometrica*, 34(1):1–17.

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.

BISIN, A. and GOTTARDI, P. (2006). Efficient Competitive Equilibria with Adverse Selection. *The Journal of Political Economy*, 114(3):485–516.

BOONEN, T., CHEUNG, K., and ZHANG, Y. (2021). Bowley Reinsurance with Asymmetric Information on the Insurer's Risk Preferences. *Scandinavian Actuarial Journal*, 2021(7):623–644.

BOONEN, T. J. and GHOSSOUB, M. (2021). Optimal Reinsurance with Multiple Reinsurers: Distortion Risk Measures, Distortion Premium Principles, and Heterogeneous Beliefs. *Insurance: Mathematics and Economics*, 101:23–37.

BOONEN, T. J. and GHOSSOUB, M. (2023). Bowley vs. Pareto Optima in Reinsurance Contracting. *European Journal of Operational Research*, 307(1):382–391.

BOONEN, T. J. and JIANG, W. (2024). Bowley Insurance with Expected Utility Maximization of the Policyholders. *North American actuarial journal*, 28(2):407–425.

Boonen, T. J. and Zhang, Y. (2021). Optimal Reinsurance Design with Distortion Risk Measures and Asymmetric Information. *ASTIN Bulletin*, 51(2):607–629.

Cai, J. and Chi, Y. (2020). Optimal Reinsurance Designs Based on Risk Measures: A Review. *Statistical theory and related fields*, 4(1):1–13.

Cai, J., Lemieux, C., and Liu, F. (2016). Optimal Reinsurance from the Perspectives of Both an Insurer and a Reinsurer. *ASTIN Bulletin*, 46(3):815–849.

Cai, J., Liu, H., and Wang, R. (2017). Pareto-Optimal Reinsurance Arrangements under General Model Settings. *Insurance: Mathematics and Economics*, 77:24–37.

Cai, J. and Tan, K. S. (2007). Optimal Retention for a Stop-Loss Reinsurance Under the VaR and CTE Risk Measures. *ASTIN Bulletin*, 37(1):93–112.

Cai, J., Tan, K. S., Weng, C., and Zhang, Y. (2008). Optimal Reinsurance under VaR and CTE Risk Measures. *Insurance: mathematics and Economics*, 43(1):185–196.

Cao, J., Li, D., Young, V., and Zou, B. (2022). Stackelberg Differential Game for Insurance under Model Ambiguity. *Insurance: Mathematics and Economics*, 106:128–145.

Cao, J., Li, D., Young, V. R., and Zou, B. (2023). Reinsurance Games with Two Reinsurers: Tree versus Chain. *European Journal of Operational Research*, 310(2):928–941.

Cao, J., Li, D., Young, V. R., and Zou, B. (2024). Short Communication: Optimal Insurance to Maximize Exponential Utility When Premium Is Computed by a Convex Functional. *SIAM Journal on Financial Mathematics*, 15(1):SC15–SC27.

Carlier, G. and Dana, R. (2003). Pareto Efficient Insurance Contracts when the Insurer's Cost Function is Discontinuous. *Economic Theory*, 21(4):871–893.

Carlier, G. and Dana, R. (2005). Rearrangement Inequalities in Non-convex Insurance Models. *Journal of Mathematical Economics*, 41(4-5):483–503.

Chade, H., Marone, V. R., Starc, A., and Swinkels, J. (2022). Multidimensional Screening and Menu Design in Health Insurance Markets. *NBER Working Paper Series*.

Chade, H. and Schlee, E. (2012). Optimal Insurance with Adverse Selection. *Theoretical Economics*, 7(3).

Chade, H. and Schlee, E. (2020). Insurance as a Lemons Market: Coverage Denials and Pooling. *Journal of Economic Theory*, 189:105085.

Chan, F.-Y. and Gerber, H. (1985). The Reinsurer's Monopoly and the Bowley Solution. *ASTIN Bulletin: The Journal of the IAA*, 15(2):141–148.

Chateauneuf, A. and Cohen, M. (1994). Risk Seeking with Diminishing Marginal Utility in a Non-Expected Utility Model. *Journal of Risk Uncertainty*, 9(1):77–91.

Chen, L. and Shen, Y. (2018). On a New Paradigm of Optimal Reinsurance: A Stochastic Stackelberg Differential Game between an Insurer and a Reinsurer. *ASTIN Bulletin*, 48(2, part 48):905–960.

Cheung, K. C., Yam, S., Yuen, F., and Zhang, Y. (2025). Optimal Design of Reinsurance Contracts under Adverse Selection with a Continuum of Types. *arXiv preprint arXiv:2504.17468*.

Cheung, K. C., Yam, S. C. P., and Zhang, Y. (2019). Risk-Adjusted Bowley Reinsurance under Distorted Probabilities. *Insurance: Mathematics and Economics*, 86:64–72.

Chi, Y. and Tan, K. S. (2013). Optimal Reinsurance with General Premium Principles. *Insurance: Mathematics and Economics*, 52(2):180–189.

Chi, Y., Tan, K. S., and Zhuang, S. C. (2020). A Bowley Solution with Limited Ceded Risk for a Monopolistic Reinsurer. *Insurance: Mathematics and Economics*, 91:188–201.

Chi, Y. and Zhuang, S. C. (2020). Optimal Insurance with Belief Heterogeneity and Incentive Compatibility. *Insurance: Mathematics and Economics*, 92:104–114.

Crocker, K. and Snow, A. (1985). The Efficiency of Competitive Equilibria in Insurance Markets with Asymmetric Information. *Journal of Public Economics*, 26(2):207–219.

d'Aspremont, C., Crémer, J., and Gérard-Varet, L.-A. (1990). Incentives and the Existence of Pareto-Optimal Revelation Mechanisms. *Journal of Economic Theory*, 51(2):233–254.

Denneberg, D. (1994). *Non-Additive Measure and Integral*. Kluwer Academic Publishers.

Deprez, O. and Gerber, H. U. (1985). On Convex Principles of Premium Calculation. *Insurance: Mathematics and Economics*, 4(3):179–189.

Dunford, N. and Schwartz, J. (1958). *Linear Operators, Part 1: General Theory*. Wiley-Interscience.

Föllmer, H. and Schied, A. (2011). *Stochastic Finance.* De Gruyter, Berlin, New York.

Gershkov, A., Moldovanu, B., Strack, P., and Zhang, M. (2023). Optimal Insurance: Dual Utility, Random Losses, and Adverse Selection. *American Economic Review*, 113(10):2581–2614.

Ghossoub, M. (2019). Optimal Insurance under Rank-Dependent Expected Utility. *Insurance: Mathematics and Economics*, 87:51–66.

Ghossoub, M. and He, X. (2021). Comparative Risk Aversion in RDEU with Applications to Optimal Underwriting of Securities Issuance. *Insurance Mathematics and Economics*, 101(1):6–22.

Ghossoub, M. and Zhu, M. B. (2024). Stackelberg Equilibria with Multiple Policyholders. *Insurance: Mathematics and Economics*, 116(1):189–201.

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

Jerez, B. (2003). A Dual Characterization of Incentive Efficiency. *Journal of Economic Theory*, 112(1):1–34.

Landsberger, M. and Meilijson, I. (1994). Monopoly Insurance under Adverse Selection When Agents Differ in Risk Aversion. *Journal of Economic Theory*, 63(2):392–407.

Landsberger, M. and Meilijson, I. (1999). A General Model of Insurance under Adverse Selection. *Economic Theory*, 14:331–352.

Li, D. and Young, V. R. (2021). Bowley Solution of a Mean–Mariance Game in Insurance. *Insurance: Mathematics and Economics*, 98:35–43.

Milgrom, P. and Segal, I. (2002). Envelope Theorems for Arbitrary Choice Sets. *Econometrica*, 70(2):583–601.

Mossin, J. (1968). Aspects of Rational Insurance Purchasing. *Journal of political economy*, 76(4, Part 1):553–568.

Prescott, E. C. and Townsend, R. M. (1984). Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard. *Econometrica*, 52(1):21–45.

Quiggin, J. (1993). *Generalized Expected Utility Theory - The Rank-Dependent Model.* Kluwer Academic Publishers.

Raviv, A. (1979). The Design of an Optimal Insurance Policy. *The American Economic Review*, 69(1):84–96.

Rothschild, M. and Stiglitz, J. (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics*, 90(4):629–649.

Shitovitz, B. (1973). Oligopoly in Markets with a Continuum of Traders. *Econometrica*, 41(3):467–501.

Stiglitz, J. (1977). Monopoly, Non-Linear Pricing and Imperfect Information: The Insurance Market. *The Review of Economic Studies*, 44(3):407–430.

Kaluszka, M. (2001). Optimal Reinsurance under Mean-Variance Premium Principles. *Insurance: Mathematics and Economics*, 28(1):61–67.

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.

Wang, S. (1996). Premium Calculation by Transforming the Layer Premium Density. *ASTIN Bulletin*, 26(1):71–92.

Wang, S. S., Young, V. R., and Panjer, H. H. (1997). Axiomatic Characterization of Insurance Prices. *Insurance: Mathematics and Economics*, 21(2):173–183.

Xu, Z. Q. and Zhou, X. Y. (2013). Optimal Stopping under Probability Distortion. *The Annals of Applied Probability*, 23(1):251–282.

Yaari, M. (1987). The Dual Theory of Choice under Risk. *Econometrica*, 55(1):95–115.

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

Young, V. R. and Browne, M. J. (2000). Equilibrium in Competitive Insurance Markets Under Adverse Selection and Yaari's Dual Theory of Risk. *The Geneva Papers on Risk and Insurance Theory*, 25(2):141–157.

Zhu, M. B., Ghossoub, M., and Boonen, T. J. (2023). Equilibria and Efficiency in a Reinsurance Market. *Insurance: Mathematics and Economics*, 113:24–49.