University of Toronto
Department of Economics

Working Paper 815

Optimal Insurance with Information Asymmetry: Nonlinear and Linear Pricing

By Xia Han, Bin Li and Yao Luo

January 22, 2026

Optimal Insurance with Information Asymmetry: Nonlinear and Linear Pricing

Xia Han* Bin Li† Yao Luo‡

January 22, 2026

Abstract

We propose a new framework for studying optimal insurance under information asymmetry within the Stackelberg game framework. In this setting, a monopolistic insurer faces uncertainty regarding a customer's loss distribution or risk attitude. The customer is assumed to follow a mean–variance preference in continuous time, while the insurer sets premiums through a risk loading based on the expected loss. An optimal menu is explicitly derived for a general class of aggregate loss models.

Our approach connects with the extensive literature on optimal insurance demand, stemming from the seminal work of Arrow (1963), and leads to an interesting finding: a nonlinear pricing structure for risk-type uncertainty versus a linear pricing structure for risk-attitude uncertainty. Specifically, if an insurer is uncertain about a customer's risk type and seeks to elicit this information, the risk loading (premium minus expected loss) is set lower for high-risk individuals to encourage them to select the corresponding contract. In contrast, if the insurer is only uncertain about the customer's risk attitude, no such discounts—in terms of risk loading—are provided. This reveals that information about customers' risk types is more valuable than information about their risk attitudes. Additionally, we compare our optimal menu with the worst-case contract derived from the maxmin expected utility, we find that our optimal menu increases the insurer's expected profit and enhances the likelihood of trading.

Keywords: Optimal insurance; Information asymmetry; Stackelberg game framework; Risk loading; Nonlinear pricing; Linear pricing; Ambiguity.

1 Introduction

1.1 Background and Motivation

Information asymmetry between insurers and customers is arguably the most fundamental problem in insurance. It is widely recognized that customers hold private information about the risks they intend to insure, such as their lifestyle with health insurance, driving habits with car insurance, or a firm owner's commitment with business interruption insurance. Although modern insurers increasingly rely on observable data—such as driving records, claims history, or credit scores—to approximate risk types, substantial information gaps persist. For example, new drivers or first-time customers lack historical data, complicating risk assessment. In the context of health insurance,

*School of Mathematical Sciences, LPMC and AAIS, Nankai University, Tianjin, P.R.China (xiahan@nankai.edu.cn)

†Department of Statistics and Actuarial Science, University of Waterloo, Waterloo ON N2L 3G1, Canada (bin.li@uwaterloo.ca)

‡Department of Economics, University of Toronto, Toronto ON M5S 3G7, Canada (yao.luo@utoronto.ca)

medical histories from small clinics or foreign providers are often inaccessible, while lifestyle-related factors (e.g., smoking or physical inactivity) are frequently underreported. Furthermore, some risk-relevant variables such as genetic information or socio-economic status may raise ethical or fairness concerns, and their use is often subject to regulatory restrictions. These types of private information, while essential for accurate risk pricing, frequently remain partially or entirely unobservable to insurers. On top of this, customers' risk attitudes play a fundamental role in insurance contract design but constitute even more private information (e.g., Finkelstein and McGarry (2006) and Cohen and Einav (2007)), posing another significant challenge for insurers to ascertain. As such, information asymmetry in insurance can be categorized into two main categories: *risk-type uncertainty* and *risk-attitude uncertainty*. The literature primarily focuses on risk-type uncertainty, often referred to as the adverse selection problem, as insurers are concerned that customers who hold private information may actually be high-risk individuals. They might purchase excessive insurance coverage, leading to significant losses for the insurer if the contracts are not priced carefully.

Based on market models, the literature on optimal insurance with information asymmetry can be categorized into *competitive markets* and *monopoly markets*, which originate from the seminal works of Rothschild and Stiglitz (1976) and Stiglitz (1977), respectively. Rothschild and Stiglitz (1976) explores a competitive market model where insurers have imperfect information about the probability of a deterministic loss occurring. It assumes that insurers share a common perspective on the binary distribution of high-risk and low-risk customers. The most notable implication is that a competitive equilibrium of the Nash-Cournot type may not exist under quite plausible conditions. On the other hand, Stiglitz (1977) examines a profit-maximizing monopoly insurer operating under similar conditions of imperfect information. An optimal menu of insurance contracts is derived, taking into account a customer's participation constraint and truth-telling constraint, the latter serving to elicit customers' true information.

Both competitive and monopoly models offer valuable insights, however, the two models do not completely align with real markets. First, insurance markets lack full competitiveness, attributed not only to the presence of explicit entry costs but also to the limited hedging power of new entrants. Established insurers have large pools of policies with similar characteristics, and they benefit from risk diversification ensured by the law of large numbers, allowing them to offer low-priced contracts. Hence, as also emphasized in Rothschild and Stiglitz (1976) and Stiglitz (1977), it is widely acknowledged that real insurance markets should "*lie in between perfect competition and monopoly*". Second, the predictions of both models are extreme, with the high-risk customer purchasing full-coverage insurance, whereas real-world contracts usually include deductibles or coinsurance. Moreover, in absence of information asymmetry, the competitive model implies zero expected profit for insurers, while the monopoly model suggests no welfare gains for customers in the insurance purchase decision. Third, customer's demand for optimal insurance coverage, a crucial aspect in real contract design and theoretical study, is not taken into consideration. This creates a significant gap with the extensive body of literature on optimal insurance coverage that originated from the seminal work of Arrow (1963).

One remedy is to model a market structure that lies between monopoly and perfect competition using a *Stackelberg game* framework. In continuous time, principal–agent problems with such hierarchical decision-making structures were first studied in the seminal work of Holmstrom and Milgrom (1987). The solution process typically involves two steps: (1) solving the agent's optimal response given a proposed contract, and (2) determining the principal's optimal contract based on that response. This framework reflects the leader–follower relationship in economic decision making, illustrating how intermediate market structures influence contract design and risk-sharing.

The application of the Stackelberg game framework to insurance is natural. The insurer acts as the principal who designs insurance contracts, while the customer assumes the role of an agent who

chooses how much coverage to purchase. A natural trade-off arises in this setting: if the insurer sets the premium too high, the agent will demand less coverage, thereby reducing the insurer's expected profit. Moreover, the customer's optimal insurance demand problem, extensively studied following Arrow (1963), corresponds precisely to the first step of the Stackelberg game. Since Chen and Shen (2018, 2019), models based on this hierarchical framework have attracted significant attention and have been widely studied in recent years. Nevertheless, a key limitation of these models is that the principal is assumed to have perfect information about the agent's risk type and preferences.

To overcome the limitation of assuming perfect information, we adopt a *Stackelberg game framework with hidden information*, which integrates hierarchical decision making with information asymmetry. In this framework, agents possess *hidden information* (e.g., skill levels or risk preferences) that the principal cannot directly observe, while the principal anticipates and optimally responds to the agents' decisions in a Stackelberg manner. Within this framework, the principal's objective is twofold: (i) to design a menu of contracts that induces each agent to self-select the contract intended for them, thereby revealing their private information, and (ii) to maximize expected profit (or utility) subject to this truth-telling constraint. Problems involving hidden types are commonly referred to as *adverse selection*, as the principal faces the risk that an agent may select a contract not intended for them, potentially resulting in losses. To the best of our knowledge, Sung (2005) is the first to study such problems in a general continuous-time principal–agent framework. These problems are widely regarded as among the most technically challenging topics in contract theory. Subsequent literature has primarily focused on firm management applications, where the principal faces uncertainty regarding the agent's productivity or skill. Nevertheless, it is surprising that the insurance literature has largely overlooked this framework, despite the fact that moral hazard and adverse selection remain two of the most fundamental challenges in the insurance industry.

1.2 Contributions

This paper proposes a new framework for studying optimal insurance design under information asymmetry within the Stackelberg game setting. Specifically, we consider a monopolistic insurer facing uncertainty about a customer's loss distribution (i.e., risk-type uncertainty) or risk preferences (i.e., risk-attitude uncertainty). The insurer acts as the principal who sets the risk loading (the premium minus expected loss), while the customer, acting as the agent, chooses the level of coverage in response. The solution procedure unfolds in four steps: (1) Customer's Problem: The customer selects an optimal coverage level based on their preferences and the offered risk loading, including the possibility of choosing contracts not tailored to their true type. (2) Truth-Telling Constraint: To elicit the customer's hidden information, the contract must satisfy a truth-telling constraint, namely, that the customer maximizes their value function by revealing their true type. This yields a first-order condition for the insurer's pricing strategy. (3) Insurer's Problem: We characterize the insurer's optimal menu of contracts, pairs of premiums and coverage levels, that maximize expected profit while satisfying the first-order condition implied by the truth-telling constraint. (4) Verification: We verify that the resulting contract menu satisfies the truth-telling constraint. This step is essential because the first-order condition is necessary but not sufficient for incentive compatibility.

Our new framework lies between classic monopoly and competitive models. With or without information asymmetry, both the insurer and the customer are better off by participating in our solved optimal menu. In contrast, under the competitive model, the insurer earns zero expected profit, while under the monopoly model, the customer derives no welfare gain. Our result more closely reflects the real market, where insurers need to earn positive profits to cover non-negligible costs on top of insurance loss, such as commissions to agents, and where buyers should benefit from

the existence of insurance markets. Additionally, our approach establishes a connection with the extensive literature on optimal insurance coverage. Notably, we address the customer's problem in the first step, allowing them to choose coverage that maximizes their expected utility, in contrast to the classic monopoly model where the buyer's demand is not considered. This shift acknowledges the significance of customers' demand for optimal coverage and establishes a connection with thousands of existing works on buyer's optimal coverage demand, stemming from the seminal work of Arrow (1963).

The new findings and implications of the paper are summarized as follows. First of all, the most intriguing result of the paper is that, in the optimal menu, we find a *nonlinear pricing* structure for risk-type uncertainty and a *linear pricing* structure for risk-attitude uncertainty in terms of *risk loading*, the premium minus the expected value of loss. Specifically, for high-risk customers, the risk loading is set lower to encourage selection of higher coverage contracts, thereby revealing their true risk type. However, in cases where the insurer is only uncertain about the customer's risk attitude, there is no necessity to provide such a *risk-loading discount* to elicit the customer's true risk attitude. Intuitively, this finding can be attributed to the difference between these two types of missing information. A customer's risk type directly affects both their claim experience and the insurer's profit, making it *more valuable information* than the customer's risk preference, which does not have an immediate impact on either claim experience or insurer profit. Therefore, to obtain the more valuable missing information—the customer's risk type—insurers need to provide discounts in pricing menus.

Second, our findings of a risk-loading discount in response to risk-type uncertainty provide new insights to the existing literature, which primarily rely on examining the price-quantity (P-Q) curve to test for quantity discounts (concave P-Q curve) or quantity premia (convex P-Q curve). In a competitive insurance market, Rothschild and Stiglitz (1976) show that insurers offer global quantity premia due to the convexity of the P-Q curve. In contrast, in a monopoly insurance market Chade and Schlee (2012) show that under mild conditions, the optimal P-Q curve takes a backward S-shape: concave at first, then convex. This departs from general monopoly models (e.g., Maskin and Riley (1984)), which typically predict concave P-Q curve. Empirical findings based on P-Q curves are also mixed, with evidence supporting the existence of both convex and concave P-Q curves in practice, see, for instance Puelz and Snow (1994) and Cawley and Philipson (1999), respectively.

Our results are consistent with these empirical observations, showing that the P-Q curve is not necessarily convex or concave (e.g., the P-Q curve in Figure 1 is convex, while the one in Figure 2 is S-shaped). This implies that testing for quantity discounts or premia solely based on the shape of the P-Q curve may be insufficient. Our theoretical analysis reveals that pricing discounts do exist, but in the sense that high-risk customers are charged a lower risk loading. Therefore, we suggest also examining the monotonicity of the loading-type curve (see the left panels of Figures 1 and 2) as a complementary approach to test for the presence of quantity discounts or premia. For clarity, the aforementioned commonalities and differences between our paper and the existing literature on whether insurers offer quality discounts/premia are summarized in Table 1.

Third, this paper also contributes to the limited theoretical literature on risk-attitude uncertainty. Most existing theoretical studies in this line focus on risk-type uncertainty. A notable exception is Landsberger and Meilijson (1994), which examines a monopoly market with a risk-neutral insurer and two types of expected-utility consumers who differ in their degree of risk aversion. To the best of our knowledge, our paper is the first to consider a continuous spectrum of risk aversion levels. This broader framework enables us to derive more analytical insights into the curvature relationships among risk loading, premium, and loss coverage, as discussed in Section 3.5. More specifically, Figures 3 and 4 show the premium-coverage curve is globally convex, an insight that

Table 1: Price-Quantity (P-Q) and Loading-Type (L-T) Curves under Risk-Type Uncertainty

Model	**P-Q Curve**	**L-T Curve**	**Implication**	**Reference**
Competitive insurance	Convex	—	Global quantity premia	Rothschild and Stiglitz (1976)
Monopoly Insurance	Backward S-shaped	—	No global premia/discount	Chade and Schlee (2012)
General Monopoly	Concave	—	Global quantity discount	Maskin and Riley (1984)
Empirical Studies	Uncertain	—	Quantity premia or discount observed	Puelz and Snow (1994), Cawley and Philipson (1999)
This Paper	Uncertain	Decreasing	1. P-Q test alone is insufficient 2. Global discount in risk loading 3. Suggest to add L-T curve test	—

cannot be captured when only two risk-attitude types are considered.

Lastly, as for comparison, we also consider an alternative approach for addressing information asymmetry by formulating it as a decision-making problem under the *worst-case scenario* through the maxmin expected utility (e.g., Gilboa and Schmeidler (1989)). As expected, the worst-case contract is a single contract, and thus, a customer's private information cannot be sorted and elicited. Moreover, compared to the worst-case contract, our optimal menu has the advantages of higher expected profit for the insurer and a greater likelihood of market existence. In contrast, if the insurer can only offer the worst-case contract, there is a non-negligible likelihood of no trade, where the insurer rejects all customers.

Our model includes several specifications. We specify that the customer's aggregate loss follows a compound Poisson process, which is arguably the most fundamental loss model for multiple periods, and the customer's risk preference is assumed to follow the dynamic mean-variance (MV) criterion. The dynamic MV criterion offers great mathematical tractability, as evidenced by the pioneering contributions of Basak and Chabakauri (2010), Björk and Murgoci (2010), and numerous subsequent studies. This results in the optimal menu being not only in *closed form* but also *independent* of the customer's wealth. While the wealth effect may play an important role in areas such as firm or fund management, discrimination based on customer characteristics such as wealth, however, is prohibited to ensure equitable treatment in insurance regulation.

Admittedly, restricting the buyer's reward function to the MV criterion is a limitation. We believe that the main implication—more loading discounts in the presence of risk-type uncertainty compared to risk-attitude uncertainty—is likely to hold in a more general setting, as information about the customer's risk type is indeed more valuable. We leave it for future work to examine this conjecture within a more general expected utility framework.

1.3 Relationship to the Literature

Optimal insurance with information asymmetry is arguably one of the most important and challenging topics in insurance. It has been significantly developed since the seminal works by Rothschild and Stiglitz (1976) and Stiglitz (1977). Later literature can be categorized into studies on competitive markets (e.g., Wambach (2000), Garcia and Tsur (2021), and Farinha Luz et al. (2023)) and

monopoly markets (e.g., Chade and Schlee (2012), Hendren (2013), and Gershkov et al. (2023)). This paper is closer to the monopoly setting but with major differences in both modeling and implications. We adopt a hierarchical principal–agent approach based on the Stackelberg game framework, separating the customer's problem and the insurer's problem into two sequential optimization problems. By considering the customer's demand for optimal coverage, our approach establishes a connection between the topic of optimal insurance with information asymmetry and thousands of existing works on optimal insurance coverage that have evolved from the seminal work by Arrow (1963). In contrast, the classic monopoly setting is a single-step insurer's optimization problem with customer's participation and truth-telling constraints, but without considering customer's demand on optimal coverage.

Our new findings also make valuable contributions to the literature. For an optimal menu solved from either competitive or monopoly model, a key feature to examine is whether the menu exhibits quantity discounts or quantity premia. Existing works analyze this through the *price-quantity curve*, where a concave (resp. convex) curve indicates quantity discounts (resp. premia). What is most surprising about the monopoly model is that, under the log-concave distribution assumption on the customer's risk type, Chade and Schlee (2012) shows that the price-quantity curve is backward S-shaped. This sharply contrasts with the expectations for a "general" monopolist, who would typically suggest quantity discounts (Maskin and Riley (1984)). In our paper, we propose a novel approach to determining whether a menu offers discounts by analyzing *risk loading (premium minus expected loss)* rather than the premium itself. Our findings are consistent across different loss distributions and customer risk type or attitude distributions: the optimal menu indeed provides risk loading discounts to high-risk individuals but not much to high-risk-averse individuals. We attribute this to the fact that a customer's risk type is more valuable information than their risk attitude, the latter only reflects their willingness to buy insurance and has no direct impact on insurance losses or the insurer's profit.

This paper is also related to the literature on continuous-time principal–agent problems with hidden information, which are among the most technically challenging topics in contract theory. Existing studies primarily focus on firm management problems, where the principal faces uncertainty about the agent's productivity or skill; see, e.g., Sung (2005), Cvitanić and Zhang (2007), Sannikov (2008), Cvitanić et al. (2013, 2017). These analyses often employ advanced backward stochastic differential equations (BSDEs), and the solutions are generally not fully explicit. To the best of our knowledge, this paper is the first to explore optimal insurance in continuous time within a framework that combines hidden information with a Stackelberg differential game. Remarkably, our optimal menu is fully explicit, and the derivations rely only on ordinary differential equations. This simplification arises naturally from the structural restrictions in insurance contracts, where premiums consist of expected loss plus a risk loading, whereas admissible contracts in firm management problems typically have no such constraints. Another key distinction is our novel finding regarding the nonlinear versus linear risk loading structure. Given the central role of information asymmetry (or adverse selection) in insurance, we believe that the explicit solutions and analysis presented here can serve as a foundation for future extensions, including cases with multi-dimensional private information (e.g., Wambach (2000), Finkelstein and McGarry (2006), Crocker and Snow (2011), and Aryal et al. (2019)), which to date have been limited to a finite number of provided contracts.

An alternative approach to addressing information asymmetry is the ambiguity framework, where a decision maker considers a collection of plausible perspectives (referred to as prior probability measures) and makes decisions based on various criteria, such as the worst-case scenario (maxmin expected utility), a weighted sum of the worst and best cases (α-maxmin expected utility), or a smooth aggregation of all perspectives (smooth ambiguity). Existing works on optimal

insurance with ambiguity include Gollier (2014), Li et al. (2016), Guan and Li (2022), among others. However, the contract obtained from an ambiguity model is a single contract, rather than a menu of contracts, and thus cannot elicit the missing information from customers. Moreover, our comparison with the single contract obtained from maxmin expected utility reveals that our optimal menu offers additional desirable features, including a higher expected profit for the insurer and a greater likelihood of market existence.

1.4 Organization of the Paper

The structure of the paper is as follows. Section 2 studies the case of risk-type uncertainty. Specifically, Section 2.1 introduces the models for aggregate loss, customer's loss coverage function, and insurer's premium structure. Section 2.2 formulates the customer's problem and the insurer's problem. Section 2.3 presents the optimal menu of contracts in explicit form. Section 2.4 examines major properties of the optimal menu. Section 2.5 considers the alternative worst-case contract derived based on the maxmin expected utility and compares it with our optimal menu. Section 2.6 investigates the risk-loading curve with respect to risk parameters and uses an example to demonstrate that the classic approach of examining discounts based on premium-quantity curve can be misleading. Section 3 provides a parallel analysis to solve and analyze the optimal menu when the insurer is uncertain about the customer's degree of risk aversion. Section 4 concludes. All proofs are relegated to the Appendix.

2 Optimal Menu with Risk-Type Uncertainty

2.1 Modelling

In the paper, the insurer (referred to as "she") takes on the role of the principal and designs an insurance contract, while the customer (referred to as "he") acts as the agent in the principal-agent problem. Without purchasing any insurance, the aggregate loss of a customer is modelled by a compound Poisson model, arguably the most fundamental loss model in continuous time, i.e.,

$$U^\theta(t) = \sum_{i=1}^{N(t)} Z_i^\theta. \tag{2.1}$$

Here, $\{N(t)\}_{t\geq 0}$ is a Poisson process with intensity $\lambda > 0$, and $\{Z_i^\theta\}_{i\in\mathbb{N}}$ is a sequence of independent and identically distributed positive random variables, independent of $\{N(t)\}_{t\geq 0}$, with common probability density function $f(y;\theta)$ for $y > 0$. The parameter θ characterizes customer *risk type*, which is his private information. For example, $f(y;\theta) = \frac{1}{\theta}$ for $y \in (0,\theta)$ if Z^θ follows a uniform distribution in $(0,\theta)$, or $f(y;\theta) = \frac{1}{\theta}e^{-y/\theta}$ for $y > 0$ if Z^θ follows an exponential distribution with parameter θ. We assume that $f(\cdot;\theta)$ is differentiable in θ. Alternatively, one can rewrite (2.1) as

$$U^\theta(t) = \int_0^t \int_0^\infty y N^\theta(\mathrm{d}s, \mathrm{d}y), \tag{2.2}$$

where $N^\theta(\mathrm{d}s, \mathrm{d}y)$ is a Poisson random measure with intensity measure $\lambda f(y;\theta)\mathrm{d}y$, denoting the random number of claims of size $[y, y+\mathrm{d}y]$ occurred within the time period $[s, s+\mathrm{d}s]$. In this section, we assume the insurer is uncertain about the customer's true risk type θ, which is modelled by a random variable with cumulative distribution function $G(\theta)$ with compact support $[\theta_L, \theta_H] \subset \mathbb{R}$.

An insurance strategy adopted by customers is a *loss coverage function* of the form $l(t,y) : [0,T] \times \mathbb{R}_+ \to \mathbb{R}$, where $y = U^\theta(t) - U^\theta(t_-)$ is the size of an instantaneous loss occurring at

time t, and $l(t,y)$ is the portion reimbursed by the insurer. For clarity and conciseness, we shall omit time variable t throughout the remainder of this paper, as the optimal loss functions are actually independent of t. This aligns with the practice in which insurance coverage remains fixed throughout the duration of a contract. Common insurance contracts include proportional insurance with $l(y) = qy$, where $q \in [0,1]$, and stop-loss insurance with $l(y) = \min\{y, M\}$, where $M > 0$ is the maximum loss amount covered. The set of all admissible insurance strategies is denoted by $\mathbb{A}_l$. In particular, $l(y) \equiv 0$ (resp. $l(y) \equiv y$) means that the customer purchases *no insurance* (resp. *full insurance*). At this point, we do not impose the restriction that $l(y) \in [0,y]$ to exclude shorting or over insurance. But it will be shown that the optimal menu of contracts satisfies this restriction.

The insurer charges an insurance premium to cover the part of the loss specified by an insurance contract. With asymmetric information, a customer may not reveal the true information of his risk type to the insurer. For risk type $\tilde{\theta}$, the expected rate of loss corresponding to an insurance strategy $l \in \mathbb{A}_l$ per unit of time is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[\int_0^t \int_0^\infty l(y) N^{\tilde{\theta}}(\mathrm{d}s, \mathrm{d}y)\right] = \lambda \int_0^\infty l(y) f(y;\tilde{\theta}) \mathrm{d}y.$$

In practice, the insurer charges a *risk loading* on top of expected loss to cover expenses and generate profit. We assume the insurer charges an insurance premium according to the variance premium principle. More specifically, if the customer adopts an insurance strategy $l \in \mathbb{A}_l$ and he chooses the type-$\tilde{\theta}$ contract, the continuously-charged premium rate paid by the customer is given by

$$p(l;\tilde{\theta}) = \lambda \int_0^\infty l(y) f(y;\tilde{\theta}) \mathrm{d}y + \frac{\eta(\tilde{\theta})}{2} \lambda \int_0^\infty l(y)^2 f(y;\tilde{\theta}) \mathrm{d}y, \tag{2.3}$$

where $\eta(\tilde{\theta}) \geq 0$ is the *risk loading factor* for type $\tilde{\theta}$. We call $p(l;\tilde{\theta})$ the *type-$\tilde{\theta}$ premium rate*, as both the risk loading factor $\eta(\tilde{\theta})$ and the claim distribution $f(\cdot;\tilde{\theta})$ are determined by $\tilde{\theta}$. The set of all admissible loading factors is denoted by

$$\mathbb{A}_\eta = \left\{\eta(\cdot) : \eta(\tilde{\theta}) \geq 0 \text{ for all } \tilde{\theta}\right\}.$$

The first term in (2.3) is the base premium, which corresponds to the expected loss undertaken by the insurer. The second term denotes the risk loading, which is determined based on the second moment of the customer's risk. As such, (2.3) is referred to as the variance premium principle. There are several reasons for adopting the variance premium principle: First, it is widely used in actuarial practice as one of the most common premium principles. Second, it aligns with the customer's mean-variance (MV) risk preference, see later (2.6). Third, it offers analytical tractability, as the customer's optimal insurance strategy under the variance premium principle is to purchase a simple proportional insurance.

Given that the customer's true risk type is θ, and he chooses the type-$\tilde{\theta}$ contract and coverage function l, the *customer's surplus* is given by

$$\begin{aligned} X_C^\theta(t,l;\tilde{\theta}) &= x_0 - p(l;\tilde{\theta})t - U^\theta(t) + \int_0^t \int_0^\infty l(y) N^\theta(\mathrm{d}s, \mathrm{d}y) \\ &= x_0 - p(l;\tilde{\theta})t - \int_0^t \int_0^\infty (y - l(y))\, N^\theta(\mathrm{d}s, \mathrm{d}y), \end{aligned} \tag{2.4}$$

where x_0 is his initial wealth level. Correspondingly, the *insurer's profit* evolves according to

$$X_I^\theta(t,l;\tilde{\theta}) = p(l;\tilde{\theta})t - \int_0^t \int_0^\infty l(y) N^\theta(\mathrm{d}s, \mathrm{d}y). \tag{2.5}$$

In X_C^θ and X_I^θ, subscript "C" stands for customer and I stands for insurer; superscript "θ" represents the customer's true risk type, while the "$\tilde{\theta}$" after the semicolon represents his chosen contract type. It is very important to note that the selected contract type $\tilde{\theta}$ only affects the premium rate and not the real claim experience, while the true risk type θ only affects the claim experience and not the premium.

Due to the insurer's superior hedging ability in managing insurance risks through risk pooling, there is a natural difference in risk appetite between insurers and customers. Customers, recognizing their limited ability to hedge risks independently, tend to be more risk-averse compared to insurers. This disparity arises from the understanding that insurers have better resources and mechanisms to effectively manage and mitigate risks through pooling and diversification strategies. As a result, customers are inclined to seek insurance coverage to protect themselves against potential losses, while insurers are better equipped to handle and absorb those risks. In this paper, we assume the customer uses the MV criterion, namely,

$$\mathbb{E}\left[X_C^\theta(T,l;\tilde{\theta})\right] - \frac{\gamma}{2}\mathrm{Var}\left[X_C^\theta(T,l;\tilde{\theta})\right], \tag{2.6}$$

where $\gamma > 0$ is the risk aversion parameter, and the insurer is risk neutral (e.g., Arrow (1963) and Rothschild and Stiglitz (1976)).

2.2 Problem Formulation

This section formulates the customer's and insurer's problems. For a customer with true risk type θ selecting a contract for risk type $\tilde{\theta}$, his objective is to maximize the MV of the surplus by selecting an optimal loss coverage function l.

Problem 2.1 (Customer's problem) *We define the value function for a type-θ customer choosing a contract for risk type $\tilde{\theta}$ as follows:*

$$V_C^\theta(\tilde{\theta}) := \sup_{l\in\mathbb{A}_l}\left\{\mathbb{E}\left[X_C^\theta(T,l;\tilde{\theta})\right] - \frac{\gamma}{2}\mathrm{Var}\left[X_C^\theta(T,l;\tilde{\theta})\right]\right\}. \tag{2.7}$$

Correspondingly, the customer's optimal loss coverage to (2.7) is denoted by $\hat{l}^\theta(\tilde{\theta}) := \{\hat{l}^\theta(y;\tilde{\theta})\}_{y>0}$.

Problem 2.1 presents a dynamic MV problem, which has the well-known issue of time inconsistency. This paper employs the mainstream equilibrium approach by considering a customer's decision-making process as a non-cooperative game involving their future selves; see a series of seminal papers by e.g., Strotz (1973), Basak and Chabakauri (2010), Björk et al. (2014), Björk et al. (2017), and Dai et al. (2021). The solution derived from this approach is formally referred to as an intra-personal equilibrium strategy, as defined in Definition A.1 in the Appendix. For simplicity, we still refer to it as an *optimal* strategy, while acknowledging it is in the equilibrium sense.

After solving the customer's problem, we turn to the insurer's problem, which encompasses two objectives. First, the menu of contracts needs to be designed in a manner that makes it optimal for the customer to select the contract aligned with their true risk type. This objective is encapsulated by the *truth-telling constraint* stated in (2.8). By achieving this, the true risk type of the customer can be effectively elicited. Second, while adhering to the truth-telling constraint, the insurer aims to maximize her *weighted* expected profit, considering all possible prior values of the true risk type θ. This objective entails optimizing the profitability of the insurance contracts while accounting for the different risk profiles of the customers. By addressing these two objectives, the insurer can effectively design a menu of contracts that not only maximizes the insurer's weighted expected profit but also elicits the customer's true risk type.

Problem 2.2 (Insurer's problem) *The insurer's weighted expected profit is given by*

$$\sup_{\eta(\cdot)\in\mathbb{A}_\eta}\int_{\theta_L}^{\theta_H}\mathbb{E}\left[X_I^\theta(T,\hat{l}^\theta(\theta);\theta)\right]\mathrm{d}G(\theta),$$

subject to the customer's truth-telling constraint

$$\theta\in\mathrm{argmax}_{\tilde{\theta}\in[\theta_L,\theta_H]}V_C^\theta(\tilde{\theta}). \tag{2.8}$$

It is also worth noting that the customer's problem 2.1 implies the automatic satisfaction of the customer's *participation constraint*. This means that the customer's optimal value function should always be higher than the value they would have if they purchase no insurance. This holds true because adopting the strategy where the loss coverage function $l\equiv 0$ is an admissible strategy for the customer.

2.3 Solution of the Optimal Menu

This subsection focuses on presenting our solution to the optimal menu of contracts. The following theorem presents the solution to the customer's problem.

Theorem 2.1 *An optimal loss coverage function to the customer's problem (2.7) is given by*

$$\hat{l}^\theta(y;\tilde{\theta})=\frac{(\gamma y+1)f(y;\theta)-f(y;\tilde{\theta})}{\eta(\tilde{\theta})f(y;\tilde{\theta})+\gamma f(y;\theta)}, \tag{2.9}$$

and the value function is given by

$$\begin{aligned}V_C^\theta(\tilde{\theta})=x_0&-\lambda T\int_0^\infty\left[\hat{l}^\theta(y;\tilde{\theta})+\frac{\eta(\tilde{\theta})}{2}\hat{l}^\theta(y;\tilde{\theta})^2\right]f(y;\tilde{\theta})\mathrm{d}y\\&-\lambda T\int_0^\infty\left[\left(y-\hat{l}^\theta(y;\tilde{\theta})\right)+\frac{\gamma}{2}\left(y-\hat{l}^\theta(y;\tilde{\theta})\right)^2\right]f(y;\theta)\mathrm{d}y.\end{aligned}$$

In particular, if $\tilde{\theta}=\theta$, we have

$$\hat{l}^\theta(y;\theta)=\frac{\gamma}{\eta(\theta)+\gamma}y, \tag{2.10}$$

which is a proportional insurance strategy, and

$$V_C^\theta(\theta)=x_0-\lambda T\int_0^\infty\left(y+\frac{\eta(\theta)\gamma}{2(\eta(\theta)+\gamma)}y^2\right)f(y;\theta)\mathrm{d}y. \tag{2.11}$$

To solve the insurer's problem, we break it down into two steps. First, Lemma 2.1 derives the form of the loading factors that satisfy the first-order condition, which is a *necessary* condition for the truth-telling constraint (2.8). Second, in Theorem 2.2, we confirm the *sufficiency* by proving that the derived menu of contracts indeed satisfies the truth-telling constraint. It is worth noting that the approach of breaking down necessity and sufficiency is commonly adopted in the literature on optimal risk-sharing with hidden information (e.g., Sung (2005) and Cvitanić and Zhang (2007)) due to the challenge of characterizing the if and only if condition for the truth-telling constraint.

For ease of notation, we denote the first two moments of Z^θ by

$$\mu(\theta):=\mathbb{E}[Z^\theta]\text{ and }\sigma(\theta):=\mathbb{E}[(Z^\theta)^2].$$

Lemma 2.1 *A necessary condition for the loading factors to satisfy the truth-telling constraint (2.8) is of the form*

$$\eta(\theta) = -\int_{\theta_L}^{\theta} e^{-\int_x^\theta a(z)\mathrm{d}z} r(x)\mathrm{d}x + Ce^{-\int_{\theta_L}^\theta a(z)\mathrm{d}z}, \tag{2.12}$$

where $a(\theta) := \frac{2\mu'(\theta)+\gamma\sigma'(\theta)}{\gamma\sigma(\theta)}$, $r(\theta) := \frac{2\mu'(\theta)}{\sigma(\theta)}$, *and* C *is a constant such that*

$$C \geq \int_{\theta_L}^{\theta_H} e^{\int_{\theta_L}^x a(z)\mathrm{d}z} r(x)\mathrm{d}x := \bar{r},$$

where the last inequality is equivalent to $\eta(\theta) \geq 0$ *for all* $\theta \in [\theta_L, \theta_H]$.

The following assumption will be invoked in Theorem 2.2. It is a natural assumption to imply that a greater value of θ represents a riskier customer type. Moreover, the moments of most parametric distributions are indeed monotone with respect to the parameters, such as exponential distributions with parameter θ and uniform distributions on $(0, \theta)$.

Assumption 2.1 *We assume* $\mu(\theta)$ *and* $\sigma(\theta)$ *are strictly increasing in* θ.

The following theorem summarizes the optimal loading factor, optimal menu of contracts with risk-type uncertainty, the insurer's expected profit from offering the optimal menu, and verifies the truth-telling constraint.

Theorem 2.2

(1) An optimal loading factor to the relaxed insurer's problem, i.e.,

$$\sup_{\eta\in\mathbb{A}_\eta} \int_{\theta_L}^{\theta_H} \mathbb{E}\left[X_I^\theta(T, \hat{l}^\theta(\theta); \theta)\right] \mathrm{d}G(\theta),$$

subject to the first-order condition (2.12) is given by

$$\hat{\eta}(\theta) = -\int_{\theta_L}^{\theta} e^{-\int_x^\theta a(z)\mathrm{d}z} r(x)\mathrm{d}x + C^* e^{-\int_{\theta_L}^\theta a(z)\mathrm{d}z}, \tag{2.13}$$

where $C^* := \mathrm{argmin}_{C\geq\bar{r}} H(C)$, *and*

$$H(C) = \int_{\theta_L}^{\theta_H} \left(\left(-\int_{\theta_L}^{\theta} e^{-\int_x^\theta a(z)\mathrm{d}z} r(x)\mathrm{d}x + Ce^{-\int_{\theta_L}^\theta a(z)\mathrm{d}z} + \gamma \right)^{-1} - \frac{1}{2\gamma} \right)^2 \sigma(\theta)\mathrm{d}G(\theta).$$

Moreover, we have

$$\gamma \vee \bar{r} \leq C^* \leq \bar{r} + \gamma e^{\int_{\theta_L}^{\theta_H} a(z)\mathrm{d}z}. \tag{2.14}$$

(2) An optimal menu of contracts with loss coverage $\hat{l}(\theta) := \hat{l}(y;\theta)$ *and premium rate* $\hat{p}(\theta) := \hat{p}(\hat{l}(\theta);\theta)$ *for* $\theta \in [\theta_L, \theta_H]$ *is given by*

$$\begin{cases} \hat{l}(\theta) = \frac{\gamma}{\hat{\eta}(\theta)+\gamma} y, \\ \hat{p}(\theta) = \frac{\lambda\gamma}{\hat{\eta}(\theta)+\gamma}\mu(\theta) + \frac{\lambda\gamma^2\hat{\eta}(\theta)}{2(\hat{\eta}(\theta)+\gamma)^2}\sigma(\theta), \end{cases} \tag{2.15}$$

where $\hat{\eta}(\theta)$ *is given in (2.13).*

(3) Correspondingly, the insurer's optimal expected profit is given by

$$\int_{\theta_L}^{\theta_H} \mathbb{E}\left[X_I^{\theta}(T,\hat{l}(\theta);\theta)\right] \mathrm{d}G(\theta) = \frac{\lambda\gamma T}{8}\int_{\theta_L}^{\theta_H} \sigma(\theta)\mathrm{d}G(\theta) - \frac{\lambda\gamma^3 T}{2}H(C^*). \tag{2.16}$$

(4) Under Assumption 2.1, the optimal menu of contracts $(\hat{l}(\theta),\hat{p}(\theta))_{\theta\in[\theta_L,\theta_H]}$ satisfies the truth-telling constraint in the sense that

$$\theta = \mathrm{argmax}_{\tilde{\theta}\in[\theta_L,\theta_H]}\left\{\mathbb{E}\left[X_C^{\theta}(T,\hat{l}(\tilde{\theta});\tilde{\theta})\right] - \frac{\gamma}{2}\mathrm{Var}\left[X_C^{\theta}(T,\hat{l}(\tilde{\theta});\tilde{\theta})\right]\right\},$$

where $\{X_C^{\theta}(t,\hat{l}(\tilde{\theta});\tilde{\theta})\}_{t\in[0,T]}$ is the surplus process of a type-θ customer who chooses the type-$\tilde{\theta}$ contract $(\hat{l}(\tilde{\theta}),\hat{p}(\tilde{\theta}))$.

In particular, if the insurer has no uncertainty on the customer's risk type θ, it is not difficult to see from Theorem 2.2(1) that function H reduces to

$$H(C) = \left(\frac{1}{C+\gamma} - \frac{1}{2\gamma}\right)^2 \sigma(\theta).$$

This implies that the optimal loading factor becomes

$$\hat{\eta}(\theta) = C^* = \mathrm{argmin}_{C\geq 0}\, H(C) = \gamma. \tag{2.17}$$

Then, by (2.15), we have the following corollary.

Corollary 2.1 *Suppose the insurer has no uncertainty about the customer's risk type θ. Then, the optimal menu contains a single contract given by*

$$\begin{cases} \hat{l} = \frac{1}{2}y, \\ \hat{p} = \frac{\lambda}{2}\mu(\theta) + \frac{\lambda\gamma}{8}\sigma(\theta). \end{cases}$$

Remark 2.1 *Theorem 2.2 characterizes the Stackelberg equilibrium menu under hidden information; that is, the set of contract pairs consisting of a loss coverage function and a premium rate $(\hat{l}(\theta),\hat{p}(\theta))_{\theta\in[\theta_L,\theta_H]}$ as given in* (2.15)*. Moreover, Part (4) of Theorem 2.2 confirms that the optimal menu $(\hat{l}(\theta),\hat{p}(\theta))_{\theta\in[\theta_L,\theta_H]}$ indeed satisfies the truth-telling constraint, meaning that it is suboptimal for a customer of true type θ to choose any contract $(\hat{l}(\tilde{\theta}),\hat{p}(\tilde{\theta}))$ with $\tilde{\theta}\neq\theta$. Consequently, customer's hidden private information regarding their risk types could be revealed through the self-selection process.*

In contrast, when there is no hidden information, the contract coincides with the Stackelberg equilibrium under perfect information, reducing the menu of contracts to a single proportional insurance contract, with both parties covering 50% *of the losses. Note that Corollary 2.1 illustrates this case, and it coincides with the main result of Chen and Shen (2019).*

2.4 Properties of the Optimal Menu

This section examines major properties of the optimal menu solved in Theorem 2.2. More specifically, Proposition 2.1 investigates the monotonicity of the characteristics of the optimal menu with respect to risk type. Proposition 2.2 assesses the welfare improvement for both parties. Proposition 2.3 demonstrates that the optimal menu effectively avoids moral hazard concerns, ensuring that the customer has no motivation to behave recklessly after choosing his contract.

Proposition 2.1 *For the optimal menu of contracts solved in Theorem 2.2, we have*
(1) $\hat{\eta}(\theta)$ is decreasing in θ;
(2) $\hat{p}(\theta)$ is increasing in θ;
(3) $\hat{l}(\theta)$ is increasing in θ.

Proposition 2.1(1) demonstrates that the optimal menu involves assigning a *lower* risk loading to *high-risk* customers. This approach aims to reveal the customer's true risk type, making it advantageous for high-risk (resp. low-risk) customers to choose high-coverage (resp. low-coverage) contracts. Notably, this menu of contracts implies a *nonlinear pricing* strategy—namely, the insurer needs to offer *discounts* to high-risk customers—by reducing the risk loading—in order to obtain the missing information about their true risk type. Empirical support for this screening mechanism can be found in Cawley and Philipson (1999), who provide evidence of such nonlinear pricing patterns in life insurance market.

This is the most intriguing finding of the paper, and it is worth noting that it holds for any customer's loss distribution Z^θ and the insurer's perspective on the distribution of risk types G. It also differs from the classic analysis which uses the premium-coverage curve to examine discounts. As we will see in Figure 1, the premium coverage curve could be misleading by suggesting a quantity premia, whereas discounts are actually provided in the risk loading!

Parts (2) and (3) show that the premium rate $\hat{p}(\theta)$ and insurance coverage $\hat{l}(\theta)$ increase with θ. This menu of contract design naturally accommodates customers' demands, because high-risk customers generally prefer to purchase more insurance than low-risk customers.

Proposition 2.2
(1) $\hat{l}(\theta) > 0$ for all $\theta \in [\theta_L, \theta_H]$;
(2) By choosing the optimal contract, the gain in type-θ customer's MV value is given by

$$\frac{\lambda T \gamma^2 \sigma(\theta)}{2(\hat{\eta}(\theta) + \gamma)} > 0.$$

(3) The insurer's expected profit, by offering this menu of contracts, is given by

$$\lambda T \gamma^2 \int_{\theta_L}^{\theta_H} \frac{\hat{\eta}(\theta)\sigma(\theta)}{2(\hat{\eta}(\theta) + \gamma)^2} \mathrm{d}G(\theta) > 0.$$

Proposition 2.2(1) establishes that our optimal menu has *no exclusion*, meaning that no risk type is excluded. Consequently, customers of any risk type can select a contract from this menu. It is important that our menu differs from the one proposed by Chade and Schlee (2012), as their menu excludes customers with low risk or low risk aversion (see Proposition 2 therein). In other words, the participants in the menu presented by Chade and Schlee (2012) are more likely to be of the high-risk type which is suboptimal for risk pooling from the insurer's perspective. On the contrary, our no-exclusion menu not only demonstrates greater efficiency from the view point of a social planner but also benefits the insurer. The absence of exclusions allows for a more diversified pool of customers, which is a vital advantage in risk management.

It is also worth noting that our model allows for the incorporation of *administrative costs* by adding a lower bound to the risk loading factor, e.g., all admissible loading factors should satisfy $\eta(\cdot) \geq K$, where the lowest admissible risk load K is currently set as 0. In the presence of costs, by the decreasing property of $\hat{\eta}(\theta)$, our menu will *exclude high-risk customers*, aligning well with insurance practices. In contrast, as discussed in Chade and Schlee (2020), the incorporation of costs into the optimal menu from monopoly models will exclude low-risk customers.

Part (2) means customers of any risk type are better off by participating in their optimal contracts, and part (3) means the insurer will also make a positive expected profit by offering this menu of contracts. Note that both parties (the customers and the insurer) continue to experience welfare gains when the insurer has no uncertainty. The menus addressed by traditional competitive or monopoly models do not exhibit this significant distinction. In the absence of information uncertainty, traditional competitive-model menus (e.g., Rothschild and Stiglitz (1976)) generate zero expected profit for the insurer, while participating in traditional monopoly-model menus (e.g., Stiglitz (1977) or Theorem 1(iii) of Chade and Schlee (2012)) offers no welfare gain for customers. However, our model incorporates a broad range of loss coverage functions. Increasing the risk loading prompts customers to purchase less insurance, potentially reducing the insurer's profit. This introduces a *natural trade-off* between supply and demand in the contract. Consequently, even within the current framework with a single insurer, she cannot fully eliminate all the surplus enjoyed by customers. This distinction is akin to introducing negotiation power or potential competitors to our model.

These appealing properties of our results can also be attributed to the utilization of the *Stackelberg game* formulation. In traditional competitive or monopoly models, these problems are formulated as one-step optimization with participation constraints, ensuring that customers or insurers are not worse off by participating. As a result, it is expected that the optimal solution is *binding (i.e.,* reached with equality). In contrast, the Stackelberg game formulation requires solving the customers' optimal demand with a given price (Theorem 2.1) in the first step, followed by determining the insurer's optimal pricing strategy in the second step. By considering the welfare of both customers and insurers in an optimal manner, our menu allows for a strict increase in their welfare upon participation.

Proposition 2.3 *Given that a type-θ contract is chosen, a customer's value function decreases if his risk type is raised. That is, for a fixed θ,*

$$\mathbb{E}\left[X_C^{\theta_0}(T,\hat{l}(\theta);\theta)\right]-\frac{\gamma}{2}\mathrm{Var}\left[X_C^{\theta_0}(T,\hat{l}(\theta);\theta)\right]$$

decreases with θ_0.

In addition to adverse selection, *moral hazard* is another significant challenge in the insurance industry. Moral hazard refers to the situation where a customer may alter their behavior after obtaining insurance, becoming more careless or even reckless. However, under a type-θ contract, the above proposition shows that a customer initially belonging to risk type θ will experience a disadvantage if their level of risk increases. Thus, our optimal menu effectively mitigates moral hazard, as customers are incentivized to minimize their own riskiness. Indeed, if the customer's riskiness decreases while the premium remains the same, the insurer's liability decreases and their profitability increases.

2.5 Comparison with the Worst-Case Contract

In the decision-making literature, another frequently used method for addressing information asymmetry is to consider the worst-case scenario through the utilization of the *maxmin expected utility*. In this section, we adhere to this principle by examining the worst-case contract and contrasting it with the optimal menu achieved in Theorem 2.2.

Suppose that the customer adopts an insurance coverage l, the insurer speculates the customer is of type θ, and the insurer charges loading factor η. We use the same premium rate as in (2.3), i.e.,

$$p(l,\eta;\theta):=\lambda\int_0^\infty l(y)f(y;\theta)\mathrm{d}y+\frac{\eta}{2}\lambda\int_0^\infty l(y)^2f(y;\theta)\mathrm{d}y. \tag{2.18}$$

It follows from (2.9) (with $\theta = \tilde{\theta}$) that the customer's demand function for optimal coverage is of the form

$$\mathring{l}(\eta) = \frac{\gamma}{\eta + \gamma} y. \tag{2.19}$$

Problem 2.3 (Insurer's worst-case problem) *Consider*

$$\sup_{\eta \geq 0} \inf_{\theta \in [\theta_L, \theta_H]} \mathbb{E}\left[X_I^{\theta}(T, \mathring{l}, \eta; \theta) \right]. \tag{2.20}$$

The optimization problem (2.3) involves two control variables, η and θ. The loading factor η is under the control of the insurer and thus shall be maximized. However, the customer's risk type θ is unknown to the insurer and beyond her control. To align with the insurer's worst-case estimation, we consider the infimum of the expected profit over all potential values of θ.

Theorem 2.3

(1) The worst-case contract to problem 2.3 is given by

$$\begin{cases} l^{wc}(y) = \frac{1}{2} y, \\ p^{wc} = \frac{\lambda}{2} \mu(\theta_L) + \frac{\lambda \gamma}{8} \sigma(\theta_L). \end{cases} \tag{2.21}$$

(2) Customers of any risk type $\theta \in [\theta_L, \theta_H]$ would be better-off by participating in the worst-case contract.

(3) The insurer is better-off offering the worst-case contract only if

$$\int_{\theta_L}^{\theta_H} \mu(\theta) \mathrm{d}G(\theta) < \mu(\theta_L) + \frac{\gamma}{4} \sigma(\theta_L). \tag{2.22}$$

We observe several disadvantages of the worst-case contract (2.21) in comparison with the optimal menu of contracts (2.15). First, the worst-case contract offers only a single option, and thus is incapable of effectively eliciting the customer's true risk type. Second, the worst-case contract exhibits low efficiency. Even though all customers are willing to participate, we see from (2.22) that it is not opportune for the insurer to offer the worst-case contract when the customer is more likely to have a high risk type, i.e. if the distribution function $G(\theta)$ has a heavy right tail near θ_H. In this case, the worst-case contract would not exist, resulting in the absence of insurance coverage for all individuals. This is consistent with the insight from the seminal work Akerlof (1970), which suggests that the worst consequence of information asymmetry is *no trade*. In contrast, the optimal menu of contracts provides the advantage of non-exclusion, ensuring that the insurer has a positive expected profit and every customer receives a portion of the risk coverage, as stated in Proposition 2.2(1). Therefore, compared with the worst-case contract, our optimal menu increases insurer's expected profit and enhances the likelihood of trading.

2.6 Curvatures

Our previous results have shown that the risk loading $\hat{\eta}(\theta)$ in the optimal menu decreases with θ, indicating that discounts are provided to high-risk customers, which holds for any customer's risk type distribution Z^{θ}. However, the traditional approach in the literature uses the premium-coverage curve to determine whether quantity discounts or premia exist. In this subsection, we use an example to demonstrate that the premium-coverage curve can be misleading and may lead to the opposite conclusion. Suppose the claim size Z^{θ} follows a uniform distribution on $[0, \theta]$,

In the first numerical study, we assume that θ follows a uniform distribution. Specifically, we consider three distributions for θ: (1) a *no-uncertainty* case with constant value $\theta = 1.5$; (2) a *low-uncertainty* case with $\theta \sim U(1.25, 1.75)$; and (3) a *high-uncertainty* case with $\theta \sim U(1, 2)$. The three scenarios share the same mean of 1.5, but differ in their *levels of uncertainty*, as reflected by the range of their supports (equivalently, their variances). In addition, we set $\lambda = 1$ and $\gamma = 0.5$.

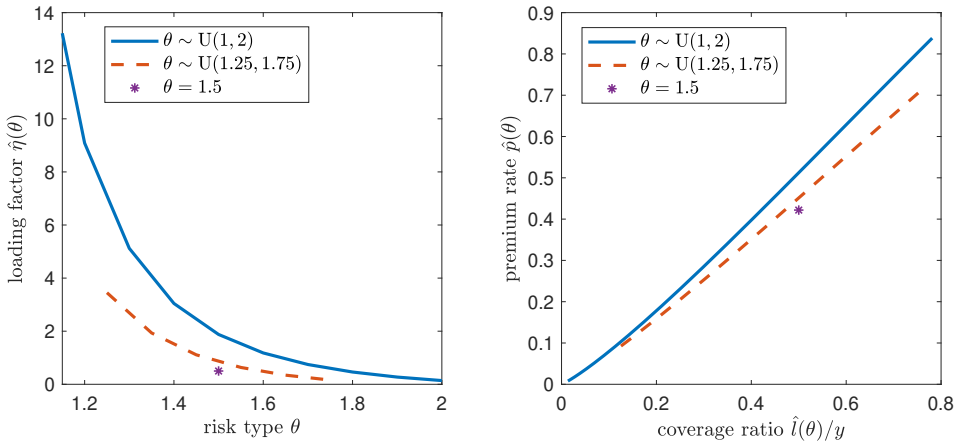

Figure 1: Risk loading function and premium-coverage curve under risk-type uncertainty for θ follows a uniform distribution

We examine two types of curves in Figure 1: the risk loading function ($\hat{\eta}(\theta)$ versus θ) in the left panel and the premium-coverage curve ($\hat{p}(\theta)$ versus $\hat{l}(\theta)/y$) in the right panel. Since the menu reduces to a single contract in the absence of uncertainty (Corollary 2.1), it appears as a star-shaped point in all panels for $\theta = 1.5$. With an increase in uncertainty, changing from Uniform $(1.25, 1.75)$ to Uniform $(1, 2)$, a broader menu is offered.

In the left panel of Figure 1, we see that the risk loading factor $\hat{\eta}(\theta)$ decreases with the risk type θ which is predicted by Proposition 2.1(1). In fact, $\hat{\eta}(\theta)$ also declines with the coverage ratio $\hat{l}(\theta)/y$. This relationship follows directly from equation (2.15), which implies

$$\hat{\eta}(\theta) = \gamma \left(\frac{1}{\hat{l}(\theta)/y} - 1 \right). \tag{2.23}$$

Therefore, the monotonic patterns imply risk loading discounts are granted to higher-risk customers and to those opting for greater coverage.

By contrast, the right depicts premium-coverage curve showing that the premium rate $\hat{p}(\theta)$ increases with the coverage ratio $\hat{l}(\theta)/y$. This is expected, as a contract with higher coverage commands a higher cost. However, it is rather surprising that the premium-cover curve appears to be *convex*, which may lead to the *opposite* inference—the insurer is charging quantity premia rather than offering discounts! Therefore, one of the main messages of the paper is to use risk loading (premium minus expected loss), rather than the premium itself, to examine the existence of discounts or premia. Indeed, the two decreasing property of risk loading shown in the two top panels hold for all loss distributions, as predicted by Proposition 2.1(1) and equation (2.23). Note that the contradictory inferences based on risk loading and premium will become even more evident in the case of risk-attitude uncertainty, as shown in Figure 3 in the next section.

Additionally, the shapes of our loading-coverage and premium-coverage curves indicate that our model distinguishes itself from classical monopoly models and actually shares some commonalities

with competitive models. Chade and Schlee (2012) demonstrates that the premium-coverage curve for monopoly insurance can exhibit a backward-S shape, first concave then convex. This distinction contrasts with findings in Maskin and Riley (1984), where a "standard" monopolist offers quantity discounts and the price-quantity curve is globally concave under mild conditions. Figure 1 shows that our premium-coverage curve can be globally convex, aligning with a feature of competitive models (Rothschild and Stiglitz (1976)). Moreover, our decreasing loading-coverage relation, equation (2.23), implies that the insurer does offer quantity discounts, but in terms of risk loading. This is more aligned with the implications from Maskin and Riley (1984) for general monopoly models.

Lastly, both panels demonstrate that *the risk loading and the premium increase* as the *level of uncertainty rises*. In particular, for the customer with true risk type $\theta_0 = 1.5$, the loading factor $\hat{\eta}(\theta_0)$ and the premium rate $\hat{p}(\theta_0)$ are the lowest in the no-uncertainty case, increase under the low-uncertainty case $\theta \sim U(1.25, 1.75)$, and increase further in the high-uncertainty case with $\theta \sim U(1, 2)$. This pattern is expected, since greater uncertainty about customers' true risk types leads the insurer to charge higher premiums and impose higher risk loadings to compensate for the increased informational risk.

To further examine the impact of uncertainty, we also model θ using truncated normal distributions supported on the interval $[1, 10]$. Similarly, we consider three distributions for θ to represent different levels of uncertainty: (1) a no-uncertainty case with constant value $\theta = 1.5$; (2) a low-uncertainty case with $\theta \sim \text{TruncN}(1.5, 0.2)$ (i.e., mean 1.5 and standard deviation 0.2); and (3) a high-uncertainty case with $\theta \sim \text{TruncN}(1.5, 0.5)$ (i.e., mean 1.5 and standard deviation 0.5). All three scenarios share the same mean of 1.5, while the standard deviation, reflecting the level of uncertainty, increases from 0 to 0.2 and then to 0.5. In addition, we set $\lambda = 1$ and $\gamma = 10$.

Notably, the right panel of Figure 2 reveals that the premium-coverage relationship is not globally convex but rather S-shaped, while the left panel of Figure 2 still shows higher-risk insureds benefit from risk-loading discounts in both cases. Again, these findings provide robust empirical evidence supporting our theoretical framework.

Moreover, for any fixed $\theta_0 \in [1, 10]$, the risk loading factor $\hat{\xi}(\theta_0)$ and the premium rate $\hat{p}(\theta_0)$ from $\text{TruncN}(1.5, 0.5)$ is higher than the one from $\text{TruncN}(1.5, 0.2)$. This observation is consistent with Figure 1 and reinforces the insight that greater uncertainty leads to higher pricing.

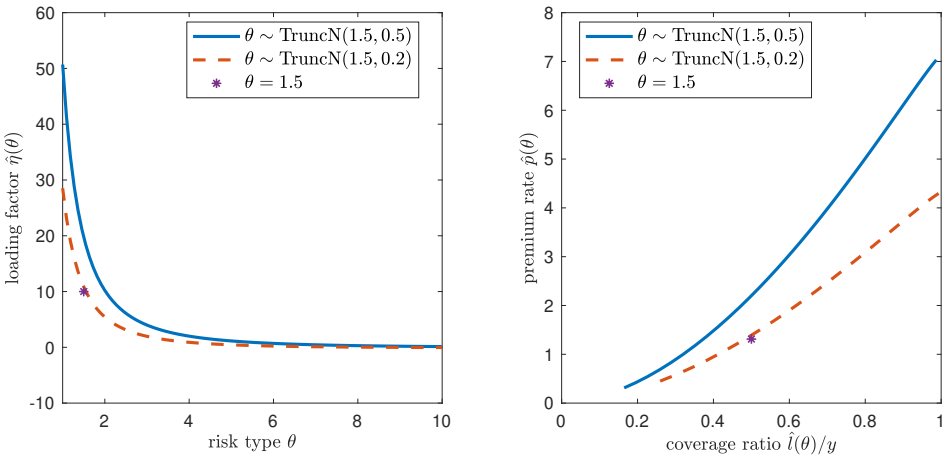

Figure 2: Risk loading function and premium-coverage curve under risk-type uncertainty for θ follows a truncated normal distribution

3 Optimal Menu with Risk-Attitude Uncertainty

3.1 Modelling and Problem Formulation

We now consider the case where the insurer knows the customers' true risk type, but is uncertain about their *true risk aversion* parameter γ. The insurer models risk aversion γ as an independent random variable with cumulative distribution function $G(\gamma)$ on compact support $[\gamma_L, \gamma_H] \subset [0, \infty)$.

If a customer selects a contract for risk aversion $\tilde{\gamma}$ and insurance coverage $l \in \mathbb{A}_l$, the continuously-charged premium rate paid by the customer at time t is given by

$$p(l; \tilde{\gamma}) = \lambda \int_0^\infty \left(l(y) + \frac{\eta(\tilde{\gamma})}{2} l(y)^2 \right) f(y) \mathrm{d}y, \tag{3.1}$$

where $\eta(\tilde{\gamma}) \geq 0$ is the loading factor for risk aversion $\tilde{\gamma}$. The set of all admissible loading factors is denoted by

$$\mathbb{A}_\eta = \{\eta : \eta(\tilde{\gamma}) \geq 0 \text{ for all } \tilde{\gamma} \in [\gamma_L, \gamma_H]\}.$$

The customer's surplus is given by

$$\begin{aligned} X_C(t, l; \tilde{\gamma}) &= x_0 - p(l; \tilde{\gamma})t - U(t) + \int_0^t \int_0^\infty l(y) N(\mathrm{d}s, \mathrm{d}y) \\ &= x_0 - p(l; \tilde{\gamma})t - \int_0^t \int_0^\infty (y - l(y)) N(\mathrm{d}s, \mathrm{d}y), \end{aligned}$$

with initial wealth level x_0. Correspondingly, the insurer's profit evolves according to

$$X_I(t, l; \tilde{\gamma}) = p(l; \tilde{\gamma})t - \int_0^t \int_0^\infty l(y) N(\mathrm{d}s, \mathrm{d}y). \tag{3.2}$$

Here, we suppress the superscript θ to the aggregate loss U^θ and the Poisson random measure N^θ defined in (2.2) because risk type is known.

We then present the customer's and insurer's problems. For a customer with true risk aversion γ choosing a contract for risk aversion $\tilde{\gamma}$, his objective is to maximize the MV of the surplus by selecting an optimal loss coverage function l.

Problem 3.1 (Customer's problem) *We define the value function for a customer with true risk aversion γ choosing a contract for risk aversion $\tilde{\gamma}$ as follows:*

$$V_C^\gamma(\tilde{\gamma}) = \sup_{l \in \mathbb{A}_l} \left\{ \mathbb{E}\left[X_C(T, l; \tilde{\gamma}) \right] - \frac{\gamma}{2} \mathrm{Var}\left[X_C(T, l; \tilde{\gamma}) \right] \right\}. \tag{3.3}$$

Correspondingly, the customer's optimal insurance strategy to (3.3) is denoted by $\hat{l}^\gamma(\tilde{\gamma}) := \{\hat{l}^\gamma(y; \tilde{\gamma})\}_{y>0}$.

The insurer aims to maximize her weighted expected profit across all priors of the customer's risk aversion, while also seeking to elicit the true information regarding the customer's risk aversion. Condition (3.4) is the corresponding truth-telling constraint for risk-attitude uncertainty.

Problem 3.2 (Insurer's problem) *Consider*

$$\sup_{\eta(\cdot) \in \mathbb{A}_\eta} \int_{\gamma_L}^{\gamma_H} \mathbb{E}\left[X_I(T, \hat{l}^\gamma(\gamma); \gamma) \right] \mathrm{d}G(\gamma)$$

such that

$$\gamma \in \mathrm{argmax}_{\tilde{\gamma} \in [\gamma_L, \gamma_H]} V_C^\gamma(\tilde{\gamma}). \tag{3.4}$$

3.2 Solution of the Optimal Menu

Given that risk type θ is known, we briefly write

$$\mu := \mu(\theta) = \mathbb{E}[Z^\theta] \text{ and } \sigma := \sigma(\theta) = \mathbb{E}[(Z^\theta)^2].$$

We first present the solution to the customer's problem. Note that the optimal loss coverage is still a proportional strategy.

Theorem 3.1 *An optimal insurance coverage to the customer's problem (3.3) is given by*

$$\hat{l}^\gamma(\tilde{\gamma}) := \hat{l}^\gamma(y;\tilde{\gamma}) = \frac{\gamma}{\eta(\tilde{\gamma})+\gamma} y, \tag{3.5}$$

and the value function is given by

$$V_C^\gamma(\tilde{\gamma}) = x_0 - \lambda T\mu - \frac{\lambda T\gamma\eta(\tilde{\gamma})\sigma}{2(\eta(\tilde{\gamma})+\gamma)}. \tag{3.6}$$

We then derive the first-order condition of the truth-telling constraint (3.4). Differentiating (3.6) with respect to $\tilde{\gamma}$ yields the first order condition, $\frac{\partial V_C^\gamma(\tilde{\gamma})}{\partial\tilde{\gamma}}\Big|_{\tilde{\gamma}=\gamma} = 0$, which further implies that

$$\eta'(\tilde{\gamma}) = 0 \quad \text{for all } \tilde{\gamma}. \tag{3.7}$$

Hence, the risk loading factor associated with truth-telling will remain constant for all degrees of risk aversion $\tilde{\gamma}$. It is intriguing to observe that, in contrast to situations involving uncertainty related to risk types, the insurer will employ a *linear pricing* approach to elicit true information regarding the customer's risk aversion. This means that the loading factor will be the same for all customers, and no discounts or premia are offered.

The next theorem summarizes the optimal loading factor, optimal menu of contracts with risk-attitude uncertainty, and verifies the truth-telling constraint.

Theorem 3.2

(1) An optimal loading factor to the relaxed insurer's problem, i.e.,

$$\sup_{\eta\in\mathbb{A}_\eta} \int_{\gamma_L}^{\gamma_H} \mathbb{E}\left[X_I(T,\hat{l}^\gamma(\gamma);\gamma)\right] \mathrm{d}G(\gamma),$$

subject to the first-order condition (3.7) is given by

$$\hat{\eta} = \mathrm{argmax}_{\eta\geq 0} \int_{\gamma_L}^{\gamma_H} \frac{\eta\gamma^2}{(\eta+\gamma)^2} \mathrm{d}G(\gamma). \tag{3.8}$$

(2) An optimal menu of contracts with loss coverage $\hat{l}(\gamma) := \hat{l}^\gamma(y;\gamma)$ and premium rate $\hat{p}(\gamma) := \hat{p}(\hat{l}(\gamma);\gamma)$ for $\gamma\in[\gamma_L,\gamma_H]$ is given by

$$\begin{cases} \hat{l}(\gamma) = \frac{\gamma}{\hat{\eta}+\gamma} y, \\ \hat{p}(\gamma) = \frac{\lambda\gamma\mu}{\hat{\eta}+\gamma} + \frac{\lambda\gamma^2\hat{\eta}\sigma}{2(\hat{\eta}+\gamma)^2}, \end{cases} \tag{3.9}$$

where $\hat{\eta}$ is given in (3.8).

(3) The insurer's optimal value function is given by

$$\int_{\gamma_L}^{\gamma_H} \mathbb{E}\left[X_I(T, \hat{l}^{\gamma}(\gamma); \gamma)\right] \mathrm{d}G(\gamma) = \frac{\lambda T \sigma}{2} \int_{\gamma_L}^{\gamma_H} \frac{\hat{\eta}\gamma^2}{(\hat{\eta}+\gamma)^2} \mathrm{d}G(\gamma).$$

(4) The optimal menu of contracts $(\hat{l}(\gamma), \hat{p}(\gamma))_{\gamma\in[\gamma_L,\gamma_H]}$ satisfies the truth-telling constraint in the sense that

$$\gamma = \mathrm{argmax}_{\tilde{\gamma}\in[\gamma_L,\gamma_H]} \left\{\mathbb{E}\left[X_C(T, \hat{l}(\tilde{\gamma}); \tilde{\gamma})\right] - \frac{\gamma}{2}\mathrm{Var}\left[X_C(T, \hat{l}(\tilde{\gamma}); \tilde{\gamma})\right]\right\},$$

where $\{X_C(t, \hat{l}(\tilde{\gamma}); \tilde{\gamma})\}_{t\in[0,T]}$ is the customer's surplus process if he chooses the contract $(\hat{l}(\tilde{\gamma}), \hat{p}(\tilde{\gamma}))$ for risk aversion $\tilde{\gamma}$.

Remark 3.1 *Theorem 3.2 characterizes the optimal menu of contracts under risk-attitude uncertainty. Similar to Theorem 2.2, the optimal menu $(\hat{l}(\gamma), \hat{p}(\gamma))_{\gamma\in[\gamma_L,\gamma_H]}$, given in* (3.9)*, consists of pairs of loss coverage functions and premium rates, and is confirmed to satisfy the truth-telling constraint. It is worth noting that, although the optimal risk loading $\hat{\eta}$ remains constant across all levels of risk aversion, both the optimal loss coverage functions $\hat{l}(\gamma)$ and the corresponding premium rate $\hat{p}(\gamma)$ increase with risk aversion parameter γ. This implies that more risk-averse customers are offered contracts with higher coverage and higher premiums, but intrinsically, the insurer applies the same risk loading $\hat{\eta}$ across all contracts.*

In particular, in the absence of information asymmetry, when the insurer knows that a customer's degree of risk aversion is γ, it follows from (3.8) *that the optimal risk loading is given by*

$$\hat{\eta} = \mathrm{argmax}_{\eta\geq 0} \frac{\eta\gamma^2}{(\eta+\gamma)^2} = \gamma.$$

Consequently, the menu concentrates to a single contract

$$\begin{cases} \hat{l} = \frac{1}{2}y, \\ \hat{p} = \frac{\lambda}{2}\mu + \frac{\lambda\gamma}{8}\sigma, \end{cases}$$

which coincides with Corollary 2.1 and Chen and Shen (2019).

3.3 Properties of the Optimal Menu

The optimal contract with risk-attitude uncertainty has several properties, which we enumerate below. The proofs are straightforward and can be obtained easily from Theorems 3.1 and 3.2.

Proposition 3.1

(1) $\hat{\eta}$ is independent of γ;

(2) $\hat{l}(\gamma)$ and $\hat{p}(\gamma)$ are increasing in γ;

(3) $\hat{l}(\gamma) > 0$ for all $\gamma \in [\gamma_L, \gamma_H]$;

(4) By choosing the optimal contract, the gain for a customer with risk aversion γ is given by

$$\frac{\lambda T \gamma^2 \sigma}{2(\hat{\eta}+\gamma)} > 0.$$

(5) The insurer's expected gain in profit by offering this menu of contracts is given by

$$\frac{\lambda T \sigma}{2} \int_{\gamma_L}^{\gamma_H} \frac{\hat{\eta}\gamma^2}{(\hat{\eta}+\gamma)^2} \mathrm{d}G(\gamma) > 0. \tag{3.10}$$

As mentioned earlier, compared to the case of risk-type uncertainty, the primary distinction of the optimal menu in the presence of risk aversion uncertainty is that the optimal loading factor $\hat{\eta}$ remains constant for all levels of risk aversion, as stated in part (1). This suggests that a *linear pricing* strategy is employed to elicit information regarding the customer's true risk aversion. The insurer does not offer loading discounts in this case because the customer's risk attitude, which represents only the willingness to purchase insurance, has no direct impact on actual claim experience or the insurer's profit, and is therefore less valuable information.

The other properties are similar to those obtained under the risk-type uncertainty. Specifically, part (2) indicates that a more risk-averse customer selects a contract with higher coverage while paying a higher premium. It provides a theoretical support for empirical studies that find *no* evidence of strong adverse selection (e.g., Cawley and Philipson (1999) and Finkelstein and McGarry (2006)), as customers opting for higher insurance coverage can be those who are more risk-averse. Part (3) denotes that there is no exclusion based on risk aversion, ensuring that insurers accommodate all levels of risk aversion. Parts (4) and (5) imply that both the customers and the insurer are better off from participating/offering this menu of contracts.

3.4 Comparison with the Worst-Case Contract

In the presence of risk-attitude uncertainty, the insurer also has the option to consider the worst-case contract. Following a similar approach based on the maxmin expected utility, the formulation of the worst-case problem is presented as follows.

Problem 3.3 (Insurer's worst-case problem) *Consider*

$$\sup_{\eta\geq 0}\inf_{\gamma\in[\gamma_L,\gamma_H]}\mathbb{E}\left[X_I(T,\mathring{l}(\eta);\gamma)\right],$$

where

$$\mathring{l}(\eta)=\frac{\gamma}{\eta+\gamma}y \tag{3.11}$$

is customer's optimal coverage function solved in (3.5).

Theorem 3.3

(1) The worst-case contract to Problem 3.3 is given by

$$\begin{cases} l^{wc}(y)=\frac{1}{2}y, \\ p^{wc}=\frac{\lambda\mu}{2}+\frac{\lambda\sigma\gamma_L}{8}. \end{cases} \tag{3.12}$$

(2) All customers are better of by participation in the worst-case contract.

(3) Insurer's expected profit by offering this worst-case contract is given by

$$V_I^{wc}=\frac{\lambda T\sigma\gamma_L}{8}.$$

However,

$$V_I^{wc}<V_I,$$

where $V_I:=\frac{\lambda T\sigma}{2}\int_{\gamma_L}^{\gamma_H}\frac{\hat{\eta}\gamma^2}{(\hat{\eta}+\gamma)^2}\mathrm{d}G(\gamma)$ *is the insurer's expected profit from offering our optimal menu of contracts as given in (3.10).*

Similar to the case with risk-type uncertainty, the worst-case contract (3.12) only has a single contract, and thus the true risk aversion of a customer cannot be elicited even though all customers are willing to participate. Moreover, the insurer's expected profit is lower compared to offering the optimal menu of contracts.

3.5 Curvatures

In this subsection, we examine the risk loading function ($\hat{\eta}(\gamma)$ versus γ) and the premium-coverage curve ($\hat{p}(\gamma)$ versus $\hat{l}(\gamma)/y$) in the presence of risk-attitude uncertainty.

Figure 3 depicts the two types of curves for the following example, but note that the key findings—the flatness of risk loading and the convexity of premium-coverage curve—generally hold for all distributions and parameters. We assume the claim size Z follows a uniform distribution on $[0, 1.5]$ and set $\lambda = 1$.

We consider three uniform distributions for the risk aversion parameter γ to reflect different levels of uncertainty: (1) a no-uncertainty case with constant value $\gamma = 5$; (2) a low-uncertainty case with $\gamma \sim U(1.5, 8.5)$; and (3) a high-uncertainty case with $\gamma \sim U(0.1, 9.9)$. They share the same mean of 5, but differ in their ranges of support and variances.

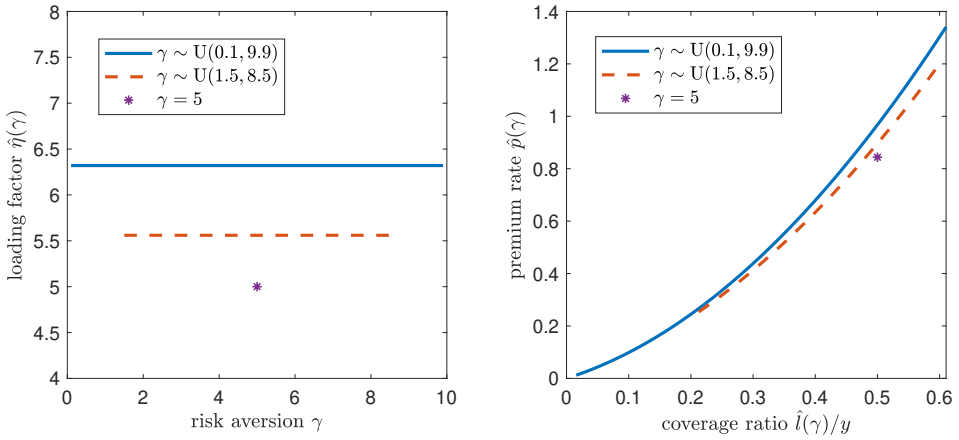

Figure 3: Risk loading function and premium-coverage curve under risk-attitude uncertainty for γ follows a uniform distribution

The left panel depicts the risk loading function $\hat{\eta}(\gamma)$. As predicted by Proposition 3.1(1), the risk loading in the optimal menu remains constant, and independent of risk aversion and the coverage ratio. Therefore, the curves are *flat* indicating that the insurer does not offer discounts or premia in risk loading.

However, it is observed that the premium-coverage curve in the right panel is evidently *globally convex*, suggesting quantity premia from the traditional perspective on the price-quantity relationship. In fact, this convexity holds generally as one can rewrite the premium rate in equation (3.9) as

$$\hat{p}(\gamma) = \lambda\mu\left(\hat{l}(\gamma)/y\right) + \frac{\lambda\hat{\eta}\sigma}{2}\left(\hat{l}(\gamma)/y\right)^2. \tag{3.13}$$

This indicates that the premium rate $\hat{p}(\gamma)$ is actually a quadratic function of the coverage ratio $\hat{l}(\gamma)/y$. This reinforces our key message that it might be more reliable to infer the presence of discounts or premia by analyzing the risk loading rather than focusing solely on the premium.

Besides, both panels demonstrate that the insurer sets higher risk loadings and premium rates when facing greater levels of uncertainty.

In Figure 4, we extend our analysis by replacing the previously assumed uniform distribution for the risk aversion parameter γ with truncated normal distributions supported on the interval $[1, 9]$. Specifically, we consider three scenarios for the distribution of γ: (1) a no-uncertainty case with constant value $\gamma = 1.5$; (2) a low-uncertainty case with $\gamma \sim \text{TruncN}(1.5, 0.5)$ (i.e., mean

1.5 and standard deviation 0.5); and (3) a high-uncertainty case with $\gamma \sim \text{TruncN}(1.5, 2)$. All three scenarios share the same mean of 1.5, while the standard deviation, reflecting the level of uncertainty, increases across the cases.

The resulting patterns remain broadly consistent with those observed under the uniform distribution, thereby demonstrating the robustness of our results to alternative distributional assumptions regarding risk preferences.

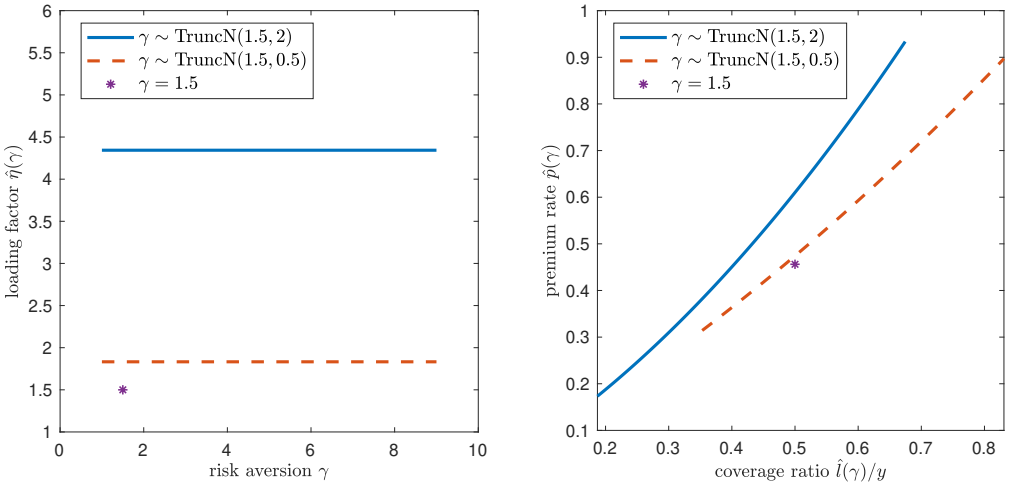

Figure 4: Risk loading function and premium-coverage curve under risk-attitude uncertainty for γ follows a truncated normal distribution

4 Concluding Remarks

In this paper, we introduce a Stackelberg optimal risk-sharing approach to study optimal insurance with information asymmetry, where the insurer has imperfect information about the customer's risk type or risk attitude. Assuming the customer adopts the MV criterion and the insurer applies the variance premium principle, we derive closed-form optimal menus of contracts for both cases. In the case of risk-type uncertainty, we find a decrease in risk loading for high-risk types, suggesting risk-loading discounts and nonlinear pricing structure. However, the insurer does not offer risk-load discounts when she is only uncertain about the customer's risk attitude. For comparison, we also derive the worst-case contract according to maxmin expected utility, and we find that our optimal menu increases the insurer's expected profit and enhances the likelihood of trading. Our key messages are twofold. (1) The information of customers' risk types is more valuable than their risk attitudes, since the latter only represents the willingness to purchase insurance but has no direct impact on claim experience. (2) We suggest analyzing the risk loading, not just the premium, to infer the presence of discounts or premia.

Several extensions can be considered for future research. Building upon the current Stackelberg game framework with hidden information, one natural and important direction is to incorporate competitive insurance markets where insurers hold heterogeneous beliefs about customer risk distributions. Such an extension would introduce endogenous market segmentation driven by belief-induced pricing differences and strategic interactions among insurers, offering richer insights into market dynamics and structure. Another promising avenue under the same framework is to explore static models with more general utility functions to capture a broader range of customer risk prefer-

ences. It would be worthwhile to incorporate customer's private actions that lie beyond the control of insurers, as exemplified in Kao et al. (2022) for business interruption insurance. A particularly challenging and intriguing problem is to involve eliciting customers' uncertainties in both risk type and risk attitude simultaneously, while current works on multi-dimensional private information are limited to a finite number of provided contracts and cannot sort all missing information. Other valuable directions include examining alternative premium principles beyond the variance principle and exploring settings in which the insurer's observation of risk outcomes is partial or costly. We hope that the explicit solutions and analysis in this paper can shed light on this cutting-edge topic.

Acknowledgments

The authors thank the editor, an associate editor, and two anonymous referees for their helpful comments. Xia Han gratefully acknowledges support from the National Natural Science Foundation of China (Grant Nos. 12301604, 12371471, and 12471449). Bin Li gratefully acknowledges support from the Natural Sciences and Engineering Research Council of Canada (Grant No. 04338), and support from the National Natural Science Foundation of China (Grant Nos. 12271171 and 12471446). Yao Luo gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada (Grant No. 430202300061).

A Proofs

A.1 Proof of Theorem 2.1

Before proving Theorem 2.1, we first present several auxiliary results including Theorem 5.1 and Lemma 5.1, which will be used in the proof.

For any $l \in \mathbb{A}_l$ and $\varphi(t,x) \in C^{1,2}([0,T]\times\mathbb{R})$, i.e., first-order continuously differentiable in t and second-order continuously differentiable in x, we define the integro-differential operator

$$\begin{aligned}\mathcal{L}_l\varphi(t,x) :&= \lim_{\varepsilon\to 0}\frac{\mathbb{E}[\varphi(t+\varepsilon, X_C^\theta(t+\varepsilon,l;\tilde{\theta}))|X_C^\theta(t,l;\tilde{\theta})=x]-\varphi(t,x)}{\varepsilon}\\ &= \frac{\partial\varphi(t,x)}{\partial t} - \frac{\partial\varphi(t,x)}{\partial x}\lambda\int_0^\infty\left(l(t,y)+\frac{\eta(\tilde{\theta})}{2}l(t,y)^2\right)f(y;\tilde{\theta})\mathrm{d}y \\ &\quad + \lambda\int_0^\infty\left[\varphi(t,x-y+l(t,y))-\varphi(t,x)\right]f(y;\theta)\mathrm{d}y.\end{aligned} \tag{A.1}$$

To address the time inconsistency for the dynamic MV problem (2.7), we adopt the usual definition of equilibrium strategy below. For $(t,x)\in[0,T]\times\mathbb{R}$, we define

$$J^\theta(t,x,l;\tilde{\theta}) = \mathbb{E}\left[X_C^\theta(T,l;\tilde{\theta})\middle| X_C^\theta(t,l;\tilde{\theta})=x\right] - \frac{\gamma}{2}\mathrm{Var}\left[X_C^\theta(T,l;\tilde{\theta})\middle| X_C^\theta(t,l;\tilde{\theta})=x\right]. \tag{A.2}$$

Definition A.1 *Given $(t,x)\in[0,T]\times\mathbb{R}$ and $\hat{l}\in\mathbb{A}_l$, we define a perturbed strategy l^ε as*

$$l^\varepsilon(s,y) = \begin{cases}\bar{l}(y), & s\in[t,t+\varepsilon),\ y\in\mathbb{R},\\ \hat{l}(s,y), & s\in[t+\varepsilon,T],\ y\in\mathbb{R},\end{cases} \tag{A.3}$$

where $\bar{l}:\mathbb{R}\to\mathbb{R}$ and $\varepsilon>0$. Suppose that, for any given $(t,x)\in[0,T]\times\mathbb{R}$, and any such perturbed strategy $l^\varepsilon\in\mathbb{A}_l$, we have

$$\liminf_{\varepsilon\to 0}\frac{J^\theta(t,x,\hat{l};\tilde{\theta})-J^\theta(t,x,l^\varepsilon;\tilde{\theta})}{\varepsilon}\geq 0.$$

Then, $\hat{l}$ is called an equilibrium strategy for the dynamic MV problem (2.7).

Theorem A.1 (Verification theorem) *Suppose there exist $V(t,x), g(t,x) \in C^{1,2}([0,T])\times\mathbb{R}$ satisfy the following conditions:*

(1) For any $(t,x)\in[0,T]\times\mathbb{R}$,

$$\sup_{l\in\mathbb{R}}\left\{\mathcal{L}_l V(t,x) - \frac{1}{2}\gamma\mathcal{L}_l g^2(t,x) + \gamma g(t,x)\mathcal{L}_l g(t,x)\right\} = 0, \tag{A.4}$$

and $\hat{l}$ denotes the optimal value to achieve the supremum in l.

(2) For any $(t,x)\in[0,T]\times\mathbb{R}$,

$$\begin{cases} \mathcal{L}_{\hat{l}} g(t,x) = 0, \\ V(T,x) = g(T,x) = x. \end{cases} \tag{A.5}$$

(3) For any $(t,x)\in[0,T]\times\mathbb{R}$, $\hat{l}(t)$, $\mathcal{L}_{\hat{l}} V(t,x)$ and $\mathcal{L}_{\hat{l}} g^2(t,x)$ are deterministic functions of t and independent of x.

Then $\hat{l}$ is the equilibrium strategy and $V(t,x) = J^\theta(t,x,\hat{l};\tilde{\theta})$ is the equilibrium value function to the dynamic MV problem (2.7). Besides, $g(t,x) = \mathbb{E}\left[X_C^\theta(T,\hat{l};\tilde{\theta})\middle| X_C^\theta(t,\hat{l};\tilde{\theta}) = x\right]$.

Proof. By (A.1), Dynkin's formula and the first relation of (A.5), we have

$$\mathbb{E}\left[g(T, X_C^\theta(T,\hat{l};\tilde{\theta}))|X_C^\theta(t,\hat{l};\tilde{\theta}) = x\right] = g(t,x) + \mathbb{E}\left[\int_t^T \mathcal{L}_{\hat{l}} g(s, X_C^\theta(s,\hat{l};\tilde{\theta}))\mathrm{d}s|X_C^\theta(t,\hat{l};\tilde{\theta}) = x\right] = g(t,x).$$

On the other hand, it follows from the second relation of (A.5) that

$$g(t,x) = \mathbb{E}\left[g(T, X_C^\theta(T,\hat{l};\tilde{\theta}))|X_C^\theta(t,\hat{l};\tilde{\theta}) = x\right] = \mathbb{E}\left[X_C^\theta(T,\hat{l};\tilde{\theta})|X_C^\theta(t,\hat{l};\tilde{\theta}) = x\right]. \tag{A.6}$$

Next, we show $V(t,x) = J^\theta(t,x,\hat{l};\tilde{\theta})$. Since the optimal value in (A.4) is achieved at $\hat{l}$, by condition (3) and the first relation of (A.5), we can rewrite (A.4) as

$$\mathcal{L}_{\hat{l}} V(t,x) - \frac{1}{2}\gamma\mathcal{L}_{\hat{l}} g^2(t,x) = 0. \tag{A.7}$$

By the second relation of (A.5), Dynkin's formula and condition (3), one obtains

$$\mathbb{E}\left[X_C^\theta(T,\hat{l};\tilde{\theta})|X_C^\theta(t,\hat{l};\tilde{\theta}) = x\right] = \mathbb{E}\left[V(T, X_C^\theta(T,\hat{l};\tilde{\theta}))|X_C^\theta(t,\hat{l};\tilde{\theta}) = x\right] = V(t,x) + \int_t^T \mathcal{L}_{\hat{l}} V(s,\cdot)\mathrm{d}s.$$

Substituting (A.7) into the last equation, we have

$$V(t,x) = \mathbb{E}\left[X_C^\theta(T,\hat{l};\tilde{\theta})|X_C^\theta(t,\hat{l};\tilde{\theta}) = x\right] - \frac{1}{2}\gamma\int_t^T \mathcal{L}_{\hat{l}} g^2(s,\cdot)\mathrm{d}s. \tag{A.8}$$

To show the right hand side of (A.8) equals $J^\theta(t,x,\hat{l};\tilde{\theta})$, we shall find a connection between $\int_t^T \mathcal{L}_{\hat{l}} g^2(s,\cdot)\mathrm{d}s$ and the variance of $X_C^\theta(T,\hat{l};\tilde{\theta})$. Again, by the second relation of (A.5), Dynkin's formula, and condition (3), we have

$$\begin{aligned}
\mathbb{E}\left[X_C^\theta(T,\hat{l};\tilde{\theta})^2|X_C^\theta(t,\hat{l};\tilde{\theta}) = x\right] &= \mathbb{E}\left[g^2(T, X_C^\theta(T,\hat{l};\tilde{\theta}))|X_C^\theta(t,\hat{l};\tilde{\theta}) = x\right] \\
&= g^2(t,x) + \int_t^T \mathcal{L}_{\hat{l}} g^2(s,\cdot)\mathrm{d}s \\
&= \left(\mathbb{E}\left[X_C^\theta(T,\hat{l};\tilde{\theta})\middle| X_C^\theta(t,\hat{l};\tilde{\theta}) = x\right]\right)^2 + \int_t^T \mathcal{L}_{\hat{l}} g^2(s,\cdot)\mathrm{d}s.
\end{aligned}$$

Note that the above equation is equivalent to

$$\mathrm{Var}\left[X_C^\theta(T,\hat{l};\tilde{\theta})\middle| X_C^\theta(t,l;\tilde{\theta})=x\right]=\int_t^T \mathcal{L}_{\hat{l}}g^2(s,\cdot)\mathrm{d}s. \tag{A.9}$$

Finally, substituting (A.9) into (A.8) yields

$$V(t,x)=\mathbb{E}\left[X_C^\theta(T,\hat{l};\tilde{\theta})|X_C^\theta(t,\hat{l};\tilde{\theta})=x\right]-\frac{\gamma}{2}\mathrm{Var}\left[X_C^\theta(T,\hat{l};\tilde{\theta})\middle| X_C^\theta(t,l;\tilde{\theta})=x\right]=J^\theta(t,x,\hat{l};\tilde{\theta}), \tag{A.10}$$

where the second equality is due to (A.2).

Finally, we show that $\hat{l}$ is an equilibrium strategy. We know from Definition A.1 that it suffices to show

$$J^\theta(t,x,l^\varepsilon;\tilde{\theta})-J^\theta(t,x,\hat{l};\tilde{\theta})\le o(\varepsilon).$$

Thus, in the proof below, we devote to derive an expression of $J^\theta(t,x,l^\varepsilon;\tilde{\theta})-J^\theta(t,x,\hat{l};\tilde{\theta})$, and then show that it is bounded above by $o(\varepsilon)$.

By (A.2) and Lemma A.1 below, we have

$$\begin{aligned}
&J^\theta(t,x,l^\varepsilon;\tilde{\theta})\\
&=\mathbb{E}\left[X_C^\theta(T,l^\varepsilon;\tilde{\theta})-\frac{1}{2}\gamma X_C^\theta(T,l^\varepsilon;\tilde{\theta})^2|X_C^\theta(t,l^\varepsilon;\tilde{\theta})=x\right]+\frac{1}{2}\gamma\left(\mathbb{E}\left[X_C^\theta(T,l^\varepsilon;\tilde{\theta})\middle| X_C^\theta(t,l^\varepsilon;\tilde{\theta})=x\right]\right)^2\\
&=\mathbb{E}\left[\mathbb{E}\left[X_C^\theta(T,\hat{l};\tilde{\theta})-\frac{1}{2}\gamma X_C^\theta(T,\hat{l};\tilde{\theta})^2|X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta})\right]|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]\\
&+\frac{1}{2}\gamma\left(\mathbb{E}\left[\mathbb{E}\left[X_C^\theta(T,\hat{l};\tilde{\theta})|X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta})\right]|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]\right)^2\\
&=\mathbb{E}\left[J^\theta(t+\varepsilon,X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta}),\hat{l};\tilde{\theta})|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]\\
&-\frac{1}{2}\gamma\mathbb{E}\left[\left(\mathbb{E}\left[X_C^\theta(T,\hat{l};\tilde{\theta})|X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta})\right]\right)^2|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]\\
&+\frac{1}{2}\gamma\left(\mathbb{E}\left[\mathbb{E}\left[X_C^\theta(T,\hat{l};\tilde{\theta})|X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta})\right]|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]\right)^2\\
&=\mathbb{E}\left[J^\theta(t+\varepsilon,X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta}),\hat{l};\tilde{\theta})|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]\\
&-\frac{1}{2}\gamma\mathbb{E}\left[g^2\left(t+\varepsilon,X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta})\right)|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]\\
&+\frac{1}{2}\gamma\left(\mathbb{E}\left[g\left(t+\varepsilon,X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta})\right)|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]\right)^2,
\end{aligned}$$

where the last step is due to (A.6). By subtracting $J^\theta(t,x,\hat{l};\tilde{\theta})$ to both sides of the above equation, and inserting the extra terms of $g^2(t,x)$, one obtains

$$J^\theta(t,x,l^\varepsilon;\tilde{\theta})-J^\theta(t,x,\hat{l};\tilde{\theta}):=H_\varepsilon$$

where

$$\begin{aligned}
H_\varepsilon:=&\mathbb{E}\left[J^\theta(t+\varepsilon,X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta}),\hat{l};\tilde{\theta})-J^\theta(t,x,\hat{l};\tilde{\theta})|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]\\
&-\frac{1}{2}\gamma\left\{\mathbb{E}\left[g^2\left(t+\varepsilon,X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta})\right)|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]-g^2(t,x)\right\}\\
&+\frac{1}{2}\gamma\left\{\left(\mathbb{E}\left[g\left(t+\varepsilon,X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta})\right)|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]\right)^2-g^2(t,x)\right\}.
\end{aligned}$$

Next we show that $H_\varepsilon \leq o(\varepsilon)$. For ease of notation, we define the operator

$$\mathcal{L}_{l,\varepsilon}\varphi(t,x) := \mathbb{E}\left[\varphi\left(t+\varepsilon, X_C^\theta(t+\varepsilon,l;\tilde{\theta})\right)|X_C^\theta(t,l;\tilde{\theta})=x\right]-\varphi(t,x),$$

where $l \in \mathbb{A}_l, \varepsilon > 0$ is a small constant, and $\varphi \in C^{1,2}([0,T]\times\mathbb{R})$. We know from (A.1) that

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathcal{L}_{l,\varepsilon}\varphi(t,x) = \mathcal{L}_l\varphi(t,x). \tag{A.11}$$

Using the notation $\mathcal{L}_{l,\varepsilon}\varphi(t,x)$, we can rewrite H_ε as

$$H_\varepsilon = \mathcal{L}_{\bar{l},\varepsilon}J^\theta(t,x,\hat{l};\tilde{\theta}) - \frac{1}{2}\gamma\mathcal{L}_{\bar{l},\varepsilon}g^2(t,x) + \frac{1}{2}\gamma\left\{\left(\mathbb{E}\left[g\left(t+\varepsilon, X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta})\right)|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]\right)^2 - g^2(t,x)\right\}. \tag{A.12}$$

Since we will mainly use the extended HJB (A.4) to show the right hand side of (A.12) is bounded above by $o(\varepsilon)$, we shall rewrite $(\mathbb{E}[g(t+\varepsilon, X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta}))|X_C^\theta(t,\bar{l};\tilde{\theta})=x])^2$ using $\mathcal{L}_{\bar{l},\varepsilon}g(t,x)$. Thus, by Dynkin's formula, we have

$$\begin{aligned}
&\mathbb{E}\left[g\left(t+\varepsilon, X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta})\right)|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]\\
&= g(t,x) + \mathbb{E}\left[\int_t^{t+\varepsilon}\mathcal{L}_{\bar{l},\varepsilon}g\left(s, X_C^\theta(s,\bar{l};\tilde{\theta}))\right)\mathrm{d}s|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right],
\end{aligned}$$

which implies that

$$\begin{aligned}
&\left(\mathbb{E}\left[g\left(t+\varepsilon, X_C^\theta(t+\varepsilon,\bar{l};\tilde{\theta})\right)|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]\right)^2 - g^2(t,x)\\
&= 2g(t,x)\mathbb{E}\left[\int_t^{t+\varepsilon}\mathcal{L}_{\bar{l},\varepsilon}g\left(s, X_C^\theta(s,\bar{l};\tilde{\theta}))\right)\mathrm{d}s|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right] + o(\varepsilon).
\end{aligned} \tag{A.13}$$

Substituting (A.13) into (A.12) yields

$$H_\varepsilon - o(\varepsilon) = \mathcal{L}_{\bar{l},\varepsilon}J^\theta(t,x,\hat{l};\tilde{\theta}) - \frac{1}{2}\gamma\mathcal{L}_{\bar{l},\varepsilon}g^2(t,x) + \gamma g(t,x)\mathbb{E}\left[\int_t^{t+\varepsilon}\mathcal{L}_{\bar{l},\varepsilon}g\left(s, X_C^\theta(s,\bar{l};\tilde{\theta}))\right)\mathrm{d}s|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]. \tag{A.14}$$

It is only left to show that the right-hand side of (A.14) is bounded above by $o(\varepsilon)$. By the extended HJB (A.4), we have

$$\mathcal{L}_{\bar{l}}V(t,x) - \frac{1}{2}\gamma\mathcal{L}_{\bar{l}}g^2(t,x) + \gamma g(t,x)\mathcal{L}_{\bar{l}}g(t,x) \leq 0. \tag{A.15}$$

Also, by (A.10), we have

$$\mathcal{L}_{\bar{l}}V(t,x) = \mathcal{L}_{\bar{l}}J^\theta(t,x,\hat{l};\tilde{\theta}).$$

Substituting the last equation into (A.15) yields

$$\mathcal{L}_{\bar{l}}J^\theta(t,x,\hat{l};\tilde{\theta}) - \frac{1}{2}\gamma\mathcal{L}_{\bar{l}}g^2(t,x) + \gamma g(t,x)\mathcal{L}_{\bar{l}}g(t,x) \leq 0.$$

This together with (A.11) implies that

$$o(\varepsilon) \geq \mathcal{L}_{\bar{l},\varepsilon}J^\theta(t,x,\hat{l};\tilde{\theta}) - \frac{1}{2}\gamma\mathcal{L}_{\bar{l},\varepsilon}g^2(t,x) + \gamma g(t,x)\mathbb{E}\left[\int_t^{t+\varepsilon}\mathcal{L}_{\bar{l},\varepsilon}g\left(s, X_C^\theta(s,\bar{l};\tilde{\theta}))\right)\mathrm{d}s|X_C^\theta(t,\bar{l};\tilde{\theta})=x\right]. \tag{A.16}$$

Finally, substituting (A.16) to (A.14) leads to

$$J^\theta(t,x,l^\varepsilon;\tilde{\theta}) - J^\theta(t,x,\hat{l};\tilde{\theta}) = H_\varepsilon \leq o(\varepsilon),$$

which implies that $\hat{l}$ is an equilibrium strategy. ■

Lemma A.1 *Under condition (3) of Theorem A.1, consider the deterministic strategy*

$$l^{\varepsilon}(s,y)=\begin{cases}\bar{l}(y), & s\in[t,t+\varepsilon),\ y\in\mathbb{R},\\ \hat{l}(s,y), & s\in[t+\varepsilon,T],\ y\in\mathbb{R},\end{cases}$$

where $\bar{l}:\mathbb{R}\to\mathbb{R}$ and $\varepsilon>0$. For any measurable function $f:\mathbb{R}\to\mathbb{R}$ such that the following conditional expectations are finite, we have

$$\mathbb{E}\left[f\left(X_C^{\theta}(T,l^{\varepsilon};\tilde{\theta})\right)|X_C^{\theta}(t,l^{\varepsilon};\tilde{\theta})=x\right]=\mathbb{E}\left[\mathbb{E}\left[X_C^{\theta}(T,\hat{l};\tilde{\theta})|X_C^{\theta}(t+\varepsilon,\bar{l};\tilde{\theta})\right]|X_C^{\theta}(t,\bar{l};\tilde{\theta})=x\right]. \quad (A.17)$$

Proof. Let $f:\mathbb{R}\to\mathbb{R}$ be a measurable function such that all the involved conditional expectations are finite. From the definition of the strategy l^{ε}, we have

$$X_C^{\theta}(t+\varepsilon,l^{\varepsilon};\tilde{\theta})\overset{\mathrm{d}}{=}X_C^{\theta}(t+\varepsilon,\bar{l};\tilde{\theta}),\quad\text{given that } X_C^{\theta}(t,l^{\varepsilon};\tilde{\theta})=X_C^{\theta}(t,\bar{l};\tilde{\theta})=x.$$

Thus, by the strong Markov property,

$$\begin{aligned}&\mathbb{E}\left[f\left(X_C^{\theta}(T,l^{\varepsilon};\tilde{\theta})\right)|X_C^{\theta}(t,l^{\varepsilon};\tilde{\theta})=x\right]\\&=\mathbb{E}\left[\mathbb{E}\left[f\left(X_C^{\theta}(T,l^{\varepsilon};\tilde{\theta})\right)|X_C^{\theta}(t+\varepsilon,l^{\varepsilon};\tilde{\theta})\right]|X_C^{\theta}(t,l^{\varepsilon};\tilde{\theta})=x\right]\\&=\mathbb{E}\left[\mathbb{E}\left[f\left(X_C^{\theta}(T,l^{\varepsilon};\tilde{\theta})\right)|X_C^{\theta}(t+\varepsilon,l^{\varepsilon};\tilde{\theta})=X_C^{\theta}(t+\varepsilon,\bar{l};\tilde{\theta})\right]|X_C^{\theta}(t,\bar{l};\tilde{\theta})=x\right]\\&=\mathbb{E}\left[\mathbb{E}\left[f\left(X_C^{\theta}(T,l^{\varepsilon};\tilde{\theta})\right)|X_C^{\theta}(t+\varepsilon,\bar{l};\tilde{\theta})\right]|X_C^{\theta}(t,\bar{l};\tilde{\theta})=x\right].\end{aligned} \quad (A.18)$$

Given $X_C^{\theta}(t+\varepsilon,l^{\varepsilon};\tilde{\theta})=X_C^{\theta}(t+\varepsilon,\hat{l};\tilde{\theta})=X_C^{\theta}(t+\varepsilon,\bar{l};\tilde{\theta})$, since the dynamics of $X_C^{\theta}(t+\varepsilon,l^{\varepsilon};\tilde{\theta})$ and $X_C^{\theta}(t+\varepsilon,\hat{l};\tilde{\theta})$ are the same for $s\in[t+\varepsilon,T]$, we have $X_C^{\theta}(T,l^{\varepsilon};\tilde{\theta})\overset{\mathrm{d}}{=}X_C^{\theta}(T,\hat{l};\tilde{\theta})$, i.e.,

$$\mathbb{E}\left[f\left(X_C^{\theta}(T,l^{\varepsilon};\tilde{\theta})\right)|X_C^{\theta}(t+\varepsilon,\bar{l};\tilde{\theta})=x\right]=\mathbb{E}\left[f\left(X_C^{\theta}(T,\hat{l};\tilde{\theta})\right)|X_C^{\theta}(t+\varepsilon,\bar{l};\tilde{\theta})=x\right].$$

Substituting it into (A.18) yields (A.17). ∎

Proof of Theorem 2.1. By (2.3) and (2.4), the customer's surplus evolves as follows:

$$\mathrm{d}X_C^{\theta}(t,l;\tilde{\theta})=-\lambda\int_0^{\infty}\left(l(t,y)+\frac{\eta(\tilde{\theta})}{2}l(t,y)^2\right)f(y;\tilde{\theta})\mathrm{d}y\mathrm{d}t-\int_0^{\infty}(y-l(t,y))N^{\theta}(\mathrm{d}t,\mathrm{d}y).$$

We aim to solve for V and g satisfying all the conditions of Theorem A.1. We use the following ansatz for the value functions:

$$V(t,x)=x+B(t)\quad\text{and}\quad g(t,x)=x+b(t), \quad (A.19)$$

with $B(T)=b(T)=0$. Substituting (A.19) into the extended HJB equation (A.4) yields

$$B'(t)-\lambda\sup_{l\in\mathbb{R}}\int_0^{\infty}\left\{\left[l(t,y)+\frac{\eta(\tilde{\theta})}{2}l(t,y)^2\right]f(y;\tilde{\theta})+\left[(y-l(t,y))+\frac{\gamma}{2}(y-l(t,y))^2\right]f(y;\theta)\right\}\mathrm{d}y=0. \quad (A.20)$$

It follows that the maximizer, that is the equilibrium strategy, is given by

$$\hat{l}^{\theta}(t,y;\tilde{\theta})=\frac{(1+\gamma y)f(y;\theta)-f(y;\tilde{\theta})}{\eta(\tilde{\theta})f(y;\tilde{\theta})+\gamma f(y;\theta)}.$$

By the boundary condition $B(T)=0$, one obtains

$$\begin{aligned}V_C^\theta(\tilde{\theta}) &= V(0,x_0) = x_0 + B(0)\\ &= x_0 - \lambda T\int_0^\infty \left[\hat{l}^\theta(t,y;\tilde{\theta}) + \frac{\eta(\tilde{\theta})}{2}\hat{l}^\theta(t,y;\tilde{\theta})^2\right] f(y;\tilde{\theta})\mathrm{d}y\\ &\quad - \lambda T\int_0^\infty \left[\left(y-\hat{l}^\theta(t,y;\tilde{\theta})\right) + \frac{\gamma}{2}\left(y-\hat{l}^\theta(t,y;\tilde{\theta})\right)^2\right] f(y;\theta)\mathrm{d}y.\end{aligned}$$

In particular, when $\tilde{\theta}=\theta$, we have

$$\hat{l}^\theta(t,y;\theta) = \frac{\gamma}{\eta(\theta)+\gamma}y$$

and

$$V_C^\theta(\theta) = x_0 - \lambda T\int_0^\infty \left(y + \frac{\eta(\theta)\gamma}{2(\eta(\theta)+\gamma)}y^2\right) f(y;\theta)\mathrm{d}y.$$

■

A.2 Proof of Lemma 2.1

Differentiating (2.9) yields

$$\begin{aligned}&\left.\frac{\partial\hat{l}^\theta(t,y;\tilde{\theta})}{\partial\tilde{\theta}}\right|_{\tilde{\theta}=\theta}\\ &= \left.\frac{-f'(y;\tilde{\theta})\left[\eta(\tilde{\theta})f(y;\tilde{\theta})+\gamma f(y;\theta)\right] - \left[(1+\gamma y)f(y;\theta)-f(y;\tilde{\theta})\right]\left[\eta'(\tilde{\theta})f(y;\tilde{\theta})+\eta(\tilde{\theta})f'(y;\tilde{\theta})\right]}{\left(\eta(\tilde{\theta})f(y;\tilde{\theta})+\gamma f(y;\theta)\right)^2}\right|_{\tilde{\theta}=\theta}\\ &= -\frac{\left[(1+\gamma y)\eta(\theta)+\gamma\right]f'(y;\theta)+\gamma y\eta'(\theta)f(y;\theta)}{(\eta(\theta)+\gamma)^2 f(y;\theta)},\end{aligned}$$

where $f'(y;\theta) := \frac{\partial f(y;\theta)}{\partial\theta}$. By this, together with (2.10), one can calculate that the first-order condition of (2.8)—namely,

$$\left.\frac{\partial V_C^\theta(\tilde{\theta})}{\partial\tilde{\theta}}\right|_{\tilde{\theta}=\theta} = 0$$

yields

$$\int_0^\infty \left[\frac{\eta'(\theta)}{2}\frac{\gamma^2}{(\eta(\theta)+\gamma)^2}y^2\right] f(y;\theta) + \left[\frac{\gamma}{\eta(\theta)+\gamma}y + \frac{\eta(\theta)\gamma^2}{2(\eta(\theta)+\gamma)^2}y^2\right] f'(y;\theta)\mathrm{d}y = 0.$$

Equivalently, we have

$$\eta'(\theta)\gamma\int_0^\infty y^2 f(y;\theta)\mathrm{d}y + \eta(\theta)\int_0^\infty \left(2yf'(y;\theta)+\gamma y^2 f'(y;\theta)\right)\mathrm{d}y + 2\gamma\int_0^\infty yf'(y;\theta)\mathrm{d}y = 0.$$

Recall that $\mu(\theta) := \mathbb{E}[Z^\theta]$ and $\sigma(\theta) := \mathbb{E}[(Z^\theta)^2]$, we obtain a first-order ordinary differential equation (ODE) of $\eta(\cdot)$; that is,

$$\eta'(\theta) + \frac{2\mu'(\theta)+\gamma\sigma'(\theta)}{\gamma\sigma(\theta)}\eta(\theta) + \frac{2\mu'(\theta)}{\sigma(\theta)} = 0, \quad \theta\in(\theta_L,\theta_H). \tag{A.21}$$

We define $a(\theta) := \frac{2\mu'(\theta)+\gamma\sigma'(\theta)}{\gamma\sigma(\theta)}$ and $r(\theta) := \frac{2\mu'(\theta)}{\sigma(\theta)}$. The general solution to (A.21) is given by

$$\eta(\theta) = -\int_{\theta_L}^{\theta} e^{-\int_x^\theta a(z)\mathrm{d}z} r(x)\mathrm{d}x + Ce^{-\int_{\theta_L}^\theta a(z)\mathrm{d}z}$$

for an arbitrary constant C. To have $\eta(\theta) \geq 0$ for all $\theta \in [\theta_L, \theta_H]$, it is easy to see that

$$C \geq \int_{\theta_L}^{\theta} e^{\int_{\theta_L}^x a(z)\mathrm{d}z} r(x)\mathrm{d}x, \quad \forall \theta \in [\theta_L, \theta_H],$$

or equivalently,

$$C \geq \int_{\theta_L}^{\theta_H} e^{\int_{\theta_L}^x a(z)\mathrm{d}z} r(x)\mathrm{d}x := \bar{r}.$$

A.3 Proof of Theorem 2.2

To prove Theorem 2.2, we need the following lemma.

Lemma A.2 *Define an auxiliary function*

$$H(C) = \int_{\theta_L}^{\theta_H} \left(\frac{1}{-\int_{\theta_L}^{\theta} e^{-\int_x^\theta a(z)\mathrm{d}z} r(x)\mathrm{d}x + Ce^{-\int_{\theta_L}^\theta a(z)\mathrm{d}z} + \gamma} - \frac{1}{2\gamma} \right)^2 \sigma(\theta)\mathrm{d}G(\theta), \quad C \geq \bar{r}. \quad \text{(A.22)}$$

Then, its minimum $C^* = \mathrm{argmin}_{C\geq\bar{r}} H(C)$ *satisfies*

$$\gamma \vee \bar{r} \leq C^* \leq \bar{r} + \gamma e^{\int_{\theta_L}^{\theta_H} a(z)\mathrm{d}z}.$$

Proof. Let

$$h(x) = \left(\frac{1}{x+\gamma} - \frac{1}{2\gamma} \right)^2, \quad x \geq 0.$$

It is straightforward to verify that h is decreasing for $x \in [0, \gamma)$ and increasing for $x \in (\gamma, \infty)$. By

$$H(C) = \int_{\theta_L}^{\theta_H} h\left(-\int_{\theta_L}^{\theta} e^{-\int_x^\theta a(z)\mathrm{d}z} r(x)\mathrm{d}x + Ce^{-\int_{\theta_L}^\theta a(z)\mathrm{d}z} \right) \sigma(\theta)\mathrm{d}G(\theta),$$

we deduce that H is increasing if

$$-\int_{\theta_L}^{\theta} e^{-\int_x^\theta a(z)\mathrm{d}z} r(x)\mathrm{d}x + Ce^{-\int_{\theta_L}^\theta a(z)\mathrm{d}z} \geq \gamma \text{ for } \forall \theta \in [\theta_L, \theta_H].$$

Equivalently,

$$C \geq \max_{\theta\in[\theta_L,\theta_H]} \left\{ \int_{\theta_L}^{\theta} e^{\int_{\theta_L}^x a(z)\mathrm{d}z} r(x)\mathrm{d}x + \gamma e^{\int_{\theta_L}^\theta a(z)\mathrm{d}z} \right\} = \bar{r} + \gamma e^{\int_{\theta_L}^{\theta_H} a(z)\mathrm{d}z}.$$

Thus,

$$C^* = \mathrm{argmin}_{C\geq\bar{r}} H(C) \leq \bar{r} + \gamma e^{\int_{\theta_L}^{\theta_H} a(z)\mathrm{d}z}.$$

To show the other side, $C^* \geq \gamma \vee \bar{r}$, it suffices to consider the non-trivial case when $\gamma > \bar{r}$. Since h is decreasing for $x \in [0, \gamma)$, we deduce that H is decreasing in C if

$$0 \leq -\int_{\theta_L}^{\theta} e^{-\int_x^\theta a(z)\mathrm{d}z} r(x)\mathrm{d}x + Ce^{-\int_{\theta_L}^\theta a(z)\mathrm{d}z} \leq \gamma, \quad \forall \theta \in [\theta_L, \theta_H],$$

or equivalently,

$$\int_{\theta_L}^{\theta} e^{\int_{\theta_L}^{x} a(z)\mathrm{d}z} r(x)\mathrm{d}x \le C \le \int_{\theta_L}^{\theta} e^{\int_{\theta_L}^{x} a(z)\mathrm{d}z} r(x)\mathrm{d}x + \gamma e^{\int_{\theta_L}^{\theta} a(z)\mathrm{d}z}, \quad \forall \theta \in [\theta_L, \theta_H].$$

Clearly, all $C \in [\bar{r}, \gamma]$ satisfy the above inequality. In other words, H is decreasing in $C \in [\bar{r}, \gamma]$, and thus the minimum $C^* = \mathrm{argmin}_{C \ge \bar{r}} H(C) \ge \gamma$. ■

Proof of Theorem 2.2. (1) By (2.3), (2.5), and (2.10), the insurer's surplus follows the dynamics

$$\begin{aligned}\mathrm{d}X_I^{\theta}(t, \hat{l}^{\theta}(\theta); \theta) &= \lambda \int_0^{\infty} \left(\frac{\gamma}{\eta(\theta)+\gamma} y + \frac{\eta(\theta)\gamma^2}{2(\eta(\theta)+\gamma)^2} y^2 \right) f(y;\theta)\mathrm{d}y\mathrm{d}t - \int_0^{\infty} \frac{\gamma}{\eta(\theta)+\gamma} y N^{\theta}(\mathrm{d}t, \mathrm{d}y) \\ &= \frac{\lambda\gamma}{\eta(\theta)+\gamma}\mu(\theta)\mathrm{d}t + \frac{\lambda\eta(\theta)\gamma^2}{2(\eta(\theta)+\gamma)^2}\sigma(\theta)\mathrm{d}t - \frac{\gamma}{\eta(\theta)+\gamma}\mathrm{d}U^{\theta}(t),\end{aligned}$$

with initial value zero. It follows that

$$\begin{aligned}&\int_{\theta_L}^{\theta_H} \mathbb{E}\left[X_I^{\theta}(T, \hat{l}^{\theta}(\theta); \theta) \right] \mathrm{d}G(\theta) \\ &= \int_{\theta_L}^{\theta_H} \frac{\lambda T \eta(\theta)\gamma^2}{2(\eta(\theta)+\gamma)^2}\sigma(\theta)\mathrm{d}G(\theta) \\ &= \lambda T \int_{\theta_L}^{\theta_H} \frac{\eta(\theta)\gamma^2}{2(\eta(\theta)+\gamma)^2}\sigma(\theta)\mathrm{d}G(\theta) \\ &= \frac{\lambda\gamma T}{8}\int_{\theta_L}^{\theta_H} \sigma(\theta)\mathrm{d}G(\theta) - \frac{\lambda\gamma^3 T}{2}\int_{\theta_L}^{\theta_H} \left(\frac{1}{\eta(\theta)+\gamma} - \frac{1}{2\gamma} \right)^2 \sigma(\theta)\mathrm{d}G(\theta). \end{aligned} \tag{A.23}$$

It follows from (2.12) that the insurer's optimal value function

$$\begin{aligned}&\sup_{\eta \in \mathbb{A}_{\eta}} \int_{\theta_L}^{\theta_H} \mathbb{E}\left[X_I^{\theta}(T, \hat{l}^{\theta}(\theta); \theta) \right] \mathrm{d}G(\theta) \\ &= \frac{\lambda\gamma T}{8}\int_{\theta_L}^{\theta_H} \sigma(\theta)\mathrm{d}G(\theta) - \frac{\lambda\gamma^3 T}{2} H(C^*), \end{aligned} \tag{A.24}$$

where $H(\cdot)$ is defined in (A.22) and $C^* = \mathrm{argmin}_{C \ge \bar{r}} H(C)$.

(2) This is an immediate result of (2.10) and (2.3).

(3) This is an immediate result of (A.24).

(4) If a type-θ customer selects the contract for type $\tilde{\theta}$, i.e., $(\hat{l}(\tilde{\theta}), \hat{p}(\tilde{\theta}))$, the customer's surplus evolves as follows:

$$\mathrm{d}X_C^{\theta}(t, \hat{l}(\tilde{\theta}); \tilde{\theta}) = -\lambda \int_0^{\infty} \left(\hat{l}(t, y; \tilde{\theta}) + \frac{\hat{\eta}(\tilde{\theta})}{2} \hat{l}(t, y; \tilde{\theta})^2 \right) f(y; \tilde{\theta})\mathrm{d}y\mathrm{d}t - \int_0^{\infty} (y - \hat{l}(t, y; \tilde{\theta})) N^{\theta}(\mathrm{d}t, \mathrm{d}y),$$

with $\hat{l}(\tilde{\theta}) = \frac{\gamma}{\hat{\eta}(\tilde{\theta})+\gamma} y$, $\hat{\eta}(\tilde{\theta}) = -\int_{\theta_L}^{\tilde{\theta}} e^{-\int_x^{\tilde{\theta}} a(z)\mathrm{d}z} r(x)\mathrm{d}x + C^* e^{-\int_{\theta_L}^{\tilde{\theta}} a(z)\mathrm{d}z}$, and initial value x_0. Correspondingly, his value function is denoted by

$$J^{\theta}(\tilde{\theta}) = \mathbb{E}\left[X_C^{\theta}(T, \hat{l}(\tilde{\theta}); \tilde{\theta}) \right] - \frac{\gamma}{2}\mathrm{Var}\left[X_C^{\theta}(T, \hat{l}(\tilde{\theta}); \tilde{\theta}) \right].$$

By Itô's formula, one can show that

$$J^{\theta}(\tilde{\theta}) = x_0 - \lambda T \int_0^{\infty} \left[\hat{l}(t,y;\tilde{\theta}) + \frac{\hat{\eta}(\tilde{\theta})}{2} \hat{l}(t,y;\tilde{\theta})^2 \right] f(y;\tilde{\theta}) \mathrm{d}y$$
$$- \lambda T \int_0^{\infty} \left[\left(y - \hat{l}(t,y;\tilde{\theta}) \right) + \frac{\gamma}{2} \left(y - \hat{l}(t,y;\tilde{\theta}) \right)^2 \right] f(y;\theta) \mathrm{d}y.$$

Substituting $\hat{l}(\tilde{\theta}) = \frac{\gamma}{\hat{\eta}(\tilde{\theta})+\gamma} y$ into the above equation yields

$$\begin{aligned}
J^{\theta}(\tilde{\theta}) &= x_0 - \lambda T \int_0^{\infty} \left[\frac{\gamma}{\hat{\eta}(\tilde{\theta})+\gamma} y + \frac{\hat{\eta}(\tilde{\theta})}{2} \frac{\gamma^2}{(\hat{\eta}(\tilde{\theta})+\gamma)^2} y^2 \right] f(y;\tilde{\theta}) \mathrm{d}y \\
&- \lambda T \int_0^{\infty} \left[\frac{\hat{\eta}(\tilde{\theta})}{\hat{\eta}(\tilde{\theta})+\gamma} y + \frac{\gamma}{2} \frac{\hat{\eta}(\tilde{\theta})^2}{(\hat{\eta}(\tilde{\theta})+\gamma)^2} y^2 \right] f(y;\theta) \mathrm{d}y \\
&= x_0 - \lambda T \left\{ \frac{\gamma \mu(\tilde{\theta})}{\hat{\eta}(\tilde{\theta})+\gamma} + \frac{\hat{\eta}(\tilde{\theta}) \gamma^2 \sigma(\tilde{\theta})}{2(\hat{\eta}(\tilde{\theta})+\gamma)^2} + \frac{\hat{\eta}(\tilde{\theta}) \mu(\theta)}{\hat{\eta}(\tilde{\theta})+\gamma} + \frac{\hat{\eta}(\tilde{\theta})^2 \gamma \sigma(\theta)}{2(\hat{\eta}(\tilde{\theta})+\gamma)^2} \right\} \\
&= x_0 - \lambda T \frac{2(\hat{\eta}(\tilde{\theta})+\gamma)\gamma\mu(\tilde{\theta}) + \hat{\eta}(\tilde{\theta})\gamma\sigma(\tilde{\theta}) + 2(\hat{\eta}(\tilde{\theta})+\gamma)\hat{\eta}(\tilde{\theta})\mu(\theta) + \hat{\eta}(\tilde{\theta})^2\gamma\sigma(\theta)}{2(\hat{\eta}(\tilde{\theta})+\gamma)^2}.
\end{aligned} \tag{A.25}$$

Differentiating with respect to $\tilde{\theta}$ yields

$$\begin{aligned}
&\frac{\partial J^{\theta}(\tilde{\theta})}{\partial \tilde{\theta}} \cdot 2(\hat{\eta}(\tilde{\theta})+\gamma)^3 \\
&= -2\gamma\hat{\eta}'(\tilde{\theta})(\hat{\eta}(\tilde{\theta})+\gamma)\mu(\tilde{\theta}) + \gamma^2\hat{\eta}'(\tilde{\theta})(\gamma - \hat{\eta}(\tilde{\theta}))\sigma(\tilde{\theta}) + 2\gamma\hat{\eta}'(\tilde{\theta})(\hat{\eta}(\tilde{\theta})+\gamma)\mu(\theta) \\
&+ 2\gamma^2\hat{\eta}(\tilde{\theta})\hat{\eta}'(\tilde{\theta})\sigma(\theta) + 2\gamma(\hat{\eta}(\tilde{\theta})+\gamma)^2\mu'(\tilde{\theta}) + \gamma^2\hat{\eta}(\tilde{\theta})(\hat{\eta}(\tilde{\theta})+\gamma)\sigma'(\tilde{\theta}) \\
&= 2\gamma\hat{\eta}'(\tilde{\theta})(\hat{\eta}(\tilde{\theta})+\gamma) \left(\mu(\theta) - \mu(\tilde{\theta}) \right) + 2\gamma^2\hat{\eta}(\tilde{\theta})\hat{\eta}'(\tilde{\theta}) \left(\sigma(\theta) - \sigma(\tilde{\theta}) \right) \\
&+ \gamma(\gamma + \hat{\eta}(\tilde{\theta}) \left\{ \gamma\hat{\eta}'(\tilde{\theta}))\sigma(\tilde{\theta}) + 2(\hat{\eta}(\tilde{\theta})+\gamma)\mu'(\tilde{\theta}) + \gamma\hat{\eta}(\tilde{\theta})\sigma'(\tilde{\theta}) \right\}.
\end{aligned}$$

It follows from (A.21) that

$$\gamma\hat{\eta}'(\tilde{\theta}))\sigma(\tilde{\theta}) + 2(\hat{\eta}(\tilde{\theta})+\gamma)\mu'(\tilde{\theta}) + \gamma\hat{\eta}(\tilde{\theta})\sigma'(\tilde{\theta}) = 0.$$

Thus,

$$\begin{aligned}
&\frac{\partial J^{\theta}(\hat{l}(\tilde{\theta});\tilde{\theta})}{\partial \tilde{\theta}} \cdot 2(\hat{\eta}(\tilde{\theta})+\gamma)^3 \\
&= 2\gamma\hat{\eta}'(\tilde{\theta})(\hat{\eta}(\tilde{\theta})+\gamma) \left(\mu(\theta) - \mu(\tilde{\theta}) \right) + 2\gamma^2\hat{\eta}(\tilde{\theta})\hat{\eta}'(\tilde{\theta}) \left(\sigma(\theta) - \sigma(\tilde{\theta}) \right).
\end{aligned}$$

By Assumption 2.1, both $\mu(\theta)$ and $\sigma(\theta)$ increase with θ. Then, it follows that

$$\frac{\partial J^{\theta}(\hat{l}(\tilde{\theta});\tilde{\theta})}{\partial \tilde{\theta}} > (\text{reps. } <)0, \quad \text{for } \tilde{\theta} < (\text{resp. } >)\theta.$$

As such,

$$\theta \in \mathrm{argmax}_{\tilde{\theta} \in [\theta_L, \theta_H]} J^{\theta}(\tilde{\theta}).$$

This completes the proof. ■

A.4 Proof of Proposition 2.1

(1) Differentiating $\hat{\eta}(\theta)$ in (2.13) yields

$$\begin{aligned}\hat{\eta}'(\theta) &= -r(\theta) + a(\theta)\int_{\theta_L}^{\theta} e^{-\int_x^\theta a(z)\mathrm{d}z} r(x)\mathrm{d}x - a(\theta)C^* e^{-\int_{\theta_L}^{\theta} a(z)\mathrm{d}z} \\ &= -r(\theta) - a(\theta)\hat{\eta}(\theta). \end{aligned} \tag{A.26}$$

By Assumption 2.1,

$$r(\theta) = \frac{2\mu'(\theta)}{\sigma(\theta)} > 0 \text{ and } a(\theta) = \frac{2\mu'(\theta) + \gamma\sigma'(\theta)}{\gamma\sigma(\theta)} > 0. \tag{A.27}$$

Further, because $\hat{\eta}(\theta) \geq 0$, we have that $\hat{\eta}'(\theta) < 0$.

(2) By (2.15) and the decreasing property of $\hat{\eta}(\theta)$, shown in part (1), it is obvious that $\hat{l}(t,y;\theta) = \frac{\gamma}{\hat{\eta}(\theta)+\gamma} y$ is increasing in θ.

(3) For ease of notation, we denote $\hat{p}'(\theta) := \frac{\partial \hat{p}(t;\theta)}{\partial\theta}$ and $\hat{l}'(t,y;\theta) := \frac{\partial \hat{l}(t,y;\theta)}{\partial\theta}$ in this proof. Differentiating $\hat{p}(\theta)$ in (2.15) with respect to θ yields

$$\begin{aligned}\hat{p}'(\theta) &= \lambda\int_0^\infty \left[\hat{l}'(t,y;\theta)f(y;\theta) + \hat{l}(t,y;\theta)f'(y;\theta)\right]\mathrm{d}y + \frac{\hat{\eta}'(\theta)}{2}\lambda\int_0^\infty \hat{l}(t,y;\theta)^2 f(y;\theta)\mathrm{d}y \\ &\quad + \frac{\hat{\eta}(\theta)}{2}\lambda\int_0^\infty \left[2\hat{l}(t,y;\theta)\hat{l}'(t,y;\theta)f(y;\theta) + \hat{l}(t,y;\theta)^2 f'(y;\theta)\right]\mathrm{d}y \\ &= -\frac{\lambda\gamma\hat{\eta}'(\theta)}{(\hat{\eta}(\theta)+\gamma)^2}\mu(\theta) - \frac{\lambda\gamma^2\hat{\eta}'(\theta)(\hat{\eta}(\theta)-\gamma)}{2(\hat{\eta}(\theta)+\gamma)^3}\sigma(\theta) + \frac{\lambda\gamma}{\hat{\eta}(\theta)+\gamma}\mu'(\theta) + \frac{\lambda\gamma^2\hat{\eta}(\theta)}{2(\hat{\eta}(\theta)+\gamma)^2}\sigma'(\theta), \end{aligned} \tag{A.28}$$

where $\hat{l}(t,y;\theta) = \frac{\gamma}{\hat{\eta}(\theta)+\gamma}y$ in the last equation. Further, by (A.26) and (A.27), one obtains

$$\hat{\eta}'(\theta) = -r(\theta) - a(\theta)\hat{\eta}(\theta) = -\frac{2\mu'(\theta)}{\sigma(\theta)} - \frac{2\mu'(\theta)+\gamma\sigma'(\theta)}{\gamma\sigma(\theta)}\hat{\eta}(\theta),$$

which implies

$$-\frac{\lambda\gamma^2\hat{\eta}'(\theta)(\hat{\eta}(\theta)-\gamma)}{2(\hat{\eta}(\theta)+\gamma)^3}\sigma(\theta) = \frac{\lambda\gamma(\hat{\eta}(\theta)-\gamma)}{(\hat{\eta}(\theta)+\gamma)^2}\mu'(\theta) + \frac{\lambda\gamma^2\hat{\eta}(\theta)(\hat{\eta}(\theta)-\gamma)}{2(\hat{\eta}(\theta)+\gamma)^3}\sigma'(\theta).$$

Substituting the last equation into (A.28) gives

$$\hat{p}'(\theta) = -\frac{\lambda\gamma\hat{\eta}'(\theta)}{(\hat{\eta}(\theta)+\gamma)^2}\mu(\theta) + \frac{2\lambda\gamma\hat{\eta}(\theta)}{(\hat{\eta}(\theta)+\gamma)^2}\mu'(\theta) + \frac{\lambda\gamma^2\hat{\eta}(\theta)^2}{(\hat{\eta}(\theta)+\gamma)^3}\sigma'(\theta).$$

By Assumption 2.1 and $\hat{\eta}'(\theta) < 0$, we have that $\hat{p}'(\theta) > 0$, and thus $\hat{p}(\theta)$ increases with θ.

A.5 Proof of Proposition 2.2

(1) By (2.15), we have $\hat{l}(t,y;\theta)/y = \frac{\gamma}{\hat{\eta}(\theta)+\gamma}$. It is known from Theorem 2.2(1) that $\hat{\eta}(\theta) \leq C^* \leq \bar{r} + \gamma e^{\int_{\theta_L}^{\theta_H} a(z)\mathrm{d}z}$. Thus, $\hat{l}(t,y;\theta)/y > 0$ for all θ.

(2) Without purchasing any insurance, the dynamics of the customer's surplus follows

$$X_C^\theta(t) = x_0 - \int_0^t\int_0^\infty y N^\theta(\mathrm{d}s,\mathrm{d}y).$$

Correspondingly, the MV value of a type-θ customer's surplus is given by

$$\begin{aligned} J^\theta &= \mathbb{E}\left[X_C^\theta(T)\right] - \frac{\gamma}{2}\mathrm{Var}\left[X_C^\theta(T)\right] \\ &= x_0 - \lambda T\int_0^\infty (y + \frac{\gamma}{2}y^2) f(y;\theta)\mathrm{d}y. \end{aligned}$$

On the other hand, by (2.11), the optimal value function of a type-θ customer is given by

$$V_C^\theta(\theta) = x_0 - \lambda T\int_0^\infty \left(y + \frac{\hat{\eta}(\theta)\gamma}{2(\hat{\eta}(\theta)+\gamma)}y^2\right) f(y;\theta)\mathrm{d}y.$$

It follows that the gain in the type-θ customer's value is given by

$$\begin{aligned} V_C^\theta(\theta) - J^\theta &= \frac{\lambda T\gamma}{2}\left(\frac{\gamma}{\hat{\eta}(\theta)+\gamma}\right)\int_0^\infty y^2 f(y;\theta)\mathrm{d}y \\ &= \frac{\lambda T\gamma^2\sigma(\theta)}{2(\hat{\eta}(\theta)+\gamma)}. \end{aligned}$$

(3) Similar to the second equation in (A.23), we have

$$\int_{\theta_L}^{\theta_H} \mathbb{E}\left[X_I^\theta(T,\hat{l}(\theta);\theta)\right]\mathrm{d}G(\theta) = \lambda T\int_{\theta_L}^{\theta_H} \frac{\hat{\eta}(\theta)\gamma^2}{2(\hat{\eta}(\theta)+\gamma)^2}\sigma(\theta)\mathrm{d}G(\theta).$$

By Proposition 2.1(1), $\hat{\eta}(\theta) > \hat{\eta}(\theta_H) \geq 0$ for any $\theta \in [\theta_L, \theta_H)$. Therefore, $V_I(t,x) > x$.

A.6 Proof of Proposition 2.3

By (A.25), we have

$$\begin{aligned} &\mathbb{E}\left[X_C^{\theta_0}(T,\hat{l}(\theta);\theta)\right] - \frac{\gamma}{2}\mathrm{Var}\left[X_C^{\theta_0}(T,\hat{l}(\theta);\theta)\right] \\ &= x_0 - \lambda T\frac{2(\hat{\eta}(\theta)+\gamma)\gamma\mu(\theta) + \hat{\eta}(\theta)\gamma^2\sigma(\theta) + 2(\hat{\eta}(\theta)+\gamma)\hat{\eta}(\theta)\mu(\theta_0) + \hat{\eta}(\theta)^2\gamma\sigma(\theta_0)}{2(\hat{\eta}(\theta)+\gamma)^2}. \end{aligned}$$

By Assumption 2.1, it can be seen that $\mathbb{E}\left[X_C^{\theta_0}(T,\hat{l}(\theta);\theta)\right] - \frac{\gamma}{2}\mathrm{Var}\left[X_C^{\theta_0}(T,\hat{l}(\theta);\theta)\right]$ decreases with θ_0.

A.7 Proof of Theorem 2.3

By (2.18) and (2.19), the premium rate is given by

$$\begin{aligned} p(\mathring{l},\eta;\theta) &= \lambda\int_0^\infty \frac{\gamma}{\eta+\gamma} y f(y;\theta)\mathrm{d}y + \frac{\eta}{2}\lambda\int_0^\infty \frac{\gamma^2}{(\eta+\gamma)^2}y^2 f(y;\theta)\mathrm{d}y \\ &= \lambda\left(\frac{\gamma}{\eta+\gamma}\mu(\theta) + \frac{\eta\gamma^2}{2(\eta+\gamma)^2}\sigma(\theta)\right). \end{aligned} \tag{A.29}$$

The dynamics of the insurer's surplus then follows

$$\begin{aligned} \mathrm{d}X_I^\theta(t,\mathring{l},\eta;\theta) &= p(\mathring{l},\eta;\theta)\mathrm{d}t - \int_0^\infty \mathring{l}(t,y)N^\theta(\mathrm{d}t,\mathrm{d}y) \\ &= \lambda\left(\frac{\gamma}{\eta+\gamma}\mu(\theta) + \frac{\eta\gamma^2}{2(\eta+\gamma)^2}\sigma(\theta)\right)\mathrm{d}t - \frac{\gamma}{\eta+\gamma}\int_0^\infty y N^\theta(\mathrm{d}t,\mathrm{d}y), \end{aligned}$$

which implies that

$$\mathbb{E}\left[X_I^\theta(T,\mathring{l},\eta;\theta)\right]=\frac{\lambda T\eta\gamma^2\sigma(\theta)}{2(\eta+\gamma)^2}.$$

By Assumption 2.1, we obtain

$$\begin{aligned}\sup_{\eta\geq 0}\inf_{\theta\in[\theta_L,\theta_H]}\mathbb{E}\left[X_I^\theta(T,\mathring{l},\eta;\theta)\right]&=\frac{\lambda T\gamma^2}{2}\sup_{\eta\geq 0}\left\{\frac{\eta}{(\eta+\gamma)^2}\right\}\sigma(\theta_L),\\&=\frac{1}{8}\lambda T\gamma\sigma(\theta_L),\end{aligned}\tag{A.30}$$

where the infimum of θ is reached at $\hat{\theta}=\theta_L$ and the supremum of η is reached at $\eta^{\mathrm{wc}}=\gamma$. It follows from (2.19) and (A.29) that the worst-case contract to Problem 2.3 is given by

$$\begin{cases} l^{\mathrm{wc}}(t,y)/y=\frac{1}{2},\\ p^{\mathrm{wc}}(t)=\frac{\lambda}{2}\mu(\theta_L)+\frac{\lambda\gamma}{8}\sigma(\theta_L).\end{cases}$$

Next, we investigate the customer's participation. By (2.4),

$$\begin{aligned}\mathrm{d}X_C^\theta(t,l^{\mathrm{wc}},\eta^{\mathrm{wc}})&=-p^{\mathrm{wc}}\mathrm{d}t-\int_0^\infty(y-l^{\mathrm{wc}}(t,y))N^\theta(\mathrm{d}t,\mathrm{d}y)\\&=-\lambda\left(\frac{1}{2}\mu(\theta_L)+\frac{\gamma}{8}\sigma(\theta_L)\right)\mathrm{d}t-\frac{1}{2}\int_0^\infty yN^\theta(\mathrm{d}t,\mathrm{d}y).\end{aligned}$$

The corresponding type-θ customer's value function is equal to

$$\begin{aligned}J^\theta(l^{\mathrm{wc}},\eta^{\mathrm{wc}})&:=\mathbb{E}\left[X_C^\theta(T,l^{\mathrm{wc}},\eta^{\mathrm{wc}};\theta_L)\right]-\frac{\gamma}{2}\mathrm{Var}\left[X_C^\theta(T,l^{\mathrm{wc}},\eta^{\mathrm{wc}};\theta_L)\right]\\&=x_0-\lambda T\left(\frac{1}{2}\mu(\theta_L)+\frac{\gamma}{8}\sigma(\theta_L)+\frac{1}{2}\mu(\theta)+\frac{\gamma}{8}\sigma(\theta)\right).\end{aligned}$$

Without participating in the worst-case contract, the value function of a type-θ customer is

$$\begin{aligned}J^\theta&:=\mathbb{E}\left[X_C^\theta(T,0;\cdot)\right]-\frac{\gamma}{2}\mathrm{Var}_{t,x}\left[X_C^\theta(T,0;\cdot)\right]\\&=x_0-\lambda T\mu(\theta)-\frac{\gamma}{2}\lambda T\sigma(\theta).\end{aligned}$$

It is clear that customers of any risk type θ are better off participating as $J^\theta(l^{\mathrm{ws}}(\eta^{\mathrm{ws}}))\geq J^\theta$.

Lastly, we examine the insurer's profit. By (2.5),

$$\begin{aligned}\mathrm{d}X_I^\theta(t,l^{\mathrm{ws}},\eta^{\mathrm{ws}})&=p^{\mathrm{ws}}\mathrm{d}t-\int_0^\infty l^{\mathrm{ws}}(s,y)N^\theta(\mathrm{d}t,\mathrm{d}y)\\&=\left(\frac{\lambda}{2}\mu(\theta_L)+\frac{\lambda\gamma}{8}\sigma(\theta_L)\right)\mathrm{d}t-\frac{1}{2}\mathrm{d}U^\theta(t).\end{aligned}$$

It follows that the insurer's expected profit is given by

$$\int_{\theta_L}^{\theta_H}\mathbb{E}\left[X_I^\theta(T,l^{\mathrm{ws}},\eta^{\mathrm{ws}})\right]\mathrm{d}G(\theta)=\frac{\lambda T}{2}\left(\mu(\theta_L)+\frac{\gamma}{4}\sigma(\theta_L)-\int_{\theta_L}^{\theta_H}\mu(\theta)\mathrm{d}G(\theta)\right).$$

If $\mu(\theta_L)+\frac{\gamma}{4}\sigma(\theta_L)<\int_{\theta_L}^{\theta_H}\mu(\theta)\mathrm{d}G(\theta)$, we have $\int_{\theta_L}^{\theta_H}\mathbb{E}\left[X_I^\theta(T,l^{\mathrm{ws}},\eta^{\mathrm{ws}})\right]\mathrm{d}G(\theta)<0$, meaning that the insurer is worse off by participating in the contract.

A.8 Proof of Theorem 3.1

Let

$$J^\gamma(t,x,l;\tilde{\gamma}) = \mathbb{E}\left[X_C(T,l;\tilde{\gamma})\middle|\, X_C(t,l;\tilde{\gamma}) = x\right] - \frac{\gamma}{2}\mathrm{Var}\left[X_C(T,l;\tilde{\gamma})\middle|\, X_C(t,l;\tilde{\gamma}) = x\right].$$

For any $l \in \mathbb{A}_l$ and $\varphi(t,x) \in C^{1,2}([0,T]\times\mathbb{R})$, i.e., first-order continuously differentiable in t and second-order continuously differentiable in x, we define the integro-differential operator

$$\begin{aligned}\mathcal{L}_l\varphi(t,x) = \frac{\partial\varphi(t,x)}{\partial t} - \frac{\partial\varphi(t,x)}{\partial x}\lambda\int_0^\infty \left(l(t,y) + \frac{\eta(\tilde{\gamma})}{2}l(t,y)^2\right) f(y)\mathrm{d}y \\ + \lambda\int_0^\infty \left[\varphi(t,x-y+l(t,y)) - \varphi(t,x)\right] f(y)\mathrm{d}y.\end{aligned}$$

Via a formal verification argument as in Theorem A.1, one can show that value functions

$$V(t,x) := J^\gamma(t,x,\hat{l};\tilde{\gamma}) \text{ and } g(t,x) := \mathbb{E}_{t,x}\left[X_C(T,\hat{l};\tilde{\gamma})\right]$$

satisfy the following extended HJB equations: for any $(t,x)\in[0,T]\times\mathbb{R}$,

$$\sup_{l\in\mathbb{R}}\left\{\mathcal{L}_l V(t,x) - \frac{1}{2}\gamma\mathcal{L}_l g^2(t,x) + \gamma g(t,x)\mathcal{L}_l g(t,x)\right\} = 0$$

and

$$\mathcal{L}_{\hat{l}} g(t,x) = 0,$$

with boundary conditions $V(T,x) = g(T,x) = x$. We can consider the same ansatz for the value functions—namely,

$$V(t,x) = x + B(t) \quad \text{and} \quad g(t,x) = x + b(t),$$

with $B(T) = b(T) = 0$. Substituting the above two forms into the extended HJB equation yields

$$\begin{aligned}0 &= B'(t) - \lambda\sup_{l\in\mathbb{R}}\int_0^\infty \left(l(t,y) + \frac{\eta(\tilde{\gamma})}{2}l(t,y)^2 + (y - l(t,y)) + \frac{\gamma}{2}\left(y - l(t,y)\right)^2\right) f(y)\mathrm{d}y \\ &= B'(t) - \lambda\sup_{l\in\mathbb{R}}\int_0^\infty \left(y + \frac{\eta(\tilde{\gamma})+\gamma}{2}l(t,y)^2 - \gamma y l(t,y) + \frac{\gamma}{2}y^2\right) f(y)\mathrm{d}y.\end{aligned}$$

It follows that the maximizer, i.e. the equilibrium strategy, is given by

$$\hat{l}^\gamma(t,y;\tilde{\gamma}) = \frac{\gamma}{\eta(\tilde{\gamma})+\gamma}y.$$

By the boundary condition $B(T) = 0$, one obtains

$$\begin{aligned}V_C^\gamma(\tilde{\gamma}) &= V(0,x_0) = x_0 + B(0) \\ &= x_0 - \lambda T\int_0^\infty \left(y + \frac{\eta(\tilde{\gamma})+\gamma}{2}\hat{l}^\gamma(t,y;\tilde{\gamma})^2 - \gamma y\hat{l}^\gamma(t,y;\tilde{\gamma}) + \frac{\gamma}{2}y^2\right) f(y)\mathrm{d}y \\ &= x_0 - \lambda T\int_0^\infty \left(y + \frac{\gamma\eta(\tilde{\gamma})}{2(\eta(\tilde{\gamma})+\gamma)}y^2\right) f(y)\mathrm{d}y \\ &= x_0 - \lambda T\mu - \frac{\lambda T\gamma\eta(\tilde{\gamma})\sigma}{2(\eta(\tilde{\gamma})+\gamma)}.\end{aligned}$$

A.9 Proof of Theorem 3.2

(1) By (3.1), (3.2), and (3.5), the insurer's surplus follows the dynamics

$$\begin{aligned}\mathrm{d}X_I(t,\hat{l}^\gamma(\gamma);\gamma) &= \lambda\int_0^\infty\left(\frac{\gamma}{\eta+\gamma}y+\frac{\eta\gamma^2}{2(\eta+\gamma)^2}y^2\right)f(y)\mathrm{d}y\mathrm{d}t-\int_0^\infty\frac{\gamma}{\eta+\gamma}yN^\theta(\mathrm{d}t,\mathrm{d}y)\\ &= \frac{\lambda\gamma\mu}{\eta+\gamma}\mathrm{d}t+\frac{\lambda\eta\gamma^2\sigma}{2(\eta+\gamma)^2}\mathrm{d}t-\frac{\gamma}{\eta+\gamma}\mathrm{d}U^\theta(t).\end{aligned}$$

It follows that

$$\begin{aligned}\int_{\gamma_L}^{\gamma_H}\mathbb{E}\left[X_I(T,\hat{l}^\gamma(\gamma);\gamma)\right]\mathrm{d}G(\gamma) &= \int_{\gamma_L}^{\gamma_H}\frac{\lambda T\eta\gamma^2\sigma}{2(\eta+\gamma)^2}\mathrm{d}G(\gamma)\\ &= \frac{\lambda T\sigma}{2}\int_{\gamma_L}^{\gamma_H}\frac{\eta\gamma^2}{(\eta+\gamma)^2}\mathrm{d}G(\gamma).\end{aligned}\tag{A.31}$$

By (3.7), it can be seen that

$$\begin{aligned}\hat{\eta} &= \mathrm{argmax}_{\eta(\cdot)\geq 0}\int_{\gamma_L}^{\gamma_H}\mathbb{E}\left[X_I(T,\hat{l}^\gamma(\gamma);\gamma)\right]\mathrm{d}G(\gamma)\\ &= \mathrm{argmax}_{\eta\geq 0}\int_{\gamma_L}^{\gamma_H}\frac{\eta\gamma^2}{(\eta+\gamma)^2}\mathrm{d}G(\gamma).\end{aligned}$$

Let $H(\eta):=\int_{\gamma_L}^{\gamma_H}\frac{\eta\gamma^2}{(\eta+\gamma)^2}\mathrm{d}G(\gamma)$, we have

$$H'(\eta):=\int_{\gamma_L}^{\gamma_H}\frac{(\gamma-\eta)}{(\eta+\gamma)^3}\gamma^2\mathrm{d}G(\gamma).$$

Clearly, $H'(\eta)>0$ for any $\eta<\gamma_L$, and $H'(\eta)<0$ for any $\eta>\gamma_H$. Therefore, $\hat{\eta}\in[\gamma_L,\gamma_H]$.

(2) This is an immediate result of (3.5) and (3.1).

(3) This is an immediate result of (A.31).

(4) If a customer with risk aversion γ selects the contract for type $\tilde{\gamma}$, i.e., $(\hat{l}(\tilde{\gamma}),\hat{p}(\tilde{\gamma}))$, the customer's surplus follows the dynamics

$$\begin{aligned}\mathrm{d}X_C(t,\hat{l}(\tilde{\gamma});\tilde{\gamma}) &= -\lambda\int_0^\infty\left(\frac{\tilde{\gamma}}{\hat{\eta}+\tilde{\gamma}}y+\frac{\hat{\eta}\tilde{\gamma}^2}{2(\hat{\eta}+\tilde{\gamma})^2}y^2\right)f(y)\mathrm{d}y\mathrm{d}t-\int_0^\infty\frac{\hat{\eta}}{\hat{\eta}+\tilde{\gamma}}yN^\theta(\mathrm{d}t,\mathrm{d}y)\\ &= -\lambda\left(\frac{\tilde{\gamma}}{\hat{\eta}+\tilde{\gamma}}\mathbb{E}[Z]+\frac{\hat{\eta}\tilde{\gamma}^2}{2(\hat{\eta}+\tilde{\gamma})^2}\mathbb{E}[Z^2]\right)\mathrm{d}t-\frac{\hat{\eta}}{\hat{\eta}+\tilde{\gamma}}\mathrm{d}U^\theta(t).\end{aligned}$$

Correspondingly, the customer's value function is equal to

$$\begin{aligned}J^\gamma(\tilde{\gamma}) &:= \mathbb{E}\left[X_C(T,\hat{l}(\tilde{\gamma});\tilde{\gamma})\right]-\frac{\gamma}{2}\mathrm{Var}\left[X_C(T,\hat{l}(\tilde{\gamma});\tilde{\gamma})\right]\\ &= x_0-\lambda T\mu-\lambda T\frac{\hat{\eta}\tilde{\gamma}^2\sigma}{2(\hat{\eta}+\tilde{\gamma})^2}-\frac{\lambda T\gamma}{2}\frac{\hat{\eta}^2\mu}{(\hat{\eta}+\tilde{\gamma})^2}\\ &= x_0-\lambda T\mu-\lambda T\sigma\frac{\hat{\eta}\tilde{\gamma}^2+\hat{\eta}^2\gamma}{2(\hat{\eta}+\tilde{\gamma})^2}.\end{aligned}$$

Differentiating with respect to $\tilde{\gamma}$ yields

$$\begin{aligned}\frac{\partial J^\gamma(\tilde{\gamma})}{\partial\tilde{\gamma}} &= -\frac{\lambda T\sigma}{2}\frac{2\hat{\eta}\tilde{\gamma}(\hat{\eta}+\tilde{\gamma})^2-2(\hat{\eta}\tilde{\gamma}^2+\hat{\eta}^2\gamma)(\hat{\eta}+\tilde{\gamma})}{(\hat{\eta}+\tilde{\gamma})^4}\\ &= \lambda T\sigma\frac{\hat{\eta}^2(\gamma-\tilde{\gamma})}{(\hat{\eta}+\tilde{\gamma})^3}.\end{aligned}$$

It is obvious to see that $\frac{\partial J^\gamma(\tilde{\gamma})}{\partial \tilde{\gamma}} > (\text{reps.}<)0, \quad \text{for } \tilde{\gamma} < (\text{resp.}>)\gamma$, and thus

$$\gamma = \text{argmax}_{\tilde{\gamma}\in[\gamma_L,\gamma_H]} J^\gamma(\tilde{\gamma}).$$

A.10 Proof of Theorem 3.3

(1) For a given $\eta \geq 0$, by (3.1) and (3.11), the premium rate is given by

$$\begin{aligned} p(\hat{l}^\gamma, \eta) &= \lambda \int_0^\infty \left(\hat{l}^\gamma(t,y) + \frac{\eta}{2}\hat{l}^\gamma(t,y)^2\right) f(y)\mathrm{d}y \\ &= \frac{\lambda\gamma\mu}{\eta+\gamma} + \frac{\lambda\eta\gamma^2\sigma}{2(\eta+\gamma)^2}. \end{aligned} \tag{A.32}$$

It follows that the insurer's profit evolves according to

$$\begin{aligned} \mathrm{d}X_I(t,\hat{l}^\gamma,\eta) &= p(t,\hat{l}^\gamma,\eta)\mathrm{d}t - \int_0^\infty \hat{l}^\gamma(t,y)N^\theta(\mathrm{d}t,\mathrm{d}y) \\ &= \left(\frac{\lambda\gamma\mu}{\eta+\gamma} + \frac{\lambda\eta\gamma^2\sigma}{2(\eta+\gamma)^2}\right)\mathrm{d}t - \frac{\gamma}{\eta+\gamma}\mathrm{d}U^\theta(t). \end{aligned}$$

Then,

$$\mathbb{E}\left[X_I(T,\hat{l}^\gamma,\eta)\right] = \frac{\lambda T\eta\gamma^2\sigma}{2(\eta+\gamma)^2}.$$

It follows that

$$\begin{aligned} \sup_{\eta\geq 0}\inf_{\gamma\in[\gamma_L,\gamma_H]} \mathbb{E}\left[X_I(T,\hat{l}^\gamma,\eta)\right] &= \frac{\lambda T\gamma_L^2\sigma}{2}\sup_{\eta\geq 0}\frac{\eta}{(\eta+\gamma_L)^2} \\ &= \frac{\lambda T\gamma_L\sigma}{8}, \end{aligned} \tag{A.33}$$

where the infimum and supremum are reached, respectively, at

$$\gamma^{\text{wc}} = \gamma_L \text{ and } \eta^{\text{wc}} = \gamma_L.$$

By (3.11) and (A.32), the worst-case contract is then given by

$$\begin{cases} l^{\text{wc}}(t,y) = \frac{\hat{\gamma}}{\hat{\eta}+\hat{\gamma}}y = \frac{1}{2}y, \\ p^{\text{wc}}(t) = \frac{\lambda\hat{\gamma}\mu}{\hat{\eta}+\hat{\gamma}} + \frac{\lambda\hat{\eta}\hat{\gamma}^2\sigma}{2(\hat{\eta}+\hat{\gamma})^2} = \frac{\lambda\mu}{2} + \frac{\lambda\sigma\gamma_L}{8}. \end{cases} \tag{A.34}$$

(2) By participating in the worst-case contract (A.34), the customer's surplus follows

$$\begin{aligned} \mathrm{d}X_C(t,l^{\text{wc}},\eta^{\text{wc}}) &= -p^{\text{wc}}\mathrm{d}t - \int_0^\infty l^{\text{wc}}(t,y)N^\theta(\mathrm{d}t,\mathrm{d}y) \\ &= -\left(\frac{\lambda\mu}{2} + \frac{\lambda\sigma\gamma_L}{8}\right)\mathrm{d}t - \frac{1}{2}\mathrm{d}U^\theta(t). \end{aligned}$$

The value function of the corresponding customer with risk aversion γ is

$$\begin{aligned} J^\gamma(l^{\text{wc}},\eta^{\text{wc}}) &:= \mathbb{E}\left[X_C(T,l^{\text{wc}},\eta^{\text{wc}})\right] - \frac{\gamma}{2}\text{Var}\left[X_C(T,l^{\text{wc}},\eta^{\text{wc}})\right] \\ &= x_0 - \lambda T\left(\frac{\mu}{2} + \frac{\gamma_L\sigma}{8}\right) - \frac{\lambda T\mu}{2} - \frac{\gamma}{2}\frac{1}{4}\lambda T\sigma \\ &= x - \lambda T\mu - \frac{\lambda T(\gamma_L+\gamma)\sigma}{8}. \end{aligned}$$

Without participating in the worst-case contract, the value function of the same γ-type customer is

$$J^{\gamma} = x - \lambda T\mu - \frac{\gamma}{2}\lambda T\sigma.$$

Then,

$$J^{\gamma}(l^{\mathrm{wc}}, \eta^{\mathrm{wc}}) - J^{\gamma} = \frac{\lambda T(3\gamma - \gamma_L)\sigma}{8} \geq 0,$$

which implies that all customers are willing to participate in the worst-case contract.

(3) Under the worst-case contract, the insurer's surplus follows

$$\begin{aligned}\mathrm{d}X_I(t, l^{\mathrm{wc}}, \eta^{\mathrm{wc}}) &= p^{\mathrm{wc}}(t)\mathrm{d}t - \int_0^{\infty} l^{\mathrm{wc}}(s, y)N(\mathrm{d}t, \mathrm{d}y) \\ &= \left(\frac{\lambda\mu}{2} + \frac{\lambda\sigma\gamma_L}{8}\right)\mathrm{d}t - \frac{1}{2}\mathrm{d}U(t).\end{aligned}$$

Then, the insurer's expected profit from offering this worst-case contract is given by

$$V_I^{\mathrm{wc}} = \int_{\gamma_L}^{\gamma_H} \mathbb{E}\left[X_I(T, l^{\mathrm{wc}}, \eta^{\mathrm{wc}})\right] \mathrm{d}G(\gamma) = \frac{\lambda T\sigma\gamma_L}{8}.$$

Subtracting V_I^{wc} from V_I, given in (3.10), yields

$$\begin{aligned}V_I - V_I^{\mathrm{wc}} &= \frac{\lambda T\sigma}{2}\left(\int_{\gamma_L}^{\gamma_H} \frac{\hat{\eta}\gamma^2}{(\hat{\eta} + \gamma)^2}\mathrm{d}G(\gamma) - \frac{\gamma_L}{4}\right) \\ &\geq \frac{\lambda T\sigma}{2}\left(\int_{\gamma_L}^{\gamma_H} \frac{\gamma_L\gamma^2}{(\gamma_L + \gamma)^2}\mathrm{d}G(\gamma) - \frac{\gamma_L}{4}\right) \\ &> \frac{\lambda T\sigma}{2}\left(\int_{\gamma_L}^{\gamma_H} \frac{\gamma_L\gamma_L^2}{(\gamma_L + \gamma_L)^2}\mathrm{d}G(\gamma) - \frac{\gamma_L}{4}\right) \\ &= 0,\end{aligned}$$

where the second step is due to $\hat{\eta} = \mathrm{argmax}_{\eta \geq 0} \int_{\gamma_L}^{\gamma_H} \frac{\eta\gamma^2}{(\eta+\gamma)^2}\mathrm{d}G(\gamma)$.

References

George A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500, 1970.

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

Gaurab Aryal, Isabelle Perrigne, Quang Vuong, and Haiqing Xu. Econometrics of insurance with multidimensional types. *Available at SSRN 2990738*, 2019.

Suleyman Basak and Georgy Chabakauri. Dynamic mean-variance asset allocation. *The Review of Financial Studies*, 23(8):2970–3016, 2010.

Tomas Björk and Agatha Murgoci. A general theory of markovian time inconsistent stochastic control problems. *Available at SSRN 1694759*, 2010.

Tomas Björk, Agatha Murgoci, and Xun Yu Zhou. Mean–variance portfolio optimization with state-dependent risk aversion. *Mathematical Finance*, 24(1):1–24, 2014.

Tomas Björk, Mariana Khapko, and Agatha Murgoci. On time-inconsistent stochastic control in continuous time. *Finance and Stochastics*, 21:331–360, 2017.

John Cawley and Tomas Philipson. An empirical examination of information barriers to trade in insurance. *American Economic Review*, 89(4):827–846, 1999.

Hector Chade and Edward Schlee. Optimal insurance with adverse selection. *Theoretical Economics*, 7(3):571–607, 2012.

Hector Chade and Edward E Schlee. Insurance as a lemons market: Coverage denials and pooling. *Journal of Economic Theory*, 189:105085, 2020.

Lv Chen and Yang Shen. On a new paradigm of optimal reinsurance: A stochastic stackelberg differential game between an insurer and a reinsurer. *Astin Bulletin*, 48(2):905–960, 2018.

Lv Chen and Yang Shen. Stochastic stackelberg differential reinsurance games under time-inconsistent mean-variance framework. *Insurance: Mathematics and Economics*, 88:120–137, 2019.

Alma Cohen and Liran Einav. Estimating risk preferences from deductible choice. *American economic review*, 97(3):745–788, 2007.

Keith J Crocker and Arthur Snow. Multidimensional screening in insurance markets with adverse selection. *Journal of Risk and Insurance*, 78(2):287–307, 2011.

Jakša Cvitanić and Jianfeng Zhang. Optimal compensation with adverse selection and dynamic actions. *Mathematics and Financial Economics*, 1:21–55, 2007.

Jakša Cvitanić, Xuhu Wan, and Huali Yang. Dynamics of contract design with screening. *Management Science*, 59(5):1229–1244, 2013.

Jakša Cvitanić, Dylan Possamaï, and Nizar Touzi. Moral hazard in dynamic risk management. *Management Science*, 63(10):3328–3346, 2017.

Min Dai, Hanqing Jin, Steven Kou, and Yuhong Xu. A dynamic mean-variance analysis for log returns. *Management Science*, 67(2):1093–1108, 2021.

Vitor Farinha Luz, Piero Gottardi, and Humberto Moreira. Risk classification in insurance markets with risk and preference heterogeneity. *Review of Economic Studies*, 90(6):3022–3082, 2023.

Amy Finkelstein and Kathleen McGarry. Multiple dimensions of private information: Evidence from the long-term care insurance market. *American Economic Review*, 96(4):938–958, 2006.

Daniel Garcia and Matan Tsur. Information design in competitive insurance markets. *Journal of Economic Theory*, 191:105160, 2021.

Alex Gershkov, Benny Moldovanu, Philipp Strack, and Mengxi Zhang. Optimal insurance: Dual utility, random losses, and adverse selection. *American Economic Review*, 113(10):2581–2614, 2023.

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

Christian Gollier. Optimal insurance design of ambiguous risks. *Economic Theory*, 57:555–576, 2014.

Guohui Guan and Bin Li. Equilibrium investment and reinsurance strategies under smooth ambiguity with a general second-order distribution. *Journal of Economic Dynamics and Control*, 143:104515, 2022.

Nathaniel Hendren. Private information and insurance rejections. *Econometrica*, 81(5):1713–1762, 2013.

Bengt Holmstrom and Paul Milgrom. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, pages 303–328, 1987.

Yuan-Mao Kao, N Bora Keskin, and Kevin Shang. Impact of information asymmetry and limited production capacity on business interruption insurance. *Management Science*, 68(4):2824–2841, 2022.

Michael Landsberger and Isaac Meilijson. Monopoly insurance under adverse selection when agents differ in risk aversion. *Journal of Economic Theory*, 63(2):392–407, 1994.

Bin Li, Danping Li, and Dewen Xiong. Alpha-robust mean-variance reinsurance-investment strategy. *Journal of Economic Dynamics and Control*, 70:101–123, 2016.

Eric Maskin and John Riley. Monopoly with incomplete information. *The RAND Journal of Economics*, 15(2):171–196, 1984.

Robert Puelz and Arthur Snow. Evidence on adverse selection: Equilibrium signaling and cross-subsidization in the insurance market. *Journal of Political Economy*, 102(2):236–257, 1994.

Michael Rothschild and Joseph Stiglitz. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4):629–649, 1976.

Yuliy Sannikov. A continuous-time version of the principal-agent problem. *The Review of Economic Studies*, 75(3):957–984, 2008.

Joseph E Stiglitz. Monopoly, non-linear pricing and imperfect information: The insurance market. *The Review of Economic Studies*, 44(3):407–430, 1977.

Robert Henry Strotz. *Myopia and inconsistency in dynamic utility maximization*. Springer, 1973.

Jaeyoung Sung. Optimal contracts under adverse selection and moral hazard: A continuous-time approach. *The Review of Financial Studies*, 18(3):1021–1073, 2005.

Achim Wambach. Introducing heterogeneity in the rothschild-stiglitz model. *Journal of Risk and Insurance*, pages 579–591, 2000.