Optimal Design of Fixed and Variable Costs in Peer-to-Peer Insurance with Heterogeneous Risk

Tim J. Boonen* Ze Chen† Wentao Hu‡

June 16, 2025

Abstract

This paper examines the optimal design of peer-to-peer insurance models, which combines outside insurance purchases with peer-to-peer risk sharing and heterogeneous risk. Participants contribute deposits to collectively cover the premium for group-based insurance against tail risks and to share uncovered losses. We analyze the cost structure by decomposing it into a fixed premium for outside coverage and a variable component for shared losses, the latter of which may be partially refunded if aggregate losses are sufficiently low. We derive closed-form solutions to the optimal sharing rule that maximizes a mean-variance objective from the perspective of a central or social planner, and we characterize its theoretical properties. Building on this foundation, we further investigate the choice of deposit for the common fund. Finally, we also provide numerical illustrations.

Keywords: Risk sharing; peer-to-peer insurance; optimal risk pooling

*Department of Statistics and Actuarial Science, School of Computing and Data Science, The University of Hong Kong, Hong Kong, People's Republic of China. E-mail: `tjboonen@hku.hk`.

†Corresponding author: China Institute of Insuance & School of Finance, Renmin University of China, Beijing, People's Republic of China. E-mail: `zechen@ruc.edu.cn`.

‡Faculty of Economics and Business, KU Leuven, Leuven, Belgium. E-mail: `wentao.hu@kuleuven.be`.

1 Introduction

Peer-to-peer (P2P) insurance has established itself as an innovative risk-sharing model in the InsurTech landscape since its introduction by Friendsurance in Germany during the 2010s. This model operates through a two-layer structure where participants collectively cover small-to-moderate losses within their community pool, while transferring excess losses beyond the pool's capacity to a conventional insurer. While this basic framework is well-documented in the existing literature, our study advances the field by examining its implementation in heterogeneous risk pools.

The distinctive feature of our approach lies in recognizing that participants bear two fundamentally different types of costs that require separate allocation rules. The fixed cost component represents each member's share of the total premium paid to the insurer, transferring risk outside the pool. In contrast, the variable costs component consists of refundable contributions to the common fund, embodying the risk-sharing nature of P2P insurance. Previous work, including Chen et al. (2023), has primarily examined these cost components in homogeneous groups where proportional sharing may suffice. However, this approach encounters limitations when applied to pools with diverse risk profiles.

Our research addresses this gap by developing a framework that ensures actuarial fairness in allocating both cost components across heterogeneous groups. Meanwhile, we determine the optimal deposit for the common fund, which plays a crucial role in determining the additional costs arising from risk loading and uncertainty derived from risk sharing. This dual approach maintains the economic viability of the P2P structure while properly accounting for differential risk exposures among participants. The practical implementation of this framework resolves a longstanding challenge in designing equitable P2P insurance schemes for diverse populations.

This paper makes the following theoretical contributions to the P2P insurance literature. We mathematically propose the actuarial fairness condition for multi-risk groups in P2P insurance. This issue incorporates the extra profit loading of group-covered insurance, which has not been explored in the literature. We then formulate and solve a mean-variance optimization problem for a central planner designing P2P insurance, which incorporates both a risk-sharing rule and a deposit. Our analysis suggests a trade-off and negative relation between the participant's nonrefundable contribution to the group insurance premium and the possibly refundable deposit to the common fund. We derive analytical solutions for the optimal design of P2P insurance both with and without the individual rationality (IR) constraints, and further characterize the theoretical properties.

Our research contributes to multiple areas of literature. First, we advance the literature on risk sharing. Prior work has extensively examined the conditional mean risk-sharing rule proposed by Denuit and Dhaene (2012), where participants share losses based on the conditional expectation of the participant's own loss, conditional on the aggregate loss. This rule satisfies some nice properties (Denuit et al., 2022; Jiao et al., 2023), and it particularly leads to an improvement of the *status quo* in the convex order. Various applications and extensions are proposed by Denuit (2019), Denuit et al. (2021), Denuit and Robert (2021a,b, 2022), and Clemente and Marano (2020), such as individual retention levels, cash backs, and stop-loss premiums. In particular, Denuit and Robert (2022) allows for outside insurance as well.

Second, this paper contributes to the literature on the optimal design of P2P insurance (see, e.g., Von Bieberstein et al., 2019; Chen and Feng, 2021; Charpentier et al., 2022; Levantesi and Piscopo, 2022). In heterogeneous settings, determining fair inter-personal loss sharing among participants is a key challenge. A number of recent studies tried to address this problem. For instance, Abdikerimova and Feng (2022) explores altruistic risk sharing in P2P insurance, a concept that is further developed in Feng et al. (2023) for flood risk pooling and extended to multi-period settings in Abdikerimova et al. (2024). Mutual aid—a simplified form of P2P insurance—has also received attention. Several studies provide formal analysis of mutual aid mechanisms and their theoretical foundations (see Chen et al., 2020; Li et al., 2023; Zhao and Zeng, 2023).

Third, by enabling the collective purchase of insurance, this work also contributes to the literature on optimal reinsurance. A rich body of research has examined optimal reinsurance design, often balancing the interests of both insurers and reinsurers (e.g., Li et al., 2014; Boonen et al., 2016; Chen et al., 2019). Some studies approach reinsurance from a risk-sharing or network perspective, such as Bäuerle and Glauner (2018). Further developments in this area can be found in Cai et al. (2017), Chi et al. (2017), Cheung et al. (2019), and Tang et al. (2022).

The subsequent sections of this paper are structured as follows. Section 2 outlines the general theoretical framework. Section 3 examines the optimal risk-sharing rule and deposit without the IR constraints, and Section 4 examines the optimal design under the IR constraints. Section 5 presents numerical examples, and Section 6 presents the conclusion. The proofs are delegated to the Appendix.

2 General framework

2.1 General setting

We consider a one-period setting. At time 0, each individual makes a payment and deposit as specified by the rule. At time 1, those who suffer a loss receive the predetermined compensation. During the period between time 0 and time 1, no participants join or leave the pool. Throughout this paper, all random variables are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We consider a group of n heterogeneous individuals. Let X_i be a non-negative, bounded random variable[1] defined on $\mathbb{R}_+ = [0, +\infty)$, representing the loss of the i-th participant. That is, $X_i \in [0, M_i]$ almost surely for some constant $M_i > 0$. If a loss occurs, any claimant will receive an amount of compensation equal to the amount of the loss. Let X denote the aggregate loss incurred by the entire group, defined as

$$X = \sum_{i=1}^{n} X_i.$$

Let M denote the upper bound of X, with cumulative distribution function (CDF) F_X, and survival function S_X. Moreover, our analysis does not require the specification of a

[1] It is reasonable to assume that X_i is bounded, as pricing or regulatory practices often introduce upper limits (e.g., compensation caps or reinsurance contracts) to bound unbounded random variables. The assumption that X_i is unbounded does not affect the main results. Indeed, the results obtained under the unbounded case can be regarded as a special instance of the present analysis, corresponding to the limit as $M_i \to +\infty$.

dependence structure among the losses.

2.2 Framework of P2P insurance

In the model of P2P insurance, participants' risks are aggregated and separated into two layers. The first layer of small-to-medium losses is retained within a common fund, and the second layer of severe losses is covered by an insurer. Specifically, at the start of the insurance period, all the participants contribute to a common fund by depositing D. At the end of the insurance period, the participants' losses are first covered by the common fund, with the insurer stepping in for any excess losses. Moreover, any remaining balance in the common fund is refunded to all participants (see Figure 1 for an illustration).

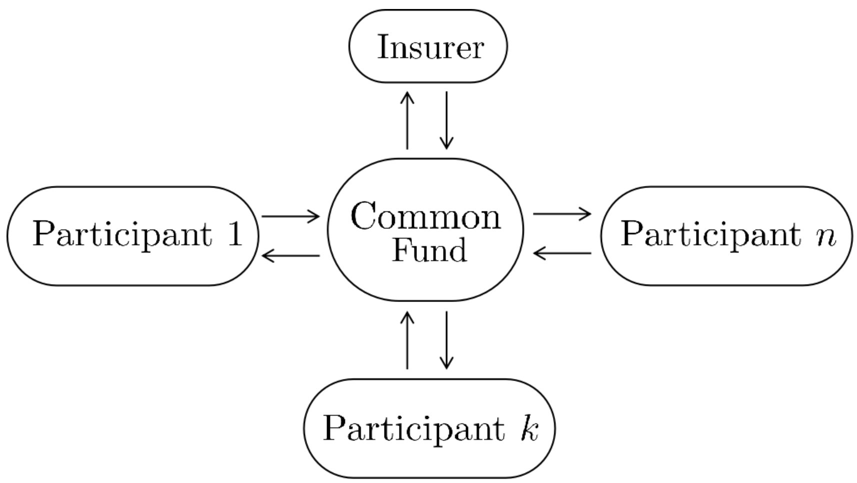

Figure 1: Overview of the framework of P2P insurance.

Under this arrangement, the insurer assumes liability for losses exceeding the deposit, i.e., for $X - D$. The common fund's risk-sharing responsibility is capped at D, ensuring limited liability for participants. Thus, the common fund must purchase insurance coverage with a deductible equal to the deposit D. In setting the premium, the insurer adds a risk loading $\theta > 0$ to the fair premium. Consequently, the total premium[2] is:

$$\Pi = (1 + \theta)\mathbb{E}[(X - D)_+]. \tag{1}$$

[2] Under the criterion of minimizing the variance of an insurer's risk exposure, Borch (1960) demonstrates that stop-loss reinsurance is optimal when the reinsurance premium is calculated by the expected value principle.

In the next section, we will decompose the costs involved in P2P insurance according to its underlying risk-sharing arrangement.

2.3 Fixed and variable costs

In P2P insurance, the entire group initially contributes a total amount of $\Pi + D$. The deposit D corresponds to the deductible, representing the portion of risk retained by the community. Due to the presence of a refund mechanism and the absence of risk loading, risk sharing within the common fund is typically more economical than the insurer's underwriting costs. However, the cost for risk sharing is variable, as it depends on the actual loss outcomes. In contrast, the insurer's underwriting cost Π is nonrefundable, generally higher, and predetermined.

We now provide a mathematical decomposition of the cost structure in P2P insurance. Let π_k and d_k for $k = 1, \ldots, n$ represent the portion of Π and D contributed by each participant, respectively. Thus, we have $D = \sum_{k=1}^{n} d_k$ and $\Pi = \sum_{k=1}^{n} \pi_k$. For simplicity of notation, we define

$$\alpha_k = \frac{\pi_k}{\Pi}, \text{ and } \beta_k = \frac{d_k}{D}.$$

By the summation condition, it follows that $\sum_{k=1}^{n} \alpha_k = 1$ and $\sum_{k=1}^{n} \beta_k = 1$. Here, α_k represents the ratio of the premium contributed by participant k to the total premium Π, and β_k represents the ratio of the deposit contributed to the total deposit D, reflecting their share of the remaining deposit in the common fund.

At time 0, participant k contributes $\pi_k + d_k$, or equivalently $\alpha_k \Pi + \beta_k D$, to the pool. At time 1, the participant receive a partial refund $\beta_k (D - X)_+$ from the common fund, where $z_+ = \max\{z, 0\}$. Under this decomposition, the cost allocation for each participant consists of a fixed premium and a variable deposit. The premium contribution, $\pi_k = \alpha_k \Pi$, is nonrefundable and fixed, whereas the deposit contribution, $d_k = \beta_k D$, is refundable depending on the ultimate loss outcome within the common fund. Using the notation of α_k and β_k, the cost for participant k in the P2P insurance is given by:

$$C_k = \alpha_k \Pi + \beta_k D - \beta_k (D - X)_+. \tag{2}$$

To summarize, the costs of P2P insurance consist of two parts (see Figure 2): (1) the *fixed contribution*, covering the group insurance premium paid to the insurer, and (2) the *variable sharing cost*, arising from the variability in the risk-sharing component managed by the common fund. This decomposition offers a flexible framework for adjusting each participant's share of both the premium and deposit according to their risk profile and risk aversion, thereby enabling a reasonable approach for risk sharing within the group.

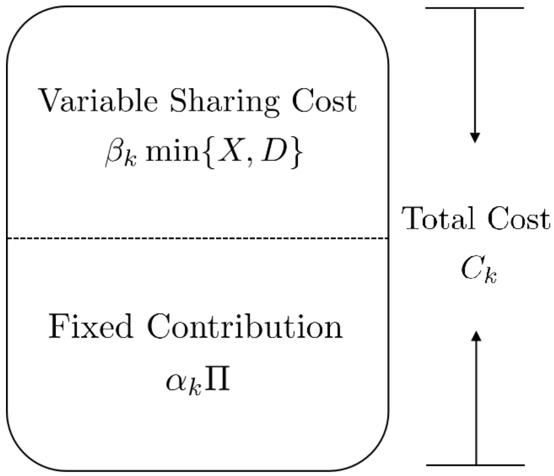

Figure 2: Decomposition of the cost in P2P insurance.

2.4 Actuarial fairness

In this section, we investigate actuarial fairness within the P2P insurance framework. Given participant heterogeneity, actuarial fairness is crucial for effectively pooling diverse risk types.

We begin with the expected cost for all participants in the P2P insurance model. Based on the participant's cost in Equation (2), we have:

$$\sum_{k=1}^{n} C_k = (1+\theta)\mathbb{E}[(X-D)_+] + \min\{D, X\}.$$

Using this result, the expected total cost of all participants is given by

$$\mathbb{E}\left[\sum_{k=1}^{n} C_k\right] = \mathbb{E}[X] + \Delta, \tag{3}$$

where $\Delta = \theta \mathbb{E}\left[(X-D)_+\right]$. Equation (3) shows that the sum of the participants' expected costs exceeds their expected income, with the additional cost component Δ arising from the insurer's risk loading.

To obtain actuarial fairness among peers, we decompose both $\mathbb{E}[X]$ and Δ among all participants. We decompose each participant's expected cost $\mathbb{E}[C_k]$ as his/her expected loss $\mathbb{E}[X_k]$ and an additional extra cost Δ_k. We have:

$$\mathbb{E}[C_k] = \mathbb{E}[X_k] + \Delta_k, \tag{4}$$

where $\sum_{k=1}^{n} \Delta_k = \Delta$. We define $\rho_k = \frac{\Delta_k}{\Delta}$ as the proportion of the total additional costs paid by participant k, and we can obtain $\sum_{k=1}^{n} \rho_k = 1$. Substituting Equation (2) into Equation (4) yields:

$$\mathbb{E}\left[\min\left\{D, X\right\}\right]\beta_k = \mathbb{E}[X_k] + \rho_k \theta \mathbb{E}[(X-D)_+] - \alpha_k(1+\theta)\mathbb{E}[(X-D)_+], \tag{5}$$

which establishes the specified relationship between α_k and β_k. It is straightforward to verify the summation condition $\sum_{k=1}^{n} \beta_k = 1$ holds naturally if $\sum_{k=1}^{n} \alpha_k = 1$.

Remark 1. *The fairness condition (4) requires that participants contribute an additional expected cost Δ_k on top of their own expected loss $\mathbb{E}[X_k]$. This differs from the traditional definition of actuarial fairness in the literature, where the additional costs incurred from external insurance are not considered, that is,*

$$\mathbb{E}[\tilde{C}_k] = \mathbb{E}[X_k],$$

where

$$\tilde{C}_k = \alpha_k \mathbb{E}[(X-D)_+] + \beta_k \min(D, X).$$

Our consideration of ρ_k stems from the fact that the additional cost Δ, arising from risk loading θ, are additional expenses borne by the policyholders. In the P2P framework, where all participants are treated equally, it is essential to ensure fairness among all participants. Particularly under heterogeneous conditions, individuals with different risk characteristics contribute differently to the aggregate risk, meaning that each participant contributes a different proportion of the extra costs. Therefore, when examining actuarial *fairness within the current framework, it is essential to account for the impact of these additional costs. Moreover, the classical condition for actuarial fairness in risk sharing (e.g., see Abdikerimova and Feng (2022)) emerges as a special case of of our proposed formulation when* $\Delta_k = 0$.

In addition, when $D = 0$, the P2P insurance model reduces to traditional insurance. In this case, there is no variable cost, and the parameter β_k becomes irrelevant. The allocation parameter α_k simplifies to:

$$\alpha_k = \frac{\mathbb{E}[X_k] + \rho_k \theta \mathbb{E}[X]}{(1+\theta)\mathbb{E}[X]},$$

which represents the distribution of the total premium among participants. Conversely, when $D = M$, participants fully share all losses, rendering α_k irrelevant. In this scenario, β_k is given by:

$$\beta_k = \frac{\mathbb{E}[X_k]}{\mathbb{E}[X]}.$$

These two extreme cases illustrate that it is meaningful to consider both α_k and β_k simultaneously only when $D \in (0, M)$. Under this condition, the relationship between α_k and β_k is characterized by the derivative:

$$\frac{\mathrm{d}\beta_k(\alpha_k)}{\mathrm{d}\alpha_k} = -\frac{(1+\theta)\mathbb{E}[(X-D)_+]}{\mathbb{E}[\min\{D, X\}]} < 0.$$

This implies a substitution effect: as participant k increases their premium share α_k, they can reduce their required deposit β_k. This trade-off aligns with intuitive expectations, as each participant aims to balance their contributions to the nonrefundable premium pool Π and the potentially refundable deposit D.

2.5 Optimization problem for central planner

In this study, we adopt a mean-variance framework, where the utility for participant k is given by $\mathbb{E}[W_k] - \gamma_k \operatorname{Var}[W_k]$, with $\gamma_k > 0$ and W_k is participant k's wealth at time 1. In the mean-variance framework, any deterministic initial wealth is irrelevant and is therefore normalized to zero. Thus, W_k can be expressed as

$$W_k = -C_k = -\alpha_k(1+\theta)\mathbb{E}[(X-D)_+] - \beta_k D + \beta_k (D-X)_+.$$

Suppose there exists a central planner whose objective is to maximize the collective mean-variance utility of all participants. As discussed earlier, adjusting the (α_k, β_k) allows for individual-level trade-offs between cost efficiency and variability, and further varying the deposit D enables this trade-off to be made at the aggregate level across all participants. Therefore, the central planner's decision variables are the risk-sharing rules (α_k, β_k) and the deposit level D. Thus, the central planner's optimization problem can be defined as follows:

$$\max_{\substack{\{\alpha_1,\ldots,\alpha_n\}\in\mathbb{R}^n \\ \{\beta_1,\ldots,\beta_n\}\in\mathbb{R}^n \\ D\in(0,M)}} \sum_{k=1}^{n} \left(\mathbb{E}[W_k] - \gamma_k \operatorname{Var}[W_k]\right).$$

We solve this optimization problem through a step-by-step approach as follows:

Step 1. Solve (α_k^*, β_k^*) for a given D, which we refer to as the optimal risk-sharing rule;

Step 2. Solve the optimal deposit $D^* \in (0, M)$, based on the results in Step 1.

In this way, we eventually determine the optimal parameters $(\alpha_1^*, \ldots, \alpha_n^*, \beta_1^*, \ldots, \beta_n^*, D^*)$.

In this paper, we analyze the optimization problem under two different scenarios: without individual rationality (IR) constraints, and with IR constraints. The IR constraint, which asserts that participation in the insurance should not be worse than maintaining individual risk, is often a foundational assumption in the optimal design of commercial insurance. In Section 3, we first derive the general optimal solution under the assumption that all participants are either willing or required to join the P2P insurance plan, thus disregarding the IR constraint. Subsequently, in Section 4, we introduce the IR constraint and characterize the resulting optimal solution.

3 Optimal design of P2P insurance without IR constraints

3.1 Optimal risk-sharing rules

First, we consider optimizing α_k and β_k for a given D. The corresponding optimization problem for Step 1 is:

Problem 1.

$$\max_{\substack{\{\alpha_1,\ldots,\alpha_n\}\in\mathbb{R}^n \\ \{\beta_1,\ldots,\beta_n\}\in\mathbb{R}^n}} \sum_{k=1}^{n} \left(\mathbb{E}[W_k] - \gamma_k \operatorname{Var}[W_k]\right), \tag{6}$$

$$s.t. \quad \sum_{k=1}^{n} \alpha_k = 1, \quad \sum_{k=1}^{n} \beta_k = 1, \tag{7}$$

$$\begin{aligned} &\mathbb{E}[X_k] + \rho_k\theta\mathbb{E}[(X-D)_+] - \alpha_k(1+\theta)\mathbb{E}[(X-D)_+] \\ &= \mathbb{E}\left[\min\{D, X\}\right]\beta_k, \quad \forall k \in \{1,\ldots,n\}. \end{aligned} \tag{8}$$

Solving Problem 1 yields the optimal solutions α_k^* and β_k^*, which are characterized as follows.

Theorem 1. *For a given D, the solution of Problem 1 is given by:*

$$\alpha_k^* = \frac{\mathbb{E}\left[X_k\right] + \rho_k\theta\mathbb{E}[(X-D)_+] - \beta_k^*\mathbb{E}\left[\min\{D,X\}\right]}{(1+\theta)\mathbb{E}[(X-D)_+]}, \tag{9}$$

$$\beta_k^* = \frac{1}{\gamma_k \sum_{l=1}^{n} \frac{1}{\gamma_l}}. \tag{10}$$

For the central planner, once the deposit D is determined, Theorem 1 provides the optimal allocation rule. It can be readily verified that the optimal β_k^* is always constrained within the interval $(0, 1)$. In contrast, no such constraint applies to α_k^* and the only requirements is that $\sum_{k=1}^{n} \alpha_k^* = 1$.

Remark 2. *In the homogeneous case, where all participants have identical risk preferences even if the risks are different, each individual equally shares the portion of risk not covered by the insurer. Consequently, the optimal deposit ratio is given by $\beta^* = 1/N$.*

Remark 3. *For the k-th participant, we have $\alpha_k^* < 0$ if*

$$\mathbb{E}\left[X_k\right] + \rho_k\theta\mathbb{E}[(X-D)_+] < \mathbb{E}\left[\beta_k^* \min\{D,X\}\right].$$

The condition suggests that $\alpha_k^* < 0$ may arise either when the participants bear a low ρ_k, indicating they bear a small share of the additional cost, or when the participants *have a large β_k^*, meaning they assume the excessive uncertain costs. Interestingly, when $\alpha_k^* < 0$, the participant may initially receive a payment from the common fund, rather than contributing to it. This situation, particularly when driven by a high β_k^*, implies that the participant requires compensation for shouldering a larger share of risk.*

The above facts reflect an interesting feature of P2P insurance that is different from traditional insurance: participants can take on the roles of both insurance 'buyer' and 'seller':

(1) As the 'buyer' type, characterized by $(\alpha_k^* > 0, \beta_k^* > 0)$, the participant pays a positive fixed cost, $\alpha_k^*\Pi$, and contributes $\beta_k^* D$ to cover variable costs.

(2) As the 'seller' type, characterized by $(\alpha_k^* < 0, \beta_k^* > 0)$, the participant receives an income cash inflow of $-\alpha_k^*\Pi$ but is required to deposit $\beta_k^* D$ into the common fund to cover variable costs.

Now, we examine the effect of parameters related to α_k^* and β_k^*. We begin by examining the effects of γ_k and γ_ℓ for $\ell \neq k$, on α_k^* and β_k^* in the following corollary.

Corollary 1. *For a given D, the pair (α_k^*, β_k^*) satisfies the following properties:*

(1) α_k^* increases with respect to γ_k, and decreases with respect to γ_ℓ for $\ell \neq k$,

(2) β_k^* decreases with respect to γ_k, and increases with respect to γ_ℓ for $\ell \neq k$.

Corollary 1 shows that, an increase in the degree of risk aversion γ_k leads to an increase in the participant's optimal fixed premium payment ratio α_k^* and a corresponding decrease in the optimal deposit ratio β_k^*. This outcome supports the idea that higher levels of risk aversion correlate with a preference for fixed cost structures.

Remark 4. *Considering a heterogeneous case where $n = 2$, we have:*

$$\frac{1}{\beta_1^*} = 1 + \frac{\gamma_1}{\gamma_2}, \quad \frac{1}{\beta_2^*} = 1 + \frac{\gamma_2}{\gamma_1}.$$

It is straightforward to verify that $\beta_1^* < \beta_2^*$ if $\gamma_1 > \gamma_2$, and vice versa. This example highlights the impact of risk aversion on β_k^*: participants with higher risk aversion contribute a *smaller share. We extend this analysis by introducing two types of participants, high-risk and low-risk, denoted by X_H and X_L, respectively. Let n_H and n_L denote the number of high-risk and low-risk participants. In this case, the optimal deposit ratios are given by:*

$$\frac{1}{\beta_H^*} = n_H + n_L\frac{\gamma_H}{\gamma_L}, \quad \frac{1}{\beta_L^*} = n_L + n_H\frac{\gamma_L}{\gamma_H}.$$

Corollary 2. *For a given D, α_k^* satisfies the following properties:*

(1) α_k^* increases with θ when $\alpha_k^* \leqslant \rho_k$,

(2) α_k^* decreases with θ when $\alpha_k^* \geqslant \rho_k$.

Corollary 2 states that α_k^* decreases with θ when ρ_k is below the threshold α_k^*. This implies that the peers are required to contribute a smaller share of the premium. Once ρ_k exceeds this threshold, α_k^* increases with θ, reflecting a greater required contribution.

Corollary 3. *For a given D, α_k^* satisfies the following properties:*

(1) α_k^* increases with D when $\beta_k^* \leq \frac{\mathbb{E}[X_k]}{\mathbb{E}[X]}$,

(2) α_k^* decreases with D when $\beta_k^* \geq \frac{\mathbb{E}[X_k]}{\mathbb{E}[X]}$.

This corollary shows that α_k^* increases with D when the deposit ratio β_k^* lower than their proportional risk exposure in the entire pool, and decreases otherwise. This conclusion aligns with the fairness condition, as it suggests that a group should contribute a larger share to the fixed premium when its deposit ratio is relatively lower, thus balancing the overall contributions within the pool.

Remark 5. *We consider a special case where $n = 2$. Suppose $\mathbb{E}[X_2] = a\mathbb{E}[X_1]$, and $\gamma_2 = b\gamma_1$. Then, if $a = 1$ and $b > 1$, we get*

$$\beta_1^* = \frac{1}{1+\frac{\gamma_1}{\gamma_2}} = \frac{1}{1+\frac{1}{b}} > \frac{1}{2} = \frac{\mathbb{E}[X_1]}{\mathbb{E}[X]}, \quad \beta_2^* = \frac{1}{1+\frac{\gamma_2}{\gamma_1}} = \frac{1}{1+b} < \frac{1}{2} = \frac{\mathbb{E}[X_2]}{\mathbb{E}[X]}.$$

In this case, α_1^* decreases with D and α_2^* increases with D. If $a > 1$ and $b = 1$, then

$$\beta_1^* = \frac{1}{1+\frac{\gamma_1}{\gamma_2}} = \frac{1}{2} > \frac{1}{1+a} = \frac{\mathbb{E}[X_1]}{\mathbb{E}[X]}, \quad \beta_2^* = \frac{1}{1+\frac{\gamma_2}{\gamma_1}} = \frac{1}{2} < \frac{1}{1+\frac{1}{a}} = \frac{\mathbb{E}[X_2]}{\mathbb{E}[X]}.$$

In this case, α_1^* decreases with D and α_2^* increases with D.

3.2 Optimal deposit

So far, we have solved (α_k^*, β_k^*) for any given D. Building on this analysis, we now proceed to determine the optimal deposit. The problem for Step 2 is formulated as follows:

Problem 2.

$$\max_{D\in(0,M)} \sum_{k=1}^{n} \left(\mathbb{E}[W_k] - \gamma_k \operatorname{Var}[W_k]\right). \tag{11}$$

Based on Theorem 1, we derive the following theorem.

Theorem 2. *The optimal deposit D^* for Problem 2 is characterized as follows: If*

$$\theta < \frac{2}{\sum_{k=1}^{n} \frac{1}{\gamma_k}} \left(M - \mathbb{E}[X]\right),$$

then D^* is the unique solution to the equation

$$\frac{\theta}{2} \sum_{k=1}^{n} \frac{1}{\gamma_k} = \int_0^{D^*} F_X(s)\,\mathrm{d}s. \tag{12}$$

Theorem 2 indicates that participants prefer to combine risk sharing with traditional insurance coverage when the risk loading factor θ is below a certain threshold. In contrast, when the insurance price is excessively high, the proof of Theorem 2 demonstrates that the objective function is increasing in D. This implies that participants are inclined to maximize their risk sharing within the common fund rather than relying on traditional insurance.

4 Optimal design of P2P insurance with IR constraints

In this section, we consider the optimization problem with the IR constraints, which ensure that participation in the P2P risk-sharing scheme is at least as favorable as retaining individual risk. Specifically, for each participant k, the following constraint must hold:

$$\mathbb{E}[W_k] - \gamma_k \operatorname{Var}[W_k] \geq \mathbb{E}[-X_k] - \gamma_k \operatorname{Var}[X_k], \quad \forall k \in \{1, \ldots, n\},$$

which implies the following condition:

$$\rho_k \leq \gamma_k \frac{\mathrm{Var}[X_k] - \beta_k^2 \,\mathrm{Var}[\min\{X, D\}]}{\theta \mathbb{E}[(X-D)_+]}, \quad \forall k \in \{1, \ldots, n\}. \tag{13}$$

Thus, the IR constraint can be equivalently formulated as a bound on the cost allocation rule ρ_k, which is inherently related to considerations of fairness.

Remark 6. *We introduce two common examples of ρ_k from a fairness perspective:*

(1) Uniform sharing principle. This approach equally distributes the excess cost $\theta\mathbb{E}[(X - D)_+]$ among all participants, such that $\rho_k = \frac{1}{n}$ for all k.

(2) Ex-ante mean proportion principle. This method allocates the additional cost in proportion to each participant's expected loss. For participant k, we have $\rho_k = \frac{\mathbb{E}[X_k]}{\mathbb{E}[X]}$.

Remark 7. *Abdikerimova and Feng (2022) propose a foundational fairness condition requiring that each participant's expected loss remains invariant before and after risk sharing. This corresponds to a specific case with $\theta = 0$, and naturally emerges in pure P2P risk-sharing pools where the absence of an insurer precludes additional costs.*

Under the IR constraints (13), we solve the central planner's problem:

$$\max_{\substack{\{\alpha_1,\ldots,\alpha_n\}\in\mathbb{R}^n \\ \{\beta_1,\ldots,\beta_n\}\in\mathbb{R}^n \\ D\in(0,M)}} \sum_{k=1}^{n} \left(\mathbb{E}[W_k] - \gamma_k \,\mathrm{Var}[W_k]\right) \tag{14}$$

by finding the parameters (α^*, β^*) for a given deposit D at first, and then calculating the optimal deposit D^*.

4.1 Optimal risk-sharing rules with IR constraints

Given $\rho_1, \ldots, \rho_n$ and D, we have the following Step 1 optimization problem:

Problem 3.

$$\max_{\substack{\{\alpha_1,\ldots,\alpha_n\}\in\mathbb{R}^n \\ \{\beta_1,\ldots,\beta_n\}\in\mathbb{R}^n}} \sum_{k=1}^{n} \left(\mathbb{E}[W_k] - \gamma_k \,\mathrm{Var}[W_k]\right), \tag{15}$$

$$s.t. \quad \sum_{k=1}^{n} \alpha_k = 1, \quad \sum_{k=1}^{n} \beta_k = 1, \tag{16}$$

$$\rho_k \leq \frac{\gamma_k \operatorname{Var}[X_k] - \gamma_k \beta_k^2 \operatorname{Var}[\min\{X, D\}]}{\theta \mathbb{E}[(X-D)_+]}, \quad \forall k \in \{1, \ldots, n\}, \tag{17}$$

$$\mathbb{E}[X_k] + \rho_k \theta \mathbb{E}[(X-D)_+] - \alpha_k (1+\theta) \mathbb{E}[(X-D)_+]$$
$$= \mathbb{E}\left[\min\{D, X\}\right] \beta_k, \quad \forall k \in \{1, \ldots, n\}. \tag{18}$$

By solving Problem 3, we obtain the following results of α_k^* and β_k^*.

Theorem 3. *Given $\rho_1, \ldots, \rho_n$ and D, if $\rho_k \leq \frac{\gamma_k \operatorname{Var}[X_k]}{\theta \mathbb{E}[(X-D)_+]}$ for any k, then no solution for Problem 3 exists. Otherwise, denote $\mathcal{J}$ be the set of all k satisfying:*

$$\rho_k > \frac{\gamma_k \operatorname{Var}[X_k] - \frac{1}{\gamma_k (\sum_{k=1}^{n} \frac{1}{\gamma_k})^2} \operatorname{Var}[\min\{X, D\}]}{\theta \mathbb{E}[(X-D)_+]}, \tag{19}$$

and $\mathcal{K}$ is given by $\mathcal{K} = \{1, \ldots, n\} \setminus \mathcal{J}$. Then,

(1) If $\mathcal{K}$ is nonempty, and for all $k \in \mathcal{K}$,

$$\rho_k \leq \frac{\gamma_k \operatorname{Var}[X_k] - \frac{\left(1 - \sum_{k \in \mathcal{J}} \sqrt{\frac{\gamma_k \operatorname{Var}[X_k] - \rho_k \theta \mathbb{E}[(X-D)_+]}{\gamma_k \operatorname{Var}[\min\{X, D\}]}}\right)^2}{\gamma_k \left(\sum_{k \in \mathcal{K}} \frac{1}{\gamma_k}\right)^2} \operatorname{Var}[\min\{X, D\}]}{\theta \mathbb{E}[(X-D)_+]}, \tag{20}$$

then the solution for Problem 3 is given by:

$$\alpha_k^* = \frac{\mathbb{E}\left[X_k\right] + \rho_k \theta \mathbb{E}[(X-D)_+] - \beta_k^* \mathbb{E}\left[\min\{D, X\}\right]}{(1+\theta) \mathbb{E}[(X-D)_+]}, \tag{21}$$

$$\beta_k^* = \frac{1 - \sum_{k \in \mathcal{J}} \beta_k^*}{\gamma_k \sum_{k \in \mathcal{K}} \frac{1}{\gamma_k}}, k \in \mathcal{K}, \tag{22}$$

$$\beta_k^* = \sqrt{\frac{\gamma_k \operatorname{Var}[X_k] - \rho_k \theta \mathbb{E}[(X-D)_+]}{\gamma_k \operatorname{Var}[\min\{X, D\}]}}, k \in \mathcal{J}, \tag{23}$$

where we note that if $\mathcal{J}$ is empty, then $\sum_{k \in \mathcal{J}} \beta_k^* = 0$.

(2) Otherwise, no solution exists for Problem 3.

Corollary 4. *If a solution exists for Problem 3, then it holds that $0 < \beta_k^* < 1$ for all $k \in \{1, \ldots, n\}$.*

For the central planner, once ρ_k and D are specified, Theorem 3 determines the optimal risk-sharing rule. When both $\mathcal{J}$ and $\mathcal{K}$ are nonempty, we observe the following distinct patterns: (1) For participants $k \in \mathcal{J}$, β_k^* decreases with respect to ρ_k. This inverse relationship indicates that the more additional cost a participant is allocated, the smaller the random cost they should bear. (2) For participants in $k \in \mathcal{K}$, β_k^* remains independent of their corresponding ρ_k, demonstrating complete independence between these variables.

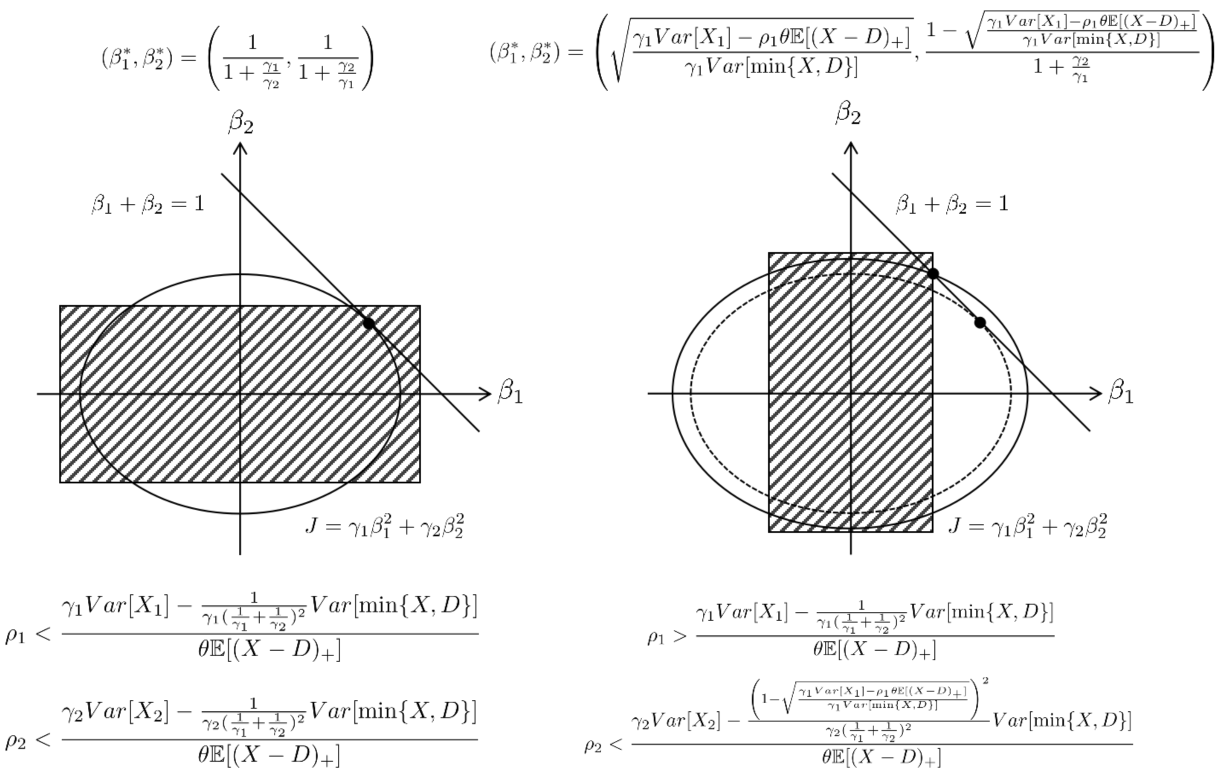

Figure 3: Schematic illustration of Theorem 3

Remark 8. *We illustrate the results of Theorem 3 using a simple example where $n = 2$. As depicted in Figure 3, the objective function forms an ellipse. The feasible solutions for (β_1, β_2) lie at the intersections of the ellipse and the constraint $\beta_1 + \beta_2 = 1$. The left panel of Figure 3 illustrates the case where $\mathcal{J}$ is empty. Here, the unconstrained minimum of the ellipse lies within the feasible region. The parameters (β_1^*, β_2^*) are given by: $\beta_1^* = \frac{1}{1+\frac{\gamma_1}{\gamma_2}}$ and $\beta_2^* = \frac{1}{1+\frac{\gamma_2}{\gamma_1}}$.*

The right panel of Figure 3 depicts the case where both $\mathcal{J}$ and $\mathcal{K}$ are nonempty. In this case, the feasible region for β_1 excludes the unconstrained minimizer. As a result,

β_1^* lies on the boundary. Although the feasible region for β_2 includes the unconstrained minimizer, the constraint $\beta_2^* = 1 - \beta_1^*$ forces β_2^* to exceed $\frac{1}{1+\frac{\gamma_2}{\gamma_1}}$ because β_1^* is below its unconstrained optimum.

4.2 Optimal deposit with IR constraints

4.2.1 Characterization of feasible region

Before we derive the optimal deposit D^*, we first characterize the feasible region $\mathcal{D}$ under IR constraints in this section.

As established in Theorem 3, when both $\mathcal{K}$ and $\mathcal{J}$ are nonempty, the index set pair $(\mathcal{K}, \mathcal{J})$ admits a total of $2^n - 2$ configurations. Including the special case where $\mathcal{J}$ is empty, the total number of configurations becomes $2^n - 1$. Each configuration corresponds to a (possibly overlapping) subinterval of the deposit domain $(0, M)$, within which the solution (α_k^*, β_k^*) exists. Given the parameters $\rho_1, \ldots, \rho_n$, let

$$\mathcal{D} = \bigcup_{t=1}^{2^n - 1} \mathcal{D}_t$$

denote the feasible region for deposits, where $\mathcal{D}_t$ denote the subinterval associated with the t-th configuration. Before proceeding with the analysis, we distinguish between two cases in Theorem 3:

- When $\mathcal{J}$ is empty: if ρ_k satisfies the following condition:

$$\rho_k \leq \frac{\gamma_k \operatorname{Var}[X_k] - \frac{1}{\gamma_k \left(\sum_{k=1}^{n} \frac{1}{\gamma_k}\right)^2} \operatorname{Var}[\min\{X, D\}]}{\theta \mathbb{E}[(X - D)_+]}, \ \forall k \in \{1, \ldots, n\}, \tag{24}$$

  then Problem 3 reduces to the unconstrained formulation in Problem 1. In this case, the solution (α_k^*, β_k^*) coincides with that in Theorem 1, where β_k^* depends only on risk aversion and is independent of D, as shown in equation (10).

- When $\mathcal{J}$ is nonempty: if instead

$$\frac{\gamma_k \operatorname{Var}[X_k] - \frac{1}{\gamma_k \left(\sum_{k=1}^{n} \frac{1}{\gamma_k}\right)^2} \operatorname{Var}[\min\{X, D\}]}{\theta \mathbb{E}[(X - D)_+]} < \rho_k \leq \frac{\gamma_k \operatorname{Var}[X_k]}{\theta \mathbb{E}[(X - D)_+]}, \ \forall k \in \mathcal{J}, \tag{25}$$

and for all $k \in \mathcal{K}$, ρ_k satisfy (20), then each β_k^* is determined at the boundary of constraint (17). As shown in equations (22) and (23), β_k^* becomes a function of D.

This distinction is essential, as the relationship between β^* and D fundamentally differs between the two cases. For convenience, we label the subinterval corresponding to inequality (24) as $\mathcal{D}_1$, and those corresponding to inequality (25) as $\mathcal{D}_t$, where $t \in \{2, \ldots, 2^n - 1\}$.

In the following sections, we examine the structure of each subinterval $\mathcal{D}_t$, and then determine the existence and uniqueness of a local optimum D_t^*, and the closed-form expression for D_t^* when it exists. Specifically, in Sections 4.2.2 and 4.2.3, we introduce the local optimal D_1^* within $\mathcal{D}_1$ and $\mathcal{D}_t$ for $t \in \{2, \ldots, 2^n - 1\}$, respectively. Finally, the global optimal deposit D^* over the entire feasible region $\mathcal{D}$ is determined as the one among these D_t^* that yields the highest objective function value.

4.2.2 Local optimal D_1^* within $\mathcal{D}_1$

In this section, we start from $\mathcal{D}_1$ where inequality (24) holds. For simplicity, we define two tool functions as:

$$\phi_k(D) = \frac{\gamma_k \operatorname{Var}[X_k] - \frac{1}{\gamma_k \left(\sum_{k=1}^n \frac{1}{\gamma_k}\right)^2} \operatorname{Var}[\min\{X, D\}]}{\theta\, \mathbb{E}[(X - D)_+]},$$

$$\psi_k(D) = -\frac{2(D - \mathbb{E}[\min\{X, D\}])\mathbb{E}[(X - D)_+] + \operatorname{Var}[\min\{X, D\}]}{\gamma_k^2 \left(\sum_{k=1}^n \frac{1}{\gamma_k}\right)^2} + \operatorname{Var}[X_k].$$

Then, we characterize the structure of $\mathcal{D}_1$ through the following lemma:

Lemma 1. *Given $\rho_1, \ldots, \rho_n$, the interval $\mathcal{D}_1$ is given by:*

$$\mathcal{D}_1 = \bigcap_{k=1}^n \tilde{\mathcal{D}}_k,$$

where each $\tilde{\mathcal{D}}_k$ is determined according to the following cases:

(1) When

$$\operatorname{Var}[X_k] - \frac{\operatorname{Var}[X]}{\gamma_k^2 \left(\sum_{j=1}^n \frac{1}{\gamma_j}\right)^2} \geq 0,$$

(a) *If* $\rho_k > \frac{\gamma_k \operatorname{Var}[X_k]}{\theta \mathbb{E}[X]}$, *there exists a unique* $\hat{D}_k$ *such that* $\phi_k(\hat{D}_k) = \rho_k$, *and we have* $\tilde{\mathcal{D}}_k = [\hat{D}_k, M)$;

(b) *If* $\rho_k \leq \frac{\gamma_k \operatorname{Var}[X_k]}{\theta \mathbb{E}[X]}$, *we have* $\tilde{\mathcal{D}}_k = (0, M)$.

(2) *When*

$$\operatorname{Var}[X_k] - \frac{\operatorname{Var}[X]}{\gamma_k^2 \left(\sum_{j=1}^n \frac{1}{\gamma_j}\right)^2} < 0,$$

then there exists a unique $\tilde{D}_k$ *such that* $\psi_k(\tilde{D}_k) = 0$. *Then,*

(a) *If* $\rho_k > \phi_k(\tilde{D}_k)$, *we have* $\tilde{\mathcal{D}}_k = \emptyset$;

(b) *If* $\frac{\gamma_k \operatorname{Var}[X_k]}{\theta \mathbb{E}[X]} < \rho_k \leq \phi_k(\tilde{D}_k)$, *there exist* $\underline{\hat{D}}_k \leq \hat{\overline{D}}_k$ *such that* $\phi_k(\underline{\hat{D}}_k) = \phi_k(\hat{\overline{D}}_k) = \rho_k$, *and we have* $\tilde{\mathcal{D}}_k = [\underline{\hat{D}}_k, \hat{\overline{D}}_k]$;

(c) *If* $\rho_k \leq \frac{\gamma_k \operatorname{Var}[X_k]}{\theta \mathbb{E}[X]}$, *we have* $\tilde{\mathcal{D}}_k = (0, \hat{D}_k]$.

In what follows, we identify the associated local optimal deposit D_1^* within $\mathcal{D}_1$, if it exists. Note that, for given $\rho_1, \ldots, \rho_n$, $\mathcal{D}_1$ is possibly empty. To proceed with the analysis, we assume that $\mathcal{D}_1$ is a nonempty closed interval, ensuring the existence of a D_1^*. Then we establish the following theorem.

Theorem 4. *Assume that* $\mathcal{D}_1$ *is a nonempty closed interval, denoted by* $\mathcal{D}_1 = [\underline{D}_1, \overline{D}_1]$. *Given* $\rho_1, \ldots, \rho_n$, D_1^* *is given by:*

(1) *If*

$$\theta - \frac{2}{\sum_{k=1}^n \frac{1}{\gamma_k}} \int_0^{\underline{D}_1} F_X(s) \mathrm{d}s \leq 0,$$

then $D_1^* = \underline{D}_1$.

(2) *If*

$$\theta - \frac{2}{\sum_{k=1}^n \frac{1}{\gamma_k}} \int_0^{\overline{D}_1} F_X(s) \mathrm{d}s \geq 0,$$

then $D_1^* = \overline{D}_1$.

(3) *If*

$$\theta - \frac{2}{\sum_{k=1}^n \frac{1}{\gamma_k}} \int_0^{\underline{D}_1} F_X(s) \mathrm{d}s > 0, \text{ and } \theta - \frac{2}{\sum_{k=1}^n \frac{1}{\gamma_k}} \int_0^{\overline{D}_1} F_X(s) \mathrm{d}s < 0,$$

then D_1^* is determined as follows:

$$D_1^* = \underline{D}_1 \cdot \mathbb{I}[D^o < \underline{D}_1] + D^o \cdot \mathbb{I}[\underline{D}_1 \leq D^o \leq \overline{D}_1] + \overline{D}_1 \cdot \mathbb{I}[D^o > \overline{D}_1], \tag{26}$$

where D^o is the unique solution to the equation:

$$\frac{\theta}{2} \sum_{k=1}^{n} \frac{1}{\gamma_k} = \int_0^{D^o} F_X(s) \, \mathrm{d}s. \tag{27}$$

4.2.3 Local optimal D_t^* within $\mathcal{D}_t$ for $t \in \{2, \ldots, 2^n - 1\}$

We now turn to $\mathcal{D}_t$ for $t \in \{2, \ldots, 2^n - 1\}$ where the inequality (25) holds. In what follows, we analyze the structure of $\mathcal{D}_t$ for a given t. We denote the index sets corresponding to the k-th configuration as $\mathcal{K}_t$ and $\mathcal{J}_t$ respectively. Similar to Lemma 1, we have

$$\mathcal{D}_t = \tilde{\mathcal{D}}_{\mathcal{J}_t} \bigcap \tilde{\mathcal{D}}_{\mathcal{K}_t},$$

where

$$\tilde{\mathcal{D}}_{\mathcal{K}_t} = \bigcap_{k \in \mathcal{K}_t} \tilde{\mathcal{D}}_k, \quad \text{and} \quad \tilde{\mathcal{D}}_{\mathcal{J}_t} = \bigcap_{k \in \mathcal{J}_t} \tilde{\mathcal{D}}_k.$$

We first discuss the structure of $\tilde{\mathcal{D}}_{\mathcal{J}_t}$. For simplicity, we define a tool function as:

$$\tilde{\phi}_k(D) = \frac{\gamma_k \operatorname{Var}[X_k]}{\theta \mathbb{E}[(X - D)_+]},$$

and then have the following lemma:

Lemma 2. *Given $\rho_1, \ldots, \rho_n$, the interval $\tilde{\mathcal{D}}_{\mathcal{J}_t}$ can be expressed as:*

$$\tilde{\mathcal{D}}_{\mathcal{J}_t} = \bigcap_{k \in \mathcal{J}_t} \tilde{\mathcal{D}}_k,$$

where each $\tilde{\mathcal{D}}_k$ is determined according to the following cases:

(1) When

$$\operatorname{Var}[X_k] - \frac{\operatorname{Var}[X]}{\gamma_k^2 \left(\sum_{j=1}^{n} \frac{1}{\gamma_j} \right)^2} \geq 0,$$

(a) If $\rho_k > \frac{\gamma_k \operatorname{Var}[X_k]}{\theta \, \mathbb{E}[X]}$, there exists a unique $\hat{D}_k$ such that $\phi_k(\hat{D}_k) = \rho_k$, and a unique $D_k^\sharp$ such that $\tilde{\phi}_k(D_k^\sharp) = \rho_k$, and then we have $\tilde{\mathcal{D}}_k = [D_k^\sharp, \hat{D}_k]$;

(b) *If* $\rho_k \le \frac{\gamma_k \operatorname{Var}[X_k]}{\theta\,\mathbb{E}[X]}$, *we have* $\tilde{\mathcal{D}}_k = \emptyset$.

(2) *When*

$$\operatorname{Var}[X_k] - \frac{\operatorname{Var}[X]}{\gamma_k^2 \left(\sum_{j=1}^n \frac{1}{\gamma_j}\right)^2} < 0,$$

then there exists a unique $\tilde{D}_k$ *such that* $\psi_k(\tilde{D}_k) = 0$. *Then,*

(a) *If* $\rho_k > \phi_k(\tilde{D}_k)$, *we have* $\tilde{\mathcal{D}}_k = [D_k^\sharp, M)$;

(b) *If* $\frac{\gamma_k \operatorname{Var}[X_k]}{\theta\,\mathbb{E}[X]} < \rho_k \le \phi_k(\tilde{D}_k)$, *then there exist* $\underline{\hat{D}}_k \le \hat{\overline{D}}_k$ *such that* $\phi_k(\underline{\hat{D}}_k) = \phi_k(\hat{\overline{D}}_k) = \rho_k$, *and we have* $\tilde{\mathcal{D}}_k = [D_k^\sharp, \underline{\hat{D}}_k]$, *or* $\tilde{\mathcal{D}}_k = [\underline{\hat{D}}_k, M)$;

(c) *If* $\rho_k \le \frac{\gamma_k \operatorname{Var}[X_k]}{\theta\,\mathbb{E}[X]}$, *we have* $\tilde{\mathcal{D}}_k = [\hat{D}_k, M)$.

Notice that the constraints associated with $\tilde{\mathcal{D}}_{\mathcal{K}_t}$ are inherently dependent on the structure of $\tilde{\mathcal{D}}_{\mathcal{J}_t}$. As shown in the proof of Theorem 3, the inequality (20) is meaningful only if the set $\tilde{\mathcal{D}}_{\mathcal{J}_t}$ is nonempty. To ensure the validity of the analysis for $\tilde{\mathcal{D}}_{\mathcal{K}_t}$, $\tilde{\mathcal{D}}_{\mathcal{J}_t}$ must be nonempty. Before delving into the structure of $\tilde{\mathcal{D}}_{\mathcal{K}_t}$, we first introduce the following tool functions:

$$\kappa_k(D) = \left(1 - \sum_{k\in\mathcal{J}} \sqrt{\frac{\gamma_k \operatorname{Var}[X_k] - \rho_k\theta\mathbb{E}[(X-D)_+]}{\gamma_k \operatorname{Var}[\min\{X,D\}]}}\right)^2 \operatorname{Var}[\min\{X,D\}],$$

$$\eta_k(D) = \frac{\gamma_k \operatorname{Var}[X_k] - \frac{\kappa_k(D)}{\gamma_k\left(\sum_{k\in\mathcal{K}} \frac{1}{\gamma_k}\right)^2}}{\theta\mathbb{E}[(X-D)_+]},$$

$$h(D) = -\frac{\frac{\mathrm{d}\kappa_k(D)}{\mathrm{d}D}\mathbb{E}[(X-D)_+]}{\gamma_k\left(\sum_{k\in\mathcal{K}} \frac{1}{\gamma_k}\right)^2} + S_X(D)\left(\gamma_k \operatorname{Var}[X_k] - \frac{\kappa_k(D)}{\gamma_k\left(\sum_{k\in\mathcal{K}} \frac{1}{\gamma_k}\right)^2}\right).$$

Then, we can establish the following lemma:

Lemma 3. *Assume that* $\tilde{\mathcal{D}}_{\mathcal{J}_t}$ *is a nonempty closed interval, denoted by* $\tilde{\mathcal{D}}_{\mathcal{J}_t} = [\underline{D}_{\mathcal{J}_t}, \overline{D}_{\mathcal{J}_t}]$. *Given* $\rho_1, \ldots, \rho_n$, *then* $\tilde{\mathcal{D}}_{\mathcal{K}_t}$ *can be expressed as:*

$$\tilde{\mathcal{D}}_{\mathcal{K}_t} = \bigcap_{k\in\mathcal{K}_t} \tilde{\mathcal{D}}_k,$$

where each $\tilde{\mathcal{D}}_k$ *is determined according to the following cases:*

(1) *When* $h_k(\underline{D}_{\mathcal{J}_t}) \geq 0$ *and* $h_k(\overline{D}_{\mathcal{J}_t}) \geq 0$*, then:*

(a) *If* $\rho_k > \eta_k(\overline{D}_{\mathcal{J}_t})$*, we have* $\tilde{\mathcal{D}}_k = \emptyset$*;*

(b) *If* $\rho_k < \eta_k(\underline{D}_{\mathcal{J}_t})$*, we have* $\tilde{\mathcal{D}}_k = [\underline{D}_{\mathcal{J}_t}, \overline{D}_{\mathcal{J}_t}]$*;*

(c) *If* $\eta_k(\underline{D}_{\mathcal{J}_t}) \leq \rho_k \leq \eta_k(\overline{D}_{\mathcal{J}_t})$*, there uniquely exists* D_k^+ *such that* $\rho_k = \eta_k(D_k^+)$*, and we have* $\tilde{\mathcal{D}}_k = [D_k^+, \overline{D}_{\mathcal{J}_t}]$*.*

(2) *When* $h_k(\underline{D}_{\mathcal{J}_t}) \leq 0$ *and* $h_k(\overline{D}_{\mathcal{J}_t}) \leq 0$*, then:*

(a) *If* $\rho_k > \eta_k(\underline{D}_{\mathcal{J}_t})$*, we have* $\tilde{\mathcal{D}}_k = \emptyset$*;*

(b) *If* $\rho_k < \eta_k(\overline{D}_{\mathcal{J}_t})$*, we have* $\tilde{\mathcal{D}}_k = [\underline{D}_{\mathcal{J}_t}, \overline{D}_{\mathcal{J}_t}]$*;*

(c) *If* $\eta_k(\overline{D}_{\mathcal{J}_t}) \leq \rho_k \leq \eta_k(\underline{D}_{\mathcal{J}_t})$*, we have* $\tilde{\mathcal{D}}_k = [\underline{D}_{\mathcal{J}_t}, D_k^+]$*.*

(3) *When* $h_k(\underline{D}_{\mathcal{J}_t}) > 0$ *and* $h_k(\overline{D}_{\mathcal{J}_t}) < 0$*, there uniquely exists* D_k^η *such that* $h_k(D_k^\eta) = 0$*, then:*

(a) *If* $\rho_k > \eta_k(D_k^\eta)$*, we have* $\tilde{\mathcal{D}}_k = \emptyset$*;*

(b) *If* $\max\{\eta_k(\overline{D}_{\mathcal{J}_t}), \eta_k(\underline{D}_{\mathcal{J}_t})\} \leq \rho_k \leq \eta_k(D_k^\eta)$*, then there exist* $\underline{D}_k^+ \leq \overline{D}_k^+$ *such that* $\rho_k = \eta_k(\underline{D}_k^+) = \eta_k(\overline{D}_k^+)$*, and we have* $\tilde{\mathcal{D}}_k = [\underline{D}_{\mathcal{J}_t}, \underline{D}_k^+]$ *or* $\tilde{\mathcal{D}}_k = [\overline{D}_k^+, \overline{D}_{\mathcal{J}_t}]$*;*

(c) *If* $\min\{\eta_k(\overline{D}_{\mathcal{J}_t}), \eta_k(\underline{D}_{\mathcal{J}_t})\} \leq \rho_k < \max\{\eta_k(\overline{D}_{\mathcal{J}_t}), \eta_k(\underline{D}_{\mathcal{J}_t})\}$*, then we have* $\tilde{\mathcal{D}}_k = [\underline{D}_{\mathcal{J}_t}, D_k^+]$ *if* $\eta_k(\overline{D}_{\mathcal{J}_t}) > \eta_k(\underline{D}_{\mathcal{J}_t})$*, and* $\tilde{\mathcal{D}}_k = [D_k^+, \overline{D}_{\mathcal{J}_t}]$ *if* $\eta_k(\overline{D}_{\mathcal{J}_t}) < \eta_k(\underline{D}_{\mathcal{J}_t})$*;*

(d) *If* $\rho_k < \min\{\eta_k(\overline{D}_{\mathcal{J}_t}), \eta_k(\underline{D}_{\mathcal{J}_t})\}$*, we have* $\tilde{\mathcal{D}}_k = \emptyset$*.*

Having characterized each $\mathcal{D}_t$ for $t \in \{2, \ldots, 2^n - 1\}$, we now identify the local optimal deposit D_t^*, when it exists. Similar to the approach in Theorem 4, we assume that $\mathcal{D}_t$ is a nonempty closed interval, which guarantees the existence of a local optimum. We obtain the following result.

Theorem 5. *Assume that* $\mathcal{D}_t$ *is a nonempty closed interval, denoted by* $\mathcal{D}_t = [\underline{D}_t, \overline{D}_t]$*. Given* $\rho_1, \ldots, \rho_n$*, if there exist* $D_t^o \in \mathcal{D}_t$ *which is the solution to the following equation:*

$$\sum_{k \in \mathcal{K}_t} \rho_k \theta S_X(D) = \frac{2}{\sum_{k \in \mathcal{K}_t} \frac{1}{\gamma_k}} \left(\sqrt{\mathrm{Var}[\min\{X, D\}]} - \sum_{k \in \mathcal{J}_t} \sqrt{\mathrm{Var}[X_k] - \frac{\rho_k}{\gamma_k} \theta \mathbb{E}[(X - D)_+]} \right)$$

$$\cdot\, S_X(D)\left(\frac{D-\mathbb{E}[\min\{X,D\}]}{\sqrt{\operatorname{Var}[\min\{X,D\}]}}-\sum_{k\in\mathcal{J}_t}\frac{\frac{\rho_k}{\gamma_k}\theta}{2\sqrt{\operatorname{Var}[X_k]-\frac{\rho_k}{\gamma_k}\theta\mathbb{E}[(X-D)_+]}}\right),$$

then $D_t^*=D_t^o$. *Otherwise,*

$$D_t^*=\arg\max_{D\in\{\underline{D}_t,\overline{D}_t\}}\sum_{k=1}^{n}\left(\mathbb{E}[W_k]-\gamma_k\operatorname{Var}[W_k]\right).$$

4.2.4 Characterization of global optimal deposit D^*

In the previous sections, we have established that the global feasible region $\mathcal{D}$ is the union of 2^n-1 subfeasible regions $\mathcal{D}_t$, each corresponding to a configuration of the index sets $(\mathcal{K},\mathcal{J})$. We also examined each $\mathcal{D}_t$ individually and identify the associated local optimal deposit D_t^*, if it exists.

Once all potentially valid local optima D_t^* are obtained, the global optimal deposit D^* is identified as the one maximizing the objective function. Specifically, $D^*=D_{t^*}^*$ with

$$t^*=\arg\max_{t\in\mathcal{T}}\sum_{k=1}^{n}\left(\mathbb{E}[W_k(D_t^*)]-\gamma_k\operatorname{Var}[W_k(D_t^*)]\right), \tag{28}$$

and

$$\mathcal{T}=\{t\in\{1,\ldots,2^n-1\}\mid D_t^*\text{ exists}\}. \tag{29}$$

In the following section, we illustrate our findings through numerical simulations.

5 Numerical illustration

In this section, we numerically illustrate the optimal design of rules, beginning with our calibration approach and followed by a detailed analysis of the results.

5.1 Parameter selection

We consider a simplified case with two participants ($n=2$): a high-risk type whose risk is denoted by X_H, and a low-risk type whose risk is denoted by X_L. We assume that X_H

and X_L are independent and follow a gamma distribution with the probability density function:

$$f(x) = \frac{1}{\Gamma(k)\eta^k} x^{k-1} e^{-\frac{x}{\eta}}, \quad x > 0,$$

where $\Gamma(k)$ is the gamma function.

In this illustration, we set the baseline parameters as $k_L = 2$, $k_H = 3$, $\eta_L = 1$, and $\eta_H = 2$, yielding $\mathbb{E}[X_H] = 6$, $\mathrm{Var}[X_H] = 12$, $\mathbb{E}[X_L] = 2$, and $\mathrm{Var}[X_L] = 2$, while setting the benchmark loading parameter at $\theta = 0.5$. We use the ex-ante mean proportion principle, which yields $\rho_H = \mathbb{E}[X_H]/\mathbb{E}[X] = 0.75$ and $\rho_L = \mathbb{E}[X_L]/\mathbb{E}[X] = 0.25$, where $\mathbb{E}[X] = \mathbb{E}[X_H] + \mathbb{E}[X_L]$ represents the aggregate expected loss. For risk aversion, we consider two scenarios:

Scenario 1. $\gamma_H = 1$, $\gamma_L = 0.25$, indicating that the low-risk group is less risk averse.

Scenario 2. $\gamma_H = 0.25$, $\gamma_L = 1$, indicating that the low-risk group is more risk averse.

5.2 Numerical results

Based on the above calibrations, we conduct a numerical analysis, presenting the optimized parameters α^*, β^*, and D^* across the considered scenarios in Table 1. In Scenario 1, α_H^* is positive, while α_L^* is negative, indicating that the low-risk type aligns with the 'seller' type discussed earlier. In Scenario 2, both α_H^* and α_L^* are positive, implying that both high- and low-risk types exhibit 'buyer' type behavior. Comparing α^*, β^*, and D^* with and without the IR constraints reveals that the IR condition is binding in Scenario 1, leading to distinct optimal solutions. In contrast, the IR condition is nonbinding in Scenario 2, resulting in identical solutions.

Table 1: α^*, β^*, and D^* in considered scenarios

	α_H^*	α_L^*	β_H^*	β_L^*	D^*
Scenario 1 (with IR)	2.0499	-1.0499	0.2105	0.7895	7.6338
Scenario 1 (without IR)	2.1781	-1.1781	0.2003	0.7997	7.6025
Scenario 2 (with IR)	0.6178	0.3822	0.7995	0.2005	7.6571
Scenario 2 (without IR)	0.6178	0.3822	0.7995	0.2005	7.6571

In the following, we further examine how the optimal parameters (α^*, β^*, D^*) vary with respect to the loading parameter θ under Scenario 1. First, Figure 4 presents the results without the IR constraint. We observe that α_H^* increases with the insurer's loading factor θ. As expected, since $\alpha_H^* + \alpha_L^* = 1$, α_L^* follows the opposite trend. Additionally, we find that β_H^* and β_L^* remain independent of θ. This result is consistent with Equation (10) in Theorem 1, as the optimal β^* depends solely on the values of γs. Moreover, D^* increases with θ, indicating the pool's greater willingness to share risk.

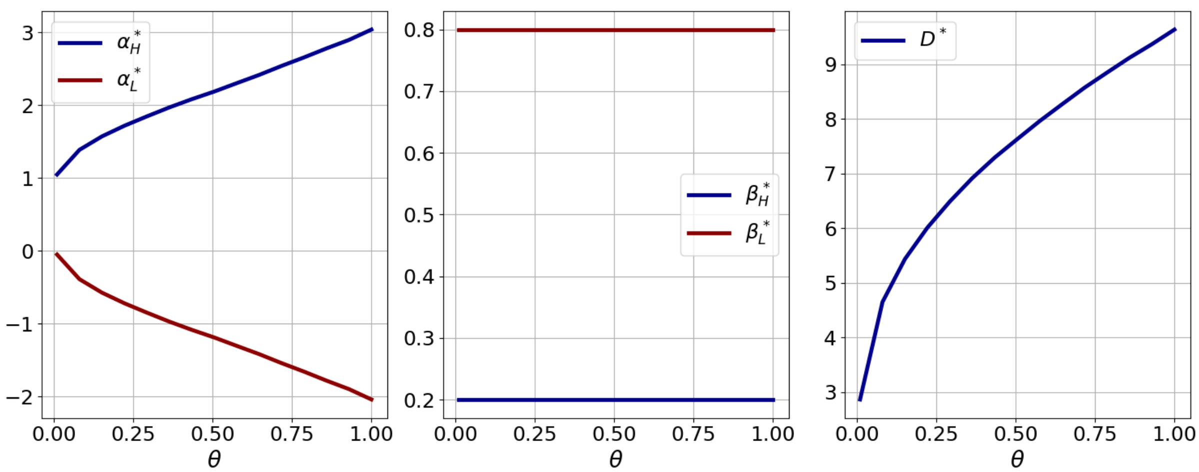

Figure 4: Sensitivity analysis of θ without IR condition.

Second, Figure 5 presents the sensitivity analysis of the optimal contract parameters (α^*, β^*, D^*) under the IR constraint with respect to the loading parameter θ. The analysis reveals three distinct behavioral patterns. First, α_H^* exhibits a nonmonotonic relationship with θ, initially increasing before decreasing, while α_L^* follows the opposite trajectory, first decreasing then increasing; this phase transition occurs precisely when θ crosses the threshold at which the IR constraint becomes binding. Second, the β^* parameters show threshold-dependent behavior: both β_H^* and β_L^* remain constant in the unconstrained regime but diverge significantly once the IR condition takes effect, with β_H^* decreasing and β_L^* increasing monotonically with θ. Finally, the optimal deductible D^* increases monotonically with θ until the IR constraint becomes binding, after which it remains relatively stable. These results collectively demonstrate how the IR condition serves as a bifurcation point that fundamentally alters the sensitivity of optimal contract parameters to the loading factor.

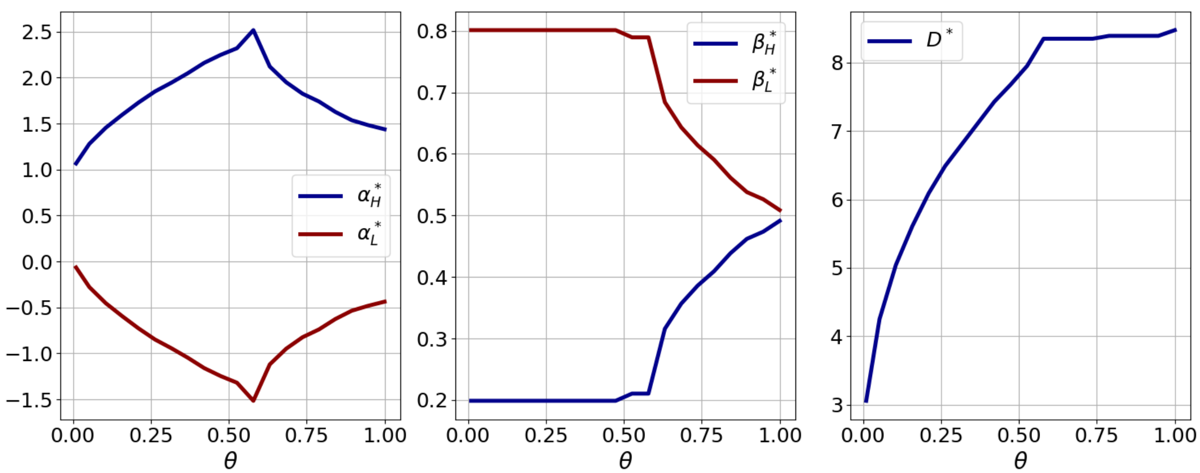

Figure 5: Sensitivity analysis of θ with IR condition.

6 Conclusion

This paper presents a comprehensive theoretical exploration of the optimal design of P2P insurance models in the context of heterogeneous risk. In this context, we have discussed actuarial fairness conditions and embarked on a step-by-step analytical journey to determine optimal design features of the P2P insurance model. We have mathematically formulated and solved an optimization problem, which includes the option to cede part of the aggregate risk to an external insurer. We have examined the trade-off between two key components of participants' costs: the sharing of group insurance policy premiums and the potential refundable deposit. We have provided analytical solutions both with and without the IR constraints, and then shed light on the intricacies of designing such insurance models.

This study has extended its analysis to encompass scenarios in which the insurer imposes an extra profit requirement for undertaking tail risks. These findings advance our understanding of designing P2P insurance systems that balance the interests of heterogeneous participants while maintaining fairness and risk-sharing principles. While this paper adopts the widely-used mean-variance objectives that dominates both academic research and practical applications, extending the results to other preference classes, such as expected utilities or risk measures, remains an important direction for future investigation. Furthermore, although our current analysis treats fairness condition as exogenously determined, future research could productively explore endogenizing it within the optimization framework, thereby extending the theoretical foundations in this study.

References

Abdikerimova, S., Boonen, T. J., and Feng, R. (2024). Multi-period peer-to-peer risk sharing. *Journal of Risk and Insurance*, 91(4):943–982.

Abdikerimova, S. and Feng, R. (2022). Peer-to-peer multi-risk insurance and mutual aid. *European Journal of Operational Research*, 299(2):735–749.

Bäuerle, N. and Glauner, A. (2018). Optimal risk allocation in reinsurance networks. *Insurance: Mathematics and Economics*, 82:37–47.

Boonen, T. J., Tan, K. S., and Zhuang, S. C. (2016). Pricing in reinsurance bargaining with comonotonic additive utility functions. *ASTIN Bulletin: The Journal of the IAA*, 46(2):507–530.

Borch, K. K. (1960). An attempt to determine the optimum amount of stop loss reinsurance. *Transactions of the 16th International Congress of Actuaries*, I:597–610.

Cai, J., Liu, H., and Wang, R. (2017). Pareto-optimal reinsurance arrangements under general model settings. *Insurance: Mathematics and Economics*, 77:24–37.

Charpentier, A., Kouakou, L., Löwe, M., Ratz, P., and Vermet, F. (2022). Collaborative insurance sustainability and network structure. `https://arxiv.org/abs/2107.02764`.

Chen, S., Liu, Y., and Weng, C. (2019). Dynamic risk-sharing game and reinsurance contract design. *Insurance: Mathematics and Economics*, 86:216–231.

Chen, Z. and Feng, R. (2021). A theory of decentralized insurance. *Working paper*.

Chen, Z., Feng, R., Hu, W., and Mao, Y. (2023). Optimal risk pooling of peer-to-peer insurance. *Working paper*.

Chen, Z., Feng, R., Wei, L., and Zhao, J. (2020). Cost-effectiveness, unfairness and adverse selection in mutual aid. *https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3925904*.

Cheung, K. C., Yam, S. C. P., and Zhang, Y. (2019). Risk-adjusted bowley reinsurance under distorted probabilities. *Insurance: Mathematics and Economics*, 86:64–72.

Chi, Y., Lin, X. S., and Tan, K. S. (2017). Optimal reinsurance under the risk-adjusted value of an insurer's liability and an economic reinsurance premium principle. *North American Actuarial Journal*, 21(3):417–432.

Clemente, G. P. and Marano, P. (2020). The broker model for peer-to-peer insurance: An analysis of its value. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 45:457–481.

Denuit, M. (2019). Size-biased transform and conditional mean risk sharing, with application to p2p insurance and tontines. *ASTIN Bulletin*, 49(3):591–617.

Denuit, M. and Dhaene, J. (2012). Convex order and comonotonic conditional mean risk sharing. *Insurance: Mathematics and Economics*, 51(2):265–270.

Denuit, M., Dhaene, J., and Robert, C. Y. (2021). Risk-sharing rules and their properties, with applications to peer-to-peer insurance. *Journal of Risk and Insurance*.

Denuit, M., Dhaene, J., and Robert, C. Y. (2022). Risk-sharing rules and their properties, with applications to peer-to-peer insurance. *Journal of Risk and Insurance*, 89(3):615–667.

Denuit, M. and Robert, C. Y. (2021a). Risk sharing under the dominant peer-to-peer property and casualty insurance business models. *Risk Management and Insurance Review*, 24(2):181–205.

Denuit, M. and Robert, C. Y. (2021b). Stop-loss protection for a large p2p insurance pool. *Insurance: Mathematics and Economics*, 100:210–233.

Denuit, M. and Robert, C. Y. (2022). Collaborative insurance with stop-loss protection and team partitioning. *North American Actuarial Journal*, 26(1):143–160.

Feng, R., Liu, C., and Taylor, S. (2023). Peer-to-peer risk sharing with an application to flood risk pooling. *Annals of Operations Research*, 321(1-2):813–842.

Jiao, Z., Kou, S., Liu, Y., and Wang, R. (2023). An axiomatic theory for anonymized risk sharing. *arXiv preprint arXiv: 2208.07533*.

Levantesi, S. and Piscopo, G. (2022). Mutual peer-to-peer insurance: the allocation of risk. *Journal of Co-operative Organization and Management*, 10(1):100154.

Li, D., Rong, X., and Zhao, H. (2014). Optimal reinsurance–investment problem for maximizing the product of the insurer's and the reinsurer's utilities under a cev model. *Journal of Computational and Applied Mathematics*, 255:671–683.

Li, J., Fang, Z., Nie, C., and Chen, S. (2023). Optimal design for network mutual aid. *Probability in the Engineering and Informational Sciences*, 37(2):567–596.

Tang, Q., Tong, Z., and Xun, L. (2022). Portfolio risk analysis of excess of loss reinsurance. *Insurance: Mathematics and Economics*, 102:91–110.

Von Bieberstein, F., Feess, E., Fernando, J. F., Kerzenmacher, F., and Schiller, J. (2019). Moral hazard, risk sharing, and the optimal pool size. *Journal of Risk and Insurance*, 86(2):297–313.

Zhao, Y. and Zeng, Y. (2023). Optimal commissions and subscriptions in mutual aid platforms. *ASTIN Bulletin: The Journal of the IAA*, 53(3):658–683.

A Appendix: Proofs

A.1 Proof for Theorem 1

Considering Problem 1, denote the objective function as $J(\cdot)$. Then we have

$$J(\cdot) = -(1+\theta)\mathbb{E}\left[(X-D)_+\right] - \mathbb{E}\left[\min\{D, X\}\right] - \mathrm{Var}\left[\min\{D, X\}\right]\sum_{k=1}^{n}\gamma_k\beta_k^2.$$

Therefore, solutions to Problem 1 coincide with solutions of the following problem:

$$\min_{\{\beta_1,\ldots,\beta_n\}\in\mathbb{R}^n}\sum_{k=1}^{n}\gamma_k\beta_k^2, \quad s.t. \ \sum_{k=1}^{n}\beta_k = 1.$$

The Lagrangian function is

$$L(\cdot) = \sum_{k=1}^{n}\gamma_k\beta_k^2 - \lambda\left(\sum_{k=1}^{n}\beta_k - 1\right).$$

Setting $\frac{\partial L(\cdot)}{\partial\lambda} = 0$ and $\frac{\partial L(\cdot)}{\partial\beta_k} = 0$ for all $k = 1, \ldots, n$ leads to

$$\beta_k^* = \frac{1}{\gamma_k\sum_{k=1}^{n}\frac{1}{\gamma_k}}.$$

Moreover, Equation (8) implies that:

$$\alpha_k = \frac{\mathbb{E}\left[X_k\right] + \rho_k\theta\mathbb{E}\left[(X-D)_+\right] - \mathbb{E}\left[\min\{D, X\}\right]\beta_k^*}{(1+\theta)\mathbb{E}\left[(X-D)_+\right]}.$$

A.2 Proof for Corollary 2

Differentiating α_k^* with respect to θ yields:

$$\begin{aligned}
\frac{\mathrm{d}\alpha_k^*}{\mathrm{d}\theta} &= \frac{\rho_k\mathbb{E}\left[(X-D)_+\right] - \mathbb{E}\left[X_k\right] + \beta_k^*\mathbb{E}\left[\min\{D, X\}\right]}{(1+\theta)^2\mathbb{E}\left[(X-D)_+\right]} \\
&= \frac{\rho_k\mathbb{E}\left[(X-D)_+\right] - \alpha_k^*(1+\theta)\mathbb{E}[(X-D)_+] + \rho_k\theta\mathbb{E}[(X-D)_+]}{(1+\theta)^2\mathbb{E}\left[(X-D)_+\right]} \\
&= \frac{\rho_k - \alpha_k^*}{1+\theta}.
\end{aligned}$$

It follows that if $\rho_k \geq \alpha_k^*$, α_k^* increases with θ; and if $\rho_k \leq \alpha_k^*$, α_k^* decreases with θ.

A.3 Proof for Corollary 3

It is straightforward to check that:

$$\frac{\mathrm{d}\alpha_k^*}{\mathrm{d}D} = \frac{S_X(D)\left[\mathbb{E}\left[X_k\right] - \beta_k^*\mathbb{E}\left[X\right]\right]}{(1+\theta)\mathbb{E}\left[(X-D)_+\right]}.$$

Thus, if $\mathbb{E}\left[X_k\right] - \beta_k^*\mathbb{E}\left[X\right] \geq 0$, then α_k^* increases with D. If $\mathbb{E}\left[X_k\right] - \beta_k^*\mathbb{E}\left[X\right] \leq 0$, then α_k^* decreases with D.

A.4 Proof for Theorem 2

Problem 2 is equivalent to maximizing the following objective function $J(D)$:

$$J(D) = \sum_{k=1}^{n} \left(\mathbb{E}[-\alpha_k^*\Pi - \beta_k^* D + \beta_k^*(D-X)_+] - \gamma_k \operatorname{Var}[\beta_k^*(D-X)_+]\right),$$

which can be rewritten as:

$$J(D) = -(1+\theta)g_1(D) - g_2(D) - \frac{g_3(D)}{\sum_{k=1}^{n}\frac{1}{\gamma_k}},$$

where the auxiliary functions are defined as:

$$g_1(D) = \mathbb{E}[(X-D)_+] = \int_D^{+\infty} x \,\mathrm{d}F_X(x) - DS_X(D),$$

$$g_2(D) = \mathbb{E}[\min\{D, X\}] = \int_0^D x \,\mathrm{d}F_X(x) + DS_X(D),$$

$$g_3(D) = \operatorname{Var}[\min\{D, X\}] = \int_0^D x^2 \,\mathrm{d}F_X(x) + D^2 S_X(D) - \left(\int_0^D x \,\mathrm{d}F_X(x) + DS_X(D)\right)^2.$$

Next, we compute the first derivative of $J(D)$ with respect to D:

$$\frac{\mathrm{d}J(D)}{\mathrm{d}D} = -(1+\theta)\frac{\mathrm{d}g_1(D)}{\mathrm{d}D} - \frac{\mathrm{d}g_2(D)}{\mathrm{d}D} - \frac{1}{\sum_{k=1}^{n}\frac{1}{\gamma_k}}\frac{\mathrm{d}g_3(D)}{\mathrm{d}D} = J_1(D)\cdot S_X(D),$$

where

$$J_1(D) = \theta - 2\left(DF_X(D) - \int_0^D x \,\mathrm{d}F_X(x)\right)\cdot\left(\frac{1}{\sum_{k=1}^{n}\frac{1}{\gamma_k}}\right) = \theta - \frac{2}{\sum_{k=1}^{n}\frac{1}{\gamma_k}}\int_0^D F_X(s)\,\mathrm{d}s.$$

Since $S_X(D) \geq 0$, the sign of $\frac{\mathrm{d}J(D)}{\mathrm{d}D}$ depends on the sign of $J_1(D)$. Therefore, we analyze the properties of $J_1(D)$. Taking the derivative of $J_1(D)$ with respect to D, we get:

$$\frac{\mathrm{d}J_1(D)}{\mathrm{d}D} = -\frac{2F_X(D)}{\sum_{k=1}^{n}\frac{1}{\gamma_k}} \leq 0,$$

which implies that $J_1(D)$ is decreasing in D. Consequently, we obtain in $(0, M)$:

$$J_1(D) < \sup J_1(D) = J_1(0) = \theta,$$
$$J_1(D) > \inf J_1(D) = J_1(M) = \theta - \frac{2}{\sum_{k=1}^{n} \frac{1}{\gamma_k}} (M - \mathbb{E}[X]).$$

We now have two cases based on the value of θ:

(1) If $\theta \geq \frac{2}{\sum_{k=1}^{n} \frac{1}{\gamma_k}} (M - \mathbb{E}[X])$, then $J_1(D) > 0$ for all D, implying that $\frac{\mathrm{d}J(D)}{\mathrm{d}D} > 0$. In this case, there are no maximums that exist in $(0, M)$.

(2) If $\theta < \frac{2}{\sum_{k=1}^{n} \frac{1}{\gamma_k}} (\sup(X) - \mathbb{E}[X])$, then there uniquely exists $\hat{D} \in (0, M)$ such that $\frac{\mathrm{d}J(D)}{\mathrm{d}D} > 0$ for $D \in (0, \hat{D})$, and $\frac{\mathrm{d}J(D)}{\mathrm{d}D} < 0$ for $D \in (\hat{D}, M)$. In this case, $\hat{D}$ is the unique solution to $J_1(D) = 0$, i.e.,

$$\frac{\theta}{2} \sum_{k=1}^{n} \frac{1}{\gamma_k} = \int_0^{\hat{D}} F_X(s) \, \mathrm{d}s.$$

Thus, the optimal deposit is given by: $D^* = \hat{D}$.

A.5 Proof for Theorem 3

Problem 3 is equivalent to the following formulation:

$$\min_{\{\beta_1, \dots, \beta_n\} \in \mathbb{R}^n} \mathrm{Var}[(D - X)_+] \sum_{k=1}^{n} \gamma_k \beta_k^2, \tag{30}$$

$$s.t. \quad \sum_{k=1}^{n} \beta_k = 1, \tag{31}$$

$$\rho_k \leq \frac{\gamma_k \mathrm{Var}[X_k] - \gamma_k \beta_k^2 \mathrm{Var}[\min\{X, D\}]}{\theta \mathbb{E}[(X - D)_+]}, \quad \forall k \in \{1, ..., n\}. \tag{32}$$

First, we consider the case where the constraint (32) is inactive for all $k \in \{1, \dots, n\}$. In this case, we solve the unconstrained version of the problem. In this case, the Lagrangian function is:

$$L = \mathrm{Var}[(D - X)_+] \sum_{k=1}^{n} \gamma_k \beta_k^2 + \lambda \left(\sum_{k=1}^{n} \beta_k - 1 \right).$$

Setting $\frac{\partial L}{\partial \lambda} = 0$ and $\frac{\partial L}{\partial \beta_k} = 0$ for all k, we derive:

$$\beta_k^* = \frac{1}{\gamma_k \sum_{k=1}^{n} \frac{1}{\gamma_k}}.$$

Since we assumed that (32) is inactive, we must ensure that, for all $k \in \{1, \ldots, n\}$,

$$\rho_k \leq \frac{\gamma_k \operatorname{Var}[X_k] - \frac{1}{\gamma_k(\sum_{k=1}^n \frac{1}{\gamma_k})^2} \operatorname{Var}[\min\{X, D\}]}{\theta \mathbb{E}[(X-D)_+]}.$$

This is holds true by (20), which concludes the proof for the case that $\mathcal{J}$ is empty.

Second, define two nonempty index sets $\mathcal{K}$ and $\mathcal{J}$ such that $\mathcal{K} \cap \mathcal{J} = \emptyset$, $\mathcal{K} \cup \mathcal{J} = \{1, \ldots, n\}$. Suppose that for all $k \in \mathcal{J}$ the constraint (17) is active , i.e.

$$\rho_k > \frac{\gamma_k \operatorname{Var}[X_k] - \frac{1}{\gamma_k(\sum_{k=1}^n \frac{1}{\gamma_k})^2} \operatorname{Var}[\min\{X, D\}]}{\theta \mathbb{E}[(X-D)_+]}.$$

This implies that all β_k^* for $k \in \mathcal{J}$ lies on the boundary. Meanwhile, we assume that for all $k \in \mathcal{K}$, the constraint (17) is inactive. Thus, the optimization problem becomes:

$$\begin{aligned}
&\min_{\{\beta_1, \ldots, \beta_n\} \in \mathbb{R}^n} \operatorname{Var}[(D-X)_+] \sum_{k=1}^n \gamma_k \beta_k^2, \\
s.t. \quad &\sum_{k \in \mathcal{K}} \beta_k + \sum_{k \in \mathcal{J}} \beta_k = 1, \\
&\rho_k = \frac{\gamma_k \operatorname{Var}[X_k] - \gamma_k \beta_k^2 \operatorname{Var}[\min\{X, D\}]}{\theta \mathbb{E}[(X-D)_+]}, k \in \mathcal{J}.
\end{aligned}$$

The Lagrangian for this problem is:

$$\begin{aligned}
L = &\operatorname{Var}[(D-X)_+] \left(\sum_{k \in \mathcal{K}} \gamma_k \beta_k^2 + \sum_{k \in \mathcal{J}} \frac{\gamma_k \operatorname{Var}[X_k] - \rho_k \theta \mathbb{E}[(X-D)_+]}{\gamma_k \operatorname{Var}[\min\{X, D\}]} \right) \\
&+ \lambda \left(\sum_{k \in \mathcal{K}} \beta_k + \sum_{k \in \mathcal{J}} \beta_k - 1 \right).
\end{aligned}$$

Setting $\frac{\partial L}{\partial \lambda} = 0$ and $\frac{\partial L}{\partial \beta_k} = 0$ for $k \in \mathcal{K}$, we can obtain:

$$\beta_k = \frac{1 - \sum_{k \in \mathcal{J}} \beta_k}{\gamma_k \sum_{k \in \mathcal{K}} \frac{1}{\gamma_k}}.$$

For $k \in \mathcal{J}$, from the binding constraint we have:

$$\frac{\gamma_k \operatorname{Var}[X_k] - \frac{1}{\gamma_k(\sum_{k=1}^n \frac{1}{\gamma_k})^2} \operatorname{Var}[\min\{X, D\}]}{\theta \mathbb{E}[(X-D)_+]} < \frac{\gamma_k \operatorname{Var}[X_k] - \gamma_k \beta_k^2 \operatorname{Var}[\min\{X, D\}]}{\theta \mathbb{E}[(X-D)_+]},$$

which leads to:

$$-1 < -\frac{\sum_{k \in \mathcal{J}} \frac{1}{\gamma_k}}{\sum_{k=1}^n \frac{1}{\gamma_k}} < \sum_{k \in \mathcal{J}} \beta_k < \frac{\sum_{k \in \mathcal{J}} \frac{1}{\gamma_k}}{\sum_{k=1}^n \frac{1}{\gamma_k}} < 1.$$

Thus, it holds for $k \in \mathcal{K}$ that $\beta_k > 0$. For $k \in \mathcal{J}$, from the constraints we can obtain:

$$\beta_k = \pm\sqrt{\frac{\gamma_k \operatorname{Var}[X_k] - \rho_k \theta \mathbb{E}[(X-D)_+]}{\gamma_k \operatorname{Var}[\min\{X, D\}]}}.$$

Now, substitute β_k, $k \in \mathcal{K}$, into the Lagrange function L, we get

$$L = \operatorname{Var}[(D-X)_+] \left(\sum_{k \in \mathcal{K}} \gamma_k \left(\frac{1 - \sum_{k \in \mathcal{J}} \beta_k}{\gamma_k \sum_{k \in \mathcal{K}} \frac{1}{\gamma_k}} \right)^2 + \sum_{k \in \mathcal{J}} \frac{\gamma_k \operatorname{Var}[X_k] - \rho_k \theta \mathbb{E}[(X-D)_+]}{\operatorname{Var}[\min\{X, D\}]} \right).$$

As we have already shown, $1 - \sum_{k \in \mathcal{J}} \beta_k > 0$. Thus, to minimize L, we should maximize $\sum_{k \in \mathcal{J}} \beta_k$, which implies that the optimal β_k for $k \in \mathcal{J}$ should be positive, i.e. ,

$$\beta_k^* = \sqrt{\frac{\gamma_k \operatorname{Var}[X_k] - \rho_k \theta \mathbb{E}[(X-D)_+]}{\gamma_k \operatorname{Var}[\min\{X, D\}]}}.$$

Together with the inequality $\sum_{k \in \mathcal{J}} \beta_k < 1$, we get that for all $k \in \mathcal{J}$, $0 < \beta_k < 1$, and hence $0 < \sum_{k \in \mathcal{J}} \beta_k < 1$. Therefore, for $k \in \mathcal{K}$, we also have $0 < \beta_k < 1$.

Additionally, for β_k^* to be real and exist for all k, we must require:

$$\rho_k \leq \frac{\gamma_k \operatorname{Var}[X_k]}{\theta \mathbb{E}[(X-D)_+]}, \ \forall k \in \{1, \dots, n\}.$$

Moreover, since

$$1 - \sum_{k \in \mathcal{J}} \beta_k > 1 - \frac{\sum_{k \in \mathcal{J}} \frac{1}{\gamma_k}}{\sum_{k=1}^n \frac{1}{\gamma_k}} = \frac{\sum_{k \in \mathcal{K}} \frac{1}{\gamma_k}}{\sum_{k=1}^n \frac{1}{\gamma_k}},$$

we get the lower bound for β_k, $k \in \mathcal{K}$, as:

$$\beta_k = \frac{1 - \sum_{j \in \mathcal{J}} \beta_j}{\gamma_k \sum_{j \in \mathcal{K}} \frac{1}{\gamma_j}} > \frac{1}{\gamma_k} \cdot \frac{1}{\sum_{j=1}^n \frac{1}{\gamma_j}}.$$

This implies that in this case, compared with the unconstrained case, β_k^* for $k \in \mathcal{K}$ become larger. To ensure that constraint (17) remains inactive for $k \in \mathcal{K}$, we must have:

$$\rho_k \leq \frac{\gamma_k \operatorname{Var}[X_k] - \frac{(1 - \sum_{k \in \mathcal{J}} \beta_k^*)^2}{\gamma_k \left(\sum_{k \in \mathcal{K}} \frac{1}{\gamma_k} \right)^2} \operatorname{Var}[\min\{X, D\}]}{\theta \mathbb{E}[(X-D)_+]}.$$

So far, (1) has been proved.

Finally, for (2), if $\mathcal{J} = \{1, \dots, n\}$, then for all $k \in \{1, \dots, n\}$, β_k^* must satisfy the binding constraint. However, it is straightforward to verify that in this case $\sum_{k=1}^n \beta_k^* < 1$, which violates the feasibility condition. Thus, no solution exists under this configuration.

A.6 Proof for Lemma 1

In the beginning, condition (24) is equivalent to $\rho_k \leq \phi_k(D)$ holds for all $k \in \{1, \ldots, n\}$. For each participant k, let $\tilde{\mathcal{D}}_k$ denote the feasible region of D for which the constraint (24) is satisfied. It then follows that:

$$\mathcal{D}_1 = \bigcap_{k=1}^{n} \tilde{\mathcal{D}}_k.$$

To analyze the structure of $\tilde{\mathcal{D}}_k$, we examine the behavior of $\phi_k(D)$. First, note that:

$$\frac{\mathrm{d}\phi_k(D)}{\mathrm{d}D} = \frac{\gamma_k S_X(D)\psi_k(D)}{\theta\,\mathbb{E}[(X-D)_+]^2},$$

and

$$\frac{\mathrm{d}\psi_k(D)}{\mathrm{d}D} = -\frac{2F_X(D)\,\mathbb{E}[(X-D)_+]}{\gamma_k^2\left(\sum_{j=1}^{n}\frac{1}{\gamma_j}\right)^2} < 0.$$

Hence, $\psi_k(D)$ is decreasing in D. Furthermore, the range of $\psi_k(D)$ satisfies:

$$\sup_D \psi_k(D) = \psi_k(0) = \mathrm{Var}[X_k], \quad \inf_D \psi_k(D) = \psi_k(M) = \mathrm{Var}[X_k] - \frac{\mathrm{Var}[X]}{\gamma_k^2\left(\sum_{j=1}^{n}\frac{1}{\gamma_j}\right)^2}.$$

Now consider two cases: First, if

$$\mathrm{Var}[X_k] - \frac{\mathrm{Var}[X]}{\gamma_k^2\left(\sum_{j=1}^{n}\frac{1}{\gamma_j}\right)^2} \geq 0,$$

then $\psi_k(D) > 0$ for all $D \in (0, M)$, and hence $\phi_k(D)$ is increasing. In this case, we have:

$$\inf_D \phi_k(D) = \phi_k(0) = \frac{\gamma_k\,\mathrm{Var}[X_k]}{\theta\,\mathbb{E}[X]}, \quad \sup_D \phi_k(D) = \phi_k(M) = +\infty.$$

According to the monotonicity of $\phi_k(D)$, (1) in Lemma 1 is shown. Second, if

$$\mathrm{Var}[X_k] - \frac{\mathrm{Var}[X]}{\gamma_k^2\left(\sum_{j=1}^{n}\frac{1}{\gamma_j}\right)^2} < 0,$$

then $\psi_k(D)$ is initially positive but eventually becomes negative as D increases. This implies that $\phi_k(D)$ is increasing over a certain interval and decreasing afterward. According to the monotonicity of $\phi_k(D)$, (2) in Lemma 1 is shown.

A.7 Proof for Theorem 4

By the proof of Theorem 2, we get that the objective function $J(\cdot)$ satisfies $\frac{\mathrm{d}J(D)}{\mathrm{d}D} = J_1(D)S_X(D)$, where

$$J_1(D) = \theta - \frac{2}{\sum_{k=1}^{n} \frac{1}{\gamma_k}} \int_0^D F_X(s)\mathrm{d}s,$$

and $\frac{\mathrm{d}J_1(D)}{\mathrm{d}D} < 0$. Thus, we get

$$\sup_D J_1(D) = J_1(\underline{D}_1) = \theta - \frac{2}{\sum_{k=1}^{n} \frac{1}{\gamma_k}} \int_0^{\underline{D}_1} F_X(s)\mathrm{d}s.$$

$$\inf_D J_1(D) = J_1(\overline{D}_1) = \theta - \frac{2}{\sum_{k=1}^{n} \frac{1}{\gamma_k}} \int_0^{\overline{D}_1} F_X(s)\mathrm{d}s.$$

Then, we have: When

$$\theta - \frac{2}{\sum_{k=1}^{n} \frac{1}{\gamma_k}} \int_0^{\underline{D}_1} F_X(s)\mathrm{d}s \leq 0,$$

then $J_1(D) < 0$ which implies $\frac{\mathrm{d}J(D)}{\mathrm{d}D} < 0$. In this case, we have $D_1^* = \underline{D}_1$. When

$$\theta - \frac{2}{\sum_{k=1}^{n} \frac{1}{\gamma_k}} \int_0^{\overline{D}_1} F_X(s)\mathrm{d}s \geq 0,$$

then $J_1(D) > 0$ which implies $\frac{\mathrm{d}J(D)}{\mathrm{d}D} > 0$. In this case, we have $D_1^* = \overline{D}_1$. When

$$\theta - \frac{2}{\sum_{k=1}^{n} \frac{1}{\gamma_k}} \int_0^{\underline{D}_1} F_X(s)\mathrm{d}s > 0, \text{ and } \theta - \frac{2}{\sum_{k=1}^{n} \frac{1}{\gamma_k}} \int_0^{\overline{D}_1} F_X(s)\mathrm{d}s < 0.$$

then there uniquely exists a $D^o > 0$ such that $\frac{\mathrm{d}J(D)}{\mathrm{d}D} > 0$ if $D \in (\underline{D}_1, D^o)$, and $\frac{\mathrm{d}J(D)}{\mathrm{d}D} < 0$ if $D \in \left(D^o, \overline{D}_1\right]$. Here, D^o is the unique solution of $J_1(D) = 0$, i.e.,

$$\frac{\theta}{2} \sum_{k=1}^{n} \frac{1}{\gamma_k} = \int_0^{D^o} F_X(s)\mathrm{d}s.$$

Therefore, in this case we can have:

$$D_1^* = \underline{D}_1 \cdot \mathbb{I}[D^o < \underline{D}_1] + D^o \cdot \mathbb{I}[\underline{D}_1 \leq D^o \leq \overline{D}_1] + \overline{D}_1 \cdot \mathbb{I}[D^o > \overline{D}_1].$$

A.8 Proof for Lemma 2

Note that $\tilde{\phi}_k(D)$ is increasing in D, and satisfies $\tilde{\phi}_k(D) \leq \phi_k(D)$, with equality if and only if $D \to 0$. Given these properties, the proof of Lemma 2 is similar to that of Lemma 1, by applying the same monotonicity arguments and structural reasoning.

A.9 Proof for Lemma 3

Consider (20) in Theorem 3, which imposes the condition $\rho_k \leq \eta_k(D)$ for all $k \in \mathcal{K}_t$. For each participant k, let $\tilde{\mathcal{D}}_k$ denote the feasible region of D that satisfies constraint (20). It then follows that:

$$\tilde{\mathcal{D}}_{\mathcal{K}_t} = \bigcap_{k \in \mathcal{K}_t} \tilde{\mathcal{D}}_k,$$

To analyze the structure of $\tilde{\mathcal{D}}_k$, we examine the behavior of $\eta_k(D)$. First, note that:

$$\begin{aligned}\frac{\mathrm{d}\eta_k(D)}{\mathrm{d}D} &= \frac{h(D)}{\theta\mathbb{E}[(X-D)_+]^2},\\ \frac{\mathrm{d}^2\eta_k(D)}{\mathrm{d}D^2} &= \frac{\frac{\mathrm{d}h(D)}{\mathrm{d}D}\mathbb{E}[(X-D)_+]^2 + 2h(D)\mathbb{E}[(X-D)_+]S_X(D)}{\theta\mathbb{E}[(X-D)_+]^4}.\end{aligned}$$

In the following, we will prove that when $\frac{\mathrm{d}\eta_k(D)}{\mathrm{d}D} = 0$, we have $\frac{\mathrm{d}^2\eta_k(D)}{\mathrm{d}D^2} < 0$. First, define a tool function $g(D)$ as:

$$g(D) = \sqrt{\mathrm{Var}[\min\{X, D\}]} - \sum_{k \in \mathcal{J}} \sqrt{\mathrm{Var}[X_k] - \frac{\rho_k}{\gamma_k}\theta\mathbb{E}[(X-D)_+]}.$$

It holds that $\kappa_k(D) = g(D)^2$, and we have

$$\frac{\mathrm{d}g(D)}{\mathrm{d}D} = S_X(D)l(D),$$

where

$$l(D) = \left(\frac{D - \mathbb{E}[\min\{X, D\}]}{\sqrt{\mathrm{Var}[\min\{X, D\}]}} - \sum_{k \in \mathcal{J}} \frac{\rho_k\theta}{2\gamma_k\sqrt{\mathrm{Var}[X_k] - \frac{\rho_k}{\gamma_k}\theta\mathbb{E}[(X-D)_+]}}\right)$$

Now, we investigate the monotonicity of $l(D)$. Apparently, the second item of $l(D)$ is increasing with respect to D. We now examine the monotonicity of the first item of $l(D)$, which we denote it as $l_1(D)$. We have

$$\frac{\mathrm{d}l_1(D)}{\mathrm{d}D} = \frac{m(D)}{\mathrm{Var}[\min\{X, D\}]^{5/2}},$$

where

$$m(D) = F_X(D)\,\mathrm{Var}[\min\{X, D\}] - S_X(D)(D - \mathbb{E}[\min\{X, D\}])^2.$$

We have

$$\frac{\mathrm{d}m(D)}{\mathrm{d}D} = f_X(D)\,\mathrm{Var}[\min\{X, D\}] + f_X(D)(D - \mathbb{E}[\min\{X, D\}]) > 0,$$

which implies that $m(D) > \inf m(D) = m(0) = 0$. Thus $m(D) > 0$, $\frac{\mathrm{d}l_1(D)}{\mathrm{d}D} > 0$, and then $l(D)$ is increasing with D. When $\frac{\mathrm{d}\eta_k(D)}{\mathrm{d}D} = 0$, we have $h(D) = 0$, which further implies:

$$\frac{\mathrm{d}g(D)}{\mathrm{d}D} = \frac{S_X(D)\left(\gamma_k \operatorname{Var}[X_k] - \frac{g(D)^2}{\gamma_k\left(\sum_{k\in\mathcal{K}}\frac{1}{\gamma_k}\right)^2}\right)}{2\mathbb{E}[(X-D)_+]\frac{g(D)}{\gamma_k\left(\sum_{k\in\mathcal{K}}\frac{1}{\gamma_k}\right)^2}}.$$

When $h(D) = 0$, we have

$$\frac{\mathrm{d}^2\eta_k(D)}{\mathrm{d}D^2} = \frac{\frac{\mathrm{d}h(D)}{\mathrm{d}D}}{\theta\mathbb{E}[(X-D)_+]^2},$$

where

$$\frac{\mathrm{d}h(D)}{\mathrm{d}D} = -\frac{2\mathbb{E}[(X-D)_+]}{\gamma_k\left(\sum_{k\in\mathcal{K}}\frac{1}{\gamma_k}\right)}\left[\left(\frac{\mathrm{d}g(D)}{\mathrm{d}D}\right)^2 - g(D)S_X(D)\frac{\mathrm{d}l(D)}{\mathrm{d}D}\right] < 0.$$

Thus we proved that when $\frac{\mathrm{d}\eta_k(D)}{\mathrm{d}D} = 0$, $\frac{\mathrm{d}^2\eta_k(D)}{\mathrm{d}D^2} < 0$. This shows that $\eta_k(D)$ is either monotonically increasing, monotonically decreasing, or first increasing and then decreasing. This is because all points where the first-order derivative is zero will be maximum points. It is easy to verify that the lemma follows from the monotonicity of $\eta_k(D)$.

A.10 Proof for Theorem 5

Denote $J(D)$ be the objective function, we have

$$\begin{aligned}
\frac{\mathrm{d}J(D)}{\mathrm{d}D} &= n\left(\sum_{k\in\mathcal{K}}\rho_k\theta S_X(D) - \frac{\frac{\mathrm{d}\kappa(D)}{\mathrm{d}D}}{\sum_{k\in\mathcal{K}}\frac{1}{\gamma_k}}\right),\\
\frac{\mathrm{d}^2J(D)}{\mathrm{d}D^2} &= n\left(-\sum_{k\in\mathcal{K}}\rho_k\theta f_X(D) - \frac{\frac{\mathrm{d}^2\kappa(D)}{\mathrm{d}D^2}}{\sum_{k\in\mathcal{K}}\frac{1}{\gamma_k}}\right).
\end{aligned}$$

In the following, we will prove that when $\frac{\mathrm{d}J(D)}{\mathrm{d}D} = 0$, we have $\frac{\mathrm{d}^2J(D)}{\mathrm{d}D^2} < 0$. Recall the tool function $g(D)$ we defined in the proof of Lemma 3; when $\frac{\mathrm{d}J(D)}{\mathrm{d}D} = 0$, we have

$$g(D) = \frac{\sum_{k\in\mathcal{K}}\rho_k\theta S_X(D)\sum_{k\in\mathcal{K}}\frac{1}{\gamma_k}}{2\frac{\mathrm{d}g(D)}{\mathrm{d}D}},$$

then we have

$$\frac{\mathrm{d}^2J(D)}{\mathrm{d}D^2} = -\frac{2}{\sum_{k\in\mathcal{K}}\frac{1}{\gamma_k}}\left[\left(\frac{\mathrm{d}g(D)}{\mathrm{d}D}\right)^2 + S_X(D)g(D)\frac{\mathrm{d}l(D)}{\mathrm{d}D}\right] < 0.$$

Thus, whenever $\frac{dJ(D)}{dD} = 0$, it follows that $\frac{d^2J(D)}{dD^2} < 0$, indicating that such points correspond to local maxima. As in the discussion of $\eta_k(D)$ in Lemma 3, the function $J(D)$ may exhibit monotonic behavior—either increasing or decreasing—across different regions. However, any point at which the first derivative equals to zero is guaranteed to be a maximum. Therefore, by exploiting the monotonicity of $J(D)$, the result follows.