A Flexible Group Benefits Framework for Pricing Deposit Rates

by

Cherie Ng

B.Sc., University of British Columbia, 2009

Project Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Statistics and Actuarial Science
Faculty of Science

SIMON FRASER UNIVERSITY
Fall 2020

Declaration of Committee

Name: **Cherie Ng**

Degree: **Master of Science**

Thesis title: **A Flexible Group Benefits Framework for Pricing Deposit Rates**

Committee: **Chair:** Joan Hu
Professor, Statistics and Actuarial Science

Jean-François Bégin
Co-Supervisor
Assistant Professor, Statistics and Actuarial Science

Barbara Sanders
Co-Supervisor
Associate Professor, Statistics and Actuarial Science

Yi Lu
Examiner
Professor, Statistics and Actuarial Science

Abstract

Currently, most flexible group benefit plans are designed and priced based on deterministic assumptions about the plan members' option selections. This can cause the adverse selection spiral, threatening the sustainability of the plan. We therefore propose a comprehensive framework with a novel pricing formula that incorporates both a model for claims and a model for plan members' enrollment decisions to prevent adverse selection. We find through simulation that our proposed pricing formula outperforms the traditional pricing practice by keeping flex plans sustainable over time. In addition to preventing the adverse selection spiral through pricing, our framework also serves as a tool to evaluate the impact of other parameters such as changes in plan designs, health costs, and member decision.

Keywords: group insurance; flexible benefits; stochastic modelling; adverse selection

Dedication

To my beloved family.

Acknowledgements

I want to thank my senior supervisor, Dr. Jean-François Bégin for his professional guidance throughout this project. JF, I have benefitted significantly from your expertise. You always reply to my e-mails promptly, and always make yourself available whenever I want to discuss the project. Thank you for investing your precious time in me.

My appreciation also goes to Professor Barbara Sanders. Barbara, your perspective and attention to detail substantially enriched my research experience, and your valuable input greatly enhanced this project. I am grateful to have you as my co-supervisor.

I also express my gratitude to Dr. Joan Hu and Dr. Yi Lu for agreeing to be on my examining committee. To the engaging and knowledgeable professors, Dr. Richard Lockhart, Dr. Gary Parker, Dr. Rachel Altman, and again Yi, Barbara, and JF, thank you for offering the interesting courses I took during my MSc program.

Of course, this research project would not have existed if it was not for the series of coincidental events that led me to step away from my career in the industry to pursue full time graduate studies. I must acknowledge the Mathematics department's graduate secretary, Christie, for suggesting that I check out the Statistics and Actuarial Science department down the hall, Rachel for immediately introducing me to JF, and subsequently the other two very welcoming professors, Barbara and Yi, who spent the time to chat with a random passerby. What are the odds that these busy professors all happened to be sitting in their office, waiting for the impromptu visitor on a random Thursday afternoon? In statistical speak, I rejected the null hypothesis that these events just occurred by random chance. To me, this was a clear sign that God had planned this path, and studying at the Simon Fraser University has proven to be much more joyful and fulfilling than I could have planned myself.

Finally, I wish to recognize and thank my loving family members. My completion of this thesis and the program certainly would not have been possible without their unconditional support.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Group benefits are health coverages offered by an organization to its members. Since the organization (referred to as the plan sponsor) is usually an employer, group benefit plans are sometimes known as employee benefits. Members within the same plan typically receive the same coverage, and the lines of benefits (i.e., coverage categories) often include basic life, accidental death and dismemberment, long term disability, extended healthcare, and dental care.[1] If the plan is insured through a contract with a carrier, then the fee paid by members in return for the coverage is referred to as the premium; otherwise, it is known as the deposit if the plan sponsor self-insures the plan. In either case, the rates are generally the same across the entire group and updated annually, and plan sponsors commonly subsidize this cost.

A flexible group benefit plan (flex plan) is a special type of group benefit plan that offers plan members the flexibility to select their coverage level annually among a set of predefined options and deposit rates. Compared to traditional plans, a flex plan has the advantage of increasing employee satisfaction (Barber et al., 1992), especially as the general workforce becomes increasingly diverse (French, 2008), while potentially giving the employer more control over their annual costs by limiting the deposit subsidy to a fixed amount (Skwire, 2016). In fact, flex plans have been gaining popularity since they were first introduced in Canada in the mid-1980s. By the early 21st century, almost half of major Canadian employers offered flexible benefits to their employees (McKay, 2007). These plans are usually self-insured by the plan sponsor through a claims administration agreement with a carrier. Under this arrangement, an administrative services only (ASO) account acts as the bank account in which regular deposits are paid into, and claims and administrative expenses are withdrawn from. Occasionally, the plan sponsor will need to make an extra lump sum de-

[1] It is common for plan sponsors to allocate members within a group plan to separate classes (e.g., based on type of employment), where each class is given a different set of coverage. Mathematically, we can treat these classes as separate plans; therefore we do not make the distinction in our report.

posit when the balance drops below a minimum requirement; likewise, they may sometimes withdraw from it when a large surplus accumulates.

Although flex plans can be more appealing than traditional plans, calculating sound deposit rates is a complex actuarial problem. In general, rates are set so that the total premium or deposits cover the total expected claims, administrative expense, allowance for reserves, and sometimes a profit margin. For a traditional group benefit plan, this requires estimating the upcoming year's claims, often based on multiplying the current year's claims by some trend factor.[2] For a flex plan, however, correctly pricing each option also requires accurately predicting the mix of employees choosing the particular options, but their choices depend on the rates offered. With this complexity, there is currently no coherent framework for calculating flex plan deposit rates. A common solution used in practice is to apply the traditional group benefits model to each option, implying that members remain in the same option over time.

Unfortunately, this simplifying assumption can be insufficient in pricing sound deposit rates, which can cause the so-called adverse selection spiral, thereby threatening the sustainability of the plan.[3] Adverse selection is a situation in which the consumer has more information about their risk than the insurer does, resulting in a product being sold to consumers of a higher risk group than intended. Specifically, in the flexible benefits context, adverse selection not only requires that members predict the distribution of their health expenses more accurately than the actuary does, but also that they rationally select an option based on this prediction (Marquis, 1992).[4] When this occurs, the actual claims are higher than the actuary's estimate, which drives an increase in the subsequent year's deposit rate. This forces healthier employees to switch to lower coverage options, thereby increasing the concentration of high-claim employees in a given option. The result is a further increase in per capita or per volume claims in the subsequent year.[5] This cycle of a deficit in the deposit account, skyrocketing rates, and decreased enrollment eventually spirals until the highest coverage option quickly becomes unaffordable and obsolete. A classic example of the adverse

[2]In group benefits, trend refers to inflationary and utilization changes (e.g., change in per visit cost and and frequency of visits for a paramedical service).

[3]An adverse selection spiral occurs when healthier insureds drop out of a plan or coverage option, increasing the concentration of high risk insureds covered and thereby causing skyrocketing insurance rates.

[4]The second condition is not specified in a general setting, where consumers are actively seeking out insurance. In group benefits, however, members are usually in a plan as a result of group membership such as employment. In this case, there are a number of reasons why members may not rationally select an option, including missing the enrollment window or a lack of understanding of the plan coverage.

[5]Rates may either be priced on a per capita or per volume basis. Volume may refer to, say, $1,000 of insurance coverage.

selection spiral observed in practice is the Harvard University employee flex plan (Cutler and Zeckhauser, 1998), which drove out the highest coverage option in only three years.

Although a number of strategies may be used in practice to prevent adverse selection, they are mostly related to restricting choice. For example, the step-up/step-down restriction limits plan members to changing coverage by only one level between any two years. A two-year premium plan lock in requires plan members in the highest coverage option to stay in the plan for a minimum of two years instead of one. A modular plan bundles different lines of benefits into one option, thereby limiting the plan members' flexibility of tailoring the plan to suit their exact needs.[6] Changing the premium subsidy arrangement from a fixed amount to a level percentage can also be considered a strategy for reducing adverse selection (Marquis and Buchanan, 1999).

The strategies mentioned above have the obvious shortcoming of either limiting flexibility for plan members or restricting budget control for plan sponsors, reducing the appeal of flex plans. An alternative way to control adverse selection is through premium pricing. In fact, various studies such as Marquis (1992) and van Winssen et al. (2018) have shown that "individual risk rating," in which each insured is given a customized rate based on their own risk, is an effective method for eliminating adverse selection. Unfortunately, this strategy is not applicable to group benefits, where the same premium or deposit rate applies to the whole group. Another premium pricing strategy involves applying cross subsidy between options, which Cave (1985) has shown to be a related to an equilibrium in premium rates and enrollment over time. However, the study has not been adapted to the flex plan context to suggest a clear premium pricing formula.

Our solution is to construct a comprehensive flexible group benefits framework and to use it to develop a pricing formula that is immune to adverse selection. While various models have been separately proposed for selected components of flex plans, none of them represent a complete framework. For example, Fuhrer and Shapiro (1992) proposed a model for flex benefits in the cross-section, but they have not analyzed the dynamic aspect of the problem; that is, how the plans evolve from one year to another. Mehta et al. (2017) proposed a dynamic model for option selection and health care consumption, but their model has not been combined with other aspects of a flex plan into a comprehensive framework. A wide range of enrollment decision models and observations have been presented in various papers (e.g., Marquis and Buchanan, 1999; Marquis, 1992; Carlin and Town, 2009; Sturman and Boudreau, 1994; Maciejewski et al., 2004; Bajari et al., 2006), but there is no consensus on

[6] Refer to McKay (2007), O'Neill (2012), and Skwire (2016) for more details on adverse selection prevention strategies used in practice.

a standard. Our proposed framework combines some of the ideas presented in those papers, incorporating both a model for claims and a model for plan members' enrollment decisions.

This comprehensive framework allows us to build a pricing formula that anticipates both the claims and the plan members' decisions so that we can prevent adverse selection. Blending the interest of plan sponsors and members with the actuarial equivalence principle of premium calculations, we define the following properties for an ideal set of flex plan deposit rates:

1. At least one member is enrolled in each option at any time,
2. The expected year-end ASO account balance should be as close to zero as possible.
3. The ASO account balance should be kept as stable as possible to avoid the need to replenish a deficit or withdraw a large surplus, and
4. The deposit rates should be kept as stable as possible from year to year.

The first goal ensures that the pricing does not defeat the purpose of offering options to members. The second goal ensures that the deposits are sufficient to cover expected claims and administrative expenses, but are also not too high. The third and fourth goals are both desirable for the plan sponsor's budgeting purpose. Finally, the fourth goal is clearly beneficial to plan members as well.

The remainder of this report is arranged as follows. Details of this framework can be found in Chapter 2. Chapter 3 outlines our assumptions and methodology used to apply this framework and study the evolution of a hypothetical flex plan over time. Key findings from the simulation study are summarized in Chapter 4. Our conclusion and suggestions for further research are in Chapter 5.

Chapter 2

Proposed Framework

We propose a comprehensive framework composed of three main components corresponding to the three main events of an annual flex plan cycle: pricing, option selection, and claims incurral. These three components are built upon a fourth, structural component which represents the plan design. The next few sections of this chapter outline each of the components.

2.1 Plan Design

2.1.1 The Foundations

We propose a framework for a flexible group benefits plan with the following features:

- K options, each fully defined through the deductible, coinsurance, coinsurance maximum, and overall maximum,
- one line of benefit,
- mandatory enrollment,
- single coverage only (i.e., no family or couple coverage),
- annual flex credits F, paid by the plan sponsor to offset the plan members' deposit payment, where $F \geq 0$, and
- a level percent subsidy s, paid by the plan sponsor to offset the plan members' deposit payments, where $0 \leq s \leq 1$.

We assume for simplicity that during each year t from 1 to T, this plan is offered to N plan members, identified from 1 to N. Note that while simplifying assumptions are made for ease of highlighting the features of our framework, it can be extended to include other common designs and situations, such as multiple lines of benefits, family coverage, and a dynamic group size. Also, although the term "deposit payment" implies the plan is self-insured, the calculations are equivalent to an insured plan with premium payments.

2.1.2 Reimbursement Functions

If the coverage of an option can be fully defined by its deductible, coinsurance, coinsurance maximum and overall maximum, then we can define plan options using the reimbursement functions below, as in Fuhrer and Shapiro (1992). The following is a restatement of their proposed function, adapted to the notation used in this report. Let $r(l,k)$ be the reimbursement for an annual loss (i.e., health expense) of l incurred under option k, where $k \in \{1, 2, ..., K\}$. Then,

$$r(l,k) = \begin{cases} 0 & \text{if } l < d_k \\ c_k(l - d_k) & \text{if } d_k \leq l < \frac{CM_k}{1-c_k} + d_k \\ l - d_k - CM_k & \text{if } \frac{CM_k}{1-c_k} + d_k \leq l < OM_k + d_k + CM_k \\ OM_k & \text{if } OM_k + d_k + CM_k \leq l \end{cases}, \tag{2.1}$$

where

- d_k is the deductible of plan k,
- c_k is the coinsurance rate of plan k,
- CM_k is the coinsurance maximum of plan k, and
- OM_k is the overall maximum of plan k.

Therefore, if a plan member enrolled in option k incurs and submits an annual loss of l to the plan, then $r(l,k)$ is the total claim amount paid from the plan sponsor's ASO account to the plan member for that year. In addition, in our model we explicitly specify that the reimbursement function parameters are chosen such that $r(l,k) \leq r(l,k')$ for all $k < k'$. In other words, the option indices are ordered from the lowest to the highest level of coverage.

2.2 Claims

2.2.1 A Model for Annual Health Status

The annual claim amount for each of the N employees depends on their health status in a given year. Therefore, we introduce a health status model, inspired by Mehta et al. (2017). Consider an individual's health status with respect to the benefit offered in the flex plan.[7] We define this health status as a numerical score in $\mathbb{R}$, where a high value represents good

[7] The benefit coverage may be as specific as paramedical coverage or as general as the aggregate extended health and dental coverage.

health and a low value describes poor health. While this measure evolves continuously over time, we propose a discrete-time version, where $H_{t,j,x}$ denotes the average health status of individual j who is x years old during year t.

This quantity can be broken down into two components: a deterministic predictable portion, $H^{(p)}_{t,x}$, and a stochastic unpredictable portion, $H^{(u)}_{t,j,x}$. We can write

$$H_{t,j,x} = H^{(p)}_{t,x} + H^{(u)}_{t,j,x},$$

where the two components evolve separately.

The predictable portion, $H^{(p)}_{t,x}$, represents the mean health status of all individuals aged x during year t, and is assumed to deteriorate at a rate of λ per year:

$$H^{(p)}_{t+1,x+1} = H^{(p)}_{t,x} - \lambda.$$

The unpredictable portion, $H^{(u)}_{t,j,x}$, represents the hidden factors (e.g., lifestyle and medical treatments) that affect an individual's health status:

$$H^{(u)}_{t+1,j,x+1} = H^{(u)}_{t,j,x} + \epsilon_{t,j}, \tag{2.2}$$

where $\epsilon_{t,j}$ is an independent normal random variable with mean zero and standard deviation σ, i.e., $\mathcal{N}(0, \sigma)$.

Combining the above dynamics, we have

$$H_{t+1,j,x+1} = H_{t,j,x} - \lambda + \epsilon_{t,j} \tag{2.3}$$

with some random variable $H^{(unc)}_{t,j,x_0}$. While birth would be a reasonable candidate for the fixed starting point x_0, in our context a more convenient choice is the starting working age, say 25. Let

$$H^{(unc)}_{t,j,x_0} \sim \mathcal{N}(H^{(p)}_{t,x_0}, \zeta),$$

where ζ is the standard deviation of x_0-year-olds' health statuses. It follows from Equation (2.2) that the unconditional distribution of $H_{t,j,x}$ is

$$H^{(unc)}_{t,j,x} \sim \mathcal{N}\left(H^{(p)}_{t,x}, \sqrt{\zeta^2 + (x - x_0)\sigma^2}\right), \quad x > x_0. \tag{2.4}$$

To alleviate notation, we drop the third index in $H_{t,j,x}$ for the remainder of this report, since an individual's age can be expressed as a function of t and j.

2.2.2 A Model for Annual Loss

Due to random noise, like accuracy of diagnosis or availability of treatment, an individual's health status does not deterministically translate into the loss incurred. Rather, plan member j's aggregate loss throughout year t with respect to the benefit may be modelled as a random variable, $L_{t,j}$, which depends on an unknown mean parameter, $\mu_{t,j} \equiv \mathbb{E}[L_{t,j}]$, and other known parameters.[8] The unknown mean parameter is modelled as a function of the unknown health status:

$$\begin{aligned} \mu_{t,j} &= g_t(H_{t,j}) \\ &= e^{\beta_{t,0}+\beta_{t,1}H_{t,j}}. \end{aligned} \tag{2.5}$$

Note that this link function offers flexibility for pricing purposes—the parameters $\beta_{t,0}$ and $\beta_{t,1}$ can be easily updated to reflect anticipated shocks (such as prescription drug "patent cliffs"), cost inflation, and utilization changes in an upcoming year. The realized loss can then be converted into an annual claim via the reimbursement functions discussed in Section 2.1.2.[9]

2.3 Modelling Plan Members' Option Selection

The plan members' decision-making process typically involves taking time to understand the plan coverage, anticipating their upcoming loss, and evaluating the K coverage options given their risk-aversion levels. Recall that our key motivation for designing the framework is to combat adverse selection, which arises when plan members rationally select their plan options. Therefore, in our framework, we assume all plan members are diligent, meaning they understand the health status dynamics, loss distributions, and make rational plan choices. Specifically, we model their decisions using expected utility theory.

2.3.1 A Model for Plan Members' Intuition of their Upcoming Health Status

Although the true values of health statuses are unknown to all parties, each plan member has an intuition of their upcoming year's health status. Let $\overline{H}_{t,j}$ represent plan member j's intuitive estimate of their health status in year t. We treat this intuition as an unbiased

[8]Note that we can reparameterize distributions so that the unknown parameter is the mean.

[9]We assume that any moral hazard is negligible; otherwise, a plan members' incurred loss may depend on their selected option.

estimate of their true health, where the estimation error follows a normal distribution. In other words, we have

$$\overline{H}_{t,j} = H_{t,j} + \nu_{t,j}, \tag{2.6}$$

and $\nu_{t,j} \sim \mathcal{N}(0, \delta)$.

2.3.2 Plan Members' Anticipated Plan Usage

As mentioned in the introduction, adverse selection occurs when members can accurately predict their plan usage. We therefore model diligent plan members as inherently knowing the conditional distribution of annual loss, given a health status. This means plan member j's projection of their loss in year t is consistent with inherently estimating the mean parameter of their loss distribution as

$$\begin{aligned}\overline{\mu}_{t,j} &= g_t(\overline{H}_{t,j}) \\ &= e^{\beta_{t,0}+\beta_{t,1}\overline{H}_{t,j}}\end{aligned}$$

and then anticipating the plausible range and probabilities of their potential loss using the true annual loss distribution given this estimated mean parameter, $\overline{\mu}_{t,j}$.

2.3.3 Utility Maximization

Choosing between the plan options involves evaluating not just the potential financial outcomes from each option, but also the utility gained from these financial outcomes. As in Fuhrer and Shapiro (1992), we can model the plan members' enrollment decisions by applying the expected utility framework, which assumes that an individual's preference under risk is consistent with choosing the option that maximizes their expected utility. Note that although utility is commonly modelled as a function of an individual's total wealth during a particular time interval, in our framework we consider only the portion of wealth associated with employment. This deviation from the typical model keeps our framework realistic; pricing actuaries would not know or estimate plan members' income from sources outside of employment.

Recall that the plan sponsor may subsidize a portion s of the deposit amount, or offer fixed annual flex credits of F to plan members. We define the annual consumption[10] associated with employment for plan member j enrolled in option k in year t as

[10]We assume that the difference in tax implications between taxable wages, non-taxable reimbursements, and any remaining credits paid out are negligible and can therefore be ignored.

$$w_{t,j} - (1-s)D_{t,k} + F + r(L_{t,j}, k) - L_{t,j},$$

where

- $w_{t,j}$ is the salary of plan member j in year t, and
- $D_{t,k}$ is the deposit rate of option k in year t.

Let $u_{t,j}(\,\cdot\,; R_{t,j})$ be the utility function and $R_{t,j}$ be the risk aversion parameter of plan member j in year t. Then plan member j's utility of consumption in year t is

$$u_{t,j}(w_{t,j} - (1-s)D_{t,k} + F + r(L_{t,j}, k) - L_{t,j}; R_{t,j})$$

given they have enrolled in option k. A diligent plan member will choose an option with the highest expected utility.

Let $U_{t,j}$ be the option which maximizes expected utility for plan member j in year t, given a vector of annual deposit rates, $\boldsymbol{D}_t = [D_{t,1}, D_{t,2}, ..., D_{t,K}]$. In other words,

$$U_{t,j} = \underset{k \in \{1,...,K\}}{\arg\max}\ \mathbb{E}\left[u_{t,j}(w_{t,j} - (1-s)D_{t,k} + F + r(L_{t,j}, k) - L_{t,j}; R_{t,j}) \,\middle|\, \overline{\mu}_{t,j}\right].$$

Since we assume that all plan members rationally maximize their expected utility, we can define $S_{t,j}(\boldsymbol{D}_t, \overline{\mu}_{t,j}) \equiv U_{t,j}$ as plan member j's selected option in year t, given the deposit rates and the plan member's expected annual loss.[11]

2.4 Calculating Deposit Rates

2.4.1 The Actuary's Estimator for Health Statuses

As mentioned in the introduction, a highlight of our framework is our pricing method, which incorporates both the annual claims and selection models. Since these models depend on the true health statuses and perceived health statuses, respectively, the pricing actuary needs to estimate both these latent variables.

Let $\widetilde{H}_{t,j}$ represent the pricing actuary's assessment of plan member j's true health status in year t. One convenient way to estimate the distribution of this quantity is to simply use the unconditional distribution in Equation (2.4).[12] However, since we assume that the actuary

[11] While $S_{t,j}(\boldsymbol{D}_t, \overline{\mu}_{t,j})$ is equivalent to $U_{t,j}$ in this section, we distinguish between these two quantities in a proposed extention of the model, as shown in Appendix C.

[12] Estimating health statuses based on the unconditional distribution is analogous to treating the group as having no credibility and thereby selecting manual rates as deposit rates. Manual rates refer to the insurer's rate table, and will be discussed again in Chapters 3 and 4.

also observes the plan member's past losses, a more accurate method of estimating these distributions is through the conditional distribution given the observed past losses. In other words, the actuary predicts the posterior distribution of $H_{t,j}$ denoted by

$$\widetilde{H}_{t,j} \equiv H_{t,j}|L_{1:t-1,j},$$

where $L_{1:t-1,j} = \{L_{i,j}\}_{i=1}^{t-1}$.

The distribution of $\widetilde{H}_{t,j}$, known as the predictive distribution in filtering, can be calculated using $H_{t-1,j}|L_{1:t-1,j}$ and the transition density of $H_{t,j}$:

$$f_{\widetilde{H}_{t,j}}(h_{t,j}|L_{1:t-1,j}) = \int_{\mathbb{R}} f_{H_{t,j}|H_{t-1,j}}(h_{t,j}|h_{t-1,j}) f_{H_{t-1,j}|L_{1:t-1,j}}(h_{t-1,j}|L_{1:t-1,j}) dh_{t-1,j}. \quad (2.7)$$

As shown in Appendix A.1, the distribution of $H_{t-1,j}|L_{1:t-1,j}$, also known as the filtering distribution, can be found recursively from the distribution of $H_{t-2,j}|L_{1:t-2,j}$ using the following formula:

$$\begin{aligned} &f_{H_{t-1,j}|L_{1:t-1,j}}(h_{t-1,j}|L_{1:t-1,j}) \quad (2.8) \\ &= \frac{\int_{\mathbb{R}} f_{L_{t-1,j}|H_{t-1,j}}(L_{t-1,j}|h_{t-1,j}) f_{H_{t-1,j}|H_{t-2,j}}(h_{t-1,j}|h_{t-2,j}) f_{H_{t-2,j}|L_{1:t-2,j}}(h_{t-2,j}|L_{1:t-2,j}) dh_{t-2,j}}{\int_{\mathbb{R}^2} f_{L_{t-1,j}|H_{t-1,j}}(L_{t-1,j}|h_{t-1,j}) f_{H_{t-1,j}|H_{t-2,j}}(h_{t-1,j}|h_{t-2,j}) f_{H_{t-2,j}|L_{1:t-2,j}}(h_{t-2,j}|L_{1:t-2,j}) dh_{t-1,j} dh_{t-2,j}}, \end{aligned}$$

where the starting distribution is the unconditional health status density in the year prior to the start of loss observation. So far, for simplicity, we assume in our notation that losses are observed starting in year 1. In general, past loss data from years prior to plan implementation are often available. Let the losses be observed starting in year $t_0 + 1$. Then the starting distribution is $f_{H^{(unc)}_{t_0,j}}(h_{t_0,j})$. Note that this distribution, as well as the density functions of $L_{t,j}|H_{t,j}$ and $H_{t,j}|H_{t-1,j}$ (all previously introduced in Section 2.2), are assumed to be known to the actuary.

Given the complicated nature of solving the above recursion, we employ a filter to obtain a numerical approximation. While a commonly used filtering algorithm is the particle filter, in our project we apply the unscented Kalman filter (Julier and Uhlmann, 1997, UKF henceforth), which turns out to be both a suitable and a more efficient algorithm for solving the above recursion. Note that by construction, the UKF assumes that the conditional distribution can be approximated by a normal distribution. We have assessed the validity of this assumption in our context by simulating health statuses and losses for a hypothetical group of flex plan members and generating estimates using both filters. Since the UKF estimates are very similar to results from the particle filter (which does not assume any particular form for the posterior distribution of the latent variable), we confirm that the normality assumption is reasonable.

Current Time	**Actuary's Prediction for**	**Filtering Distribution**	**Predictive Distribution**
End of Year 0	Year 1	$H_{0,j}^{(unc)} \dot\sim \mathcal{N}(\widetilde{h}_{0,j},\ \rho_{0,j})$	$\widetilde{H}_{1,j} \dot\sim \mathcal{N}\left(\widetilde{h}_{0,j} - \lambda,\ \sqrt{\rho_{0,j}^2 + \sigma^2}\right)$
End of Year 1	Year 2	$H_{1,j} \vert L_{1,j} \dot\sim \mathcal{N}(\widetilde{h}_{1,j},\ \rho_{1,j})$	$\widetilde{H}_{2,j} \dot\sim \mathcal{N}\left(\widetilde{h}_{1,j} - \lambda,\ \sqrt{\rho_{1,j}^2 + \sigma^2}\right)$
End of Year 2	Year 3	$H_{2,j} \vert L_{1:2,j} \dot\sim \mathcal{N}(\widetilde{h}_{2,j},\ \rho_{2,j})$	$\widetilde{H}_{3,j} \dot\sim \mathcal{N}\left(\widetilde{h}_{2,j} - \lambda,\ \sqrt{\rho_{2,j}^2 + \sigma^2}\right)$

Table 2.1: **First Three Years of Predictive Health Status Distributions when $t_0 = 0$.**

Therefore, using the UKF, we can fully define the posterior distribution by two parameters: the mean and the variance. At each iteration, the algorithm takes the estimated mean and variance of a plan member's health status from the previous year and the loss in the current year to estimate the mean and variance of the plan member's health status in the current year. For example, the first iteration of the UKF requires three starting values: the mean and variance estimates of the year t_0 unconditional health status distribution, which can be found in Equation (2.4), and the loss amount in year $t_0 + 1$. Based on these inputs, the algorithm then estimates the health status distribution in year $t_0 + 1$ by estimating the posterior mean (which we denote as $\widetilde{h}_{t_0+1,j}$) and posterior variance (which we denote as $\rho^2_{t_0+1,j}$). Finally, the actuary predicts the plan member's distribution of true health status in the upcoming year as a normal distribution with mean $\widetilde{h}_{t_0+1,j} - \lambda$ and variance $\rho^2_{t_0+1,j} + \sigma^2$. The second iteration then uses $\widetilde{h}_{t_0+1,j}$, $\rho^2_{t_0+1,j}$ and the loss amount in year t_0+2 to calculate the posterior mean and variance estimates of the year t_0+2 health status distribution. Based on these new estimates, which we denote as $\widetilde{h}_{t_0+2,j}$ and $\rho_{t_0+2,j}$, respectively, the actuary can then predict the plan member's year t_0+3 health status distribution as a normal distribution with mean $\widetilde{h}_{t_0+2,j} - \lambda$ and variance $\rho^2_{t_0+2,j} + \sigma^2$. Throughout the years, the actuary refines and updates these mean and variance estimates as new annual loss amounts are observed, so that the approximated health status distributions become increasingly precise through time.[13] Table 2.1 illustrates the process of updating these parameter estimates in the first three years.

This means we can approximate $\widetilde{H}_{t,j}$ as a normal distribution with mean $\widetilde{h}_{t-1,j} - \lambda$ and standard deviation $\sqrt{\rho^2_{t-1,j} + \sigma^2}$. In other words,

$$\widetilde{H}_{t,j} \dot\sim \mathcal{N}\left(\widetilde{h}_{t-1,j} - \lambda,\ \sqrt{\rho^2_{t-1,j} + \sigma^2}\right), \tag{2.9}$$

[13]For more details about the UKF, refer to Appendix A.2.

where the symbol $\dot{\sim}$ denotes an approximate distribution. Using this conditional distribution of $H_{t,j}$, we can now estimate the conditional distribution of $\overline{H}_{t,j}$ based on the relationship between the $H_{t,j}$ and the $\overline{H}_{t,j}$ models from Equation (2.6):

$$\overline{H}_{t,j} | H_{t,j} = h \sim \mathcal{N}(h, \delta). \tag{2.10}$$

We have shown how the distributions of the latent variables $H_{t,j}$ and $\overline{H}_{t,j}$ can be approximated. Next, we show how we can transform these into approximate distributions of the mean annual loss parameters.

Let $\widetilde{\mu}_{t,j}$ be the transformation of $\widetilde{H}_{t,j}$ based on the link function in Equation (2.5). Then we can approximate $\widetilde{\mu}_{t,j}$ by a lognormal distribution with parameters as follows:

$$\widetilde{\mu}_{t,j} \,\dot{\sim}\, \log\mathcal{N}\left(\beta_{t,0} + \beta_{t,1}(\widetilde{h}_{t-1,j} - \lambda),\ \beta_{t,1}\sqrt{\rho^2_{t-1,j} + \sigma^2}\right), \tag{2.11}$$

where the first and second parameters represent the mean and standard deviation of the corresponding normal distributions.

It then follows from Equations (2.5) and (2.10) that the estimated distribution of the plan member's perceived mean loss parameter given a true parameter value of v is

$$\overline{\mu}_{t,j} | \mu_{t,j} = v \ \sim\ \log\mathcal{N}(\ln v,\ \beta_{t,1}\delta). \tag{2.12}$$

The above approximate distributions of the mean annual loss parameters allow the actuary to estimate the annual loss and member selection models. The next subsection outlines how our pricing method incorporates these two models to calculate optimal deposit rates.

2.4.2 A Constrained Optimization Problem

We use the actuarial equivalence principle of premium calculations as a starting point to construct the objective of our pricing problem. Under the equivalence principle, $\boldsymbol{D}_t$ should be chosen so that the expected annual deposit amount for each option is the expected annual claims in that option.[14] In other words, the cash inflow should breakeven with the cash outflow in each option's ASO deposit account.[15]

[14] We assume for simplicity that the loading factor is zero, but the calculations can easily be extended to incorporate positive loading factors.

[15] We assume that all incurred claims are immediately reported, so that there is no float requirement.

The total amount deposited into an option's ASO account depends on the number of plan members in that option and therefore the plan members' selections given $\boldsymbol{D}_t$. Meanwhile, the total cash outflow of the account depends both on the mix of plan members selecting that option as well as their claim amounts. The total cash inflow and outflow of option k's account are

$$D_{t,k}\left(\sum_{j=1}^{N} 1_{\{S_{t,j}(\boldsymbol{D}_t,\overline{\mu}_{t,j})=k\}}\right)$$

and

$$\sum_{j=1}^{N} 1_{\{S_{t,j}(\boldsymbol{D}_t,\overline{\mu}_{t,j})=k\}} r(L_{t,j},k),$$

respectively. This means that deposit rates should satisfy the following system of equations:

$$D_{t,k}\mathbb{E}\left[\sum_{j=1}^{N} 1_{\{S_{t,j}(\boldsymbol{D}_t,\overline{\mu}_{t,j})=k\}} \;\middle|\; L_{1:t-1,j}\right] = \mathbb{E}\left[\sum_{j=1}^{N} 1_{\{S_{t,j}(\boldsymbol{D}_t,\overline{\mu}_{t,j})=k\}} r(L_{t,j},k) \;\middle|\; L_{1:t-1,j}\right], \quad \forall k \in \{1,...,K\}.$$

However, this system may not have a solution given the complexity of these equations with respect to $\boldsymbol{D}_t$. Therefore, we relax it into a multi-objective optimization problem. Specifically, we minimize the squared difference between the expected total deposits and the expected total claims in each option. Assuming an equal importance of minimizing the said objective for each option, we further combine the set of objectives into one function through a summation, so that our objective function is

$$\min_{\boldsymbol{D}_t} \sum_{k=1}^{K}\left(\sum_{j=1}^{N}\mathbb{E}\left[1_{\{S_{t,j}(\boldsymbol{D}_t,\overline{\mu}_{t,j})=k\}}\left(D_{t,k}-r(L_{t,j},k)\right) \;\middle|\; L_{1:t-1,j}\right]\right)^2, \tag{2.13}$$

where

$$\begin{aligned}
&\mathbb{E}\left[1_{\{S_{t,j}(\boldsymbol{D}_t,\overline{\mu}_{t,j})=k\}}\left(D_{t,k}-r(L_{t,j},k)\right) \;\middle|\; L_{1:t-1,j}\right] \\
&= \int_{\mathbb{R}_+^2}\left[1_{\{S_{t,j}(\boldsymbol{D}_t,u)=k\}}\left(D_{t,k}-r(l,k)\right)\right] \times f_{L_{t,j},\overline{\mu}_{t,j}|L_{1:t-1,j}}(l,u|L_{1:t-1,j})\,du\,dl \\
&= \int_{\mathbb{R}_+^3}\left[1_{\{S_{t,j}(\boldsymbol{D}_t,u)=k\}}\left(D_{t,k}-r(l,k)\right)\right] \\
&\qquad \times f_{L_{t,j},\overline{\mu}_{t,j}|L_{1:t-1,j},\mu_{t,j}}(l,u\,|\,L_{1:t-1,j},v) f_{\mu_{t,j}|L_{1:t-1,j}}(v|L_{1:t-1,j})\,dv\,du\,dl \\
&= \int_{\mathbb{R}_+^3}\left[1_{\{S_{t,j}(\boldsymbol{D}_t,u)=k\}}\left(D_{t,k}-r(l,k)\right)\right] \times f_{L_{t,j},\overline{\mu}_{t,j}|\mu_{t,j}}(l,u\,|\,v) f_{\tilde{\mu}_{t,j}}(v)\,dv\,du\,dl \\
&= \int_{\mathbb{R}_+^3}\left[1_{\{S_{t,j}(\boldsymbol{D}_t,u)=k\}}\left(D_{t,k}-r(l,k)\right)\right] \times f_{L_{t,j}|\mu_{t,j}}(l\,|\,v) f_{\overline{\mu}_{t,j}|\mu_{t,j}}(u\,|\,v) f_{\tilde{\mu}_{t,j}}(v)\,dv\,du\,dl.
\end{aligned}$$

The loss density, $f_{L_{t,j}|\mu_{t,j}}(l\,|\,v)$, is assumed to be known, while the densities $f_{\overline{\mu}_{t,j}|\mu_{t,j}}(u\,|\,v)$ and $f_{\widetilde{\mu}_{t,j}}(v)$ can be computed based on Equations (2.11) and (2.12), respectively.[16]

Without a constraint, the optimal solution to the above optimization problem may result in at least one option with zero expected enrollment.[17] Since this defeats the purpose of offering a flex plan to members, we introduce a constraint to eliminate any such impractical solutions.

Including a constraint also serves as an opportunity to help guide plan members to make a wise decision. While the enrollment choices ultimately lie with plan members, it is typically ideal for plan members to select coverage levels that correspond to their loss levels. That is, we want high risk members to choose a high coverage option so that in the event of a large loss, they will not find themselves underinsured. Thus, we allocate members into K groups ex ante based on the actuary's estimates of each member's expected annual loss, where group $G_{t,k}$ contains n_k members with expected loss within some predetermined range, and $\sum_{k=1}^{K} n_k = N$. For example, the plan sponsor may prefer an equal number of plan members in each option, so that the end points of the k^{th} predetermined range are the $\frac{k-1}{K}^{\text{th}}$ and $\frac{k}{K}^{\text{th}}$ percentiles of the annual loss estimates.[18] Although these allocations only represent what the actuary hopes to be the plan members' actual choice, we can implement the constraint below to maximize the chance that these hypothetical allocations are realized:

$$\underset{k' \in \{1,2,...,K\}}{\arg\max} \ \mathbb{P}(S_{t,j}(\boldsymbol{D}_t, \overline{\mu}_{t,j}|L_{1:t-1,j}) = k') = k, \quad \forall j \in G_{t,k}, \quad \forall k \in \{1, 2, ..., K\}.$$

That is, based on the approximate distributions of $\{\overline{\mu}_{t,j}\}_{j=1}^{N}$, the actuary can select deposit rates that maximize the probability that a plan member j's selected option, k', is in fact the actuary's intended option, k. However, considering the amount of complexity the above constraint adds to the optimization problem and the fact that these allocations merely act as guides built into the deposit rates when the enrollment decisions ultimately lie with plan members, we propose a simpler, more practical constraint:

$$S_{t,j}(\boldsymbol{D}_t, \mathbb{E}[\overline{\mu}_{t,j}|L_{1:t-1,j}]) = k, \quad \forall j \in G_{t,k}, \quad \forall k \in \{1, 2, ..., K\}.$$

[16] Note that the $f_{L_{t,j},\overline{\mu}_{t,j}|L_{1:t-1,j},\mu_{t,j}}$ in the third line above is equivalent to $f_{L_{t,j},\overline{\mu}_{t,j}|\mu_{t,j}}$ in the fourth line since $\mu_{t,j}$ contains more information than $L_{1:t-1,j}$. This density is split into a product of $f_{L_{t,j}|\mu_{t,j}}$ and $f_{\overline{\mu}_{t,j}|\mu_{t,j}}$ in the fifth line since $L_{t,j}$ and $\overline{\mu}_{t,j}$ are independent when conditioned on $\mu_{t,j}$.

[17] For example, there exist trivial solutions of $\boldsymbol{D}_t$ so that all plan members select the same option, leaving other options with no enrollment.

[18] Note that the choice of $\{n_k\}_{k=1}^{K}$ will affect the expected annual change in ASO account balance; therefore, keeping the total ASO account balance as close to zero as possible through time requires selecting a specific rather than an arbitrary set of $\{n_k\}_{k=1}^{K}$. This will be further discussed in Chapter 4.

In this proposed constraint, the approximate distributions of $\{\overline{\mu}_{t,j}\}_{j=1}^{N}$ are replaced by their expectations, making the selection functions deterministic and therefore ensuring the feasibility of computing the constrained optimization problem. Combining the objective function with this constraint, we have the following constrained optimization problem for calculating deposit rates:

$$\min_{\boldsymbol{D}_t} \sum_{k=1}^{K} \left(\sum_{j=1}^{N} \int_{\mathbb{R}_+^3} \left[1_{\{S_{t,j}(\boldsymbol{D}_t,u)=k\}} \left(D_{t,k} - r(l,k)\right) \right] f_{L_{t,j}|\mu_{t,j}}(l \,|\, v) f_{\overline{\mu}_{t,j}|\mu_{t,j}}(u \,|\, v) f_{\tilde{\mu}_{t,j}}(v)\, dv\, du\, dl \right)^2 \tag{2.14}$$

subject to $S_{t,j}(\boldsymbol{D}_t, \mathbb{E}[\overline{\mu}_{t,j}|L_{1:t-1,j}]) = k, \ \forall j \in G_{t,k}, \ \forall k \in \{1, 2, ..., K\}$, where $j \in G_{t,k}$ if $\mathbb{E}[\widetilde{\mu}_{t,j}]$ is within some predetermined range, and the mutually exclusive sets $G_{t,k}$ span $\{1, ..., N\}$ so that $G_{t,k} \cap G_{t,k'} = \varnothing, \ \forall k \neq k'$, and $\cup_{k=1}^{K} G_{t,k} = \{1, ..., N\}$.

2.5 Tying the Components Together Through the Annual Cycle of a Flex Plan

Notation	Description
$H_{t,j}$	True health status of plan member j in year t
$\overline{H}_{t,j}$	Plan member j's intuition at the end of year $t-1$ about their health status in the upcoming year
$\widetilde{H}_{t,j}$	Actuary's estimator of plan member j's health status in year t
$\widetilde{h}_{t,j}$	Mean of the actuary's approximated distribution of plan member j's health status in year t
$\rho_{t,j}$	Variance of the actuary's approximated distribution of plan member j's health status in year t
$H_{t,j}^{(unc)}$	Unconditional distribution of plan member j's health status in year t
$\mu_{t,j}$	True expected annual loss of plan member j in year t
$\overline{\mu}_{t,j}$	Plan member j's expectation of their annual loss in year t
$\widetilde{\mu}_{t,j}$	Actuary's expectation of plan member j's annual loss in year t
$L_{t,j}$	Plan member j's annual loss in year t
$S_{t,j}$	Plan member j's selected option in year t
$\boldsymbol{D}_t$	Deposit rates in year t
K	Number of options in the flex plan
$G_{t,k}$	Actuary's anticipated group of plan members enrolled in option k in year t
N	Number of plan members in the flex plan
n_k	Actuary's intended number of plan members enrolled in option k

Table 2.2: **Table of Notations.**

We have now outlined each of the four pieces of our framework: the plan design, which forms the structure of a flex plan, and three main components—the claims model, selections model, and the pricing method—which correspond to the three main events in the annual cycle of a flex plan:

1. At the end of year $t-1$ the actuary prices each option's deposit rates for an upcoming year. Specifically, the pricing actuary estimates each plan member's distribution of

health status and option selection in year t given their losses up to and including year $t-1$. This information is then used to set year t deposit rates, which are then communicated to the plan members.

2. Given the rates, the plan members each select a coverage option for the upcoming year t.

3. Finally, throughout the benefit year (i.e., year t), as each plan member's losses are reported, claims (and associated administrative fees) are withdrawn from the ASO account based on each employee's selected option. Deposits are also regularly paid into the same account. The actuary then uses the individual annual loss amounts to calculate the subsequent year's deposit rates, initiating the next annual cycle of a flex plan.

Figure 2.1 summarizes the relationship between the various parts of the framework. For ease of communicating the pieces of each component, we have temporarily adopted the notations $\widehat{H}_{t,j}$ and $\widehat{\mu}_{t,j}$ to represent the actuary's approximations of the perceived health status and the perceived mean loss parameter. These notations each combine two steps done in the deposit rate calculations; $\widehat{\mu}_{t,j}$ is an implied quantity found by integrating the distribution of $\overline{\mu}_{t,j}|\mu_{t,j}$ in Equation (2.12) over the approximate $\widetilde{\mu}_{t,j}$ distribution in Equation (2.11). Likewise, $\widehat{H}_{t,j}$, which is included in the diagram for completion, is an implied quantity that can be found by integrating the distribution of $\overline{H}_{t,j}|H_{t,j}$ in Equation (2.10) over the approximate $\widetilde{H}_{t,j}$ distribution in Equation (2.9). For convenience, we have also included Table 2.2, which summarizes the key notation introduced thus far.

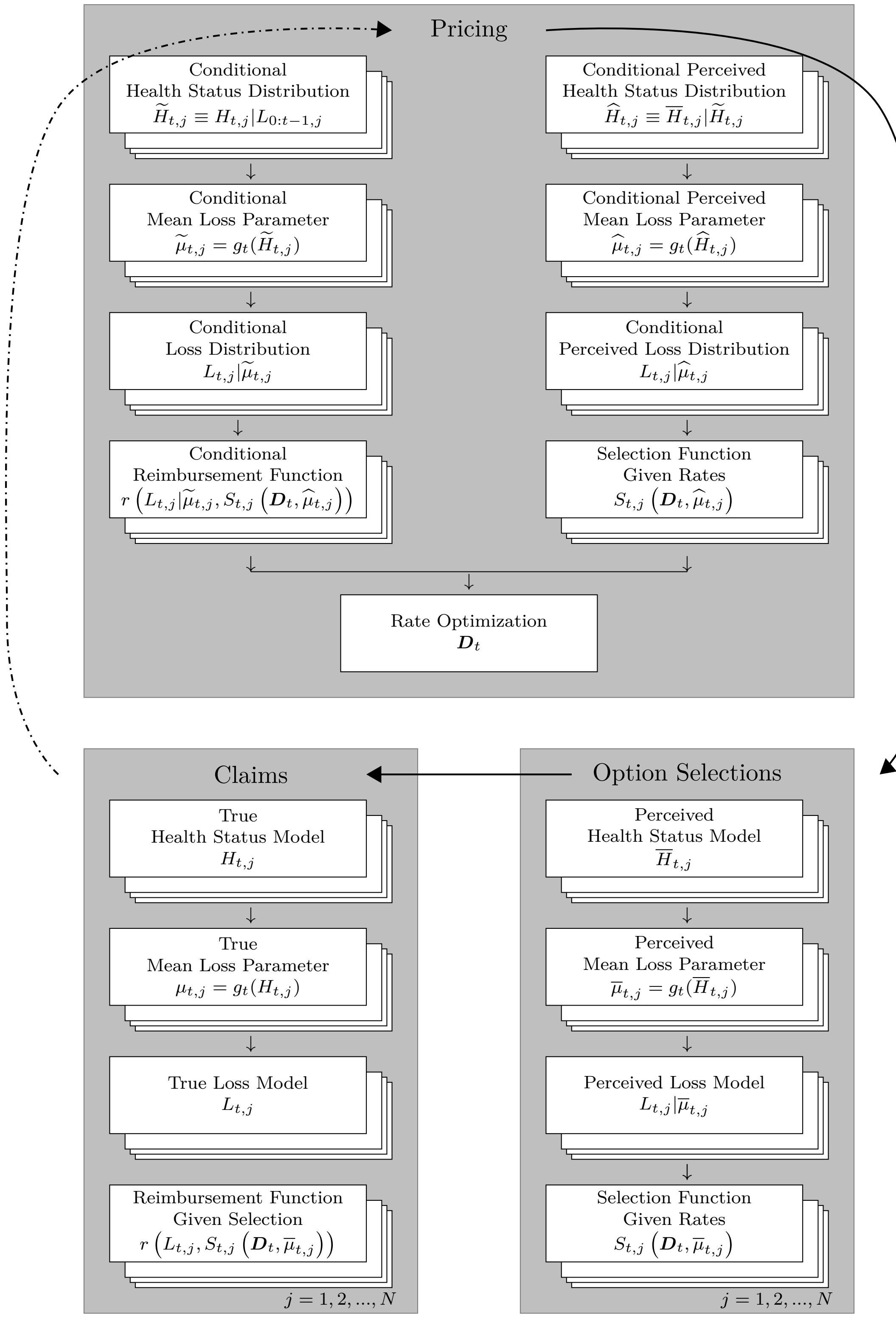

Figure 2.1: **Relationship Between the Three Main Components of an Annual Flex Plan Cycle.**

Chapter 3

Simulation Study: Assumptions and Methodology

We conduct a simulation study to compare the performance of a flex plan priced using our proposed formula against the same plan priced with the traditional practice. In addition, since we are interested in eliminating adverse selection—which requires asymmetric information—we also compare the plan's dynamics under various information sets when priced using our proposed formula. Finally, to test the robustness of our proposed method, we repeat the simulation study using various alternative parameters. In this chapter, we outline our assumptions and simulation methodology, and in the next chapter we present our key findings.

3.1 Assumptions

3.1.1 Base Case Parameters

We consider a hypothetical group benefit with $K = 2$ options: "low coverage" and "high coverage." The low coverage option has a deductible of \$25, a coinsurance rate of 50%, and a coinsurance maximum of \$2,500. The high coverage option has no deductible, a coinsurance rate of 90%, and a coinsurance maximum of \$2,000. In addition, both options have an overall maximum of \$1,000,000. The plan sponsor provides $F = \$500$ in flex credits, and any excess unused amount can be either cashed out or used towards a group retirement or savings plan. In other words, we have the following plan design parameters:

- $[d_1, c_1, CM_1, OM_1]$ = [\$25, 50%, \$2,500, \$1,000,000],
- $[d_2, c_2, CM_2, OM_2]$ = [\$0, 90%, \$2,000, \$1,000,000], and
- $s = 0\%$.

We study this flex plan for $T = 15$ years.

We set up a hypothetical group of $N = 400$ plan members with stationary group demographics. Specifically, we assume that all plan members join the group at age 25 and retire at age 64, so that at any point in time, we have 10 members of each age from 25 to 64. This is to remove any demographic change effects from our results. These plan members' utility functions are assumed to be exponential, with a constant absolute risk aversion parameter. We also assume that all plan members have the same constant annual salary of $w_{t,j} =$ \$50,000 over time so that our calculations will be net of any inflationary effects. Therefore, the functions and parameters related to the option selection model are:

$$u_{t,j}(c) = -e^{-R_{t,j}c}, \quad \forall\, t, j,$$

where $R_{t,j} = 0.00125$.[19]

We assume that in any year t, the health statuses of 25-year-old plan members are normally distributed with a mean of $H^{(p)}_{t,25} = 100$ units and a standard deviation of $\zeta = 1$ unit. We can interpret 100 health units as a convenient reference point for our health scale, much like the boiling and freezing points for water in a Celsius temperature scale. We further assume that each plan member's health status deteriorates at an average rate of $\lambda = 1$ unit per year, subject to a normally distributed random term with a standard deviation of $\sigma = 0.5$ units. In our simulations, our plan members have an intuition of their upcoming year's true health status, subject to a normally distributed estimation error with standard deviation $\delta = 1$.

Each plan member's loss distribution is assumed to be gamma with a known rate parameter, $\beta = 0.1$, that is constant for all plan members and through time. The link function parameters, which depend on the health status scale, are assumed to be constant through time so that we eliminate the impact of any health cost trends from our results. For our simulations we have $\beta_{t,0} = 13$ and $\beta_{t,1} = -0.1$. This means, for example, that a plan member with a health status of 100 units will have an expected annual loss of $e^{13-0.1(100)} = \$20$, while a plan member with a health status of 61 units will have an expected annual loss of $e^{13-0.1(61)} = \$992$. We assume that the most recent three years' annual losses incurred at or after age 25 are available by individual; that is, $t_0 + 1 = -2$.[20]

Table 3.1 summarizes the parameter values used in the simulation study.

[19] This is a representative value among absolute risk aversion parameters found in the literature (Babcock et al., 1993)

[20] Typically, a new carrier requests three years' historical claims data from the group.

Parameter	**Value**
K	2
F	\$500
s	0
$[d_1, c_1, CM_1, OM_1]$	[\$25, 50%, \$2,500, \$1,000,000]
$[d_2, c_2, CM_2, OM_2]$	[\$0, 90%, \$2,000, \$1,000,000]
T	15
N	400
$u_{t,j}(c)$	$-e^{-R_{t,j}c}, \forall\, t, j$
$R_{t,j}$	0.00125
x_0	25
$H^{(p)}_{t,25}$	100
ζ	1
λ	1
σ	0.5
δ	1
β	0.1
$\beta_{t,0}$	13
$\beta_{t,1}$	-0.1
t_0	-3

Table 3.1: **Summary of Parameter Values in the Simulation Study.**

3.1.2 Robustness Test Settings

To test the robustness of our model, we repeat each set of simulations nine times in addition to the base case under alternative parameters, as outlined in the list of alternative settings below.

- Changes in Plan Design

  1. Decrease the Difference in Coverage Between The Two Options
     - We decrease the difference in coverage between the two options by changing the coinsurance levels from $c_1 = 50\%$ and $c_2 = 90\%$ to $c_1 = 65\%$ and $c_2 = 80\%$.
  2. Increase the Number of Options
     - We add a third option—"Medium Coverage"—with a deductible of 0, coinsurance of 75%, coinsurance maximum of \$2,500, and an overall maximum of \$1,000,000. To have equal intended enrollment levels in each of the three options, we change the group size to $N = 360$ so that the intended enrollment levels are $n_1 = n_2 = n_3 = 120$. We also maintain a stationary population, with nine plan members of each age between 25 and 64.

- Changes in Group Size

  3. Decrease in Group Size

    - We maintain a stationary population but decrease the group size from 400 to 200 so that there are five plan members of each age between 25 and 64.

 4. Increase in Group Size
    - We maintain a stationary population but increase the group size from 400 to 600 so that there are fifteen plan members of each age between 25 and 64.

- Changes in Premium Subsidy Arrangement
 5. Replace Flex Credits with Moderate Level of Percent Premium Subsidy
    - We replace flex credits with a 50% premium subsidy arrangement.
 6. Replace Flex Credits with High Level of Percent Premium Subsidy
    - We replace flex credits with a 75% premium subsidy arrangement.

- Changes in Intended Enrollment Levels
 7. Change in Intended Enrollment Levels
    - We change the ex ante option allocation from $n_1 = n_2 = 200$ to $n_1 = 300$ and $n_2 = 100$.
 8. Change in Both the Intended Enrollment Levels and Premium Subsidy Arrangement
    - We change the ex ante option allocation from $n_1 = n_2 = 200$ to $n_1 = 300$ and $n_2 = 100$. In addition, we also replace flex credits with a 75% premium subsidy arrangement.
 9. Change in Both the Number of Options and Intended Enrollment Proportions
    - We add a third option as described in Robustness Test 2 above, but with a group size of $N = 400$ so that the intended enrollment levels are $n_1 = 134$ and $n_2 = n_3 = 133$.

3.2 Simulation Methodology

We simulate 10,000 paths of true health statuses, annual losses, and members' perceived health statuses based on the models and parameters described in the previous sections, and analyze the distribution of deposit rates, enrollment figures, and ASO account balances of the flex plan under various settings.

Specifically, for each of the 10,000 paths, we generate 18 years' true health statuses (i.e., $\{H_{t,j}\}_{t=-2}^{15}$) for members $j = 1, 2, ..., 400$. From each of these simulated health statuses, plan members' perceived health statuses (i.e., $\{\overline{H}_{t,j}\}_{t=1}^{15}$) and annual losses (i.e., $\{L_{t,j}\}_{t=-2}^{15}$) are generated for each member $j = 1, 2, ..., 400$.[21] We then simulate the evolution of the 15 year benefit plan from year 1 to year 15 for each of the 10,000 paths under each setting.

3.2.1 Calculating Deposit Rates Under the Traditional Pricing Practice

Under the traditional pricing practice, we assume that the actuary's estimate of the group's annual losses in the upcoming year is equal to the group's observed losses in the current year.[22] As we wish to observe the start of any adverse selection spiral, we assume the flex plan is not implemented until year 1, meaning no prior option selection data are available. Therefore, the actuary initially assumes that the 200 members with lower estimated losses will select the low coverage option, and that the remaining 200 members will select the high coverage option. Each upcoming year's deposit rates are then set to be equal to the current year's per capita claims within each option.

3.2.2 Calculating Deposit Rates Using Our Proposed Pricing Method

We also simulate the flex plan evolution with deposit rates calculated based on our proposed method outlined in Section 2. Considering the long computing time required to simulate a large number of potential flex plan dynamics, we streamline the optimizations by approximating the distributions of both the plan members' perceived mean loss (i.e., $\overline{\mu}_{t,j}$) and the true mean loss (i.e., $\mu_{t,j}$) with the point estimate $\mathbb{E}[\widetilde{\mu}_{t,j}]$ when calculating deposit rates. This means that our objective function is approximated by

$$
\begin{aligned}
&\min_{\boldsymbol{D_t}} \sum_{k=1}^{K} \left(\sum_{j=1}^{N} \mathbb{E}\left[1_{\{S_{t,j}(\boldsymbol{D}_t, \overline{\mu}_{t,j})=k\}} \left(D_{t,k} - r(L_{t,j}, k)\right) \,\Big|\, L_{1:t-1,j} \right] \right)^2 \\
\approx &\min_{\boldsymbol{D_t}} \sum_{k=1}^{K} \left(\sum_{j=1}^{N} 1_{\{S_{t,j}(\boldsymbol{D}_t, \mathbb{E}[\mu_{t,j}|L_{1:t-1,j}])=k\}} \left(D_{t,k} - \mathbb{E}\Big[r(L_{t,j}, k) \,\Big|\, \mathbb{E}[\mu_{t,j}|L_{1:t-1,j}] \Big] \right) \right)^2 \\
= &\min_{\boldsymbol{D_t}} \sum_{k=1}^{K} \left(\sum_{j=1}^{N} 1_{\{S_{t,j}(\boldsymbol{D}_t, \mathbb{E}[\widetilde{\mu}_{t,j}])=k\}} \left(D_{t,k} - \mathbb{E}\Big[r(L_{t,j}, k) \,\Big|\, \mathbb{E}[\widetilde{\mu}_{t,j}] \Big] \right) \right)^2,
\end{aligned}
$$

and that our constraint becomes

$$
\begin{aligned}
&S_{t,j}(\boldsymbol{D}_t, \mathbb{E}[\overline{\mu}_{t,j}|L_{1:t-1,j}]) = k \\
\approx &S_{t,j}(\boldsymbol{D}_t, \mathbb{E}[\mu_{t,j}|L_{1:t-1,j}]) = k
\end{aligned}
$$

[21] Only 15, 16, and 17 years' true health statuses and annual losses were generated for members aged 25, 26, or 27 in year 1, respectively, since we treat historical data prior to age 25 as unavailable.

[22] In practice, the actuary's estimate of the upcoming year's losses is equal to the previous year's losses multiplied by one plus a trend factor. As previously mentioned, the trend factor is assumed to be always zero in our study.

$$= S_{t,j}(\boldsymbol{D}_t, \mathbb{E}[\widetilde{\mu}_{t,j}]) = k, \quad \forall j \in G_{t,k}, \quad \forall k \in \{1, 2, ..., K\}.$$

These approximations reduce the problem into a quadratic optimization and even yield a closed-form solution in the two-options case. This significantly improves the efficiency of running the simulations while producing very similar deposit rates compared to the full proposed optimization problem in Equation (2.14).[23] As will be shown in Section 4, this also turns out to be a good approximation in terms of meeting our four goals.

The point estimate $\mathbb{E}[\widetilde{\mu}_{t,j}]$ is calculated based on the distribution of $\tilde{\mu}_{t,j}$, which is estimated using the UKF algorithm as explained in Section 2. The UKF parameters used in our simulation can be found in Appendix A.2.

The groups $G_{t,k}$ are chosen so that an equal number of plan members are hypothetically allocated to each option ex ante.

3.2.3 Closed Form Solution in the Two-Options Case

Recall that for our simulation study, we have the following assumptions:

- plan member j's losses in year t follow a gamma distribution with a rate parameter of β and a shape parameter of $\alpha_{t,j} = \mu_{t,j}\beta$, $\forall\, t, j$,
- plan member j's utility function is $u_{t,j}(c) = -e^{-R_{t,j}c}$, $\forall\, t, j$, and
- the reimbursement function of option k is as defined in Equation (2.1).

Given these special case assumptions, a closed-form solution to our proposed pricing formula is available if each plan members' true health statuses are known to both the actuary and the member.

Proposition 1. *Let the special case assumptions hold. Then plan member j's expected total claim amount during year t is*

$$\begin{aligned}
EC(\alpha_{t,j}, k) =& \frac{c_k \alpha_{t,j}}{\beta} \left[F_L\left(\frac{CM_k}{1-c_k} + d_k; \alpha_{t,j}+1, \beta\right) - F_L(d_k; \alpha_{t,j}+1, \beta) \right] \\
&- c_k d_k \left[F_L\left(\frac{CM_k}{1-c_k} + d_k; \alpha_{t,j}, \beta\right) - F_L(d_k; \alpha_{t,j}, \beta) \right] \\
&+ \frac{\alpha_{t,j}}{\beta} \left[F_L(OM_k + d_k + CM_k; \alpha_{t,j}+1, \beta) - F_L\left(\frac{CM_k}{1-c_k} + d_k; \alpha_{t,j}+1, \beta\right) \right] \\
&- (d_k + CM_k) \left[F_L(OM_k + d_k + CM_k; \alpha_{t,j}, \beta) - F_L\left(\frac{CM_k}{1-c_k} + d_k; \alpha_{t,j}, \beta\right) \right]
\end{aligned}$$

[23] We have simulated one 15-year flex plan evolution and found almost identical results between calculating deposit rates using the full proposed optimization problem and the approximated version.

$$+ OM_k[1 - F_L(OM_k + d_k + CM_k; \alpha_{t,j}, \beta)],$$

where $F_L(l; \alpha, \beta)$ denotes the cumulative distribution function of a gamma distribution with shape parameter α and rate parameter β, evaluated at point l.

Proof. Refer to Appendix B.1 □

Proposition 2. *Let the special case assumptions hold and assume that each plan members' true health statuses are known to both the actuary and the member. Then the constraint in the pricing optimization problem can be simplified to*

$$LB_t \leq D_{t,2} - D_{t,1} \leq UB_t,$$

where the upper and lower bounds are given by

$$LB_t = \max_{j \in G_{t,1}} \frac{1}{(1-s)R_{t,j}} \, ln\left(\frac{h(\alpha_{t,j},1)}{h(\alpha_{t,j},2)}\right) > 0,$$

$$UB_t = \min_{j \in G_{t,2}} \frac{1}{(1-s)R_{t,j}} \, ln\left(\frac{h(\alpha_{t,j},1)}{h(\alpha_{t,j},2)}\right) > 0,$$

and

$$\begin{aligned} h(\alpha_{t,j}, k) =& \left(\frac{\beta}{\beta - R_{t,j}}\right)^{\alpha_{t,j}} F_L(d_k; \alpha_{t,j}, \beta - R_{t,j}) \\ &+ \left[\frac{\beta}{(c_k - 1)R_{t,j} + \beta}\right]^{\alpha_{t,j}} e^{(c_k d_k)R_{t,j}} \\ &\times \left[F_L\left(\frac{CM_k}{1-c_k} + d_k; \alpha_{t,j}, (c_k-1)R_{t,j} + \beta\right) - F_L(d_k; \alpha_{t,j}, (c_k-1)R_{t,j} + \beta)\right] \\ &+ e^{(d_k + CM_k)R_{t,j}}\left[F_L(OM_k + d_k + CM_k; \alpha_{t,j}, \beta) - F_L\left(\frac{CM_k}{1-c_k} + d_k; \alpha_{t,j}, \beta\right)\right] \\ &+ \left(\frac{\beta}{\beta + R_{t,j}}\right)^{\alpha_{t,j}} e^{-OM_k R_{t,j}}[1 - F_L(OM_k + d_k + CM_k; \alpha_{t,j}, \beta + R_{t,j})]. \end{aligned}$$

Proof. Refer to Appendix B.3. □

Intuitively, if the cost difference between enrolling in the low and high coverage options is within the bounds, then a portion of the plan members are willing to choose the more expensive option for increased coverage. However, when the cost to upgrade to the high coverage option exceeds a certain upper bound, then at least one of the plan members intended by the actuary to select the high coverage option will no longer find it worthwhile to do so. Similarly, when the cost to upgrade to the high coverage option is low enough, then at least one of the plan members allocated to $G_{t,1}$ will choose the more expensive option instead.

Proposition 3. *Let the special case assumptions hold and assume that each plan members' true health statuses are known to both the actuary and the member. Then the closed-form solution to our deposit rate calculation formula is*

$$\boldsymbol{D}_t = \begin{cases} \boldsymbol{D}_t^{LB} = [D_{t,1}^{LB}, D_{t,2}^{LB}] & \text{if } D_{t,2}^* - D_{t,1}^* < LB_t \\ \boldsymbol{D}_t^* = [D_{t,1}^*, D_{t,2}^*] & \text{if } LB_t \leq D_{t,2}^* - D_{t,1}^* \leq UB_t \\ \boldsymbol{D}_t^{UB} = [D_{t,1}^{UB}, D_{t,2}^{UB}] & \text{if } UB_t < D_{t,2}^* - D_{t,1}^* \end{cases} \tag{3.1}$$

where

$$\begin{aligned} D_{t,k}^* =& \frac{\sum_{j \in G_{t,k}} EC(\alpha_{t,j}, k)}{n_k} \quad \forall k \in \{1, 2\}, \\ D_{t,1}^{LB} =& \frac{n_1 \sum_{j \in G_{t,1}} EC(\alpha_{t,j}, 1) + n_2 \sum_{j \in G_{t,2}} EC(\alpha_{t,j}, 2) - n_2^2 LB_t}{n_1^2 + n_2^2}, \\ D_{t,2}^{LB} =& LB_t + D_{t,1}^{LB}, \\ D_{t,1}^{UB} =& \frac{n_1 \sum_{j \in G_{t,1}} EC(\alpha_{t,j}, 1) + n_2 \sum_{j \in G_{t,2}} EC(\alpha_{t,j}, 2) - n_2^2 UB_t}{n_1^2 + n_2^2}, \\ D_{t,2}^{UB} =& UB_t + D_{t,1}^{UB}, \end{aligned}$$

and LB_t, UB_t, and $EC(\alpha_{t,j}, k)$ are as defined in the previous propositions.

Proof. Refer to Appendix B.4. □

Note that the solution $\boldsymbol{D}_t^*$ is the set of average expected claim amounts for the two options. Using this set of deposit rates means there is no built-in cross subsidy across the options, which is the most ideal scenario. However, if this solution is not feasible, meaning it does not yield the desired enrollment levels in each option, then this set of rates is adjusted to either $\boldsymbol{D}_t^{LB}$ or $\boldsymbol{D}_t^{UB}$. Specifically, the low coverage option is modified to a weighted average of the original rate, D_1^*, and an "appropriately distanced" rate of either $D_2^* - LB_t$ or $D_2^* - UB_t$, where the weights are proportional to the squared intended number of members in each group. The cost difference between the low and high coverage option rates are then adjusted to a specific amount—LB_t or UB_t.

We therefore obtain a closed-form solution to the approximate pricing formula stated in Section 3.2.2 by replacing $\alpha_{t,j} = \mu_{t,j}\beta$ with $\widetilde{\alpha}_{t,j} = \mathbb{E}[\widetilde{\mu}_{t,j}]\beta$, $\forall\, t, j$ in Equation (3.1).

3.2.4 Simplification of Our Proposed Approximate Formula when $K \geq 3$

While it is tedious to obtain a closed form solution when there are are more than two options, under the special case assumptions we can still efficiently simulate a large number of potential flex plan dynamics by treating our constrained optimization problem as a quadratic optimization problem. That is, if each plan member's true health status is known to both the actuary and the plan member, we can simplify the objective function and rewrite it in matrix form as follows:

$$\frac{1}{2}\boldsymbol{D_t}^T\boldsymbol{Q}\boldsymbol{D_t} + \boldsymbol{v_t}^T\boldsymbol{D_t}, \tag{3.2}$$

where

- $\boldsymbol{Q}$ is a $K \times K$ diagonal matrix with entry $q_{k,k} = 2n_k^2$,
- $\boldsymbol{v}$ is a size K column vector with entry $v_k = -2n_k \sum_{j \in G_{t,k}} EC(\alpha_{t,j}, k)$, and
- n_k is the number of plan members enrolled in option k, as defined in Section 2.4.2.

Extending the derivations in Appendix B.3, we can write the constraint in the form

$$\boldsymbol{A}\,\boldsymbol{D_t} \leq \boldsymbol{b_t},$$

where $\boldsymbol{A}$ is a $\binom{K}{2} \times K$ matrix and $\boldsymbol{b}_t$ is a column vector of length $\binom{K}{2}$. In the case of three options, we have

$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \text{ and } \boldsymbol{b}_t = \begin{bmatrix} -LB_{2,1,t} \\ UB_{2,1,t} \\ -LB_{3,1,t} \\ UB_{3,1,t} \\ -LB_{3,2,t} \\ UB_{3,2,t} \end{bmatrix},$$

where $\forall\, k_2 > k_1 \in \{1, 2, ..., K\}$,

$$LB_{k_2,k_1,t} = \max_{j \in G_{k_1}} \frac{1}{(1-s)R} \ln\left(\frac{h(\alpha_{t,j}, k_1)}{h(\alpha_{t,j}, k_2)}\right) > 0, \text{ and} \tag{3.3}$$

$$UB_{k_2,k_1,t} = \min_{j \in G_{k_2}} \frac{1}{(1-s)R} \ln\left(\frac{h(\alpha_{t,j}, k_1)}{h(\alpha_{t,j}, k_2)}\right) > 0. \tag{3.4}$$

The approximate pricing formula stated in Section 3.2.2 can then be simplified in the three options case by replacing $\alpha_{t,j} = \mu_{t,j}\beta$ with $\widetilde{\alpha}_{t,j} = \mathbb{E}[\widetilde{\mu}_{t,j}]\beta$, $\forall\, t, j$, in Equations (3.2), (3.3) and (3.4).

3.2.5 Calculating Manual Rates Using Our Proposed Pricing Method

In addition to the realistic information setting where neither the actuary nor the plan members know the true health statuses, we also simulate the flex plan dynamics under other information sets to evaluate the effect of available information on the flex plan evolution.[24] One such information set used is the manual rate setting. Using the manual rate as deposit rates refers to the case where the UKF is not applied to track each plan member's health status. That is, the actuary would use the unconditional density of $\mu_{t,j}^{(unc)} = e^{\beta_{t,0}+\beta_{t,1}H_{t,j}^{(unc)}}$ in place of $\widetilde{\mu}_{t,j}$ in the pricing formula:

$$
\begin{aligned}
&\mathbb{E}\left[1_{\{S_{t,j}(\boldsymbol{D}_t,\overline{\mu}_{t,j})=k\}}\left(D_{t,k}-r(L_{t,j},k)\right)\right]\\
&=\int_{\mathbb{R}+}\int_{\mathbb{R}+}\left[1_{\{S_{t,j}(\boldsymbol{D}_t,u)=k\}}\left(D_{t,k}-r(l,k)\right)\right]\times f_{L_{t,j},\overline{\mu}_{t,j}}(l,u)\,du\,dl\\
&=\int_{\mathbb{R}+}\int_{\mathbb{R}+}\int_{\mathbb{R}+}\left[1_{\{S_{t,j}(\boldsymbol{D}_t,u)=k\}}\left(D_{t,k}-r(l,k)\right)\right]\times f_{L_{t,j},\overline{\mu}_{t,j}}(l,u|\mu_{t,j}=v)f_{\mu_{t,j}^{(unc)}}(v)\,dv\,du\,dl\\
&=\int_{\mathbb{R}_+^3}\left[1_{\{S_{t,j}(\boldsymbol{D}_t,u)=k\}}\left(D_{t,k}-r(l,k)\right)\right]\times f_{L_{t,j}}(l|\mu_{t,j}=v)f_{\overline{\mu}_{t,j}}(u|\mu_{t,j}=v)f_{\mu_{t,j}^{(unc)}}(v)\,dv\,du\,dl.
\end{aligned}
$$

With the approximation explained in the Section 3.2.2, this means that the unconditional expected losses, $\mathbb{E}[\mu_{t,j}^{(\text{unc})}]$ is used in place of $\mathbb{E}[\widetilde{\mu}_{t,j}]$ to calculate the deposit rates, so that the optimization problem in our simulations becomes:

$$
\min_{\boldsymbol{D_t}}\sum_{k=1}^{K}\left(\sum_{j=1}^{N}1_{\left\{S_{t,j}\left(\boldsymbol{D}_t,\mathbb{E}\left[\mu_{t,j}^{(\text{unc})}\right]\right)=k\right\}}\left(D_{t,k}-\mathbb{E}\left[r(L_{t,j},k)\,\middle|\,\mathbb{E}[\mu_{t,j}^{(\text{unc})}]\right]\right)\right)^2, \tag{3.5}
$$

$$
\text{subject to } S_{t,j}\left(\boldsymbol{D}_t,\mathbb{E}\left[\mu_{t,j}^{(\text{unc})}\right]\right)=k,\quad \forall j\in G_{t,k},\quad \forall k\in\{1,2,...,K\}.
$$

Note that if the group has a stationary population, that is, when the unordered set of unconditional expected losses, $\left\{\mathbb{E}\left[\mu_{t,j}^{(\text{unc})}\right]\right\}_{j=1}^{N}$, are static through time, then the resulting manual rates are constant through time as well.

[24] We list all the simulated information sets in the next chapter.

Chapter 4

Simulation Study: Key Findings

We evaluate the performance of flex plans based on how close the plan is to reaching the four goals related to option enrollment, ASO account balance, and deposit rates as listed in Chapter 1.

4.1 Simulated Results Under the Traditional Pricing Practice

Under the traditional pricing practice, we observe skyrocketing rates, especially for the high coverage option, which becomes obsolete within three years in all 10,000 of our simulated scenarios. Note that this result is consistent with the fate of Harvard University's employee flex plan mentioned in Chapter 1.

Figure 4.1 illustrates how the distribution of deposit rates evolve, increasing from between \$390 and \$477 in year 1 to between \$863 and \$1,456 in year 2 in the simulated scenarios. By year 3, the high coverage option is driven out in most of the scenarios. Of the 471 remaining cases, the deposit rate further increases to a range of \$1,422 to \$3,747, at which point no plan member remains in the high coverage option in any of our simulated scenarios. This adverse selection spiral is expected, as we assume that all plan members have an accurate enough estimate of their loss distribution and diligently select the option that maximizes their utilities.

Unsurprisingly, in all of our simulated scenarios, the accumulated account balance plunges as the deposit rates try but fail to catch up to the skyrocketing average claims in the first three years, and the balance stays negative thereafter. Clearly, the flex plan priced under the traditional practice is far from reaching any of the goals defined in Chapter 1.

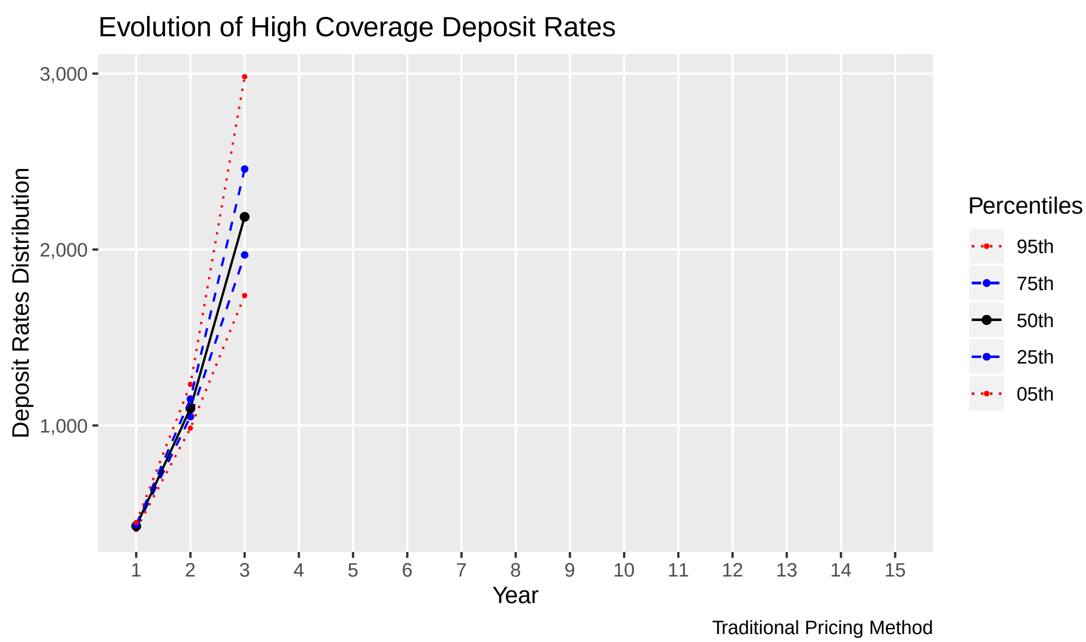

Figure 4.1: **Evolution of High Coverage Deposit Rates Under the Traditional Pricing Method.**

4.2 Simulated Results Under Our Proposed Pricing Method

Contrary to the flex plan dynamics observed when the traditional pricing practice is applied, our proposed pricing formula keeps the flex plan stable and sustainable, outperforming the plan priced under the traditional practice in terms of each of the four goals:

1. The first goal of having no empty enrollment in any option is satisfied each year in all 10,000 of our simulated scenarios under our proposed pricing method. This is an expected result; the constraint in our optimization problem ensures that if the actuary has an accurate estimate of the plan members' perceived health statuses, then n_k members enroll in option k. In fact, the distribution of option selection is stable, as illustrated in Figure 4.2. Note that we only show the low coverage option enrollment since it implies that the remaining plan members are enrolled in the high coverage option.

2. The distribution of the total ASO account balance is centred close to zero over time, as illustrated in Figure 4.3. Note that this second goal of having a year-end total ASO account balance as close to zero as possible each year means having a net annual change in total ASO account balance as close to zero as possible. While this objective is not the same as the one in our proposed optimization problem, they are highly related. Recall that our proposed objective in Equation (2.13) was constructed based on relaxing the optimal goal of having a zero balance in each option's ASO account. If

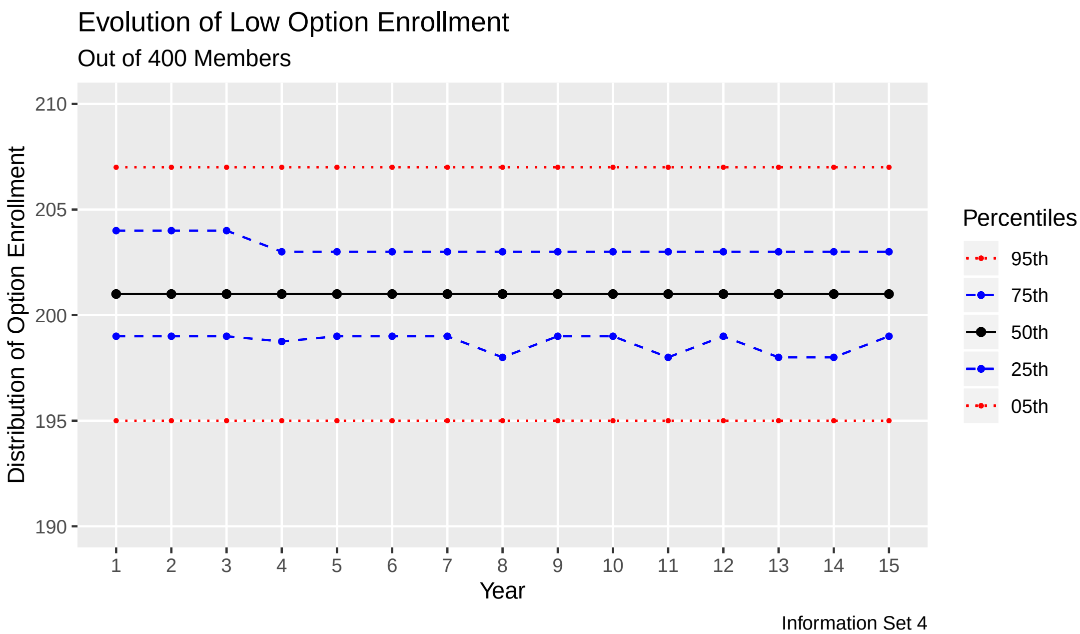

Figure 4.2: **Evolution of Low Coverage Option Enrollment Distribution Under the Proposed Pricing Formula.**

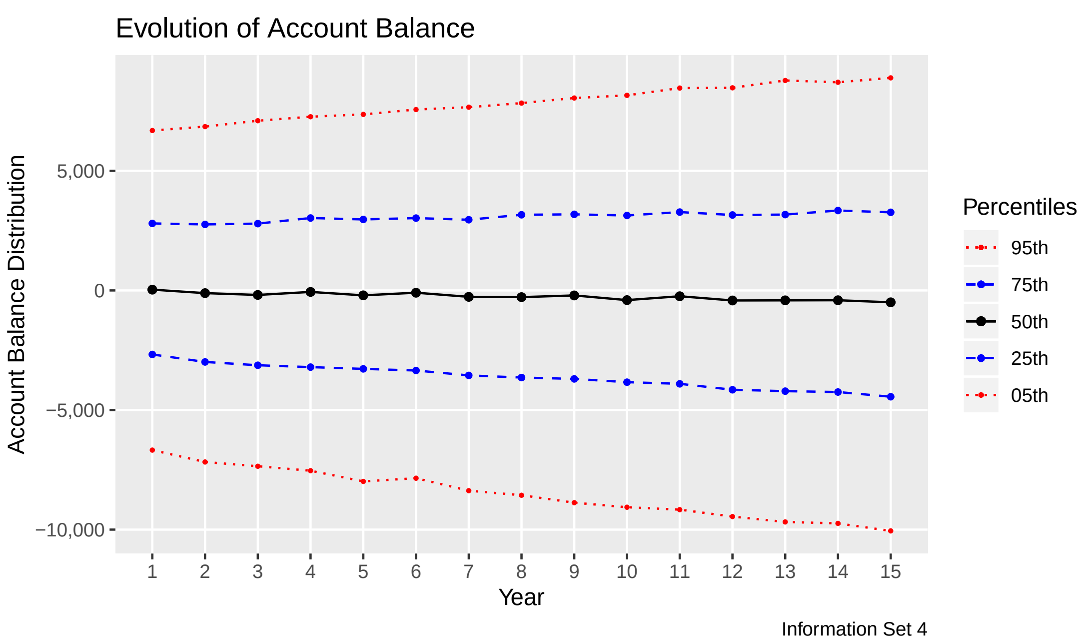

Figure 4.3: **Evolution of ASO Account Balance Distribution Under the Proposed Pricing Formula.**

each option's ASO account balance is zero, then the sum of account balances in each option is also zero. Likewise, if each option's squared ASO account balance is zero, then we also have a total ASO account balance of zero.

It turns out that having a total ASO account balance centred close to zero over time is a result of our equal allocation of plan members into the two options. Not only is it a natural choice, but it also leads to an expected total ASO account balance of exactly zero under the theoretical limiting case where both the actuary and the plan members know the true health statuses, as stated in Proposition 4 below. Under our realistic information setting, where the actuary and plan members can only estimate the true health statuses, this proposition acts as a useful guide.

Proposition 4. *Let the special case assumptions made in Propositions 1 to 3 hold and also assume that each plan members' true health statuses are known to both the actuary and the member. Then the expected total ASO account balance is zero if* $n_1 = n_2$.

Proof. Refer to Appendix B.5. □

Interestingly, pricing a flex plan using our proposed formula yields a cross subsidy between the two options each year in all 10,000 of our simulated scenarios. In other words, the case $LB_t \leq D^*_{t,2} - D^*_{t,1} \leq UB_t$ from Equation (3.1)—in which deposit rates equal the average of members' expected claim amounts in an option—is not satisfied in any of our simulations. Instead, in our simulations the low coverage deposit rates are between \$147 and \$221 above the average expected claim amount in that option, while the high coverage deposit rates are set between \$131 and \$207 below the average expected claim amount in that option. Figure 4.4 shows an example of the deposit rate and expected claims in one of the simulated paths of the flex plan. Note that this finding supports Cave's (1985) theoretical results regarding cross subsidy mentioned in Chapter 1.

3. In all of the scenarios simulated under our proposed pricing formula, the change in the ASO account balance through time is minimal in comparison with total annual deposits, which range between approximately \$80,000 and \$100,000 over the 15 years in each simulated path. As shown in Figure 4.5, in each year, over half of the 10,000 paths experience less than \$1,000 in year-to-year account balance change, and almost all paths experience less than \$3,000 change in account balance in any given year. The average of these annual changes in ASO account balance is also close to zero, ranging from -\$49 to -\$20 through the 15 simulated years. This is a direct result of our choice of n_1 and n_2, as explained in the second point above. Again, the fact that the range is close to but does not cover zero can be explained by the fact that the actuary and

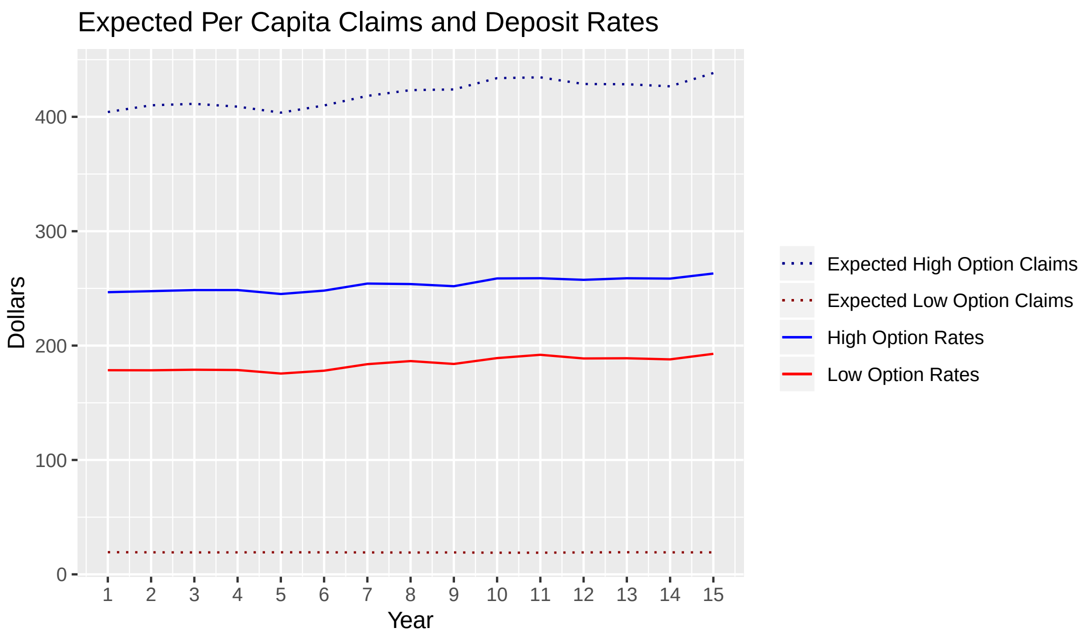

Figure 4.4: **Example of the Cross Subsidy Effect Observed in One Sample Path.**

plan members only have an estimate of the true health status.

In fact, even if we have $n_1 \neq n_2$, we still expect the annual change in total ASO account balance to be constant (albeit non-zero) through time.[25]

4. The deposit rates are stable from year to year. As shown in Figures 4.6 and 4.7, about half of the year-over-year rate changes are within 1% for both the high and low coverage options, and almost all of the year-over-year changes are within 3%.

   This observation is again not surprising; a careful examination of Equation (2.14) reveals that having constant expected deposit rates is just a by-product of our proposed pricing formula if the group's set of expected health costs is constant through time. Hypothetically, if the unordered set $\{\widetilde{\mu}_{t,j}\}_{j=1}^{N}$ is static through time, then the optimization problem and therefore the calculated deposit rates solution is also constant each year. Similar to our explanation in Point 3 above, although this theoretical condition cannot be achieved in a realistic setting (except when we are pricing with manual rates), our stationary population and static relationship between a health status and the expected loss yields a reasonable approximation. In our simulations, the distri-

[25]Recall that our population demographics is stationary and the relationship between health status and expected losses is static over time. Therefore, we expect the unordered set $\{\mathbb{E}[\widetilde{\mu}_{t,j}]\}_{j=1}^{N}$—which is used to approximate $\{\mu_{t,j}\}_{j=1}^{N}$—to be constant over the years. Hypothetically, if the unordered set $\{\mu_{t,j}\}_{j=1}^{N}$ is constant over time, then so is the unordered set $\{\alpha_{t,j}\}_{j=1}^{N}$, meaning the calculated values in Equations (B.2) and (B.3) are also constant through time.

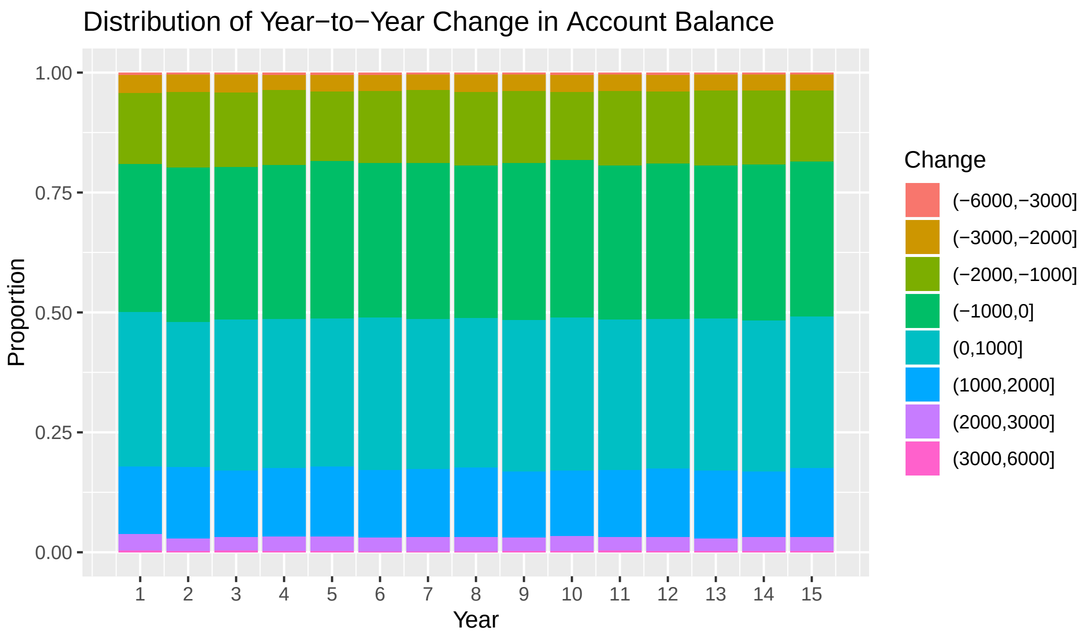

Figure 4.5: **Distribution of Year-to-Year Change in Account Balance Under the Proposed Pricing Formula.**

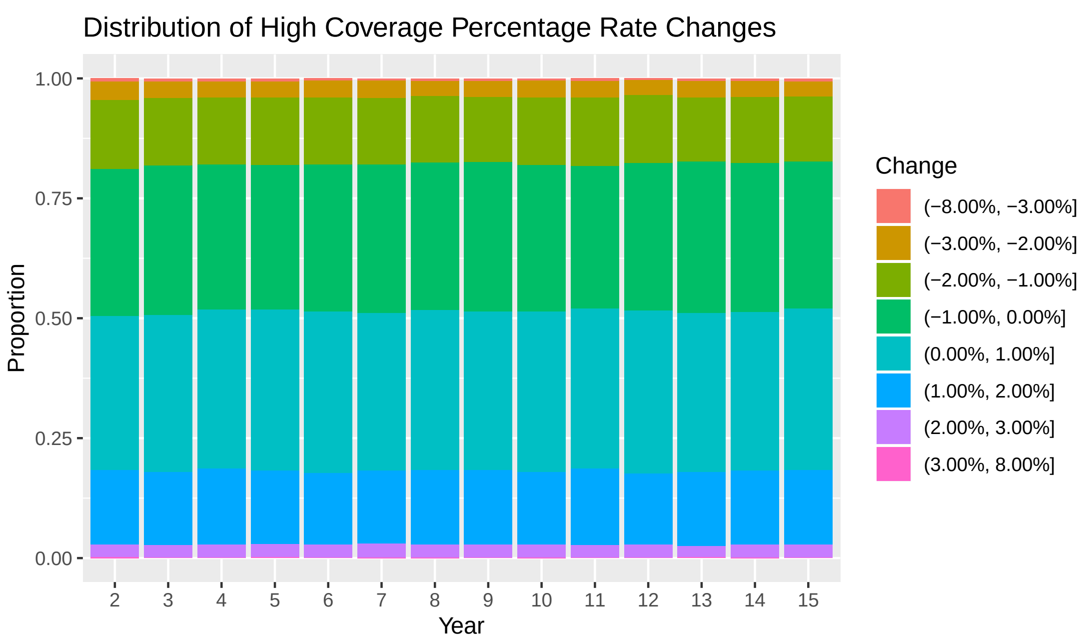

Figure 4.6: **Distribution of Year-to-Year Percent Change in High Coverage Deposit Rates.**

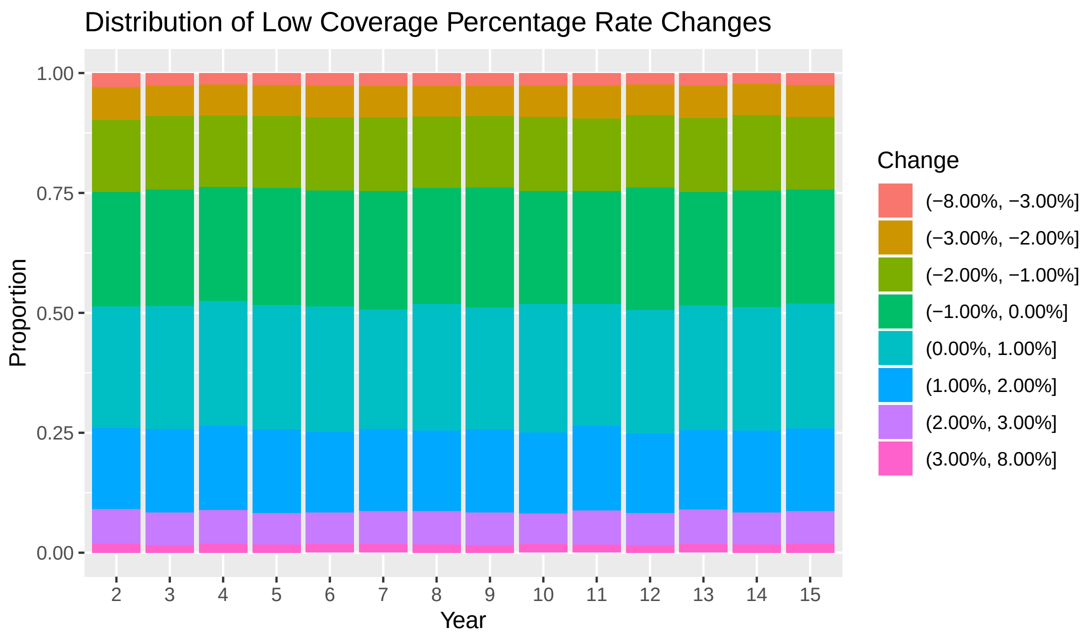

Figure 4.7: **Distribution of Year-to-Year Percent Change in Low Coverage Deposit Rates.**

butions $\{\widetilde{\mu}_{t,j}\}_{j=1}^{N}$ are approximated by their expectations, $\{\mathbb{E}[\widetilde{\mu}_{t,j}]\}_{j=1}^{N}$; therefore, our simulated year-to-year volatility in deposit rates is dictated by the year-to-year change in the set of estimated mean losses. Given our stationary population assumption, this volatility is small.

4.3 Effect of Information

We compare the flex plans under various information sets to examine the effect of information on the plan dynamics. Table 4.1 summarizes the different information sets used in our simulation study, and Figure 4.8 illustrates the relative level of information between the different settings. As indicated in the table, Information Set 4 is the realistic case, which we have shown results for in the previous section. The theoretical limiting case, which was referenced in the previous section, is Information Set 1.

4.3.1 Actuary's Information

A comparison of the flex plan evolution under Information Sets 2, 4, 5, and 6 can reveal the importance of the actuary's accuracy in estimating the health status distribution. We find that even if the actuary knows the true health status, the results only show a negligible improvement. For example, as shown in Figure 4.11, under Information Set 2, the variability in the final ASO account balance distribution only marginally improves when compared to that of Information Set 4. Another observation we make is that the lack of historical claims

Information Setting	Health Status Information Available to Actuary	Health Status Information Available to Plan Member j
Information Set 1 All Parties Know the True Health Statuses	$\{H_{t,j}\}_{j=1}^{N}$	$H_{t,j}$
Information Set 2 Actuary Knows the True Health Statuses	$\{H_{t,j}\}_{j=1}^{N}$	$\overline{H}_{t,j}$
Information Set 3 Plan Members Know Their True Health Statuses	$\{\widetilde{H}_{t,j}\}_{j=1}^{N}$ (based on $L_{-2:t-1,j}$)	$H_{t,j}$
Information Set 4 Realistic Base Case	$\{\widetilde{H}_{t,j}\}_{j=1}^{N}$ (based on $L_{-2:t-1,j}$)	$\overline{H}_{t,j}$
Information Set 5 Realistic Case with No Prior Claims Data	$\{\widetilde{H}_{t,j}\}_{j=1}^{N}$ (based on $L_{1:t-1,j}$)	$\overline{H}_{t,j}$
Information Set 6 Manual Rates	$\{H_{t,j}^{(unc)}\}_{j=1}^{N}$	$\overline{H}_{t,j}$

Table 4.1: **Description of Various Information Sets.**

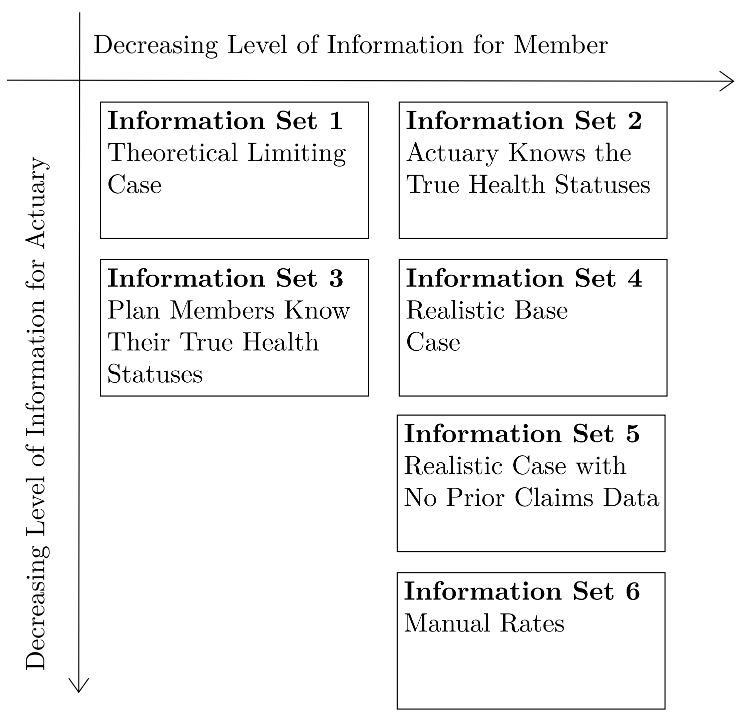

Figure 4.8: **Information Sets Ordered by Level of Information.**

data prior to the flex plan implementation does not have much of a negative impact on the plan performance. The only notable impact is a temporary instability in deposit rates in the first couple of years under Information Set 5 as the deposit rates adjust from manual rates of $185.24 and $259.15 in the first year to adapted rates based on observed losses. Nonetheless, as shown in Figures 4.9 and 4.10, the distribution stabilizes quickly.

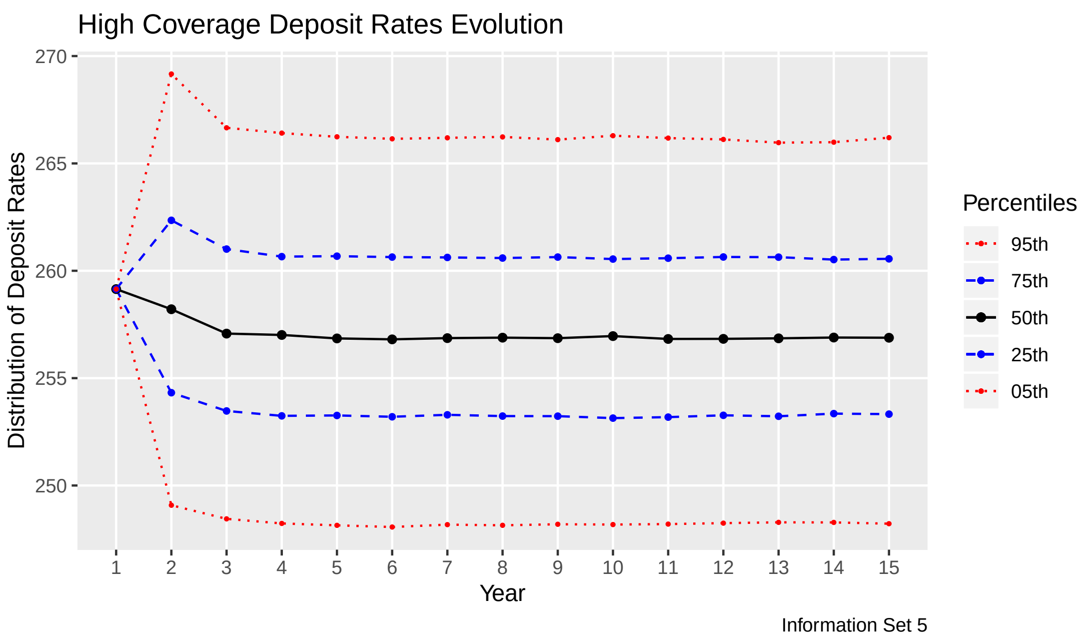

Figure 4.9: **Distribution of High Coverage Deposit Rates Under Information Set 5.**

In the interest of deposit rate stability, it may be tempting to use manual rates since they are constant through time given our assumptions; however, this deviation from our suggestion of tracking health statuses drastically increases the variance of the final ASO account balance, as already shown in the Figure 4.11. Plan sponsors likely do not view this as a sensible trade-off.

4.3.2 Members' Information

We also examined the effect of plan members' accuracy in predicting their health status by comparing Information Sets 3 and 4.[26] Our results show that such a change has negligible

[26]The case where the members' uncertainty about their health statuses is large is less relevant as the members need to have a good intuition about their health to create the potential for adverse selection and hence the spiral. We have confirmed via simulation that the adverse selection spiral does not occur under the traditional pricing method when plan members' uncertainty increases from $\delta = 1$ to $\delta = 10$.

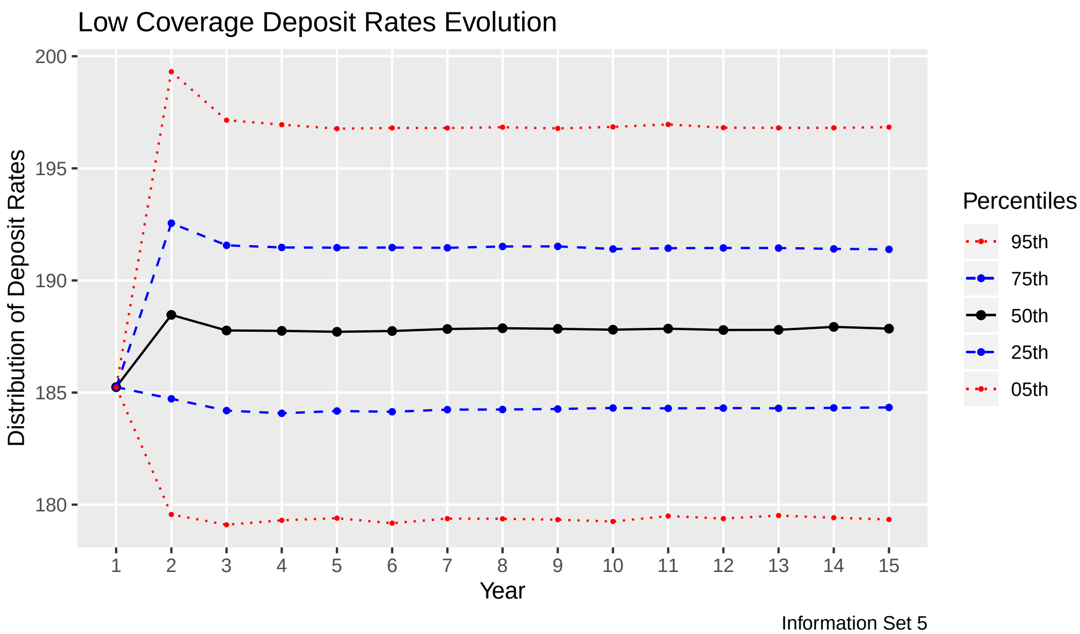

Figure 4.10: **Distribution of Low Coverage Deposit Rates Under Information Set 5.**

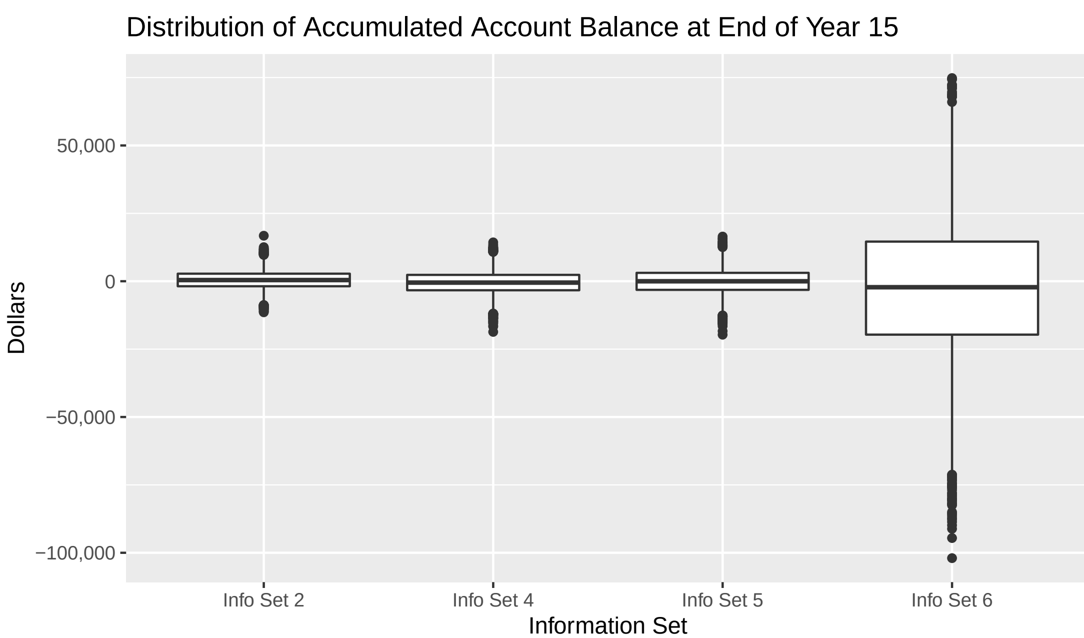

Figure 4.11: **Final ASO Account Balance at the End of Year 15 Under Varying Information Settings for the Actuary.**

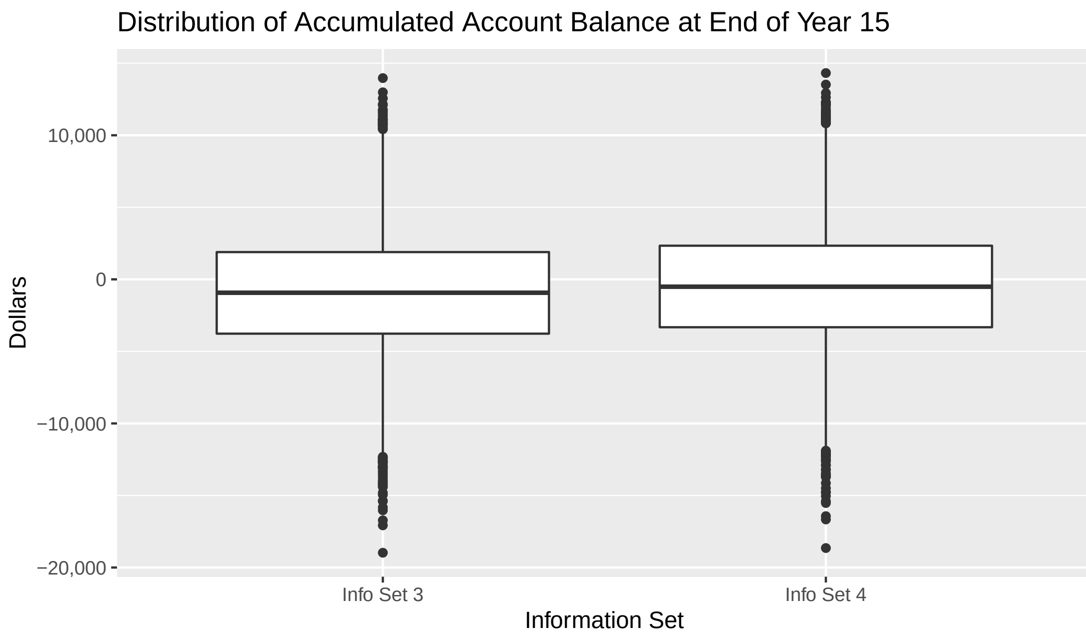

Figure 4.12: **Final ASO Account Balance at the End of Year 15 Under Varying Information Settings for the Members.**

impact on the flex plan dynamics. As shown in Figure 4.12, the final ASO account balance is slightly more negative in the case of Information Set 3, where plan members have the advantage of knowing their true health status in the upcoming year. Again, considering that the total annual deposits range between approximately $80,000 and $100,000, this small downward shift in the final ASO account balance distribution is likely immaterial from plan sponsors' perspective. Otherwise, the distribution of option enrollment is extremely similar to that of the realistic information setting (i.e., Information Set 4), and the deposit rates are unaffected by the change in plan members' knowledge of their health status since the pricing formula depends solely on the information known to the actuary.

4.3.3 Theoretical Limiting Case: Perfect Information Setting

In addition, we test the flex plan under Information Set 1 to examine whether the simulated results align with the theoretical derivations. For example, the cross subsidy observed under the realistic information setting (i.e., Information Set 4) was such that the difference between the deposit rates and expected claims in the low coverage option approximately offsets that of the high coverage option. Under Information Set 1, we observe an exact offset, with the expected ASO sub-account balance ranging from $145.40 to $196.17 in the low coverage option and from -$196.17 and -$145.40 in the high coverage option.

Also, as expected, we observe the lowest variance in option enrollment in this theoretical limiting case, where both the actuary and the plan members know the true health statuses. Figure 4.13 compares the low option enrollment under various information sets. Note that due to rounding, there are occasional cases where the low option enrollment level is not exactly $n_1 = 200$, even in this perfect information setting.[27] Nonetheless, we observe in Figure 4.14 that the ASO balance distribution is not sensitive to this rounding detail under any of the main information sets. Note that this insensitivity to the small deviation from the optimal setting of $n_1 = n_2$ is an important and desirable property, since in practice we cannot ensure that a group size is always divisible by the number of options.

Finally, under Information Set 1, 93% and 81% of the high and low coverage year-to-year deposit rate changes are within 2%, respectively. Compared to the realistic information setting, where 90% and 78% of the high and low coverage year-to-year deposit rate changes are within 2%, we find only a slight improvement in deposit rate stability. In conclusion, the flex plan performs well enough in the realistic information setting; improved accuracy about health statuses does not have any significant impact on any of the plan dynamics.

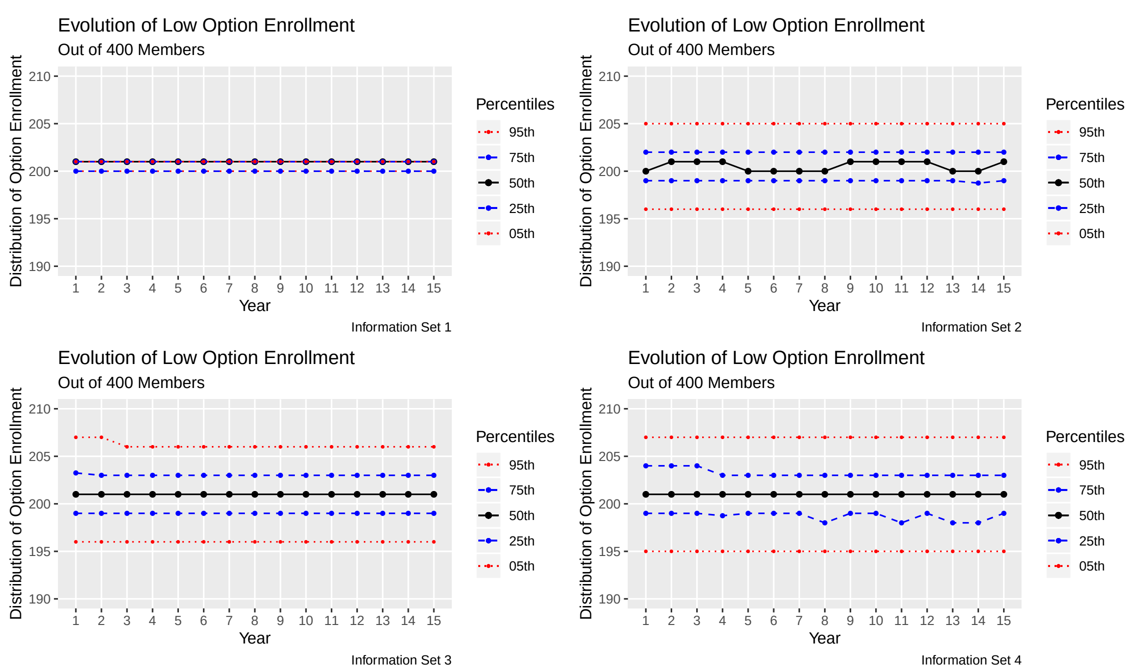

Figure 4.13: **Evolution of Low Coverage Option Enrollment Distribution Under Various Information Settings.**

[27]For example, there may exist two plan members whose rounded expected annual losses are both equal to the median, in which case the two options will have 201 and 199 enrollees.

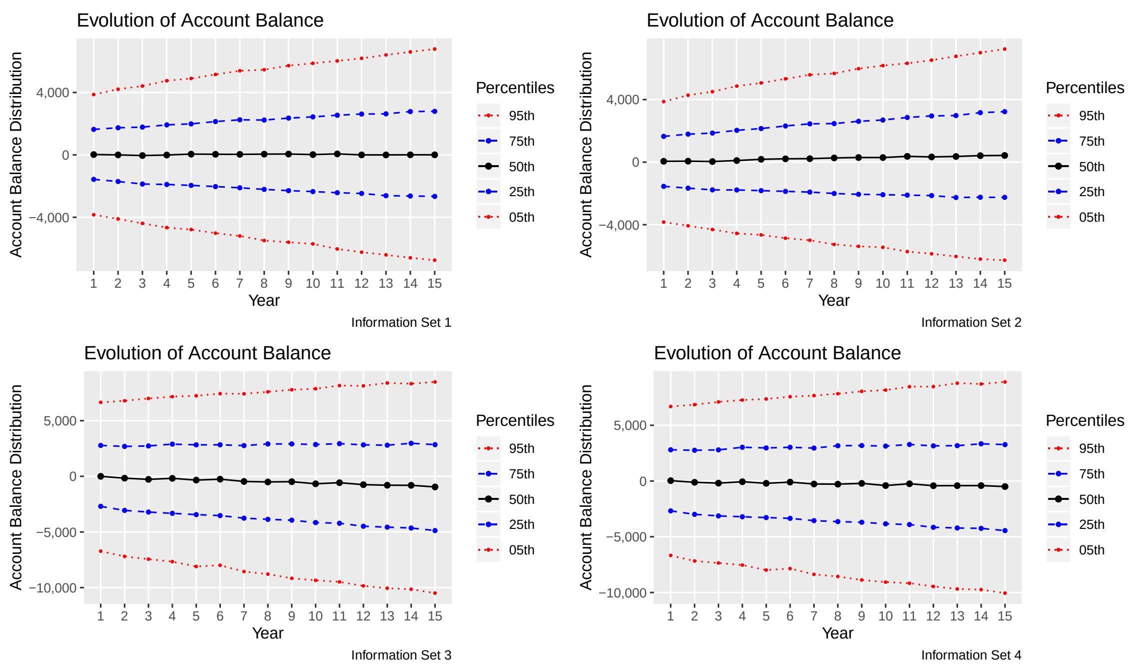

Figure 4.14: **Evolution of Account Balances Under Various Information Settings.**

4.4 Robustness Test Results

4.4.1 General Observations: Similar to the Base Case

In addition to investigating the effect of information on the plan performance, we also examine the robustness of our proposed formula by revising some parameters according to the nine alternative settings outlined in Section 3.1.2. We find our proposed formula to be robust in that no adverse selection spiral is observed in any of the simulated flex plan settings when our pricing method is applied. In all of our simulated scenarios, both the enrollment levels and deposit rates are stable through time. The ASO account balance is also stable and centred around zero as long as appropriate anticipated enrollment figures (i.e., $\{n_k\}_{k=1}^{K}$) are used in setting deposit rates, and the plan offers the typical premium subsidy arrangement of flex credits instead of a level percent premium subsidy.

Most of our observations from the robustness tests are extremely similar to what we have already shown in the graphs for the base case. For example, when the coinsurance levels are adjusted, the plan dynamics behave almost identically to that of the base case, except the average deposit rates reflect the change in coverage in each option. A change in group size also results in findings similar to those shown in previous sections for the base case. Although an increase in group size from 400 to 600 leads to a higher variability in the annual change in total ASO account balance per dollar of annual deposits, the variability is still within reason. Similarly, the deposit rate is slightly more volatile when the group size is decreased from

400 to 200, but still acceptable. Table 4.2 outlines some figures to show these small impacts.

		Group Size		
		$N = 200$	$N = 400$	$N = 600$
Range of Net Annual Change in ASO Account Balance, as a Proportion of Annual Deposits in the Simulated Scenarios		-3.6% to 3.6%	-5.9% to 5.3%	-7.9% to 7.3%
Proportion of Year-to-Year Rate Changes that are Within 2% in the Simulated Scenarios	Low Coverage Option	66%	82%	90%
	High Coverage Option	81%	93%	98%

Table 4.2: **Impact of Group Size on Various Results.**

4.4.2 Immaterial Constant Drift in ASO Account Balance

The only notable deviation from the base case dynamics arises when there is a change in subsidy arrangement or when the intended enrollment levels are not equal in each option. In these cases, even though we do not observe an adverse selection spiral, the ASO account balance drifts more significantly.

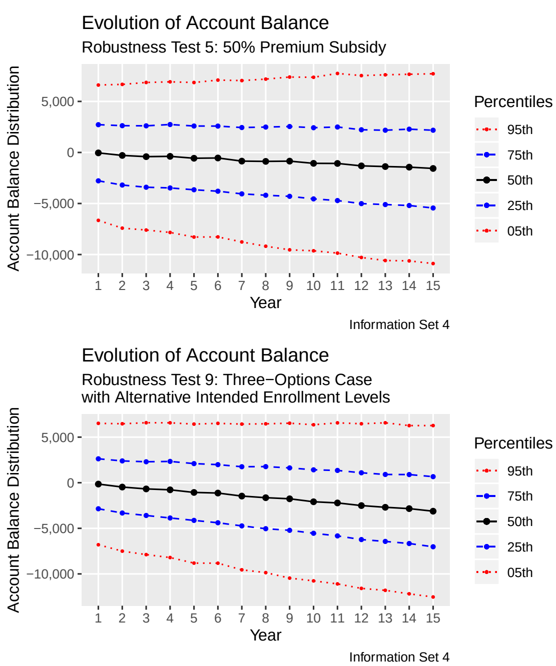

Figure 4.15: **Evolution of Account Balances Under Robustness Tests 5, 6, and 9.**

The three-options plan with slightly unequal intended enrollment levels (i.e., Robustness Test 9, where $n_1 = 134$ and $n_2 = n_3 = 133$) also experiences a similar magnitude of linear drift in the total ASO account balance under both Information Sets 1 and 4. As this drift is absent when the intended enrollment levels are equal in each of the three options (i.e., Robustness Test 2, where $n_1 = n_2 = n_3 = 120$), our conjecture is that the drift amount is sensitive to a small but consistent deviation from the equal intended enrollment levels condition. Nonetheless, the magnitude of this detectable drift is immaterial compared to the annual deposit amounts. Also, the distribution of final account balance (i.e., at the end of year 15) still covers a range of positive and negative values.

In short, we find that with appropriate intended enrollment levels, any detectable drift in the ASO account balance distribution caused by rounding is immaterial. Figure 4.15 illustrates the evolution of account balances in Robustness Tests 5, 6, and 9.

4.4.3 Significant Constant Drift in ASO Account Balance

We therefore run Robustness Tests 7 and 8 to evaluate the impact of choosing inappropriate intended enrollment levels, based on the two-options plan. Under the more extreme setting of $n_1 = 300$ and $n_2 = 100$, while the option enrollment and deposit rates are still stable, the drift in the ASO account balance can be material over time. This is an expected outcome; again using Equations (B.2) and (B.3) as an approximation for the ASO account balance drift under the realistic information setting, we see that this choice of n_1 and n_2 drastically increases the coefficient and therefore the drift by a factor of approximately 60 when compared to our base case. In fact, in contrast to all the other robustness test results, these drifts are so significant that the simulated final ASO account balances are either always positive or always negative, as illustrated in Figure 4.16.

Recall that in practice, plan sponsors make lumpsum deposits or withdrawals to keep the ASO account balance at a suitable level. Therefore, although it is not ideal to have a significant linear drift in the ASO account balance, plan sponsors may still view the regular account withdrawals or required replenishments as acceptable for budgeting purposes, as they are predictable and expected to be constant each year. The alternative solution is to eliminate or minimize such a drift by finding and implementing the optimal intended enrollment figures for each option when calculating deposit rates, as is the case in all the other robustness tests.

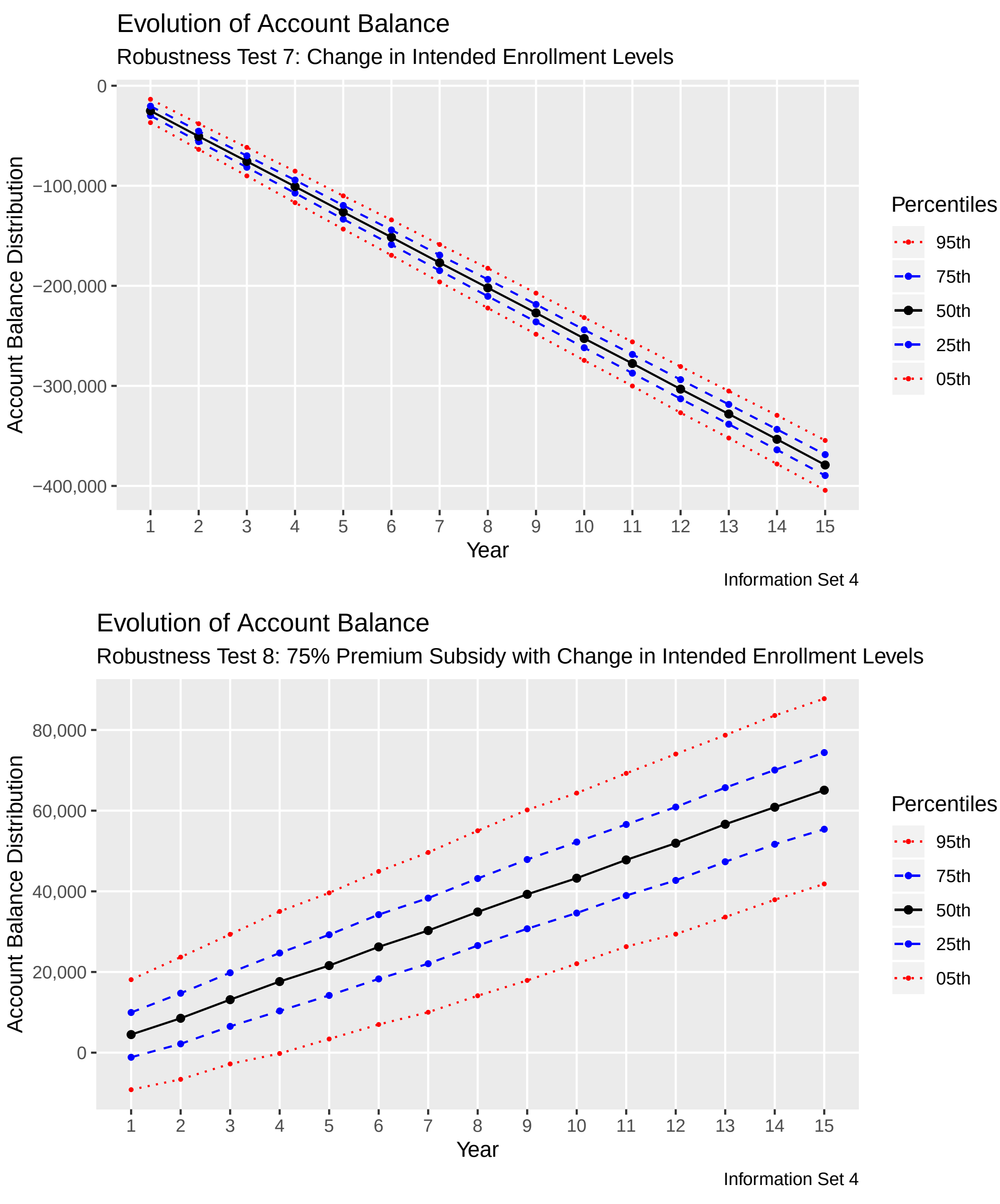

Figure 4.16: **Evolution of ASO Account Balance Under Alternative Split of Option Allocation.**

Chapter 5

Conclusions and Suggestions for Further Research

Over the years, a number of solutions have been suggested or implemented to prevent adverse selection in flex plans. Unfortunately, even though some of these fixes can be effective, none of them are ideal or practical in the flexible group benefits context. A more suitable solution to prevent the adverse selection spiral is through pricing. We therefore propose a comprehensive flexible group benefits framework with a novel pricing method that incorporates a claim model and an option selection model. We present both theoretical derivations and simulation study results to show that flex plans which are prone to the adverse selection spiral under the traditional pricing practice can become sustainable if they are priced with our proposed pricing formula instead. Specifically, our proposed method outperforms the traditional method by keeping all plan options available, maintaining the year-end total ASO account balance as close to zero and as stable as possible, and keeping deposit rates as stable as possible through time.

In addition to constructing an effective and practical pricing formula, our project contributes to the group benefits industry by proposing a comprehensive framework, which allows actuaries to better understand the dynamics and relationships between each component of a flex plan. An interesting finding from our simulation study is that the ideal set of deposit rates never equals the set of expected claims in each option, even though the total annual deposits equal the total expected annual claims in a plan. In other words, a portion of the deposits for an option subsidizes claims in another option, which is consistent with Cave's (1985) theoretical results.[28] This suggests that having pure deposits rates equal to the expected claims is at least difficult to achieve. It will therefore be helpful in future research to further

[28] In our context and given our simulation assumptions, this means that the case $LB_t \leq D^*_{t,2} - D^*_{t,1} \leq UB_t$ is never satisfied in any of our simulated scenarios.

analyze the relationship between the bounds and D^* to better understand this phenomenon.

We have also shown that maintaining a sustainable ASO account balance requires choosing a suitable set of intended enrollment levels for each plan option. Given the plan assumptions in our study, we theoretically derived and showed through simulations that the ideal scenario of zero drift in ASO account balance is achieved when $n_1 = n_2$ in the two-options setting. In future research, it will be interesting to derive a closed-form solution for deposit rates and to confirm our conjecture that for the case of three or more options, it is still optimal to choose equal intended enrollment levels in each option.

Other suggestions for further research are outlined below.

1. Extend the models in our proposed framework to account for more general flex plans. For example:
    - Having multiple lines of benefits with different reimbursement functions. This will require that each plan member has multiple correlated health statuses for each line of benefit.
    - Allowing for dependents' coverage. A simple way of incorporating family or couples coverage in addition to single coverage is to include a family status factor in each individual's health status.
    - Allowing for the group size or group demographics to change each year.
2. Refine the proposed models. For example:
    - Accounting for moral hazard in the claim model.
    - Further dividing the risk classification groups in the health status model to include more variables in addition to age, such as gender and geographical location. The health deterioration rate, λ, may also vary with age and through time; i.e., an x year old individual's health deterioration rate in year t can be $\lambda_{t,x}$.
    - Accounting for biased intuitions in each plan member's estimate of their own health status (i.e., whether particular individuals tend to be more optimistic or pessimistic about their estimated health status), so that the mean of $\nu_{t,j}$ is nonzero. The accuracy of intuition, δ, can also depend on the individual and through time, so that we have $\delta_{t,j}$ as the variance of $\nu_{t,j}$ in Equation (2.6).
3. Explore alternative models within our framework. For example:
    - Using an alternative objective function in the constrained optimization problem. Examples include:

- Minimizing the squared total ASO balance (which is the plan sponsor's goal) instead of the sum of squared individual sub-account balances (which is constructed based on relaxing the more ambitious goal of satisfying the actuarial equivalence principle in every option).
- Minimizing the squared expected difference between the actual and target loss ratios in each option. Assuming no loading factor, this means the following objective function:

$$\sum_{k=1}^{K} \left(\mathbb{E} \left[\left(\frac{\sum_{j=1}^{N} 1_{\{S_{t,j}(\boldsymbol{D}_t, \overline{\mu}_{t,j})=k\}} r(L_{t,j}, k)}{\sum_{j=1}^{N} 1_{\{S_{t,j}(\boldsymbol{D}_t, \overline{\mu}_{t,j})=k\}} D_{t,k}} \right) - 1 \,\middle|\, L_{1:t-1,j} \right] \right)^2 .$$

- Changing the loss distribution, e.g., to a Tweedie distribution.
- Changing the utility function, e.g., to one with a constant relative risk aversion level. That is,

$$u_{t,j} = \frac{[w_{t,j} - (1-s)D_{t,k} + F + r(L_{t,j}) - L_{t,j}]^{1-R_{t,j}}}{1 - R_{t,j}}$$

  where $R_{t,j} > 1$.
- Using an alternative option selection model. Examples include:
  - Introducing a random component to our proposed model. Appendix C outlines how the model can be extended.
  - Applying another decision theory, such as prospect theory (Kahneman and Tversky, 1979).

Bibliography

Babcock, B. A., E. K. Choi, and E. Feinerman. 1993. Risk and probability premiums for CARA utility functions. *Journal of Agricultural and Resource Economics* 18:17–24.

Bajari, P., H. Hong, and A. Khwaja. 2006. Moral hazard, adverse selection and health expenditures: A semiparametric analysis. Tech. rep., National Bureau of Economic Research.

Barber, A. E., R. B. Dunham, and R. A. Formisano. 1992. The impact of flexible benefits on employee satisfaction: A field study. *Personnel Psychology* 45:55–74.

Carlin, C., and R. Town. 2009. Adverse selection, welfare, and optimal pricing of employer sponsored health plans. *Working Paper*.

Cave, J. A. 1985. Subsidy equilibrium and multiple-option insurance markets. *Advances in Health Economics and Health Services Research* 6:27.

Creal, D. 2012. A survey of sequential Monte Carlo methods for economics and finance. *Econometric reviews* 31:245–296.

Cutler, D. M., and R. J. Zeckhauser. 1998. Adverse selection in health insurance. In *Forum for Health Economics & Policy*, vol. 1. De Gruyter.

French, M. 2008. Flex and match. URL `https://www.benefitscanada.com/benefits/health-benefits/flex-and-match-8332`. [Date accessed: 09/03/2020].

Fuhrer, C. S., and A. F. Shapiro. 1992. Modeling flexible benefit selection. *Transactions of Society of Actuaries* 44:135–160.

Garnick, D., A. Hendricks, K. Thorpe, J. Newhouse, K. Donelan, and B. Blendon. 1993. How well do Americans understand their health coverage? *Health Affairs* 12:204–212.

Julier, S. J., and J. K. Uhlmann. 1997. New extension of the Kalman filter to nonlinear systems. In *Signal processing, sensor fusion, and target recognition VI*, vol. 3068, pp. 182–193. International Society for Optics and Photonics.

Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47:263–292.

Maciejewski, M. L., B. Dowd, and H. O'Connor. 2004. Multiple prior years of health expenditures and Medicare health plan choice. *International Journal of Health Care Finance and Economics* 4:247–261.

Marquis, M. S. 1992. Adverse selection with a multiple choice among health insurance plans: a simulation analysis. *Journal of Health Economics* 11:129–151.

Marquis, M. S., and J. L. Buchanan. 1999. Simulating the effects of employer contributions on adverse selection and health plan choice. *Health Services Research* 34:813–837.

McKay, R. J. (ed.). 2007. *Canadian handbook of flexible benefits*. John Wiley & Sons.

Mehta, N., J. Ni, K. Srinivasan, and B. Sun. 2017. A dynamic model of health insurance choices and healthcare consumption decisions. *Marketing Science* 36:338–360.

O'Neill, A. 2012. *Group benefits plan management*. The Institute for Advanced Financial Education.

Skwire, D. D. 2016. *Group insurance*. Actex Publications.

Sturman, M. C., and J. W. Boudreau. 1994. Out-of-pocket costs and the flexible benefits decision: Do employees make effective health care choices? *Working Paper*.

van Winssen, K., R. van Kleef, and W. van de Ven. 2018. Can premium differentiation counteract adverse selection in the Dutch supplementary health insurance? A simulation study. *European Journal of Health Economics* 19:757–768.

Wan, E. A., and R. Van Der Merwe. 2000. The unscented Kalman filter for nonlinear estimation. In *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*, pp. 153–158.

Appendix A

Proofs Related to Tracking Health Statuses

A.1 Deriving the Conditional Density Recursion Formula Using Bayes' Rule

The following proof adapts Creal's (2012) derivations to our notation in this report. For ease of readability, the time index has been moved forward by one year compared to Equation (2.8), and subscripts on the density functions have been omitted in the intermediate steps:

$$
\begin{aligned}
& f_{H_{t,j}|L_{1:t,j}}(h_{t,j}|L_{1:t,j}) \\
=& \frac{f(L_{1:t,j}, h_{t,j})}{f(L_{1:t,j})} \\
=& \frac{f(L_{t,j}|h_{t,j}, L_{1:t-1,j}) f(h_{t,j}|L_{1:t-1,j}) f(L_{1:t-1,j})}{f(L_{1:t,j})} \\
=& \frac{f(L_{t,j}|h_{t,j}, L_{1:t-1,j}) f(h_{t,j}|L_{1:t-1,j})}{\frac{f(L_{1:t,j})}{f(L_{1:t-1,j})}} \\
=& \frac{f(L_{t,j}|h_{t,j}) f(h_{t,j}|L_{1:t-1,j})}{f(L_{t,j}|L_{1:t-1,j})} \\
=& \frac{f(L_{t,j}|h_{t,j}) f(h_{t,j}|L_{1:t-1,j})}{\int_{\mathbb{R}} f(L_{t,j}|h_{t,j}, L_{1:t-1,j}) f(h_{t,j}|L_{1:t-1,j}) dh_{t,j}} \\
=& \frac{f(L_{t,j}|h_{t,j}) \int_{\mathbb{R}} f(h_{t,j}|h_{t-1,j}) f(h_{t-1,j}|L_{1:t-1,j}) dh_{t-1,j}}{\int_{\mathbb{R}} f(L_{t,j}|h_{t,j}, L_{1:t-1,j}) \int_{\mathbb{R}} f(h_{t,j}|h_{t-1,j}) f(h_{t-1,j}|L_{1:t-1,j}) dh_{t-1,j} dh_{t,j}} \\
=& \frac{\int_{\mathbb{R}} f_{L_{t,j}|H_{t,j}}(L_{t,j}|h_{t,j}) f_{H_t|H_{t-1}}(h_{t,j}|h_{t-1,j}) f_{H_{t-1,j}|L_{1:t-1,j}}(h_{t-1,j}|L_{1:t-1,j}) dh_{t-1,j}}{\int_{\mathbb{R}} \int_{\mathbb{R}} f_{L_{t,j}|H_{t,j}}(L_{t,j}|h_{t,j}) f_{H_t|H_{t-1}}(h_{t,j}|h_{t-1,j}) f_{H_{t-1,j}|L_{1:t-1,j}}(h_{t-1,j}|L_{1:t-1,j}) dh_{t,j} dh_{t-1,j}}.
\end{aligned}
$$

A.2 Unscented Kalman Filter

The unscented Kalman filter is an algorithm that tracks latent variables, given signals observed in discrete time. The problem involves a nonlinear discrete-time dynamics system, which consists of:

1. The dynamics of a latent variable ($H_{t,j}$ in our case), written as a function of the previous state of this latent variable ($H_{t-1,j}$) and the process noise ($\epsilon_{t,j}$), and
2. The measurement function, which is the observed signal ($L_{t,j}$) written as a function of the latent variable and a measurement noise ($z_{t,j}$).

Our latent variable dynamics is stated in Equation (2.3). To rewrite the distribution of observed losses as a measurement function, we introduce a random variable, $z_{t,j}$, which follows a standard normal distribution. Our measurement function is therefore

$$L_{t,j} = f_{measurement}(H_{t,j}, z_{t,j}) = f_L^{-1}(\Phi(z_{t,j}); g(H_{t,j}))$$

where

- $f_L^{-1}(\cdot)$ is the inverse cumulative distribution function of the annual loss,
- $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution, and
- $g(\cdot)$ is the link function that transforms the health status into the mean parameter, $\mu_{t,j}$.

We restate the steps for one iteration of the unscented Kalman filter algorithm as outlined in Wan and Van Der Merwe (2000), but adapted to our context where appropriate:

1. Begin the iteration with the previous mean and variance estimates (i.e., $\tilde{h}_{t-1,j}$ and $\rho^2_{t-1,j}$, respectively) of the approximated health status distribution, as well as the new observed signal (i.e., $L_{t,j}$).
2. Construct a column vector $\hat{\boldsymbol{x}}^a_{t-1}$ of the means of the previous latent variable estimate, the process noise, and the measurement noise. Moreover, create a diagonal matrix $\boldsymbol{P}^a_{t-1}$ of the variances of the estimated latent variable distribution, process noise, and measurement noise. Given that our process noise is $\epsilon_{t,j} \sim \mathcal{N}(0, \sigma)$ and our measurement noise is $z_{t,j} \sim \mathcal{N}(0, 1)$, we have

$$\hat{\boldsymbol{x}}^a_{t-1} = \begin{bmatrix} \tilde{h}_{t-1,j} \\ 0 \\ 0 \end{bmatrix} \text{ and } \boldsymbol{P}^a_{t-1} = \begin{bmatrix} \rho^2_{t-1,j} & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

3. Calculate the previous year's sigma points, which are representative points in the approximated distributions, by creating a matrix

$$\boldsymbol{\chi}^a_{t-1} = \left[\hat{\boldsymbol{x}}^a_{t-1}\mathbf{1}^\top - \sqrt{(L+\lambda_{UKF})\boldsymbol{P}^a_{t-1}} \quad \hat{\boldsymbol{x}}^a_{t-1} \quad \hat{\boldsymbol{x}}^a_{t-1}\mathbf{1}^\top + \sqrt{(L+\lambda_{UKF})\boldsymbol{P}^a_{t-1}}\right]$$

$$= \begin{bmatrix} \boldsymbol{\chi}_{t-1}^{(h)} \\ \boldsymbol{\chi}_{t-1}^{(pn)} \\ \boldsymbol{\chi}_{t-1}^{(mn)} \end{bmatrix},$$

where $\mathbf{1}$ is a size-three column vector of ones. The row vectors $\boldsymbol{\chi}_{t-1}^{(h)}$, $\boldsymbol{\chi}_{t-1}^{(pn)}$, and $\boldsymbol{\chi}_{t-1}^{(mn)}$ are the sigma points for the distributions of the latent variable, process noise, and measurement noise, respectively. The square roots of the matrices refer to matrices containing the square roots of each element of the original matrices.

4. Perform time updates. This involves calculating the current year's sigma points for the latent distribution and for the signal by inputting the previous year's sigma points to Equation (2.3) and the measurement function. We denote these quantities using vectors $\boldsymbol{\chi}_{t|t-1}^{(h)}$ and $\boldsymbol{y}_{t|t-1}$, respectively. Weighting the updated sigma points then yields the anticipated mean and variance (denoted by $\hat{x}_t^-$ and P_t^-) of the latent variable, as well as the anticipated signal (denoted by $\hat{y}_t^-$). The calculations are as follows:

$$\begin{aligned} \boldsymbol{\chi}_{t|t-1}^{(h)} &= \boldsymbol{\chi}_{t-1}^{(h)} - \lambda\mathbf{1} + \boldsymbol{\chi}_{t-1}^{(pn)}, \\ \boldsymbol{y}_{t|t-1} &= f_{measurement}(\boldsymbol{\chi}_{t|t-1}^{(h)}, \boldsymbol{\chi}_{t-1}^{(mn)}), \\ \hat{x}_t^- &= \sum_{i=0}^{2L} W_i^{(m)} \chi_{i,t|t-1}^{(h)}, \\ P_t^- &= \sum_{i=0}^{2L} W_i^{(c)} (\chi_{i,t|t-1}^{(h)} - \hat{x}_t^-)^2, \\ \hat{y}_t^- &= \sum_{i=0}^{2L} W_i^{(m)} y_{i,t|t-1}, \end{aligned}$$

where $L = 3$ is the dimension of vector $\hat{\boldsymbol{x}}_{t-1}^a$, the scaling parameter is $\lambda_{UKF} = \alpha_{UKF}^2(L + \kappa_{UKF}) - L$, the weights are

- $W_0^{(m)} = \frac{\lambda_{UKF}}{L+\lambda_{UKF}}$,
- $W_0^{(c)} = W_0^{(m)} + (1 - \alpha_{UKF}^2 + \beta_{UKF})$,
- $W_i^{(m)} = W_i^{(c)} = \frac{1}{2(L+\lambda_{UKF})}, \forall i \in \{1, ..., 2L\}$,

and $\chi_{i,t|t-1}^{(h)}$ and $y_{i,t|t-1}$ are the i^{th} element of vectors $\boldsymbol{\chi}_{t|t-1}^{(h)}$ and $\boldsymbol{y}_{t|t-1}$, respectively, $\forall i \in \{0, 1, ..., 2L\}$. In our simulation study, we apply the following UKF parameter values based on the suggestions from Wan and Van Der Merwe's paper: $\alpha_{UKF} = 10^{-0.5}$, $\beta_{UKF} = 2$, and $\kappa_{UKF} = 0$.

5. Perform measurement updates, which means adjusting the estimates from Step 4 given the current year's observed signal. The updated mean and variance of the approximate latent variable distribution are therefore $\tilde{h}_{t,j}$ and $\rho_{t,j}^2$, respectively:

$$P_{\tilde{y}_k\tilde{y}_k} = \sum_{i=0}^{2L} W_i^{(c)} (y_{i,t|t-1} - \hat{y}_t^-)^2,$$

$$
\begin{aligned}
P_{x_k y_k} &= \sum_{i=0}^{2L} W_i^{(c)} (\chi_{i,t|t-1} - \hat{x}_t^-)(y_{i,t|t-1} - \hat{y}_t^-), \\
\kappa &= P_{x_k y_k} P_{\tilde{y}_k \tilde{y}_k}^{-1}, \\
\tilde{h}_{t,j} &= \hat{x}_t^- + \kappa (L_{t,j} - \hat{y}_t^-), \\
\rho_{t,j}^2 &= P_t^- - \kappa^2 P_{\tilde{y}_k \tilde{y}_k}.
\end{aligned}
$$

Appendix B

Derivations for the Two-Options Plan

In all the derivations in this appendix, we assume the following:

- that both the actuary and the plan member know the true annual mean loss parameter, μ_j,
- that plan member j's annual loss, L_j, follows a gamma distribution with shape parameter $\alpha_j = \mu_j\beta$ and rate parameter β. The distribution has density function

$$f_{L_j}(l) = \frac{\beta^{\alpha_j} l^{\alpha_j - 1} e^{-\beta l}}{\Gamma(\alpha_j)}$$

  and cumulative distribution function $F_{L_j}(l; \alpha_j, \beta)$,
- that all plan members' utility functions are exponential, with an absolute risk aversion parameter of R. That is,

$$u_j(c) = u(c) = -e^{-Rc},\ \forall j,$$

- that $\beta - R > 0$, and
- that the reimbursement function is as defined in Equation (2.1).

We also omit the time index, t, in the first four derivations to alleviate the notation.

B.1 Expected Claims in the Special Case

Let $EC(\alpha_j, k)$ be plan member j's expected claims in option k, given that their true mean loss parameter is μ_j (and therefore the shape parameter of their loss distribution is $\alpha_j = \mu_j\beta$). Then

$$EC(\alpha_j, k) = \mathbb{E}[r(L_j, k)|\alpha_j]$$

$$= EC_1 + EC_2 + EC_3 + EC_4$$

where

$$\begin{aligned} EC_1 &= \int_0^{d_k} 0 \times \frac{\beta^{\alpha_j} l^{\alpha_j - 1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\ &= 0, \end{aligned}$$

$$\begin{aligned} EC_2 &= \int_{d_k}^{\frac{CM_k}{1-c_k}+d_k} c_k(l - d_k) \times \frac{\beta^{\alpha_j} l^{\alpha_j - 1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\ &= c_k \int_{d_k}^{\frac{CM_k}{1-c_k}+d_k} \frac{\beta^{\alpha_j} l^{(\alpha_j+1) - 1} e^{-\beta l}}{\Gamma(\alpha_j)} dl - c_k d_k \int_{d_k}^{\frac{CM_k}{1-c_k}+d_k} \frac{\beta^{\alpha_j} l^{\alpha_j - 1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\ &= \frac{c_k \Gamma(\alpha_j + 1)}{\beta \Gamma(\alpha_j)} \int_{d_k}^{\frac{CM_k}{1-c_k}+d_k} \frac{\beta^{\alpha_j+1} l^{(\alpha_j+1) - 1} e^{-\beta l}}{\Gamma(\alpha_j + 1)} dl - c_k d_k \int_{d_k}^{\frac{CM_k}{1-c_k}+d_k} \frac{\beta^{\alpha_j} l^{\alpha_j - 1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\ &= \frac{c_k \alpha_j}{\beta} \int_{d_k}^{\frac{CM_k}{1-c_k}+d_k} \frac{\beta^{\alpha_j+1} l^{(\alpha_j+1) - 1} e^{-\beta l}}{\Gamma(\alpha_j + 1)} dl - c_k d_k \int_{d_k}^{\frac{CM_k}{1-c_k}+d_k} \frac{\beta^{\alpha_j} l^{\alpha_j - 1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\ &= \frac{c_k \alpha_j}{\beta} \left[F_L\left(\frac{CM_k}{1-c_k} + d_k; \alpha_j + 1, \beta \right) - F_L(d_k; \alpha_j + 1, \beta) \right] \\ &\quad - c_k d_k \left[F_L\left(\frac{CM_k}{1-c_k} + d_k; \alpha_j, \beta \right) - F_L(d_k; \alpha_j, \beta) \right], \end{aligned}$$

$$\begin{aligned} EC_3 &= \int_{\frac{CM_k}{1-c_k}+d_k}^{OM_k+d_k+CM_k} (l - d_k - CM_k) \times \frac{\beta^{\alpha_j} l^{\alpha_j - 1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\ &= \int_{\frac{CM_k}{1-c_k}+d_k}^{OM_k+d_k+CM_k} \frac{\beta^{\alpha_j} l^{(\alpha_j+1) - 1} e^{-\beta l}}{\Gamma(\alpha_j)} dl - (d_k + CM_k) \int_{\frac{CM_k}{1-c_k}+d_k}^{OM_k+d_k+CM_k} \frac{\beta^{\alpha_j} l^{\alpha_j - 1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\ &= \frac{\Gamma(\alpha_j + 1)}{\beta \Gamma(\alpha_j)} \int_{\frac{CM_k}{1-c_k}+d_k}^{OM_k+d_k+CM_k} \frac{\beta^{\alpha_j+1} l^{(\alpha_j+1) - 1} e^{-\beta l}}{\Gamma(\alpha_j + 1)} dl \\ &\quad - (d_k + CM_k) \int_{\frac{CM_k}{1-c_k}+d_k}^{OM_k+d_k+CM_k} \frac{\beta^{\alpha_j} l^{\alpha_j - 1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\ &= \frac{\alpha_j}{\beta} \int_{\frac{CM_k}{1-c_k}+d_k}^{OM_k+d_k+CM_k} \frac{\beta^{\alpha_j+1} l^{(\alpha_j+1) - 1} e^{-\beta l}}{\Gamma(\alpha_j + 1)} dl - (d_k + CM_k) \int_{\frac{CM_k}{1-c_k}+d_k}^{OM_k+d_k+CM_k} \frac{\beta^{\alpha_j} l^{\alpha_j - 1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\ &= \frac{\alpha_j}{\beta} \left[F_L(OM_k + d_k + CM_k; \alpha_j + 1, \beta) - F_L\left(\frac{CM_k}{1-c_k} + d_k; \alpha_j + 1, \beta \right) \right] \\ &\quad - (d_k + CM_k) \left[F_L(OM_k + d_k + CM_k; \alpha_j, \beta) - F_L\left(\frac{CM_k}{1-c_k} + d_k; \alpha_j, \beta \right) \right], \end{aligned}$$

$$EC_4 = \int_{OM_k+d_k+CM_k}^{\infty} OM_k \times \frac{\beta^{\alpha_j} l^{\alpha_j - 1} e^{-\beta l}}{\Gamma(\alpha_j)} dl$$

$$
\begin{aligned}
&= OM_k \int_{OM_k+d_k+CM_k}^{\infty} \frac{\beta^{\alpha_j} l^{\alpha_j-1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\
&= OM_k[1 - F_L(OM_k + d_k + CM_k; \alpha_j, \beta)],
\end{aligned}
$$

and $F_L(l, \alpha, \beta)$ denotes the cumulative distribution function of a gamma distribution with shape parameter α and rate parameter β, evaluated at point l.

B.2 Expected Utility of Selecting Option *k*

Let $EU(\alpha_j, k)$ be plan member j's expected utility, given that their true mean loss parameter is μ_j (and therefore the shape parameter of their loss distribution is $\alpha_j = \mu_j \beta$) and that they select option k. Also, let $y_{j,k} = w_j - (1-s)D_k + F$, where

- w_j is the annual salary of plan member j,
- s is the percent of premium subsidy paid by the plan sponsor,
- D_k is the deposit rate of option k, and
- F is the annual flex credit.

Then

$$
\begin{aligned}
EU(\alpha_j, k) &= \mathbb{E}[u(y_{j,k} + r(L_j) - L_j)|\alpha_j] \\
&= \mathbb{E}[-e^{-R(y_{j,k}+r(L_j)-L_j)}|\alpha_j] \\
&= EU_1 + EU_2 + EU_3 + EU_4
\end{aligned}
$$

where

$$
\begin{aligned}
EU_1 &= \int_0^{d_k} -e^{-(y_{j,k}+0-l)R} \frac{\beta^{\alpha_j} l^{\alpha_j-1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\
&= -\int_0^{d_k} \frac{\beta^{\alpha_j} l^{\alpha_j-1} e^{-y_{j,k}R+(R-\beta)l}}{\Gamma(\alpha_j)} dl \\
&= -\beta^{\alpha_j} e^{-y_{j,k}R} \int_0^{d_k} \frac{l^{\alpha_j-1} e^{-(\beta-R)l}}{\Gamma(\alpha_j)} dl \\
&= -\frac{\beta^{\alpha_j} e^{-y_{j,k}R}}{(\beta-R)^{\alpha_j}} \int_0^{d_k} \frac{(\beta-R)^{\alpha_j} l^{\alpha_j-1} e^{-(\beta-R)l}}{\Gamma(\alpha_j)} dl \\
&= -\left(\frac{\beta}{\beta-R}\right)^{\alpha_j} e^{-y_{j,k}R} F_L(d_k; \alpha_j, \beta-R),
\end{aligned}
$$

$$
\begin{aligned}
EU_2 &= \int_{d_k}^{\frac{CM_k}{1-c_k}+d_k} -e^{-(y_{j,k}+c_k(l-d_k)-l)R} \frac{\beta^{\alpha_j} l^{\alpha_j-1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\
&= -\int_{d_k}^{\frac{CM_k}{1-c_k}+d_k} \frac{\beta^{\alpha_j} l^{\alpha_j-1} e^{-y_{j,k}R+c_k d_k R-(c_k-1)Rl-\beta l}}{\Gamma(\alpha_j)} dl \\
&= -\frac{\beta^{\alpha_j} e^{(c_k d_k - y_{j,k})R}}{[(c_k-1)R+\beta]^{\alpha_j}} \int_{d_k}^{\frac{CM_k}{1-c_k}+d_k} \frac{[(c_k-1)R+\beta]^{\alpha_j} l^{\alpha_j-1} e^{-[(c_k-1)R+\beta]l}}{\Gamma(\alpha_j)} \\
&= -\left[\frac{\beta}{(c_k-1)R+\beta}\right]^{\alpha_j} e^{(c_k d_k - y_{j,k})R} \\
&\quad \times \left[F_L\left(\frac{CM_k}{1-c_k}+d_k;\alpha_j,(c_k-1)R+\beta\right) - F_L(d_k;\alpha_j,(c_k-1)R+\beta)\right],
\end{aligned}
$$

$$
\begin{aligned}
EU_3 &= \int_{\frac{CM_k}{1-c_k}+d_k}^{OM_k+d_k+CM_k} -e^{-(y_{j,k}+l-d_k-CM_k-l)R} \frac{\beta^{\alpha_j} l^{\alpha_j-1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\
&= -e^{-(y_{j,k}-d_k-CM_k)R} \int_{\frac{CM_k}{1-c_k}+d_k}^{OM_k+d_k+CM_k} \frac{\beta^{\alpha_j} l^{\alpha_j-1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\
&= -e^{(d_k+CM_k-y_{j,k})R} \left[F_L(OM_k+d_k+CM_k;\alpha_j,\beta) - F_L\left(\frac{CM_k}{1-c_k}+d_k;\alpha_j,\beta\right)\right],
\end{aligned}
$$

$$
\begin{aligned}
EU_4 &= \int_{OM_k+d_k+CM_k}^{\infty} -e^{-(y_{j,k}+OM_k-l)R} \frac{\beta^{\alpha_j} l^{\alpha_j-1} e^{-\beta l}}{\Gamma(\alpha_j)} dl \\
&= -\frac{\beta^{\alpha_j} e^{-(y_{j,k}+OM_k)R}}{(\beta+R)^{\alpha_j}} \int_{OM_k+d_k+CM_k}^{\infty} \frac{(\beta+R)^{\alpha_j} l^{\alpha_j-1} e^{-(\beta+R)l}}{\Gamma(\alpha_j)} dl \\
&= -\left(\frac{\beta}{\beta+R}\right)^{\alpha_j} e^{-(y_{j,k}+OM_k)R}[1-F_L(OM_k+d_k+CM_k;\alpha_j,\beta+R)].
\end{aligned}
$$

B.3 Inequality Constraint For Calculating the Deposit Rates in the Special Case with $K = 2$ Options

As described in Section 2.4.2, we assume the actuary splits the group of members ex ante so that those with a mean loss parameter μ_j (and analogously, α_j) below a certain predetermined threshold are allocated into the low risk group, G_1, and those with μ_j (and analogously, α_j) above this threshold are placed into the high risk group, G_2. The threshold, for example, might be the median of all the μ_j's, so that the plan members are evenly placed into the two groups. Then the feasible set of deposit rates $\{D_1, D_2\}$ are constrained by the following inequalities:

$$
\begin{aligned}
&EU(\alpha_j, 2) < EU(\alpha_j, 1), \forall j \in G_1, \text{ and} \\
&EU(\alpha_j, 1) < EU(\alpha_j, 2), \forall j \in G_2,
\end{aligned}
$$

where the expected utility is derived in Appendix B.2. Note that we can write $EU(\alpha_j, k) = -e^{-(w_j-(1-s)D_k+F)R} \times h(\alpha_j, k)$, where

$$
\begin{aligned}
h(\alpha_j, k) =& \left(\frac{\beta}{\beta - R}\right)^{\alpha_j} F_L(d_k; \alpha_j, \beta - R) \qquad \text{(B.1)}\\
&+ \left[\frac{\beta}{(c_k - 1)R + \beta}\right]^{\alpha_j} e^{(c_k d_k)R}\\
&\times \left[F_L\left(\frac{CM_k}{1 - c_k} + d_k; \alpha_j, (c_k - 1)R + \beta\right) - F_L(d_k; \alpha_j, (c_k - 1)R + \beta)\right]\\
&+ e^{(d_k + CM_k)R}\left[F_L(OM_k + d_k + CM_k; \alpha_j, \beta) - F_L\left(\frac{CM_k}{1 - c_k} + d_k; \alpha_j, \beta\right)\right]\\
&+ \left(\frac{\beta}{\beta + R}\right)^{\alpha_j} e^{-OM_k R}[1 - F_L(OM_k + d_k + CM_k; \alpha_j, \beta + R)].
\end{aligned}
$$

Then the above inequalities can be written as

$$
\begin{aligned}
e^{(1-s)R(D_2 - D_1)} h(\alpha_j, 2) >& h(\alpha_j, 1), \forall j \in G_1, \text{ and}\\
e^{(1-s)R(D_1 - D_2)} h(\alpha_j, 1) >& h(\alpha_j, 2), \forall j \in G_2.
\end{aligned}
$$

The following are two useful properties of $h(\alpha_j, k)$:

1. Given $\beta - R > 0$, $\beta > 0$, and $R > 0$, we have $h(\alpha_j, k) > 0$.

   Proof. We show that each of the four terms in Equation (B.1) is positive.

   (a) The coefficient in the first term is positive because $\beta - R > 0$ and $\beta > 0$. We also know that the cumulative distribution function, $F_L(d_k; \alpha_j, \beta - R)$, is in the interval (0,1). This implies that first term is positive.

   (b) The coefficient in the second term of Equation (B.1) contains the parameter c_k, which is in the interval (0,1). This implies that $(c_k - 1)R \in (-R, 0)$, which means $(c_k - 1)R + \beta > 0$.
   Also, we know that $\frac{CM_k}{1-c_k} + d_k > d_k$, which means

   $$F_L\left(\frac{CM_k}{1 - c_k} + d_k; \alpha_j, (c_k - 1)R + \beta\right) > F_L(d_k; \alpha_j, (c_k - 1)R + \beta).$$

   Since both the coefficient and the difference in cumulative distribution functions on the second and third lines are positive, the second term is also positive.

   (c) We know that $OM_k + d_k + CM_k > \frac{CM_k}{1-c_k} + d_k$, which implies

   $$F_L(OM_k + d_k + CM_k; \alpha_j, \beta) > F_L\left(\frac{CM_k}{1 - c_k} + d_k; \alpha_j, \beta\right).$$

   Since both the coefficient and the difference in cumulative distribution functions is positive, the third term is positive as well.

(d) Finally, we know that the cumulative distribution function $F_L(OM_k + d_k + CM_k; \alpha_j, \beta + R)$ is contained in the interval (0,1), which implies that

$$1 - F_L(OM_k + d_k + CM_k; \alpha_j, \beta + R)$$

is positive. The fourth and final term is therefore positive.

Since $h(\alpha_j, k)$ is a sum of four positive terms, it is positive. □

2. Also, we have $h(\alpha_j, 1) > h(\alpha_j, 2)$.

Proof. Clearly, if the deposit rates of the two options are equal, then any plan member j will choose the higher coverage option. In other words, if $D_1 = D_2 = D$, then we have

$$\begin{aligned} & EU(\alpha_j, 2) > EU(\alpha_j, 1) \\ \iff & -e^{-(w_j-(1-s)D+F)R} h(\alpha_j, 2) > -e^{-(w_j-(1-s)D+F)R} h(\alpha_j, 1) \\ \iff & h(\alpha_j, 2) < h(\alpha_j, 1). \end{aligned}$$

Since $h(\alpha_j, k)$ does not depend on the deposit rates, it follows that $h(\alpha_j, 1) > h(\alpha_j, 2)$ always holds. □

Therefore, by the first property above, the constraint inequalities can be combined as follows:

$$LB < D_2 - D_1 < UB,$$

where the upper and lower bounds are

$$LB = \max_{j \in G_1} \frac{1}{(1-s)R} \ln\left(\frac{h(\alpha_j, 1)}{h(\alpha_j, 2)}\right) \text{ and}$$

$$UB = \min_{j \in G_2} \frac{1}{(1-s)R} \ln\left(\frac{h(\alpha_j, 1)}{h(\alpha_j, 2)}\right).$$

Note that these bounds are positive since we have $\ln\left(\frac{h(\alpha_j,1)}{h(\alpha_j,2)}\right) > 0$ by the second property above and $s \in (0, 1)$.

Finally, since the expected utilities are continuous random variables, we can rewrite the strict inequalities into the following constraint:

$$LB \leq D_2 - D_1 \leq UB.$$

B.4 Calculating the Deposit Rates in the Special Case with $K = 2$ Options

Let $S_j(\boldsymbol{D}, \mu_j)$ and n_k be as defined in Section 2.

Then the ideal deposit rates are the solution to the optimization problem below, subject to the inequality constraint of $LB \leq D_2 - D_1 \leq UB$:

$$
\begin{aligned}
\min_{\boldsymbol{D}} \sum_{k=1}^{K} \Bigg(& D_k \mathbb{E}\left[\sum_{j=1}^{N} 1_{\{S_j(\boldsymbol{D}, \overline{\mu}_j)=k\}} \mid \overline{\mu}_j = \mu_j\right] \\
& - \mathbb{E}\left[\sum_{j=1}^{N} 1_{\{S_j(\boldsymbol{D}_t, \overline{\mu}_j)=k\}} r(L_j, k) \mid \overline{\mu}_j = \mu_j, \tilde{\mu}_j = \mu_j\right]\Bigg)^2 \\
= \min_{D_1, D_2} \sum_{k=1}^{2} & \left(D_k \sum_{j=1}^{N} \mathbb{E}(1_{S_j(\boldsymbol{D}, \mu_j)=k}) - \sum_{j=1}^{N} \mathbb{E}\left[1_{S_j(\mu_j, \vec{D})=k} r(L_j, k)\right] \right)^2 \\
= \min_{D_1, D_2} \sum_{k=1}^{2} & \left[D_k n_k - \sum_{j \in G_k} EC(\alpha_j, k) \right]^2
\end{aligned}
$$

Note that the objective function is convex because we have $\frac{\partial^2 Obj}{\partial D_k^2} = 2n_k > 0$ for $k = 1, 2$ and $\frac{\partial^2 Obj}{\partial D_k D_{k'}} = 0$ for $k \neq k'$.

Solving this optimization using Lagrange multipliers, we obtain the solution of

$$
\boldsymbol{D} = \begin{cases} \boldsymbol{D}^{LB} = [D_1^{LB}, D_2^{LB}] & \text{if } D_2^* - D_1^* < LB \\ \boldsymbol{D}^* = [D_1^*, D_2^*] & \text{if } LB \leq D_2^* - D_1^* \leq UB. \\ \boldsymbol{D}^{UB} = [D_1^{UB}, D_2^{UB}] & \text{if } UB < D_2^* - D_1^* \end{cases}
$$

where

$$
\begin{aligned}
D_k^* &= \frac{\sum_{j \in G_k} EC(\alpha_j, k)}{n_k} \quad \forall k \in \{1, 2\}, \\
D_1^{UB} &= \frac{n_1 \sum_{j \in G_1} EC(\alpha_j, 1) + n_2 \sum_{j \in G_2} EC(\alpha_j, 2) - n_2^2 UB}{n_1^2 + n_2^2}, \\
D_2^{UB} &= UB + D_1^{UB}, \\
D_1^{LB} &= \frac{n_1 \sum_{j \in G_1} EC(\alpha_j, 1) + n_2 \sum_{j \in G_2} EC(\alpha_j, 2) - n_2^2 LB}{n_1^2 + n_2^2}, \quad \text{and} \\
D_2^{LB} &= LB + D_1^{LB}.
\end{aligned}
$$

B.5 Expected Total Annual Change in ASO Account Balance

The expected annual change in total ASO account balance is

$$
\begin{aligned}
&\mathbb{E}[\text{Annual change in total ASO balance}]\\
=&\sum_{k=1}^{K}\sum_{j=1}^{N}\mathbb{E}\left[1_{\{S_{t,j}(\boldsymbol{D}_t,\mu_{t,j})=k\}}\left(D_{t,k}-r(L_{t,j},k)\right)\right]\\
=&\sum_{k=1}^{K}\left(n_k D_{t,k}-\sum_{j\in G_{t,k}}EC(\alpha_{t,j},k)\right)\\
=&\sum_{k=1}^{K}n_k\left(D_{t,k}-\frac{\sum_{j\in G_{t,k}}EC(\alpha_{t,j},k)}{n_k}\right)\\
=&n_1\left(D_{t,1}-\frac{\sum_{j\in G_{t,1}}EC(\alpha_{t,j},1)}{n_1}\right)+n_2\left(D_{t,2}-\frac{\sum_{j\in G_{t,2}}EC(\alpha_{t,j},2)}{n_2}\right).
\end{aligned}
$$

Given the solution presented in Appendix B.4, we have three cases:

1. If $LB_t\le D^*_{t,2}-D^*_{t,1}\le UB_t$, then

$$
\begin{aligned}
&n_1\left(D^*_{t,1}-\frac{\sum_{j\in G_{t,1}}EC(\alpha_{t,j},1)}{n_1}\right)+n_2\left(D^*_{t,2}-\frac{\sum_{j\in G_{t,2}}EC(\alpha_{t,j},2)}{n_2}\right)\\
=&n_1(0)+n_2(0)\\
=&0.
\end{aligned}
$$

2. If $D^*_{t,2}-D^*_{t,1}<LB_t$, then

$$
n_1\left(D^{(LB)}_{t,1}-\frac{\sum_{j\in G_{t,1}}EC(\alpha_{t,j},1)}{n_1}\right)+n_2\left(D^{(LB)}_{t,2}-\frac{\sum_{j\in G_{t,2}}EC(\alpha_{t,j},2)}{n_2}\right)
$$

$$
\begin{aligned}
=&n_1\left(\frac{n_1\sum_{j\in G_{t,1}}EC(\alpha_{t,j},1)+n_2\sum_{j\in G_{t,2}}EC(\alpha_{t,j},2)-n_2^2LB_t}{n_1^2+n_2^2}-\frac{\sum_{j\in G_{t,1}}EC(\alpha_{t,j},1)}{n_1}\right)\\
&+n_2\left(LB_t+\frac{n_1\sum_{j\in G_{t,1}}EC(\alpha_{t,j},1)+n_2\sum_{j\in G_{t,2}}EC(\alpha_{t,j},2)-n_2^2LB_t}{n_1^2+n_2^2}\right.\\
&\qquad\left.-\frac{\sum_{j\in G_{t,2}}EC(\alpha_{t,j},2)}{n_2}\right)
\end{aligned}
$$

$$
\begin{aligned}
=&\frac{n_1^2\sum_{j\in G_{t,1}}EC(\alpha_{t,j},1)+n_1n_2\sum_{j\in G_{t,2}}EC(\alpha_{t,j},2)-n_1n_2^2LB_t}{n_1^2+n_2^2}\\
&+\frac{-(n_1^2+n_2^2)\sum_{j\in G_{t,1}}EC(\alpha_{t,j},1)+(n_1^2n_2+n_2^3)LB_t}{n_1^2+n_2^2}
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{n_1 n_2 \sum_{j\in G_{t,1}} EC(\alpha_{t,j},1) + n_2^2 \sum_{j\in G_{t,2}} EC(\alpha_{t,j},2) - n_2^3 LB_t}{n_1^2+n_2^2} \\
&-\frac{(n_1^2+n_2^2)\sum_{j\in G_{t,2}} EC(\alpha_{t,j},2)}{n_1^2+n_2^2} \\
=&\frac{(n_1 n_2 - n_2^2)\sum_{j\in G_{t,1}} EC(\alpha_{t,j},1) + (n_1 n_2 - n_1^2)\sum_{j\in G_{t,2}} EC(\alpha_{t,j},2)}{n_1^2+n_2^2} \\
&+\frac{(n_1^2 n_2 - n_1 n_2^2) LB_t}{n_1^2+n_2^2} \\
=&\frac{n_1 n_2 (n_1 - n_2)\frac{\sum_{j\in G_{t,1}} EC(\alpha_{t,j},1)}{n_1} + n_1 n_2 (n_2 - n_1)\frac{\sum_{j\in G_{t,2}} EC(\alpha_{t,j},2)}{n_2}}{n_1^2+n_2^2} \\
&+\frac{n_1 n_2 (n_1 - n_2) LB_t}{n_1^2+n_2^2} \\
=&\frac{n_1 n_2 (n_1 - n_2)}{n_1^2+n_2^2}\left(\frac{\sum_{j\in G_{t,1}} EC(\alpha_{t,j},1)}{n_1} + LB_t - \frac{\sum_{j\in G_{t,2}} EC(\alpha_{t,j},2)}{n_2}\right). \qquad \text{(B.2)}
\end{aligned}
$$

3. Similarly, if $UB_t < D^*_{t,2} - D^*_{t,1}$, then

$$
\begin{aligned}
&n_1\left(D^{(UB)}_{t,1} - \frac{\sum_{j\in G_{t,1}} EC(\alpha_{t,j},1)}{n_1}\right) + n_2\left(D^{(UB)}_{t,2} - \frac{\sum_{j\in G_{t,2}} EC(\alpha_{t,j},2)}{n_2}\right) \\
=&\frac{n_1 n_2 (n_1 - n_2)}{n_1^2+n_2^2}\left(\frac{\sum_{j\in G_{t,1}} EC(\alpha_{t,j},1)}{n_1} + UB_t - \frac{\sum_{j\in G_{t,2}} EC(\alpha_{t,j},2)}{n_2}\right). \qquad \text{(B.3)}
\end{aligned}
$$

Clearly, regardless of which of the three cases hold, if $n_1 = n_2$ then

$$\mathbb{E}[\text{Annual change in total ASO balance}] = 0.$$

It also follows that if $n_1 = n_2$ then

$$\mathbb{E}[\text{Year-end total ASO balance}] = 0, \quad \forall \text{ year } t.$$

Appendix C

Including a Random Component in the Selection Function

Under our assumption that all plan members rationally select their upcoming year's options, the expected utility theory described in Section 2.3 suffices as a complete model for individual preference under financial uncertainty. In general, however, not all plan members sufficiently understand their benefits coverage, or even spend the time to make a selection during the annual enrollment window (Garnick et al., 1993). A plan member may either

- make an informed decision consistent with expected utility theory (as we have assumed in our model),
- make an uninformed decision, or
- not respond.

Therefore, to relax the assumption of all plan members being diligent (which is a requirement for adverse selection to occur), we can include a plan design feature that automatically enrolls a plan member into a default option in case the plan member does not make a selection. This default can be set to either a predetermined option k^* or the plan member's current option.

With this additional plan design feature defined, we can then add a probabilistic layer to the expected utility model to extend our framework to more general settings. Let

- $p_{t,j}^{(i)}$ be the probability that plan member j makes an informed decision and selects $U_{t,j}$ in year t, and
- $p_{t,j}^{(n)}$ be the probability that plan member j does not make a selection, so that depending on the plan set up, the selection defaults to either this plan member's current option or k^*.

We further assume that if the plan member makes an uninformed decision, this plan member chooses an option at random with equal probabilities of $\frac{1-p_{t,j}^{(i)}-p_{t,j}^{(n)}}{K}$.

Let $S_{t,j} \in \{1, \ldots , \mathrm{K}\}$ be plan member j's enrolled option in year t. Then $S_{i,j}$ follows a multinomial distribution with the number of trials parameter equal to 1 and the probability parameters determined using $p_{t,j}^{(i)}$ and $p_{t,j}^{(n)}$.