



Ulm University

Institute of Insurance Science

# **Life Insurance Companies: Product Design, Marketing, and Investment**

Cumulative doctoral thesis

Submitted to the Faculty of Mathematics and Management of Ulm University

For the academic degree of Doctor of Management and Economics Dr. rer. pol.

By

**Yusha Chen**

Born in Yibin, China

2024

**Dean in office:** Prof. Dr. Gunter Löffler

**Supervisor:** Prof. Dr. An Chen

**Evaluator:** Prof. Dr. Xian Xu

**Date of the oral examination:** April 10, 2024

# Acknowledgments

Completing this doctoral dissertation has been a journey filled with challenges and growth, and it is with immense gratitude that I express my appreciation to the individuals who have played pivotal roles in this endeavor.

First and foremost, I am profoundly thankful for the invaluable guidance and support provided by my esteemed supervisor, Prof. Dr. An Chen. Her professional and meticulous supervision, coupled with patient explanations and constructive feedbacks, has been instrumental in shaping my academic journey. Beyond academia, I deeply appreciate for the tremendous help she has extended to me in navigating life in a foreign country. She has made me feel at home in what was initially an unfamiliar environment. Her kindness and always-sympathetic mind have been a source of solace during challenging times. During our time together, she has not only imparted academic knowledge but has also shared many invaluable experiences and wisdom on navigating the complexities of the world. In every sense, Prof. Dr. An Chen is a wonderful supervisor, and I feel incredibly lucky to have her as my mentor. Her impact on my academic and personal growth is immeasurable, and I express my heartfelt gratitude to her.

Next, I would like to express my sincere thanks to Prof. Dr. Xian Xu, for not only reviewing this dissertation but also for inspiring and endorsing my pursuit of doctoral studies, as well as contributing to our collaborative projects.

Further, I am fortunate to have had the privilege of working with outstanding co-authors. My sincere thanks go to Prof. Dr. Thai Nguyen, Prof. Dr. Manuel Rach and Dr. Wei Xu, Prof. Finbarr Murphy, and Prof. Dr. Gazi Uddin. Together, we have navigated the complexities of research, and your contributions have significantly enriched the content and depth of this dissertation.

To my family, whose steadfast support and understanding have been the bedrock of my academic pursuits, I am extremely grateful. Your belief in me has been a constant motivation to strive for excellence. Additionally, I owe a debt of gratitude to my boyfriend, for his unwavering encouragement, patience, and understanding. His presence has been a source of strength, making the challenges more manageable and the successes more joyous.

In addition, I would like to thank my colleagues at the Institute of Insurance Science. The intellectual exchange, shared experiences, and collaborative spirit within our academic community have been invaluable.

---

In closing, this dissertation stands as a testament to the collective effort of a supportive network of individuals. Each of you has left an indelible mark on this academic milestone, and for that, I am truly thankful. Your belief in me has been a driving force, and I look forward to carrying the lessons and relationships forged during this journey into the next chapter of my professional pursuits.

With heartfelt appreciation,

Ulm, April 2024

Yusha Chen



# Contents

<b>Overview of Research Papers</b>	<b>vii</b>
<b>Co-Authorship</b>	<b>ix</b>
<b>Research Context and Summary of Research Papers</b>	<b>1</b>
1    Field of Research . . . . .	1
2    Motivation and Objectives . . . . .	3
3    Summary of Research Papers . . . . .	6
 <b>Research Papers</b>	
<b>1 Care-dependent tontines</b>	<b>21</b>
<b>2 Dynamic tonuity: Adapting retirement benefits to a changing environment</b>	<b>45</b>
<b>3 How does the insurer's mobile application sales strategy perform?</b>	<b>89</b>
<b>4 Investing in green bonds with goal-oriented preferences</b>	<b>125</b>
<b>Curriculum Vitae</b>	<b>157</b>



# Overview of Research Papers

## Research papers included in this dissertation

1. Chen, A., Chen, Y. & Xu, X. (2022). Care-dependent tontines. *Insurance: Mathematics and Economics*, 106, 69–89.
2. Chen, A., Chen, Y. & Rach, M. (2024). Dynamic tonuity: Adapting retirement benefits to a changing environment. Submitted to *ASTIN Bulletin: The Journal of the IAA* (under review).
3. Chen, A., Chen, Y., Murphy, F., Xu, W. & Xu, X. (2023). How does the insurer’s mobile application sales strategy perform?. *Journal of Risk and Insurance*, 90, 487–519.
4. Chen, A., Chen, Y., Nguyen, T. & Uddin, G. S. (2024). Investing in green bonds with goal-oriented preferences. Working paper.



# Co-Authorship

## **An Chen**

An Chen has earned her doctoral degree from the University of Bonn, Germany, in 2007. Since 2012, she has served as a full professor and the head of the Institute of Insurance Science at Ulm University.

## **Xian Xu**

Xian Xu finished his doctoral study in 2010, graduated from the University of Karlsruhe, Germany. Now he works as a full professor and the head of Department of Risk Management and Insurance at Fudan University, China.

## **Manuel Rach**

Manuel Rach successfully completed his doctoral studies at Ulm University, Germany, in 2020. From August 2022, he has been an assistant professor in Risk Management and Insurance, University of St. Gallen, Switzerland.

## **Wei Xu**

Wei Xu has obtained his Ph.D. degree from the University of Limerick, Ireland. Now he works as a junior researcher in China Insurance and Social Security Research Center at Fudan University.

## **Thai Nguyen**

Thai Nguyen has received his Ph.D. degree from the University of Rouen, France. At present, he works as an assistant professor in the actuarial school at the University of Laval, Canada.

## **Finbarr Murphy**

Finbarr Murphy finished his Ph.D. study in 2011 at the University of Limerick, Ireland. Currently, he is the executive dean and associate professor in quantitative finance and emerging risk at the University of Limerick.

## **Gazi Salah Uddin**

Gazi Salah Uddin has got his doctoral degree in 2016 at Linköping University, Sweden. Now he is the senior associate professor of Financial Economics at Linköping University.



# Research Context and Summary of Research Papers

## 1 Field of Research

This dissertation centers on retirement product design, marketing strategies, and investment management within the life insurance industry. Each of these aspects discusses a separate issue and is therefore related to different research topics.

In the dynamic world of life insurance, strategic capabilities, innovative product design, compelling marketing strategies and astute investment management converge to determine the trajectory of success. The adaptability and foresight of life insurance companies, as custodians of financial protection and long-term planning, are critical in this space.

The global demographic landscape is undergoing a significant transformation marked by an aging population. Advances in medical and social security services have led to increased life expectancy and growing percentage of older people. Over the past few decades, across OECD countries, the average life expectancy at the age of 65 has risen by 6 years from 1970 to 2021; and the proportion of individuals aged 65 and above has averagely doubled, rising from under 9% in 1960 to approximately 18% in the year 2021 (OECD, 2023). Along with population aging come challenges. One prominent issue revolves around retirement funding. As the increasing longevity risks place pressure on existing models, it highlights the need for innovative retirement products that can provide sustainable financial solutions for individuals in their later years. Additionally, the question of elderly care poses a substantial concern, as the increasing number of seniors necessitates a responsive and comprehensive approach to healthcare and support services. A survey conducted in 2019 (AIC and IPLE, 2020) has revealed a growing prevalence of Activities of Daily Living (ADL) disabilities among the elderly in China. And in the year 2021, the estimated number of people with dementia in OECD countries exceeded 21 million (OECD, 2023). The strain on resources, both financial and human is palpable, which requires innovative solutions to ensure a high quality of life for the elderly.

In response to the aging trend and challenge of elderly care, the first and second paper of this doctoral dissertation put forward innovative retirement products which have attractive characteristics. Among the different retirement products, tontines stand out due to their diminished

exposure to longevity risk and lower safety loadings compared to annuities (e.g., Sabin, 2010; Newfield, 2014; Fullmer and Sabin, 2019). In the context of tontines, policyholders shoulder the primary burden of longevity risks. Put differently, the ongoing payments to surviving members are augmented by the “mortality credit” generated from those who have passed away. The first paper of this dissertation introduces different ways to integrate the long-term care (LTC) contingency into retirement tontines, specifically care-dependent tontines. We explore optimal payment structures for these care-dependent tontines that maximize individuals’ expected lifetime utility. Accordingly, we calculate the risk loadings based on data from the China Health and Retirement Longitudinal Study and align them with the capital requirements stipulated by the China Risk Oriented Solvency System and Solvency II. And further, we compare the care-dependent tontines with care-dependent annuities.

Additionally, in order to integrate the appealing features of tontines into the retirement product, Chen et al. (2019) have suggested a composite offering, namely the (regular) tonuity. It blends the attractive elements of a tontine and an annuity; and it offers tontine-like returns until a specified age, at which point it transitions to annuity-like payouts. The second paper identifies limitations of the regular tonuity, particularly in the face of unforeseen events such as an unexpected mortality shock (e.g., the recent Covid-19 pandemic) or an unanticipated longevity shock (e.g., medical advancements). To address the potential drawbacks of the regular tonuity, we have introduced an enhanced product known as dynamic tonuity. This innovative product integrates a dynamic switching condition that remains visible to policyholders at all times, triggering an automatic switch when the condition is met. Utilizing data from the Human Mortality Database and UK Continuous Mortality Investigation, we assess the performance of the suggested dynamic tonuity in contrast to the regular tonuity within a changing environment.

In tandem with the demographic shift, the insurance sector is undergoing a digital revolution. Consumer preferences are shifting towards mobile devices, with a predilection for mobile applications (MAs) (Cognizant, 2015). MAs provide added convenience, quicker usability (e.g., streamlined one-hand operation with fewer clicks and minimal time), enhanced hardware functionalities (e.g., sensors, Bluetooth), and swifter responses (Wisniewski, 2011; Lazaris et al., 2015) compared to their web counterparts. Consequently, insurers are placing substantial value on the MA sales strategy, anticipating it to exert a more profound influence on insurance sales (Eling and Lehmann, 2018). The third paper explores the influence of the MA sales strategy on an insurer’s sales performance, where two key performance metrics, namely the quantity of insurance purchases and premiums received from new policies, are considered. In addition, we also identify intermediate contributors to the impact of the strategy.

Dramatic changes in times have posed a number of challenges to the insurance industry, climate change being one of them, apart from the two mentioned above. In one instance, it has triggered a call for corporate social responsibility (CSR) investments. Notably, life insurance companies



are actively engaging in CSR investments. Investing in green bonds has emerged as a choice for these companies. In 2021, AXA raised 1 billion Euro by issuing green bonds (AXA, 2022). And as of November 30, 2023, the volume of Allianz green bond fund reached 863.63 million Euro<sup>1</sup>. In the fourth paper, we develop models that integrate the CSR investment goal of institutional investors into their utility functions, indicating goal-based preferences. The optimization of terminal wealth and asset holdings, which includes green bonds, is achieved by maximizing the goal-based utility.

## 2 Motivation and Objectives

As populations around the world are experiencing a significant increase in the proportion of elderly individuals, it poses a multifaceted challenge in contemporary society. This demographic shift brings forth various societal, economic, and healthcare challenges. With longer life expectancy, there is a growing demand for comprehensive healthcare services to address age-related issues. In particular, the demand for LTC services is increasing. However, the future affordability of the public LTC insurance is questionable due to the increasing number of the disabled elderly (Morrow and Röger, 2003). Further, the LTC insurance plans offered to retired individuals in the private LTC insurance market are often either restricted in coverage or come with a relatively high safety loading (Brown and Finkelstein, 2009), aiming to mitigate the risks shouldered by insurers. On basis of this, some researchers put forward care-dependent annuities, which is the combination of LTC contingency and annuities (e.g., Murtaugh et al., 2001; Brown and Warshawsky, 2013; Vidal-Melia et al., 2016; Pitacco, 2015; Hoermann and Ruß, 2008; Ramsay and Oguledo, 2020; Chen et al., 2022).

While care-dependent annuities are generally more cost-effective than purchasing life annuities and LTC products separately (Murtaugh et al., 2001), their market penetration remains limited, constituting less than 10% of the voluntary annuity market in the United States (Webb, 2009). This restrained adoption can be attributed to persistent adverse selection challenges within the care-dependent annuity market (Zhou-Richter and Gründl, 2011). Additionally, insurers may face difficulties in accurately determining reserves for these products, potentially resulting in elevated administration and risk charges. This leads to the first question of this dissertation, which is examined by the first paper:

1. How can one craft an LTC insurance product tailored for retirees with a relatively lower safety loading? Is integrating the LTC coverage into tontines a superior option compared to the current blend of LTC contingency and annuities? What potential ways exist for combining LTC contingency with tontines?

---

<sup>1</sup>The data is accessed on December 30, 2023, sourced from <https://www.finanzen.net/fonds/allianz-global-investors-fund-allianz-green-bond-lu1542252181>.

Next, the increasing prevalence of aging populations contributes to a growing need for retirement products. Tontines are brought to the forefront of recent discussions, as they can be appealing to insurance companies due to a lower longevity risk exposure and lower safety loadings in contrast to annuities (e.g., Piggott et al., 2005; Sabin, 2010; Richter and Weber, 2011; Newfield, 2014; Fullmer and Sabin, 2019). However, the highly fluctuating payoffs of tontines are less desirable for older ages (e.g., Gemmo et al., 2020; Weinert and Gründl, 2021) and result in lower expected lifetime utilities than annuities under actuarially fair pricing (Milevsky and Salisbury, 2015). In response, solutions like partial tontinization or combined products, such as tontines with minimum guarantees and tonuities, have been proposed, presenting certain advantages (e.g., Donnelly and Young, 2017; Chen et al., 2019; Chen and Rach, 2019; Weinert and Gründl, 2021). Tonuity, proposed by Chen et al. (2019), has been identified as nearly optimal in the analysis of optimal portfolios involving tontines and annuities (Chen et al., 2020). Nevertheless, its fixed switching time when the contract is initiated does not take into account unforeseen mortality or longevity shocks, such as those caused by the Covid-19 pandemic or medical advances. Besides, the regular tonuity models require policyholders to predetermine switching times based on their risk preferences, which can be challenging due to lack of financial literacy or time constraints. Then we have the second question of this dissertation, which is investigated by the second paper, formulated as follows:

2. In the face of changing circumstances, does the regular tonuity also perform well from the policyholder's perspective? If not, what approaches can be employed to craft a tonuity product that is more appealing in such scenarios?

In addition to a good product design, the success of an insurance product is also dependent on a product marketing or sales strategy that keeps pace with the times. While extensive research has examined the impact of the Internet and mobile Internet on the sales performance of physical goods within marketing and retailing fields (e.g., Luo et al., 2014; Huang et al., 2016), the insurance sector lags behind in a thorough analysis. In particular, academic research on how Internet distribution strategies influence insurers' sales performance is notably scarce, with a specific gap in studies focusing on MA sales strategies. Existing studies, such as Brown and Goolsbee (2002), highlight evidence supporting the search theory, revealing a decline in term life insurance prices after the introduction of insurance comparison websites (Stahl, 1989). A recent global survey by Benlagha and Hemrit (2020) underscores the positive impact of Internet use on non-life insurance demand. However, key distinctions from our work emerge: these studies predominantly focus on the traditional PC Internet channel, overlooking the emerging mobile Internet channel, and adopt a consumer-centric perspective, emphasizing insurance demand rather than the insurer's standpoint, specifically sales performance. In essence, there is limited literature studying the performance of the MA sales strategy from the insurer's perspective. Given this context, we pose the following question, elucidated in the third paper:

3. How does the insurer's MA sales strategy perform? Through which mechanism does the strategy exert an impact? Are there additional effects brought by this strategy?

The dynamic nature of financial markets, coupled with the need for sustainable and secure returns, necessitates a comprehensive analysis of the investment strategies employed by life insurance companies in the context of retirement products. In the realm of investments, there is a noticeable movement towards CSR investments, with a significant rise in the appeal of green bonds, due to the climate change. This shift is particularly evident among institutional investors, such as life insurance companies who prioritize CSR investment objectives. The existing literature on green bonds primarily revolves around empirical research. For instance, Han and Li (2022) and Braga et al. (2021) analyze how portfolios containing green bonds perform. The “greenium” is explored, indicating that green bonds may be priced lower than risk-paired conventional corporate bonds (e.g., Karpf and Mandel, 2018; Gianfrate and Peri, 2019; Nanayakkara and Colombage, 2019; Hachenberg and Schiereck, 2018; Bachelet et al., 2019; Partridge and Medda, 2020; Tang and Zhang, 2020). Some scholars examine the connection between green bonds and other financial instruments (e.g., Reboredo, 2018; Reboredo and Ugolini, 2020; Park et al., 2020; Nguyen et al., 2021). Aside from these, other researchers concentrate on assessing the determinants influencing the supply of green bonds (e.g., Chiesa and Barua, 2019; Li et al., 2020). However, upon reviewing the literature, there is a distinct lack of theoretical research, especially regarding the risk preferences and investment behaviors of institutional investors with CSR goals. Concerning this, the fourth research question emerges, which is replied by the fourth paper:

4. How to model the risk preference of insurance companies who have a CSR investment goal? How do they allocate their assets? Does a higher target lead to more investments in the green bond?

This dissertation embarks on a comprehensive examination of the interplay between retirement product design, marketing strategies, and investment approaches in life insurance companies. By unraveling the complexities inherent in these domains, this dissertation aims to contribute valuable insights to academic scholarship, industry practitioners, and policymakers seeking to enhance the efficacy and sustainability of retirement solutions offered by life insurance companies.

In the next section, we summarize the four research papers respectively.

### 3 Summary of Research Papers

#### Research Paper 1: Care-dependent tontines

In the first research paper, we seek the help of insurance tools to meet the challenges of an aging society and LTC issues. In particular, we proposed a new type of retirement tontine called “care-dependent tontines” that combines LTC business with retirement tontines. It transfers a portion of longevity risk to policyholders while offers enhanced payments when they become care-dependent. By applying data from China Health and Retirement Longitudinal Study and referring to the capital requirement of China Risk Oriented Solvency System, we compute the transition probabilities and risk loadings on a realistic basis. Further, we show the attractiveness of the care-dependent tontines through numerical comparison with care-dependent annuities. The results are robust in a different regulation scheme – Solvency II. Our research contributes to advancing the adoption of elderly care insurance by introducing compelling care-dependent insurance products, thereby addressing the challenges posed by the deepening aging trend.

This paper is a joint work with An Chen and Xian Xu. It has been accepted for publication in the *Insurance: Mathematics and Economics*. Overall, this paper answers the first question.

Long-term care (LTC) insurance emerges as a critical component in responding to the escalating demand for care in later life. While public health and LTC insurance schemes cover some aspects of LTC expenses, concerns about the sustainability of these services persist, given the increasing number of disabled elderly individuals. The development of a private LTC insurance market becomes imperative, yet the efficacy of stand-alone LTC insurance is questionable. Existing practices from OECD countries indicate a limited market for private LTC insurance, largely due to high premiums and policy exclusions that deter potential buyers.

Addressing these challenges, scholars have proposed integrating life annuities and LTC insurance into a single product. This integration not only expands the accessibility of LTC coverage, especially for those traditionally rejected by stand-alone LTC insurance, but also contributes to the enrichment of the old-age insurance market. Moreover, combining these products reduces costs for both insurers and consumers. Empirical evidence suggests that such bundled products significantly reduce premiums relative to stand-alone offerings.

Despite the cost-effectiveness of care-dependent annuities compared to separate purchases of life annuities and LTC products, the market for these integrated products remains limited. Adverse selection and challenges in determining reserves for these products contribute to this constraint. To address these issues, the concept of care-dependent tontines has been introduced, where policyholders bear most of the longevity risks. This paper proposes two designs for care-dependent tontines and explores their attractiveness to policyholders through fair premium calculations and a utility framework.

Analyzing various care-dependent products, the paper finds that care-dependent annuities are the optimal choice under actuarially fair pricing. However, considering realistic risk loadings, care-dependent tontines prove more attractive to policyholders, offering lower gross premiums while maintaining the same expected lifetime utility. The study also identifies preferences based on risk aversion coefficients, highlighting the suitability of different care-dependent products for individuals with varying risk profiles. Additionally, the findings suggest that constructing a care-dependent tontine in a two-pool structure may be the most cost-efficient approach for insurers. The study's insights contribute valuable perspectives on designing care-dependent insurance contracts, fostering the growth of the elderly care insurance market and meeting the diverse demands for care insurance products.

## **Research Paper 2: Dynamic tonuity: Adapting retirement benefits to a changing environment**

The second paper presents a potential solution to address the growing need for innovative retirement products. In this paper, we numerically demonstrate that the regular tonuity designed by Chen et al. (2019) may result in a loss of policyholders' utility in the event of an unforeseen mortality or longevity shock. This is attributed to its fixed switching time at the contract initiation.

This collaborative effort is jointly undertaken with An Chen and Manuel Rach. In general, this paper responds to the second research question.

To address concerns related to the regular tonuity, we propose a modified design known as "dynamic tonuity". It automatically switches from tontine-specific payoffs to annuity-specific payoffs based on a predetermined condition linked to mortality experiences. The switching time adjusts dynamically to changes in mortality occurring after the contract's initiation, and the switching condition can be transparent to policyholders. For instance, it might be triggered by the one-year survival probability of a reference population surpassing a specified threshold index.

Assuming policyholders' utility follows a constant relative risk aversion (CRRA) function, we analytically solve optimization problems to obtain optimal payoff structures for dynamic tonuity. Utilizing a two-population mortality model and mortality data from UK Continuous Mortality Investigation and Human Mortality Database, we estimate correlated mortality rates for both the national population (reference population) and the insurer's annuitant pool (book population). We explore and compare different tonuity products under various scenarios, including a base case and two stress cases involving unexpected mortality and longevity shocks respectively.

Numerical results highlight that a well-designed dynamic tonuity can outperform regular tonuity in a changing environment. The policyholder's risk aversion plays a crucial role, influencing contract preferences in stress cases. In instances where the policyholder's risk aversion coefficient surpasses 1, a more favorable perception of the dynamic tonuity contract in the stress case involving the unexpected mortality shock is associated with a diminished level of risk aversion. This inclination likely stems from an increased tolerance for the dynamic tonuity's fluctuating payoffs. Additionally, sensitivity analysis suggests that a reduced pool size can increase the attractiveness of the dynamic tonuity contract compared to the regular tonuity contract during the unexpected mortality shock. This pattern may be influenced by the allure of tontine-like payoffs in the dynamic tonuity, particularly when a smaller pool size coincides with an unexpected mortality shock triggering a later dynamic switching time. In the stress case featuring an unforeseen longevity shock, for policyholders with a risk aversion coefficient exceeding 1, the outcomes closely parallel those observed in the stress case with an unexpected mortality shock. The main difference locates in the possibility that a reduced pool size may make the dynamic tonuity contract less favorable during an unexpected longevity shock, as it prompts an earlier dynamic switching time.

When the policyholder's risk aversion coefficient falls within the interval  $(0, 1)$ , consistent results are observed regarding the unexpected mortality shock. However, a deviation in contract attractiveness arises concerning the unexpected longevity shock. In the event of an unforeseen longevity shock, the regular tonuity contract outperforms the dynamic tonuity contract in attractiveness. This can be attributed to policyholders with a risk aversion coefficient in the range of  $(0, 1)$  expecting a tontine contract. Note that payoffs of the regular and dynamic tonuity contract are determined under base-case assumptions. For such policyholders, the regular tonuity contract has already become a tontine as it is optimal not to undergo switching in the base case and it remains as a tontine when an unexpected longevity shock strikes. However, the dynamic tonuity contract's switching point is independent of the policyholder's preference but is determined by an external event, influencing the premium distribution between annuity and tontine components in the base case. The unforeseen longevity shock results in an earlier-than-expected shift in dynamic switching time. Throughout the period from the initial switching time to the stressed switching time, the initial tontine-like payoffs are substituted with annuity-like payoffs, reducing the attractiveness of the dynamic tonuity contract in comparison to the regular one.

### **Research Paper 3: How does the insurer's mobile application sales strategy perform?**

In response to the challenge of digital transformation in the insurance industry, insurance companies have adopted the MA sales strategy in product marketing. The third paper empirically

investigates the impact of MA sales strategy on the sales performance of insurers, shedding light on underlying mechanisms.

This paper is a joint work with An Chen, Finbarr Murphy, Wei Xu and Xian Xu. It has been accepted for publication in the *Journal of Risk and Insurance*. In general, the third paper delivers answers to the third research question.

The focus of this paper is on the insurance purchase quantity and premiums received from new policies, serving as key performance measures. The choice of these metrics is justified by insurers' interest in both the total number of policies and premiums received, which are commonly used indicators in prior studies (e.g., Hammond et al., 1967; Mantis and Farmer, 1968; Ferber and Lee, 1980; Browne and Kim, 1993; Li et al., 2007; Kjosevski, 2012).

To address the third research question, a unique dataset of term life insurance purchase records from 2015 to 2019 is analyzed, provided by a representative Chinese life insurance company. The data is transformed into a panel dataset at the prefecture-day level to account for time trends and regional variations. A two-way fixed effect regression model is employed to isolate the impact of the MA sales strategy from unobservable factors.

The results indicate a significant increase in the insurance purchase quantity and premiums received from new policies, following the implementation of the MA sales strategy. Channel accessibility and cost reduction are identified as contributing factors to the growth, with the MA channel exhibiting higher accessibility and lower costs compared to the offline channel. Heterogeneity analyses reveal non-monotonic relationships, such as the inverted U-shape effect of education level on the strategy's impact, which diverges from the conclusions drawn in the literature (e.g., Truett and Truett, 1990; Browne and Kim, 1993; Kjosevski, 2012). Financial professionals and self-insured individuals respond more positively, while gender differences also play a role in the strategy's effectiveness.

Moreover, the study explores impulsive purchasing behavior and its correlation with the MA sales strategy. The strategy is found to reduce customers' impulsive purchases, leading to a lower policy surrender ratio during the cancellation period, which is consistent with findings of the existing literature (e.g., Gilly and Wolfinbarger, 2000; Wolfinbarger and Gilly, 2001; Hu and Tang, 2014). While a significant body of marketing science literature suggests that the inclusion of the Internet channel leads to substitution effects, impacting both consumer demand and salesperson motivation (e.g., Cather and Howe, 1989; Garven, 2002; Sharma and Gassenheimer, 2009; Torkestani et al., 2018), our findings present a contrasting perspective. Our results indicate that the substitution effect of the MA sales strategy on insurer's sales performance is time-constrained, occurring only in the introductory year and even reversing within one year thereafter. This outcome is reassuring, suggesting that concerns among insurance practitioners about the overstated substitution effect of introducing the MA sales strategy may be unwarranted.

In summary, this paper comprehensively analyzes the impact of the MA sales strategy on the insurer's sales performance, taking into account various factors and providing valuable insights for insurance practitioners.

## **Research Paper 4: Investing in green bonds with goal-oriented preferences**

In the contemporary landscape, climate change has risen to the forefront as a critical concern. This has prompted a discernible shift within the investment sector towards CSR investments, notably marked by a surge in the popularity of green bonds. Institutional investors, in particular, are prominently engaged in this trend, aligning their portfolios with CSR investment objectives. Despite the growing importance of green bonds, the existing literature lacks comprehensive research on the risk preferences of institutional investors focused on green bonds and holding CSR goals. A representative study, Pástor et al. (2021), have introduced equilibrium models incorporating non-pecuniary benefits into investors' utility, emphasizing the positive effects of stronger environmental, social, and governance (ESG) characteristics. In contrast, the fourth paper explores the impact of green bond institutional investors' (abbreviated as green bond investors) risk preferences from a distinct perspective.

This paper is a collaborative effort involved with An Chen, Thai Nguyen and Gazi Salah Uddin. On the whole, this paper responds to the fourth research question.

In this paper, we introduce a unique approach by considering green bond investors who harbor CSR investment targets, implying goal-based preferences. Drawing on the works of Kőszegi and Rabin (2006, 2009), our study aims to investigate whether setting a more ambitious goal correlates with increased investments in green bonds. Specifically, a CSR investment goal is embedded in the green bond investor's utility, where exceeding the goal results in utility gains, while falling short leads to utility losses.

Our model incorporates the assumption that investors exhibit loss aversion, with a stronger aversion to utility losses than gains. The utility component of terminal wealth is modeled using a CRRA utility with a specified CRRA level denoted as  $\gamma$ . The gain-or-loss component employs a piece-wise linear function.

Building upon Merton's asset allocation framework, we extend the model by integrating CSR investment targets into investors' risk preferences. This extension enables the determination of optimal wealth fractions invested in green bonds and stocks.

Our numerical results demonstrate that when green bond and stock prices exhibit negative correlations, investors significantly increase their green bond holdings as maturity approaches due to the diminished volatility of green bonds. Additionally, reduced loss aversion prompts



investors to retain larger portions in both green bonds and stocks. The effect of CSR investment goals on green bond investors' holdings varies based on risk aversion. Investors with  $\gamma > 1$  increase their holdings with higher CSR targets, prioritizing utility gain-or-loss measurements in their overall utility. In contrast, those with  $\gamma \in (0, 1)$  decrease their allocations, emphasizing the utility component derived from terminal wealth.

We further explore the influence of positive correlations between stock and green bond prices. Notably, only a strong correlation, such as 80%, generates significant variations. Green bond investors adopt a strategy of short selling the green bond while simultaneously taking a long position in the stock. Those exhibiting heightened loss aversion engage in short selling a smaller fraction of their wealth in green bonds and invest a correspondingly reduced fraction in stocks. When faced with an elevated CSR investment goal, green bond investors with  $\gamma > 1$  increase their short selling of green bonds and amplify their investments in stocks. Conversely, investors with  $\gamma \in (0, 1)$  exhibit the opposite pattern, reducing short selling of green bonds and decreasing investments in stocks as the CSR investment goal rises.



## References

- AIC and IPLE (2020). 2018-2019 China long-term care research report. Technical report.
- AXA (2022). 2021 EUR 1bn green bond. Technical report.
- Bachelet, M. J., Becchetti, L., and Manfredonia, S. (2019). The green bonds premium puzzle: The role of issuer characteristics and third-party verification. *Sustainability*, 11(4):1098.
- Benlagha, N. and Hemrit, W. (2020). Internet use and insurance growth: Evidence from a panel of OECD countries. *Technology in Society*, 62:101289.
- Braga, J. P., Semmler, W., and Grass, D. (2021). De-risking of green investments through a green bond market—empirics and a dynamic model. *Journal of Economic Dynamics and Control*, 131:104201.
- Brown, J. and Warshawsky, M. (2013). The life care annuity: A new empirical examination of an insurance innovation that addresses problems in the markets for life annuities and long-term care insurance. *Journal of Risk and Insurance*, 80(3):677–704.
- Brown, J. R. and Finkelstein, A. (2009). The private market for long-term care insurance in the United States: A review of the evidence. *Journal of Risk and Insurance*, 76(1):5–29.
- Brown, J. R. and Goolsbee, A. (2002). Does the internet make markets more competitive? Evidence from the life insurance industry. *Journal of Political Economy*, 110(3):481–507.
- Browne, M. J. and Kim, K. (1993). An international analysis of life insurance demand. *Journal of Risk and Insurance*, 60(4):616–634.
- Cather, D. A. and Howe, V. (1989). Conflict and channel management in property-liability distribution systems. *Journal of Risk and Insurance*, 56(3):535–543.
- Chen, A., Fuino, M., Sehner, T., and Wagner, J. (2022). Valuation of long-term care options embedded in life annuities. *Annals of Actuarial Science*, 16(1). 68–94.
- Chen, A., Hieber, P., and Klein, J. K. (2019). Tonuity: A novel individual-oriented retirement plan. *ASTIN Bulletin: The Journal of the IAA*, 49(1):5–30.
- Chen, A. and Rach, M. (2019). Options on tontines: An innovative way of combining annuities and tontines. *Insurance: Mathematics and Economics*, 89:182–192.

- Chen, A., Rach, M., and Sehner, T. (2020). On the optimal combination of annuities and tontines. *ASTIN Bulletin: The Journal of the IAA*, 50(1):95–129.
- Chiesa, M. and Barua, S. (2019). The surge of impact borrowing: The magnitude and determinants of green bond supply and its heterogeneity across markets. *Journal of Sustainable Finance and Investment*, 9(2):138–161.
- Cognizant (2015). Insurers use social mobile apps to increase digital value. Technical report, Cognizant.
- Donnelly, C. and Young, J. (2017). Product options for enhanced retirement income. *British Actuarial Journal*, 22(3):636–656.
- Eling, M. and Lehmann, M. (2018). The impact of digitalization on the insurance value chain and the insurability of risks. *The Geneva papers on risk and insurance-issues and practice*, 43(3):359–396.
- Ferber, R. and Lee, L. C. (1980). Acquisition and accumulation of life insurance in early married life. *Journal of Risk and Insurance*, 47(4):713–734.
- Fullmer, R. K. and Sabin, M. J. (2019). Individual tontine accounts. *Journal of Accounting and Finance*, 19(8).
- Garven, J. R. (2002). On the implications of the Internet for insurance markets and institutions. *Risk Management and Insurance Review*, 5(2):105–116.
- Gemmo, I., Rogalla, R., and Weinert, J. H. (2020). Optimal portfolio choice with tontines under systematic longevity risk. *Annals of Actuarial Science*, 14(2):302–315.
- Gianfrate, G. and Peri, M. (2019). The green advantage: Exploring the convenience of issuing green bonds. *Journal of Cleaner Production*, 219:127–135.
- Gilly, M. C. and Wolfinbarger, M. (2000). A comparison of consumer experiences with online and offline shopping. *Consumption, Markets and Culture*, 4(2):187–205.
- Hachenberg, B. and Schiereck, D. (2018). Are green bonds priced differently from conventional bonds? *Journal of Asset Management*, 19(6):371–383.
- Hammond, J., Houston, D. B., and Melander, E. R. (1967). Determinants of household life insurance premium expenditures: An empirical investigation. *Journal of Risk and Insurance*, 34(3):397–408.
- Han, Y. and Li, J. (2022). Should investors include green bonds in their portfolios? Evidence for the USA and Europe. *International Review of Financial Analysis*, 80:101998.
- Hoermann, G. and Ruß, J. (2008). Enhanced annuities and the impact of individual underwriting on an insurer’s profit situation. *Insurance: Mathematics and Economics*, 43(1):150–157.

- Hu, Y. J. and Tang, Z. (2014). The impact of sales tax on Internet and catalog sales: Evidence from a natural experiment. *International Journal of Industrial Organization*, 32:84–90.
- Huang, L., Lu, X., and Ba, S. (2016). An empirical study of the cross-channel effects between web and mobile shopping channels. *Information and Management*, 53(2):265–278.
- Karpf, A. and Mandel, A. (2018). The changing value of the “green” label on the US municipal bond market. *Nature Climate Change*, 8(2):161–165.
- Kjosevski, J. (2012). The determinants of life insurance demand in central and southeastern Europe. *International Journal of Economics and Finance*, 4(3):237–247.
- Kőszegi, B. and Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4):1133–1165.
- Kőszegi, B. and Rabin, M. (2009). Reference-dependent consumption plans. *American Economic Review*, 99(3):909–36.
- Lazaris, C., Vrechopoulos, A. P., Doukidis, G. I., and Fraidaki, A. (2015). Mobile apps for omnichannel retailing: Revealing the emerging showroom phenomenon. In *MCIS 2015 Proceedings*, page 12.
- Li, D., Moshirian, F., Nguyen, P., and Wee, T. (2007). The demand for life insurance in OECD countries. *Journal of Risk and Insurance*, 74(3):637–652.
- Li, Z., Tang, Y., Wu, J., Zhang, J., and Lv, Q. (2020). The interest costs of green bonds: Credit ratings, corporate social responsibility, and certification. *Emerging Markets Finance and Trade*, 56(12):2679–2692.
- Luo, X., Andrews, M., Fang, Z., and Phang, C. W. (2014). Mobile targeting. *Management Science*, 60(7):1738–1756.
- Mantis, G. and Farmer, R. N. (1968). Demand for life insurance. *Journal of Risk and Insurance*, 35(2):247–256.
- Milevsky, M. A. and Salisbury, T. S. (2015). Optimal retirement income tontines. *Insurance: Mathematics and Economics*, 64:91–105.
- Morrow, K. M. and Röger, W. (2003). Economic and financial market consequences of ageing populations. Technical Report 182.
- Murtaugh, C. M., Spillman, B. C., and Warshawsky, M. J. (2001). In sickness and in health: An annuity approach to financing long-term care and retirement income. *Journal of Risk and Insurance*, 68(2):225–254.
- Nanayakkara, M. and Colombage, S. (2019). Do investors in green bond market pay a premium? Global evidence. *Applied Economics*, 51(40):4425–4437.

- Newfield, P. (2014). The tontine: An improvement on the conventional annuity? *The Journal of Retirement*, 1(3):37–48.
- Nguyen, T. T. H., Naeem, M. A., Balli, F., Balli, H. O., and Vo, X. V. (2021). Time-frequency comovement among green bonds, stocks, commodities, clean energy, and conventional bonds. *Finance Research Letters*, 40:101739.
- OECD (2023). Health at a glance: OECD indicators. Technical report.
- Park, D., Park, J., and Ryu, D. (2020). Volatility spillovers between equity and green bond markets. *Sustainability*, 12(9):3722.
- Partridge, C. and Medda, F. R. (2020). The evolution of pricing performance of green municipal bonds. *Journal of Sustainable Finance and Investment*, 10(1):44–64.
- Pástor, L., Stambaugh, R. F., and Taylor, L. A. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2):550–571.
- Piggott, J., Valdez, E. A., and Detzel, B. (2005). The simple analytics of a pooled annuity fund. *Journal of Risk and Insurance*, 72(3):497–520.
- Pitacco, E. (2015). Guarantee structures in life annuities: A comparative analysis. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 41(1):78–97.
- Ramsay, C. M. and Oguledo, V. I. (2020). Doubly enhanced annuities (DEANs) and the impact of quality of long-term care under a multi-state model of activities of daily living (ADL). *North American Actuarial Journal*, 24(1):57–99.
- Reboredo, J. C. (2018). Green bond and financial markets: Co-movement, diversification and price spillover effects. *Energy Economics*, 74:38–50.
- Reboredo, J. C. and Ugolini, A. (2020). Price connectedness between green bond and financial markets. *Economic Modelling*, 88:25–38.
- Richter, A. and Weber, F. (2011). Mortality-indexed annuities managing longevity risk via product design. *North American Actuarial Journal*, 15(2):212–236.
- Sabin, M. J. (2010). Fair tontine annuity. Available at SSRN 1579932.
- Sharma, D. and Gassenheimer, J. B. (2009). Internet channel and perceived cannibalization: Scale development and validation in a personal selling context. *European Journal of Marketing*, 43(7/8):1076–1091.
- Stahl, D. O. (1989). Oligopolistic pricing with sequential consumer search. *The American Economic Review*, 79(4):700–712.
- Tang, D. Y. and Zhang, Y. (2020). Do shareholders benefit from green bonds? *Journal of Corporate Finance*, 61:101427.

- Torkestani, M. S., Zandmehr, M., and Afsahizadeh, M. (2018). Investigating the relationship between website characteristics and impulsive purchase on the Internet. *Journal of Business Administration Researches*, 10(19):1–17.
- Truett, D. B. and Truett, L. J. (1990). The demand for life insurance in Mexico and the United States: A comparative study. *Journal of Risk and Insurance*, 57(2):321–328.
- Vidal-Melia, C., Ventura-Marco, M., and Pla-Porcel, J. (2016). Life care annuities (LCA) embedded in a notional defined contribution (NDC) framework. *ASTIN Bulletin*, 46(2):331–363.
- Webb, D. C. (2009). Asymmetric information, long-term care insurance, and annuities: The case for bundled contracts. *Journal of Risk and Insurance*, 76(1):53–85.
- Weinert, J. H. and Gründl, H. (2021). The modern tontine: An innovative instrument for longevity risk management in an aging society. *European Actuarial Journal*, 11(1):49–86.
- Wisniewski, J. (2011). Control-shift-mobile that works for your library. *Online-Medford*, 35(1):54.
- Wolfinbarger, M. and Gilly, M. C. (2001). Shopping online for freedom, control, and fun. *California Management Review*, 43(2):34–55.
- Zhou-Richter, T. and Gründl, H. (2011). Life care annuities - Trick or treat for insurance companies? Available at SSRN 1856994.





## **Research Papers**



# 1 Care-dependent tontines

**Source:**

Chen, A., Chen, Y., & Xu, X. (2022). Care-dependent tontines. *Insurance: Mathematics and Economics*, 106, 69–89.

URL: <https://www.sciencedirect.com/science/article/pii/S0167668722000609>

DOI: <https://doi.org/10.1016/j.insmatheco.2022.05.002>

© 2022 Elsevier B.V. All rights reserved.

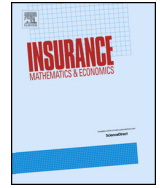




Contents lists available at ScienceDirect

## Insurance: Mathematics and Economics

www.elsevier.com/locate/ime



## Care-dependent tontines

An Chen<sup>a</sup>, Yusha Chen<sup>a,\*</sup>, Xian Xu<sup>b</sup><sup>a</sup> Institute of Insurance Science, Ulm University, Helmholtzstr. 20, 89069 Ulm, Germany<sup>b</sup> School of Economics, Fudan University, Guoquan Road 600, 200433 Shanghai, China

## ARTICLE INFO

## Article history:

Received August 2021

Received in revised form November 2021

Accepted 9 May 2022

Available online 13 May 2022

## JEL classification:

G22

G13

## Keywords:

Tontine

Care-dependent

Lifetime utility

Optimal retirement product

Annuity

## ABSTRACT

With the gradual deepening of aging, the affordability of long-term care (LTC) services in aging societies will become increasingly questionable. Both private stand-alone LTC insurance and care-dependent annuity do not seem to provide efficient solutions, due to their high risk charges. In this article, we propose two ways of combining the long-term care business with retirement tontines, i.e. care-dependent tontines, which shift a part of longevity risk to policyholders and provide increased payments in care-dependent state. We determine the optimal payment structures of these products that maximize the individuals' expected lifetime utility. Using realistic risk loadings based on data from China Health and Retirement Longitudinal Study (CHARLS) and capital requirement of China Risk Oriented Solvency System (C-ROSS), it is shown that both care-dependent tontines present as better choices for individuals, comparing to care-dependent annuities. The results are robust in a different regulation scheme – Solvency II. Our findings contribute to improving the penetration of elderly care insurance with appealing care-dependent insurance products.

© 2022 Elsevier B.V. All rights reserved.

## 1. Introduction

The aging wave has been sweeping the world. With continuous improvement of medical and social security services, people have been living longer in expectation. Although this is great progress for humanity, it raises plenty of aging problems in many countries. For instance, the population aging speed in China is one of the fastest among countries. Since entering an aging era in 2000, the degree of population aging in China has been rising. According to a report published by AIC and IPLE (2020), in 2019, among the Chinese people aged over 60 in the survey area, there were 7% in moderate Activities of Daily Living (ADL)<sup>1</sup> disability status and 4.8% in severe ADL disability. Besides, the prediction statistics from Statista (Statista, 2021) have pointed out that China's population over 60 will rise from 17.4% of the total population in 2020, to 29.9% in 2040 with a rather low fertility rate.

Long-term care (LTC) insurance is born out of the long-term need for care, especially in later life. It assists retirees to cover expenses they may need with care at home or in a facility when they cannot perform daily living activities. Although some types of LTC expenditure have already been covered in the public health/LTC insurance scheme, the increasing number of the disabled elderly has prompted extensive concerns about the future affordability of LTC services (Morrow and Röger, 2003). In this sense, the development of a private LTC insurance market has become even more important and urgent. However, industry practice experiences from OECD countries have proved that the private long-term care insurance is not an effective solution as it only accounts for a small market. Not much LTC insurance is issued to a relatively high percentage of potential buyers at the retirement ages, for the sake of reducing the self-bearing risks of the issuers.<sup>2</sup> These policy exclusions (Brown and Warshawsky, 2013) exactly keep those who need the LTC insurance away. Besides, LTC policies are offered at premiums that are substantially higher than the actuarially fair level. According to Brown and Finkelstein (2009), the loading can be enormously high, particularly when taking account of the lapse of policies. In the U.S. market, it raises the loading of 18 cents on the dollar to 51 cents on the dollar.

\* Corresponding author.

E-mail addresses: [an.chen@uni-ulm.de](mailto:an.chen@uni-ulm.de) (A. Chen), [yusha.chen@uni-ulm.de](mailto:yusha.chen@uni-ulm.de) (Y. Chen).<sup>1</sup> ADL refers to people's daily self-care activities such as bathing, dressing, eating, continence, toileting, and transferring.<sup>2</sup> The LTC insurance is intensively underwritten by its issuers (Brown and Warshawsky, 2013).

On the basis of these facts, Murtaugh et al. (2001) put forward empirically the idea of integrating the life annuity and LTC insurance into one product. Other scholars also come up with different ways to combine lifetime annuities and LTC insurance (Brown and Warshawsky, 2013; Vidal-Melia et al., 2016; Pitacco, 2015; Hoermann and Ruß, 2008; Ramsay and Oguledo, 2020; Chen et al., 2021, etc.). In fact, by combining the life annuity and LTC coverage, it may make the long-term insurance product available to more people, especially those who are currently rejected by the stand-alone LTC insurance (e.g. those with poor health or unhealthy lifestyles). And also, more available insurance products may contribute to enriching products in the old age insurance market as well as increasing the penetration of elderly care insurance. In addition, it reduces the cost of both stand-alone insurance products. Through the estimation of Murtaugh et al. (2001), such bundled products significantly reduce premiums by 3–5 percent relative to the stand-alone products.

Even if the care-dependent annuities are typically less expensive than buying life annuity and long-term care products separately, the market for these products is still rather limited. According to Webb (2009), these bundled products only make up less than 10% of the voluntary annuity market in the United States. One reason is that adverse selection still exists in the care-dependent annuity market (Zhou-Richter and Gründl, 2011).<sup>3</sup> Another possible reason is that it is rather challenging for insurers to determine reserves for these products, which might also lead to high administration and risk charges. One possible way for the insurer to transfer a part of these risks to policyholders is to combine the long-term care business with retirement tontines, see e.g. Hieber and Lucas (2020) putting forward a care-dependent tontine scheme. Under the framework of tontines, it is policyholders that bear most of the longevity risks, stated differently, the “mortality credit” of the dead provides more payments for the alive. For the LTC coverage, it requires a higher payment when the policyholder moves to a severely sick state, while the mortality rate in this state will also increase. Thus, the increased care-dependent payment could be viewed as an advance of additional “mortality credit”.<sup>4</sup> The focus of Hieber and Lucas (2020) is on how to distribute the mortality credit such that fairness can be achieved among the individuals.

In contrast to Hieber and Lucas (2020), in this article, we propose two ways of designing the care-dependent tontines (CDT): (i) consider that all insured members (both healthy ones and severely sick ones) are in one pool; (ii) the insured members are divided into two groups: the healthy, and the severely sick. At each time  $t$ , we reallocate the ones whose states get changed into the corresponding groups. Considering the fact that there are three main categories of LTC insurance products currently sold in the private LTC insurance market, i.e. predetermined benefits, reimbursement benefits and service benefits (Denuit et al., 2019). Our products can be regarded as care insurance products with predetermined benefits. Although we do not explicitly model the care costs in the benefits, in designing our products, we do take account of the fact that people in care-dependent state have increasing liquidity need for care costs. Individuals with various risk aversion can choose those payments maximizing their own expected lifetime utility. In other words, this article stands on the policyholder's side and describes the optimal decisions of payment streams that the policyholder would prefer. Based on fair premium calculations and a utility framework (Chen and Rach, 2019; Chen et al., 2020; Chen and Rach, 2021), we determine the structures of the CDTs that maximize the policyholder's expected lifetime utility with constant relative risk aversion and no bequest motives. Furthermore, in this article, we compare different care-dependent products analytically and numerically.

When considering an actuarially fair premium, we find that the care-dependent annuity (CDA) is the best choice among different care-dependent products. Adding care-dependent payoffs to the regular retirement products does not change the preference order of the tontine and the annuity under actuarially fair pricing (Milevsky and Salisbury, 2015; Chen et al., 2019). In order to conduct a realistic comparison among the three care-dependent products, we rely on the data of China Health and Retirement Longitudinal Study (CHARLS) and China Risk Oriented Solvency System (C-ROSS). We compute the risk loadings following C-ROSS for various care-dependent products under consideration. Our results reveal that the two types of CDTs are more attractive than CDA for policyholders, when taking account of the realistic risk loadings. More specifically, we show that the two CDTs lead to lower gross premiums than the CDA product, while ensuring the same expected lifetime utility to the policyholders as the CDA product. Besides, as the pool size grows larger, the advantage of CDTs over CDA becomes more substantial. Further, we find that the two-pool CDT is more attractive to those who are less risk-averse, while one-pool CDT is a better choice for those with larger risk aversion coefficients. Finally, we detect that for insurers, under our baseline parameter setting, constructing a care-dependent tontine in a two-pool structure may be the most cost-efficient way to add the care-dependent coverage among the three products studied in this article. In order to examine the possible robustness of our results under different regulation regimes, we also compute the risk loadings with reference to the capital requirements of Solvency II, the European insurance regulation framework. Using the baseline parameter setting, it still shows that the CDTs are preferable to CDA, which means our results are to some extent robust across regulation regimes. Our findings provide the insurers with considerable views on the design of care-dependent insurance contracts, which may boost the development of elderly care insurance market and contribute to fulfilling the increasing demand of diversified care insurance products.

The remainder of this article is structured as follows: in Section 2, we introduce our transition probabilities used throughout the entire article. In Section 3, we derive the optimal payoffs of the care-dependent tontines from two different angles respectively. Additionally, for the sake of product comparison later, we also give the optimal payoffs of the care-dependent annuities under our model settings. Section 4 sets up the framework for the product comparison under a realistic setting. And Section 5 presents the numerical analysis, describing the used statistics and results. In the last Section, we conclude the article. Several proofs, a statistical procedure to determine the transition probabilities, the risk loading computation following the capital requirement of Solvency II, and additional tables concerning the sensitivity analysis are listed in the appendix.

## 2. Transition probabilities

In this section, we introduce the transition probabilities used throughout the entire article on the basis of a three-state model. The transition probabilities reflect the policyholder's health condition.

<sup>3</sup> An adverse selection is said to exist in the care-dependent annuity market when people with both high longevity and morbidity risk are more likely to purchase such products, which will force the insurers to increase their premiums accordingly. The adverse selection in annuities is reduced (Murtaugh et al., 2001; Spillman et al., 2003; Brown and Warshawsky, 2013), but not eliminated in the care-dependent annuity market (Zhou-Richter and Gründl, 2011).

<sup>4</sup> Mortality of the severely sick ones occurs more likely than that of healthy ones, and the corresponding “mortality credits” are larger. The additional “mortality credit” here represents the excess part of sick ones' “mortality credit” over the healthy ones'.

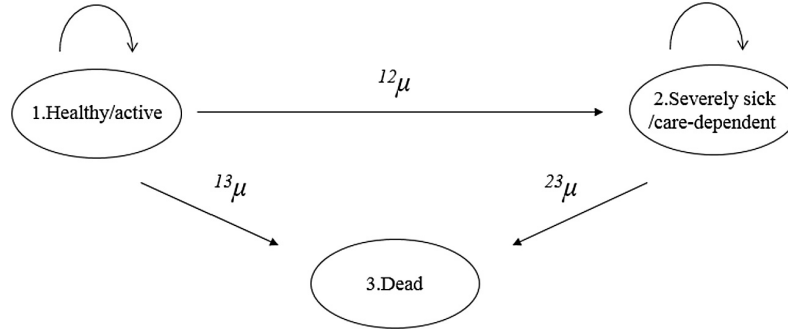


Fig. 2.1. State set and transition intensities.

Let  $x$  be the initial age, and  $S_{x+t}$  be a Markovian process describing the development of a single policy in continuous time. Then the LTC insurance is modeled by a multi-state model with state space  $\mathcal{S} = \{1 = \text{healthy/active}, 2 = \text{severely sick/care-dependent}, 3 = \text{dead}\}$ , and a set of transitions according to Fig. 2.1. We assume the policyholder is healthy at the initial age, i.e.  $S_x \equiv 1$ .

Given that a policyholder is in state  $i$  at age  $y$ , the transition probabilities of this policyholder being in state  $j$  at age  $y + t$ , are defined as:

$${}^i_t p_y = \mathbb{P}\{S_{y+t} = j \mid S_y = i\} \quad i, j \in \mathcal{S}, i \neq j. \quad (2.1)$$

The corresponding transition intensities could be written as (Haberman and Pitacco, 1998):

$${}^i_j \mu_y = \lim_{t \rightarrow 0} \left( \frac{{}^i_j p_y}{t} \right), \quad i, j \in \mathcal{S}, i \neq j. \quad (2.2)$$

Furthermore, the probability of a policyholder being in state  $i$  at age  $y$  and remaining in state  $i$  till age  $y + t$  is:

$${}^i_i p_y = \mathbb{P}\{S_{y+t} = i, \forall t \mid S_y = i\}, \quad i \in \mathcal{S}. \quad (2.3)$$

Given the transition intensities of the multi-state model, we obtain the probabilities (2.1) and (2.3) by Kolmogorov forward differential equations (Haberman and Pitacco, 1998).

$${}^1_1 p_y = \exp \left\{ - \int_0^t [{}^{12}\mu_{y+s} + {}^{13}\mu_{y+s}] ds \right\}, \quad (2.4)$$

$${}^1_2 p_y = \int_0^t [{}^1_1 p_y \cdot {}^{12}\mu_{y+s} \cdot {}^{22}p_{y+s}] ds, \quad (2.5)$$

$${}_t p_y = {}^1_1 p_y + {}^1_2 p_y, \quad (2.6)$$

$${}^2_2 p_y = \exp \left\{ - \int_0^t {}^{23}\mu_{y+s} ds \right\}. \quad (2.7)$$

- ${}^1_1 p_y$  is the occupancy probability of a healthy individual aged  $y$  staying healthy at age  $y + t$ ;
- ${}^1_2 p_y$  denotes the probability of a healthy individual aged  $y$ , becoming severely sick between  $(0, t)$  and remaining severely sick until  $t$ ;
- ${}_t p_y$  is the  $t$ -year survival probability of a healthy individual aged  $y$ , which includes the case that he remains healthy for  $t$  years and the case that he becomes severely sick during the  $t$ -year time;
- while  ${}^2_2 p_y$  represents the occupancy probability of a  $y$ -year-old sick individual staying severely sick at age  $y + t$ .

Here, we disregard the possibility of recovery from the severely sick state, i.e. the transition  $\{2 \rightarrow 1\}$  is not considered (Pitacco, 2016; Hanewald et al., 2019; Chen et al., 2021).

### 3. Care-dependent products

Consider a retiree at age  $x$  endowed with a wealth level  $v > 0$  at his retirement date (in our framework at time  $t = 0$ ). Following Yaari (1965)'s setting, we assume a continuous-time stream of payments for the care-dependent tontines. In a care-dependent tontine contract, we further assume that the risks are shared among homogeneous policyholders who are assumed to be identical copies of each other. There are  $n$  policyholders initially. To construct the care-dependent tontine, we put forward two ways of designing the product: (i) consider that all insured members (both healthy ones and severely sick ones) are in one pool; (ii) the insured members are divided into two groups: the healthy, and the severely sick. We will introduce these two kinds of care-dependent tontines in the following subsections respectively.

## 3.1. One-pool care-dependent tontine

Firstly, we consider all the members in only one pool at each time  $t$ . Here the one pool means that only alive and dead persons are identified at each time  $t$ , while different payment streams are provided to the healthy and the severely sick respectively. We use the notations  $d_1^{oc_1}(t)$  and  $d_2^{oc_1}(t)$  for the payment streams to the healthy and the severely sick separately, which are pre-determined when the policy is underwritten. Correspondingly,  $b_1^{oc_1}(t)$  and  $b_2^{oc_1}(t)$  are the payoffs for the healthy members and the severely sick members. In accordance to the tontine scheme stated in Milevsky and Salisbury (2015), we define the care-dependent tontine payoff for the policyholders in the healthy state as

$$b_1^{oc_1}(t) := \begin{cases} \mathbb{1}_{\{S_{x+t}=1\}} \frac{nd_1^{oc_1}(t)}{N(t)}, & \text{if } N(t) > 0, \\ 0, & \text{else} \end{cases}, \quad (3.1)$$

where  $N(t)$  denotes the number of alive policyholders at time  $t$ , containing both healthy and severely sick ones.

In another case, when the policyholder is severely sick, we define the care-dependent tontine payoff as:

$$b_2^{oc_1}(t) := \begin{cases} \mathbb{1}_{\{S_{x+t}=2\}} \frac{nd_2^{oc_1}(t)}{N(t)}, & \text{if } N(t) > 0, \\ 0, & \text{else.} \end{cases} \quad (3.2)$$

As individuals usually have a higher demand for liquidity when becoming severely sick, it might be more reasonable to expect  $d_2^{oc_1}(t) > d_1^{oc_1}(t)$ . In later sections, a risk-averse policyholder will choose  $d_1^{oc_1}(t)$  and  $d_2^{oc_1}(t)$  optimally to maximize his expected lifetime utility (Equation (3.13)). The results are then given in Theorem 3.1. After these optimal payments are derived, we will examine the relation between  $d_1^{oc_1}(t)$  and  $d_2^{oc_1}(t)$ .

Next, we determine the actuarially fair premium of this care-dependent tontine under an actuarial pricing framework. To focus on the effects of mortality and morbidity risks, for simplicity, we ignore the financial risks.<sup>5</sup> Besides, the policyholder under consideration is assumed to pay a single premium at the beginning of the contract. Thus, we assume that the premium earns a constant and continuously compounded risk-free rate  $r \in R$ . This assumption of the single premium is made not only for convenience of computation, but also for the fact that single premium has been widely used for retirement products. In 2018, sales of single-premium immediate annuity in the United States amounted to 9.7 billion dollars.<sup>6</sup> Meanwhile, insurers including OneAmerica, Global Atlantic Financial Group, and etc., have already provided the single-premium LTC plus annuity products in the market.<sup>7</sup> The actuarially fair premium for the considered one-pool care-dependent tontine is given by

$$\begin{aligned} P_0^{oc_1} &:= \mathbb{E} \left[ \int_0^\infty e^{-rt} b_1^{oc_1}(t) dt + \int_0^\infty e^{-rt} b_2^{oc_1}(t) dt \right] \\ &= \mathbb{E} \left[ \int_0^\infty e^{-rt} \mathbb{1}_{\{S_{x+t}=1\}} \frac{nd_1^{oc_1}(t)}{N(t)} dt \right] + \mathbb{E} \left[ \int_0^\infty e^{-rt} \mathbb{1}_{\{S_{x+t}=2\}} \frac{nd_2^{oc_1}(t)}{N(t)} dt \right] \\ &= \int_0^\infty e^{-rt} \cdot {}^{11}p_x \sum_{k=0}^{n-1} \frac{nd_1^{oc_1}(t)}{k+1} \binom{n-1}{k} ({}_t p_x)^k (1 - {}_t p_x)^{n-1-k} dt \\ &\quad + \int_0^\infty e^{-rt} \cdot {}^{12}p_x \sum_{k=0}^{n-1} \frac{nd_2^{oc_1}(t)}{k+1} \binom{n-1}{k} ({}_t p_x)^k (1 - {}_t p_x)^{n-1-k} dt \\ &= \int_0^\infty e^{-rt} \frac{{}^{11}p_x}{{}_t p_x} \sum_{k=1}^n \binom{n}{k} ({}_t p_x)^k (1 - {}_t p_x)^{n-k} \cdot d_1^{oc_1}(t) dt + \int_0^\infty e^{-rt} \frac{{}^{12}p_x}{{}_t p_x} \sum_{k=1}^n \binom{n}{k} ({}_t p_x)^k (1 - {}_t p_x)^{n-k} \cdot d_2^{oc_1}(t) dt \\ &= \int_0^\infty e^{-rt} \frac{{}^{11}p_x}{{}_t p_x} \left( 1 - (1 - {}_t p_x)^n \right) \cdot d_1^{oc_1}(t) dt + \int_0^\infty e^{-rt} \frac{{}^{12}p_x}{{}_t p_x} \left( 1 - (1 - {}_t p_x)^n \right) \cdot d_2^{oc_1}(t) dt. \end{aligned} \quad (3.3)$$

For lines three and four, in the case that the policyholder is still alive,  $N(t)$  is at least 1. Thus, the overall pool size satisfies  $N(t) - 1 \sim \text{Bin}(n-1, {}_t p_x)$ .

<sup>5</sup> When defining the predetermined benefits for the care-dependent tontine products, we neglect the financial market risk. In other words, we do not consider the interest rate risk, nor are the benefits unit-linked. Incorporating equity risk can make the benefits more volatile. If we compare pure mortality-linked products with unit-linked products, it is to expect that the unit-linked products probably are more preferable. However, it is unclear what the preference order will turn out if we compare a unit-linked care-dependent annuity to a unit-linked care-dependent tontine. We leave this question for future research.

<sup>6</sup> LIMRA Secure Retirement Institute: Total Annuity Sales Have Best Quarter in Nearly 10 Years. <https://www.limra.com/en/newsroom/news-releases/2019/limra-secure-retirement-institute-total-annuity-sales-have-best-quarter-in-nearly-10-years>.

<sup>7</sup> <https://www.ltcinsuranceconsultants.com/long-term-care-annuity>.



The policyholder's utility is assumed to be given by the following state-dependent utility function

$$U(b_i(t)) := \sum_{i \in \mathcal{S}} u_i(b_i(t)), \quad (3.4)$$

where  $b_i$  is the payoff of care-dependent tontine in state  $i$ ,  $i \in \mathcal{S}$ . And  $u_i(b_i(t))$ ,  $i \in \mathcal{S}$  is strictly increasing and concave functions in  $b_i$ .

It is usually hard to determine a person's utility function exactly. Thus, a possible alternative is to establish one's utility function according to some plausible properties (e.g. people are non-satiated, and typically show decreasing absolute risk aversion in wealth). The power utility satisfies the above two properties and is abundantly used in theoretical research regarding the long-term care insurance (Brown and Finkelstein, 2008; Davidoff, 2010; Ameriks et al., 2020; Chen et al., 2021, etc.) because of its nice analytical tractability. After setting up the model for utility, one can fit it with real-world data to estimate the relevant parameters (e.g. risk aversion coefficient). For example, Friedman (1974) has estimated the policyholder's relative risk aversion coefficient w.r.t. power utility applying the data from health insurance. Szpiro (1986) uses the property and liability insurance data of the U.S. to estimate the relative risk aversion coefficient. Therefore, we assume that the policyholder evaluates payoffs through a power utility with a constant relative risk aversion coefficient  $\gamma \in (0, \infty)$ ,  $\gamma \neq 1$ . Furthermore, considering the fact that different values of  $\gamma$  may affect our results, a sensitivity analysis is given w.r.t.  $\gamma$  later on. Then, for a healthy individual, his utility is

$$u_1(\omega) = \frac{\omega^{1-\gamma}}{1-\gamma}, \quad (3.5)$$

where  $\omega$  is the payoff for healthy participants.

We introduce a payment weighting factor  $\alpha$ , with

$$\alpha = \begin{cases} \in (0, 1], & \text{if } \gamma > 1, \\ > 1, & \text{if } \gamma \in (0, 1) \end{cases}, \quad (3.6)$$

and the utility function for the severely sick is defined as

$$u_2(\omega) = \frac{(\alpha \cdot \omega)^{1-\gamma}}{1-\gamma}. \quad (3.7)$$

We have known from the existing literature that the severe sickness may increase the marginal utility of payments (e.g. Evans and Viscusi (1991)). Moreover, for the purpose of designing attractive products, we would like to compensate the severely sick people with higher payments. In this sense, the marginal utility in the severely sick state is assumed to be not less than that in the healthy state. Note that the marginal utility in the healthy state and the severely sick state are  $\frac{\partial u_1(\omega)}{\partial \omega} = \omega^{-\gamma}$  and  $\frac{\partial u_2(\omega)}{\partial \omega} = \alpha^{1-\gamma} \omega^{-\gamma}$  respectively. In order to achieve this assumption,  $\alpha^{1-\gamma} \geq 1$  needs to be met. Thus, different values of  $\alpha$  are needed, due to the negative and positive sign of  $(1-\gamma)$  for  $\gamma > 1$  and  $\gamma \in (0, 1)$  respectively. For  $\alpha = 1$ , the utility function of the severely sick is identical to that of the healthy.

Here, for this one-pool case, the objective function can be written as

$$\sup_{d_1^{oc_1}(t), d_2^{oc_1}(t)} \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( u_1 \left( \frac{nd_1^{oc_1}(t)}{N(t)} \right) \mathbb{1}_{\{S_{x+t}=1\}} + u_2 \left( \frac{nd_2^{oc_1}(t)}{N(t)} \right) \mathbb{1}_{\{S_{x+t}=2\}} \right) dt \right] \quad \text{s.t. } P_0^{oc_1} \leq v, \quad (3.8)$$

where  $\rho$  is the individual's subjective discount rate and  $v$  the initial wealth the policyholder owns to invest in the care-dependent tontine. In Theorem 3.1, we solve the optimization problem (3.8).

**Theorem 3.1.** Assume that the policyholder's preferences for payments can be described by (3.8). For a one-pool care-dependent tontine product, maximizing the expected state-dependent lifetime utility (3.8), the optimal payment stream functions are given by

$$d_1^{oc_1*}(t) = \frac{e^{1/\gamma(r-\rho)t}}{(\lambda_{oc_1}^*)^{1/\gamma}} \cdot \frac{(\kappa_{n,\gamma}(t p_x))^{1/\gamma}}{(1 - (1 - t p_x)^n)^{1/\gamma}}, \quad (3.9)$$

$$d_2^{oc_1*}(t) = \alpha^{1/\gamma-1} \cdot \frac{e^{1/\gamma(r-\rho)t}}{(\lambda_{oc_1}^*)^{1/\gamma}} \cdot \frac{(\kappa_{n,\gamma}(t p_x))^{1/\gamma}}{(1 - (1 - t p_x)^n)^{1/\gamma}}, \quad (3.10)$$

$$\Rightarrow d_2^{oc_1*} = \alpha^{1/\gamma-1} \cdot d_1^{oc_1*}, \quad (3.11)$$

where  $\kappa_{n,\gamma}(t p_x) = \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma (t p_x)^k (1 - t p_x)^{n-k}$ .

Furthermore, the optimal Lagrangian multiplier  $\lambda_{oc_1}^* > 0$  is given by

$$\begin{aligned} \lambda_{oc_1}^* = & \left( \frac{1}{v} \left( \int_0^\infty e^{\left(\frac{r-\rho}{\gamma}-r\right)t} \cdot \frac{\left( \frac{11}{t} p_x \cdot \kappa_{n,\gamma}(t p_x) \right)^{1/\gamma}}{\left( \frac{11}{t} p_x \cdot (1 - (1 - t p_x)^n) \right)^{1/\gamma-1}} dt \right. \right. \\ & \left. \left. + \int_0^\infty e^{\left(\frac{r-\rho}{\gamma}-r\right)t} \cdot \alpha^{1/\gamma-1} \cdot \frac{\left( \frac{12}{t} p_x \cdot \kappa_{n,\gamma}(t p_x) \right)^{1/\gamma}}{\left( \frac{12}{t} p_x \cdot (1 - (1 - t p_x)^n) \right)^{1/\gamma-1}} dt \right) \right)^\gamma. \end{aligned} \quad (3.12)$$

And then, the expected state-dependent lifetime utility is given by

$$U_0^{oc_1*} = \int_0^\infty e^{-\rho t} \frac{{}^{11}p_x}{t p_x} \sum_{k=1}^n \binom{n}{k} \left(\frac{k}{n}\right)^\gamma ({}_t p_x)^k (1 - {}_t p_x)^{n-k} \cdot u_1(d_1^{oc_1*}(t)) dt \\ + \int_0^\infty e^{-\rho t} \frac{{}^{12}p_x}{t p_x} \sum_{k=1}^n \binom{n}{k} \left(\frac{k}{n}\right)^\gamma ({}_t p_x)^k \cdot (1 - {}_t p_x)^{n-k} \cdot u_2(d_2^{oc_1*}(t)) dt. \quad (3.13)$$

As can be seen in (3.11),  $\alpha$  and  $\gamma$  play crucial roles in the relation between  $d_2^{oc_1*}$  and  $d_1^{oc_1*}$ . For  $\alpha = 1$ , we return to the original tontine setting of Chen et al. (2019) and  $d_2^{oc_1*} = d_1^{oc_1*}$ ; our optimal tontine payments take the same form, with the exclusive difference that the critical Lagrangian multiplier  $\lambda_{oc_1}^*$  depends on the various transition probabilities.

As shown by Equation (3.11), the optimal care-dependent tontine payment  $d_2^{oc_1*}$  for the severely sick is  $\alpha^{1/\gamma-1}$  times of  $d_1^{oc_1*}$  for the healthy. According to the different value arranges of  $\alpha$  when  $\gamma > 1$  and  $\gamma \in (0, 1)$  (see (3.6)), it always holds  $\alpha^{1/\gamma-1} > 1$ . It means, the sick members will receive more payments, which coincides with our aim – to provide higher liquidity to cover the medical care cost.

### 3.2. Two-pool care-dependent tontine

In another case, we distinguish two groups of policyholders at each time  $t > 0$ , i.e. the pool of healthy members, and the pool of severely sick members. At each time  $t$ , we reallocate the ones who become severely sick into the corresponding pool.

Then the payoff for the healthy members is defined as:

$$b_1^{oc_2}(t) := \begin{cases} \mathbb{1}_{\{S_{x+t}=1\}} \frac{nd_1^{oc_2}(t)}{N_1(t)}, & \text{if } N_1(t) > 0, \\ 0, & \text{else} \end{cases} \quad (3.14)$$

And the payoff for the severely sick members is defined as:

$$b_2^{oc_2}(t) := \begin{cases} \mathbb{1}_{\{S_{x+t}=2\}} \frac{nd_2^{oc_2}(t)}{N_2(t)}, & \text{if } N_2(t) > 0, \\ 0, & \text{else} \end{cases} \quad (3.15)$$

Here  $N_1(t)$  represents the number of healthy policyholders, and  $N_2(t)$  is that in the pool of the severely sick.

For a two-pool care-dependent tontine, the actuarially fair premium is given by:

$$P_0^{oc_2} := \mathbb{E} \left[ \int_0^\infty e^{-rt} b_1^{oc_2}(t) dt + \int_0^\infty e^{-rt} b_2^{oc_2}(t) dt \right] \\ = \mathbb{E} \left[ \int_0^\infty e^{-rt} \mathbb{1}_{\{S_{x+t}=1\}} \frac{nd_1^{oc_2}(t)}{N_1(t)} dt \right] + \mathbb{E} \left[ \int_0^\infty e^{-rt} \mathbb{1}_{\{S_{x+t}=2\}} \frac{nd_2^{oc_2}(t)}{N_2(t)} dt \right] \\ = \int_0^\infty e^{-rt} {}^{11}p_x \sum_{k=0}^{n-1} \frac{nd_1^{oc_2}(t)}{k+1} \binom{n-1}{k} ({}_t p_x)^k (1 - {}_t p_x)^{n-1-k} dt \\ + \int_0^\infty e^{-rt} {}^{12}p_x \sum_{k=0}^{n-1} \frac{nd_2^{oc_2}(t)}{k+1} \binom{n-1}{k} ({}_t p_x)^k (1 - {}_t p_x)^{n-1-k} dt \\ = \int_0^\infty e^{-rt} \sum_{k=1}^n \binom{n}{k} ({}_t p_x)^k (1 - {}_t p_x)^{n-k} \cdot d_1^{oc_2}(t) dt + \int_0^\infty e^{-rt} \sum_{k=1}^n \binom{n}{k} ({}_t p_x)^k (1 - {}_t p_x)^{n-k} \cdot d_2^{oc_2}(t) dt \\ = \int_0^\infty e^{-rt} \left( 1 - (1 - {}^{11}p_x)^n \right) \cdot d_1^{oc_2}(t) dt + \int_0^\infty e^{-rt} \left( 1 - (1 - {}^{12}p_x)^n \right) \cdot d_2^{oc_2}(t) dt. \quad (3.16)$$

In the third line, conditional on  $S_{x+t} = 1$ , there is at least one alive member in the healthy pool. Then the healthy pool size is distributed as  $N_1(t) - 1 \sim \text{Bin}(n-1, {}^{11}p_x)$ . In addition, in line four, conditional on  $S_{x+t} = 2$ , the number in the severely sick pool is not less than 1. Then  $N_2(t) - 1 \sim \text{Bin}(n-1, {}^{12}p_x)$ .

Here again we write down the objective function:

$$\sup_{d_1^{oc_2}(t), d_2^{oc_2}(t)} \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( u_1 \left( \frac{nd_1^{oc_2}(t)}{N_1(t)} \right) \mathbb{1}_{\{S_{x+t}=1\}} + u_2 \left( \frac{nd_2^{oc_2}(t)}{N_2(t)} \right) \mathbb{1}_{\{S_{x+t}=2\}} \right) dt \right] \quad \text{s.t. } P_0^{oc_2} \leq v. \quad (3.17)$$

In Theorem 3.2, we solve the objective function (3.17).

**Theorem 3.2.** Assume that the policyholder's preferences for payments can be described by (3.17). For a two-pool care-dependent tontine product, maximizing the expected state-dependent lifetime utility (3.17), the optimal payment stream functions are given by

$$d_1^{oc_2^*}(t) = \frac{e^{1/\gamma(r-\rho)t}}{(\lambda_{oc_2}^*)^{1/\gamma}} \cdot \frac{(\kappa_{n,\gamma}({}^{11}_t p_x))^{1/\gamma}}{\left(1 - (1 - {}^{11}_t p_x)^n\right)^{1/\gamma}}, \quad (3.18)$$

$$d_2^{oc_2^*}(t) = \alpha^{1/\gamma-1} \cdot \frac{e^{1/\gamma(r-\rho)t}}{(\lambda_{oc_2}^*)^{1/\gamma}} \cdot \frac{(\kappa_{n,\gamma}({}^{12}_t p_x))^{1/\gamma}}{\left(1 - (1 - {}^{12}_t p_x)^n\right)^{1/\gamma}}. \quad (3.19)$$

The optimal Lagrangian multiplier  $\lambda_{oc_2}^* > 0$  is given by

$$\lambda_{oc_2}^* = \left( \frac{1}{v} \left( \int_0^\infty e^{\left(\frac{r-\rho}{\gamma}-r\right)t} \left[ \frac{(\kappa_{n,\gamma}({}^{11}_t p_x))^{1/\gamma}}{\left(1 - (1 - {}^{11}_t p_x)^n\right)^{1/\gamma-1}} + \frac{\alpha^{1/\gamma-1} \cdot (\kappa_{n,\gamma}({}^{12}_t p_x))^{1/\gamma}}{\left(1 - (1 - {}^{12}_t p_x)^n\right)^{1/\gamma-1}} \right] dt \right) \right)^\gamma. \quad (3.20)$$

And then, the expected state-dependent lifetime utility is given by

$$\begin{aligned} U_0^{oc_2} &= \int_0^\infty e^{-\rho t} \sum_{k=1}^n \binom{n}{k} \left(\frac{k}{n}\right)^\gamma ({}^{11}_t p_x)^k (1 - {}^{11}_t p_x)^{n-k} \cdot u_1(d_1^{oc_2^*}(t)) dt \\ &\quad + \int_0^\infty e^{-\rho t} \sum_{k=1}^n \binom{n}{k} \left(\frac{k}{n}\right)^\gamma ({}^{12}_t p_x)^k (1 - {}^{12}_t p_x)^{n-k} \cdot u_2(d_2^{oc_2^*}(t)) dt. \end{aligned} \quad (3.21)$$

In comparison with the work of Hieber and Lucas (2020), which mainly focuses on the insurer's perspective, we start from the policyholder's angle. Our first design of care-dependent tontine coincides with the 3-state case of Hieber and Lucas (2020), i.e. we both pool together two types of people (the healthy and the severely sick). The difference is how we generate the optimal payoffs. Under the setting of Hieber and Lucas (2020), the authors compensate the sick members' extra mortality credits to themselves owing to higher mortality, thus the payoff for the severely sick gets higher than that for the healthy; additionally, the payoff for the severely sick will vary with the extra mortality credits. While in our case, a rational policyholder will choose the optimal payoffs that maximize his utility, under the budget constraint. In this sense, pricing is not the main purpose of this paper. The payoff for the severely sick is determined by the policyholder's payment weighting factor and the overall utility. It gets higher than the payoff for the healthy due to the payment weighting factor.

### 3.3. Care-dependent annuity

In order to compare different care-dependent products, we write down the optimal payoff functions of a care-dependent annuity using our model framework. For a policyholder that is healthy, he will receive a regular annuity payoff of  $c_1^{ac}(t) > 0$ . Once he becomes care-dependent, a different payoff  $c_2^{ac}(t) > 0$  will be provided. It is also expected that  $c_2^{ac}(t) > c_1^{ac}(t)$  for a higher liquidity is desired when the policyholder gets severely sick. This will be examined later.

Before going towards the optimal payoffs, let us first write down the actuarially fair premium for the care-dependent annuity:

$$P_0^{ac} := \mathbb{E} \left[ \int_0^\infty e^{-rt} \mathbb{1}_{\{S_{x+t}=1\}} c_1^{ac}(t) dt \right] + \mathbb{E} \left[ \int_0^\infty e^{-rt} \mathbb{1}_{\{S_{x+t}=2\}} c_2^{ac}(t) dt \right] = \int_0^\infty e^{-rt} {}^{11}_t p_x c_1^{ac}(t) dt + \int_0^\infty e^{-rt} {}^{12}_t p_x c_2^{ac}(t) dt. \quad (3.22)$$

Analogously, the policyholder's utility is given by

$$\sup_{c_1^{ac}(t), c_2^{ac}(t)} \mathbb{E} \left[ \int_0^\infty e^{-\rho t} (u_1(c_1^{ac}(t)) \mathbb{1}_{\{S_{x+t}=1\}} + u_2(c_2^{ac}(t)) \mathbb{1}_{\{S_{x+t}=2\}}) dt \right] \quad \text{s.t. } P_0^{ac} \leq v. \quad (3.23)$$

Then we maximize the utility optimization problem (3.23) by Theorem 3.3.

**Theorem 3.3.** For a care-dependent annuity product, the optimal payoffs are given by

$$c_1^{ac^*}(t) = \left( \lambda_{ac}^* \cdot e^{(\rho-r)t} \right)^{-1/\gamma} \quad (3.24)$$

$$c_2^{ac^*}(t) = \alpha^{1/\gamma-1} \cdot \left( \lambda_{ac}^* \cdot e^{(\rho-r)t} \right)^{-1/\gamma}. \quad (3.25)$$

By budget constraint, we obtain

$$\lambda_{ac}^* = \left( \frac{1}{v} \int_0^\infty e^{\left(\frac{r-\rho}{\gamma}-r\right)t} ({}^{11}_t p_x + \alpha^{1/\gamma-1} \cdot {}^{12}_t p_x) dt \right)^\gamma. \quad (3.26)$$

The policyholder's expected state-dependent lifetime utility is given by

$$U_0^{ac*} = \int_0^\infty e^{-\rho t} {}^{11}p_x \cdot u_1(c_1^{ac*}(t)) dt + \int_0^\infty e^{-\rho t} {}^{12}p_x \cdot u_2(c_2^{ac*}(t)) dt. \quad (3.27)$$

Similarly as in one-pool care-dependent tontine case,  $\alpha = 1$  leads to a regular annuity, which does not lead to an increase in payoff when becoming care-dependent, i.e.  $c_1^{ac*}(t) = c_2^{ac*}(t)$ .

In case that actuarially fair premiums are adopted for care-dependent tontines and care-dependent annuities as in Equations (3.3), (3.16) and (3.22), the relation between the care-dependent annuities and care-dependent tontines can be explored theoretically.

**Proposition 3.4.** By results reached from problems (3.8), (3.17), and (3.23), we have

$$U_0^{ac*} \geq U_0^{oc1*}, \quad (3.28)$$

$$U_0^{ac*} \geq U_0^{oc2*}. \quad (3.29)$$

**Proof.** See Appendix A.4.  $\square$

Adding care-dependency payoffs to the regular retirement products does not change the preference order of the tontine and the annuity under actuarially fair pricing. The optimal care-dependent annuities always deliver a higher expected lifetime utility level than the optimal care-dependent tontines. However, between the two care-dependent tontines, no clear relation can be detected.

#### 4. Product comparison under a realistic setting

So far, we have obtained the optimal payoffs as well as the corresponding utility of each product with actuarially fair premiums on a net basis. In the following, we take the realistic risk loadings into account, in order to explore how the attractiveness of the care-dependent tontines and care-dependent annuities would change in the real world.<sup>8</sup> We compute the gross premium which consists of the net premium and the risk loading (Section 4.1), and the expected life-time utility of each product. A product leading to a higher utility level (but costing more) is not necessarily better than a product leading to a lower utility level (but costing less). In Section 4.2, we will show how to make the considered care-dependent products reach the same expected utility level, such that we can focus on the comparison of the gross premiums of the products.

##### 4.1. Calculation of risk loadings

By ignoring any administration or acquisition charges, we assume a single charged risk loading, i.e.  $F_0^{oc1} \geq 0$  for the 1-pool care-dependent tontine,  $F_0^{oc2} \geq 0$  for the 2-pool care-dependent tontine and  $F_0^{ac} \geq 0$  for care-dependent annuity. Then we can specify the initial gross premium for each product by

$$\hat{P}_0^{oc1} = P_0^{oc1} + F_0^{oc1}, \quad (4.1)$$

$$\hat{P}_0^{oc2} = P_0^{oc2} + F_0^{oc2}, \quad (4.2)$$

$$\hat{P}_0^{ac} = P_0^{ac} + F_0^{ac}. \quad (4.3)$$

Now, in order to determine the risk loading, we further consider the longevity risk and disease risk<sup>9</sup> following China Risk Oriented Solvency System (C-ROSS)<sup>10</sup> regulation about minimum capital requirement in China (see CBIRC (2020b)).

According to the Insurance Company Solvency Supervision Rules No.5, the minimum capital requirement in retirement insurance section is computed by a scenario-comparison method (CBIRC, 2020b). It is worthwhile to mention that the base scenario assumptions are used by the insurers to calculate their best-estimate liabilities. The capital requirement is the change of the present value (PV) between the unfavorable scenario and the base scenario, and it should not be negative, i.e.

$$MC = \max(PV_{unf} - PV_{bas}, 0), \quad (4.4)$$

<sup>8</sup> With reference to Chen et al. (2019), one could notice that when adding a reasonable risk loading for longevity risk, tontines become more attractive than annuities. However, it is not the case without a consideration about risk loading.

<sup>9</sup> As for the disease risk, in fact, it contains not only the disease incidence risk, but also the disease deterioration risk; these two types of disease risks are correlated (CBIRC, 2020b). Nevertheless, dealing with this may distract from our main points that we want to convey in this part. As our focus in this section is to illustrate the possible choice between two care-dependent tontines and the care-dependent annuity in view of a gross premium. Under actuarially fair premium setting, results of Proposition 3.4 have shown the care-dependent annuity is always more attractive than either of the care-dependent tontine. Thus, we ignore the disease deterioration risk faced by the severely sick members in the following contents. Disease incidence risk refers to the risk of insurance companies suffering unexpected losses due to the actual experience of the disease incidence being higher than expected. Disease deterioration risk means the risk that the disease deterioration trend is higher than expected and finally causes non-expected losses to the insurers.

<sup>10</sup> On 13 February 2015, the China Insurance Regulatory Commission (now called China Banking and Insurance Regulatory Commission (CBIRC)) released the rules of the new solvency regime, which is known as China Risk Oriented Solvency System (C-ROSS). It adopts a regulatory framework of 'three pillars', reshapes it according to the characteristics of China's insurance market to ensure that it is viable and reflects the realities of the emerging market, and makes sure it is comparable with other representative solvency regimes in the world (e.g. Solvency II in Europe and Risk-based Capital Phase 2 in Singapore) in terms of its "three-pillar" structure and specific regulatory standards and requirements.

where  $MC$  is the minimum capital requirement for overall insurance risk in retirement insurance business. The reinsurance factor is left out here, and  $PV_{bas}$  represents the PV under base scenario assumptions at time 0, and  $PV_{unf}$  is that under the unfavorable scenario at time 0. As there are actually two types of risks here, we will need to calculate the minimum capital under each kind of risks separately. We denote the  $MC_{long}$  as the minimum capital requirement for longevity risk and  $MC_{morb}$  for disease incidence risk. Then, they can be computed by

$$MC_{long} = \max(PV_{long} - PV_{bas}, 0), \quad (4.5)$$

$$MC_{morb} = \max(PV_{morb} - PV_{bas}, 0), \quad (4.6)$$

with  $PV_{long}$  and  $PV_{morb}$  representing the PV in unfavorable situation of longevity and disease incidence risk, respectively. The longevity risk and disease incidence risk are uncorrelated according to CBIRC (2020b). Thus we get the overall required minimum insurance capital by

$$MC_{ins} = \sqrt{MC_{long}^2 + MC_{morb}^2}. \quad (4.7)$$

Referring to CBIRC (2020b), the assumptions for one-year probability under the unfavorable scenario are defined to be the assumptions for one-year probability under the base scenario multiplied by certain shock factors. The unfavorable scenario assumption is equal to the base scenario assumption  $\times (1 + SF)$ , where  $SF$  is the unfavorable scenario factor, the proportional shift upward or downward of the underlying assumption (e.g. the best-estimate survival probabilities).

In the aspect of longevity risk,  $SF$  of longevity risk is based on the proportional shift downward of the base mortality assumption at all future policy dates. We denote the unfavorable scenario factor of longevity risk by  $SF_t^{long}$ , and it takes the values according to the policy duration as follows (CBIRC, 2020b):

$$SF_t^{long} = \begin{cases} (1 - 3\%)^t - 1 & 0 < t \leq 10 \\ (1 - 3\%)^{10} \cdot (1 - 2\%)^{t-10} - 1 & 10 < t \leq 20 \\ (1 - 3\%)^{10} \cdot (1 - 2\%)^{10} \cdot (1 - 1\%)^{t-20} - 1 & 20 < t \leq 30 \\ (1 - 3\%)^{10} \cdot (1 - 2\%)^{10} \cdot (1 - 1\%)^{10} - 1 & t > 30 \end{cases}. \quad (4.8)$$

In the aspect of disease incidence risk, the unfavorable scenario factor  $SF_t^{morb}$  is based on the proportional shift upward of the base morbidity assumption at all future policy dates. We set it to be 20%, which is stated by CBIRC (2020b), i.e.

$$SF_t^{morb} = 20\%, \quad \forall t. \quad (4.9)$$

In the unfavorable scenario, we denote  $p^{long}$  for probability in consideration with a longevity risk and  $p^{morb}$  is applied to represent the unfavorable probability with the disease incidence risk. Then, for different care-dependent products, the PVs of cash flows for different scenarios at time 0 are given by

$$PV_i^{oc1} = \int_0^\infty e^{-rt} \frac{{}_{11}p_x^i}{t p_x^i} \left(1 - \left(1 - {}_t p_x^i\right)^n\right) \cdot d_1^{oc1*}(t) dt + \int_0^\infty e^{-rt} \frac{{}_{12}p_x^i}{t p_x^i} \left(1 - \left(1 - {}_t p_x^i\right)^n\right) \cdot d_2^{oc1*}(t) dt, \quad (4.10)$$

$$PV_i^{oc2} = \int_0^\infty e^{-rt} \left(1 - \left(1 - {}_{11}p_x^i\right)^n\right) \cdot d_1^{oc2*}(t) dt + \int_0^\infty e^{-rt} \left(1 - \left(1 - {}_{12}p_x^i\right)^n\right) \cdot d_2^{oc2*}(t) dt, \quad (4.11)$$

$$PV_i^{ac} = \int_0^\infty e^{-rt} {}_{11}p_x^i c_1^{ac*}(t) dt + \int_0^\infty e^{-rt} {}_{12}p_x^i c_2^{ac*}(t) dt, \quad (4.12)$$

$$i = bas, \quad long, \quad morb,$$

where  $i = bas$  represents the base scenario assumption,  $i = long$  is the unfavorable scenario of longevity risk, and  $i = morb$  is the unfavorable scenario of disease incidence risk.

According to the Insurance Company Solvency Supervision Rules No.3 (CBIRC, 2020a), the risk margin  $RM$  actually measures the difference of future liabilities in an unfavorable scenario and a base scenario with a regulation proportion, which reflects as a risk compensation to the insurers or re-insurers. Thus, we take it as an alternative to risk loading for longevity and disease incidence risks (Chen et al., 2019; Bauer et al., 2010). Then, according to CBIRC (2020a), the risk margin  $RM$  is defined by

$$RM = MC_{ins} \cdot \frac{G^{-1}(85\%)}{G^{-1}(99.5\%)}, \quad (4.13)$$

where  $MC_{ins}$  is the overall required minimum capital defined by (4.7), and  $G^{-1}(x\%)$  represents the quantile of a normal distribution under probability  $x\%$ , and  $x = 85$  is prescribed by the regulator (CBIRC, 2020a).  $G(\cdot)$  is the distribution of best-estimate liabilities.<sup>11</sup> Furthermore, the initial single loading for longevity and disease incidence risks is then set to be  $F_0 = RM$ .

<sup>11</sup> The definition of  $G(\cdot)$  is not given clearly in CBIRC (2020a) as it is still being tested by the insurance industry. Here we take it as the distribution of best-estimate liabilities with reference to the quantile method for risk margin computation (Zheng et al., 2013). In addition, as determining the distribution of best-estimate liabilities is not our focus in this article, dealing with the additional issues might cause too much distraction from the key points we try to convey. Thus, we apply the standard normal distribution, and get  $\frac{G^{-1}(85\%)}{G^{-1}(99.5\%)} \approx 0.403$ . We also use different normal distributions and get corresponding coefficients, but it turns out that different coefficients  $\frac{G^{-1}(85\%)}{G^{-1}(99.5\%)}$  have little impact on our conclusions.

In this paper, the risk loading in our setting is purely driven by solvency capital requirement. In the real world, typically a higher risk loading than required by the solvency regulation is charged.

#### 4.2. Product comparison with utility indifference number

Taking account of the risk loadings for the various care-dependent products, we cannot purely compare the expected discounted lifetime utility resulting from the corresponding optimal payoffs. A product leading to a higher utility level might cost in total more initially, i.e. the gross premium of this product might be higher. To further compare the attractiveness of two care-dependent tontines and the care-dependent annuity and ease our comparison, we first make one dimension of the two quantities (utility and gross premium) identical. More specifically, we compute the number of care-dependent tontines that one should purchase, in order to obtain the same utility yielded by one care-dependent annuity.  $Q_1$  and  $Q_2$  are used to denote the number of one-pool care-dependent tontines and two-pool ones respectively (as has been applied by Chen et al. (2019)). Then we have:

$$U_0^{ac} = \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( u_1 \left( Q_1 \frac{nd_1^{oc_1}(t)}{N(t)} \right) \mathbb{1}_{\{S_{x+t}=1\}} + u_2 \left( Q_1 \frac{nd_2^{oc_1}(t)}{N(t)} \right) \mathbb{1}_{\{S_{x+t}=2\}} \right) dt \right] = Q_1^{1-\gamma} \cdot U_0^{oc_1}, \quad (4.14)$$

$$U_0^{ac} = \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( u_1 \left( Q_2 \frac{nd_1^{oc_2}(t)}{N_1(t)} \right) \mathbb{1}_{\{S_{x+t}=1\}} + u_2 \left( Q_2 \frac{nd_2^{oc_2}(t)}{N_2(t)} \right) \mathbb{1}_{\{S_{x+t}=2\}} \right) dt \right] = Q_2^{1-\gamma} \cdot U_0^{oc_2}. \quad (4.15)$$

The  $Q_1$  and  $Q_2$  are solved by:

$$Q_1 = \left( \frac{U_0^{ac}}{U_0^{oc_1}} \right)^{\frac{1}{1-\gamma}}, \quad (4.16)$$

$$Q_2 = \left( \frac{U_0^{ac}}{U_0^{oc_2}} \right)^{\frac{1}{1-\gamma}}. \quad (4.17)$$

$Q_1$  and  $Q_2$  respectively represent the number of 1-pool CDT and 2-pool CDT with actuarially fair premium  $P_0^{oc_1} = P_0^{oc_2} = v = 10000$  that one needs to purchase, in order to receive the same expected discounted lifetime utility as from a CDA with actuarially fair premium  $P_0^{ac} = v = 10000$ . In other words, the policyholder becomes indifferent between one CDA and  $Q_1$  1-pool CDTs (or  $Q_2$  2-pool CDTs). Among the care-dependent products with identical expected discounted lifetime utility, the one with lowest comparable gross premium would be most attractive to the policyholder. Hence, we next will compare the gross premium of  $Q_1$  1-pool care-dependent tontines (i.e.  $Q_1 \cdot \hat{P}^{oc_1}$ ),  $Q_2$  2-pool care-dependent tontines (i.e.  $Q_2 \cdot \hat{P}^{oc_2}$ ) with one care-dependent annuity (i.e.  $\hat{P}^{ac}$ ).

To quantify this part of results, we shall rely on numerical techniques. In this case, we apply the data from China Health and Retirement Longitudinal Study (CHARLS) and rely on China Risk Oriented Solvency System (C-ROSS) to calculate the realistic risk loadings and make product comparison.

### 5. Numerical analysis

In this section, we compare the three different care-dependent products numerically. First, we estimate the transition probabilities. After that, we make the comparison among various care-dependent products from two aspects: (i) we compute the corresponding risk loadings, (ii) and we generate the utility indifference numbers to compare the attractiveness of different products. Finally, sensitivity analyses are given, to examine how the results will change with different inputs (i.e. risk aversion coefficient  $\gamma$ , payment weighting factor  $\alpha$ , and risk-free rate  $r$ ).

#### 5.1. Estimates of transition probabilities

To estimate the transition probabilities, we use the data from China Health and Retirement Longitudinal Study (CHARLS). This database covers the period from 2011 to 2018, and has 4 waves in total. It contains information of Chinese residents ages 45 and older. The baseline national wave of CHARLS is being fielded in 2011 and includes about 10,257 households and 17,708 individuals in 150 counties/districts and 450 villages/resident committees. The survey data includes detailed demographic characteristics (such as age, education level, marital status, etc.), family economic status, health status (self-evaluated health status, chronic disease status, fundamental living ability and cognitive ability, etc.), participation in medical insurance, medical service utilization and community basic information.

The state of health is defined by the number of elderly people who lose their ability to perform daily activities. The ability of daily living (ADL) is originally proposed by Katz et al. (1963) and has been widely used in academia. There are 6 indicators of ADL: eating, dressing, bathing, getting into and out of bed, using the toilet, and controlling urination and defecation. Consistent with the basic situation in Hu et al. (2016)'s research and Chinese LTC insurance practice,<sup>12</sup> we define individuals as severely sick if they have difficulty with three or more (i.e., 3+) ADLs.

To fully utilize the available data, we construct an unbalanced panel dataset with sample weights considered and missing records deleted. Among four waves of the survey 2011, 2013, 2015, 2018, each individual should have at least two consecutive observations.

<sup>12</sup> LTC insurance products in many Chinese insurance companies (e.g. China Life Insurance Company, Kunlun Healthy Insurance Company and etc.), define that one becomes hard doing three or more ADLs as the trigger condition of making LTC payments.

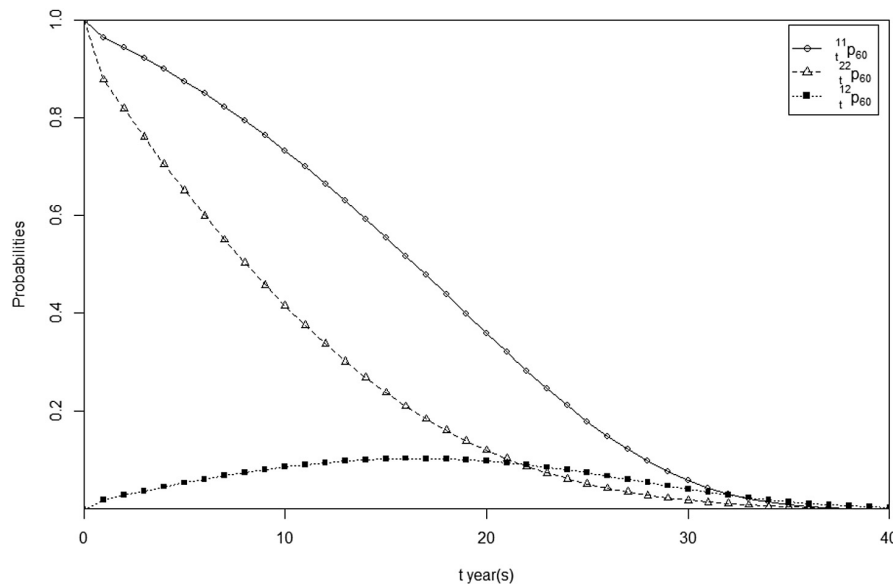


Fig. 5.1.  $t$ -year probabilities of a 60-year-old male ( ${}^{11}_t p_{60}$ ,  ${}^{22}_t p_{60}$  and  ${}^{12}_t p_{60}$ ).

To calculate the transition intensities among different age groups, we first calculate the crude transition intensities with reference to the work of Hanewald et al. (2019). When going towards a specific transition situation, the crude transition intensity for an individual aged  $y$  is given by

$$\widetilde{\mu}_{ij}^y = \frac{{}^{ij}C_y}{{}^iY_y}, \quad i, j \in \mathcal{S}, i \neq j, \quad (5.1)$$

where  ${}^{ij}C_y$  represents the total transition number from state  $i$  to state  $j$  at age  $y$ . And  ${}^iY_y$  describes the overall years of exposure (e.g. sum of years for which the  $y$ -year-old healthy individuals stay healthy). More concretely, the crude transition intensity of healthy to severely sick at age 60, is defined as the total number of transitions from healthy to severely sick at age 60 divided by the total number of years of risk exposure for the 60-year-old at a healthy state. The crude transition intensity directly reflects the number of transitions from the previous state to the next state in unit time. The higher the crude transition intensity, the greater the probability of occurrence of the corresponding transition.

On the basis of resulting crude transition intensities, we then apply a GLM approach (see Appendix A.5) to smooth the trend of transition intensity going with age (Haberman and Renshaw, 2009; Fong and Feng, 2016). By Equations (2.4), (2.5) and (2.7) defined in Section 2, we approximate the  $t$ -year probabilities  ${}^{11}_t p_y$ ,  ${}^{12}_t p_y$  and  ${}^{22}_t p_y$  by age and gender. As gender difference is not our focus in this article, we only demonstrate the results for the 60-year-old males in the following contents. Fig. 5.1 displays the estimated  $t$ -year probabilities for a male aged 60.

The courses of the depicted curves follow the natural expectations: in Fig. 5.1, the occupancy probability of being healthy  ${}^{11}_t p_{60}$  starts from one, goes down with  $t$  and approaches zero for larger  $t$  as the policyholder is going to die at an older age. When looking at the curve with hollow triangles, we notice that the occupancy probability of being severely sick  ${}^{22}_t p_{60}$  descends rapidly in the first two decades and tends to be smooth to zero when he is about to die. Last but not least, the third curve describes the  $t$ -year transition probability of a 60-year-old male, i.e.  ${}^{12}_t p_{60}$ . The transition probability goes up at first then down for longer years. As transition probability is in fact affected by forces from two different directions (see Equation (2.5)): on the one hand, the increasing probability of becoming severely sick with age, tends to pull up the transition probability; while on the other hand, the decreasing occupancy probability of being severely sick, or say, the decreasing survival probability in severely sick state. In the about first two decades, the first type of force dominates; while later, the second type of force plays a leading role in the overall influence to the probability  ${}^{12}_t p_{60}$ .

## 5.2. Care-dependent tontine vs care-dependent annuity

In this subsection, we calculate the risk loadings for different care-dependent products, and use utility indifference numbers to show policyholders' preference for these products in a realistic scenario. But before that, we need to fix some further parameters (see Table 5.1).

With reference to the 5-year average growth rate of CPI in China (Statistics Bureau of the P.R. China, 2020), we fix the risk-free rate  $r$  at 2%. Then we let  $\rho = r$  at first, and the cases where  $\rho > r$  and  $\rho < r$  will be further examined (see e.g. Chen et al. (2021)). As for the risk aversion coefficient, we set  $\gamma = 2$  firstly (see e.g. Havranek et al. (2015)), and we will vary it later for sensitivity analysis.

### 5.2.1. Risk loadings

Referring to our calculation method for risk loadings (see Section 4.1), as well as the baseline parameters (see Table 5.1), we can compute the minimum capital requirements (MCs) and risk loadings ( $F'_0$ s) for different care-dependent products under consideration. The risk loadings for each care-dependent product are shown in Table 5.2. Generally, the risk loadings for care-dependent tontine products



**Table 5.1**  
Additional parameters for further computation.

Variable	Notation	Value
Net premium	$P_0^j = v, j = oc_1, oc_2, ac$	10000
Pool size	$n$	1000
Risk-free rate	$r$	0.02
Subjective discount rate	$\rho$	0.02
Risk aversion coefficient	$\gamma$	2
Initial age	$x$	60
Payment weighting factor	$\alpha$	0.5

**Table 5.2**  
Minimum capitals and risk loadings for various care-dependent products.

$MC_{oc_1}^{long}$	$MC_{oc_2}^{long}$	$MC_{ac}^{long}$	$MC_{oc_1}^{morb}$	$MC_{oc_2}^{morb}$	$MC_{ac}^{morb}$
51.6	0.633	1477	77.7	0	0
$MC_{oc_1}^{ins}$	$MC_{oc_2}^{ins}$	$MC_{ac}^{ins}$	$F_0^{oc_1}$	$F_0^{oc_2}$	$F_0^{ac}$
93.3	0.633	1477	37.589	0.255	595.118

**Table 5.3**  
Risk loading  $F_0$  for different pool sizes  $n$  using the baseline parameter setting from Table 5.1. Net premium  $P_0^{oc_1} = P_0^{oc_2} = P_0^{ac} = v = 10000$ , subjective discount rate  $\rho = 0.02$ , risk-free rate  $r = 0.02$ , initial age  $x = 60$ , risk aversion coefficient  $\gamma = 2$ , and payment weighting factor  $\alpha = 0.5$ .

$n$	$F_0^{oc_1}$	$F_0^{oc_2}$	$F_0^{ac}$
10	144	270	595
100	41	14.9	595
500	37.7	0.978	595
1000	37.6	0.255	595
2000	37.6	0.0648	595
5000	37.6	0.0071	595

are much lower than that of a care-dependent annuity. The care-dependent annuity requires the highest risk loading, which amounts to a relevant share of 5.62% ( $= \frac{595.118}{10000+595.118}$ ) of the initial contribution. It could be explained from two aspects: (i) the insurer completely bears the longevity risk for this annuity-like product; (ii) this big amount of risk loading partly comes from the unfavorable assumption for longevity risk (see Equation (4.8)) provided by CBIRC (2020b). For instance,  $SF_t^{long}$  takes a large value as time goes by,<sup>13</sup> which means it is assumed there is a huge longevity risk in the long run.

We also observe the risk loading of the 1-pool CDT, i.e. 37.589, accounts for 0.37% ( $= \frac{37.589}{10000+37.589}$ ) of the gross premium. This is substantially greater than that of the 2-pool CDT, which is negligible. The difference comes from the first row in the Table 5.2. (i). The 1-pool CDT requires a larger minimum capital for longevity risk ( $MC^{long}$ ). Let us take a closer look at the liability structures of these two CDTs (see Equation (4.10) and (4.11)). With a strike from longevity risk, the changed probabilities will have a greater impact on the 1-pool CDT than the 2-pool one. The term  $(1 - (1 - p)^n)$  tends to 1 as  $n$  becomes larger.<sup>14</sup> Thus, the longevity risk does not substantially affect the two-pool CDT, especially in a large pool case. (ii) The 1-pool CDT also requires a larger minimum capital for disease incidence risk ( $MC^{morb}$ ). The insurer takes over a relatively high disease incidence risk for 1-pool CDT. While for the 2-pool CDT, on one side, the increase of disease incidence risk causes  ${}^{11}_t p_x^{morb}$  to decline over years; on the other side, an increased disease incidence risk brings about a greater  ${}^{12}_t p_x^{morb}$ . In our case, the effect of disease incidence risk gets canceled out and even drives down the future cash flows. By definition of the corresponding minimum capital (4.6), the minimum capital for disease incidence risk ( $MC_{oc_2}^{morb}$ ) of the 2-pool CDT is equal to zero (This explanation also holds for  $MC_{ac}^{morb} = 0$  of CDA).

In addition to risk loading computation w.r.t. the Chinese regulation regime (C-ROSS), we also carry out the relevant computation under a different regulation regime, i.e. the capital requirements of Solvency II, the insurance regulation in EU countries. The risk loadings obtained by the Solvency II do not deviate much from those by C-ROSS. The care-dependent annuity still requires the highest risk loading, while under the baseline parameter setting, the two-pool care-dependent tontine requires the least (see Appendix A.6).

Regarding different pool sizes, we also choose the baseline parameters from Table 5.1 to compute the corresponding risk loadings. Table 5.3 shows that for small pool sizes, say  $n = 10$ , the loadings for the CDTs are relatively large. The advantage of the CDTs over the CDA is less substantial. A larger  $n$  leads to less risk loadings for both one-pool CDT and two-pool CDT, but that of the one-pool CDT stops decreasing when the pool size  $n$  reaches some certain level. We could observe that its risk loading remains the same when  $n \geq 1000$ . This is due to the fact that the term  $(1 - (1 - p)^n)$  tends to 1 as  $n$  becomes larger as mentioned at an early place. There is no other term related to  $n$  in the calculation for the one-pool CDT's risk loading, which means the influence of  $n$  fades away as  $n$  grows. In case of two-pool CDT, for a substantially large pool size, the loading for care-dependent tontines is negligible.

<sup>13</sup> According to Equation (4.8), when  $t > 30$ ,  $SF_t^{long} = (1 - 3\%)^{10} \cdot (1 - 2\%)^{10} \cdot (1 - 1\%)^{10} - 1 = -45.5\%$ .

<sup>14</sup> Here,  $p$  represents the probability, e.g.  ${}^{11}_t p_x^{long}$ ,  ${}^{11}_t p_x^{long}$ .



**Table 5.4**

Sensitivity analysis w.r.t. risk aversion coefficient  $\gamma$  and payment weighting factor  $\alpha$ . Other parameters are taken from the baseline setting table (i.e. Table 5.1). More concretely, net premium  $P_0^{oc1} = P_0^{oc2} = P_0^{ac} = v = 10000$ , pool size  $n = 1000$ , risk-free rate  $r = 0.02$ , subjective discount rate  $\rho = 0.02$ , and initial age  $x = 60$ .

$Q_1 * \hat{P}_0^{oc1}$				$Q_2 * \hat{P}_0^{oc2}$			$\hat{P}_0^{ac}$		
$\alpha \in (0, 1), \gamma > 1$									
$\alpha \backslash \gamma$	2	5	8	2	5	8	2	5	8
0.2	10113.74	10224.02	10269.52	10039.65	10126.93	10211.94	10729.28	10916.33	10969.9
0.5	10046.59	10088.6	10110.18	10030.36	10082.98	10133.68	10595.12	10650.96	10666.18
0.8	10019.41	10038.94	10052.78	10026.6	10066.86	10105.5	10541.05	10555.52	10559.25
$\alpha > 1, \gamma \in (0, 1)$									
$\alpha \backslash \gamma$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
1.5	10266.52	10047.26	10013.38	10006.57	10008.14	10010.84	11070.1	10610.15	10538.91
2	10533.97	10087.92	10021.03	10010.43	10009.6	10011.28	11641.77	10692.76	10554.36
5	10840.26	10265.98	10048.25	10014.86	10015.99	10012.83	12301.26	11064.11	10609.34

### 5.2.2. Product comparison with utility indifference number

As introduced in Section 4.2, we will compute the utility indifference numbers  $Q_1$  and  $Q_2$  for one-pool and two-pool care-dependent tontine separately. Using parameters assumed in Table 5.1, we obtain  $Q_1$ ,  $Q_2$  and comparable gross premiums for different care-dependent products:

$$\begin{aligned}
 Q_1 &= 1.000896 \text{ and } Q_2 = 1.003010, \\
 Q_1 \cdot \hat{P}^{oc1} &= 1.000896 \times (10000 + 37.589) = 10046.583, \\
 Q_2 \cdot \hat{P}^{oc2} &= 1.003010 \times (10000 + 0.255) = 10030.356, \\
 \hat{P}^{ac} &= 10000 + 595.118 = 10595.118.
 \end{aligned}$$

Obviously, from results above, it is easy to find out that the 2-pool CDT becomes the best choice for the policyholder and then is the 1-pool CDT, using the baseline parameters. The CDA's gross is taken off by its high risk loading. To conclude, incorporating an initial risk loading required by the insurance regulator, we find that under our baseline setting, the care-dependent tontines are much more attractive than care-dependent annuities for both cases, and the 2-pool CDT gets more appealing than the 1-pool one.

In the following, we explore the sensitivities of the results towards different inputs, i.e. risk aversion coefficient  $\gamma$ , payment weighting factor  $\alpha$ , and risk-free rate  $r$  respectively. Based on the utility indifference number, we calculate the comparable gross premiums for different care-dependent products, which is seen as the basis for people to choose a product.<sup>15</sup> Generally, as can be seen from Table 5.4 and Table 5.5, it is noted that CDTs are always better choices than CDA.

We first numerically examine the sensitivity towards the policyholder's risk aversion, i.e.  $\gamma$ , and other parameters are taken from the baseline setting table (i.e. Table 5.1). We notice from Theorem 3.1, 3.2 and 3.3, that  $\gamma$  exerts an impact on the policyholder's utility as well as risk loadings through different payoffs. For example, changing  $\gamma$  brings about the changes in the payment streams for the severely sick ( $d_2^{oc1}$ ,  $d_2^{oc2}$ , and  $d_2^{ac}$ ) and the utility indifference number ( $Q_1$  and  $Q_2$ ), thus it affects the comparable gross premiums of care-dependent tontines. In the case that  $\gamma > 1$  and  $\alpha \in (0, 1)$ , the rows of Table 5.4 reveal that the more risk-averse policyholders regard CDTs as a little bit more expensive than those who are less risk-averse. To our surprise, we find that for those who are assigned a bigger risk aversion coefficient, e.g.  $\gamma \geq 8$  when  $\alpha = 0.5$ , the 1-pool CDT becomes the best option while 2-pool CDT comes in second in terms of choices. In the case that  $\gamma \in (0, 1)$  and  $\alpha > 1$ , the comparable gross premium of the 1-pool CDT drops with  $\gamma$ . With different value ranges of  $\gamma$  and  $\alpha$ , the payment streams for the severely sick vary in different direction with  $\gamma$ .<sup>16</sup> The variation in the payment stream leads to changes in the risk loading, and then the comparable gross premium. The risk aversion coefficient  $\gamma$  plays a critical role in the payment functions of these two CDTs, thus a slight change of  $\gamma$  may bring about relatively large difference in comparable gross premiums.

Next, let us fix the risk aversion coefficient and look at the columns of Table 5.4. And also, the other parameters are taken from the baseline setting table (i.e. Table 5.1). Generally speaking, the change of the payment weighting factor does not lead to 2-pool CDT being less attractive. The payment weighting factor  $\alpha$  has an effect on the comparable gross premiums of CDTs through utility indifference numbers and risk loadings. It can be detected that a greater  $\alpha$  slightly decreases the comparable gross premiums of CDTs when  $\gamma > 1$  and  $\alpha \in (0, 1)$ , while increases those in the case that  $\gamma \in (0, 1)$  and  $\alpha > 1$ . Likewise, the difference comes from the fact that the payment streams for the severely sick vary in different directions with  $\alpha$ , when  $\alpha$  lies in different value ranges (see Equation (3.10), (3.19) and (3.25)). In the case that  $\gamma > 1$  and  $\alpha \in (0, 1)$ ,  $\alpha^{1/\gamma-1}$  goes down with  $\alpha$ . Thus, corresponding payments for the severely sick drops with  $\alpha$ .<sup>17</sup> Further, risk loadings reduce as  $\alpha$  increases. In the case that  $\gamma \in (0, 1)$  and  $\alpha > 1$ ,  $\alpha^{1/\gamma-1}$  goes up with  $\alpha$ . Thus, the corresponding payments for the severely sick and then risk loadings increase with  $\alpha$ ; meanwhile, the advantage of the 2-pool CDT becomes more substantial.<sup>18</sup>

<sup>15</sup> For additional reference, the corresponding sensitivities of risk loadings and utility indifference numbers are shown in Table A.2, Table A.3, Table A.4, and Table A.5 respectively.

<sup>16</sup> By Equation (3.10), (3.19) and (3.25), when  $\gamma > 1$  and  $\alpha \in (0, 1)$ , greater  $\gamma$  leads to a rise in the term  $\alpha^{1/\gamma-1}$ , while a decline in  $\alpha^{1/\gamma-1}$  when  $\gamma \in (0, 1)$  and  $\alpha > 1$ . More concretely,  $\frac{\partial \alpha^{1/\gamma-1}}{\partial \gamma} = -\frac{1}{\gamma^2} \cdot \ln \alpha \cdot \alpha^{1/\gamma-1}$ , with  $\alpha^{1/\gamma-1} > 0$  and  $-\frac{1}{\gamma^2} < 0$ . Thus, the monotony depends on the value of  $\alpha$ .

<sup>17</sup> But the payments for the severely sick are still higher than those for the healthy.

<sup>18</sup>  $\frac{\partial \alpha^{1/\gamma-1}}{\partial \alpha} = (\frac{1}{\gamma} - 1)\alpha^{1/\gamma-2}$ , with  $\alpha^{1/\gamma-2} > 0$ . Thus, the monotony depends on the value of  $\gamma$ .

**Table 5.5**

Sensitivity analysis w.r.t. risk-free rate  $r$ . Other parameters are taken from the baseline setting table (i.e. Table 5.1). Net premium  $P_0^{oc1} = P_0^{oc2} = P_0^{ac} = v = 10000$ , pool size  $n = 1000$ , subjective discount rate  $\rho = 0.02$ , initial age  $x = 60$ , risk aversion coefficient  $\gamma = 2$ , and payment weighting factor  $\alpha = 0.5$ .

$r$	$Q_1 * \hat{P}_0^{oc1}$	$Q_2 * \hat{P}_0^{oc2}$	$\hat{P}_0^{ac}$	$Q_1 * \hat{P}_0^{oc1}$	$Q_2 * \hat{P}_0^{oc2}$	$\hat{P}_0^{ac}$
	$\gamma = 2, \alpha = 0.5$			$\gamma = 0.5, \alpha = 2$		
0.00	10050.93	10033.5	10690.53	10076.65	10008.1	10515.29
0.01	10048.7	10031.86	10641.07	10082.07	10008.79	10597.38
0.02	10046.59	10030.36	10595.12	10087.92	10009.6	10692.76
0.03	10044.59	10028.97	10552.47	10094.2	10010.56	10803.27
0.04	10042.7	10027.7	10512.9	10100.92	10011.7	10930.82
0.05	10040.91	10026.53	10476.22	10108.06	10013.04	11077.43

**Table 5.6**

Cost of care-dependent coverage. Other parameters are taken from the baseline setting table (i.e. Table 5.1). Net premium  $P_0^{oc1} = P_0^{oc2} = P_0^{ac} = v = 10000$ , pool size  $n = 1000$ , subjective discount rate  $\rho = 0.02$ , risk-free rate  $r = 0.02$ , initial age  $x = 60$ , risk aversion coefficient  $\gamma = 2$ , and payment weighting factor  $\alpha = 0.5$ .

Risk Loading of 1-pool CDT	37.589
Risk Loading of 2-pool CDT	0.255
Risk Loading of CDA	595.118
Risk Loading of Optimal Tontine	0.011
Risk Loading of Optimal Annuity	518.502
Cost of Care-dependent Coverage (1-pool CDT)	37.578
Cost of Care-dependent Coverage (2-pool CDT)	0.244
Cost of Care-dependent Coverage (CDA)	76.616

Finally, as mentioned before, we are about to analyze the cases that  $r < \rho$  and  $r > \rho$ . With  $\rho = 2\%$  keeping fixed, we vary the risk-free interest rate  $r$  by 0 to 0.05, and comparable gross premiums correspondingly. By Table 5.5, analogously, the comparable gross premiums of CDTs decrease in risk-free interest rate  $r$  when  $\gamma > 1$  and  $\alpha \in (0, 1)$ , and increase with  $r$  when  $\gamma \in (0, 1)$  and  $\alpha > 1$ . However, this effect is moderate.

### 5.3. Further results

Regarding the risk loadings computed above, we can further explore how much it costs to add the care-dependent coverage to the original retirement products. In this case, we split the risk loading of a care-dependent tontine/annuity into the risk loading of optimal tontine/annuity (in the form of optimal tontine/annuity introduced by Chen et al. (2019)) and cost of care-dependent coverage. By setting the weighting factor  $\alpha = 1$  for 1-pool CDT and CDA, while keeping the other parameters the same in Table 5.1, we get the risk loading required by optimal tontine for 0.011, and optimal annuity for 518.502 (see Table 5.6). Subtracting the risk loading of optimal tontine/annuity from that of care-dependent tontine/annuity, the cost for care-dependent coverage of each product is clear. The results indicate that for the insurers, under the baseline parameter setting, constructing a care-dependent tontine in a two-pool structure may be the most cost-efficient way to add the care-dependent coverage among the three insurance products under consideration. Due to the rather high cost of care-dependent coverage, the insurer of the one-pool care-dependent tontines do not only play an administrative role as regular tontines, but serves more as a regular care insurance provider.

## 6. Conclusion and discussion

This article comes up with two different ways to evaluate the care-dependent tontines with a policyholder's view. We first compute the actuarially fair premiums and then based on a utility framework, we determine the optimal payment stream structures of the care-dependent tontines that maximize the policyholder's expected lifetime utility. Further, by taking account of the realistic risk loadings, we are able to make product comparison in a setting closer to the realistic world. Our results reveal that by purchasing a care-dependent tontine, the policyholders benefit from lower risk loadings comparing with the care-dependent annuity proposed by predecessors. The results are robust when following the capital requirement of Solvency II, the insurance regulation in EU countries. In addition, two-pool care-dependent tontines draw attractions to the policyholder with a smaller risk aversion coefficient, while one-pool care-dependent tontines are more appealing to the more risk-averse policyholders. Further, we find that for insurers, under the baseline parameter setting, constructing a care-dependent tontine in a one-pool structure may cost more for adding the care-dependent coverage compared with the two-pool case. The (care) insurance for the old age is an increasingly concerning topic for now and future, our findings enrich the related existing literature and are helpful to improve the penetration of elderly care insurance with appealing care-dependent insurance products.

In this article, in order to get analytical forms of the product structure as well as simplify the analysis, we construct our products by pooling the policyholders of homogeneous cohorts. It will also be interesting to go with pooling heterogeneous cohorts from the policyholder's side. We will leave it for future research.

### Declaration of competing interest

The authors declare that they have no competing interests.

# Appendix A

## A.1. Proofs for care-dependent tontines in one-pool case

To obtain the optimal payout functions  $d_1^{oc_1}(t)$  and  $d_2^{oc_1}(t)$  to the care-dependent tontine, we maximize the expected state-dependent utility (3.8) subject to (3.3). First, we compute the policyholder's expected utility  $U_0^{oc_1}$ . In case of power utility,  $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$ ,  $\gamma > 0$  and  $\gamma \neq 1$ . We find that

$$\begin{aligned} U_0^{oc_1} &= \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( u_1 \left( \frac{nd_1^{oc_1}(t)}{N(t)} \right) \mathbb{1}_{\{S_{x+t}=1\}} + u_2 \left( \frac{nd_2^{oc_1}(t)}{N(t)} \right) \mathbb{1}_{\{S_{x+t}=2\}} \right) dt \right] \\ &= \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( u_1 \left( \frac{nd_1^{oc_1}(t)}{N(t)} \right) \mathbb{1}_{\{S_{x+t}=1\}} \right) dt \right] + \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( u_2 \left( \frac{nd_2^{oc_1}(t)}{N(t)} \right) \mathbb{1}_{\{S_{x+t}=2\}} \right) dt \right] \\ &= \int_0^\infty e^{-\rho t} \mathbb{E} \left[ {}^{11}_t p_x u_1 \left( \frac{nd_1^{oc_1}(t)}{N(t)} \right) \right] dt + \int_0^\infty e^{-\rho t} \mathbb{E} \left[ {}^{12}_t p_x u_2 \left( \frac{nd_2^{oc_1}(t)}{N(t)} \right) \right] dt \\ &= \int_0^\infty e^{-\rho t} \frac{{}^{11}_t p_x}{{}_t p_x} \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma ({}_t p_x)^k (1 - {}_t p_x)^{n-k} \cdot u(d_1^{oc_1}(t)) dt \\ &\quad + \alpha^{1-\gamma} \cdot \int_0^\infty e^{-\rho t} \frac{{}^{12}_t p_x}{{}_t p_x} \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma ({}_t p_x)^k (1 - {}_t p_x)^{n-k} \cdot u(d_2^{oc_1}(t)) dt. \end{aligned} \quad (A.1)$$

Then, we write down the Lagrangian function as follows. For easy reading, we simplify the notations  $d_1^{oc_1}(t)$  and  $d_2^{oc_1}(t)$  as  $d_1(t)$ ,  $d_2(t)$  respectively.

$$\begin{aligned} L(d_1(t), d_2(t), \lambda_{oc_1}) &= \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( u_1 \left( \frac{nd_1(t)}{N(t)} \right) \mathbb{1}_{\{S_{x+t}=1\}} + u_2 \left( \frac{nd_2(t)}{N(t)} \right) \mathbb{1}_{\{S_{x+t}=2\}} \right) dt \right] + \lambda_{oc_1} (v - P_0^{oc_1}) \\ &= \int_0^\infty e^{-\rho t} \frac{{}^{11}_t p_x}{{}_t p_x} \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma ({}_t p_x)^k (1 - {}_t p_x)^{n-k} \cdot u(d_1(t)) dt \\ &\quad + \alpha^{1-\gamma} \int_0^\infty e^{-\rho t} \frac{{}^{12}_t p_x}{{}_t p_x} \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma ({}_t p_x)^k (1 - {}_t p_x)^{n-k} \cdot u(d_2(t)) dt \\ &\quad + \lambda_{oc_1} \left( v - \left( \int_0^\infty e^{-rt} \frac{{}^{11}_t p_x}{{}_t p_x} \left( 1 - (1 - {}_t p_x)^n \right) \cdot d_1(t) dt + \int_0^\infty e^{-rt} \frac{{}^{12}_t p_x}{{}_t p_x} \left( 1 - (1 - {}_t p_x)^n \right) \cdot d_2(t) dt \right) \right). \end{aligned} \quad (A.2)$$

In order to obtain the maximum of Lagrangian function, we take derivatives with respect to  $d_1(t)$  and  $d_2(t)$ , i.e.

$$\begin{aligned} \frac{\partial L(d_1(t), d_2(t), \lambda_{oc_1})}{\partial d_1(t)} &= e^{-\rho t} \frac{{}^{11}_t p_x}{{}_t p_x} \sum_{k=1}^n \binom{n}{k} ({}_t p_x)^k (1 - {}_t p_x)^{n-k} \left( \frac{nd_1(t)}{k} \right)^{-\gamma} \\ &\quad - \lambda_{oc_1} \cdot e^{-rt} \frac{{}^{11}_t p_x}{{}_t p_x} \left( 1 - (1 - {}_t p_x)^n \right) \stackrel{!}{=} 0, \end{aligned} \quad (A.3)$$

$$\frac{\partial L(d_1(t), d_2(t), \lambda_{oc_1})}{\partial d_2(t)} = \alpha^{1-\gamma} \cdot e^{-\rho t} \frac{{}^{12}_t p_x}{{}_t p_x} \sum_{k=1}^n \binom{n}{k} ({}_t p_x)^k \cdot (1 - {}_t p_x)^{n-k} \left( \frac{nd_2(t)}{k} \right)^{-\gamma} \quad (A.4)$$

$$- \lambda_{oc_1} \cdot e^{-rt} \frac{{}^{12}_t p_x}{{}_t p_x} \left( 1 - (1 - {}_t p_x)^n \right) \stackrel{!}{=} 0. \quad (A.5)$$

Denote  $\kappa_{n,\gamma}({}_t p_x) = \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma ({}_t p_x)^k (1 - {}_t p_x)^{n-k}$ .

The Lagrangian function takes the global optima when

$$d_1^{oc_1*}(t) = \frac{e^{1/\gamma(r-\rho)t}}{(\lambda_{oc_1}^*)^{1/\gamma}} \cdot \frac{(\kappa_{n,\gamma}({}_t p_x))^{1/\gamma}}{(1 - (1 - {}_t p_x)^n)^{1/\gamma}}, \quad (A.6)$$

$$d_2^{oc1*}(t) = \alpha^{1/\gamma-1} \cdot \frac{e^{1/\gamma(r-\rho)t}}{(\lambda_{oc1}^*)^{1/\gamma}} \cdot \frac{(\kappa_{n,\gamma}(t p_x))^{1/\gamma}}{(1 - (1 - t p_x)^n)^{1/\gamma}}. \quad (A.7)$$

The  $\lambda_{oc1}^* > 0$  is chosen satisfying the budget constraint, i.e.

$$\begin{aligned} \lambda_{oc1}^* = & \left( \frac{1}{v} \left( \int_0^\infty e^{\left(\frac{r-\rho}{\gamma}-r\right)t} \cdot \frac{\left(\frac{{}^{11}t p_x}{t p_x} \cdot \kappa_{n,\gamma}(t p_x)\right)^{1/\gamma}}{\left(\frac{{}^{11}t p_x}{t p_x} \cdot (1 - (1 - t p_x)^n)\right)^{1/\gamma-1}} dt \right. \right. \\ & \left. \left. + \int_0^\infty e^{\left(\frac{r-\rho}{\gamma}-r\right)t} \cdot \alpha^{1/\gamma-1} \cdot \frac{\left(\frac{{}^{12}t p_x}{t p_x} \cdot \kappa_{n,\gamma}(t p_x)\right)^{1/\gamma}}{\left(\frac{{}^{12}t p_x}{t p_x} \cdot (1 - (1 - t p_x)^n)\right)^{1/\gamma-1}} dt \right) \right)^\gamma. \end{aligned} \quad (A.8)$$

#### A.2. Proofs for care-dependent tontines in two-pool case

When we distinguish two pools at each time  $t > 0$ , the policyholder's state-dependent expected utility  $U_0^{oc2}$  could be written as

$$\begin{aligned} U_0^{oc2} = & \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( u_1 \left( \frac{nd_1^{oc2}(t)}{N_1(t)} \right) \mathbb{1}_{\{S_{x+t}=1\}} + u_2 \left( \frac{nd_2^{oc2}(t)}{N_2(t)} \right) \mathbb{1}_{\{S_{x+t}=2\}} \right) dt \right] \\ = & \int_0^\infty e^{-\rho t} \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma ({}^{11}t p_x)^k (1 - {}^{11}t p_x)^{n-k} \cdot u(d_1^{oc2}(t)) dt \\ & + \alpha^{1-\gamma} \cdot \int_0^\infty e^{-\rho t} \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma ({}^{12}t p_x)^k (1 - {}^{12}t p_x)^{n-k} \cdot u(d_2^{oc2}(t)) dt. \end{aligned} \quad (A.9)$$

Afterwards, we write down the Lagrangian function. Here, we also simplify the notations  $d_1^{oc2}(t)$  and  $d_2^{oc2}(t)$  as  $d_1(t)$ ,  $d_2(t)$  respectively.

$$\begin{aligned} L(d_1(t), d_2(t), \lambda_{oc2}) = & \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( u_1 \left( \frac{nd_1(t)}{N_1(t)} \right) \mathbb{1}_{\{S_{x+t}=1\}} + u_2 \left( \frac{nd_2(t)}{N_2(t)} \right) \mathbb{1}_{\{S_{x+t}=2\}} \right) dt \right] + \lambda_{oc2} (v - P_0^{oc2}) \\ = & \int_0^\infty e^{-\rho t} \kappa_{n,\gamma}({}^{11}t p_x) u(d_1(t)) dt + \alpha^{1-\gamma} \cdot \int_0^\infty e^{-\rho t} \kappa_{n,\gamma}({}^{12}t p_x) u(d_2(t)) dt \\ & + \lambda_{oc2} \left( v - \left( \int_0^\infty e^{-rt} \left( 1 - (1 - {}^{11}t p_x)^n \right) d_1(t) dt + \int_0^\infty e^{-rt} \left( 1 - (1 - {}^{12}t p_x)^n \right) d_2(t) dt \right) \right). \end{aligned} \quad (A.10)$$

As denoted,  $\kappa_{n,\gamma}({}^{ij}t p_x) = \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma ({}^{ij}t p_x)^k (1 - {}^{ij}t p_x)^{n-k}$ ,  $i, j \in \mathcal{S}$ .

Then we take derivatives with respect to  $d_1(t)$  and  $d_2(t)$  to maximize the Lagrangian function

$$\frac{\partial L(d_1(t), d_2(t), \lambda_{oc2})}{\partial d_1(t)} = e^{-\rho t} \kappa_{n,\gamma}({}^{11}t p_x) (d_1(t))^{-\gamma} - \lambda_{oc2} e^{-rt} \left( 1 - (1 - {}^{11}t p_x)^n \right) \stackrel{!}{=} 0, \quad (A.11)$$

$$\frac{\partial L(d_1(t), d_2(t), \lambda_{oc2})}{\partial d_2(t)} = \alpha^{1-\gamma} \cdot e^{-\rho t} \kappa_{n,\gamma}({}^{12}t p_x) (d_2(t))^{-\gamma} - \lambda_{oc2} e^{-rt} \left( 1 - (1 - {}^{12}t p_x)^n \right) \stackrel{!}{=} 0. \quad (A.12)$$

The Lagrangian function takes the global optima when

$$d_1^*(t) = \frac{e^{1/\gamma(r-\rho)t}}{(\lambda_{oc2}^*)^{1/\gamma}} \cdot \frac{(\kappa_{n,\gamma}({}^{11}t p_x))^{1/\gamma}}{(1 - (1 - {}^{11}t p_x)^n)^{1/\gamma}}, \quad (A.13)$$

$$d_2^*(t) = \alpha^{1/\gamma-1} \cdot \frac{e^{1/\gamma(r-\rho)t}}{(\lambda_{oc2}^*)^{1/\gamma}} \cdot \frac{(\kappa_{n,\gamma}({}^{12}t p_x))^{1/\gamma}}{(1 - (1 - {}^{12}t p_x)^n)^{1/\gamma}}. \quad (A.14)$$

The  $\lambda_{oc_2}^* > 0$  is chosen satisfying the budget constraint, i.e.

$$\lambda_{oc_2}^* = \left( \frac{1}{v} \left( \int_0^\infty e^{\left(\frac{r-\rho}{\gamma}-r\right)t} \left[ \frac{(\kappa_{n,\gamma}({}^{11}_t p_x))^{1/\gamma}}{\left(1 - (1 - {}^{11}_t p_x)^n\right)^{1/\gamma-1}} + \frac{\alpha^{1/\gamma-1} \cdot (\kappa_{n,\gamma}({}^{12}_t p_x))^{1/\gamma}}{\left(1 - (1 - {}^{12}_t p_x)^n\right)^{1/\gamma-1}} \right] dt \right) \right)^\gamma. \quad (A.15)$$

### A.3. Proofs for care-dependent annuities

For simplicity, we use  $c_1(t)$  and  $c_2(t)$  to represent  $c_1^{ac}(t)$  and  $c_2^{ac}(t)$  separately. Then, the Lagrangian function is given by

$$L(c_1(t), c_2(t), \lambda_{ac}) = \int_0^\infty e^{-\rho t} {}^{11}_t p_x \cdot u(c_1(t)) dt + \alpha^{1-\gamma} \cdot \int_0^\infty e^{-\rho t} {}^{12}_t p_x \cdot u(c_2(t)) dt + \lambda_{ac} \cdot \left( v - \int_0^\infty e^{-rt} {}^{11}_t p_x c_1(t) dt - \int_0^\infty e^{-rt} {}^{12}_t p_x c_2(t) dt \right). \quad (A.16)$$

Next, by optimal conditions  $\partial L(c_1(t), c_2(t), \lambda_{ac}) / \partial c_1(t) \stackrel{!}{=} 0$  and  $\partial L(c_1(t), c_2(t), \lambda_{ac}) / \partial c_2(t) \stackrel{!}{=} 0$ , the optimal payoffs for care-dependent annuities are yielded, i.e.

$$c_1^{ac*}(t) = \left( \lambda_{ac}^* \cdot e^{(\rho-r)t} \right)^{-1/\gamma}, \quad (A.17)$$

$$c_2^{ac*}(t) = \alpha^{1/\gamma-1} \cdot \left( \lambda_{ac}^* \cdot e^{(\rho-r)t} \right)^{-1/\gamma}. \quad (A.18)$$

Through the budget constraint, we obtain

$$\lambda_{ac}^* = \left( \frac{1}{v} \int_0^\infty e^{\left(\frac{r-\rho}{\gamma}-r\right)t} ({}^{11}_t p_x + \alpha^{1/\gamma-1} \cdot {}^{12}_t p_x) dt \right)^\gamma. \quad (A.19)$$

### A.4. Proofs for utility inequalities

By Equations (3.27), (3.13) and (3.21), and let  $P_0^{ac} = P_0^{oc_1} = P_0^{oc_2} = v$ , then we have

$$\begin{aligned} U_0^{ac*} &= \int_0^\infty e^{-\rho t} {}^{11}_t p_x \cdot u_1(c_1^{ac*}(t)) dt + \int_0^\infty e^{-\rho t} {}^{12}_t p_x \cdot u_2(c_2^{ac*}(t)) dt \\ &= \frac{1}{1-\gamma} \int_0^\infty e^{-\rho t} {}^{11}_t p_x \cdot (c_1^{ac*}(t))^{1-\gamma} dt + \frac{\alpha^{1/\gamma-1}}{1-\gamma} \int_0^\infty e^{-\rho t} {}^{12}_t p_x \cdot (c_2^{ac*}(t))^{1-\gamma} dt \\ &= \frac{\lambda_{ac}^* v}{1-\gamma}, \end{aligned} \quad (A.20)$$

$$\begin{aligned} U_0^{oc_1*} &= \int_0^\infty e^{-\rho t} \frac{{}^{11}_t p_x}{t p_x} \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma ({}_t p_x)^k (1 - {}_t p_x)^{n-k} \cdot u_1(d_1^{oc_1*}(t)) dt \\ &\quad + \int_0^\infty e^{-\rho t} \frac{{}^{12}_t p_x}{t p_x} \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma ({}_t p_x)^k \cdot (1 - {}_t p_x)^{n-k} \cdot u_2(d_2^{oc_1*}(t)) dt \\ &= \frac{1}{1-\gamma} \int_0^\infty e^{-\rho t} \frac{{}^{11}_t p_x}{t p_x} \kappa_{n,\gamma}({}_t p_x) (d_1^{oc_1*})^{1-\gamma} dt + \frac{\alpha^{1/\gamma-1}}{1-\gamma} \int_0^\infty e^{-\rho t} \frac{{}^{12}_t p_x}{t p_x} \kappa_{n,\gamma}({}_t p_x) (d_2^{oc_1*})^{1-\gamma} dt \\ &= \frac{\lambda_{oc_1}^* v}{1-\gamma}, \end{aligned} \quad (A.21)$$

$$\begin{aligned} U_0^{oc_2*} &= \int_0^\infty e^{-\rho t} \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma ({}^{11}_t p_x)^k (1 - {}^{11}_t p_x)^{n-k} \cdot u_1(d_1^{oc_2*}(t)) dt \\ &\quad + \int_0^\infty e^{-\rho t} \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma ({}^{12}_t p_x)^k \cdot (1 - {}^{12}_t p_x)^{n-k} \cdot u_2(d_2^{oc_2*}(t)) dt \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{1-\gamma} \int_0^\infty e^{-\rho t} \kappa_{n,\gamma}({}^{11}_t p_x) (d_1^{oc_2*})^{1-\gamma} dt + \frac{\alpha^{1/\gamma-1}}{1-\gamma} \int_0^\infty e^{-\rho t} \kappa_{n,\gamma}({}^{12}_t p_x) (d_2^{oc_2*})^{1-\gamma} dt \\
&= \frac{\lambda_{oc_2}^* v}{1-\gamma}.
\end{aligned} \tag{A.22}$$

To prove (3.28), we only need to demonstrate by

$$U_0^{ac*} \geq U_0^{oc_1*} \Leftrightarrow \begin{cases} \lambda_{ac}^* \geq \lambda_{oc_1}^*, & \text{if } \gamma \in (0, 1) \\ \lambda_{ac}^* \leq \lambda_{oc_1}^*, & \text{if } \gamma > 1 \end{cases}. \tag{A.23}$$

Before going towards the proofs, we define

$$\begin{aligned}
a &:= \left( \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1 \right), \\
b &:= (1, 1, \dots, 1, 1).
\end{aligned}$$

Such that

$$\begin{aligned}
\|a\|_{L^\infty} &= 1, & \|a\|_{L^1} &= p, \\
\|b\|_{L^\infty} &= 1, & \|b\|_{L^1} &= 1 - (1-p)^n,
\end{aligned}$$

the  $L^\gamma$ -norm of  $x$  is written as

$$\|x\|_{L^\gamma} := \left[ \sum_{k=1}^n x_k^\gamma \binom{n}{k} p^k (1-p)^{n-k} \right]^{1/\gamma}. \tag{A.24}$$

Here we take the proof for (3.28) as an example:

**Proof.** For  $\gamma > 1$ , define  $\tilde{\gamma} = \frac{\gamma}{\gamma-1}$ , s.t.  $\frac{1}{\gamma} + \frac{1}{\tilde{\gamma}} = 1$ .

$$\begin{aligned}
&\lambda_{ac}^* \leq \lambda_{oc_1}^* \\
&\Leftrightarrow {}^{11}_t p_x + \alpha^{1/\gamma-1} \cdot {}^{12}_t p_x \leq \frac{\left( \frac{{}^{11}_t p_x}{{}_t p_x} \cdot \kappa_{n,\gamma}({}_t p_x) \right)^{1/\gamma}}{\left( \frac{{}^{11}_t p_x}{{}_t p_x} \cdot (1 - (1-{}_t p_x)^n) \right)^{1/\gamma-1}} + \alpha^{1/\gamma-1} \cdot \frac{\left( \frac{{}^{12}_t p_x}{{}_t p_x} \cdot \kappa_{n,\gamma}({}_t p_x) \right)^{1/\gamma}}{\left( \frac{{}^{12}_t p_x}{{}_t p_x} \cdot (1 - (1-{}_t p_x)^n) \right)^{1/\gamma-1}} \\
&\Leftrightarrow {}^{11}_t p_x + \alpha^{1/\gamma-1} \cdot {}^{12}_t p_x \leq \left( {}^{11}_t p_x + \alpha^{1/\gamma-1} \cdot {}^{12}_t p_x \right) \cdot \frac{(1 - (1-{}_t p_x)^n)}{{}_t p_x} \cdot \frac{(\kappa_{n,\gamma}({}_t p_x))^{1/\gamma}}{(1 - (1-{}_t p_x)^n)^{1/\gamma}} \\
&\Leftrightarrow {}_t p_x \leq (\kappa_{n,\gamma}({}_t p_x))^{1/\gamma} \cdot (1 - (1-{}_t p_x)^n)^{1-1/\gamma} \\
&\xLeftrightarrow{\text{Hoelder Inequality}} {}_t p_x = \|a\|_{L^1} \leq \|a\|_{L^\gamma} \|b\|_{L^{\tilde{\gamma}}}.
\end{aligned}$$

And for  $\gamma \in (0, 1)$ ,

$$\begin{aligned}
&\lambda_{ac}^* \geq \lambda_{oc_1}^* \\
&\Leftrightarrow {}_t p_x \geq \|a\|_{L^\gamma} \|b\|_{L^{\tilde{\gamma}}} \\
&\xLeftrightarrow{\text{Hoelder Inequality}} {}_t p_x = \|a\|_{L^1} \geq \|a\|_{L^\gamma} \|b\|_{L^{\tilde{\gamma}}}. \quad \square
\end{aligned}$$

Analogously, for (3.29), it can be proved that

$$U_0^{ac*} \geq U_0^{oc_2*} \Leftrightarrow \begin{cases} \lambda_{ac}^* \geq \lambda_{oc_2}^*, & \text{if } \gamma \in (0, 1) \\ \lambda_{ac}^* \leq \lambda_{oc_2}^*, & \text{if } \gamma > 1 \end{cases}. \tag{A.25}$$

#### A.5. GLM approximation

After obtaining the crude transition intensities, we use a GLM model to smooth them. We build up estimations for mentioned  ${}^{12}\mu_y$ ,  ${}^{13}\mu_y$ , and  ${}^{23}\mu_y$ . To avoid lengthy, here we take  ${}^{12}\mu_y$  for example.

##### GLM model

The GLM model is composited by three parts: the link function, linear predictor, and probability distribution.

**Table A.1**

Risk loadings for various care-dependent products w.r.t. the capital requirement of Solvency II, using the baseline parameter setting from Table 5.1. Net premium  $\hat{p}_0^{oc1} = \hat{p}_0^{oc2} = \hat{p}_0^{ac} = v = 10000$ , pool size  $n = 1000$ , subjective discount rate  $\rho = 0.02$ , risk-free rate  $r = 0.02$ , initial age  $x = 60$ , risk aversion coefficient  $\gamma = 2$ , and payment weighting factor  $\alpha = 0.5$ .

CoC	$F_0^{oc1}$	$F_0^{oc2}$	$F_0^{ac}$
0.06	58.6	0.38	653

*Link function:* Here we adopt a log link function  $g(\cdot)$  (Hanewald et al., 2019).

$$g(^{12}\mu_y) = \ln(^{12}\mu_y) = \eta_y, \quad (\text{A.26})$$

where  $\eta_y$  is the linear predictor.

*Linear predictor:* Suppose the transition intensity is only relevant to age and gender. For each gender, the linear predictor can be estimated by

$$\eta_y = \sum_{m=0}^k \beta_m y^m = \beta_0 + \beta_1 x + \beta_2 y^2 + \cdots + \beta_k y^k. \quad (\text{A.27})$$

With reference to Fong and Feng (2016), we set  $k \leq 3$  to avoid model overfitting.

*Probability distribution:* We further assume the transition intensity is constant for each one-year age group in a given time interval. Suppose transition number follows an independent Poisson distribution between two consecutive survey waves, i.e.

$$^{12}C_y \sim \text{Poisson}(^1Y_y \cdot ^{12}\mu_y), \quad (\text{A.28})$$

where  $^{12}C_y$  is the transition number that one transiting from healthy to severely sick at age  $y$ . And  $^1Y_y$  represents the year(s) of risk exposure in healthy state at age  $y$ .

#### Estimation and model selection

Moreover, we use Maximization Likelihood Estimation (MLE) to approximate the coefficients  $\beta_m$ . The log-likelihood function is given by

$$\ln F = \sum_y -^1Y_y ^{12}\mu_y + ^{12}C_y \ln(^1Y_y ^{12}\mu_y) + A, \quad (\text{A.29})$$

where  $A$  is a constant and  $F$  represents the likelihood function.

Finally, we use AIC and BIC criteria to decide the optimal  $k$ . Based on the estimated  $\beta_m$ , we can forecast the transition intensity for every agegroup, then the transition probabilities can be further computed.

#### A.6. Risk loading computation w.r.t. solvency II

We further compute the risk loadings following the capital requirements of Solvency II. Applying a cost of capital (CoC) ratio of 0.06 which is currently used by Solvency II (see, e.g. EIOPA (2014)), we get the risk loadings for different care-dependent products. Similar to the results computed under the requirement of C-ROSS, care-dependent annuity charges the highest risk loading, while under the baseline parameter setting (see Table 5.1), the two-pool care-dependent tontine still requires the least (see Table A.1).

As already known from subsection 5.2.2, under the baseline parameter setting (Table 5.1), the utility indifference number  $Q_1 = 1.000896$  and  $Q_2 = 1.003010$ . Thus, combined with risk loadings in Table A.1, the comparable gross premiums for different care-dependent products are given by

$$Q_1 \cdot \hat{p}^{oc1} = 1.000896 \times (10000 + 58.6) = 10067.613,$$

$$Q_2 \cdot \hat{p}^{oc2} = 1.003010 \times (10000 + 0.38) = 10030.481,$$

$$\hat{p}^{ac} = 10000 + 653 = 10653.$$

From the above results we can learn that regarding the comparable gross premiums, people's preference order for these three care-dependent products stays the same as previously computed under the C-ROSS, i.e. two-pool CDT > one-pool CDT > CDA. This holds of course only for given parameters in Table 5.1.

#### A.7. Additional tables of sensitivity analysis

*How do the risk loadings vary with different coefficients ( $\alpha$ ,  $\gamma$ ,  $r$ )?*

**Table A.2**

Sensitivity analysis of risk loadings w.r.t. risk aversion coefficient  $\gamma$  and payment weighting factor  $\alpha$ . Other parameters are taken from the baseline setting table (i.e. Table 5.1). Net premium  $P_0^{oc1} = P_0^{oc2} = P_0^{ac} = v = 10000$ , pool size  $n = 1000$ , risk-free rate  $r = 0.02$ , subjective discount rate  $\rho = 0.02$ , and initial age  $x = 60$ .

$F_0^{oc1}$				$F_0^{oc2}$			$F_0^{ac}$		
$\alpha \in (0, 1), \gamma > 1$									
$\alpha \backslash \gamma$	2	5	8	2	5	8	2	5	8
0.2	103.09943	192.84282	218.4911	0.2477858	0.3088327	0.3682938	729.2807	916.3349	969.8961
0.5	37.58904	65.09438	72.68008	0.2552838	0.3308881	0.4024555	595.1179	650.9564	666.1752
0.8	11.06583	18.20009	20.07213	0.2583212	0.3390165	0.4148757	541.0479	555.5173	559.2484
$\alpha > 1, \gamma \in (0, 1)$									
$\alpha \backslash \gamma$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
1.5	264.9986	44.9172	10.00598	0.191409	0.2187796	0.2288318	1070.098	610.1457	538.9051
2	531.6991	85.31641	17.58199	0.1657416	0.2148421	0.2280768	1641.7743	692.7645	554.3612
5	837.1008	262.20912	44.52996	0.1376781	0.1975901	0.2253912	2301.2606	1064.105	609.3379

**Table A.3**

Sensitivity analysis of risk loadings w.r.t. risk-free rate  $r$ . Other parameters are taken from the baseline setting table (i.e. Table 5.1). Net premium  $P_0^{oc1} = P_0^{oc2} = P_0^{ac} = v = 10000$ , pool size  $n = 1000$ , subjective discount rate  $\rho = 0.02$ , initial age  $x = 60$ , risk aversion coefficient  $\gamma = 2$ , and payment weighting factor  $\alpha = 0.5$ .

$r$	$F_0^{oc1}$	$F_0^{oc2}$	$F_0^{ac}$	$F_0^{oc1}$	$F_0^{oc2}$	$F_0^{ac}$
	$\gamma = 2, \alpha = 0.5$			$\gamma = 0.5, \alpha = 2$		
0.00	40.28555	0.339966	690.5326	74.80001	0.1196744	515.2861
0.01	38.91467	0.2948221	641.068	79.88006	0.160809	597.377
0.02	37.58904	0.2552838	595.1179	85.31641	0.2148421	692.7645
0.03	36.30882	0.2207213	552.4653	91.1043	0.2852828	803.2663
0.04	35.07392	0.1905645	512.9012	97.23027	0.3763815	930.8164
0.05	33.88404	0.1642992	476.225	103.67105	0.4932094	1077.4339

How do the utility indifference numbers vary with different coefficients ( $\alpha, \gamma, r$ )?

**Table A.4**

Sensitivity analysis of utility indifference numbers w.r.t. risk aversion coefficient  $\gamma$  and payment weighting factor  $\alpha$ . Other parameters are taken from the baseline setting table (i.e. Table 5.1). Net premium  $P_0^{oc1} = P_0^{oc2} = P_0^{ac} = v = 10000$ , pool size  $n = 1000$ , risk-free rate  $r = 0.02$ , subjective discount rate  $\rho = 0.02$ , and initial age  $x = 60$ .

$Q_1$				$Q_2$		
$\alpha \in (0, 1), \gamma > 1$						
$\alpha \backslash \gamma$	2	5	8	2	5	8
0.2	1.001053	1.003059	1.004994	1.00394	1.012662	1.021157
0.5	1.000896	1.002335	1.003723	1.00301	1.008264	1.013327
0.8	1.000833	1.00207	1.003265	1.002634	1.006652	1.010508
$\alpha > 1, \gamma \in (0, 1)$						
$\alpha \backslash \gamma$	0.2	0.5	0.8	0.2	0.5	0.8
1.5	1.000149	1.000233	1.000337	1.000638	1.000792	1.001061
2	1.000215	1.000258	1.000344	1.001027	1.000938	1.001105
5	1.000292	1.000368	1.000371	1.001472	1.001579	1.00126

**Table A.5**

Sensitivity analysis of utility indifference numbers w.r.t. risk-free rate  $r$ . Other parameters are taken from the baseline setting table (i.e. Table 5.1). Net premium  $P_0^{oc1} = P_0^{oc2} = P_0^{ac} = v = 10000$ , pool size  $n = 1000$ , subjective discount rate  $\rho = 0.02$ , initial age  $x = 60$ , risk aversion coefficient  $\gamma = 2$ , and payment weighting factor  $\alpha = 0.5$ .

$r$	$Q_1$	$Q_2$	$Q_1$	$Q_2$
	$\gamma = 2, \alpha = 0.5$		$\gamma = 0.5, \alpha = 2$	
0.00	1.00106	1.003316	1.000184	1.000798
0.01	1.000974	1.003157	1.000217	1.000863
0.02	1.000896	1.00301	1.000258	1.000938
0.03	1.000825	1.002875	1.000307	1.001027
0.04	1.00076	1.002751	1.000365	1.001132
0.05	1.0007	1.002636	1.000434	1.001255



## References

- AIC, IPLE, 2020. 2018–2019 China long-term care research report. Technical report.
- Ameriks, J., Briggs, J., Caplin, A., Shapiro, M.D., Tonetti, C., 2020. Long-term-care utility and late-in-life saving. *Journal of Political Economy* 128 (6), 2375–2451.
- Bauer, D., Börgers, M., Ruß, J., 2010. On the pricing of longevity-linked securities. *Insurance, Mathematics & Economics* 46 (1), 139–149.
- Brown, J., Warshawsky, M., 2013. The life care annuity: a new empirical examination of an insurance innovation that addresses problems in the markets for life annuities and long-term care insurance. *The Journal of Risk and Insurance* 80 (3), 677–704.
- Brown, J.R., Finkelstein, A., 2008. The interaction of public and private insurance: medicaid and the long-term care insurance market. *The American Economic Review* 98 (3), 1083–1102.
- Brown, J.R., Finkelstein, A., 2009. The private market for long-term care insurance in the United States: a review of the evidence. *The Journal of Risk and Insurance* 76 (1), 5–29.
- CBIRC, 2020a. Supervision rules for the second phase of the C-ROSS (draft for comments) || assessment of life insurance contract liability.
- CBIRC, 2020b. Supervision rules for the second phase of the C-ROSS (draft for comments) || minimum capital requirement for insurance risk (life insurance business).
- Chen, A., Fuino, M., Sehner, T., Wagner, J., 2021. Valuation of long-term care options embedded in life annuities. *Annals of Actuarial Science*, 1–27.
- Chen, A., Hieber, P., Klein, J.K., 2019. Tontuity: a novel individual-oriented retirement plan. *ASTIN Bulletin: The Journal of the IAA* 49 (1), 5–30.
- Chen, A., Hieber, P., Rach, M., 2020. Optimal retirement products under subjective mortality beliefs. *Insurance, Mathematics & Economics* 101, 55–69.
- Chen, A., Rach, M., 2019. Options on tontines: an innovative way of combining annuities and tontines. *Insurance, Mathematics & Economics* 89, 182–192.
- Chen, A., Rach, M., 2021. Bequest-embedded annuities and tontines. *Asia-Pacific Journal of Risk and Insurance* 16 (1), 1–46.
- Davidoff, T., 2010. Home equity commitment and long-term care insurance demand. *Journal of Public Economics* 94 (1–2), 44–49.
- Denuit, M., Lucas, N., Pitacco, E., 2019. Pricing and reserving in LTC insurance. In: *Actuarial Aspects of Long-Term Care*. Springer, pp. 129–158.
- EIOPA, 2014. Technical specification for the Solvency II preparatory phase (part I). Technical report.
- Evans, W.N., Viscusi, W.K., 1991. Estimation of state-dependent utility functions using survey data. *Review of Economics and Statistics*, 94–104.
- Fong, J.H., Feng, J., 2016. Patterns of functional disability in the oldest adults in China. *Journal of the American Geriatrics Society* 64 (9), 1890–1894.
- Friedman, B., 1974. Risk aversion and the consumer choice of health insurance option. *Review of Economics and Statistics* 56 (2), 209–214.
- Haberman, S., Pitacco, E., 1998. *Actuarial Models for Disability Insurance*. Chapman and Hall, London.
- Haberman, S., Renshaw, A., 2009. On age-period-cohort parametric mortality rate projections. *Insurance, Mathematics & Economics* 45 (2), 255–270.
- Hanewald, K., Li, H., Shao, A.W., 2019. Modelling multi-state health transitions in China: a generalised linear model with time trends. *Annals of Actuarial Science* 13 (1), 145–165.
- Havranek, T., Horvath, R., Irsova, Z., Rusnak, M., 2015. Cross-country heterogeneity in intertemporal substitution. *Journal of International Economics* 96 (1), 100–118.
- Hieber, P., Lucas, N., 2020. Modern Life-Care Tontines. Available at SSRN 3688386.
- Hoermann, G., Ruß, J., 2008. Enhanced annuities and the impact of individual underwriting on an insurer's profit situation. *Insurance, Mathematics & Economics* 43 (1), 150–157.
- Hu, X., Chen, B., Zhu, W., 2016. Long-term care insurance pricing on the basis of household survey data. *Insurance Studies* 4, 57–67.
- Katz, S., Ford, A.B., Moskowitz, R.W., Jackson, B.A., Jaffe, M.W., 1963. Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function. *JAMA* 185 (12), 914–919.
- Milevsky, M.A., Salisbury, T.S., 2015. Optimal retirement income tontines. *Insurance, Mathematics & Economics* 64, 91–105.
- Morrow, K.M., Röger, W., 2003. Economic and financial market consequences of ageing populations. Technical Report 182.
- Murtaugh, C.M., Spillman, B.C., Warshawsky, M.J., 2001. In sickness and in health: an annuity approach to financing long-term care and retirement income. *The Journal of Risk and Insurance* 68 (2), 225–254.
- Pitacco, E., 2015. Guarantee structures in life annuities: a comparative analysis. *The Geneva Papers on Risk and Insurance, Issues and Practice* 41 (1), 78–97.
- Pitacco, E., 2016. Premiums for long-term care insurance packages: sensitivity with respect to biometric assumptions. *Risks* 4 (1), 3.
- Ramsay, C.M., Oguledo, V.I., 2020. Doubly enhanced annuities (deans) and the impact of quality of long-term care under a multi-state model of activities of daily living (ADL). *North American Actuarial Journal* 24 (1), 57–99.
- Spillman, B.C., Murtaugh, C.M., Warshawsky, M.J., 2003. Policy implications of an annuity approach to integrating long-term care financing and retirement income. *Journal of Aging and Health* 15 (1), 45–73.
- Statista, 2021. Share of population aged 60 and older in China from 1950 to 2010 with forecasts until 2100. Available online at: <https://www.statista.com/statistics/251529/share-of-persons-aged-60-and-older-in-the-chinese-population/>. (Accessed 6 April 2021).
- Statistics Bureau of the P.R. China, 2020. *China Statistical Yearbook 2015–2019*. China Statistics Press, Beijing.
- Szpiro, G.G., 1986. Measuring risk aversion: an alternative approach. *Review of Economics and Statistics* 68 (1), 156–159.
- Vidal-Melia, C., Ventura-Marco, M., Pla-Porcel, J., 2016. Life care annuities (LCA) embedded in a notional defined contribution (NDC) framework. *ASTIN Bulletin* 46 (2), 331–363.
- Webb, D.C., 2009. Asymmetric information, long-term care insurance, and annuities: the case for bundled contracts. *The Journal of Risk and Insurance* 76 (1), 53–85.
- Yaari, M.E., 1965. Uncertain lifetime, life insurance, and the theory of the consumer. *The Review of Economic Studies* 32 (2), 137–150.
- Zheng, S., Xu, J., Xiong, L., 2013. Measuring the risk margin of life insurance liability in light of the newly issued China accounting standards and solvency II. *Management Review* 25 (3), 159–170.
- Zhou-Richter, T., Gründl, H., 2011. Life care annuities - trick or treat for insurance companies? Available at SSRN 1856994.



## **2 Dynamic tonuity: Adapting retirement benefits to a changing environment**

**Source:**

Chen, A., Chen, Y., & Rach, M. (2024). Dynamic tonuity: Adapting retirement benefits to a changing environment. Submitted to *ASTIN Bulletin: The Journal of the IAA* (under review).

URL: <https://ssrn.com/abstract=4810735>



# Dynamic tonuity: Adapting retirement benefits to a changing environment

An Chen\*, Yusha Chen\* and Manuel Rach<sup>§</sup>

\* Institute of Insurance Science, Ulm University, Helmholtzstr. 20, 89081 Ulm, Germany.  
E-mails: an.chen@uni-ulm.de; cys.0416@outlook.com.

<sup>§</sup> Institute of Insurance Economics, University of St. Gallen, Tannenstrasse 19, 9000 St. Gallen, Switzerland. E-Mail: manuel.rach@unisg.ch

## Abstract

The tonuity, proposed by Chen et al. (2019), is a combination of an immediate tontine and a deferred annuity. However, its switching time from tontine to annuity is fixed at the moment the contract is closed, possibly becoming sub-optimal if mortality changes over time. This paper introduces an alternative tonuity product, wherein a dynamic switching condition is pivotal, relying on the observable mortality trends within a reference population. The switching from tontine to annuity then occurs automatically once the condition is satisfied. Using data from the Human Mortality Database and UK Continuous Mortality Investigation, we demonstrate that, in a changing environment, where an unforeseen mortality/longevity shock leads to an unexpected increase/decrease in mortality rates, the proposed dynamic tonuity contract can be preferable to the regular tonuity contract.

**Keywords:** Retirement plan, tonuity, dynamic switching time, two-population stochastic mortality

**JEL:** G22, G52, J32

# 1 Introduction

Different from annuities, in a tontine contract, longevity risk is shared among a pool of policyholders. Tontines can be attractive as they require lower safety loadings compared to annuities (e.g., Piggott et al., 2005; Sabin, 2010; Richter and Weber, 2011; Newfield, 2014; Fullmer and Sabin, 2019). However, highly-fluctuating payoffs of tontines are not desirable for old ages (e.g., Gemmo et al., 2020; Weinert and Gründl, 2021), and generate lower expected lifetime utilities than annuities when considering actuarially fair pricing (Milevsky and Salisbury, 2015). Under this background, partial tontinization or combined products, such as tontines with minimum guarantees and tonuities are proposed and have demonstrated advantages compared to annuities and tontines (e.g., Donnelly and Young, 2017; Chen et al., 2019; Chen and Rach, 2019; Weinert and Gründl, 2021).

The tonuity (Chen et al., 2019) is a combination of an immediate tontine and a deferred annuity, providing tontine-like payoffs till some fixed age, and annuity-like payoffs afterwards. It is found to be almost optimal in the analysis of optimal portfolios of tontines and annuities (Chen et al., 2020). However, its current design exhibits some potential flaws and disadvantages which are to be explained in the following. First, the switching time is fixed at the moment the contract is closed. Nevertheless, the Covid-19 pandemic and medical advancements raise the question how an unanticipated mortality/longevity shock affects policyholders' decision on the optimal switching time ex post. In such a changing circumstance, where the unforeseen mortality/longevity shock results in significant deviations in the predicted survival probabilities, the policyholder may suffer a loss in utility. Note that, in this paper, the terms "mortality shock" and "longevity shock" are used in a narrow sense to denote instances of mortality worsening and mortality improvement, respectively. The second possible drawback of the regular tonuity is that it requires policyholders to choose a switching time according to their risk preferences at the beginning. However, they might not possess the requisite financial knowledge to make informed decisions or may simply lack the time needed to make sound judgments. Of course, the insurance company can help policyholders make the decision. Then another question relevant to transparency pops up: How can policyholders validate whether the insurance company's choice of switching time is optimal for them or not? In most fields of financial regulation, customer protection emphasizes on the importance of transparency, including the insurance market regulation (Schwarcz, 2013). This calls for insurance products to be designed in a more transparent way.

In order to address these two concerns, we come up with a modified design which we call

dynamic tonuity. This system transitions from tontine-style payouts to annuity-style payouts upon meeting specific switching conditions based on mortality trends experienced in some reference population (e.g., national population). This design ensures that the switching time automatically adjusts to changes in mortality happening after the beginning of the contract. Furthermore, we can make the switching condition visible to policyholders. As an example, the switching might occur when the one-year survival probability of a reference population at some age passes a certain threshold index. The one-year survival probability is accessible in public sources, and the threshold index is specified on the contract by the insurer at the outset, which is also viewable.

Assuming that the policyholder's utility on the product payoffs is described by a constant relative risk aversion (CRRA) utility function, we solve the maximization problems and obtain analytic expressions for the optimal payoff structures of the dynamic tonuity. Applying a two-population mortality model, together with mortality data from UK Continuous Mortality Investigation (CMI) and Human Mortality Database (HMD), we estimate and simulate the correlated mortality rates for two populations, i.e. the reference population and the book population. The book population can be referred to the insurer's annuitant pool, which is smaller than the reference population. Here, we distinguish among different cases: (i) The base case, where mortality rates develop as expected; and (ii) The stress case, where after the contract is made under the base-case assumptions, some unforeseen mortality/longevity shock strikes. To illustrate the unexpected mortality shock, we use a data set containing data of some pandemic years to estimate and project the mortality rates; while for the base case, we use a subset of the data which does not include the pandemic years' data to obtain the mortality rates. As for the unanticipated longevity shock, it cannot be generated in a similar way as the unexpected mortality shock due to the data limitation. Instead, we create a fictional stress case with the longevity shock. Drawing insights from the mortality data and our analytical findings, we make a comprehensive discussion and comparative analysis of various tonuity products, considering them from the standpoint of the policyholder.

Our numerical findings demonstrate that a properly designed dynamic tonuity can be preferable to the regular tonuity in a changing environment. The policyholder's risk aversion is a crucial factor influencing contract preference in the stress case. Specifically, when the policyholder's risk aversion coefficient exceeds 1 in the stress cases, policyholders with a lower level of risk aversion prefer the dynamic tonuity contract, as the less risk averse policyholders have greater tolerance for the dynamic tonuity contract's more-volatile payoffs.

When the policyholder's risk aversion coefficient falls within the interval  $(0, 1)$ , the results regarding the unforeseen mortality shock remain consistent. An anomaly in contract attractiveness emerges concerning the unexpected longevity shock. In the event of an unforeseen longevity shock, the regular tonuity contract surpasses the dynamic tonuity contract in attractiveness. This can be attributed to policyholders with a risk aversion coefficient in the range of  $(0, 1)$  expecting a tontine contract. And this is realized by the regular tonuity contract, as it is optimal not to undergo switching for such policyholders. However, the dynamic tonuity contract's switching point is independent of the policyholder's preference but is determined by an external event, influencing the premium distribution between annuity and tontine components. The unexpected longevity shock causes a shift in the switching time to an earlier point, resulting in tontine-specific payoffs being disbursed for a shorter period, while annuity-specific payoffs extend over a longer term. This restructuring of the payoffs may diminish the appeal of the dynamic tonuity contract.

The remainder of this paper is structured as follows: Section 2 reviews the construction of a regular tonuity and demonstrates the disadvantage of a regular tonuity with numerical exercises. Section 3 gives the model setup for the dynamic tonuity. Moreover, we derive the premium and optimal payoffs of the dynamic tonuity. In Section 4, with reference to some thresholds, we provide numerical illustrations to the dynamic switching time and optimal payoffs of the dynamic tonuity. Afterwards, we compare it with the regular one in different cases. In the last section, we conclude the main findings and implications of this paper. The mortality models and several proofs are collected in the appendix.

## 2 A Review of the Regular Tonuity

Before turning to the introduction of the dynamic tonuity, let us give a brief overview of a regular tonuity. In order to align more closely with the content of this paper, we employ the mortality assumption specified within this study for the valuation of the regular tonuity. We start by reviewing the premium and optimal payoffs of the regular tonuity. Afterwards, we point out the potential disadvantage(s) of the regular tonuity.



## 2.1 The Regular Tonuity

We consider  $n$  identical policyholders, each endowed with an initial wealth amounting to  $v > 0$  who use their entire wealth to purchase some retirement plan. We assume that the insurers design regular tonuity contracts based on mortality rates of their policyholder pool, i.e. the *book population*. We use  $T^{(b)}$  for the random remaining lifetime of an individual in the book population, and  $\mu_{x,h}^{(b)}$  to denote the corresponding stochastic force of mortality of an  $x$ -year old in year  $h$ . Correspondingly,  $\mathbb{E}^{\mathbb{P}} \left[ e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right]$  is the best estimate  $t$ -year survival probability of an  $x$ -year-old in calendar year  $h$  under the real world measure  $\mathbb{P}$ . In addition, we assume that at the moment when the insurance contract is initiated, the insurer prices the retirement plan by some chosen risk-neutral probability measure  $\mathbb{Q}$ , which takes both unsystematic and systematic mortality risk into account. In order to be prudent, it is reasonable to assume the survival probabilities under the pricing measure  $\mathbb{Q}$  are larger than under the real world measure  $\mathbb{P}$ , i.e.  $\mathbb{E}^{\mathbb{P}} \left[ e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right] < \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right]$  for all  $t > 0$ .<sup>1</sup> Furthermore, throughout this paper, we assume the absence of financial market risks to focus exclusively on mortality risks.

The tonuity is originally introduced by Chen et al. (2019), which works as a tontine before a fixed switching time ( $\tau$ ) and becomes an annuity after  $\tau$ . As shown by Chen et al. (2020) in an expected lifetime utility framework, the optimal way to invest in tontines and annuities is to put part of the wealth in tontines and annuities respectively. In particular, they show that the optimal tonuity performs rather well and delivers just a slightly lower utility level than the optimal portfolio of annuities and tontines. However, the fixed switching time  $\tau$  may not be desirable when some unexpected mortality shock (e.g., the Covid-19 pandemic) takes place, resulting in significant deviations of survival probabilities from the assumptions used to determine  $\tau$  at the beginning. We will continue the discussion in Subsection 2.2, before which we shortly review the construction of the regular tonuity.

Mathematically, the payoff of the regular tonuity can be expressed as

$$b_{[\tau]}(t) = \mathbb{1}_{\{0 \leq t < \min\{\tau, T^{(b)}\}\}} \frac{n}{N^{(b)}(t)} d_{[\tau]}(t) + \mathbb{1}_{\{\tau \leq t < T^{(b)}\}} c_{[\tau]}(t), \quad (2.1)$$

where the tontine-specific payoff  $d_{[\tau]}(t)$  and annuity-specific payoff  $c_{[\tau]}(t)$  are determined at the beginning of the contract.  $\mathbb{1}_B$  denotes an indicator function which equals 1 if  $B$

<sup>1</sup>In the following contents, all the notations under  $\mathbb{Q}$  are marked with a superscript  $\mathbb{Q}$ .

occurs and zero otherwise, and  $N^{(b)}(t)$  denotes the number of policyholders in the book population alive after  $t$  years. In order to make various settings mentioned in the paper comparable with each other, we rely on the risk-neutral pricing approach, different from Chen et al. (2019, 2020). More specifically, insurers determine premiums as the expected present values of future benefits under some risk-neutral probability measure  $\mathbb{Q}$ . To be precise, the premium under  $\mathbb{Q}$  can be written as

$$P_0^{(RT, \mathbb{Q})} = \int_0^\tau e^{-r_f t} \mathbb{E}^\mathbb{Q} \left[ 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right)^n \right] d_{[\tau]}(t) dt + \int_\tau^\infty e^{-r_f t} \mathbb{E}^\mathbb{Q} \left[ e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right] c_{[\tau]}(t) dt, \quad (2.2)$$

where  $r_f$  is a risk-free interest rate.

To evaluate the payoffs, the policyholders are assumed to be risk-averse towards the uncertain lifetime. With reference to Yaari (1965), the policyholder's expected discounted lifetime utility is introduced as

$$U(\{b(t)\}_{t \geq 0}) := \mathbb{E}^\mathbb{P} \left[ \int_0^\infty e^{-\eta t} u(b(t)) \mathbb{1}_{\{T^{(b)} > t\}} dt \right], \quad (2.3)$$

where  $\eta$  denotes the subjective discount rate of the policyholder,  $u(\cdot)$  is a utility function and  $b(t)$  represents the product payoff to a living policyholder. In this paper, we adopt the constant relative risk aversion (CRRA) utility function, which is typically used in the life cycle literature (e.g., Gomes et al., 2008; Cocco and Gomes, 2012; Hubener et al., 2016; Kim et al., 2016). That is, we assume

$$u(y) = \frac{y^{1-\gamma}}{1-\gamma}, \quad (2.4)$$

with a payoff stream  $y \geq 0$ , and a constant relative risk aversion coefficient  $\gamma \in (0, \infty)$  adhering to  $\gamma \neq 1$ . By incorporating the characteristics of the regular tonuity into Equation (2.3), the policyholder chooses the optimal tonuity by addressing the optimization problem below:

$$\begin{aligned} \max_{d_{[\tau]}(t), c_{[\tau]}(t)} \quad & \mathbb{E}^\mathbb{P} \left[ \int_0^\infty e^{-\eta t} \left( \mathbb{1}_{\{0 \leq t < \min\{\tau, T^{(b)}\}\}} u\left(\frac{nd_{[\tau]}(t)}{N^{(b)}(t)}\right) + \mathbb{1}_{\{\tau \leq t < T^{(b)}\}} u(c_{[\tau]}(t)) \right) dt \right] \\ \text{s.t.} \quad & P_0^{(RT, \mathbb{Q})} = v. \end{aligned} \quad (2.5)$$

Theorem 2.1 presents the solution to the optimization problem (2.5).

**Theorem 2.1** *The solution to problem (2.5) is given by*

$$d_{[\tau]}^*(t) = \left( \lambda_{[\tau]} e^{(\eta - r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right)^n \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathcal{K}_{n, \gamma} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right) \right]} \right)^{-\frac{1}{\gamma}}, \quad t \in [0, \tau), \quad (2.6)$$

$$c_{[\tau]}^*(t) = \left( \lambda_{[\tau]} e^{(\eta - r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right]}{\mathbb{E}^{\mathbb{P}} \left[ e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right]} \right)^{-\frac{1}{\gamma}}, \quad t \in [\tau, \infty). \quad (2.7)$$

*The optimal Lagrangian multiplier is given by*

$$\begin{aligned} \lambda_{[\tau]} = & \left( \frac{1}{v} \int_0^\tau e^{(\frac{r_f - \eta}{\gamma} - r_f)t} \frac{\mathbb{E}^{\mathbb{P}} \left[ \mathcal{K}_{n, \gamma} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right) \right]}{\mathbb{E}^{\mathbb{Q}} \left[ 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right)^n \right]^{\frac{1}{\gamma} - 1}} dt \right. \\ & \left. + \frac{1}{v} \int_\tau^\infty e^{(\frac{r_f - \eta}{\gamma} - r_f)t} \frac{\mathbb{E}^{\mathbb{P}} \left[ e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right]^{\frac{1}{\gamma}}}{\mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right]^{\frac{1}{\gamma} - 1}} dt \right)^\gamma, \end{aligned}$$

where

$$\mathcal{K}_{n, \gamma} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right) := \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right)^k \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right)^{n-k}.$$

*Then the policyholder's optimal expected utility is*

$$EU_{[\tau]} = \frac{\lambda_{[\tau]}}{1 - \gamma} \cdot v.$$

Proof: See Appendix B.1.

Note that Theorem 2.1 presents only the optimal payout functions for a given switching time. As pointed out in Chen et al. (2019, 2020), it is not possible to derive an explicit representation for the optimal switching time, but we can still rely on the numerical way to determine it. The idea is to calculate the optimal payoffs  $d_{[\tau]}^*(t)$  and  $c_{[\tau]}^*(t)$ , along with determining the expected utility for each potential switching time ( $\tau$ ). Subsequently, the switching time that yields the highest expected utility is deemed optimal, denoted as  $\tau^*$ .

## Numerical Realization

We apply the Age-Period-Cohort model with reference to (Currie, 2006) to estimate and forecast the mortality rates for our book population under the real world measure  $\mathbb{P}$ . To focus on our main research question, the detailed approach and related data used for estimation and projection are placed in the Appendixes A.2 and A.5. To develop the forces of mortality  $\mu_{x,h}^{(b,\mathbb{Q})}$  under the risk neutral measure  $\mathbb{Q}$ , we add the longevity risk premiums to the mortality trends (e.g., Cairns et al., 2006; Li, 2018). The longevity risk premiums for  $\mu_{x,h}^{(b,\mathbb{Q})}$  are further calibrated with reference to the Sharpe ratio approach (e.g., Milevsky et al., 2006; Chen et al., 2022a), where the Sharpe ratio is defined as the difference between the expected present value of the total annuity payment under  $\mathbb{Q}$  and that under  $\mathbb{P}$  divided by the standard deviation of the total annuity payment's present value under  $\mathbb{P}$ . With a given value of the Sharpe ratio, the mortality rates under  $\mathbb{Q}$  can be obtained. More details can be found in Appendixes A.3 and A.6.

To obtain the optimal payoffs with respect to different switching point  $\tau$  as described in Theorem 2.1, we need to make some additional assumptions on parameters. Our model is calibrated such that a plausible set of benchmark parameters is obtained. Let us consider that 1000 identical male policyholders join the regular tonuity plan at the beginning of 2019 (at age 65). The pool size of 1000 is based on the findings of Qiao and Sherris (2013), where a minimum pool size of 1000 is recommended to be used in group self-annuitization schemes. Each of them invests their initial wealth of 10000 to this plan, and is assumed to have a risk aversion coefficient of  $\gamma = 8$ . Since different values of  $\gamma$  are used in the literature (e.g., Blake et al., 2003; Steinorth and Mitchell, 2015; Chen et al., 2019), we will carry out sensitivity analyses with respect to different  $\gamma$ , and choose a relatively high value as a benchmark for illustrative purposes. The risk-free rate is chosen to be 2% according to the average annual interest rate of the United Kingdom from 2019 to 2022 (Bank of England, 2023). The subjective discount rate  $\eta$  is assumed to be equal to the risk-free rate. In particular, this choice results in an annual subjective discount factor in the benchmark being  $e^{-0.02} \approx 0.98$ , which is equal to the value chosen in Cocco and Gomes (2012).<sup>2</sup> Furthermore, as mentioned above, we use the Sharpe ratio approach to calibrate probabilities under  $\mathbb{Q}$ . Referring to Bauer et al. (2010), who have found that the Sharpe ratio of the UK pension annuities is between 4% and 12%, we set the Sharpe ratio as 8% and different values will be examined later in sensitivity analyses. Applying the risk neutral pricing measure  $\mathbb{Q}$  to the optimal regular tonuity, the resulting safety loading for the optimal tontine (when  $\tau \rightarrow \infty$ ) and for the

<sup>2</sup>We have also examined the cases that  $\eta < r_f$  and  $\eta > r_f$  while keeping the other parameters the same as in Table 2.1, and we find that the effect from varying  $\eta$  is only moderate.

optimal annuity (when  $\tau = 0$ ) are 0.019%, and 1.31% respectively. Here we see that, using a prudent life table, the optimal tontine requires a much lower safety loading than the optimal annuity, which implies that the optimal tontine contains less longevity risk from the insurer's perspective. The assumed parameters are summarized in Table 2.1.

Variable	Notation	Value
Initial wealth	$v$	10000
Pool size	$n$	1000
Risk-free rate	$r_f$	0.02
Subjective discount rate	$\eta$	0.02
Risk aversion coefficient	$\gamma$	8
Sharpe ratio	-	8%
Initial age	$x$	65
Calendar year	$h$	2019

Table 2.1: Benchmark parameters.

Applying the projected mortality rates and assumed parameters provided in Table 2.1, we numerically find the optimal switching point that maximizes the policyholder's expected utility,  $\tau^* = 15$ , i.e. when the policyholder is at 80.

## 2.2 Disadvantage(s) of A Regular Tonuity

In a typical regular tonuity arrangement, the switching time  $\tau$  (from tontine to annuity) is fixed at contract initiation and is determined based on the mortality assumptions established at the outset. Nevertheless, unexpected mortality shocks (such as the Covid-19 pandemic), or unexpected longevity shocks (e.g., a medical treatment breakthrough) can significantly impact human lives. If the tonuity contract is initiated shortly before the onset of the shock, the predetermined switching time may no longer be ideal for the policyholders. Below, we analyze different cases to determine the utility losses experienced by individual policyholders when they enter into a standard regular tonuity contract, particularly in the presence of unforeseen mortality/longevity shocks.

We specify the following cases:

- (i) **The base case**, where mortality assumptions do not change and there is no re-assessment of the contract. The contract is optimized at the beginning. Meanwhile,

the optimal switching time ( $\tau^*$ ) as well as the corresponding payoff streams ( $d_{[\tau^*]}^*(t)$  and  $c_{[\tau^*]}^*(t)$ ) are determined and fixed.

(ii) **The stress case**, when an unforeseen mortality or longevity shock occurs shortly after the contract is made (at time  $t = 0$ ). Below, we elaborate on the construction of the stress case.

(ii-M) **The stress case (M)**. As we know that the pandemic takes away more lives than expected, especially the elderly (Yanez et al., 2020), mortality rates of elderly policyholders are expected to increase. In the process of generating mortality rates for the base case and the stress case, we integrate pandemic data into the mortality rates applied in the stress case. In contrast, for the base case, we assume mortality rates to evolve based on the initial mortality assumptions, which do not account for the influence of the pandemic. To clarify, the stressed mortality rates are estimated and projected using a dataset encompassing calendar years that include some pandemic years; while the base case mortality rates are derived from a subset of data that excludes the years affected by the pandemic. More details can be found in Appendix A.5.

(ii-L) **The stress case (L)**. In contrast to the unexpected mortality shock, we create a stress case in terms of an unexpected longevity shock. Due to data limitation, we are not able to create such a case using real mortality data. Instead, we generate a fictitious case by introducing a safety buffer to the mortality trends derived from the base-case mortality model (calibrated using the subset of data excluding pandemic years). Given that mortality improvements tend to occur gradually, driven by factors such as medical advancements, the unexpected longevity shock is constructed to be less pronounced compared to the unforeseen mortality shock. Refer to the Appendix A.5 for details.

Figure 2.1 depicts the expected  $t$ -year survival probabilities for the book population in (i), (ii-M), and (ii-L). As explained earlier, the expected  $t$ -year book survival probabilities in the stress case (M) are lower than those in the base case. Not surprisingly, the survival probabilities under the stress case (L) are the highest among these three cases.

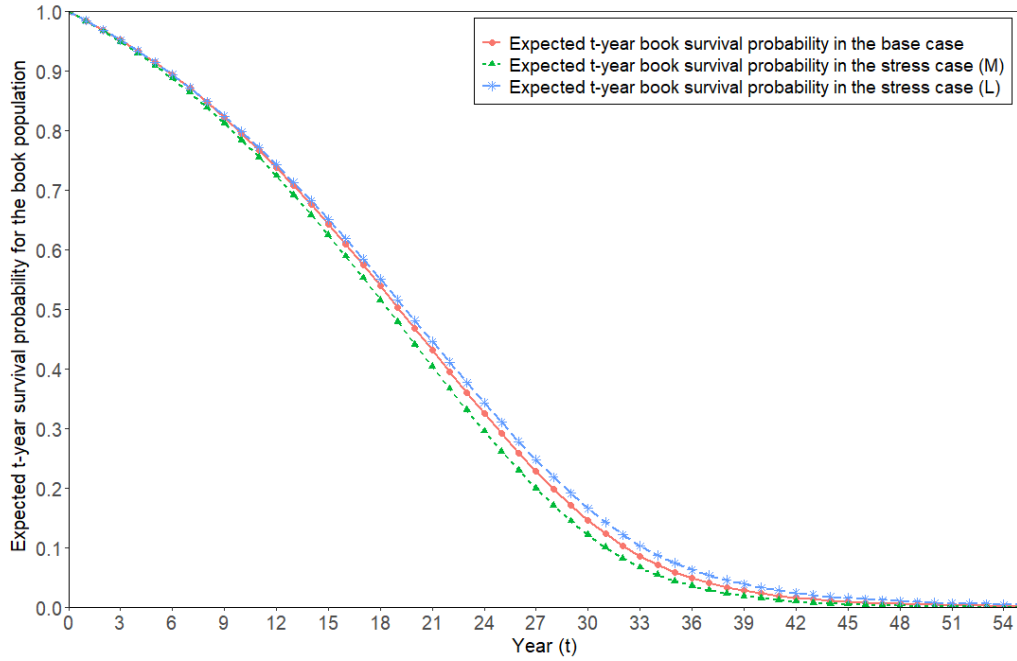


Figure 2.1: Expected  $t$ -year survival probabilities for the book population in different cases. The red, green and blue curve represents the expected  $t$ -year book survival probabilities in the base case, stress case (M) and stress case (L).

**Expected utility of the original contract in the stress case:** Knowing the (mortality or longevity) shock at time  $t = 0$ , the policyholders reassess their original contract (set up under the base-case assumptions) immediately, where the original contract suggests switching at  $\tau^*$  and the corresponding payoff streams are given by  $d_{[\tau^*]}^*(t)$  and  $c_{[\tau^*]}^*(t)$ . Then, an individual's expected utility is given by:

$$EU_{\{bookstr, \tau^*\}} = \mathbb{E}^{\mathbb{P}} \left[ \int_0^{\tau^*} \mathbb{1}_{\{\hat{T}^{(b)} > t\}} e^{-\eta t} u \left( \frac{nd_{[\tau^*]}^*}{\hat{N}^{(b)}(t)} \right) dt \right] + \mathbb{E}^{\mathbb{P}} \left[ \int_{\tau^*}^{\infty} \mathbb{1}_{\{\hat{T}^{(b)} > t\}} e^{-\eta t} u \left( c_{[\tau^*]}^*(t) \right) dt \right], \quad (2.8)$$

where  $\hat{N}^{(b)}(t)$  is the number of pool members in case (ii-M or ii-L), and  $\hat{T}^{(b)}$  is the remaining lifetime after the unforeseen mortality/longevity shock.

**Expected utility of the re-optimized contract in the stress case:** Given the stress cases (ii-M) and (ii-L) described above, where mortality rates increase in case of an unexpected mortality shock (or decreases with an unexpected longevity shock), at time

$t = 0$ , policyholders reassess their contract immediately. Unlike (2.8), they now alter their payoffs and the switching time according to the shock. The recomputed optimal switching time  $\hat{\tau}$  is given by maximizing the expected utility of the policyholder given the information of the shock. And we mark the corresponding optimal payoff streams as  $\hat{d}_{[\hat{\tau}]}(t)$  and  $\hat{c}_{[\hat{\tau}]}(t)$ . Then the expected utility can be calculated by:

$$EU_{\{bookreopt, \hat{\tau}\}} = \mathbb{E}^{\mathbb{P}} \left[ \int_0^{\hat{\tau}} \mathbb{1}_{\{\hat{T}^{(b)} > t\}} e^{-\eta t} u \left( \frac{n\hat{d}_{[\hat{\tau}]}}{\hat{N}^{(b)}(t)} \right) dt \right] + \mathbb{E}^{\mathbb{P}} \left[ \int_{\hat{\tau}}^{\infty} \mathbb{1}_{\{\hat{T}^{(b)} > t\}} e^{-\eta t} u \left( \hat{c}_{[\hat{\tau}]}(t) \right) dt \right]. \quad (2.9)$$

### Numerical Example

To keep the example straightforward as well as simple, we assume that the unexpected mortality/longevity shock takes place immediately after the contract is made, and is instantly noticed by the policyholder, i.e. at time 0. Applying the parameters from Table 2.1, we find that the re-optimized switching time ( $\hat{\tau}$ ) with an unforeseen mortality shock such as the Covid-19 pandemic, is 32 years after the age 65. In other words, the policyholder optimally switches from tontine-like payoffs to annuity-like payoffs at age 97, instead of 80 (from the base case). In addition, the re-optimized switching time with an unforeseen longevity shock is at age 66. The reason for this is that when an unforeseen mortality/longevity shock occurs, resulting in mortality worsening/improving, the tontine/annuity component of the regular tonuity will be preferable.<sup>3</sup>

Additionally, we calculate the expected utilities using equations (2.8) and (2.9) as outlined earlier. It is observed that in the event of an unforeseen Covid-19 pandemic occurring shortly after the contract's effective date, compared to the one who promptly re-optimizes the contract following the change in circumstances, the policyholder adhering to the original contract experiences a utility loss (i.e., 8.298% of the expected utility (2.9) concerning a re-optimized contract in case (ii-M)). In the stress case with an unforeseen longevity shock, the utility loss is calculated to be 1.424% of the expected utility (2.9) relative to a re-optimized contract in case (ii-L). Because we intentionally design the unexpected longevity shock to be moderate, accounting for the gradual improvement of mortality rates, the percentage of utility loss in case (ii-L) is comparatively modest.

---

<sup>3</sup>The expected utility of the tontine increases/declines if death probabilities in the pool are higher/lower, as reported, e.g., by Chen et al. (2021). Therefore, the optimal switch happens later ( $\hat{\tau} > \tau^*$ ) or earlier ( $\hat{\tau} < \tau^*$ ).



### 3 Dynamic Tonuity

As mentioned before, the regular tonuity has some drawbacks. On the one hand, the predetermined “optimal” switching time for the regular tonuity, established using initial information, may not be advantageous for the policyholder when faced with evolving circumstances. On the other hand, some policyholders may find it challenging to determine the optimal switching time. To address these problems, we propose a dynamic tonuity, which relates the insurer’s portfolio (i.e. the book population) to some “bigger” portfolio (i.e. the reference population). In this case, the switching time is not fixed at the beginning. Instead, the product switches automatically from tontine to annuity after the event that triggers the switch concerning the reference population takes place. In the following, we will define the switching criterion and payoff structure, calculate the premium and determine the optimal payoff streams of the dynamic tonuity.

#### 3.1 The Dynamic Switching Time and Payoff Structure

As mentioned before, the proposed dynamic tonuity involves two populations. Consistent with notations of the regular tonuity, we use  $T^{(b)}$  to mark the random remaining lifetime, and  $\mu_{x,h}^{(b)}$  to denote the stochastic forces of mortality for the insurer’s portfolio, i.e. the book population. Additionally, we introduce  $T^{(r)}$  and  $\mu_{x,h}^{(r)}$  for the reference population to set up the automatic switching criterion. In particular,  $\mu_{x,h}^{(b)}$  and  $\mu_{x,h}^{(r)}$  are stochastic and correlated.

In our design of the dynamic tonuity, the switch from tontine to annuity occurs when the one-year survival probability among the reference population goes over a certain threshold index. Here, we use  $e^{-\int_0^1 \mu_{x+t+s,h+t+s}^{(r)} ds}$  to represent the one-year realized survival curve of an  $(x+t)$ -year-old policyholder of the cohort  $h-x$  among the reference population. The retirement payments of the dynamic tonuity switch at the random switching time

$$\Lambda := \inf \left\{ t \geq 0 : e^{-\int_0^1 \mu_{x+t+s,h+t+s}^{(r)} ds} > p_{x+t,h+t}^{thr} \right\}, \quad (3.1)$$

where  $\{p_{x+t,h+t}^{thr}\}_{t \geq 0}$  is a threshold vector. It can be set based on, e.g., the latest officially published survival probabilities. This switching condition is fully observable to the policyholders. When the survival probability of the reference population exceeds the predetermined threshold  $\{p_{x+t,h+t}^{thr}\}_{t \geq 0}$  for the first time, the payment structure transits from tontine-like to annuity-like payments.

When varying the threshold values for  $\{p_{x+t,h+t}^{thr}\}_{t \geq 0}$ , the switching time  $\Lambda$  can have values of 0,  $\infty$ , or other sensible options. Establishing an exceedingly high threshold makes it almost unattainable, thereby converting the dynamic tonuity into a tontine-like arrangement. Conversely, an excessively low threshold triggers the switch right from the beginning, transforming the dynamic tonuity into an annuity-like arrangement. It is of vital importance to make a judicious choice of the threshold to ensure the emergence of reasonable values for  $\Lambda$ . Numerical illustrations on the dynamic switching time can be found in Subsection 4.1.

Note that the random variable  $\Lambda$  represents the realized switching time in the future. Standing at present, we can only make predictions to get possible realizations of  $\Lambda$ . Subsequently, we can write down the random payoff of the dynamic tonuity given a realized switching time  $\Lambda$  as follows:

$$\tilde{b}_{[\Lambda]}(t) = \mathbb{1}_{\{0 \leq t < \min\{\Lambda, T^{(b)}\}\}} \frac{n}{N^{(b)}(t)} d_{[\bar{\Lambda}]}(t) + \mathbb{1}_{\{\Lambda \leq t < T^{(b)}\}} c_{[\bar{\Lambda}]}(t), \quad (3.2)$$

where  $d_{[\bar{\Lambda}]}(t)$  and  $c_{[\bar{\Lambda}]}(t)$  respectively represent the deterministic tontine-like payoff and the annuity-like payoff, which are determined by maximizing the policyholder's expected lifetime utility. By this way, these payoffs depend on expectations of functions of  $\Lambda$  (cf. Theorem 3.1).

### 3.2 Premium Calculation and Optimal Payoff Streams

Similar to the regular tonuity, we obtain the premium of the dynamic tonuity as the expected present value under the risk-neutral measure  $\mathbb{Q}$ . More specifically, let us define  $\mu_{x+t,h+t}^{(\mathbb{Q})} := (\mu_{x+t,h+t}^{(b,\mathbb{Q})}, \mu_{x+t,h+t}^{(r,\mathbb{Q})})$ , and the realized switching time under  $\mathbb{Q}$  is given by

$$\Lambda^{\mathbb{Q}} := \inf \left\{ t \geq 0 : e^{-\int_0^t \mu_{x+s,h+s}^{(r,\mathbb{Q})} ds} > p_{x+t,h+t}^{thr} \right\}.$$

For the risk-neutral pricing, the first-hitting time which triggers the switch in the dynamic tonuity will be determined based on this measure. Accordingly, we obtain the risk-neutral price of the dynamic tonuity as follows:

$$\begin{aligned} P_0^{(DT,\mathbb{Q})} &= \mathbb{E}^{\mathbb{Q}} \left[ \int_0^\infty e^{-r_f t} \tilde{b}_{[\Lambda]}(t) dt \right] \\ &= \mathbb{E}^{\mathbb{Q}} \left[ \int_0^\infty e^{-r_f t} \left( \mathbb{1}_{\{0 \leq t < \min\{\Lambda^{\mathbb{Q}}, T^{(b,\mathbb{Q})}\}\}} \frac{n d_{[\bar{\Lambda}]}(t)}{N^{(b,\mathbb{Q})}(t)} + \mathbb{1}_{\{\Lambda^{\mathbb{Q}} \leq t < T^{(b,\mathbb{Q})}\}} c_{[\bar{\Lambda}]}(t) \right) dt \right] \end{aligned}$$

$$\begin{aligned}
&= \int_0^\infty e^{-r_f t} \mathbb{E}^\mathbb{Q} \left[ \mathbb{E}^\mathbb{Q} \left[ \mathbb{1}_{\{0 \leq t < \Lambda^\mathbb{Q}\}} \mathbb{1}_{\{0 \leq t < T^{(b, \mathbb{Q})}\}} \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b, \mathbb{Q})}(t)} \right. \right. \\
&\quad \left. \left. + \mathbb{1}_{\{\Lambda^\mathbb{Q} \leq t\}} \mathbb{1}_{\{t < T^{(b, \mathbb{Q})}\}} c_{[\bar{\Lambda}]}(t) \mid \{\mu_{x+t, h+t}^{(\mathbb{Q})}\}_{t \geq 0} \right] \right] dt \\
&= \int_0^\infty e^{-r_f t} \mathbb{E}^\mathbb{Q} \left[ \underbrace{\mathbb{1}_{\{0 \leq t < \Lambda^\mathbb{Q}\}} \mathbb{E}^\mathbb{Q} \left[ \mathbb{1}_{\{0 \leq t < T^{(b, \mathbb{Q})}\}} \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b, \mathbb{Q})}(t)} \mid \{\mu_{x+t, h+t}^{(\mathbb{Q})}\}_{t \geq 0} \right]}_{\mathcal{A}} \right] dt \\
&\quad + \int_0^\infty e^{-r_f t} \mathbb{E}^\mathbb{Q} \left[ \underbrace{\mathbb{1}_{\{\Lambda^\mathbb{Q} \leq t\}} \mathbb{E}^\mathbb{Q} \left[ \mathbb{1}_{\{t < T^{(b, \mathbb{Q})}\}} c_{[\bar{\Lambda}]}(t) \mid \{\mu_{x+t, h+t}^{(\mathbb{Q})}\}_{t \geq 0} \right]}_{\mathcal{B}} \right] dt.
\end{aligned}$$

Note that the premium can be split into a tontine and an annuity part. The first conditional expectation  $\mathcal{A}$  is given by

$$\begin{aligned}
&\mathbb{E}^\mathbb{Q} \left[ \mathbb{1}_{\{0 \leq t < T^{(b, \mathbb{Q})}\}} \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b, \mathbb{Q})}(t)} \mid \{\mu_{x+t, h+t}^{(\mathbb{Q})}\}_{t \geq 0} \right] \\
&= e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \cdot \mathbb{E}^\mathbb{Q} \left[ \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b, \mathbb{Q})}(t)} \mid \{\mu_{x+t, h+t}^{(\mathbb{Q})}\}_{t \geq 0}, T^{(b, \mathbb{Q})} > t \right] \\
&= e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \cdot d_{[\bar{\Lambda}]}(t) \sum_{k=0}^{n-1} \frac{n}{k+1} \binom{n-1}{k} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right)^k \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right)^{n-1-k} \\
&= d_{[\bar{\Lambda}]}(t) \sum_{k=0}^{n-1} \binom{n}{k+1} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right)^{k+1} \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right)^{n-1-k} \\
&= d_{[\bar{\Lambda}]}(t) \sum_{k=1}^n \binom{n}{k} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right)^k \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right)^{n-k} \\
&= d_{[\bar{\Lambda}]}(t) \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right)^n \right).
\end{aligned}$$

The second conditional expectation  $\mathcal{B}$  is given by

$$\mathbb{E}^\mathbb{Q} \left[ \mathbb{1}_{\{t < T^{(b, \mathbb{Q})}\}} c_{[\bar{\Lambda}]}(t) \mid \{\mu_{x+t, h+t}^{(\mathbb{Q})}\}_{t \geq 0} \right] = c_{[\bar{\Lambda}]}(t) \mathbb{E}^\mathbb{Q} \left[ \mathbb{1}_{\{t < T^{(b, \mathbb{Q})}\}} \mid \{\mu_{x+t, h+t}^{(\mathbb{Q})}\}_{t \geq 0} \right] = c_{[\bar{\Lambda}]}(t) e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds}.$$

Putting everything together, we obtain the following expression for the premium of the

dynamic tonuity under  $\mathbb{Q}$ :

$$\begin{aligned} P_0^{(DT, \mathbb{Q})} &= \int_0^\infty e^{-r_f t} \mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda^{\mathbb{Q}}\}} d_{[\bar{\Lambda}]}(t) \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds \right)^n \right) \right] dt \\ &\quad + \int_0^\infty e^{-r_f t} \mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{\Lambda^{\mathbb{Q}} \leq t\}} c_{[\bar{\Lambda}]}(t) e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right] dt. \end{aligned} \quad (3.3)$$

Depending on the precise structure of  $\mu_{x+t, h+t}^{(b, \mathbb{Q})}$  and  $\mu_{x+t, h+t}^{(r, \mathbb{Q})}$ , the expression may be further explored. However, due to the fact that mortality rates of the book and reference population are correlated and based on our two-population stochastic mortality model (placed in Appendix A to avoid distraction), we are unable to simplify (3.3) further at this point.

The policyholder's objective function in the framework of the dynamic tonuity can then be translated to the following mathematical expression:

$$\begin{aligned} \max_{d_{[\bar{\Lambda}]}(t), c_{[\bar{\Lambda}]}(t)} \quad & \mathbb{E}^{\mathbb{P}} \left[ \int_0^\infty e^{-\eta t} \left( \mathbb{1}_{\{0 \leq t < \min\{\Lambda, T^{(b)}\}} u \left( \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b)}(t)} \right) + \mathbb{1}_{\{\Lambda \leq t < T^{(b)}\}} u(c_{[\bar{\Lambda}]}(t)) \right) dt \right] \\ \text{s.t.} \quad & P_0^{(DT, \mathbb{Q})} = v. \end{aligned} \quad (3.4)$$

**Theorem 3.1** *The solution to problem (3.4) is given by*

$$d_{[\bar{\Lambda}]}^*(t) = \left( \lambda_{[\bar{\Lambda}]} e^{(\eta - r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda^{\mathbb{Q}}\}} \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds \right)^n \right) \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n, \gamma} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right) \right]} \right)^{-\frac{1}{\gamma}}, \quad (3.5)$$

$$c_{[\bar{\Lambda}]}^*(t) = \left( \lambda_{[\bar{\Lambda}]} e^{(\eta - r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{\Lambda^{\mathbb{Q}} \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right]} \right)^{-\frac{1}{\gamma}}, \quad (3.6)$$

where

$$\begin{aligned} \lambda_{[\bar{\Lambda}]} &= \left[ \frac{1}{v} \int_0^\infty e^{-r_f t} \left( e^{(\eta - r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda^{\mathbb{Q}}\}} \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds \right)^n \right) \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n, \gamma} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right) \right]} \right)^{-\frac{1}{\gamma}} \right. \\ &\quad \cdot \mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda^{\mathbb{Q}}\}} \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds \right)^n \right) \right] dt \\ &\quad \left. + \frac{1}{v} \int_0^\infty e^{-r_f t} \left( e^{(\eta - r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{\Lambda^{\mathbb{Q}} \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right]} \right)^{-\frac{1}{\gamma}} \cdot \mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{\Lambda^{\mathbb{Q}} \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right] dt \right]^\gamma. \end{aligned}$$

Here  $\mathcal{K}_{n,\gamma} \left( e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right)$  is defined as in Theorem 2.1. And the optimal level of expected utility is then given by

$$\frac{\lambda_{[\bar{\Lambda}]}}{1-\gamma} \cdot v.$$

Proof: See Appendix B.2.

## 4 Numerical Illustration

In this section, we present numerical examples to elucidate our proposed dynamic tonuity and compare it to the regular tonuity in the base and stress cases (as defined in Subsection 2.2). We start by introducing the thresholds for the dynamic switching condition, followed by the computation of optimal payoffs of the dynamic tonuity. Subsequently, we conduct a product comparison.

### 4.1 Threshold and Optimal Payoffs

For the dynamic tonuity, we assume that the switch from tontine to annuity is triggered, when the one-year survival probability of the reference population becomes higher than the threshold index. Note that the threshold index should be chosen carefully. It should make sure that, when the insurer forecasts the future possibilities at the beginning (to create the contract), there are chances that the dynamic switching time  $\Lambda$  could be equal to 0 and chances that the switching time approaches infinity. The switching times we mention here and in the following text are under the real world measure  $\mathbb{P}$ . In this paper, we take the 1954 cohort as an example (they turn 65 in 2019), and set the threshold index vector as the 0.99th power of the best-estimate one-year survival probability of the  $(65+t)$ -year-old policyholder in the reference population (i.e.  $p_{65+t,2019+t}^{thr} = \left( \mathbb{E} \left[ e^{-\int_0^1 \mu_{65+t+s,2019+t+s}^{(r)} ds} \right] \right)^{0.99}$ ).<sup>4</sup>

The switch occurs, when the simulated reference one-year survival probability at age 65 and older gets larger than the threshold index for the first time. To capture various possibilities in the future, we deploy Monte Carlo simulation methods to generate 10 000

<sup>4</sup>We choose 0.99 as an example for illustrative purposes, reflecting a minor buffer with respect to longevity risk of the realized survival probabilities compared to the threshold index.

mortality rate paths for both populations in the base case, where we denote the realized switching time by  $\Lambda_{\{base\}}$ . In Figure 4.1, we first exhibit three scenarios in the base case to help us better understand the switching mechanism of the dynamic tonuity. The three scenarios are:

- Scenario A: The dynamic tonuity switches from tontine to annuity at age 80 (i.e.  $\Lambda_{\{base\}} = 15$ ). In Panel A,  $e^{-\int_0^1 \mu_{65+t+s, 2019+t+s}^{(r)} ds}$  becomes higher than the threshold index in year 15, implying  $\Lambda_{\{base\}} = 15$ . This is comparable to the optimal switching time of the regular tonuity in the base case (cf. the numerical example presented in Section 2.1).
- Scenario B: The dynamic tonuity switches immediately after purchase (i.e.  $\Lambda_{\{base\}} = 0$ ). Panel B displays the scenario that  $e^{-\int_0^1 \mu_{65+t+s, 2019+t+s}^{(r)} ds}$  is greater than the threshold index at the beginning (at age 65), such that  $\Lambda_{\{base\}} = 0$ .
- Scenario C: The dynamic tonuity suggests no switching (i.e.  $\Lambda_{\{base\}} = \infty$ ). Panel C represents that  $e^{-\int_0^1 \mu_{65+t+s, 2019+t+s}^{(r)} ds}$  will never hit the threshold index, thus the dynamic tonuity becomes a tontine.

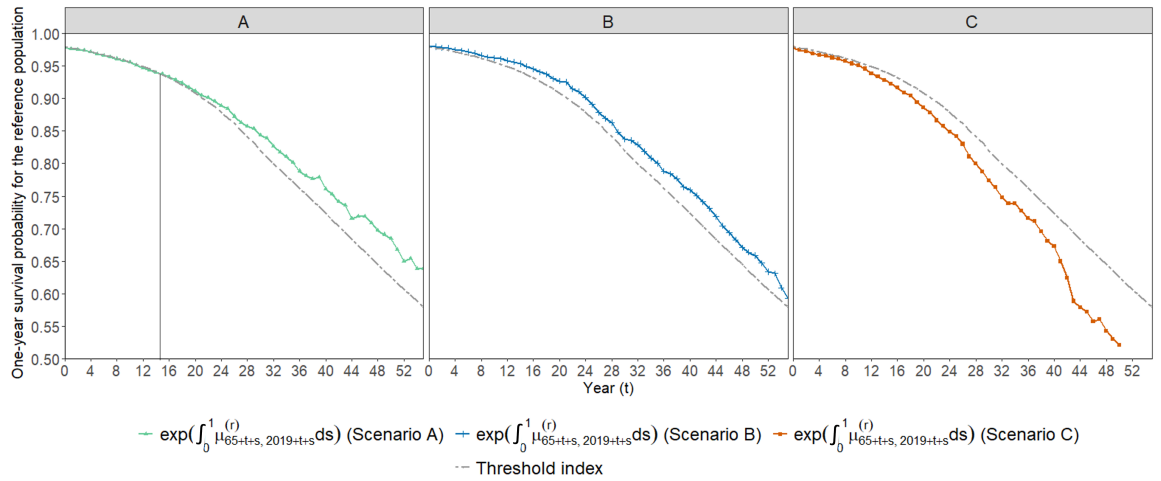


Figure 4.1: Selected paths of the reference one-year survival probability for Scenario A, B and C. The  $x$ -axis represents the year(s) after age 65.

Based on the parameters stated in Table 2.1, we illustrate optimal payoffs for the dynamic tonuity in Figure 4.2. As indicated by Theorem 3.1, the optimal payoffs of the dynamic tonuity are deterministic, which depend on expectations of functions of the dynamic

switching time. More specifically, for Scenario A, the tontine-like payoffs (the red solid curve) are delivered to the policyholder until the 14th year, and starting from year 15, the annuity-like payoffs are provided (the green dotted curve on the right-hand side of the vertical line in Figure 4.2); for Scenario B, since it suggests an immediate switching, the dynamic tonuity becomes an annuity and the payoffs are described completely by the green dotted curve; and for Scenario C, the dynamic tonuity remains as a tontine, such that the payoffs are represented by the red solid curve.

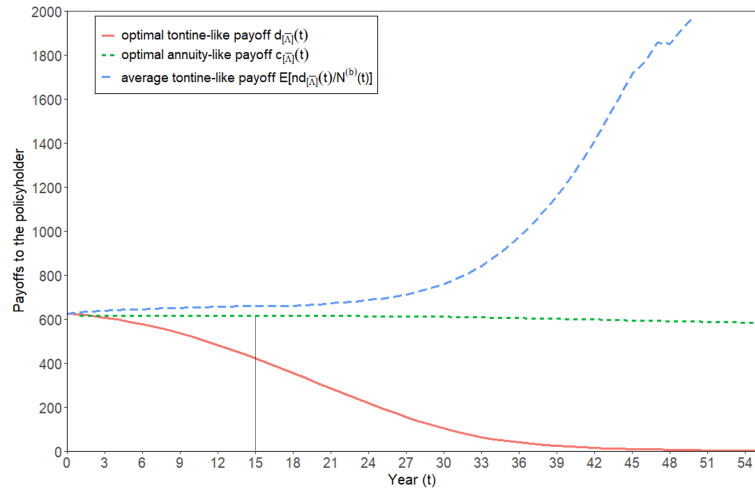


Figure 4.2: Optimal payoffs of the dynamic tonuity in the base case. The  $x$ -axis represents the year(s) after age 65.

## 4.2 Product Comparison

### Base case

To compare the optimal dynamic tonuity with the optimal regular tonuity, we compute the certainty equivalence quantity, which signifies the quantity of optimal dynamic tonuities required to achieve an equivalent expected utility as that produced by a single optimal regular tonuity. The certainty equivalence quantity in the base case is defined by:

$$\begin{aligned}
 EU_{\{bookbase, \tau^*\}} &= \int_0^\infty e^{-\eta t} \cdot \mathbb{E}^\mathbb{P} \left[ \mathbb{1}_{\{0 \leq t < \min\{\Lambda_{\{base\}}, T^{(b)}\}\}} u \left( CEQ_{\{dty, base\}} \cdot \frac{nd_{[\bar{\Lambda}_{\{base\}}]}^*(t)}{N^{(b)}(t)} \right) \right] dt \\
 &+ \int_0^\infty e^{-\eta t} \cdot \mathbb{E}^\mathbb{P} \left[ \mathbb{1}_{\{\Lambda_{\{base\}} \leq t < T^{(b)}\}} u \left( CEQ_{\{dty, base\}} \cdot c_{[\bar{\Lambda}_{\{base\}}]}^*(t) \right) \right] dt
 \end{aligned}$$

$$\begin{aligned}
&= CEQ_{\{dty, base\}}^{1-\gamma} \cdot EU_{\{dty, base\}}, \\
CEQ_{\{dty, base\}} &= \left( \frac{EU_{\{bookbase, \tau^*\}}}{EU_{\{dty, base\}}} \right)^{\frac{1}{1-\gamma}}.
\end{aligned} \tag{4.1}$$

Here,  $EU_{\{dty, base\}}$  and  $EU_{\{bookbase, \tau^*\}}$  represent the expected utilities of the optimal dynamic tonuity and the optimal regular tonuity in the base case respectively.

Under the base-case assumptions, we obtain  $CEQ_{\{dty, base\}} = 0.998979$ , which reflects that the expected utilities of the optimal dynamic tonuity and optimal regular tonuity are rather close to each other, and the former can be slightly more favorable than the latter. It is found that the optimal dynamic tonuity has a lower safety loading (0.604%) than the optimal regular tonuity (1.176%); while the payoffs provided by the optimal dynamic tonuity are more volatile over years compared to those of the optimal regular tonuity. We calculate the annual expected payoffs for these two optimal retirement products, and the volatility of the expected payoffs<sup>5</sup> is computed to be 30.440 for the optimal dynamic tonuity and 12.414 for the optimal regular tonuity.

### Stress case

Furthermore, we want to explore how the dynamic tonuity contract performs in a changing environment, where the mortality deteriorates significantly, or stated differently, the survival probability decreases substantially. To achieve this, we source to the “stress case” described in Subsection 2.2. In the stress case, we consider that when an unforeseen mortality/longevity shock strikes, the switching time of the dynamic tonuity may change (marked by  $\Lambda_{\{str(M)\}}$  and  $\Lambda_{\{str(L)\}}$ ). To demonstrate the importance of an automatically adjusting switching time, we assume the policyholders keep to the payoff streams in the original contract, i.e. the payoff streams for both the regular and the dynamic tonuity contract in the stress case remain the same as in the base case.<sup>6</sup> The tontine-like payoffs are delivered until the new switching time ( $\Lambda_{\{str(M)\}}$  or  $\Lambda_{\{str(L)\}}$ ), and then the annuity-like payoffs start to be provided. In other words, information on

<sup>5</sup>We compute the volatility of the expected payoffs as:  $\sqrt{\frac{\int_0^\zeta (\text{Epayoff}(t) - \overline{\text{Epayoff}}) dt}{\zeta}}$ , where  $\text{Epayoff}(t)$  means the expected payoff in year  $t$ ,  $\zeta$  is the maximum lifetime of a 65-year-old male person born in 1954, and  $\overline{\text{Epayoff}}$  represents the average value of expected payoffs from year 0 to  $\zeta$ . In our numerical analyses, we take  $\zeta = 55$ .

<sup>6</sup>Note that, in practice, upon noticing the shock, the policyholder might be able to change the switching time of the regular tonuity, as contract changes are a common practice in the insurance industry. However, it might be costly to change his contract due to possible fees required by the insurance company. Furthermore, the policyholder then still might not know what the optimal switching time in the future will be, especially as there might be other mortality shocks.



the unpredictable mortality shock is reflected in the new switching time, but not included in the payoffs.

Let us first have a look at the dynamic switching time in the stress case. Since we can only forecast future outcomes in the present moment, we simulate 10000 paths as well for the stress case resulting in 10000 possible values of the stressed switching time. In the stress case (M), as described by Table 4.1, compared to the base case, the chances of an immediate switching (i.e. switching at time 0) decrease, while the chances of no switching ( $\infty$ ) increase. This suggests that the unpredictable Covid-19 mortality shock makes the dynamic tonuity more likely to switch at a later time point, or remain as a tontine. Contrarily, in the stress case (L), immediate switching is more frequently recommended, whereas the suggestion of no switching is less common.

Chances / Extreme values	0	$\infty$
$\Lambda_{\{base\}}$	34.07%	10.25%
$\Lambda_{\{str(M)\}}$	0	46.19%
$\Lambda_{\{str(L)\}}$	39.27%	5.64%

Table 4.1: Chances of the extreme cases (the switching time being equal to 0 or  $\infty$ ) in 10000 simulations in the base and stress case. In 10000 base case simulations, an immediate switching ( $\Lambda_{\{base\}} = 0$ ) results, i.e. dynamic tonuity becomes an annuity) for 3407 times, and no switching is obtained for 1025 times ( $\Lambda_{\{base\}} \rightarrow \infty$ , i.e. dynamic tonuity becomes a tontine). Among the 10000 simulations for the stress case (M), none of the simulations recommends the immediate switching ( $\Lambda_{\{str(M)\}} = 0$ ), and 4619 simulations suggest no switching ( $\Lambda_{\{str(M)\}} \rightarrow \infty$ ). As for the stress case (L), 3927 simulations exhibit the immediate switching ( $\Lambda_{\{str(L)\}} = 0$ ), and 564 simulations show no switching ( $\Lambda_{\{str(L)\}} \rightarrow \infty$ ). All the reported switching times are under the real-world measure  $\mathbb{P}$ .

To make product comparison in the stress case, we refer to the terminologies the “dynamic tonuity contract” and “regular tonuity contract”, since the contracts are not optimal anymore now. Then the expected utility of the dynamic tonuity contract in the stress case can be specified as:

$$\begin{aligned}
EU_{\{dty, str\}} = & \int_0^\infty e^{-\eta t} \mathbb{E}^\mathbb{P} \left[ \mathbb{1}_{\{0 \leq t < \min\{\Lambda_{\{str\}}, \hat{T}^{(b)}\}\}} u \left( \frac{nd_{[\bar{\Lambda}_{\{base\}]}^*](t)}^*}{\hat{N}^{(b)}(t)} \right) \right] dt \\
& + \int_0^\infty e^{-\eta t} \mathbb{E}^\mathbb{P} \left[ \mathbb{1}_{\{\Lambda_{\{str\}} \leq t < \hat{T}^{(b)}\}} u \left( c_{[\bar{\Lambda}_{\{base\}]}^*}^*(t) \right) \right] dt.
\end{aligned}$$

Again, we compute the corresponding certainty equivalence quantity, i.e.

$$EU_{\{bookstr, \tau^*\}} = (CEQ_{\{dty, str\}})^{1-\gamma} EU_{\{dty, str\}}, \quad (4.2)$$

where  $EU_{\{bookstr, \tau^*\}}$  is the expected utility of the regular tonuity contract in the stress case, which is specified in Subsection 2.2. When  $CEQ_{\{dty, str\}} < 1$ , it suggests that the dynamic tonuity contract is more favorable than the regular tonuity contract in the stress case. Further, we use  $CEQ_{\{dty, str(M)\}}$  and  $CEQ_{\{dty, str(L)\}}$  to denote the certainty equivalence quantity in the stress cases with an unexpected mortality shock and an unforeseen longevity shock respectively.

Using the parameters given in Table 2.1, we find that in terms of an unexpected mortality shock, the value of  $CEQ_{\{dty, str(M)\}}$  approximates 0.973, and  $CEQ_{\{dty, str(L)\}}$  is around 0.993 in case of an unforeseen longevity shock. This indicates that the dynamic tonuity contract can be more attractive than the regular tonuity contract in the stress cases. The reason for this is the dynamic switch of the dynamic tonuity contract, which adjusts automatically to alterations in the reference survival probabilities in the stress case.

### 4.3 Sensitivity Analysis

In this subsection, we carry out sensitivity analysis with respect to the risk aversion coefficient  $\gamma$ , the pool size  $n$  and the Sharpe ratio that we used to calibrate the probabilities under the risk neutral measure  $\mathbb{Q}$ .

Our results in Table 4.2 show that a smaller pool size  $n$  and a larger risk aversion coefficient  $\gamma$  lead to an earlier optimal switching time of the regular tonuity, which agrees with the findings of Chen et al. (2019, 2020). As for the product comparison, we notice that the optimal dynamic tonuity can sometimes be slightly less favorable than the optimal regular tonuity in the base case. But overall, the policyholder's preference for the optimal dynamic tonuity and optimal regular tonuity are rather similar. For policyholders with different risk aversion coefficients  $\gamma$ , the optimal dynamic tonuity is more attractive to the less risk-averse ones than the optimal regular tonuity (reflected by a smaller  $CEQ_{\{dty, base\}}$ ). It is possibly because the less risk-averse policyholders can be more tolerant to the volatile payoffs than those who are more risk averse. With respect to the pool size  $n$ , an increased pool size tends to generate slightly higher expected payoffs associated with the optimal dynamic tonuity over the years, and thus, compared to the optimal regular tonuity, the preference for the optimal dynamic tonuity

intensifies. Furthermore, a higher Sharpe ratio enhances the attractiveness of the optimal dynamic tonuity, potentially attributed to the increased prominence of the safety loading advantage.

Parameter	Base case	
	Regular tonuity ( $\tau^*$ )	$CEQ_{\{dyt,base\}}$
Sensitivity analysis w.r.t. $\gamma$ ( $n = 1000$ , Sharpe ratio = 8%)		
$\gamma = 8$	15	0.998979
$\gamma = 6$	21	0.998398
$\gamma = 3$	33	0.997702
$\gamma = 0.5$	$\infty$	0.962485
$\gamma = 0.3$	$\infty$	0.918100
Sensitivity analysis w.r.t. $n$ ( $\gamma = 8$ , Sharpe ratio = 8%)		
$n = 100$	0	1.001012
$n = 800$	13	0.998981
$n = 1000$	15	0.998979
Sensitivity analysis w.r.t. the Sharpe ratio ( $\gamma = 8$ , $n = 1000$ )		
Sharpe ratio = 6%	0	1.000170
Sharpe ratio = 8%	15	0.998979
Sharpe ratio = 10%	20	0.998257

Table 4.2: Sensitivity analysis of the certainty equivalence quantity with respect to different parameters ( $n$ , the Sharpe ratio and  $\gamma$ ) in the base case.

As indicated by Table 4.3, regarding the stress cases, with other parameters given in Table 2.1, variations in pool sizes and Sharpe ratios do not alter the fact that the dynamic tonuity contract remains more appealing than the regular tonuity contract (i.e.,  $CEQ_{\{dyt,str(M)\}} < 1$  and  $CEQ_{\{dyt,str(L)\}} < 1$ ). Notably, we observe an interesting trend where a smaller pool size enhances the attractiveness of the dynamic tonuity contract over the regular one in the context of an unexpected mortality shock, whereas a larger pool size renders the dynamic tonuity contract more preferable than the regular one with an unanticipated longevity shock. This observation may be attributed to the impact of payoffs, which are determined under base-case assumptions. In our numerical analyses, as the pool size ( $n$ ) increases, the annuity-like payoffs of the dynamic tonuity become

more enticing. An unexpected longevity shock amplifies this attractiveness, given that it reduces  $\Lambda_{\{str(L)\}}$ , while an unexpected mortality shock leads to a larger  $\Lambda_{\{str(M)\}}$ . Consequently,  $CEQ_{\{dyt, str(L)\}}$  decreases with increasing  $n$ , whereas  $CEQ_{\{dyt, str(M)\}}$  increases with  $n$ .

Furthermore, for policyholders with  $\gamma = 8$  (see Table 2.1), it has been observed that a higher Sharpe ratio renders the dynamic tonuity contract more advantageous compared to the regular tonuity contract in the presence of unforeseen longevity or mortality shocks. A greater Sharpe ratio indicates that the contracts bear increased safety loadings within the same gross premium. The potential advantage of a reduced safety loading in the dynamic contract may be accentuated with a higher Sharpe ratio. We have also examined the sensitivity of results with respect to  $n$  and Sharpe ratio given  $\gamma \in (0, 1)$ . As  $n$  increases,  $CEQ_{\{dyt, str(M)\}}$  increases, while  $CEQ_{\{dyt, str(L)\}}$  decreases. Additionally, a higher Sharpe ratio leads to smaller  $CEQ_{\{dyt, str(M)\}}$  and  $CEQ_{\{dyt, str(L)\}}$ . These results align with our observations for policyholders with risk aversion coefficients greater than 1.

	Stress case (M)	Stress case (L)
Parameter	$CEQ_{\{dyt, str(M)\}}$	$CEQ_{\{dyt, str(L)\}}$
Sensitivity analysis w.r.t. $n$ ( $\gamma = 8$ , Sharpe ratio = 8%)		
$n = 100$	0.970157	0.996803
$n = 800$	0.971587	0.994024
$n = 1000$	0.973060	0.993343
Sensitivity analysis w.r.t. the Sharpe ratio ( $\gamma = 8$ , $n = 1000$ )		
Sharpe ratio = 6%	0.974622	0.996261
Sharpe ratio = 8%	0.973060	0.993343
Sharpe ratio = 10%	0.971724	0.989697

Table 4.3: Sensitivity analysis of the certainty equivalence quantity with respect to different parameters ( $n$  and the Sharpe ratio) in the stress case.

According to the results presented in Table 4.4, when an unexpected mortality shock strikes, the dynamic tonuity contract is particularly appealing to policyholders with lower risk aversion, indicating a greater willingness to accept fluctuating payoffs. This mostly holds for the unexpected longevity shock as well, however, an exception arises when the policyholder's risk aversion coefficient lies in the interval  $(0, 1)$  (e.g.,  $\gamma = 0.5$ ).

In such cases, the regular tonuity contract collapses to a tontine contract (due to an infinite switching time), and it surpasses the dynamic tonuity contract in attractiveness during a stress case involving an unforeseen longevity shock.

A plausible explanation for this anomaly is that policyholders with a risk aversion coefficient in the range of  $(0, 1)$  may prefer an optimal tontine initially, anticipating the entire premium allocation to be directed towards the tontine component. Contrarily, the dynamic tonuity contract takes into account the probability of various switching points. In essence, the switching time of the dynamic contract is not determined by policyholders' preferences but is externally contingent on the trigger condition. This may influence the distribution of the premium between the annuity and tontine components. In other words, although the dynamic tonuity contract may be advantageous because of its dynamic switching condition, its payoffs are determined under base-case assumptions. The unexpected longevity shock leads to a shift in dynamic switching time, occurring earlier than anticipated (i.e.  $\Lambda_{\{str(L)\}} < \Lambda_{\{base\}}$ ). This shift spans from the initial switching time ( $\Lambda_{\{base\}}$ ) to the stressed switching time ( $\Lambda_{\{str(L)\}}$ ). During this period, the original tontine-like payoffs are replaced with annuity-like payoffs, diminishing the appeal of the dynamic tonuity contract compared to the regular one.

	Stress case (M)	Stress case (L)
Parameter	$CEQ_{\{dyl, str(M)\}}$	$CEQ_{\{dyl, str(L)\}}$
Sensitivity analysis w.r.t. $\gamma$ ( $n = 1000$ , Sharpe ratio = 8%)		
$\gamma = 8$	0.973060	0.993343
$\gamma = 6$	0.968707	0.986717
$\gamma = 3$	0.941704	0.963698
$\gamma = 0.5$	0.607263	1.249141
$\gamma = 0.3$	0.419417	1.089043

Table 4.4: Sensitivity analysis of the certainty equivalence quantity with respect to  $\gamma$  in the stress case.

## 5 Conclusion

While the tonuity introduced by Chen et al. (2019) has turned out to be a sound retirement product combining the advantages of annuities and tontines (Chen et al., 2020),

it does exhibit some drawbacks in the light of an unexpected mortality shock (e.g., the recent Covid-19 pandemic) or an unanticipated longevity shock (e.g., medical advancements). Most importantly, the switching time has to be chosen at contract initiation, requiring policyholders to make a decision for the entire retirement phase based on information on mortality available at time 0. We suggest a modified version of the tonuity which contains an automatic switching time depending on the mortality experienced in a reference population, which we call dynamic tonuity. We derive the utility-maximizing payoff structures of this product for a general (random) switching time and demonstrate numerically how a properly designed dynamic tonuity can lead to higher levels of expected utility than a regular tonuity in a changing environment with mortality or longevity shocks.

While we disregard any kind of financial market risks in this paper, a question for future research could be whether unit-linked versions of tonuities might be attractive retirement plans. For example, we could combine the unit-linked tontine design introduced in Chen et al. (2022b) with some unit-linked annuity design, such as the one considered in Chen and Rach (2023), for example, to form a unit-linked tonuity. This unit-linked tonuity could then operate either with a fixed switching time (as in Chen et al. (2019)) or with a random switching time as in the present paper.

## Acknowledgments

We want to thank the participants of the MAF conference, taking place in Le Havre from April 4 to April 6, 2024, for fruitful discussions and comments that helped us improve earlier versions of this manuscript.

## References

- Bank of England (2023). Official bank rate history data from 1694. [https://www.bankofengland.co.uk/monetary-policy/the-interest-rate-bank-rate,\[2023-11-18\]](https://www.bankofengland.co.uk/monetary-policy/the-interest-rate-bank-rate,[2023-11-18]).
- Bauer, D., Börger, M., and Ruß, J. (2010). On the pricing of longevity-linked securities. *Insurance: Mathematics and Economics*, 46(1):139–149.
- Blake, D., Cairns, A. J., and Dowd, K. (2003). Pensionmetrics 2: Stochastic pension

- plan design during the distribution phase. Insurance: Mathematics and Economics, 33(1):29–47.
- Cairns, A. J., Blake, D., and Dowd, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. Journal of Risk and Insurance, 73(4):687–718.
- Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., and Khalaf-Allah, M. (2011a). Mortality density forecasts: An analysis of six stochastic mortality models. Insurance: Mathematics and Economics, 48(3):355–367.
- Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., and Balevich, I. (2009). A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. North American Actuarial Journal, 13(1):1–35.
- Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Khalaf-Allah, M., et al. (2011b). Bayesian stochastic mortality modelling for two populations. ASTIN Bulletin: The Journal of the IAA, 41(1):29.
- Chen, A., Hieber, P., and Klein, J. K. (2019). Tonuity: A novel individual-oriented retirement plan. ASTIN Bulletin: The Journal of the IAA, 49(1):5–30.
- Chen, A., Hieber, P., and Rach, M. (2021). Optimal retirement products under subjective mortality beliefs. Insurance: Mathematics and Economics, 101:55–69.
- Chen, A., Li, H., and Schultze, M. B. (2022a). Tail index-linked annuity: A longevity risk sharing retirement plan. Scandinavian Actuarial Journal, 2022(2):139–164.
- Chen, A., Nguyen, T., and Sehner, T. (2022b). Unit-linked tontine: Utility-based design, pricing and performance. Risks, 10(4):78.
- Chen, A. and Rach, M. (2019). Options on tontines: An innovative way of combining annuities and tontines. Insurance: Mathematics and Economics, 89:182–192.
- Chen, A. and Rach, M. (2023). Who chooses which retirement income? A CPT-based analysis. Review of Behavioral Economics, 10(3):203–227.
- Chen, A., Rach, M., and Sehner, T. (2020). On the optimal combination of annuities and tontines. ASTIN Bulletin: The Journal of the IAA, 50(1):95–129.

- Cocco, J. F. and Gomes, F. J. (2012). Longevity risk, retirement savings, and financial innovation. Journal of Financial Economics, 103(3):507–529.
- Currie, I. D. (2006). Smoothing and forecasting mortality rates with p-splines. Paper given at the Institute of Actuaries. <https://www.actuaries.org.uk/system/files/documents/pdf/currie.pdf>.
- Donnelly, C. and Young, J. (2017). Product options for enhanced retirement income. British Actuarial Journal, 22(3):636–656.
- Dowd, K., Cairns, A. J., Blake, D., Coughlan, G. D., and Khalaf-Allah, M. (2011). A gravity model of mortality rates for two related populations. North American Actuarial Journal, 15(2):334–356.
- Fullmer, R. K. and Sabin, M. J. (2019). Individual tontine accounts. Journal of Accounting and Finance, 19(8).
- Gemmo, I., Rogalla, R., and Weinert, J. H. (2020). Optimal portfolio choice with tontines under systematic longevity risk. Annals of Actuarial Science, 14(2):302–315.
- Gomes, F. J., Kotlikoff, L. J., and Viceira, L. M. (2008). Optimal life-cycle investing with flexible labor supply: A welfare analysis of life-cycle funds. American Economic Review, 98(2):297–303.
- Hubener, A., Maurer, R., and Mitchell, O. S. (2016). How family status and social security claiming options shape optimal life cycle portfolios. The Review of Financial Studies, 29(4):937–978.
- Jacobsen, R., Keiding, N., and Lynge, E. (2002). Long term mortality trends behind low life expectancy of Danish women. Journal of Epidemiology & Community Health, 56(3):205–208.
- Kim, H. H., Maurer, R., and Mitchell, O. S. (2016). Time is money: Rational life cycle inertia and the delegation of investment management. Journal of financial economics, 121(2):427–447.
- Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting U.S. mortality. Journal of the American Statistical Association, 87(419):659–671.
- Li, H. (2018). Dynamic hedging of longevity risk: The effect of trading frequency. ASTIN Bulletin: The Journal of the IAA, 48(1):197–232.



- Milevsky, M. A., Promislow, S. D., and Young, V. R. (2006). Killing the law of large numbers: Mortality risk premiums and the Sharpe ratio. Journal of Risk and Insurance, 73(4):673–686.
- Milevsky, M. A. and Salisbury, T. S. (2015). Optimal retirement income tontines. Insurance: Mathematics and Economics, 64:91–105.
- Newfield, P. (2014). The tontine: An improvement on the conventional annuity? The Journal of Retirement, 1(3):37–48.
- Osmond, C. (1985). Using age, period and cohort models to estimate future mortality rates. International Journal of Epidemiology, 14(1):124–129.
- Piggott, J., Valdez, E. A., and Detzel, B. (2005). The simple analytics of a pooled annuity fund. Journal of Risk and Insurance, 72(3):497–520.
- Qiao, C. and Sherris, M. (2013). Managing systematic mortality risk with group self-pooling and annuitization schemes. Journal of Risk and Insurance, 80(4):949–974.
- Renshaw, A. E. and Haberman, S. (2006). A cohort-based extension to the lee–carter model for mortality reduction factors. Insurance: Mathematics and Economics, 38(3):556–570.
- Richter, A. and Weber, F. (2011). Mortality-indexed annuities managing longevity risk via product design. North American Actuarial Journal, 15(2):212–236.
- Sabin, M. J. (2010). Fair tontine annuity. Available at SSRN: <https://ssrn.com/abstract=1579932>.
- Schwarcz, D. (2013). Transparency opaque: Understanding the lack of transparency in insurance consumer protection. UCLA L. Rev., 61:394.
- Steinorth, P. and Mitchell, O. S. (2015). Valuing variable annuities with guaranteed minimum lifetime withdrawal benefits. Insurance: Mathematics and Economics, 64:246–258.
- Weinert, J. H. and Gründl, H. (2021). The modern tontine: An innovative instrument for longevity risk management in an aging society. European Actuarial Journal, 11(1):49–86.
- Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. The Review of Economic Studies, 32(2):137–150.

Yanez, N. D., Weiss, N. S., Romand, J.-A., and Treggiari, M. M. (2020). Covid-19 mortality risk for older men and women. *BMC public health*, 20(1):1–7.

## A Stochastic Mortality Model

In practice, the expected remaining lifetime and mortality rates can be estimated using the data of different populations, e.g., the general population (or the reference population), and the annuitant pool (or the book population). The book population is usually a proportion of the reference population, and survival probabilities among the book population are usually greater than those in the other one due to the adverse selection. To distinguish the estimates from different populations, we use  $\mu_x^{(r)}$  and  $\mu_x^{(b)}$  to represent the stochastic forces of mortality for the reference population and book population respectively. The mortality rates of these two populations are correlated. To obtain the estimates, we need to first introduce our stochastic mortality models. Here, we add an additional calendar year factor  $h$  to them. To be more concrete, we apply  $\mu_{\{x,h\}}^{(i)}$ ,  $i = r, b$  for population  $i$ , where the central death rate is marked as  $m_{\{x,h\}}^{(i)}$ .

### A.1 Two-population Mortality Model

Due to its tractable form and inclusion of a cohort effect, the Age-Period-Cohort (APC) model is widely used to model the stochastic mortality rate (Osmond, 1985; Jacobsen et al., 2002; Renshaw and Haberman, 2006). We follow Dowd et al. (2011) and Chen et al. (2022a) to establish the two-population stochastic mortality model. The key idea is to estimate the parameters for mortality models of the reference and book population respectively, using the historical data; and based on the fitted models, when carrying out the mortality projections, we then connect the mortality rates of these two populations by adding specific components to the dynamics of the period effect and cohort effect of the book population. Further details are stated in the next Subsection A.2. Here we start by introducing the mortality models of the two populations to fit the historical data.

According to Currie (2006), we set up our stochastic mortality model for the reference population by:

$$\log m_{\{x,h\}}^{(r)} = \beta_x^{(r)} + n_a^{-1} \kappa_h^{(r)} + n_a^{-1} \omega_{h-x}^{(r)}, \quad (\text{A.1})$$

where  $r$  denotes the reference population,  $x$  represents the age,  $h$  means the calendar year of observation and  $h - x$  is thus the cohort year of birth.  $\beta_x^{(r)}$ ,  $\kappa_h^{(r)}$  and  $\omega_{h-x}^{(r)}$  are described as the age-specific, period-related and cohort-related stochastic factors for the reference population respectively. Correspondingly,  $n_a$  represents the number of ages in the sample under investigation.

Similarly, the mortality model for the book population is also assumed to be given by an APC model as well, i.e.

$$\log m_{\{x,h\}}^{(b)} = \beta_x^{(b)} + n_a^{-1} \kappa_h^{(b)} + n_a^{-1} \omega_{h-x}^{(b)}, \quad (\text{A.2})$$

where  $b$  marks the book population, and  $\beta_x^{(b)}$ ,  $\kappa_h^{(b)}$  and  $\omega_{h-x}^{(b)}$  are the age, period and cohort stochastic factors for the book population respectively.

Furthermore, by taking the assumptions that the forces of mortality ( $\mu_{\{x,h\}}^{(i)}$ ,  $i = r, b$ ) remain constant over each year of integer age and over each calendar year, and the size of these two populations at all ages remains constant over time (Cairns et al., 2009), we get

$$m_{\{x,h\}}^{(i)} = \mu_{\{x,h\}}^{(i)}, \quad (\text{A.3})$$

$$q_{\{x,h\}}^{(i)} = 1 - \exp\left(-\mu_{\{x,h\}}^{(i)}\right) = 1 - \exp\left(-m_{\{x,h\}}^{(i)}\right), \quad i = r, b. \quad (\text{A.4})$$

And then the survival curve can be written as:

$$p_{\{x,h\}}^{(i)} = \exp\left(-m_{\{x,h\}}^{(i)}\right), \quad (\text{A.5})$$

$${}_t p_{\{x,h\}}^{(i)} = \exp\left(-\sum_{s=0}^t \mu_{\{x,h+s\}}^{(i)}\right), \quad i = r, b. \quad (\text{A.6})$$

## A.2 Mortality Projection under the Real-world Measure $\mathbb{P}$

In order to make projections to the mortality rates, for the reference population, we assume the dynamics of the period effect ( $\kappa_h^{(r)}$ ) can be described by a random walk with drift  $\alpha^{(\kappa,r)}$ , and the cohort effect ( $\omega_{h-x}^{(r)}$ ) follows  $ARIMA(1, 1, 0)$  with drift  $\alpha^{(\omega,r)}$

$$\kappa_h^{(r)} - \kappa_{h-1}^{(r)} = \alpha^{(\kappa,r)} + \Theta^{(\kappa,r)} Z_h^{(\kappa,r)}, \quad (\text{A.7})$$

$$\omega_{h-x}^{(r)} - \omega_{h-x-1}^{(r)} = \phi^{(\omega,r)}(\omega_{h-x-1}^{(r)} - \omega_{h-x-2}^{(r)}) + (1 - \phi^{(\omega,r)})\alpha^{(\omega,r)} + \Theta^{(\omega,r)} Z_{h-x}^{(\omega,r)}. \quad (\text{A.8})$$

Here,  $Z_h^{(\kappa,r)}$  and  $Z_{h-x}^{(\omega,r)}$  are standard normal random variables under the real-world measure  $\mathbb{P}$ .

It is reasonable to assume that the book population (e.g., the insurer's annuity portfolio) has lower mortality rates than the reference population (e.g., the national population) (Cairns et al., 2011b). Furthermore, the mortality rates of the reference population and book population are correlated. We incorporate these features via additional specific spread patterns in the predictive distributions of the period factor ( $\kappa_h^{(b)}$ ) and cohort factor ( $\omega_{h-x}^{(b)}$ ) for the book population's mortality model. With reference to Dowd et al. (2011), the dynamics of the period-specific effects can be written as:

$$\kappa_h^{(b)} - \kappa_{h-1}^{(b)} = \psi^{(\kappa)}(\kappa_{h-1}^{(r)} - \kappa_{h-1}^{(b)}) + \alpha^{(\kappa,b)} + \Theta^{(\kappa,b1)} Z_h^{(\kappa,r)} + \Theta^{(\kappa,b2)} Z_h^{(\kappa,b)}, \quad (\text{A.9})$$

$$\begin{aligned} \omega_{h-x}^{(b)} - \omega_{h-x-1}^{(b)} &= \phi^{(\omega,b)}(\omega_{h-x-1}^{(b)} - \omega_{h-x-2}^{(b)}) + \psi^{(\omega)}(\omega_{h-x-1}^{(r)} - \omega_{h-x-1}^{(b)}) \\ &\quad + (1 - \phi^{(\omega,b)})\alpha^{(\omega,b)} + \Theta^{(\omega,b1)} Z_{h-x}^{(\omega,r)} + \Theta^{(\omega,b2)} Z_{h-x}^{(\omega,b)}. \end{aligned} \quad (\text{A.10})$$

Similarly,  $Z_h^{(\kappa,b)}$  and  $Z_{h-x}^{(\omega,b)}$  are standard normal random variables under  $\mathbb{P}$  in terms of the book population.  $(Z_h^{(\kappa,b)}, Z_h^{(\kappa,r)})$  are assumed to be independent and identically distributed, and so are the pair of  $(Z_{h-x}^{(\omega,b)}, Z_{h-x}^{(\omega,r)})$ . The spread terms are specified as  $(\psi^{(\kappa)}(\kappa_{h-1}^{(r)} - \kappa_{h-1}^{(b)}))$  for the period effect and  $(\psi^{(\omega)}(\omega_{h-x-1}^{(r)} - \omega_{h-x-1}^{(b)}))$  for the cohort effect. Positive values of  $\psi^{(\kappa)}$  and  $\psi^{(\omega)}$  will pull the spread towards 0. More detailed explanations can be found in Dowd et al. (2011).

### A.3 Mortality Projection under the Risk Neutral Measure $\mathbb{Q}$

With reference to the existing literature (e.g., Cairns et al., 2006; Chen et al., 2022a), we specify the risk neutral measure by incorporating the longevity risk premiums to the mortality trends. More concretely, we add some constants to drift terms of period effects for both populations, i.e.

$$\tilde{Z}_h^{(\kappa,r)} = g^{(r)} + Z_h^{(\kappa,r)}, \quad \tilde{Z}_h^{(\kappa,b)} = g^{(b)} + Z_h^{(\kappa,b)}.$$

It is assumed that  $(\tilde{Z}_h^{(\kappa,r)}, \tilde{Z}_h^{(\kappa,b)})$  are two dimensional i.i.d. normal random variables under the risk neutral measure  $\mathbb{Q}$ .  $g^{(r)}$  and  $g^{(b)}$  represent the market prices of the longevity risk related to  $Z_h^{(\kappa,r)}$  and  $Z_h^{(\kappa,b)}$  respectively. Correspondingly, the longevity risk premium for the reference population is given by  $g^{(r)}$ ; and that for the book population is composed by  $g^{(r)}$  and  $g^{(b)}$ . Then, we further write down the dynamics of the

period effects of both populations under  $\mathbb{Q}$ :

$$\kappa_h^{(r)} - \kappa_{h-1}^{(r)} = \underbrace{\alpha^{(\kappa,r)} + \Theta^{(\kappa,r)} g^{(r)}}_{\tilde{\alpha}^{(\kappa,r)}} + \Theta^{(\kappa,r)} Z_h^{(\kappa,r)}, \quad (\text{A.11})$$

$$\kappa_h^{(b)} - \kappa_{h-1}^{(b)} = \psi^{(\kappa)}(\kappa_{h-1}^{(r)} - \kappa_{h-1}^{(b)}) + \underbrace{\alpha^{(\kappa,b)} + \Theta^{(\kappa,b1)} g^{(r)} + \Theta^{(\kappa,b2)} g^{(b)}}_{\tilde{\alpha}^{(\kappa,b)}} + \Theta^{(\kappa,b1)} Z_h^{(\kappa,r)} + \Theta^{(\kappa,b2)} Z_h^{(\kappa,b)}. \quad (\text{A.12})$$

## A.4 Data Introduction

In this paper, we acquire the historical data of two populations from 1961 to 2020 for ages 56 to 98. England and Wales (EW) males experience is used as the reference population data, which is generated by the general population of England and Wales, and can be obtained from the Human Mortality Database (HMD). As for the book population, we use the datasets of male pensioners from UK Continuous Mortality Investigation (CMI). The pensioner datasets are composed of the ones that are able to participate in the pension scheme. In this sense, it is reasonable to infer that comparing to males from HMD, males from CMI are possibly to live longer because of adverse selection. Moreover, the CMI population is (mostly) a sub-population in roughly 10% in size of the EW population.

## A.5 Mortality Rates for the Base and Stress Case under the Real World Measure $\mathbb{P}$

To estimate parameters of the mortality models, we apply a two-stage method, which is widely used by researchers (e.g., Lee and Carter, 1992; Cairns et al., 2009, 2011a). More concretely, at stage 1, age, period and cohort effects regardless of their underlying dynamics are estimated; at stage 2, the mortality rates can be forecast by fitting appropriate stochastic processes to the period and cohort effects. Due to the fact that our obtained historical data is only for ages till 98, we extrapolate the age effects  $\beta_x^{(r)}$  and  $\beta_x^{(b)}$  to 120, and simulate the future mortality rates by extrapolating the period effect and cohort effect applying their corresponding time-series processes.

In the base case (i), where there is no unforeseen shock, we apply the data from 1961 to 2018 to estimate and forecast the mortality rates. In the stress case with an unexpected mortality shock (ii-M), we apply the entire data set from 1961 to 2020 for the mortality estimation and projection. And in the stress case with an unexpected longevity shock

(ii-L), we add a safety buffer to mortality trends in the mortality model estimated under the base-case assumptions. To elaborate further, in the context of a longevity shock, we assume a 10% increase in the period parameter ( $\kappa_h^{(r)}$ ) for the base-case reference mortality model. Simultaneously, the period parameter ( $\kappa_h^{(b)}$ ) for the base-case book mortality model experiences a larger inflation, specifically, a 25% increase. This adjustment accounts for the notion that the book population is generally wealthier than the reference population. Furthermore, it considers the possibility that a medical treatment breakthrough may initially benefit wealthier individuals more, given their ability to afford associated expenses.

In the following contents, we take the cohort 1954 as an example. To obtain a relatively stable prediction of the mortality rates, we have simulated 10000 mortality paths for both populations.

#### **Obtained mortality data for the base case:**

1. Reference population ( $DATA_{\{refbase\}}$ ): 10000 simulated paths starting from 2019 for cohort-1954 males.
2. Book population ( $DATA_{\{bookbase\}}$ ): 10000 simulated paths starting from 2019 for cohort-1954 males.

In the stress case (ii-M), where an unforeseen mortality shock strikes right after the contract is made, the prediction of mortality starts from 2021. Here, we construct the data as: mortality rates for ages 65 and 66 of cohort-1954 (in years 2019 and 2020) are calculated based on real data; and mortality rates over 67 (included) of the cohort-1954 (in 2021 and afterwards) are predicted.

#### **Obtained mortality data for the stress case (ii-M):**

1. Reference population ( $DATA_{\{refstr(M)\}}$ ): Real mortality rates for years 2019-2020, and 10000 simulated paths starting from 2021 for cohort-1954 males. In other words, for 10000 paths of the mortality rates in the stress case, they have the same mortality rates for years 2019-2020.
2. Book population ( $DATA_{\{bookstr(M)\}}$ ): Real mortality rates for years 2019-2020, and 10000 simulated paths starting from 2021 for cohort-1954 males.

In the stress case (ii-L), shortly after the contract begins, an unforeseen longevity shock takes place. As elaborate above, we adjust the base-case mortality model to get the mortality rates under the longevity shock. The mortality projection starts from 2019.

**Obtained mortality data for the stress case (ii-L):**

1. Reference population (  $DATA_{\{refstr(L)\}}$  ): 10000 (adjusted) simulated paths starting from 2019 for cohort-1954 males.
2. Book population (  $DATA_{\{bookstr(L)\}}$  ): 10000 (adjusted) simulated paths starting from 2019 for cohort-1954 males.

## **A.6 Parameter Calibration for the Risk Neutral Measure $\mathbb{Q}$**

To calibrate the market prices of the longevity risk, we refer to the approach described in Chen et al. (2022a), which relies on the Sharpe ratio approach (Milevsky et al., 2006). In this paper, we obtain the values for  $g^{(r)}$  and  $g^{(b)}$  by setting the Sharpe ratio of the insurer's annuity portfolio as 8% (Bauer et al., 2010)), where the derivations of the Sharpe ratio for the book and reference population are directly from Chen et al. (2022a).

## B Proofs

### B.1 Proof of Theorem 2.1

In order to obtain the optimal payoff stream functions to the regular tonuity, we need to maximize the policyholder's expected utility subject to the condition  $P_0^{\text{RT}}(\mathbb{Q}) = v$ . In this case, we can write down the Lagrangian function, i.e.

$$\begin{aligned}
L &= \mathbb{E}^{\mathbb{P}} \left[ \int_0^\infty e^{-\eta t} \left( \mathbb{1}_{\{0 \leq t < \min\{\tau, T^{(b)}\}\}} u \left( \frac{nd_{[\tau]}(t)}{N^{(b)}(t)} \right) + \mathbb{1}_{\{\tau \leq t < T^{(b)}\}} u(c_{[\tau]}(t)) \right) dt \right] \\
&\quad + \lambda_{[\tau]} \left( v - \int_0^\tau e^{-r_f t} \mathbb{E}^{\mathbb{Q}} \left[ 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right)^n \right] d_{[\tau]}(t) dt \right. \\
&\quad \left. - \int_\tau^\infty e^{-r_f t} \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right] c_{[\tau]}(t) dt \right) \\
&= \int_0^\tau e^{-\eta t} u(d_{[\tau]}(t)) \mathbb{E}^{\mathbb{P}} \left[ \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right)^k \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right)^{n-k} \right] dt \\
&\quad + \int_\tau^\infty e^{-\eta t} u(c_{[\tau]}(t)) \mathbb{E}^{\mathbb{P}} \left[ e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right] dt \\
&\quad + \lambda_{[\tau]} \left( v - \int_0^\tau e^{-r_f t} \mathbb{E}^{\mathbb{Q}} \left[ 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right)^n \right] d_{[\tau]}(t) dt \right. \\
&\quad \left. - \int_\tau^\infty e^{-r_f t} \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right] c_{[\tau]}(t) dt \right) \\
&= \int_0^\tau e^{-\eta t} u(d_{[\tau]}(t)) \mathbb{E}^{\mathbb{P}} \left[ \mathcal{K}_{n, \gamma} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right) \right] dt + \int_\tau^\infty e^{-\eta t} u(c_{[\tau]}(t)) \mathbb{E}^{\mathbb{P}} \left[ e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right] dt \\
&\quad + \lambda_{[\tau]} \left( v - \int_0^\tau e^{-r_f t} \mathbb{E}^{\mathbb{Q}} \left[ 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right)^n \right] d_{[\tau]}(t) dt \right. \\
&\quad \left. - \int_\tau^\infty e^{-r_f t} \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right] c_{[\tau]}(t) dt \right).
\end{aligned}$$

Denote  $\mathcal{K}_{n, \gamma} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right) := \sum_{k=1}^n \binom{n}{k} \left( \frac{k}{n} \right)^\gamma \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right)^k \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right)^{n-k}$ .



In order to obtain the maximum of Lagrangian function, we take derivatives with respect to  $d_{[\tau]}(t)$  and  $c_{[\tau]}(t)$ :

$$\begin{aligned}\frac{\partial L}{\partial d_{[\tau]}(t)} &= \int_0^\tau e^{-\eta t} u'(d_{[\tau]}(t)) \mathbb{E}^\mathbb{P} \left[ \mathcal{K}_{n,\gamma} \left( e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right) \right] dt \\ &\quad - \lambda_{[\tau]} \int_0^\tau e^{-r_f t} \mathbb{E}^\mathbb{Q} \left[ 1 - \left( 1 - e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right)^n \right] d_{[\tau]}(t) dt \stackrel{!}{=} 0. \\ \frac{\partial L}{\partial c_{[\tau]}(t)} &= \int_\tau^\infty e^{-\eta t} u'(c_{[\tau]}(t)) \mathbb{E}^\mathbb{P} \left[ e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right] dt \\ &\quad - \lambda_{[\tau]} \int_\tau^\infty e^{-r_f t} \mathbb{E}^\mathbb{Q} \left[ e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right] c_{[\tau]}(t) dt \stackrel{!}{=} 0.\end{aligned}$$

Therefore, we have

$$\begin{aligned}d_{[\tau]}^*(t) &= \left( \lambda_{[\tau]} e^{(\eta - r_f)t} \frac{\mathbb{E}^\mathbb{Q} \left[ 1 - \left( 1 - e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right)^n \right]}{\mathbb{E}^\mathbb{P} \left[ \mathcal{K}_{n,\gamma} \left( e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right) \right]} \right)^{-\frac{1}{\gamma}}, \forall t \in [0, \tau), \\ c_{[\tau]}^*(t) &= \left( \lambda_{[\tau]} e^{(\eta - r_f)t} \frac{\mathbb{E}^\mathbb{Q} \left[ e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right]}{\mathbb{E}^\mathbb{P} \left[ e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right]} \right)^{-\frac{1}{\gamma}}, \forall t \in [\tau, \infty).\end{aligned}$$

The  $\lambda_{[\tau]} > 0$  is chosen to meet the budget constraint, such that

$$\begin{aligned}\lambda_{[\tau]} &= \left( \frac{1}{v} \int_0^\tau e^{(\frac{r_f - \eta}{\gamma} - r_f)t} \frac{\mathbb{E}^\mathbb{P} \left[ \mathcal{K}_{n,\gamma} \left( e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right) \right]^{\frac{1}{\gamma}}}{\mathbb{E}^\mathbb{Q} \left[ 1 - \left( 1 - e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right)^n \right]^{\frac{1}{\gamma} - 1}} dt \right. \\ &\quad \left. + \frac{1}{v} \int_\tau^\infty e^{(\frac{r_f - \eta}{\gamma} - r_f)t} \frac{\mathbb{E}^\mathbb{P} \left[ e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right]^{\frac{1}{\gamma}}}{\mathbb{E}^\mathbb{Q} \left[ e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right]^{\frac{1}{\gamma} - 1}} dt \right)^\gamma.\end{aligned}$$

## B.2 Proof of Theorem 3.1

Let  $\mu_{x+t,h+t} := (\mu_{x+t,h+t}^{(b)}, \mu_{x+t,h+t}^{(r)})$ . The expected utility of the dynamic tonuity can be written as

$$\begin{aligned}
& \mathbb{E}^{\mathbb{P}} \left[ \int_0^\infty e^{-\eta t} \left( \mathbb{1}_{\{0 \leq t < \min\{\Lambda, T^{(b)}\}\}} u \left( \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b)}(t)} \right) + \mathbb{1}_{\{\Lambda \leq t < T^{(b)}\}} u(c_{[\bar{\Lambda}]}(t)) \right) dt \right] \\
&= \int_0^\infty e^{-\eta t} \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \min\{\Lambda, T^{(b)}\}\}} u \left( \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b)}(t)} \right) + \mathbb{1}_{\{\Lambda \leq t < T^{(b)}\}} u(c_{[\bar{\Lambda}]}(t)) \right] dt \\
&= \int_0^\infty e^{-\eta t} \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathbb{1}_{\{0 \leq t < T^{(b)}\}} u \left( \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b)}(t)} \right) + \mathbb{1}_{\{\Lambda \leq t\}} \mathbb{1}_{\{t < T^{(b)}\}} u(c_{[\bar{\Lambda}]}(t)) \right] dt \\
&= \int_0^\infty e^{-\eta t} \mathbb{E}^{\mathbb{P}} \left[ \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathbb{1}_{\{0 \leq t < T^{(b)}\}} u \left( \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b)}(t)} \right) + \mathbb{1}_{\{\Lambda \leq t\}} \mathbb{1}_{\{t < T^{(b)}\}} u(c_{[\bar{\Lambda}]}(t)) \mid \{\mu_{x+t,h+t}\}_{t \geq 0} \right] \right] dt \\
&= \int_0^\infty e^{-\eta t} \mathbb{E}^{\mathbb{P}} \left[ \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathbb{1}_{\{0 \leq t < T^{(b)}\}} u \left( \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b)}(t)} \right) \mid \{\mu_{x+t,h+t}\}_{t \geq 0} \right] \right. \\
&\quad \left. + \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} \mathbb{1}_{\{t < T^{(b)}\}} u(c_{[\bar{\Lambda}]}(t)) \mid \{\mu_{x+t,h+t}\}_{t \geq 0} \right] \right] dt \\
&= \int_0^\infty e^{-\eta t} \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < T^{(b)}\}} u \left( \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b)}(t)} \right) \mid \{\mu_{x+t,h+t}\}_{t \geq 0} \right] \right. \\
&\quad \left. + \mathbb{1}_{\{\Lambda \leq t\}} \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{t < T^{(b)}\}} u(c_{[\bar{\Lambda}]}(t)) \mid \{\mu_{x+t,h+t}\}_{t \geq 0} \right] \right] dt \\
&= \int_0^\infty e^{-\eta t} \mathbb{E}^{\mathbb{P}} \left[ \underbrace{\mathbb{1}_{\{0 \leq t < \Lambda\}} \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < T^{(b)}\}} u \left( \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b)}(t)} \right) \mid \{\mu_{x+t,h+t}\}_{t \geq 0} \right]}_{\mathcal{C}} \right] dt \\
&\quad + \int_0^\infty e^{-\eta t} \mathbb{E}^{\mathbb{P}} \left[ \underbrace{\mathbb{1}_{\{\Lambda \leq t\}} \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{t < T^{(b)}\}} u(c_{[\bar{\Lambda}]}(t)) \mid \{\mu_{x+t,h+t}\}_{t \geq 0} \right]}_{\mathcal{D}} \right] dt.
\end{aligned}$$

The first conditional expectation  $\mathcal{C}$  is given by

$$\begin{aligned}
& \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < T^{(b)}\}} u \left( \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b)}(t)} \right) \mid \{\mu_{x+t,h+t}\}_{t \geq 0} \right] \\
&= \mathbb{E}^{\mathbb{P}} \left[ e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \cdot u \left( \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b)}(t)} \right) \mid \{\mu_{x+t,h+t}\}_{t \geq 0}, T^{(b)} > t \right] \\
&= e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \cdot \mathbb{E}^{\mathbb{P}} \left[ u \left( \frac{nd_{[\bar{\Lambda}]}(t)}{N^{(b)}(t)} \right) \mid \{\mu_{x+t,h+t}\}_{t \geq 0}, T^{(b)} > t \right] \\
&= e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \cdot u(d_{[\bar{\Lambda}]}(t)) \sum_{k=0}^{n-1} \left( \frac{n}{k+1} \right)^{1-\gamma} \binom{n-1}{k} \left( e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right)^k \left( 1 - e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right)^{n-1-k}
\end{aligned}$$

$$\begin{aligned}
&= u(d_{[\bar{\Lambda}]}(t)) \sum_{k=0}^{n-1} \left(\frac{k+1}{n}\right)^\gamma \binom{n}{k+1} \left(e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds}\right)^{k+1} \left(1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds}\right)^{n-1-k} \\
&= u(d_{[\bar{\Lambda}]}(t)) \sum_{k=1}^n \left(\frac{k}{n}\right)^\gamma \binom{n}{k} \left(e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds}\right)^k \left(1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds}\right)^{n-k} \\
&=: u(d_{[\bar{\Lambda}]}(t)) \mathcal{K}_{n, \gamma} \left(e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds}\right).
\end{aligned}$$

The second conditional expectation  $\mathcal{D}$  is given by

$$\begin{aligned}
\mathbb{E}^\mathbb{P} \left[ \mathbb{1}_{\{t < T^{(b)}\}} u(c_{[\bar{\Lambda}]}(t)) \mid \{\mu_{x+t, h+t}\}_{t \geq 0} \right] &= u(c_{[\bar{\Lambda}]}(t)) \mathbb{E}^\mathbb{P} \left[ \mathbb{1}_{\{t < T^{(b)}\}} \mid \{\mu_{x+t, h+t}\}_{t \geq 0} \right] \\
&= u(c_{[\bar{\Lambda}]}(t)) e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds}.
\end{aligned}$$

Putting everything together, we obtain the following expression for the expected lifetime utility of the dynamic tonuity:

$$\begin{aligned}
&\int_0^\infty e^{-\eta t} \mathbb{E}^\mathbb{P} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} u(d_{[\bar{\Lambda}]}(t)) \mathcal{K}_{n, \gamma} \left(e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds}\right) \right] dt \\
&\quad + \int_0^\infty e^{-\eta t} \mathbb{E}^\mathbb{P} \left[ \mathbb{1}_{\{\Lambda \leq t\}} u(c_{[\bar{\Lambda}]}(t)) e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right] dt \\
&= \int_0^\infty e^{-\eta t} u(d_{[\bar{\Lambda}]}(t)) \mathbb{E}^\mathbb{P} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n, \gamma} \left(e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds}\right) \right] dt \\
&\quad + \int_0^\infty e^{-\eta t} u(c_{[\bar{\Lambda}]}(t)) \mathbb{E}^\mathbb{P} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right] dt.
\end{aligned}$$

A policyholder starting with an initial wealth  $v$  then faces the following optimization problem:

$$\begin{aligned}
&\max_{c_{[\bar{\Lambda}]}(t), d_{[\bar{\Lambda}]}(t)} \int_0^\infty e^{-\eta t} u(d_{[\bar{\Lambda}]}(t)) \mathbb{E}^\mathbb{P} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n, \gamma} \left(e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds}\right) \right] dt \\
&\quad + \int_0^\infty e^{-\eta t} u(c_{[\bar{\Lambda}]}(t)) \mathbb{E}^\mathbb{P} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right] dt
\end{aligned}$$

subject to

$$\begin{aligned}
v &= \int_0^\infty e^{-r_f t} d_{[\bar{\Lambda}]}(t) \mathbb{E}^\mathbb{Q} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \left(1 - \left(1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds}\right)^n\right) \right] dt \\
&\quad + \int_0^\infty e^{-r_f t} c_{[\bar{\Lambda}]}(t) \mathbb{E}^\mathbb{Q} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right] dt.
\end{aligned}$$

The first order conditions are given by

$$\begin{aligned} e^{-\eta t} u' \left( d_{[\bar{\Lambda}]}(t) \right) \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n,\gamma} \left( e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right) \right] &= \lambda e^{-r_f t} \mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right)^n \right) \right] \\ e^{-\eta t} u' \left( c_{[\bar{\Lambda}]}(t) \right) \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right] &= \lambda e^{-r_f t} \mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right]. \end{aligned}$$

These deliver the optimal payout functions:

$$\begin{aligned} d_{[\bar{\Lambda}]}^*(t) &= \left( \lambda e^{(\eta-r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right)^n \right) \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n,\gamma} \left( e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right) \right]} \right)^{-\frac{1}{\gamma}} \\ c_{[\bar{\Lambda}]}^*(t) &= \left( \lambda e^{(\eta-r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right]} \right)^{-\frac{1}{\gamma}}. \end{aligned}$$

Plugging these expressions back into the budget constraint delivers an expression for the Lagrangian multiplier  $\lambda$ :

$$\begin{aligned} v &= \int_0^\infty e^{-r_f t} \left( \lambda e^{(\eta-r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right)^n \right) \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n,\gamma} \left( e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right) \right]} \right)^{-\frac{1}{\gamma}} \\ &\quad \cdot \mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right)^n \right) \right] dt \\ &\quad + \int_0^\infty e^{-r_f t} \left( \lambda e^{(\eta-r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right]} \right)^{-\frac{1}{\gamma}} \\ &\quad \cdot \mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right] dt, \end{aligned}$$

which is equivalent to

$$\begin{aligned} \lambda^{\frac{1}{\gamma}} &= \frac{1}{v} \int_0^\infty e^{-r_f t} \left( e^{(\eta-r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right)^n \right) \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n,\gamma} \left( e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right) \right]} \right)^{-\frac{1}{\gamma}} \\ &\quad \cdot \mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right)^n \right) \right] dt \\ &\quad + \frac{1}{v} \int_0^\infty e^{-r_f t} \left( e^{(\eta-r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s,h+s}^{(b)} ds} \right]} \right)^{-\frac{1}{\gamma}} \\ &\quad \cdot \mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s,h+s}^{(b,\mathbb{Q})} ds} \right] dt. \end{aligned}$$

For the optimal level of expected utility, we obtain

$$\begin{aligned}
& \int_0^\infty e^{-\eta t} u \left( \left( \lambda e^{(\eta-r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds \right)^n \right) \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n, \gamma} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds \right)} \right]} \right)^{-\frac{1}{\gamma}} \right) \\
& \quad \cdot \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n, \gamma} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right) \right] dt \\
& + \int_0^\infty e^{-\eta t} u \left( \left( \lambda e^{(\eta-r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right]} \right)^{-\frac{1}{\gamma}} \right) \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right] dt \\
& = \frac{\lambda^{1-\frac{1}{\gamma}}}{1-\gamma} \left[ \int_0^\infty e^{-\eta t} \left( \left( e^{(\eta-r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds \right)^n \right) \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n, \gamma} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right)} \right]} \right)^{-\frac{1}{\gamma}} \right)^{1-\gamma} \right. \\
& \quad \cdot \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n, \gamma} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right) \right] dt \\
& \quad \left. + \int_0^\infty e^{-\eta t} \left( \left( e^{(\eta-r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right]} \right)^{-\frac{1}{\gamma}} \right)^{1-\gamma} \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right] dt \right] \\
& = \frac{\lambda^{1-\frac{1}{\gamma}}}{1-\gamma} \left[ \int_0^\infty e^{-\eta t} \left( \left( e^{(\eta-r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \left( 1 - \left( 1 - e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds \right)^n \right) \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n, \gamma} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right)} \right]} \right)^{-\frac{1}{\gamma}} \right)^{1-\frac{1}{\gamma}} \right. \\
& \quad \cdot \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{0 \leq t < \Lambda\}} \mathcal{K}_{n, \gamma} \left( e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right) \right] dt \\
& \quad \left. + \int_0^\infty e^{-\eta t} \left( \left( e^{(\eta-r_f)t} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b, \mathbb{Q})} ds} \right]}{\mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right]} \right)^{-\frac{1}{\gamma}} \right)^{1-\frac{1}{\gamma}} \mathbb{E}^{\mathbb{P}} \left[ \mathbb{1}_{\{\Lambda \leq t\}} e^{-\int_0^t \mu_{x+s, h+s}^{(b)} ds} \right] dt \right] \\
& = \frac{\lambda^{1-\frac{1}{\gamma}}}{1-\gamma} \cdot \lambda^{\frac{1}{\gamma}} \cdot v = \frac{\lambda}{1-\gamma} \cdot v.
\end{aligned}$$

□



### **3 How does the insurer's mobile application sales strategy perform?**

**Source:**

Chen, A., Chen, Y., Murphy, F., Xu, W., & Xu, X. (2023). How does the insurer's mobile application sales strategy perform?. *Journal of Risk and Insurance*, 90, 487–519.

URL: <https://onlinelibrary.wiley.com/doi/full/10.1111/jori.12424>

DOI: <https://doi.org/10.1111/jori.12424>

© 2023 American Risk and Insurance Association. Reprinted with kind permission.







## ORIGINAL ARTICLE

Journal of Risk and Insurance

# How does the insurer's mobile application sales strategy perform?

An Chen<sup>1</sup> | Yusha Chen<sup>1</sup>  | Finbarr Murphy<sup>2</sup> | Wei Xu<sup>2</sup> | Xian Xu<sup>3</sup> 

<sup>1</sup>Institute of Insurance Science, Ulm University, Ulm, Germany

<sup>2</sup>Kemmy Business School, University of Limerick, Limerick, Ireland

<sup>3</sup>School of Economics, Fudan University, Shanghai, China

**Correspondence**

Yusha Chen, Institute of Insurance Science, Ulm University, Helmholtzstr. 20, 89069 Ulm, Germany.  
Email: [yusha.chen@uni-ulm.de](mailto:yusha.chen@uni-ulm.de)

**Abstract**

While the impact of an Internet-based sales strategy on sales performance has been well studied, there is little academic research that examines the impact of a mobile application (MA) sales strategy on the sales performance of insurers. Using a unique data set for term life insurance policies from a Chinese life insurer, we study the impact of implementing this strategy on insurance purchases. We find a significant growth in the insurance purchase quantity and somewhat lower growth in premiums received from new policies. This paper determines that this is due to improved channel accessibility and the cost reduction of the MA channel. Although sales of traditional distribution channels are cannibalized in the short term by the MA distribution strategy, this substitution effect does not persist in the long run. In addition, we find that this strategy reduces impulsive purchases.

**KEYWORDS**

heterogeneity analysis, insurance distribution, mobile application, sales performance

**JEL CLASSIFICATION**

G22, D12, M31

## 1 | INTRODUCTION

The insurance industry is undergoing a digital transformation. A digital strategy helps insurers increase their market valuations (Fritzsche et al., 2021), improve their business performance (Bohnert et al., 2019), and enhance their risk management capability (Fritzsche & Bohnert, 2021). The introduction of the Internet sales channel in recent years has led to significant growth in insurance sales. From 2014 to 2019, the received premiums of online life insurance in China grew annually by 39.4% in conjunction with the development of the Internet infrastructure.<sup>1</sup> In contrast to Internet websites, consumers are spending more time on mobile devices and prefer to use mobile applications (MAs) (Cognizant, 2015). MAs can offer more convenience, faster speed of use (e.g., fewer clicks featuring one-hand operation and minimal time), more hardware functions (e.g., sensors and bluetooth), and quicker responses (Lazaris et al., 2015; Wisniewski, 2011) than Internet websites. Therefore, the MA sales strategy has been valued and expected by insurers to generate more substantial influence on insurance sales (Eling & Lehmann, 2018).

Research on the impact of the Internet/mobile Internet on sales performance of physical goods is not new in the marketing and retailing fields (e.g., Huang et al., 2016; Luo et al., 2014). However, in the insurance field, apart from some early studies that highlighted the potential of Internet-based channels for expanding the insurance business (Eastman et al., 2002; Garven, 2002), there are very few academic studies that empirically analyze how Internet distribution strategies can influence insurers' sales performance and hardly any studies that focus specifically on MA sales strategies. For instance, Brown and Goolsbee (2002) have found evidence consistent with the search theory that the price of the term life insurance falls after the introduction of insurance comparison websites (Stahl, 1989). A more recent global survey finds the positive influence of Internet use on nonlife insurance demand (Benlagha & Hemrit, 2020). There are two main differences between these studies and our work. First, they focus on the traditional PC Internet channel instead of the emerging mobile Internet channel. Second, these studies look more from the consumer's point of view (insurance demand) rather than from the insurer's point of view (e.g., the insurer's sales performance). In general, there are few studies that provide microscopic evidence of the effects of the MA sales strategy on the insurer's sales performance. To address these gaps, this paper empirically examines the effects of the MA sales strategy on the insurer's sales performance and identifies the underlying mechanisms.

When it comes to studying the extent and manner in which the MA sales strategy affects the insurer's performance, the insurance purchase quantity and premiums received from new policies (PRNPs) serve as two performance measures for two reasons. On the one hand, insurers are interested in the total number of policies and total premiums received (Berry-Stölzle et al., 2010), which correspond to the aggregated policy quantity and PRNP, respectively. On the other hand, the policy quantity and PRNP have been widely used in prior studies to measure life insurers' sales performance (e.g., Browne & Kim, 1993; Ferber & Lee, 1980; Hammond et al., 1967; Kjosovski, 2012; D. Li et al., 2007; Mantis & Farmer, 1968). In general, an increase in purchase quantity indicates the success of the MA sales strategy, if it is associated with an increase in PRNP. Conversely, if it is associated with a significant decrease in PRNP, it could indicate a failure of the MA sales strategy.

The research question is investigated through the use of a unique data set containing the term life insurance purchase records from 2015 to 2019, provided by a representative Chinese life

<sup>1</sup>The data are sourced from the Insurance Association of China (<http://www.iachina.cn/>).

insurance company. Considering the influence of time trends and different prefectures, we reconstruct the obtained data into a panel data set by calculating the purchase quantity and PRNP<sup>2</sup> at the prefecture-day level. The panel data allow us to evaluate the impact of the MA sales strategy using a two-way fixed-effect regression model that excludes the influence of unobservable factors that vary across prefectures (e.g., average insurance literacy of citizens) or time (e.g., seasonality).

Our results show that the implementation of the MA sales strategy significantly increases the insurance purchase quantity by 151.7% and PRNP by 68.5% based on 2017–2018 data sample, while recording 229.7% and 48.1% for the quantity and PRNP, respectively, when using 2015–2019 data sample. Two components are helpful to explain the different increases in the insurance purchase quantity and PRNP: (i) channel accessibility and (ii) cost reduction. Specifically, prefectures with higher MA channel accessibility exhibit larger growth effects on insurance purchases; the cost net of actuarially fair premiums of the MA channel is 64% lower than that of the offline channel. Overall, the MA sales strategy leads to growth in both the insurance purchase quantity and PRNP. The accessibility of the MA channel and lower cost both contribute to the growth in quantity, while the cost reduction in the MA channel also leads to lower growth in PRNP compared with quantity growth.

Furthermore, we carry out a series of heterogeneity analyses to study the impact of MA sales strategies: the influence of gender, education, profession, age, and whether to buy insurance for oneself. We find that the highest growth in terms of purchase quantity is among those who are bachelor's and junior college graduates, while the highest growth in PNRP is among high school graduates. This implies that the influence of education level on the effect of the MA sales strategy is nonmonotonic. It is not a simple positive relationship, but an inverted-U-shape, which deviates from the findings of the existing literature (Browne & Kim, 1993; Kjosevski, 2012; Truett & Truett, 1990). Better-educated people are likely to react more strongly towards the MA implementation as they are more knowledgeable of insurance literacy (Lin et al., 2019). Beyond a certain degree of education (e.g., bachelor's degree), they become relatively less responsive to the MA sales strategy. This may be because those with higher educational attainment are more insurance conscious and will purchase insurance regardless of sales strategies. Furthermore, we find that financial professionals and self-insureds are more responsive to the MA sales strategy. These may be explained by job relevance and the focus on oneself, respectively. Moreover, the MA sales strategy influences women more than men, possibly because women are more receptive to promotion campaigns than men (e.g., Harmon & Hill, 2003; Kwon & Kwon, 2007).

In addition, consistent with prior literature on impulsive purchasing behavior (Gilly & Wolfenbarger, 2000; Hu & Tang, 2014; Wolfenbarger & Gilly, 2001), we show that the MA sales strategy significantly reduces customers' impulsive purchases, leading to a lower policy surrender ratio during the cancellation period. Even though much literature in marketing science has found that the addition of the Internet channel precipitates substitution by both diverging consumer demand and discouragement to salespersons (Cather & Howe, 1989; Garven, 2002; Sharma & Gassenheimer, 2009; Torkestani et al., 2018), our results show that the substitution effect of the MA sales strategy on the insurer's performance is time-limited (only in the introduction year) and even reverses 1 year after introduction. This result is encouraging, as it implies that insurance practitioners' fear of the substitution effect of introducing the MA sales strategy could be overstated.

The remainder of this paper is structured as follows: Section 2 provides more institutional background about the MA and insurance. This is followed by a section introducing the data set,

<sup>2</sup>More specifically, the prefecture-day PRNP is calculated as the sum of all received (periodic or up-front) premiums from purchased insurance policies per prefecture per day.

trends overview, and empirical strategy. Section 3 provides the empirical results, and two mechanisms to explain the growth effects in insurance purchases are proposed and examined, including channel accessibility and cost reduction. Section 4 provides robustness checks and Section 5 gives the heterogeneity analysis across different customer groups. Further results regarding the impact of the MA sales strategy on the insurance surrender and the substitution effect on offline sales are analyzed in Section 6. Section 7 concludes the paper and discusses the implications for future research. The appendix provides further evidence to support the statements in the main text.

## 2 | DATA AND METHODOLOGY

### 2.1 | Background

As mentioned above, with customers shifting their attention from Internet websites towards MAs (Cognizant, 2015), insurers are investing more in the development of MAs and portals to provide customers with convenient digital self-service. The MA sales strategy has been particularly popular among life insurers in China. According to the statistics from the Insurance Association of China (Securities Daily, 2016), 18 of the 55 life insurers that have disclosed their domestic Internet insurance business launched MAs to sell insurance. The earliest one among them has been running the MA sales strategy for more than 10 years. The digital customer self-service on these launched MAs is very comprehensive, usually covering the entire service journey from quotation to claim, combined with additional health services, such as online diagnosis and fit logs. Therefore, the established market cultivation of the MA sales strategy in China provides us with rich data to examine the actual performance of this sales strategy.

### 2.2 | Data introduction

The term life insurance data set used in this paper is obtained from a Chinese life insurer who accounts for a considerable share of the life insurance market in China. This insurer has been operating in China for many years, with various branches and business locations distributed nationwide as well as a huge service network of life insurance agents, providing customers with a full cycle of life insurance products and services. Its business operation model and development path, to some extent, are representative of the Chinese life insurance industry. This insurer's MA was released before 2014, which can be freely downloaded by users from an MA store. At the beginning of the release, the insurer promoted its MA through various activities such as placing the QR code of the MA on advertisements and encouraging insurance agents to guide prospective clients to download the MA.<sup>3</sup> Insurance products initially sold on the MA included accident insurance and endowment insurance. Later, medical insurance, term life insurance, and other insurance types started to be sold on the MA.

The investigated term life insurance<sup>4</sup> was sold from January 1, 2015, to December 31, 2019. Throughout the sales period, in addition to the ever-present offline agent channel, the MA sales

<sup>3</sup>For instance, the agents need to achieve some goal(s) of downloading frequencies when the MA was released.

<sup>4</sup>Term life insurance is a life insurance policy that pays benefits on death within a fixed insured duration, such as 10 years, 20 years, or until the insured reaches a certain age. If the insured dies during the insured duration, the insurance company pays the insurance benefit; at the end of the insured duration, if the insured is still alive, the insurance company does not pay the insurance benefit and usually does not refund the insurance premium, and the insurance contract is terminated.

strategy was initially launched as an additional channel on April 24, 2018, and ended temporarily on June 11, 2018. This was a commercial trial run for selling life insurance products on the MA, only open to customers of the insurer. An adjustment for term life insurance products on the MA (i.e., suspension of term life insurance sales on the MA) was conducted from June 12, 2018 (49 days after the first launch) to August 12, 2018 (111 days after the first launch). After the adjustment period, term life insurance can be bought through both the MA and offline agent channels, and was open to all customers. The underwriting process of term life insurance is different between the MA and offline agent channels. Unlike offline agents who persuade customers and manually check policy and policyholder information (e.g., gender, age, desired premium payment term, etc.), the MA allows customers to independently choose the preferred term life insurance product and automatically check the corresponding information.

There are in total four term life insurance products sold by the insurer during the 2015–2019 sales period. In practice, the insurer offers customers the right to choose their preferred duration, insured amount, premium payment term, as well as premium payment frequency, and then calculates a specific (up-front/level) premium accordingly. In this sense, policy plan characteristics could vary with different personal choices. Although the product difference might be a factor affecting the purchase decision of the customers, it has been recorded in the literature, for example, Brown and Goolsbee (2002), that term life insurance is to some extent homogeneous. Among all the policies sold during the 5-year sales period, 32.3% of policies are sold through the MA platform, 64.7% are sold offline (e.g., by the insurer's individual or bank agents), and 3.0% through Internet websites.

Our data set is composed of insurance purchase records with both the policyholder and policy information. All the data has already been desensitized. More specifically, there are two main categories of data for each insurance purchase record: (i) the policyholder's demographic information, for example, policyholder number, birth date, gender, and profession; (ii) insurance purchase information, which includes policy number, underwriting date, product type, premiums paid, premium payment term, premium payment frequency (e.g., monthly and annually), and so on.

As noted in the last section, we are interested in the extent and manner in which the MA sales strategy affects the insurer's performance, focusing on the total number of policies sold and total premium income. In addition, there are many unobservable factors (e.g., insurance awareness) that vary from prefecture to prefecture and may also influence insurance sales. These unobservable factors cannot be ruled out by prefecture fixed effects if they are not at the prefecture level. Therefore, we aggregate the insurance purchase quantity as well as the daily PRNP at the prefecture-day level. By producing panel data at the prefecture-day level, we can analyze how the impact of the MA sales strategy changes over time. Understanding this change is also important because it can provide insight into whether the estimated effect on insurance sales is permanent or temporary.

For the insurance purchase quantity, considering the use of big data and the retention of as much data information as possible, we keep all the records except those with missing data. For the analysis of the PRNP, to make the products homogeneous, we compute premiums for the sample of term life insurance policies with no additional riders (*rider* = 0). Due to the limitation of the data structure, we are only able to compute premiums for those with policy status “still valid,” “payment fulfilled,” or “expired.” Moreover, we exclude the policy records with missing data or value zero on the premium. Therefore, the sample for the PRNP analysis is much smaller than that



for the quantity analysis, and the variable “rider” is not used in the PRNP's regression model. We also run the regressions on the quantity using the same sample as the PRNP analysis and the results are consistent.

Table 1 shows the descriptive statistics of the corresponding variables at the prefecture level. It shows that, in the sample for the analysis of the insurance purchase quantity, the average proportion of female policyholders per prefecture per day is 49.1%; the average proportion of finance-related professionals is 26.1%; the average proportion of self-insured's is 85.6%; the proportion of middle-aged and middle-educated persons is 93.6% and 79.5%, respectively. The proportion of policies with one rider per prefecture per day is 25.7%. In comparison, the sample for the PRNP analysis has a higher proportion of female policyholders (57.3%) and financial professionals (32.7%), while the remainder is similar to the sample for the quantity analysis.

### 2.3 | Trends overview

To gain an intuition about the impact of the MA sales strategy implementation on the insurance purchase quantity and PRNP, we plot the trends of the aggregated purchase quantity and PRNP across policies for each day in each year from 2015 to 2019. Figure 1 illustrates the trends in the logarithmic daily purchase quantity for the term life insurance between the years 2015 and 2019. Periods during which the term life insurance was available on the MA platform are shaded (see Figure 1). It shows that from 2015 to 2018, the trends in the daily purchase quantity for the uncovered days without the MA sales strategy are largely parallel. A substantial difference in trends emerges right after the dates of implementing the MA sales strategy in 2018, driven by the substantial growth of the insurance purchase quantity. A notable observation is that the trend difference disappears immediately after shutting the first implementation of the MA sales strategy (on June 11, 2018), implying the causal effect of the MA sales strategy. The growth effect induced by the MA sales strategy continued in 2019—the year after the MA implementation, as visually presented by a significant rise in the logarithmic daily purchase quantity. It should also be noticed that the insurance purchase quantity increases sharply on the last day of every month. This is due to the extra efforts made by salespersons in response to the monthly performance targets of the insurer. This also applies to the MA channel, as insurance agents are also encouraged to ask potential customers to purchase via the MA.<sup>5</sup> As for PRNP, similar parallel trends and clear growth in PRNP following the MA implementation can be observed in Figure 2.

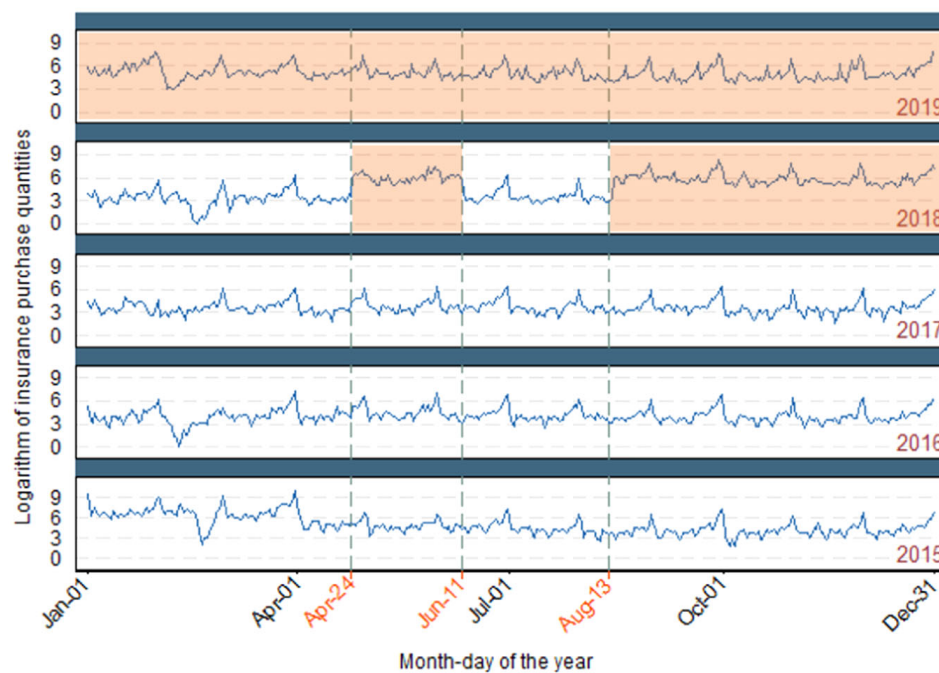
### 2.4 | Empirical strategy

To explore the impact of the MA sales strategy, we compare the insurer's sales performance before and after implementing the MA sales strategy. Recall that we use the insurance purchase quantity and PRNP as two important indicators to measure the performance of this sales strategy. These two indicators may perform differently across prefectures because of possible

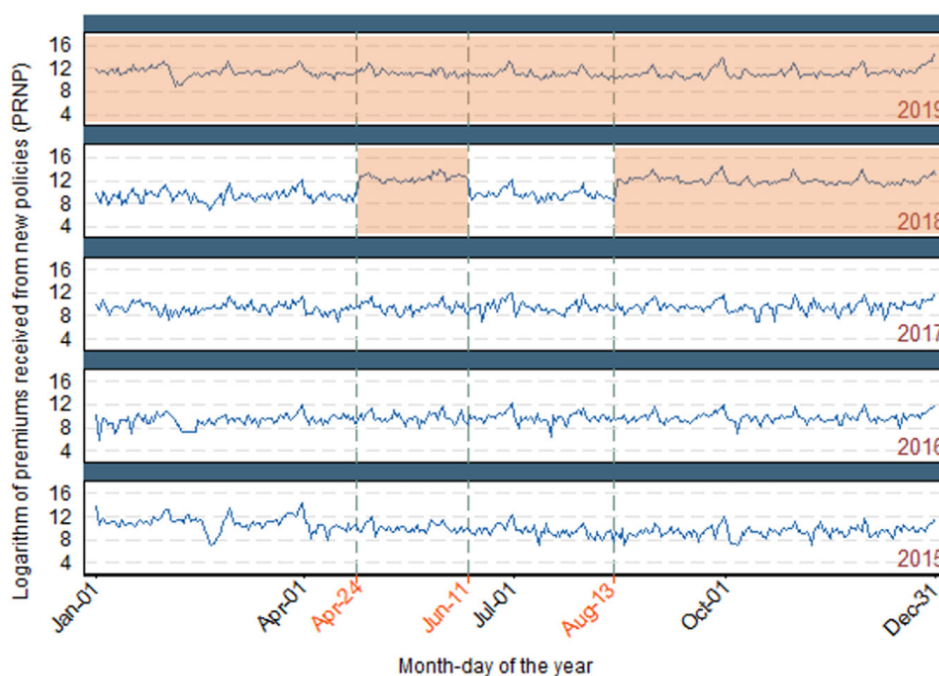
<sup>5</sup>We plotted the trend for purchases on the MA and offline, respectively, and found for both cases, spikes are observed at the end of each month.

TABLE 1 Descriptive statistics for the data sample from 2015 to 2019.

Variable name	Description	Sample for the insurance quantity analysis			Sample for the PRNP analysis		
		Observation (1)	Mean (2)	Std. Dev. (3)	Observation (4)	Mean (5)	Std. Dev. (6)
<i>log of the quantity</i>	The logarithm of the insurance purchase quantity in prefecture $i$ on day $d$ .	93,580	0.712	0.979			
<i>log of PRNP</i>	The logarithm of premiums received from new policies in prefecture $i$ on day $d$ .				51,145	7.066	1.147
<i>female</i>	The proportion of female policyholders in prefecture $i$ on day $d$ .	93,580	0.491	0.414	51,145	0.573	0.422
<i>fiprof</i>	The proportion of policyholders with finance-related occupations in prefecture $i$ on day $d$ .	93,580	0.261	0.363	51,145	0.327	0.405
<i>self</i>	The proportion of policies that are purchased for oneself in prefecture $i$ on day $d$ .	93,580	0.856	0.299	51,145	0.893	0.278
<i>midage</i>	The proportion of policyholders aged between 21 and 50 in prefecture $i$ on day $d$ .	93,580	0.936	0.210	51,145	0.953	0.187
<i>midedu</i>	The proportion of policyholders with a level of education higher than "technical secondary school," but lower than "master degree" in prefecture $i$ on day $d$ .	93,580	0.795	0.348	51,145	0.798	0.354
<i>prem_term</i>	The average premium payment term of policies in prefecture $i$ on day $d$ .	93,580	17.734	5.403	51,145	17.945	4.947
<i>sum_ins</i>	The average insured amount of policies in prefecture $i$ on day $d$ .	93,580	296,723	232,320	51,145	295,685	183,642
<i>rider</i>	The proportion of policies with a rider in prefecture $i$ on day $d$ .	93,580	0.257	0.392	51,145	0.000	0.000
<i>duration</i>	The average insurance duration of policies in prefecture $i$ on day $d$ .	93,580	20.289	5.076	51,145	20.525	4.373



**FIGURE 1** Trends in the logarithmic daily insurance purchase quantity (2015–2019), computed on a daily basis. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 2** Trends in logarithmic daily premiums received from new policies (PRNP) (2015–2019), computed on a daily basis. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



heterogeneity across prefectures in terms of average wages, gross domestic product per capita, population density, mobile Internet infrastructure, and so forth. If such heterogeneity is not fully considered, the empirical results may be biased. The two-way fixed-effects model with both time and prefecture effects can deal with the heterogeneity well, and is a widely used model to estimate the effect of a treatment on the outcome of interest (De Chaisemartin & D'Haultfoeuille, 2022). Moreover, by comparing the daily trend of the purchase quantity and PRNP (cf. Figures 1 and 2) in the same period in different years, the increased quantity and PRNP in 2018 and 2019 (after the MA implementation) are almost parallel to corresponding periods when the MA was not present. Accordingly, we make the assumption that the growth effect of the MA sales strategy is constant over time.

By aggregating the quantity and PRNP of sold policies for each prefecture on each day, we construct a panel data set of  $i = 1, \dots, N$  prefectures for  $d = 1, \dots, T$  days. The setting of the model is to examine the average difference in the insurer's sales performance before and after the implementation of the MA sales strategy. The empirical specification is given by

$$\ln(Y_{id}) = \beta_0 + \beta_1 \times App + month + year + day\_of\_month + prefecture_i \times year + \epsilon_{id}, \quad (1)$$

where  $Y_{id}$  denotes the insurance purchase quantity or PRNP in prefecture  $i$  on date  $d$ . *App* is the dummy variable of interest indicating the MA launch, which takes 1 from April 24, 2018, to June 11, 2018, or after August 13, 2018; otherwise, it takes 0. If *App* takes 0, it means the insurer keeps the old sales strategy without the MA channel. We also include the *day\_of\_month*, *month*, and *year* fixed effects: *day\_of\_month* is composed of 30- or 31-day-of-month dummies to exclude the influence from end-of-month effects; *month* consists of 12-month-of-year dummies to absorb seasonality; *year* denotes the year dummy to absorb time trends. Agencies located in different prefectures usually adjust their sales targets on an annual basis. In this sense, we use the interactive fixed effect  $prefecture_i \times year$  to absorb the possible impact of these fluctuations on our results.

### 3 | RESULTS

In this section, we start by providing the results of our model, presented in (1). It is followed by the mechanism analyses on channel accessibility and cost reduction.

#### 3.1 | Preliminary results

Table 2 presents the results of estimating the effects of MA implementation with our main specification.<sup>6</sup> The left side of the table presents the estimated effects on the insurance purchase quantity and the right side, the estimated effects on insurance PRNPs. On each side, we present results for both sample periods of 2 and 5 years, respectively, allowing for examining both the short-term and long-term effects. The introduction of the MA sales strategy, on average, drastically increases the purchases of term life insurance. The results are all statistically significant and positive for the purchase quantity. Specifically,

<sup>6</sup>The corresponding results using the same sample for the quantity and PRNP analysis are presented by Table A1 in Appendix A, which are similar to Table 2.

TABLE 2 Preliminary results of the MA sales strategy effects.

	Insurance purchase quantity		PRNP	
	(2017–2018)	(2015–2019)	(2017–2018)	(2015–2019)
	(1)	(2)	(3)	(4)
<i>App</i>	0.923*** (0.0725)	1.193*** (0.0744)	0.522*** (0.119)	0.393*** (0.130)
<i>_cons</i>	−0.199*** (0.0512)	0.796*** (0.0310)	6.934*** (0.115)	6.830*** (0.0370)
<i>prefecture × year</i>	Yes	Yes	Yes	Yes
<i>time effect<sup>a</sup></i>	Yes	Yes	Yes	Yes
<i>N</i>	30,068	93,580	19,127	51,145
Adjusted <i>R</i> <sup>2</sup>	0.469	0.403	0.369	0.340

Note: Standard errors in parentheses are clustered on the prefecture level.

Abbreviations: MA, mobile application; PRNP, premiums received from new policy.

<sup>a</sup>Time effect includes the fixed effects of *month*, *year*, and *day\_of\_month*.

\**p* < 0.1.

\*\**p* < 0.05.

\*\*\**p* < 0.01.

Columns (1) and (2) show that daily insurance purchase quantity rises by 151.7% ( $=e^{0.923} - 1$ ) and 229.7% ( $=e^{1.193} - 1$ ) after implementing the MA sales strategy for the 2- and 5-year sample, respectively. Positive effects of 68.5% ( $=e^{0.522} - 1$ ) and 48.1% ( $=e^{0.393} - 1$ ) growth, presented in Columns (3) and (4), respectively, are also found for the daily PRNP during the sample periods of 2 and 5 years.

There are two noteworthy observations from the above results. First, the effect of the MA sales strategy on the insurance purchase quantity is substantially larger than the effect on PRNP. One possible source is cost reduction. With the widespread use of the Internet, reduced search cost (Goldfarb & Tucker, 2019) tends to push down the price of term life insurance (Brown & Goolsbee, 2002). Another likely explanation is flexible underwriting conditions for offline sales. In practice, customers need to fulfill the health disclosure before underwriting.<sup>7</sup> For offline insurance sales, insurance agents usually are allowed to underwrite certain risks above health requirements with additional fees. However, this is not the case for online insurance sales where it is not possible to flexibly add fees. Due to this fact, the received premium per policy through the MA channel may be lower than the offline channel, which, in turn, leads to a lower positive impact of MA sales strategy on PRNP than on the insurance purchase quantity. To examine the decline in the received premium per policy and clarify the corresponding reasons, we provide further analysis in Section 3.2.

The second observation is the contrast that the growth effect estimated by the 5-year sample is higher than that estimated by the 2-year sample for the purchase quantity, while it is reversed for PRNP. Specifically, for the insurance purchase quantity, the growth effect of the MA launch from the 5-year sample is 78 percentage points higher than that from the 2-year sample ( $=229.7\% - 151.7\%$ ), while for PRNP, the growth effect of the MA launch from the 5-year sample

<sup>7</sup>The health condition requirement for term life insurance is not as strict as for health insurance.

is 20 percentage points lower than that from the 2-year sample (=48.1%–68.5%). The probable reason is that the 5-year sample period covers a longer period of the second MA launch (e.g., the year 2019) than the 2-year sample period. More new customers, who might prefer a lower insured amount, gain access to purchasing insurance on the MA through the second launch; while the target customers of the trial run are regular customers who have established a long-term trust relationship with the insurer and are therefore willing to buy insurance with a relatively higher insured amount.

To identify the source of the insurance purchase growth, we tabulate the average insurance purchase quantity and PRNP per prefecture per day for different periods during the 2-year sample period (see Table 3). Column (1) provides the average values of the period when the MA sales strategy was not implemented, Column (2) is after the first launch while before the adjustment, Column (3) is the adjustment period, and Column (4) delivers the values after the second launch. After the first MA launch, the insurance purchase quantity through the MA channel grows from nothing to an average of 5.587 per prefecture per day. After the second MA launch, the increase in the quantity through the MA strategy is again comparably large (from 0 to 4.709), but smaller than the increase caused by the first MA launch. In this period, the offline purchase quantity increases slightly too, from 2.282 to 2.814. The changes in PRNP in these four periods are expressed in the bottom half of Table 3. Different from the quantity change, after the first launch (Column 2), the average PRNP per prefecture per day on the MA is higher than that on the offline channel, while that on the MA becomes lower than that on the offline channel after the second launch (Column 4). This may be the result of a combination of quantity and price changes. Further explorations are provided in Section 3.2.

### 3.2 | Mechanism analysis

In the previous analysis, we have demonstrated that the MA implementation leads to significant growth in both the insurance purchase quantity and PRNPs for term life insurance. In this section, we explore the mechanisms explaining these positive effects.

**TABLE 3** Average insurance purchase quantity and PRNP per prefecture per day for different periods during the 2-year sample period (2017–2018).

	Jan 1, 2017–Apr 24, 2018 (1)	Apr 24, 2018–Jun 11, 2018 (2)	Jun 11, 2018–Aug 13, 2018 (3)	Aug 13, 2018–Dec 31, 2018 (4)
<i>Average insurance purchase quantity per prefecture per day</i>				
MA	0	5.587	0	4.709
Offline	2.324	2.044	2.282	2.814
<i>Average PRNP per prefecture per day</i>				
MA	0	2556.207 <sup>a</sup>	0	2467.653
Offline	2236.13	2226.256	2125.649	2704.872

Abbreviations: MA, mobile application; PRNP, premiums received from new policy.

<sup>a</sup>Due to the fact that the insurer received a few huge premiums during this period, which makes the average value much higher than in other periods. To exclude this effect, we calculate the average value for all the values within the 95% quantile during this sample period.

### 3.2.1 | Channel accessibility

Faster Internet connection promotes insurance sales and increases the demand for insurance (Butler, 2021). Similarly, increasing accessibility by the MA channel saves customers' time and effort involved in searching for a term life insurance product (H. Li et al., 1999). In this sense, prefectures with higher channel accessibility and better facilitation should correspond to greater increases in insurance purchases after the MA implementation, therefore leading to larger growth effects on insurance purchases. To examine this, we apply the Internet penetration ratio, the cellphone ratio, and the digitization index as proxies to measure increased channel accessibility by the MA sales strategy for each prefecture. The prefecture-year data of Internet penetration ratio and cellphone ratio are obtained from Statistics Bureau of the P.R. China (2020), and the prefecture-year data of digitization index is from Digital Finance Research Center of Peking University (2021). We can detect how the increased channel accessibility influences the growth effects of the MA sales strategy introduction by adding interaction terms between the accessibility proxies and the independent variable *App* into the baseline model as follows:

$$\ln(Y_{id}) = \beta_0 + \beta_1 \times App + \beta_2 \times App \times access_{iy}^{(j)} + month + year + day\_of\_month + prefecture_i \times year + \epsilon_{id}, \quad (2)$$

where  $access_{iy}^{(j)}$ ,  $j = 1, 2, 3$ , denotes three different proxies for channel accessibility of prefecture  $i$  in year  $y$ , that is, the Internet penetration ratio, the cellphone ratio, and the digitization index.  $\beta_2$  measures how much channel accessibility affects the growth in insurance purchases given the MA sales strategy ( $App = 1$ ).

Tables 4 and 5 report the results of the channel effects on insurance purchase quantity and PRNP. Overall, the estimated coefficients of the intersection term (the second row) are all significant and, as anticipated, indicating that the higher the channel accessibility, the greater the growth in the insurance purchase quantity and PRNP of the term life insurance. The magnitudes of the positive coefficients of the intersection terms differ when various proxies are used, which is due to the different scales of the proxies.<sup>8</sup>

### 3.2.2 | Cost reduction

As discussed in the previous text, the much larger increase in the insurance purchase quantity as compared with PRNP may result from the reduction in price. In turn, the price reduction could also be a driver for the found growth effect. When it comes to insurance, the price consists of the actuarially fair premium and cost. The actuarially fair premium is computed based on the mortality expectation of each risk group as well as the interest rate expectation. In other words, the actuarially fair premium should be consistent within a

<sup>8</sup>In the 5-year sample of quantity analysis, the average Internet penetration ratio per prefecture per day is 0.369, the cellphone ratio per prefecture per day is on average 1.621, and the corresponding digitization index is averagely 266.742; while the 5-year sample for PRNP analysis has higher average values of these three proxies, that is, 0.415, 1.767, and 276.650, respectively. As for the 2-year sample of quantity analysis, the average values per prefecture per day are 0.413, 1.748, and 278.689; and for the 2-year sample of PRNP analysis, they are 0.427, 1.801, and 282.688, respectively.

TABLE 4 Estimation results of channel accessibility effects: insurance purchase quantity.

	<i>access</i> <sup>(1)</sup> = <i>Internet_ratio</i>		<i>access</i> <sup>(2)</sup> = <i>Cellphone_ratio</i>		<i>access</i> <sup>(3)</sup> = <i>Digitization_index</i>	
	(2017–2018)	(2015–2019)	(2017–2018)	(2015–2019)	(2017–2018)	(2015–2019)
	(1)	(2)	(3)	(4)	(5)	(6)
<i>App</i>	0.397*** (0.0973)	0.687*** (0.0995)	0.580*** (0.126)	0.866*** (0.131)	−4.260*** (1.088)	−3.736*** (1.081)
<i>App</i> × <i>access</i> <sup>(i)</sup>	1.176*** (0.187)	1.140*** (0.182)	0.180*** (0.0658)	0.173*** (0.0643)	0.0179*** (0.00385)	0.0171*** (0.00382)
<i>_cons</i>	0.175*** (0.0409)	1.222*** (0.0330)	0.175*** (0.0403)	1.222*** (0.0328)	0.175*** (0.0393)	1.226*** (0.0321)
<i>prefecture</i> × <i>year</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>time effect</i> <sup>a</sup>	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	28,797	88,167	28,814	88,382	29,302	90,490
Adjusted <i>R</i> <sup>2</sup>	0.481	0.411	0.482	0.411	0.478	0.407

Note: Standard errors in parentheses are clustered on the prefecture level.

<sup>a</sup>Time effect includes the fixed effects of *month*, *year*, and *day\_of\_month*.

\**p* < 0.1.

\*\**p* < 0.05.

\*\*\**p* < 0.01.

TABLE 5 Estimation results of channel accessibility effects: PRNP.

	<i>access</i> <sup>(1)</sup> = <i>Internet_ratio</i>		<i>access</i> <sup>(2)</sup> = <i>Cellphone_ratio</i>		<i>access</i> <sup>(3)</sup> = <i>Digitization_index</i>	
	(2017–2018)	(2015–2019)	(2017–2018)	(2015–2019)	(2017–2018)	(2015–2019)
	(1)	(2)	(3)	(4)	(5)	(6)
<i>App</i>	−0.260*	−0.386***	−0.0116	−0.137	−7.345***	−7.499***
	(0.139)	(0.136)	(0.163)	(0.171)	(2.007)	(1.999)
<i>App</i> × <i>access</i> <sup>(i)</sup>	1.586***	1.588***	0.231***	0.231***	0.0269***	0.0270***
	(0.239)	(0.236)	(0.0817)	(0.0808)	(0.00705)	(0.00704)
<i>_cons</i>	7.738***	6.975***	7.751***	6.980***	7.746***	6.991***
	(0.0683)	(0.0367)	(0.0677)	(0.0366)	(0.0679)	(0.0363)
<i>prefecture</i> × <i>year</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>time effect</i> <sup>a</sup>	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	18,428	48,587	18,440	48,770	18,687	49,878
Adjusted <i>R</i> <sup>2</sup>	0.378	0.349	0.380	0.349	0.376	0.344

Note: Standard errors in parentheses are clustered on the prefecture level.

Abbreviation: PRNP, premiums received from new policy.

<sup>a</sup>Time effect includes the fixed effects of *month*, *year*, and *day\_of\_month*.

\**p* < 0.1.

\*\**p* < 0.05.

\*\*\**p* < 0.01.

risk group for term life insurance policies that have the same duration. Hence, the price reduction in term life insurance in fact refers to the decrease in cost. In this subsection, we will decompose the total premium of each policy into the actuarially fair premium and cost.

First, we construct the total premium ( $\hat{P}^{(0)}$ ) at the time of insurance purchasing (i.e., day  $d$ ) by

$$\hat{P}_d^{(0)} = \sum_{t=0}^{T_1-1} \text{Level\_premium} \cdot {}_t p_x \cdot v^t, \quad (3)$$

where we have used the interest rate  $i$  to compute the discount factor  $v (= \frac{1}{1+i})$ .  $T_1$  is the premium payment term. *Level\_premium* is paid at the beginning of each year  $t$ . The actuarially fair premium  $P^{(0)}$  of the term life insurance contract paying upon death with a duration  $T_2$  is given by

$$P^{(0)} = \sum_{t=1}^{T_2} v^t \cdot {}_{t-1} p_x \cdot q_{x+t-1} \cdot F. \quad (4)$$

Note that the insured duration of the policy  $T_2$  can differ from the level premium payment term  $T_1$ . Further,  $F$  marks the face value of the policy, that is, the insured amount.  ${}_{t-1} p_x$  is the  $(t-1)$ -year survival probability of an  $x$ -year-old, and  $q_{x+t-1}$  represents the probability of an  $(x+t-1)$ -year-old dies within 1 year. We do not have the insurer's annual mortality rates for each risk group. In this case, we obtain the mortality rates from the experienced life tables of different genders with reference to the industry standards (China Banking and Insurance Regulatory Commission, 2016). The interest rate is assumed to be a constant 2% referring to the 5-year average growth rate of CPI in China (Statistics Bureau of the P.R. China, 2020).

And by subtracting the actuarially fair premium from the total premium, we obtain the cost for each policy, that is,

$$\text{Cost}_{idj} = \hat{P}_{idj}^{(0)} - P_j^{(0)}. \quad (5)$$

$\hat{P}_{idj}^{(0)}$  is the total premium of policy  $i$  in risk group  $j$  on day  $d$ , while  $P_j^{(0)}$  represents the actuarially fair premium of risk group  $j$ . Since here we are interested in the price change and need to estimate the cost per policy, we use policy-level data for our analysis. To explore the impact of the MA sales strategy on the cost rate ( $\frac{\text{Cost}}{P^{(0)}}$ ), we have carried out the following pooled ordinary least squares regression. The regression specification is given by

$$\begin{aligned} \text{Cost\_rate}_{idj} &= \frac{\text{Cost}_{idj}}{P_j^{(0)}} \\ &= \beta_0 + \beta_1 \times \text{if\_on\_App} + \text{month} + \text{year} + \text{day\_of\_month} \\ &\quad + \text{prefecture}_i \times \text{year} + \epsilon_{idj}. \end{aligned} \quad (6)$$

Here, instead of using the previous date-distinguished explanatory variable *App*, we distinguish the channel by *if\_on\_App*, as we want to explore the cost difference between



purchasing on the MA and purchasing offline. It takes 1 if the policy is sold through the MA, and 0 otherwise.

For simplicity, we look at the premiums of the policies purchased by the 20-, 25-, 30-, 35-, 40-, 45-, 50-, and 55-year olds for 32 possible combinations of the gender, age, and selected insured duration.<sup>9</sup> We keep the policies that are still “valid” for premium calculation. Moreover, we exclude the samples with riders because data limitations prevent us from separating the premiums or insured amounts for primary insurance and riders. Further, we only keep the policies that are purchased for oneself, as we need the insured's information for computing the actuarially fair premium.<sup>10</sup>

The corresponding results are presented in Table 6, from which we can see that the cost rate is largely reduced through the MA channel. From Columns (1) and (2) in Table 6, the cost rate decreases by roughly 0.940 and 0.909 in the value, respectively. With the average cost rate of non-MA channel policies (0.980 during the period of 2017–2018 and 1.089 during the period of 2015–2019<sup>11</sup>) as a basis, we find that the magnitude of the decrease in the estimated cost rate is substantial. This is reasonable, as agent commissions are usually high,<sup>12</sup> which are embedded into the insurance product price undertaken by customers. The MA channel emerges after the implementation of the MA sales strategy, such that a lower cost rate of the term life insurance on the MA platform leads to a lower increase in PRNP after the implementation.

## 4 | ROBUSTNESS TESTS

In this section, we address concerns about causality by providing evidence supporting our main results and excluding other possible inherent factors that could lead to the growth effects we find.

### 4.1 | More controls

There is an assumption implied by our empirical strategy specified in Section 2.4, that is, when fixing the time and prefecture effects, the MA sales strategy is the only factor that will affect the purchase quantity and PRNP. To verify this assumption as well as eliminate the concern of missing variables, we add more control variables to identify possible sources of impacts.

- *holiday*: There are two reasons for adding this control. The first one is that we observe obvious spikes during holidays in Figures 1 and 2. Taking the National Day of China, for

<sup>9</sup>Since most policies have a duration of 20 or 30 years, we selected a sample with a 20- or 30-year duration for our analysis.

<sup>10</sup>As cost rates are estimated using industry-standard mortality tables and the assumed interest rate, which may differ from the insurer's actuarial assumptions, there may be some estimation bias in the estimates relative to the true values. Approximately 1% of our observations has a cost rate over 2, which are considered outliers and are excluded from our analysis sample.

<sup>11</sup>In our case, the estimated cost rate includes: (i) sales costs, such as acquisition, administrative, and distribution charges; (ii) the risk loading determined by the insurer in addition to the mortality risk embedded in the industry-standard mortality table; and (iii) deviations from the true cost rate due to interest rates we use that may differ from the insurer's interest rate assumptions. These explain why the average estimated cost rate presents almost 100% of the actuarially fair premium.

<sup>12</sup>According to the insurance regulatory regulation (China Banking and Insurance Regulatory Commission, 2011), the commissions for the agents could be as high as 4%–5% of the total premium.



**TABLE 6** Results of the MA sales strategy effects on cost rate.

	(2017–2018) (1)	(2015–2019) (2)
<i>if_on_App</i>	−0.940*** (0.00964)	−0.909*** (0.00686)
<i>_cons</i>	0.825*** (0.229)	1.731*** (0.167)
<i>prefecture × year</i>	Yes	Yes
<i>time effect<sup>a</sup></i>	Yes	Yes
<i>N</i>	11,388	25,686
Adjusted <i>R</i> <sup>2</sup>	0.742	0.790

Abbreviation: MA, mobile application.

<sup>a</sup>Time effect includes the fixed effects of *month*, *year*, and *day\_of\_month*.

\**p* < 0.1.

\*\**p* < 0.05.

\*\*\**p* < 0.01.

example, the insurance purchase quantity and PRNP on October 1 of each year is higher than in previous months. The second reason is that “*App*,” the date variable of interest indicating the period before and after introducing the MA sales strategy, may be correlated with holidays and thus bias our estimates. A negative correlation would lead to underestimation bias while a positive correlation would lead to overestimation bias. For example, a potential source of the negative correlation is that there are different holidays in the periods with and without the MA sales strategy. Specifically, for the 2015–2019 sample period, the period before the introduction of MA (from January 1, 2015, to April 24, 2018, and from June 16, 2018, to August 12, 2018) has more holidays than the period after the introduction of MA (from April 24, 2018, to June 15, 2018, and after August 13, 2018).

- *new\_version*: As mentioned, the existence of the adjustment period divides MA insurance sales into two periods. In this case, we add a variable “*new\_version*” to absorb the difference between the two launches of the MA sales strategy.
- *Additional control variables*: As our analysis is based on also prefecture-day level, we need prefecture-day variables to capture the time-variant characteristics of different prefectures. This is different from using the “*prefecture<sub>i</sub>*” as fixed effects to control the prefecture-invariant characteristics. Here, we compute the proportion of policyholder's characteristics (e.g., the percentage of females, financial professions, the self-insured, the middle-aged, and the middle-education-level) and policy features (e.g., premium payment, insured amount, insurance duration, and percentage of policies with a rider/riders<sup>13</sup>) in each prefecture on each day. These prefecture-day feature proportions serve as proxies to describe the time-variant characteristics of prefectures.

<sup>13</sup>This variable is only for the quantity analysis, as the sample of PRNP analysis has excluded the records with riders to make the product homogeneous.

The coefficients of *holiday* are estimated to be significantly positive for both quantity and PRNP analysis, which shows that holidays lead to increases in the policy quantity and PRNP. This is in line with our observation of spikes during holidays. A possible explanation is that offline agents in China on average make more sales per day during holidays than on weekdays. This has two reasons. The first is that as insurance companies in China usually do not carry out strict attendance assessment on offline agents but mainly rely on performance assessment, offline agents in China usually work flexible hours, rather than fixed 8-h schedules typical in developed countries. The second is that consumers are more available to meet offline agents during holidays than on weekdays, which prompts insurance agents to work on holidays. Because offline insurance sales services are usually conducted face-to-face, requiring consumers' availability due to necessary transportation and meeting. After separating the effect of holidays from the overall growth effect, the current estimates of the MA introduction (coefficients of *App* in Tables 7 and 8) are closer to the estimates of the actual growth effect. The positive signs of them validate the robustness of our preliminary results.

Additionally, the estimated coefficients of *new\_version* are insignificantly positive in the analysis of quantity, while significantly negative in the analysis of PRNP. This indicates that the growth effect on the purchase quantity has no significant difference between the first and second launch; while for the analysis of PRNP, the negative coefficients of *new\_version* do not mean that the MA implementation leads to a decrease in PRNP, but rather that the growth effect by the MA is higher during the first than the second launch period. This may be due to the reduced price of the term life insurance on the MA after the second launch. In general, adding more control variables does not change the fact that the MA sales strategy results in significant increases in the insurance purchase quantity and PRNP, suggesting that our preliminary results are robust.

## 4.2 | Pretrends

Another major concern with our identification strategy might be the possible inherent growth in insurance purchases for the days before the MA introduction in 2018, which could also yield significant estimates even without the introduction of the MA sales strategy.

To address this, we examine the effects of leads and lags of the first launch date of the MA sales strategy as illustrated in the equation below:

$$\ln(Y_{id}) = \sum_{j=-100}^{200} \beta_j \mathbb{1}_{\{K_{id}=j\}} + \text{holiday} + \text{new\_version} + \text{month} + \text{year} + \text{day\_of\_month} + \text{prefecture}_i \times \text{year} + X_{id} + \epsilon_{id}, \quad (7)$$

where the subscript  $j$  denotes  $j$  days after the first implementation date of the MA sales strategy if positive and  $-j$  days before the first implementation if negative.  $K_{id}$  denotes the relative distance between the date  $d$  and the date of the first insurance purchase on the MA platform in prefecture  $i$ .  $\mathbb{1}_{\{K_{id}=j\}}$  is a dummy equal to 1 if  $K_{id} = j$ , otherwise 0. The coefficients of interest  $\beta_j$  indicate pretrends and dynamic effects of the MA sales strategy.  $X_{id}$ 's are the policyholder-related and policy-related control variables per prefecture per day. We set the investigated sample period as 2017–2018 for regression.

The results of the quantity analysis with confidence levels are graphed in Figure 3. Consistent with the main results, the estimates before the first implementation date and during

TABLE 7 Robustness test for the insurance purchase quantity with more controls.

	Insurance purchase quantity			
	(2017–2018)	(2015–2019)	(2017–2018)	(2015–2019)
	(1)	(2)	(3)	(4)
<i>App</i>	0.892*** (0.0782)	1.147*** (0.0712)	0.835*** (0.0735)	1.040*** (0.0667)
<i>holiday</i>	0.0958*** (0.0164)	0.115*** (0.0103)	0.0873*** (0.0164)	0.109*** (0.0103)
<i>new_version</i>	0.0409 (0.0399)	0.0582 (0.0396)	0.0886** (0.0402)	0.0995** (0.0400)
<i>female</i>			−0.0236** (0.0116)	−0.00312 (0.00747)
<i>fiprof</i>			0.0427*** (0.0127)	0.0540*** (0.00961)
<i>self</i>			0.0243 (0.0147)	0.0870*** (0.00975)
<i>midage</i>			0.0244 (0.0157)	0.0441*** (0.0110)
<i>midedu</i>			0.0284 (0.0185)	0.0400*** (0.0150)
<i>rider</i>			−0.0299* (0.0165)	−0.151*** (0.0124)
<i>prem_term</i>			−0.0128*** (0.00122)	−0.00874*** (0.000904)
<i>sum_ins</i>			−5.45e−08** (2.25e−08)	−0.000000151*** (2.06e−08)
<i>duration</i>			0.00995*** (0.00119)	0.00692*** (0.00108)
<i>_cons</i>	−0.201*** (0.0514)	0.793*** (0.0311)	−0.264*** (0.0548)	0.741*** (0.0386)
<i>prefecture × year</i>	Yes	Yes	Yes	Yes
<i>time effect<sup>a</sup></i>	Yes	Yes	Yes	Yes
<i>N</i>	30,068	93,580	30,068	93,580
Adjusted $R^2$	0.470	0.405	0.474	0.413

Note: Standard errors in parentheses are clustered on the prefecture level.

<sup>a</sup>Time effect includes the fixed effects of *month*, *year*, and *day\_of\_month*.

\* $p < 0.1$ .

\*\* $p < 0.05$ .

\*\*\* $p < 0.01$ .

TABLE 8 Robustness test for PRNP with more controls.

	PRNP			
	(2017–2018)	(2015–2019)	(2017–2018)	(2015–2019)
	(1)	(2)	(3)	(4)
<i>App</i>	0.694*** (0.121)	0.660*** (0.128)	0.515*** (0.122)	0.519*** (0.129)
<i>holiday</i>	0.173*** (0.0307)	0.211*** (0.0174)	0.166*** (0.0301)	0.202*** (0.0161)
<i>new_version</i>	−0.371*** (0.0576)	−0.408*** (0.0441)	−0.194*** (0.0547)	−0.237*** (0.0426)
<i>female</i>			−0.540*** (0.0225)	−0.525*** (0.0151)
<i>fiprof</i>			0.0690*** (0.0198)	0.101*** (0.0116)
<i>self</i>			−0.366*** (0.0306)	−0.331*** (0.0172)
<i>midage</i>			−0.196*** (0.0549)	0.0684* (0.0363)
<i>midedu</i>			0.0835*** (0.0228)	0.115*** (0.0161)
<i>prem_term</i>			−0.0333*** (0.00187)	−0.0314*** (0.00107)
<i>sum_ins</i>			0.00000141*** (0.000000100)	0.00000142*** (8.17e−08)
<i>duration</i>			0.0359*** (0.00206)	0.0288*** (0.00125)
<i>_cons</i>	6.798*** (0.112)	6.788*** (0.0378)	6.831*** (0.131)	6.839*** (0.0517)
<i>prefecture × year</i>	Yes	Yes	Yes	Yes
<i>time effect<sup>a</sup></i>	Yes	Yes	Yes	Yes
<i>N</i>	19,127	51,145	19,127	51,145
Adjusted <i>R</i> <sup>2</sup>	0.373	0.347	0.450	0.425

Note: Standard errors in parentheses are clustered on the prefecture level.

Abbreviation: PRNP, premiums received from new policy.

<sup>a</sup>Time effect includes the fixed effects of *month*, *year*, and *day\_of\_month*.

\**p* < 0.1.

\*\**p* < 0.05.

\*\*\**p* < 0.01.

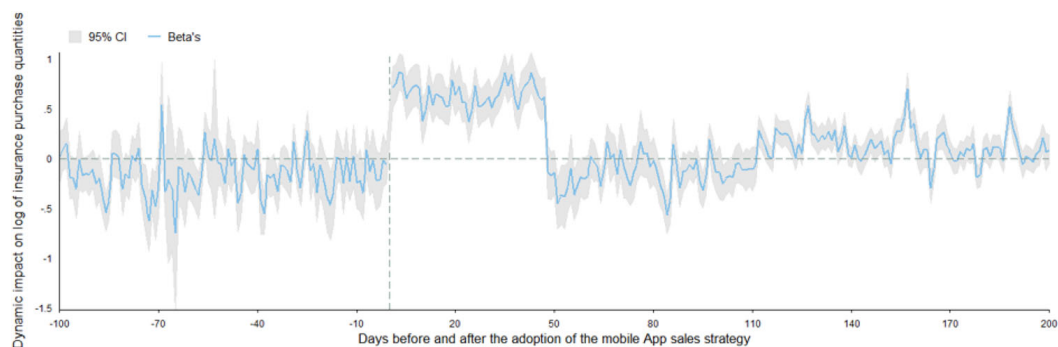
the adjustment period are statistically insignificant, but the growth effects on the insurance purchase quantity are significant for most days after MA implementation. The results of PRNP, graphed in Figure 4, also comply with preliminary results, but indicate lower growth induced by the MA sales strategy than in Figure 3.

Overall, we verify that there are no pretrends that could lead to the found growth effects of the MA sales strategy.

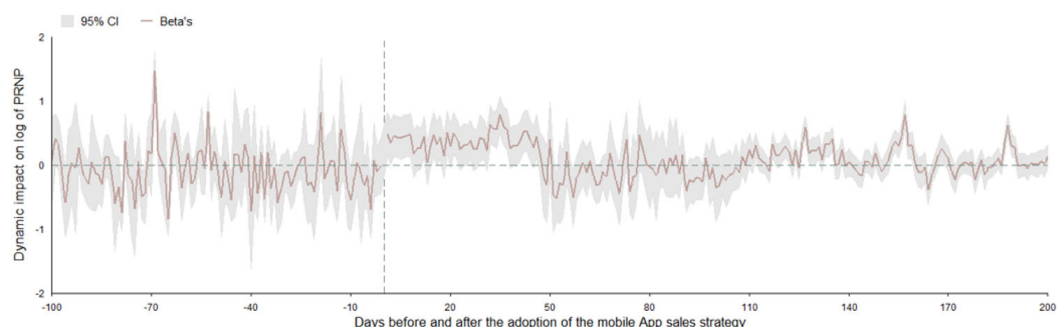
### 4.3 | Placebo tests

Insurance purchases could grow due to some unobservable reasons other than the MA sales strategy. To exclude this factor, we consider placebo tests by repeating the same regressions on samples from the years without the MA sales strategy.

For example, we regress on the subsample of 2016–2017 (the first 2 years before the introduction of the MA sales strategy). In this case, we essentially assume a virtual MA introduction on April 24, 2017, while the real MA introduction date is April 24, 2018. The results would be significantly positive if there was indeed inherent growth on the introduction date of MA implementation. In practice, we conduct placebo tests on 2-year subsamples (2016–2017, 2015–2016, and 2014–2015) and 5-year subsamples (2013–2017 and 2012–2016), respectively.



**FIGURE 3** Results of the 2-year sample (2017–2018) using event study method: insurance purchase quantity. CI, confidence interval. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 4** Results of the 2-year sample (2017–2018) using event study method: PRNP. CI, confidence interval; PRNP, premiums received from new policy. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

TABLE 9 Robustness tests using 2-year subsamples.

	Insurance purchase quantity			PRNP		
	(2016–2017)	(2015–2016)	(2014–2015)	(2016–2017)	(2015–2016)	(2014–2015)
	(1)	(2)	(3)	(4)	(5)	(6)
<i>App_vir</i> <sup>a</sup>	−0.0291 (0.0286)	−0.0848** (0.0366)	−0.716*** (0.0209)	−0.0609** (0.0307)	−0.0548 (0.0483)	−0.0140 (0.0574)
<i>_cons</i>	0.0696 (0.0426)	0.609*** (0.0492)	1.008*** (0.0496)	7.049*** (0.0583)	6.906*** (0.0724)	6.747*** (0.0608)
<i>prefecture</i> × <i>year</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>time effect</i> <sup>b</sup>	Yes	Yes	Yes	Yes	Yes	Yes
<i>additional controls</i> <sup>c</sup>	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	21,483	23,000	80,936	6789	6083	27,827
Adjusted <i>R</i> <sup>2</sup>	0.421	0.431	0.493	0.278	0.263	0.383

Note: Standard errors in parentheses are clustered on the prefecture level.

Abbreviation: PRNP, premiums received from new policy.

<sup>a</sup>*App\_vir* represents a dummy variable. Taking the subsample from 2016 to 2017, for example, *App\_vir* takes 1 between April 24, 2017 and June 11, 2017; and it also takes 1 starting from August 13, 2017; otherwise, it takes 0.

<sup>b</sup>Time effect includes the fixed effects of *month*, *year*, and *day\_of\_month*.

<sup>c</sup>Additional controls refer to *holiday*, *new\_version*, as well as the policyholder-related and policy-related control variables. Here, the variable *new\_version* is also constructed virtually in accordance with the sample window. For instance, it takes 1 after August 13, 2017 for this 2016–2017 subsample; otherwise 0.

\**p* < 0.1.

\*\**p* < 0.05.

\*\*\**p* < 0.01.

The results, presented in Tables 9 and 10, allay the concern over an inherent growth on the MA launch date, strengthening the evidence supporting the exogeneity of the MA implementation.<sup>14</sup> In other words, the growth effects we find only occurred in the MA implementation year 2018, not in any other years.

## 5 | HETEROGENEITY ANALYSIS

It has been also well documented that different customers with various characteristics respond differently to the MA sales strategy. Typically documented characteristics include gender (Chan et al., 2015), age (Nagra & Gopal, 2013), education background (Burke, 2002),

<sup>14</sup>One of the 5-year placebo tests (2013–2017) shows a significantly positive though small estimated coefficient. This suggests inherent growth on the MA launch date or a specific shock that occurred before the virtual implementation date (i.e., April 24, 2016). To further examine this thesis, we perform the same event study specification as in 7 using this 2013–2017 sample. The results, presented in Appendix B, clearly show that there was no sudden positive shock on the day of the virtual implementation date but there was such a shock roughly 20 days before the virtual implementation date. The latter leads to the significantly positive result for the 2013–2017 placebo test. On the virtual implementation date, April 24, 2016, no significant increase is observed, while some significant increases are observed afterward.

TABLE 10 Robustness tests using 5-year subsamples.

	Insurance purchase quantity		PRNP	
	(2013–2017)	(2012–2016)	(2013–2017)	(2012–2016)
	(1)	(2)	(3)	(4)
<i>App_vir</i> <sup>a</sup>	0.131*** (0.0239)	−0.779*** (0.0202)	−0.00311 (0.0320)	0.00497 (0.0504)
<i>_cons</i>	0.820*** (0.0377)	0.474*** (0.0333)	6.576*** (0.0469)	6.424*** (0.0474)
<i>prefecture</i> × <i>year</i>	Yes	Yes	Yes	Yes
<i>time effect</i> <sup>b</sup>	Yes	Yes	Yes	Yes
<i>additional controls</i> <sup>c</sup>	Yes	Yes	Yes	Yes
<i>N</i>	143,682	166,794	47,809	54,646
Adjusted <i>R</i> <sup>2</sup>	0.445	0.480	0.374	0.364

Note: Standard errors in parentheses are clustered on the prefecture level.

Abbreviation: PRNP, premiums received from new policy.

<sup>a</sup>*App\_vir* represents a dummy variable. For the 5-year subsamples, taking the one from 2013 to 2017, for example, *App\_vir* takes 1 between April 24, 2016 and June 11, 2016; and it also takes 1 starting from August 13, 2016; otherwise, it takes 0.

<sup>b</sup>Time effect includes the fixed effects of *month*, *year*, and *day\_of\_month*.

<sup>c</sup>Additional controls refer to *holiday*, *new\_version*, as well as the policyholder-related and policy-related control variables. Here, the variable *new\_version* is also constructed virtually in accordance with the sample window. For instance, it takes 1 after August 13, 2016 for this 2013–2017 subsample; otherwise 0.

\**p* < 0.1.

\*\**p* < 0.05.

\*\*\**p* < 0.01.

and profession (Nagra & Gopal, 2013). In addition, whether one is buying for oneself or someone else makes a difference too (Hansen & Jensen, 2009). In this section, we explore the heterogeneous effects of the introduction of the MA sales strategy among different genders, professions, types of the insured (i.e., whether the insurance policy is purchased for oneself), education levels, and ages.

We categorize data samples by different genders, professions, types of the insured (oneself or others), age groups, and education levels. We use dummy indicators to identify these various groups and limit the analyses to the 2-year sample. More specifically, “Indicator” refers to the group indicators of different gender groups (i.e., female [=1] and male [=0]), profession groups (i.e., finance-related [=1] and nonfinance-related [=0]), insured groups (i.e., oneself [=1] and others [=0]). As for age groups, we divide policyholders into nine age groups, that is, below 20, 20–25, 26–30, ..., 51–55, and above 55 years old. We estimate by adding interaction terms between *App* and each age group dummy. For education level groups, we have corresponding dummies with respect to primary school and below, secondary school, technical secondary school, and so forth. We estimate by adding interaction terms between *App* and each education level group dummy.

The estimated coefficients of interaction terms (“*App* × *Indicator*” or “*App* × *Dummies*”) are values of interest. For the heterogeneity analyses on gender, profession, or the insured's



type, we directly use the interaction term to capture the growth effect difference between the two groups. For the heterogeneity analyses on age or education level, as there are more than two age (or education level) groups, in operation we estimate the growth effect for each group by adding all interaction terms. Moreover, for all heterogeneity analyses, we also add interaction terms between group indicators and control variables to take into account the differences in coefficients of control variables across groups.

The estimated results for different groups (of gender, professions, and types of the insured) are displayed in Table 11. Columns (1)–(3) of Table 11 present the estimated coefficients for the insurance purchase quantity. We find that the MA sales strategy introduction induces higher growth in the insurance purchase quantity of females, of finance-related professionals, and of those purchasing for self-insurance. Columns (4)–(6) show the corresponding results for PRNP. The findings are similar to those from the quantity analysis, except that the estimate for professions shows no significance for PRNP, which indicates the MA impact on PRNP does not differ much by profession.

Reasons behind these findings may be: (i) Females are more prone to promotion activities than males (e.g., Harmon & Hill, 2003; Kwon & Kwon, 2007); (ii) As for the type of the insured, when applying for the term life insurance, a variety of information is usually required and people know more about their own needs than they do about the needs of others. In this case, the MA sales strategy may generate a higher motivation for policyholders who buy insurance to insure themselves than for policyholders who buy insurance to insure others; (iii) Policyholders with finance-related

**TABLE 11** Heterogeneity of effects on the insurance purchase quantity and PRNP w.r.t. gender, profession, and the type of the insured.

	Insurance purchase quantity			PRNP		
	Gender	Profession	Type of the insured	Gender	Profession	Type of the insured
	(1)	(2)	(3)	(4)	(5)	(6)
<i>App</i> × <i>Indicator</i> <sup>a</sup>	0.321*** (0.0334)	0.221*** (0.0383)	0.361*** (0.0493)	0.158** (0.0758)	0.0330 (0.0800)	0.268*** (0.0994)
<i>prefecture</i> × <i>year</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>time effect</i> <sup>b</sup>	Yes	Yes	Yes	Yes	Yes	Yes
<i>additional controls</i> <sup>c</sup>	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	40,104	38,891	33,941	25,236	24,655	20,883
Adjusted <i>R</i> <sup>2</sup>	0.415	0.419	0.469	0.414	0.404	0.435

Note: Standard errors in parentheses are clustered on the prefecture level. The analyses are based on the 2-year sample.

Abbreviation: PRNP, premiums received from new policy.

<sup>a</sup>“*App* × *Indicator*” is the interaction term of *App* and various group indicators, that is, “*Indicator*.” The group indicators include different gender groups (i.e., female [=1] and male [=0]), profession groups (i.e., finance-related [=1] and nonfinance-related [=0]), and insured groups (i.e., oneself [=1] and others [=0]).

<sup>b</sup>Time effect includes the fixed effects of *month*, *year*, and *day\_of\_month*.

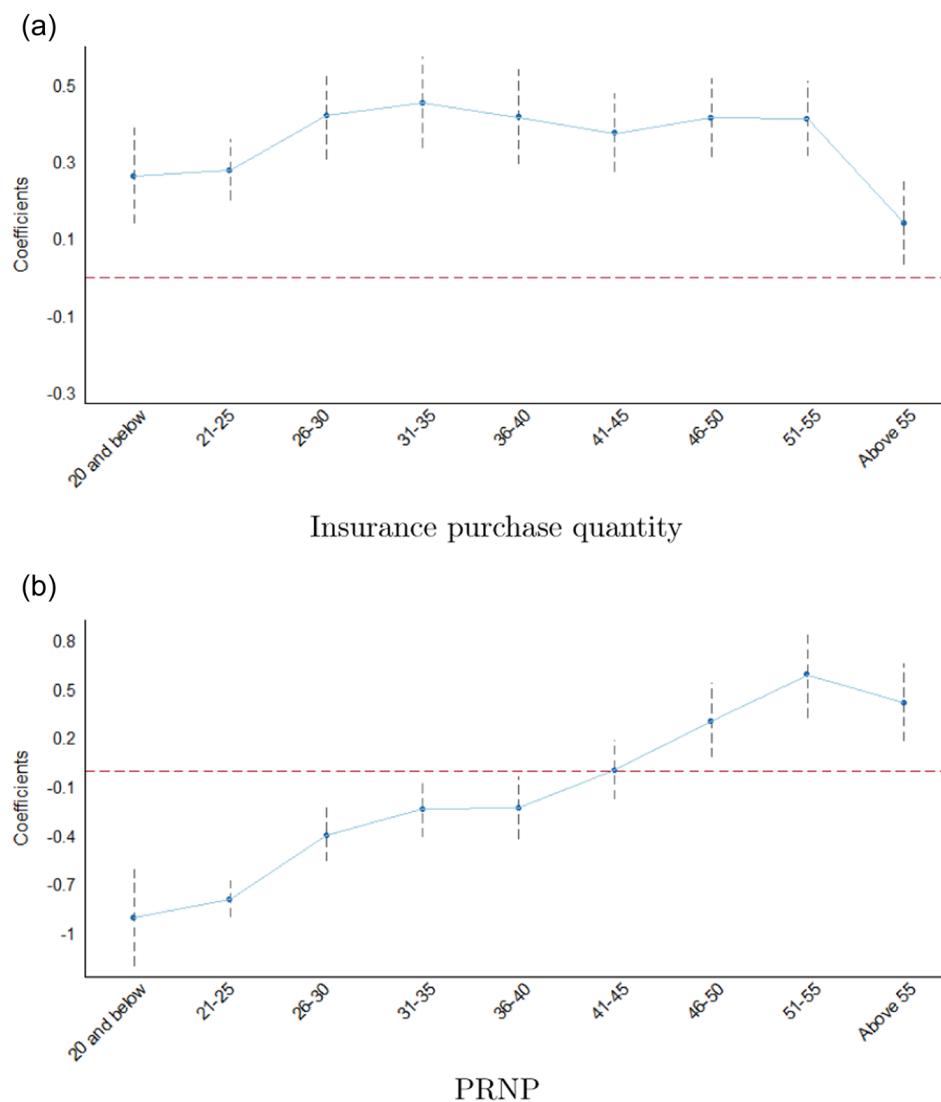
<sup>c</sup>Additional controls refer to “*holiday*,” “*new\_version*,” as well as the policyholder-related and policy-related control variables.

\**p* < 0.1.

\*\**p* < 0.05.

\*\*\**p* < 0.01.





**FIGURE 5** Heterogeneity of effects on the insurance purchase quantity and PRNP with respect to different age groups (a) Insurance purchase quantity, (b) PRNP. PRNP, premiums received from new policy. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

professions may receive more insurance-related information and have more diverse insurance purchasing channels due to job relevance, but this does not necessarily lead to them spending more money on insurance.

In Figure 5, we observe explicit heterogeneity across different age groups. As can be seen in Figure 5, Panel 5a, policyholders aged from 26 to 55 react strongly to the MA sales strategy in terms of the insurance purchase quantity, except younger groups ( $\leq 25$ ) and those above 55 having weaker responses. Starting from 26, people start to build their family and give birth to babies (Ren, 2021).<sup>15</sup> For the sake of their children, they may have strong motivations to purchase term life insurance. But with the increase in living expenses, they become more sensitive to the price, and the MA implementation driving down the price exactly meets their needs. As for the group over 55, they may

<sup>15</sup>According to the China Fertility Report (Ren, 2021), the average age of Chinese people having their first child in 2015 was 26.3 years old.

be less adept at spending on cellphones, and as a result, their response to the MA implementation is relatively weak. When we take a look at the corresponding PRNP analysis in Panel 5b, significant negative effects of the MA implementation on PRNP are observed for groups up to 40. This is partly due to the different changes in the purchasing behavior<sup>16</sup> of age groups after the implementation of the MA sales strategy. More specifically, after the MA implementation, younger groups (up to 40) tend to purchase insurance with a shorter duration and a longer premium payment term, which results in a reduction in the first premium received from underwriting their policies, thus a decrease in PRNP. Their selected insured amounts increase slightly, but not enough to drive an increase in PRNP. As for those over 40 years old, they choose to buy insurance with a longer duration and a comparably higher insured amount, which leads to the growth of PRNP.

By Figure 6, Panel 6a, a nonmonotonic and inverted-U-shaped effect of education level on the insurance purchase quantity is observed. Policyholders with educational achievement between a high school diploma and a bachelor's degree are more responsive to sales after implementing the MA sales strategy, compared with those with lower and higher education levels.<sup>17</sup> Generally, the MA sales channel lacks consultation services that the offline channels provide for complex insurance clauses. Typically, better-educated people tend to have more insurance literacy (Lin et al., 2019). Compared with less-educated ones, it is easier for them to purchase through the MA channel, and therefore better-educated people react more strongly to the MA implementation. However, we observe weaker responses from those holding master's degrees, which is presumably because they are less sensitive to the MA sales strategy. In other words, these people with advanced degrees may have high income and insurance awareness and would have purchased insurance regardless of the MA implementation, and thus would not have been affected much by MA implementation. In the analysis of PRNP (seen in Figure 6, Panel 6b), an inverted-U-shape is also shown from "secondary school" to higher degrees. Differently, people with lower educational attainments ("secondary school" and "technical secondary school") or higher education levels ("bachelor" and "master") respond negatively to the MA implementation. This is partly because large proportions of people in these education groups are aged up to 40, and Panel 5b in Figure 5 shows that people aged up to 40 have negative responses towards the MA implementation. As for those with a "primary school and below" diploma, we have observed that they have stronger responses towards the MA implementation in terms of PRNP. This may be the bias caused by the small sample size of this education level.

## 6 | FURTHER RESULTS

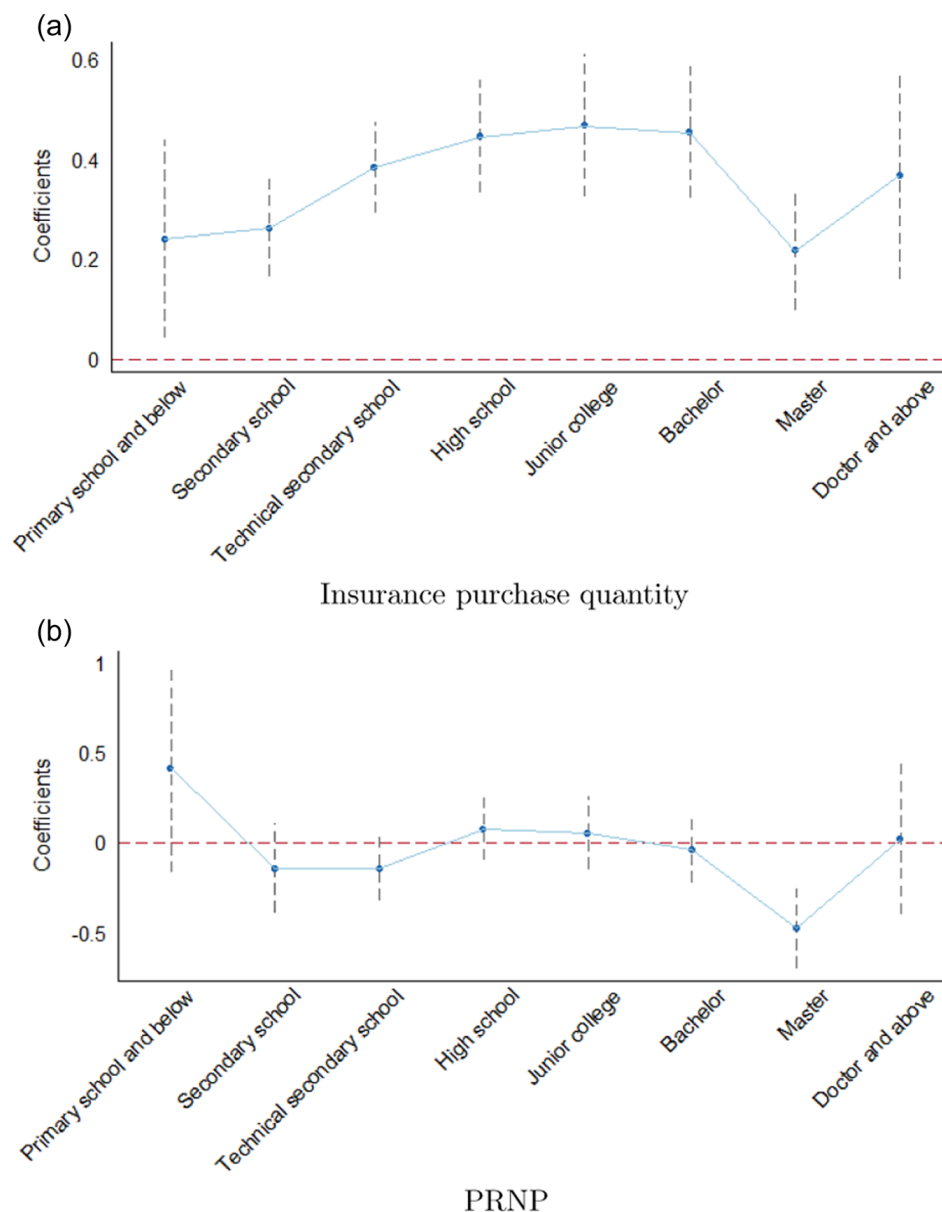
### 6.1 | Impact on insurance surrender ratio

For most term life insurance products,<sup>18</sup> policyholders are permitted to cancel the policy during the cancellation period with little or no penalty. This surrender behavior can be regarded as an impulsive purchase where the policyholder regrets the purchase. On average 10.4% of term life insurance policies are surrendered during the cancellation period. To explore how the MA sales strategy affects the impulsive purchase of insurance, we conduct a regression analysis by replacing  $Y_{it}$  in Equation (1) with the proportion of surrendered policies for day  $t$  in prefecture  $i$ . This proportion, that is,

<sup>16</sup>Here, the purchasing behavior means how the policyholder selects her policy's characteristics, that is, duration, premium payment term, and insured amount.

<sup>17</sup>As for the policyholders with a doctor or a higher degree, the corresponding effects are not well observed due to insufficient observations from these groups.

<sup>18</sup>In our 5-year data set, all term life insurance policies have a cancellation period of 10–20 days.



**FIGURE 6** Heterogeneity of effects on the insurance purchase quantity and PRNP with respect to education level (a) Insurance purchase quantity, (b) PRNP. PRNP, premiums received from new policy. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

insurance surrender ratio, is calculated as  $\frac{\text{Number of surrendered insurance policies}_{it}}{\text{Number of insurance policies exposed to surrender}_{it}}$ , which is similar to the calculation of a lapse ratio in Fang and Kung (2021).

The results, presented in Columns (1) and (2) of Table 12, indicate that the introduction of the MA sales strategy reduces the insurance surrender ratio by 5.25% for the 2017-to-2018 sample and by 6.26% for the 2015-to-2019 sample. We also further narrow the examination window of the sample period to 1 month before and after the introduction date for both the first and second MA implementation to investigate the difference in effect from such a shorter period. Column (3) shows the short-term effect of the first MA implementation and Column (4), the short-term effect of the second MA implementation. Overall, our results demonstrate that the proportion of surrendered

TABLE 12 Impact of the MA sales strategy on insurance surrender ratio.

	(2017–2018) (1)	(2015–2019) (2)	First launch (3)	Second launch (4)
<i>App</i>	−0.0525*** (0.00941)	−0.0626*** (0.00713)	−0.0329** (0.0160)	−0.0676*** (0.0199)
<i>holiday</i>	0.0120** (0.00593)	0.00838*** (0.00285)		
<i>new_version</i>	0.0174** (0.00861)	0.00576 (0.00542)		
<i>prefecture × year</i>	Yes	Yes	Yes	Yes
<i>time effect<sup>a</sup></i>	Yes	Yes	Yes	Yes
<i>additional controls<sup>b</sup></i>	Yes	Yes	Yes	Yes
<i>N</i>	30,068	93,580	2967	3697
Adjusted <i>R</i> <sup>2</sup>	0.091	0.093	0.082	0.086

Note: Standard errors in parentheses are clustered on the prefecture level.

Abbreviation: MA, mobile application.

<sup>a</sup>Time effect includes the fixed effects of *month*, *year*, and *day\_of\_month*.

<sup>b</sup>Additional controls refer to “holiday,” “new\_version,” as well as the policyholder-related and policy-related control variables.

\**p* < 0.1.

\*\**p* < 0.05.

\*\*\**p* < 0.01.

policies is significantly reduced due to the MA implementation. It implies that impulsive purchases of insurance are alleviated by the MA sales strategy, which is consistent with findings in existing research (e.g., Deleersnyder et al., 2002; Torkestani et al., 2018).

## 6.2 | Substitution effect on offline insurance purchases

Research on electronic retailing typically highlights the risk of substitution—the shift of sales from entrenched channels to the new Internet channel and losing the support from offline intermediaries in selling the firm's products (Coughlan et al., 2001). In view of this, we investigate if the introduction of the MA sales strategy crowds out offline insurance sales. We use offline policies as the sample for regression.

Table 13 shows the results of the regression analysis on offline insurance purchases. Again, we focus on the difference between the long-term (and 5 years) and short-term (2 years and 2 months<sup>19</sup>) effects. Basically, the short-term effects on the offline insurance purchase quantity, presented in Columns (1) and (3), are significantly negative, while also negative in Column (4) but nonsignificant. On the contrary, over a longer period, the negative effect of the MA sales strategy weakens and even

<sup>19</sup>The 2-month period includes 1 month before and 1 month after each launch date.

TABLE 13 Impact of the MA sales strategy on offline insurance purchases.

	Insurance purchase quantity				PRNP			
	(2017–2018) (1)	(2015–2019) (2)	First launch (3)	Second launch (4)	(2017–2018) (5)	(2015–2019) (6)	First launch (7)	Second launch (8)
<i>App</i>	−0.116*** (0.0237)	0.191*** (0.0248)	−0.118*** (0.0373)	−0.0219 (0.0337)	−0.0398 (0.0384)	−0.0595* (0.0352)	−0.0340 (0.0839)	−0.0488 (0.0684)
<i>holiday</i>	0.0989*** (0.0190)	0.0916*** (0.0127)			0.144*** (0.0434)	0.189*** (0.0190)		
<i>new_version</i>	0.152*** (0.0461)	0.186*** (0.0574)			0.0614 (0.0472)	0.0310 (0.0355)		
<i>prefecture × year</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>time effect</i> <sup>a</sup>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>additional controls</i> <sup>b</sup>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	16,636	62,691	1286	1210	6268	22,124	454	491
Adjusted <i>R</i> <sup>2</sup>	0.382	0.396	0.400	0.318	0.261	0.307	0.295	0.253

Note: Standard errors in parentheses are clustered on the prefecture level.

Abbreviations: MA, mobile application; PRNP, premiums received from new policy.

<sup>a</sup>Time effect includes the fixed effects of *month*, *year*, and *day\_of\_month*.

<sup>b</sup>Additional controls refer to “holiday,” “new\_version,” as well as the policyholder-related and policy-related control variables.

\**p* < 0.1.

\*\*\**p* < 0.05.

\*\*\*\**p* < 0.01.

turns positive. Specifically, as shown, the result in Column (2) becomes significantly positive. However, we only find the estimate from the 5-year sample is weakly significant for PRNP.

We find that the introduction of the MA sales strategy only crowds out the offline insurance purchase quantity in the short term. It is likely that the MA sales strategy yields both the negative crowding-out effect and the positive promotion effect. The mobile Internet channel accelerates brand communication by building more contacts with customers, forging a synergy of promoting offline purchases. This promotion effect likely takes some time to manifest. Therefore, the negative crowding-out effect dominates in the short term, while the positive promotion effect dominates in the long term.

## 7 | REMARKS AND CONCLUSIONS

This paper conducts an empirical analysis of the effect of introducing an MA sales strategy on the sale performance of the insurer by studying the purchase quantity and PRNPs. Using a two-way fixed-effect regression specification, we find that the MA sales strategy leads to an increase in both the insurance purchase quantity and the premium received from new policies. We find that the growth effect of the MA sales strategy on life insurance purchases can be explained by the increased accessibility of the distribution channel and lower costs. The MA distribution channel makes insurance products more accessible to customers, facilitating their purchases. The lower costs net of actuarially fair premiums results in a lower price for the MA channel, which incentivizes growth in insurance purchase quantity and PRNP. And it also accounts for the growth difference between PRNP and the quantity in response to the MA sales strategy.

Females, finance-related professionals, self-insured policyholders, and middle-educated policyholders are relatively more sensitive to the MA sales strategy in terms of both purchase quantity and PRNP. Our analysis also shows that customers of the MA channel are less impulsive and less likely to cancel their contracts during the cancellation period than offline customers. Additionally, we find that the adoption of the MA sales strategy displaces the offline insurance purchase quantity only in the adoption year (e.g., 2018), showing a temporary substitution effect.

Overall, our evidence indicates that the MA sales strategy is a successful marketing tool for an insurer. On the one hand, it substantially promotes insurance purchases, making more people aware of the insurance company's brand; on the other hand, it generates few negative impacts in the short run, but positive impacts on the offline channel sales in the long run. For insurance customers, this sales strategy is also beneficial for making insurance products more accessible, reducing cost and price as well as mitigating impulsive purchases. Our analysis is based on the policy data from a single insurer in China, which may limit the external validity of the results. More empirical evidence from other countries and insurers is needed for comparison. Second, the conclusions of this paper are based on life insurance. Whether they persist in nonlife insurance remains a question for future research.

## ACKNOWLEDGMENTS

We would like to thank the editor, the three anonymous referees, and Sebastian Kranz for the guidance, valuable comments, and suggestions.

## ORCID

Yusha Chen  <https://orcid.org/0000-0002-1115-5556>

Xian Xu  <https://orcid.org/0000-0002-7965-161X>



## REFERENCES

- Benlagha, N., & Hemrit, W. (2020). Internet use and insurance growth: Evidence from a panel of OECD countries. *Technology in Society*, 62, 101289.
- Berry-Stölzle, T. R., Hoyt, R. E., & Wende, S. (2010). Successful business strategies for insurers entering and growing in emerging markets. *The Geneva Papers on Risk and Insurance—Issues and Practice*, 35, 110–128.
- Bohnert, A., Fritzsche, A., & Gregor, S. (2019). Digital agendas in the insurance industry: The importance of comprehensive approaches. *The Geneva Papers on Risk and Insurance—Issues and Practice*, 44(1), 1–19.
- Brown, J. R., & Goolsbee, A. (2002). Does the internet make markets more competitive? Evidence from the life insurance industry. *Journal of Political Economy*, 110(3), 481–507.
- Browne, M. J., & Kim, K. (1993). An international analysis of life insurance demand. *Journal of Risk and Insurance*, 60(4), 616–634.
- Burke, R. R. (2002). Technology and the customer interface: What consumers want in the physical and virtual store. *Journal of the Academy of Marketing Science*, 30(4), 411–432.
- Butler, R. J. (2021). Information access and homeowners insurance purchases. *The Geneva Papers on Risk and Insurance—Issues and Practice*, 46(4), 649–663.
- Cather, D. A., & Howe, V. (1989). Conflict and channel management in property-liability distribution systems. *Journal of Risk and Insurance*, 56(3), 535–543.
- Chan, T. K., Cheung, C. M., Shi, N., & Lee, M. K. (2015). Gender differences in satisfaction with Facebook users. *Industrial Management and Data Systems*, 115(1), 182–206.
- China Banking and Insurance Regulatory Commission. (2011). *Notice on issues related to the regulation of life insurance business operation*. [https://www.cbirc.gov.cn/cn/view/pages/ItemDetail\\_gdsj.html?docId=10670&docType=2](https://www.cbirc.gov.cn/cn/view/pages/ItemDetail_gdsj.html?docId=10670&docType=2)
- China Banking and Insurance Regulatory Commission. (2016). *Circ notice on the release of “Life table of China's life insurance industry experience (2010–2013)”*. <https://www.cbirc.gov.cn/cn/view/pages/ItemDetail.html?docId=372677&itemId=925&generaltype=0>
- Cognizant. (2015). *Insurers use social mobile apps to increase digital value* [Technical Report, Cognizant]. <https://app.glueup.com/resources/protected/edm/7430/5087e006-cee2-4f6d-8366-b7d83260ac56.pdf>
- Coughlan, A. T., Anderson, E., Stern, L. W., & El-Ansary, A. I. (2001). Marketing channels. In A. T. Coughlan (Ed.), *Handbooks in transport*. Prentice Hall.
- De Chaisemartin, C., & D'Haultfoeulle, X. (2022). *Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey* [Technical Report, National Bureau of Economic Research]. <https://www.nber.org/papers/w29691>
- Deleersnyder, B., Geyskens, I., Gielens, K., & Dekimpe, M. G. (2002). How cannibalistic is the internet channel? A study of the newspaper industry in the United Kingdom and the Netherlands. *International Journal of Research in Marketing*, 19(4), 337–348.
- Digital Finance Research Center of Peking University. (2021). *The Peking University digital financial inclusion index of China* [Technical Report]. <https://idf.pku.edu.cn/docs/20210421101507614920.pdf>
- Eastman, K. L., Eastman, J. K., & Eastman, A. D. (2002). Issues in marketing online insurance products: An exploratory look at agents' use, attitudes, and views of the impact of the internet. *Risk Management and Insurance Review*, 5(2), 117–134.
- Eling, M., & Lehmann, M. (2018). The impact of digitalization on the insurance value chain and the insurability of risks. *The Geneva Papers on Risk and Insurance—Issues and Practice*, 43(3), 359–396.
- Fang, H., & Kung, E. (2021). Why do life insurance policyholders lapse? The roles of income, health, and bequest motive shocks. *Journal of Risk and Insurance*, 88(4), 937–970.
- Ferber, R., & Lee, L. C. (1980). Acquisition and accumulation of life insurance in early married life. *Journal of Risk and Insurance*, 47(4), 713–734.
- Fritzsche, S., Scharner, P., & Weiß, G. (2021). Estimating the relation between digitalization and the market value of insurers. *Journal of Risk and Insurance*, 88(3), 529–567.
- Fritzsche, A., & Bohnert, A. (2021). Implications of bundled offerings for business development and competitive strategy in digital insurance. *The Geneva Papers on Risk and Insurance—Issues and Practice*, 47, 817–834.
- Garven, J. R. (2002). On the implications of the internet for insurance markets and institutions. *Risk Management and Insurance Review*, 5(2), 105–116.
- Gilly, M. C., & Wolfenbarger, M. (2000). A comparison of consumer experiences with online and offline shopping. *Consumption, Markets and Culture*, 4(2), 187–205.

- Goldfarb, A., & Tucker, C. (2019). Digital economics. *Journal of Economic Literature*, 57(1), 3–43.
- Hammond, J., Houston, D. B., & Melander, E. R. (1967). Determinants of household life insurance premium expenditures: An empirical investigation. *Journal of Risk and Insurance*, 34(3), 397–408.
- Hansen, T., & Jensen, J. M. (2009). Shopping orientation and online clothing purchases: The role of gender and purchase situation. *European Journal of Marketing*, 43(9/10), 1154–1170.
- Harmon, S. K., & Hill, C. J. (2003). Gender and coupon use. *Journal of Product and Brand Management*, 12(3), 166–179.
- Hu, Y. J., & Tang, Z. (2014). The impact of sales tax on internet and catalog sales: Evidence from a natural experiment. *International Journal of Industrial Organization*, 32, 84–90.
- Huang, L., Lu, X., & Ba, S. (2016). An empirical study of the cross-channel effects between web and mobile shopping channels. *Information and Management*, 53(2), 265–278.
- Kjosevski, J. (2012). The determinants of life insurance demand in central and southeastern Europe. *International Journal of Economics and Finance*, 4(3), 237–247.
- Kwon, K.-N., & Kwon, Y. J. (2007). Demographics in sales promotion proneness: A socio-cultural approach. *ACR North American Advances*, 34, 288–294.
- Lazaris, C., Vrechopoulos, A. P., Doukidis, G. I., & Fraidaki, A. (2015). Mobile apps for omnichannel retailing: Revealing the emerging showroom phenomenon. In *The 9th Mediterranean Conference on Information Systems (MCIS)*, 3–5 October, 2015. Samos, Greece. (p. 12). <https://doi.org/10.13140/RG.2.1.1080.5364>
- Li, D., Moshirian, F., Nguyen, P., & Wee, T. (2007). The demand for life insurance in OECD countries. *Journal of Risk and Insurance*, 74(3), 637–652.
- Li, H., Kuo, C., & Russell, M. G. (1999). The impact of perceived channel utilities, shopping orientations, and demographics on the consumer's online buying behavior. *Journal of Computer-Mediated Communication*, 5(2), JCMC521.
- Lin, X., Bruhn, A., & William, J. (2019). Extending financial literacy to insurance literacy: A survey approach. *Accounting and Finance*, 59(1), 685–713.
- Luo, X., Andrews, M., Fang, Z., & Phang, C. W. (2014). Mobile targeting. *Management Science*, 60(7), 1738–1756.
- Mantis, G., & Farmer, R. N. (1968). Demand for life insurance. *Journal of Risk and Insurance*, 35(2), 247–256.
- Nagra, G., & Gopal, R. (2013). An study of factors affecting on online shopping behavior of consumers. *International Journal of Scientific and Research Publications*, 3(6), 1–4.
- Ren, Z. (2021). *China fertility report* [Technical Report]. <https://app.dahecube.com/nweb/news/20210416/95321n5d09a8fa41e6.htm?artid=95321>
- Securities Daily. (2016). 18 Life insurance companies have developed mobile apps, and most have opened the WeChat service numbers. <https://www.ljzfin.com/info/28719.jsp>
- Sharma, D., & Gassenheimer, J. B. (2009). Internet channel and perceived cannibalization: Scale development and validation in a personal selling context. *European Journal of Marketing*, 43(7/8), 1076–1091.
- Stahl, D. O. (1989). Oligopolistic pricing with sequential consumer search. *The American Economic Review*, 79(4), 700–712.
- Statistics Bureau of the P.R. China. (2020). *China statistical yearbook 2015–2019*. China Statistics Press.
- Torkestani, M. S., Zandmehr, M., & Afsahizadeh, M. (2018). Investigating the relationship between website characteristics and impulsive purchase on the internet. *Journal of Business Administration Researches*, 10(19), 1–17.
- Truett, D. B., & Truett, L. J. (1990). The demand for life insurance in Mexico and the United States: A comparative study. *Journal of Risk and Insurance*, 57(2), 321–328.
- Wisniewski, J. (2011). Control-shift-mobile that works for your library. *Online-Medford*, 35(1), 54.
- Wolfenbarger, M., & Gilly, M. C. (2001). Shopping online for freedom, control, and fun. *California Management Review*, 43(2), 34–55.

**How to cite this article:** Chen, A., Chen, Y., Murphy, F., Xu, W., & Xu, X. (2023). How does the insurer's mobile application sales strategy perform? *Journal of Risk and Insurance Review*, 1–33. <https://doi.org/10.1111/jori.12424>



## APPENDIX A

See Table A1.

**TABLE A1** Robust preliminary results of the MA sales strategy effects, using the same data sample for the quantity and PRNP analysis.

	Insurance purchase quantity		PRNP	
	(2017–2018) (1)	(2015–2019) (2)	(2017–2018) (3)	(2015–2019) (4)
<i>App</i>	1.082*** (0.109)	1.150*** (0.113)	0.522*** (0.119)	0.393*** (0.130)
<i>_cons</i>	−0.264** (0.102)	0.371*** (0.0341)	6.934*** (0.115)	6.830*** (0.0370)
<i>prefecture × year</i>	Yes	Yes	Yes	Yes
<i>time effect<sup>a</sup></i>	Yes	Yes	Yes	Yes
<i>N</i>	19,127	51,145	19,127	51,145
Adjusted <i>R</i> <sup>2</sup>	0.479	0.405	0.369	0.340

Note: Standard errors in parentheses are clustered on the prefecture level.

Abbreviations: MA, mobile application; PRNP, premiums received from new policy.

<sup>a</sup>Time effect includes the fixed effects of *month*, *year*, and *day\_of\_month*.

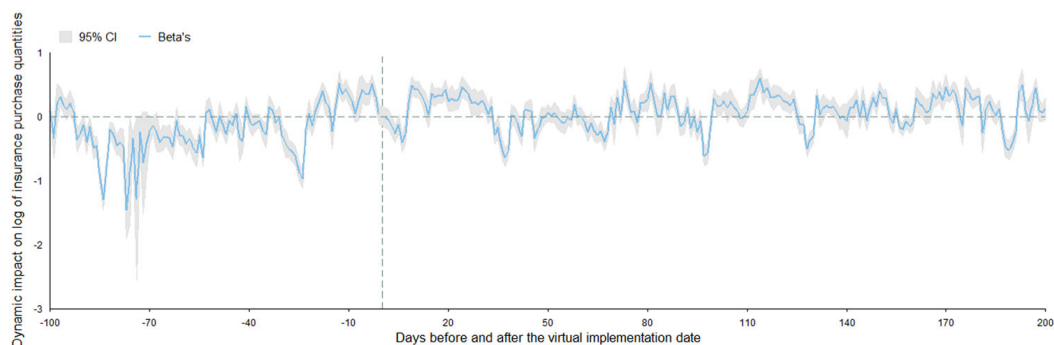
\* $p < 0.1$ .

\*\* $p < 0.05$ .

\*\*\* $p < 0.01$ .

## APPENDIX B

See Figure B1.

**FIGURE B1** Results of the 5-year sample (2013–2017) using event study method: insurance purchase quantity. CI, confidence interval. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



## **4 Investing in green bonds with goal-oriented preferences**

**Source:**

Chen, A., Chen, Y., Nguyen, T., & Uddin, G. S. (2024). Investing in green bonds with goal-oriented preferences. Working paper.



# Investing in Green Bonds with Goal-Oriented Preferences

An Chen<sup>\*</sup>, Yusha Chen<sup>\*</sup>, Thai Nguyen<sup>#</sup> and Gazi Salah Uddin<sup>‡</sup>

<sup>\*</sup> Institute of Insurance Science, Ulm University, Helmholtzstr. 20, 89069 Ulm, Germany.

<sup>#</sup> École d'Actuariat, Université Laval, 2425 rue de l'Agriculture, G1V 0A6 Québec, Canada.

<sup>‡</sup> Department of Management and Engineering, Linköping University, SE-58183 Linköping, Sweden.

E-Mail: an.chen@uni-ulm.de;

yusha.chen@uni-ulm.de;

thai.nguyen@act.ulaval.ca;

gazi.salah.uddin@liu.se

## Abstract

Coping with climate change calls for a shift to low-carbon and climate-resilient investments. Investing in green bonds has become one of the options for many companies, especially those with corporate social responsibility (CSR) investment targets. In this article, we incorporate the CSR investment goal of institutional investors in their utility function, indicating the goal-based preference. Their overall utility exhibits a non-smooth nature, encompassing a constant relative risk aversion (CRRA) utility component derived from the terminal wealth and a piece-wise linear function that measures the utility gain or loss associated with achieving or falling short of the goal. Relying on the Lagrangian approach, the optimal terminal wealth and asset holdings, including in green bonds, are determined by maximizing the expected overall utility function. Interestingly, a more ambitious target does not consistently result in a greater allocation of funds towards green bonds. This allocation significantly hinges on investors' level of risk aversion. For investors with a relative risk aversion greater than 1, an increased target leads to a higher fraction of wealth being invested or short sold in green bonds, while the opposite holds for investors with a relative risk aversion falling between 0 and 1.

**Keywords:** portfolio planning, CSR investment goals, utility gains, utility losses

**JEL:** G11, G22, G41

## 1 Introduction

The concept of corporate social responsibility (CSR) has been proposed in the literature as early as the 1930s and 1940s (Carroll, 1999), which emphasizes that business has social obligations (Smith, 2003). Under the modern definition, it is specified as “the commitment of corporations to contribute to sustainable development” (e.g., Petkoski and Twose, 2003; Chandler, 2019). It has gained increasing attention as a tool to support and accelerate the transition process to a more sustainable economy. Engagements in CSR activities or projects can usually help companies’ brands stand out from their competitors, because these activities or projects can improve their reputation (Islam et al., 2021), increase customers’ retention and loyalty (Afiuc et al., 2020), and boost employees’ morale to achieve greater productivity (Branco and Rodrigues, 2006). One crucial way to achieve CSR goals is to support and finance sustainable projects by carrying out green bond investments. Different from the conventional corporate bond, a green bond is designed as a fixed-income instrument financing or re-financing “low-carbon and climate-resilient” projects (Kennedy and Corfee-Morlot, 2012). In recent years, the market of green bonds has developed rapidly. In 2014, the issuance of green bonds amounted to 37 billion dollars worldwide, and by 2022, it had surged to 487 billion dollars (Statista, 2023) (see Figure 1).

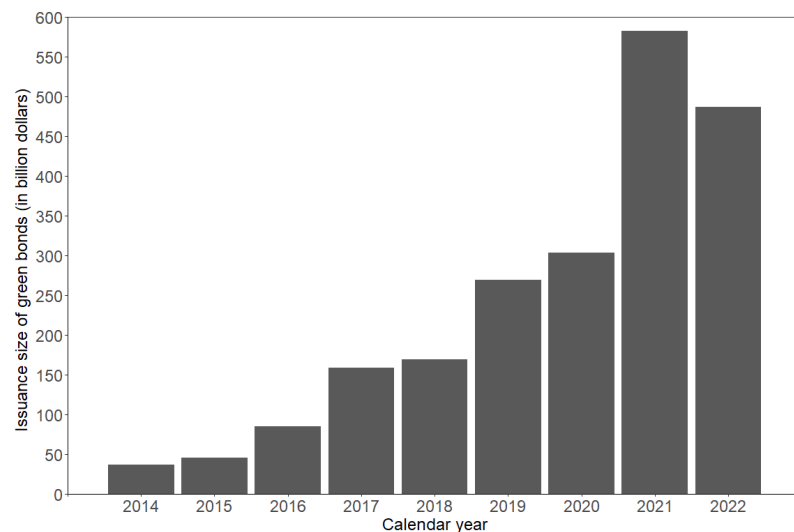


Figure 1: The value of green bonds issued worldwide (in billion dollars). Data is sourced from Statista (2023).

Going through the literature, most research on green bonds is empirical and focuses mainly on four aspects: (i) The performance of portfolios with green bonds. Empirical

results from (Han and Li, 2022) indicate that in most cases on both the European and the U.S. markets, portfolios containing green bonds outperform those with conventional bonds in terms of increased returns and decreased volatility (Braga et al., 2021). In addition, the issuance of green bonds can reduce the risk associated with green investments when compared to conventional energy bonds, which tend to be more volatile due to fluctuations in oil prices (Braga et al., 2021). (ii) Green bond premium (or greenium). Different from risk premiums, green bond premium means the green bond may be priced lower than the risk-paired conventional corporate bonds. Applying data from various countries, time spans or markets, researchers reach different conclusions. Some studies present evidence that there are green bond premiums (e.g., Karpf and Mandel, 2018; Gianfrate and Peri, 2019; Nanayakkara and Colombage, 2019), while some report controversial results (e.g., Hachenberg and Schiereck, 2018; Bachelet et al., 2019; Partridge and Medda, 2020; Tang and Zhang, 2020). (iii) The connection between green bonds and other financial instruments. For instance, Reboredo (2018) explores the co-movement between green bonds and treasury (or corporate) bonds. Strong price connectedness exists among these instruments (Reboredo and Ugolini, 2020). And it is discovered that the prices of green bonds and stocks are correlated (e.g., Park et al., 2020; Nguyen et al., 2021). (iv) Supply analysis of green bonds. This refers to explorations on determinants and influential factors of green bond supply, such as issuance size (Chiesa and Barua, 2019), financing cost (Li et al., 2020).

Nevertheless, there is a scarcity of theoretical research dedicated to green bond investments. In particular, there is limited literature attempting to model the risk preferences of green bond institutional investors (abbreviated as green bond investors) bearing CSR goals. Such research could offer valuable insights into the investment behavior of these institutional investors. A significant contribution to this field comes from Pástor et al. (2021). They have developed an equilibrium model for asset pricing that incorporates non-pecuniary benefits in investors' utility. This means that holding assets with stronger environmental, social and governance (ESG) characteristics leads to increased non-pecuniary benefits for investors.

Different from Pástor et al. (2021), we want to explore the impact of green bond investors' risk preferences from another angle. In the present article, the green bond investors bear a CSR investment target in mind, that is, they have a goal-based preference. In this sense, we construct the utility of green bond investors relying on the works of Kőszegi and Rabin (2006, 2009). Through this, our objective is to address whether a more ambitious goal established by the investors results in an increased investment in green

bonds. Specifically, a CSR investment goal is to be set. An investment in green bonds that exceeds the CSR investment goal leads to utility gains for the investor, otherwise to utility losses. Consequently, the overall utility of green bond investors depends not only on their final wealth after the investment but also on the utility gains/losses resulting from satisfying/missing the CSR investment goal in the green bond. We assume that the investors have a stronger aversion to utility losses from missing the target than utility gains from hitting the target, i.e., demonstrating loss aversion. In this current investigation, the utility component of terminal wealth is modeled as a constant relative risk aversion (CRRA) utility with a specified CRRA level denoted as  $\gamma$ . Additionally, the gain-or-loss component is characterized using a piece-wise linear function.

We build on the literature on the asset allocation problem, which dates back to Merton (1969). Merton has made two crucial assumptions to achieve analytical solutions. The first assumption is that the underlying risky investment generates log-normally distributed risky payoffs. The second one is that the risk preferences of the optimizing agent are characterized by hyperbolic absolute risk aversions (HARA), such as CRRA and exponential utility.<sup>1</sup> In the current study, we extend Merton's framework by incorporating the CSR investment target in the investors' risk preferences. Consequently, we determine the optimal fractions of the wealth invested in the green bond and stock.

When the prices of the green bond and stock are negatively correlated, our results numerically show that green bond investors significantly increase their asset holding in the green bond as time approaches maturity, which is attributed to its diminishing volatility. Moreover, we observe that the reduced loss aversion leads the investors to retain a larger portion in both the green bond and stock. As for the effect of the CSR investment goal on green bond investors' holdings of the green bond and stock, it differs between two scenarios: (i) Green bond investors with a risk aversion coefficient  $\gamma > 1$  will increase their holdings in these assets with a higher CSR investment target since their expected utility is largely influenced by the utility gain-or-loss measurement on achieving the CSR investment target or not. However, (ii) green bond investors with  $\gamma \in (0, 1)$  invest lower fractions of wealth in the green bond and stock with a higher CSR investment goal. A possible explanation is that the investors with  $\gamma > 1$  place greater importance on the utility gain achieved by reaching the target. Conversely, the investors with  $\gamma \in (0, 1)$  assign greater weight to the utility derived from the terminal wealth in their overall util-

---

<sup>1</sup>Analytical solutions can still be achieved if the HARA utility is extended to the so-called symmetric asymptotic HARA utility functions, which incorporate exponential and CRRA utility as limiting cases (e.g., Chen et al., 2011; Chen and Vellekoop, 2017).



ity assessment. Particularly, these investors allocate significant portions of their wealth to both the green bond and stock, potentially resulting in substantial financial losses. In consequence, they opt to diminish their holdings in these two assets to increase their overall expected utility.

Furthermore, we numerically investigate the influence of a positive correlation between stock and green bond prices. Our findings indicate that only a significantly strong correlation, such as 80%, can generate significant variations in our results. The green bond investors short sell the green bond while taking a long position in the stock. Investors with a greater loss aversion short sell a lower fraction of wealth in the green bond and invest a smaller fraction in the stock. With a higher CSR investment goal, green bond investors with  $\gamma > 1$  short sell a larger fraction of wealth in the green bond and buy a larger fraction in the stock; while the opposite holds true for the investors with  $\gamma \in (0, 1)$ .

The remainder of this article is structured as follows: In Section 2, we build up the models for our financial market and the utility function of green bond investors. Next, in Section 3, we introduce the asset allocation optimization problem of green bond investors and then solve the problem to obtain their optimal terminal wealth and trading strategies. This is followed by Section 4, where we illustrate and analyze the results numerically. In the last section, we give some remarks on our findings and conclude the article. Several proofs are provided in the Appendix.

## 2 Description of the Model

This section aims to introduce the model setup by specifying the underlying financial market and the utility function of green bond investors.

### 2.1 The Financial Market

Consider a green bond investor's investment operating on the time horizon  $[0, T]$ . We consider a complete continuous-time economy in which a bank account, a zero-coupon green bond and a stock are traded. Further, we assume that the processes for the zero-coupon green bond and stock are defined under some probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ , where  $\mathbb{P}$  denotes the data generating probability measure and  $(\mathcal{F}_t)_{0 \leq t \leq T}$  the natural filtration generated by the stock and green bond. More precisely, the stochastic bank

account corresponding to an accumulation factor is defined by

$$B(t) = \exp \left\{ \int_0^t r(u) du \right\}, \quad dB(t) = r(t)B(t)dt, \quad (1)$$

where  $r(t)$  is the instantaneous risk-free interest rate at time  $t$ . The term structure of the interest rate is determined by  $r(t)$  and a zero-coupon green bond with maturity  $T$ , which is modeled by the Vasiček (1977) term structure model, i.e.

$$dr(t) = (b - ar(t))dt + \sigma dW_1(t), \quad (2)$$

where  $a, b, \sigma > 0$ , and  $W_1(t)$  is assumed to be a one-dimensional Brownian motion. The parameter  $\sigma$  gives the volatility of the short rate,  $a$  represents the mean-reverting speed factor and  $\frac{b}{a}$  is the long-run equilibrium value.

Under the data generating probability measure, it is assumed that the stock and green bond are driven by some Itô processes. In particular,

$$dG(t, T) = G(t, T) (r(t) + \sigma_G(t, T)\kappa) dt + G(t, T)\sigma_G(t, T)dW_1(t), \quad (3)$$

$$dS(t) = S(t) (r(t) + \sigma_S\psi) dt + S(t)\sigma_S \left( \rho_{GS}dW_1(t) + \sqrt{1 - \rho_{GS}^2}dW_2(t) \right), \quad (4)$$

where  $W_1(t)$  is the same Brownian motion as given in Equation (2), and is independent of the other one-dimensional Brownian motion  $W_2(t)$ . The correlation coefficient  $\rho_{GS} \in (-1, 1)$  represents the correlation between the stock and green bond. Here,  $\sigma_G(t, T)$  and  $\sigma_S$  give the volatility of the green bond and stock, respectively and correspondingly,  $\kappa$  and  $\psi$  are marked as the market prices of risks. Fixed market prices of risks have been assumed for tractability. In the Vasiček's work (Vasicek, 1977), the volatility of the green bond is given by  $\sigma_G(t, T) = \frac{\sigma}{a}(1 - e^{-a(T-t)})$ .

With fixed market prices of risks  $\kappa$  and  $\psi$ , there exists a unique state price density process,  $\delta_t$ :

$$\delta_t = \delta_0 \cdot \exp \left\{ - \int_0^t r(u)du - \int_0^t \kappa dW_1(u) - \int_0^t \tilde{\kappa} dW_2(u) - \frac{1}{2} \int_0^t \kappa^2 du - \frac{1}{2} \int_0^t \tilde{\kappa}^2 du \right\}, \quad (5)$$

where  $\delta_0 = 1$ , and  $\tilde{\kappa} := \frac{\psi - \rho_{GS}\kappa}{\sqrt{1 - \rho_{GS}^2}}$ .

The state price density describes the scarceness of economic states and reflects the status

of the entire economy. The value of  $\delta_t$  is high in an unfavorable economic trend and low in a rosy economic situation. Later, the functional relationship between the optimal wealth and the state price density will be used to interpret some numerical results.

## 2.2 Utility Modelling of Green Bond Investors

When setting up their investment strategies, we assume that green bond investors bear a certain CSR investment target in mind. They care about how much green bond investment they have made and whether the green assets they have invested in have achieved the CSR investment target. Inspired by Kőszegi and Rabin (2006, 2009), it is assumed that the overall utility of the green bond investor depends not only on her final wealth after the investment, but also on whether the CSR investment target in green bonds is achieved or not.

Let  $x$  be the terminal wealth,  $\underline{x}^G \in (0, \infty)$  be a subjective CSR investment target that the green bond investor has in mind for her investment in the green bond. We assume that the green bond investor's risk preference is described by the following utility function:

$$\tilde{U}(x) = \tilde{U}(x, \underline{x}^G) := U(x) - g(U(\underline{x}^G) - U(mx)). \quad (6)$$

The overall utility  $\tilde{U}$  consists of two components. The first component ( $U(x)$ ) measures satisfaction with the final wealth position. Here,  $U(x)$  is assumed to be a constant relative risk aversion (CRRA) utility function in the form of  $U(x) = \frac{x^{1-\gamma}}{1-\gamma}$ ,  $\gamma > 0$  and  $\gamma \neq 1$ , where  $\gamma$  is the risk aversion parameter. It is an increasing and concave utility function satisfying the Inada's condition  $\lim_{x \rightarrow 0} U'(x) = +\infty$  and  $\lim_{x \rightarrow \infty} U'(x) = 0$ . The second component ( $g(\cdot)$ ) measures the utility gain and loss that the green bond investor experiences through investing in the green bond, which depends on the utility difference between the target amount that the investor desires to invest in the green bond ( $\underline{x}^G$ ) and the actual investment amount.<sup>2</sup> It can be observed that there is an increase in the green bond investor's utility if the outcome of the invested amount in the green bond is higher than the target amount that the investor desires to invest in the green bond. In case it is smaller than the desirable amount, the green bond investor suffers a loss in utility. The size of utility loss or gain is measured by the function  $g$ .

<sup>2</sup>Intuitively, the real investment amount in green bonds may be the average of actual amount during the entire investment period, or the final amount invested in the green bond in the last period. However, either of these cases would make our optimization problem much more complicated.

For simplicity, we compare  $mx$  with the CSR investment goal  $\underline{x}^G$  to determine the utility gains and losses, where  $m$  is an exogenously determined parameter. Due to the CRRA utility which is defined for the non-negative real line,  $m < 0$  does not make sense. In our subsequent analysis, we consider the nontrivial case  $m > 0$ . Later, we will come to the solutions to our optimization problem using a two-step approach. First, we obtain the solutions with a given  $m$  value (see Subsection 3.1). Second, we discuss how  $m$  can be chosen optimally by considering additional risk management criteria (see Subsection 3.2).

Back to the definition of the utility loss or gain function  $g$ , we assume that the investor has stronger aversion to negative deviations (losses) from the desirable investment in the green bond relative to positive deviations (gains), i.e. demonstrating loss aversion. This “loss aversion” feature can be incorporated in the model by assuming  $g$  to be a piecewise linear function

$$g(y) = y(\epsilon \mathbb{1}_{y \geq 0} + \iota \mathbb{1}_{y < 0}), \quad (7)$$

where  $\epsilon > \iota > 0$  are some scaling parameters (Kőszegi, 2006). For a given  $\iota$ , the higher the value of  $\epsilon$ , the more loss averse the investor is, where loss refers to the negative utility difference between the real and target amount set for the green bond investment ( $\underline{x}^G$ ).

Figure 2 gives an intuitive description to the overall utility of the green bond investor. When the investment amount in the green bond ( $mx$ ) fails to meet the subjective CSR investment target ( $\underline{x}^G$ ), the green bond investor suffers a utility loss, and vice versa. For the green bond investor, the size of her utility loss is larger than that of the utility gain. In this case, the utility curve of the green bond investor is not smooth, but has a kink at the point of  $x = 100$ .

### 3 Optimization Problem

We assume that the green bond investor owns an initial wealth of  $X_0$  for investment. We use  $\xi = \{\xi_t\}_{t \in [0, T]}$  and  $\beta = \{\beta_t\}_{t \in [0, T]}$  to denote the fractions of wealth invested in the stock and green bond respectively. The remaining wealth is invested in the bank account. We further assume that the processes  $\xi_t$  and  $\beta_t$  are  $\mathcal{F}$ -adapted. In the environment with the stochastic interest rate, the evolution of her wealth process  $(X_t)_{t \in [0, T]}$  can be

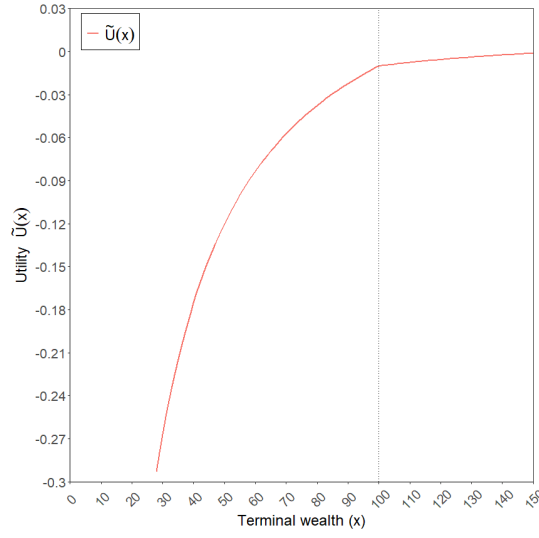


Figure 2: Utility of the green bond investor. Here we set  $m = 0.3$ ,  $\epsilon = 3$ ,  $\iota = 0.5$ ,  $\gamma = 2$ ,  $\underline{x}^G = 30$ .

written as:

$$dX_t = X_t (r(t) + \xi_t \sigma_S \psi + \beta_t \sigma_G(t, T) \kappa) dt + \xi_t X_t \sigma_S \left( \rho_{GS} dW_1(t) + \sqrt{1 - \rho_{GS}^2} dW_2(t) \right) + \beta_t X_t \sigma_G(t, T) dW_1(t). \quad (8)$$

With the utility specified by (6) in Section 2.2, the green bond investor is about to maximize her expected utility with respect to the terminal wealth over the set of admissible strategies

$$\mathcal{A}(X_0) := \left\{ (\xi, \beta) \mid (\xi, \beta) \text{ is progressively measurable, } \int_0^T (\xi_t^2 + \beta_t^2) dt < \infty \right\}.$$

Therefore, the optimization objective function of the green bond investor's asset allocation problem can be expressed by

$$\sup_{(\xi, \beta) \in \mathcal{A}(X_0)} \mathbb{E} \left[ \tilde{U}(X_T) \right], \quad \text{s.t. (8)}. \quad (9)$$

Following Cox and Huang (1989), the dynamic asset allocation problem can be solved by

a so-called Lagrangian approach, i.e. by determining the terminal wealth  $X_T$  directly:

$$\sup_{X_T} \mathbb{E} \left[ \tilde{U}(X_T) \right], \quad \text{s.t. } \mathbb{E} [X_T \delta_T] \leq X_0. \quad (10)$$

### 3.1 Optimal Solutions with A Given $m$ Value

Note that by introducing the risk preference considering a CSR investment goal, the utility function is not everywhere differentiable, hence the classical martingale approach (e.g., Cox and Huang, 1989) cannot be applied directly. However, we can still rely on the Lagrangian approach to determine the optimal terminal wealth analytically. The key idea is to determine several locally optimal solutions and the global optimal solution is obtained by comparing the local maximizers. Before moving to the solutions to the optimization problem, we define  $I^\iota(\cdot)$  and  $I^\epsilon(\cdot)$  as the inverse marginal utility functions, and they take the forms as follows

$$I^\iota(x) := \left( \frac{x}{1 + \iota m^{1-\gamma}} \right)^{-1/\gamma}, \quad I^\epsilon(x) := \left( \frac{x}{1 + \epsilon m^{1-\gamma}} \right)^{-1/\gamma}. \quad (11)$$

For any  $x > 0$ , it is obvious that  $I^\epsilon > I^\iota$  as  $\epsilon > \iota$  and both  $I^\epsilon$  and  $I^\iota$  are decreasing in  $x$ . Further, it is assumed that the integrability condition is satisfied, i.e. for any  $\lambda > 0$ , it holds that

$$\mathbb{E}[U(I^i(\lambda \delta_T))] < \infty \quad \text{and} \quad \mathbb{E}[\delta_T I^i(\lambda \delta_T)] < \infty, \quad i = \iota, \epsilon, \quad (12)$$

which clearly holds for a CRRA utility function. Then, the optimal wealth of the green bond investor at time  $T$  can be derived by the following Theorem 3.1.

**Theorem 3.1** (Optimal terminal wealth of a green bond investor). *The optimal solution to (10), with the green bond investor's utility function defined by (6) and (7), is given by*

$$X_T^* = \begin{cases} I^\iota(\lambda \delta_T), & \text{if } \delta_T < \bar{\delta}^\iota \\ \frac{\underline{x}^G}{m}, & \text{if } \bar{\delta}^\iota \leq \delta_T < \bar{\delta}^\epsilon \\ I^\epsilon(\lambda \delta_T), & \text{if } \delta_T \geq \bar{\delta}^\epsilon \end{cases}. \quad (13)$$

Moreover, two critical points  $\bar{\delta}^\iota$  and  $\bar{\delta}^\epsilon$  are defined by

$$\bar{\delta}^\iota := (1 + \iota(m)^{1-\gamma})U'(\underline{x}^G/m)/\lambda \quad \text{and} \quad \bar{\delta}^\epsilon := (1 + \epsilon(m)^{1-\gamma})U'(\underline{x}^G/m)/\lambda, \quad (14)$$

where the Lagrangian multiplier  $\lambda > 0$  solves  $\mathbb{E}(\delta_T X_T^*) = X_0$ , and  $U'(\underline{x}^G/m) = (\underline{x}^G/m)^{-\gamma}$ .

PROOF: Proof can be found in Appendix B.

According to different economic states at the final point  $T$ , the optimal terminal wealth  $X_T^*$  is classified to three parts. If the final state price  $\delta_T$  is below some threshold  $\bar{\delta}^\iota$  (in the following, we name this situation as a good economic region), the optimal terminal wealth is  $I^\iota(\lambda\delta_T)$ , which decreases with  $\delta_T$ . Then, if it is the situation that  $\bar{\delta}^\iota \leq \delta_T < \bar{\delta}^\epsilon$  (named as an intermediate economic region), the optimal terminal wealth corresponds to  $\frac{x^G}{m}$ , which does not vary with  $\delta_T$ . Eventually, if the final state price goes high,  $\delta_T \geq \bar{\delta}^\epsilon$  (named as a bad economic region), the optimal terminal wealth becomes even lower, i.e.  $I^\epsilon(\lambda\delta_T)$ , which again, declines as the final economy worsens.

**Theorem 3.2** (Optimal wealth of a green bond investor at time  $t \in (0, T)$ ). *The optimal wealth of the green bond investor at any time  $t \in (0, T)$  takes the form of*

$$\begin{aligned} X_t^* = & \left( \frac{\lambda\delta_t}{1 + \iota(m)^{1-\gamma}} \right)^{-\frac{1}{\gamma}} \exp \left\{ \left(1 - \frac{1}{\gamma}\right)\mu_{\delta_t} + \frac{1}{2}\left(1 - \frac{1}{\gamma}\right)^2\sigma_{\delta_t}^2 \right\} \Phi \left( \frac{\ln \frac{\bar{\delta}^\iota}{\delta_t} - \mu_{\delta_t} - (1 - \frac{1}{\gamma})\sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right) \\ & + \frac{x^G}{m} \exp \left\{ \mu_{\delta_t} + \frac{1}{2}\sigma_{\delta_t}^2 \right\} \left[ \Phi \left( \frac{\ln \frac{\bar{\delta}^\epsilon}{\delta_t} - \mu_{\delta_t} - \sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right) - \Phi \left( \frac{\ln \frac{\bar{\delta}^\iota}{\delta_t} - \mu_{\delta_t} - \sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right) \right] \\ & + \left( \frac{\lambda\delta_t}{1 + \epsilon(m)^{1-\gamma}} \right)^{-\frac{1}{\gamma}} \exp \left\{ \left(1 - \frac{1}{\gamma}\right)\mu_{\delta_t} + \frac{1}{2}\left(1 - \frac{1}{\gamma}\right)^2\sigma_{\delta_t}^2 \right\} \left[ 1 - \Phi \left( \frac{\ln \frac{\bar{\delta}^\epsilon}{\delta_t} - \mu_{\delta_t} - (1 - \frac{1}{\gamma})\sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right) \right], \end{aligned} \quad (15)$$

where  $\mu_{\delta_t}$  and  $\sigma_{\delta_t}^2$  are respectively the mean and variance of the  $\mathcal{F}_t$ -conditional normal variable  $\ln \frac{\delta_T}{\delta_t}$  (see Appendix A).

PROOF: See Appendix C. □

**Corollary 3.3** (Optimal trading strategies of a green bond investor). *For a green bond investor, the optimal fractions of the wealth invested in the stock and green bond at time  $t$  ( $t > 0$ ) are given by:*

$$\xi_t = \frac{-R_\delta(\delta, t)\delta_t\tilde{\kappa}}{X_t^*\sigma_S\sqrt{1 - \rho_{GS}^2}}, \quad (16)$$

$$\beta_t = \frac{-R_\delta(\delta, t)\delta_t\kappa}{X_t^*\sigma_G(t, T)} + \frac{R_\delta(\delta, t)\delta_t\tilde{\kappa}\rho_{GS}}{X_t^*\sigma_G(t, T)\sqrt{1 - \rho_{GS}^2}}, \quad (17)$$

with

$$\begin{aligned}
R_\delta(\delta, t) := & \frac{\partial X_t^*}{\partial \delta_t} = -\frac{1}{\delta_t \gamma} \left( \frac{\lambda \delta_t}{1 + \iota(m)^{1-\gamma}} \right)^{-\frac{1}{\gamma}} \exp \left\{ \left( 1 - \frac{1}{\gamma} \right) \mu_{\delta_t} + \frac{1}{2} \left( 1 - \frac{1}{\gamma} \right)^2 \sigma_{\delta_t}^2 \right\} \\
& \cdot \Phi \left( \frac{\ln \frac{\bar{\delta}_t^\iota}{\delta_t} - \mu_{\delta_t} - \left( 1 - \frac{1}{\gamma} \right) \sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right) \\
& - \frac{1}{\delta_t \sigma_{\delta_t}} \left( \frac{\lambda \delta_t}{1 + \iota(m)^{1-\gamma}} \right)^{-\frac{1}{\gamma}} \exp \left\{ \left( 1 - \frac{1}{\gamma} \right) \mu_{\delta_t} + \frac{1}{2} \left( 1 - \frac{1}{\gamma} \right)^2 \sigma_{\delta_t}^2 \right\} \varphi \left( \frac{\ln \frac{\bar{\delta}_t^\iota}{\delta_t} - \mu_{\delta_t} - \left( 1 - \frac{1}{\gamma} \right) \sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right) \\
& - \frac{\underline{x}^G}{\delta_t \sigma_{\delta_t} m} \exp \left\{ \mu_{\delta_t} + \frac{1}{2} \sigma_{\delta_t}^2 \right\} \left[ \varphi \left( \frac{\ln \frac{\bar{\delta}_t^\epsilon}{\delta_t} - \mu_{\delta_t} - \sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right) - \varphi \left( \frac{\ln \frac{\bar{\delta}_t^\epsilon}{\delta_t} - \mu_{\delta_t} - \sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right) \right] \\
& - \frac{1}{\delta_t \gamma} \left( \frac{\lambda \delta_t}{1 + \epsilon(m)^{1-\gamma}} \right)^{-\frac{1}{\gamma}} \exp \left\{ \left( 1 - \frac{1}{\gamma} \right) \mu_{\delta_t} + \frac{1}{2} \left( 1 - \frac{1}{\gamma} \right)^2 \sigma_{\delta_t}^2 \right\} \left[ 1 - \Phi \left( \frac{\ln \frac{\bar{\delta}_t^\epsilon}{\delta_t} - \mu_{\delta_t} - \left( 1 - \frac{1}{\gamma} \right) \sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right) \right] \\
& + \frac{1}{\delta_t \sigma_{\delta_t}} \left( \frac{\lambda \delta_t}{1 + \epsilon(m)^{1-\gamma}} \right)^{-\frac{1}{\gamma}} \exp \left\{ \left( 1 - \frac{1}{\gamma} \right) \mu_{\delta_t} + \frac{1}{2} \left( 1 - \frac{1}{\gamma} \right)^2 \sigma_{\delta_t}^2 \right\} \varphi \left( \frac{\ln \frac{\bar{\delta}_t^\epsilon}{\delta_t} - \mu_{\delta_t} - \left( 1 - \frac{1}{\gamma} \right) \sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right),
\end{aligned}$$

where  $\varphi(\cdot)$  is the density function of the standard normal distribution, i.e.  $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ .

PROOF: We let  $R(\delta, t)$  denote the optimal wealth  $X_t^*$  given in Equation (15). Applying the Itô's lemma and comparing the coefficients of  $dW_1(t)$  and  $dW_2(t)$  with those in (8), we obtain

$$\begin{aligned}
-R_\delta(\delta, t) \delta_t \kappa &= (\beta \sigma_G(t, T) + \xi \rho_{GS} \sigma_S) X_t^*, \\
-R_\delta(\delta, t) \delta_t \tilde{\kappa} &= (\xi \sigma_S \sqrt{1 - \rho_{GS}^2}) X_t^*,
\end{aligned}$$

and the closed form of the optimal investment strategy follows.  $\square$

### 3.2 Optimal Ratio $m^*$

So far, we have assumed that  $m$  is an exogenous parameter which controls the utility gain-or-loss measurement. A natural question arises regarding how the green bond investor selects this parameter in her investment decision. Below, we demonstrate that this parameter can be addressed by examining the subsequent maximization problem:

$$\sup_{m \in [\underline{m}, \bar{m}]} \sup_{X_T} \mathbb{E} \left[ \tilde{U}(X_T) \right], \quad \text{s.t. } \mathbb{E}[X_T \delta_T] \leq X_0, \quad (18)$$



where  $\underline{m}, \bar{m}$  are the lower bound and upper bound of  $m$  which, as discussed below, are suggested by an additional risk control step. Hence, given the optimal wealth  $X_T^*$  defined in Theorem 3.1, we want to solve

$$\sup_{m \in [\underline{m}, \bar{m}]} \mathbb{E}[\tilde{U}(X_T^*)] =: \sup_{m \in [\underline{m}, \bar{m}]} \mathcal{M}(m).$$

Since the state price density  $\delta_T$  is atomless (it has continuous distribution), it can be shown directly (see e.g. Lemma A11 of Ng and Nguyen (2023)) that by integrability condition (12) and dominated convergence theorem that  $\mathcal{M}$  is continuously differentiable. Therefore,  $\sup_{m \in [\underline{m}, \bar{m}]} \mathcal{M}(m)$  exists and can be found among  $\mathcal{M}(\underline{m}), \mathcal{M}(\bar{m})$  and  $\mathcal{M}(m^*)$ , where  $m^*$  is a critical value that solves the first order condition  $\mathcal{M}'(m) = 0$ .

Let us give a simple criterion for the selection of  $[\underline{m}, \bar{m}]$ . We define  $\alpha(m)$  as the default probability that an  $m$  proportion of real investment fails to meet the CSR investment goal. More specifically,

$$\alpha(m) := P(mX_T^* < \underline{x}^G) = 1 - \Phi\left(\frac{\ln h(m) - \mu_{\delta_T}}{\sigma_{\delta_T}}\right)$$

where  $h(m) = (\underline{x}^G)^{-\gamma} \frac{m^{\gamma+\epsilon m}}{\lambda(m)}$ . Note that  $\delta_T$  is log-normally distributed and independent of  $m$ . We want to guarantee that the default probability  $\alpha(m)$  falls to some given interval, i.e.  $\alpha(m) \in [\underline{\alpha}, \bar{\alpha}]$ , which in turn gives a reasonable interval of  $m$ .

## 4 Numerical Illustration

In this section, we conduct numerical exercises to illustrate the optimal terminal wealth and trading strategies given by Theorem 3.1 and Corollary 3.3 respectively. Insurance companies, as important institutional investors, are increasingly active in CSR investments. In the following analysis, we consider an insurance company as a green bond investor, assuming that she is risk-averse with 100 monetary unit initially ( $X_0 = 100$ ). She is ready to allocate her assets over 20 years. Her CSR investment goal is set to be 30% of the initial wealth  $X_0$ , i.e.,  $\underline{x}^G$  is 30 monetary unit. According to the investor's utility function's "loss aversion" feature,  $\epsilon > \iota > 0$ . We choose  $\iota = 0.8$  and  $\epsilon = 2$  to reflect that the investor gains an increase in utility of 80% of the actual size while suffering a utility loss of 2 times the actual size. We set the lower bound of the default probability ( $\underline{\alpha}$ ), as 9.5%. This corresponds to the regulated insurer's annual default probability

0.5% by the capital requirements of Solvency II on Pillar I (Fischer and Schlütter, 2015), which means that the default probability for an investment period of 20 years becomes  $1 - (1 - 0.5\%)^{20} = 9.5\%$ . As for the upper bound of the default probability ( $\bar{\alpha}$ ), we choose a larger annual default probability of 1%, which will lead to a 20-year default probability of  $1 - (1 - 1\%)^{20} = 18.2\%$ .

Furthermore, we choose the parameters for asset prices as presented in Table 1 concerning Munk et al. (2004) and Brennan and Xia (2002). Considering the economic development, we adjust the market prices of the stock and green bond to be slightly higher than those estimated in these studies. Following the finding of Nguyen et al. (2021), we choose a negative correlation between the green bond and stock ( $\rho_{GS}$ ).

Parameter	Description	Value
<b>Characteristic parameters for investing</b>		
$X_0$	Initial wealth	100
$\gamma$	Risk aversion coefficient	5*
$T$	Investment period	20
$\underline{x}^G$	Reference level of the green bond investment	30
$\iota$	Scaling parameter for utility gains	0.8
$\epsilon$	Scaling parameter for utility losses	2
$\underline{\alpha}$	Lower bound of the default probability	9.5%
$\bar{\alpha}$	Upper bound of the default probability	18.2%
<b>Parameters for asset prices</b>		
$\psi$	Market price of the stock risk	0.4
$\sigma_S$	Volatility of the stock price	0.1468
$\kappa$	Market price of the green bond risk	0.3
$\sigma_G$	Volatility of the green bond price	0.2696
$\rho_{GS}$	Correlation between the stock and green bond	-0.2
$a$	Mean-reverting speed factor of the interest rate	0.0395
$\frac{b}{a}$	Long-run mean of the interest rate	0.0369
$\sigma$	Volatility of the short rate	0.0195

\* We have also carried out similar numerical analyses for the investors with  $\gamma = 0.8 \in (0, 1)$ .

The corresponding discussion is presented in Subsection 4.3.

Table 1: Baseline parameters. Characteristic parameters for investing are assumed and a subsequent sensitivity analysis will be provided to examine the impact of some key parameters on our results. Asset price related parameters are chosen with reference to the literature (Munk et al., 2004; Brennan and Xia, 2002; Nguyen et al., 2021).

## 4.1 Base-case Analysis

### Optimal Terminal Wealth

With the baseline parameters given in Table 1, we obtain the reasonable range for  $m$  as  $[17.14\%, 21.48\%]$ . By solving the maximization problem (18), the optimal ratio  $m^*$  is found to be around 21.48%; and the corresponding Lagrangian multiplier approximates to be  $1.76 \times 10^{-8}$ .

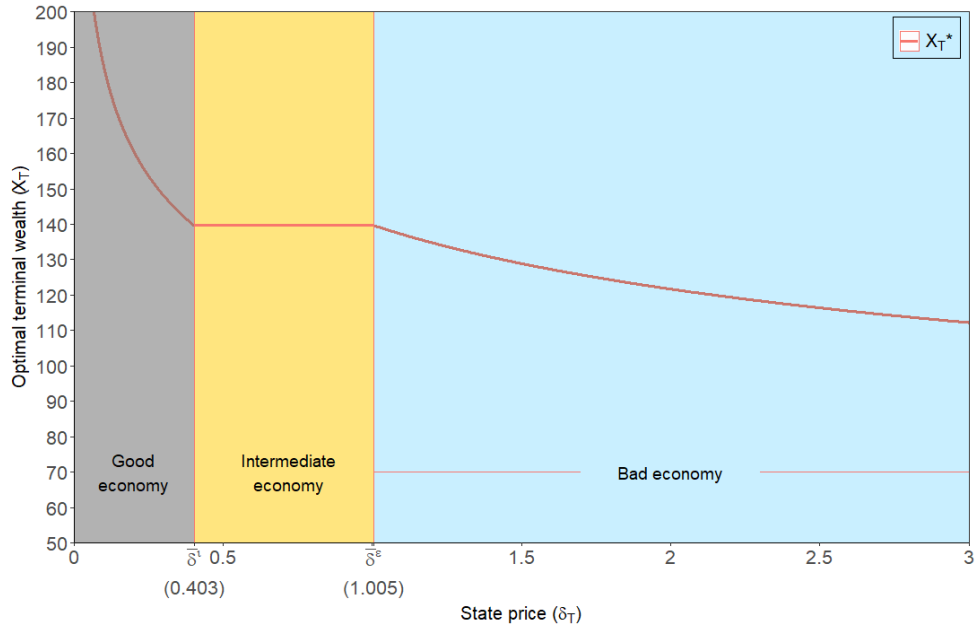


Figure 3: Optimal terminal wealth of a green bond investor. Parameters are chosen as presented in Table 1. Using the baseline parameters, two critical points are calculated to be  $\bar{\delta}^l = 0.403$  and  $\bar{\delta}^e = 1.005$  respectively. And the optimal ratio  $m$  is found to be 21.48%.

Figure 3 depicts the optimal terminal wealth of the green bond investor. Low values of the state price represent the economy in good states and vice versa. As shown in Theorem 3.1, the optimal terminal wealth of the green bond investor shows a three-region feature. In the good economic region (shaded in gray, before the economy at time  $T$  hits the first critical point  $\delta_T < \bar{\delta}^l = 0.403$ ), the optimal final wealth of the green bond investor declines as economic condition worsens. When the final economy falls in the intermediate economic region (shaded in yellow,  $\bar{\delta}^l \leq \delta_T < \bar{\delta}^e = 1.005$ ), the green bond investor's optimal terminal wealth stays at a constant terminal wealth level ( $\underline{x}^G/m \approx 139.665$ ), which does not change with the economic state. Then, if the final economy appears in

the bad economic region (shaded in blue,  $\delta_T \geq \bar{\delta}^\epsilon$ ), the green bond investor's wealth path again decreases with the state price.

### Optimal Trading Strategies

After exhibiting the optimal terminal wealth, let us examine the trading strategies concerning the optimal allocations to various assets. Given that trading strategies differ with the investment year  $t$ , we employ Monte Carlo simulation to iterate through 100,000 trials, generating state prices at each time  $t$ . We then derive the anticipated trading strategies, both in terms of amounts and fractions of wealth over the years, using the parameters outlined in the provided Table 1. Figure 4 illustrates the evolution of the

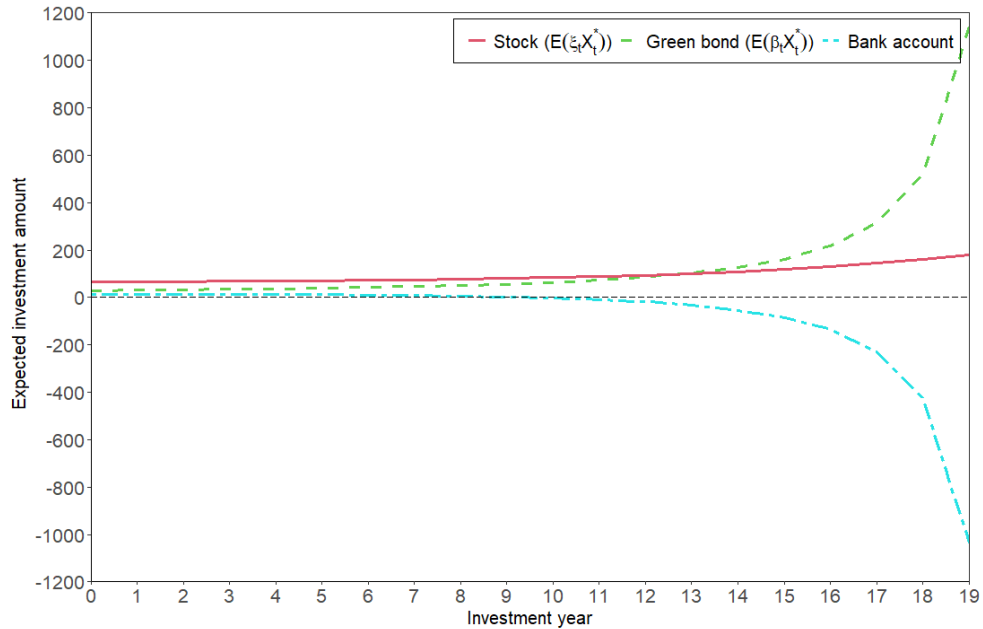


Figure 4: Evolution of the trading strategies expected in amounts. Parameters are chosen as presented in Table 1. The red solid, green dashed and blue dotted-dashed curve in the figure represent respectively the expected amount invested in the stock, green bond and bank account.

expected amounts invested in the green bond (c.f. the green dashed curve), stock (c.f. the red solid curve) and the bank account (c.f. the blue dotted-dashed curve). During the first 9 years, the investor takes long positions in all the three assets, while allocates most in the stock and least in the bank account. The long position in the stock increases smoothly with time; while the expected amount invested in the green bond augments more rapidly, and the predominance of the long-term green bond in the portfolio is ob-

served during the last 5 years. This is primarily influenced by its declining volatility (c.f. Corollary 3.3) and the fact that the investor is heavily investing in the green bond. Indeed, as it can be observed in Figure 5 that the expected fraction of wealth invested in the green bond has risen from 68.16% in the 15th year to 385.10% in the last year. Holding the green bond becomes more advantageous as the investment approaches maturity. The shorter the time to maturity, the more investment fraction of the green bond becomes. Given an extremely short time to maturity, the green bond can be regarded as a nearly risk-free asset. To finance the investments in the stock and green bond, the long position in the bank account diminishes in year 9 and transits into a rapid increase in short position.

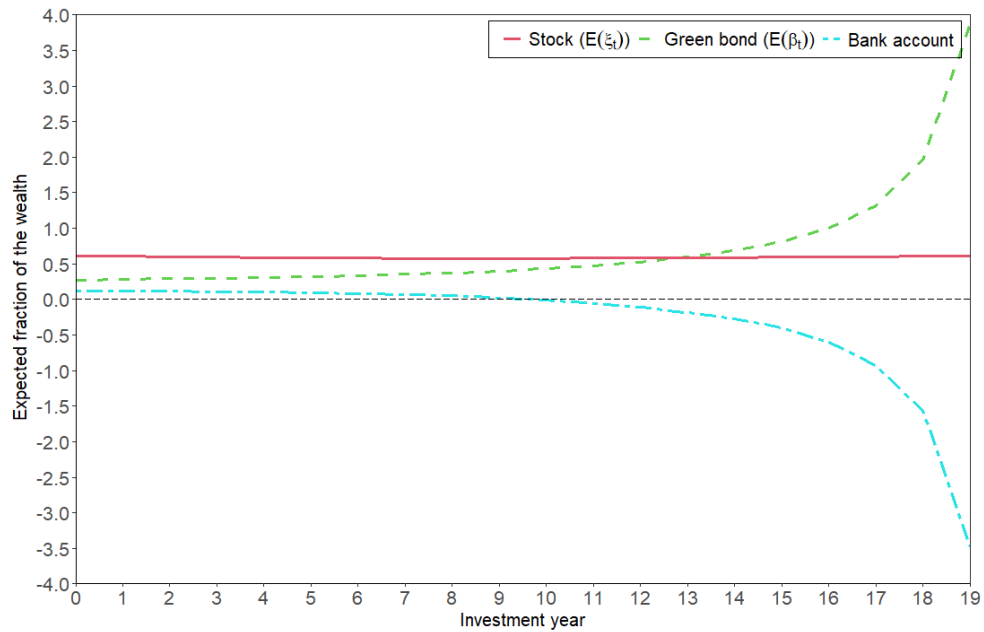


Figure 5: Trading strategies expected in fractions of the wealth over years. Parameters are chosen as presented in Table 1. The red solid, green dashed and blue dotted-dashed curve in the figure represent respectively the expected fraction of the wealth invested in the stock, green bond and bank account.

## 4.2 Sensitivity Analysis

Since the investment amounts exhibit similar sensitivity towards investigated parameters as the investment fractions do, and to enhance the clarity of our presentation, we exclusively display the outcomes related to the percentages of wealth allocated to the green

bond and stock. Note that we compute the optimal ratio  $m^*$  with respect to different values of the investigated parameters (i.e. the scaling parameter for utility losses  $\epsilon$  and the CSR investment goal  $\underline{x}^G$ ).

Figure 6 illustrates the sensitivity of the optimal investment fractions towards the scaling parameter for utility losses  $\epsilon$ , representing the loss aversion of the green bond investor. With parameters given in Table 1, we find that  $\epsilon$  has a negative impact on the fractions of wealth invested in the stock and green bond. That is, a larger  $\epsilon$  leads to lower percentages of wealth invested in both the stock and green bond. Since a greater  $\epsilon$  represents that the green bond investor is more loss averse, she is likely to invest less in the green bond and stock to protect herself against losses.

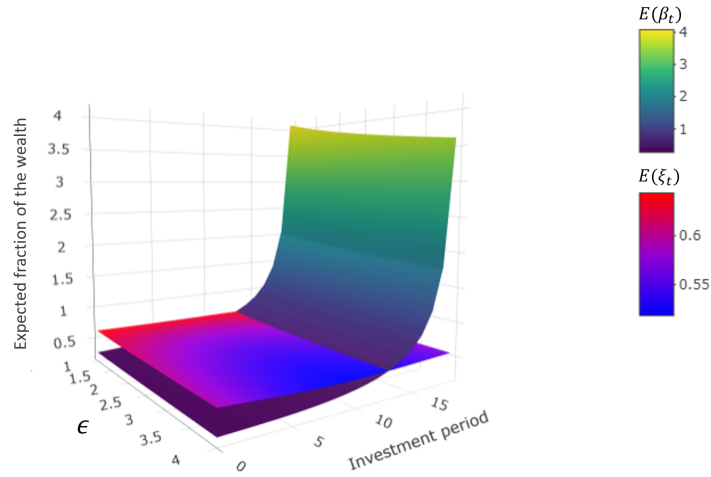


Figure 6: Expected investment fractions in the green bond and stock over years for various  $\epsilon$ . Parameters are chosen as presented in Table 1. The scaling parameter of utility losses ( $\epsilon$ ) varies from 1 to 4.

As for the CSR investment goal ( $\underline{x}^G$ ), using parameters from Table 1 where  $\gamma = 5$ , Figure 7 shows a positive effect of  $\underline{x}^G$  on optimal holdings in the green bond and stock. And a higher  $\underline{x}^G$  results in higher fractions of wealth being invested in the stock and green bond. Moreover, for the given set of parameters, it is observed that  $\underline{x}^G$  only influences the holdings in the green bond and stock to a small extent. For instance, the expected fraction of wealth invested in the green bond in year 10 increases from 43.42% to 44.29% with  $\underline{x}^G$  rising from 5 to 95 (i.e. from 5% to 95% of the initial wealth), and correspondingly the expected fraction invested in the stock increases from 57.68% to

58.84% (see Figure 8). One possible reason is that the expected utility of the green bond investor (who has a relative risk aversion of 5) is largely affected by the utility gain-or-loss in her goal-based preference structure, hence, failing to meet the goal may lead to a significant utility loss for the green bond investor with respect to this measurement. Further, it's also noteworthy that the consideration of default probability in determining  $m^*$ , a parameter responsive to changes in  $\underline{x}^G$ , suggests that fluctuations in the goal level  $\underline{x}^G$  are not anticipated to have a significant impact on the composition of both stock and green bond holdings.

To assess the influence of risk aversion on our results, we expand our examination to include cases where the risk aversion parameter  $\gamma$  takes values of 2, 3, and 4. Notably, our initial findings remain consistent across these scenarios. However, a significant departure becomes evident when the risk aversion parameter  $\gamma$  is limited to the interval  $(0, 1)$ . In this specific range, we observe distinct variations in the results, particularly concerning changes in the CSR investment goal  $\underline{x}^G$ . More details on these observations are provided in the following subsection.

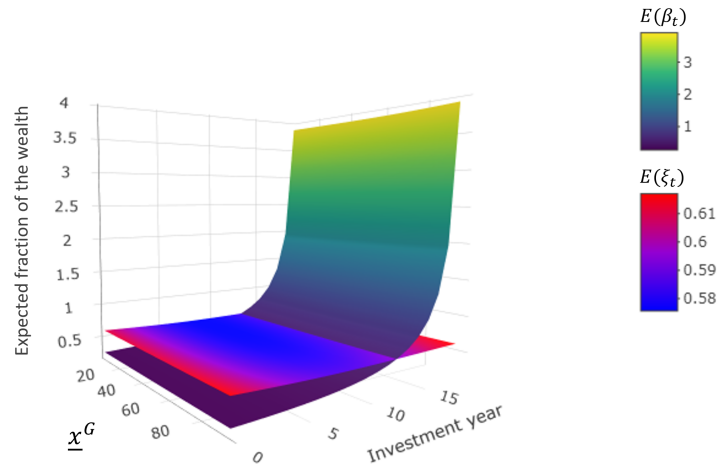


Figure 7: Expected investment fractions in the green bond and stock over years for various  $\underline{x}^G$ . Parameters are chosen as presented in Table 1. The CSR investment goal ( $\underline{x}^G$ ) changes from 5% to 95% of the initial wealth.

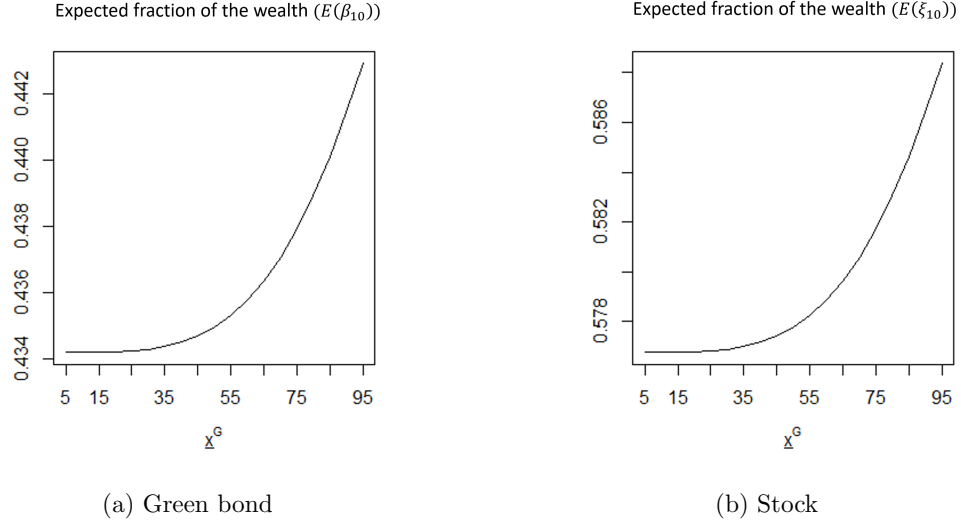


Figure 8: Expected fractions of the wealth invested in the green bond (Panel (a)) and stock (Panel (b)) in the 10th investment year for various  $\underline{x}^G$ . Other parameters are from Table 1. The CSR investment goal ( $\underline{x}^G$ ) changes from 5% to 95% of the initial wealth.

### 4.3 Numerical Results for Investors with $\gamma \in (0, 1)$

To showcase the outcomes of sensitivity analysis in the context where  $\gamma$  is within the range  $(0, 1)$ , we set  $\gamma = 0.8$ . Then, we conduct the same numerical analyses as presented in Subsections 4.1 and 4.2. Similarly, we also compute the optimal  $m^*$  with respect to different values of parameters under investigation. Most results are similar to the case of  $\gamma > 1$ ; while in the case of  $\gamma \in (0, 1)$ , the green bond investor invests large fractions of wealth in the green bond and stock. For example, the green bond investor with  $\gamma = 0.8$  allocates 300.35% and 398.99% of the wealth in the green bond and stock when  $\underline{x}^G = 30$  and  $t = 10$  (see Figure 9). Furthermore, interestingly, we find that the CSR investment goal ( $\underline{x}^G$ ) negatively affects the green bond investor's optimal holding in the green bond in such a case.

To find out the reasons behind, we have done further explorations on the optimal ratio  $m^*$ , and numerically find  $m^*$  is determined by the upper bound  $\bar{m}$ , which is given as the lower bound of the default probability. Keeping in mind that by incorporating an additional risk management, the probability of not achieving the CSR investment goal is fixed in our analysis, which is done for both cases  $\gamma > 1$  and  $\gamma \in (0, 1)$ . This would suggest that, for the green bond investor with  $\gamma \in (0, 1)$ , her expected utility may be



dominantly affected by the first component (i.e. the expected utility on her final wealth). Higher investment fractions in the green bond and stock may result in a larger probability of monetary losses, and thus a lower expected utility on the final wealth. In this case, when  $\underline{x}^G$  increases, she reduces her holdings in the green bond and stock to realize an overall higher expected utility. In addition, the impact of the  $\underline{x}^G$  on the fractions of wealth invested in the green bond and stock is found to be small as well.

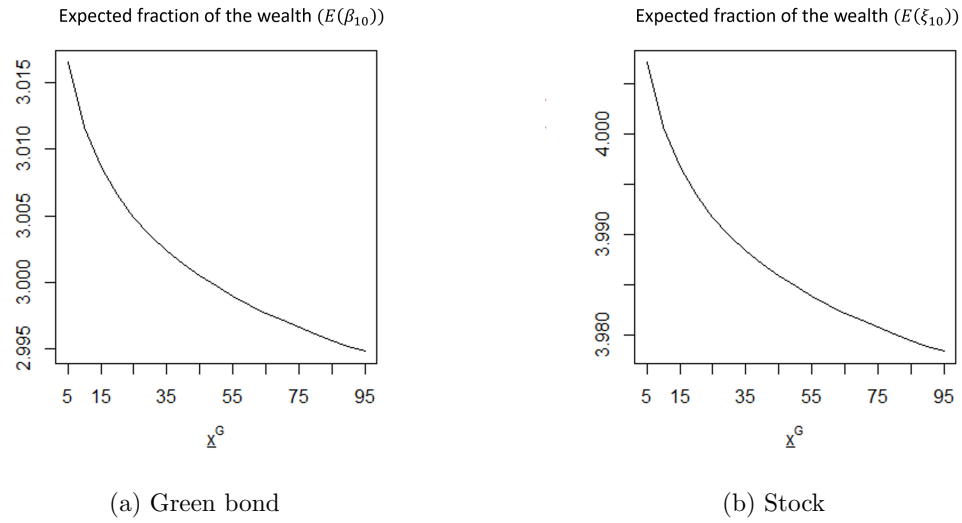


Figure 9: Expected fractions of the wealth invested in the green bond (Panel (a)) and stock (Panel (b)) in the 10th investment year for various  $\underline{x}^G$  when  $\gamma = 0.8$ . Other parameters are from Table 1. The CSR investment goal ( $\underline{x}^G$ ) changes from 5% to 95% of the initial wealth.

#### 4.4 Impact of the Correlation

So far, the analysis is conducted for a correlation coefficient  $\rho_{GS} = -0.2$ . As indicated by Corollary 3.3, the correlation may exert influence on the optimal fractions of wealth invested in the green bond and stock, and the correlation may vary by markets and countries (Gao et al., 2023). Therefore, we have also examined various correlations, where most of them generate similar results and do not affect our conclusions above. However, when the green bond and stock is strongly positively correlated (e.g.,  $\rho_{GS} = 0.8$ <sup>3</sup>), the results display some differences. For example, when  $\rho_{GS}$  is set as 0.8, the optimal holding

<sup>3</sup>Gao et al. (2023) find that the correlation between the global green bond (S&P green bond index) and stock (S&P500 composite index) is 0.638; and the correlation between Chinese green bond (China green bond index) and stock (S&P500 composite index) is 0.92 during the investigation period of 2012–2022.

in the green bond becomes negative (see Figure 10). In other words, in such a case, the investor will short sell the green bond and initiate a long position in the stock.

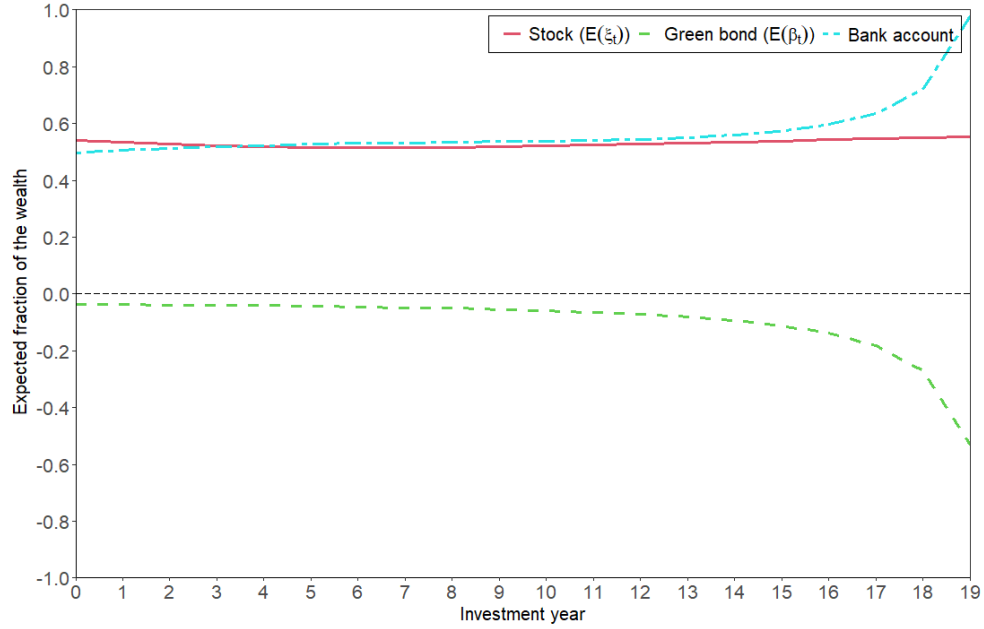


Figure 10: Trading strategies expected in fractions of the wealth over years when  $\rho_{GS} = 0.8$ . Other parameters are chosen as presented in Table 1. The red solid, green dashed and blue dotted-dashed curve in the figure represent respectively the expected fraction of the wealth invested in the stock, green bond and bank account.

Moreover, with prices of the green bond and stock strongly positively correlated, a higher  $\epsilon$  results in short selling a lower fraction of wealth in the green bond and holding a lower fraction in the stock due to the green bond investor's loss aversion; and the greater the  $\underline{x}^G$ , the larger fraction of wealth is short sold in the green bond (and held in the stock) in terms of an investor with  $\gamma > 1$ , while the smaller the fraction of wealth is short sold in the green bond (and invested in the stock) if the investor has a relative risk aversion  $\gamma \in (0, 1)$ .

## 5 Concluding Remark

Climate change has emerged as a paramount concern in the 21st century. Within the investment sector, there is a notable shift towards CSR investments, with a particular surge in the popularity of green bonds. This trend is especially pronounced among institutional investors with CSR investment goals. In this article, we model the risk preferences of green bond investors bearing a CSR investment goal by incorporating a component of achieving the goal into their utility function. We derive the optimal terminal wealth as well as their optimal trading strategies. The green bond investors' optimal terminal wealth exhibits a three-region feature. A more ambitious target does not consistently increase funds allocated to green bonds: this allocation depends on investors' risk aversion as well as the correlation between the prices of the green bond and stock. Our results suggest potential avenues for further research into the investment behavior of green bond investors.

For future research, we might relax some of the assumptions on our goal-based utility function to capture more realistic investment behavior. For example, the green bond investor might not only want to maximize her expected utility in terms of her final wealth, but also the average investment in green bonds during the investment period. In addition, some intangible assets, such as reputation (Wren-Lewis, 2013), may be taken into account if the CSR target is achieved.

## References

- Afiuc, O., Bonsu, S. K., Manu, F., Knight, C. B., Panda, S., and Blankson, C. (2020). Corporate social responsibility and customer retention: Evidence from the telecommunication industry in Ghana. *Journal of Consumer Marketing*, 38(1):15–26.
- Bachelet, M. J., Becchetti, L., and Manfredonia, S. (2019). The green bonds premium puzzle: The role of issuer characteristics and third-party verification. *Sustainability*, 11(4):1098.
- Barua, S. and Chiesa, M. (2019). Sustainable financing practices through green bonds: What affects the funding size? *Business Strategy and the Environment*, 28(6):1131–1147.
- Braga, J. P., Semmler, W., and Grass, D. (2021). De-risking of green investments through a green bond market—empirics and a dynamic model. *Journal of Economic Dynamics and Control*, 131:104201.
- Branco, M. C. and Rodrigues, L. L. (2006). Corporate social responsibility and resource-based perspectives. *Journal of Business Ethics*, 69(2):111–132.
- Brennan, M. J. and Xia, Y. (2002). Dynamic asset allocation under inflation. *The Journal of Finance*, 57(3):1201–1238.
- Carroll, A. B. (1999). Corporate social responsibility: Evolution of a definitional construct. *Business & Society*, 38(3):268–295.
- Chandler, D. (2019). *Strategic corporate social responsibility: Sustainable value creation*. Sage Publications.
- Chen, A., Pelsser, A., and Vellekoop, M. (2011). Modeling non-monotone risk aversion using SAHARA utility functions. *Journal of Economic Theory*, 146(5):2075–2092.
- Chen, A. and Vellekoop, M. (2017). Optimal investment and consumption when allowing terminal debt. *European Journal of Operational Research*, 258(1):385–397.
- Chiesa, M. and Barua, S. (2019). The surge of impact borrowing: The magnitude and determinants of green bond supply and its heterogeneity across markets. *Journal of Sustainable Finance & Investment*, 9(2):138–161.
- Cox, J. C. and Huang, C.-f. (1989). Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory*, 49(1):33–83.

- Fischer, K. and Schlütter, S. (2015). Optimal investment strategies for insurance companies when capital requirements are imposed by a standard formula. *The Geneva Risk and Insurance Review*, 40:15–40.
- Gao, L., Guo, K., and Wei, X. (2023). Dynamic relationship between green bonds and major financial asset markets from the perspective of climate change. *Frontiers in environmental science*, 10:1109796.
- Gianfrate, G. and Peri, M. (2019). The green advantage: Exploring the convenience of issuing green bonds. *Journal of Cleaner Production*, 219:127–135.
- Hachenberg, B. and Schiereck, D. (2018). Are green bonds priced differently from conventional bonds? *Journal of Asset Management*, 19(6):371–383.
- Han, Y. and Li, J. (2022). Should investors include green bonds in their portfolios? Evidence for the USA and Europe. *International Review of Financial Analysis*, 80:101998.
- Islam, T., Islam, R., Pitafi, A. H., Xiaobei, L., Rehmani, M., Irfan, M., and Mubarak, M. S. (2021). The impact of corporate social responsibility on customer loyalty: The mediating role of corporate reputation, customer satisfaction, and trust. *Sustainable Production and Consumption*, 25:123–135.
- Karpf, A. and Mandel, A. (2018). The changing value of the “green” label on the US municipal bond market. *Nature Climate Change*, 8(2):161–165.
- Kennedy, C. and Corfee-Morlot, J. (2012). Mobilising investment in low carbon, climate resilient infrastructure. Working paper. <https://www.oecd-ilibrary.org/docserver/5k8zm3gxxmnq-en.pdf?expires=1681895096&id=id&accname=ocid43023316&checksum=F338B410E271D09103AFD555DF8B9B4C>.
- Kőszegi, B. (2006). Emotional agency. *The Quarterly Journal of Economics*, 121(1):121–155.
- Kőszegi, B. and Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4):1133–1165.
- Kőszegi, B. and Rabin, M. (2009). Reference-dependent consumption plans. *American Economic Review*, 99(3):909–36.
- Li, Z., Tang, Y., Wu, J., Zhang, J., and Lv, Q. (2020). The interest costs of green bonds: Credit ratings, corporate social responsibility, and certification. *Emerging Markets Finance and Trade*, 56(12):2679–2692.

- Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 247–257.
- Munk, C., Sørensen, C., and Vinther, T. N. (2004). Dynamic asset allocation under mean-reverting returns, stochastic interest rates, and inflation uncertainty: Are popular recommendations consistent with rational behavior? *International Review of Economics & Finance*, 13(2):141–166.
- Nanayakkara, M. and Colombage, S. (2019). Do investors in green bond market pay a premium? Global evidence. *Applied Economics*, 51(40):4425–4437.
- Ng, T. W. and Nguyen, T. (2023). Portfolio performance under benchmarking relative loss and portfolio insurance: From omega ratio to loss aversion. *ASTIN Bulletin: The Journal of the IAA*, 53(1):149–183.
- Nguyen, T. T. H., Naeem, M. A., Balli, F., Balli, H. O., and Vo, X. V. (2021). Time-frequency comovement among green bonds, stocks, commodities, clean energy, and conventional bonds. *Finance Research Letters*, 40:101739.
- Park, D., Park, J., and Ryu, D. (2020). Volatility spillovers between equity and green bond markets. *Sustainability*, 12(9):3722.
- Partridge, C. and Medda, F. R. (2020). The evolution of pricing performance of green municipal bonds. *Journal of Sustainable Finance & Investment*, 10(1):44–64.
- Pástor, L., Stambaugh, R. F., and Taylor, L. A. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2):550–571.
- Petkoski, D. and Twose, N. (2003). Public policy for corporate social responsibility. *WBI Series on Corporate Responsibility*. 7–25.
- Reboredo, J. C. (2018). Green bond and financial markets: Co-movement, diversification and price spillover effects. *Energy Economics*, 74:38–50.
- Reboredo, J. C. and Ugolini, A. (2020). Price connectedness between green bond and financial markets. *Economic Modelling*, 88:25–38.
- Smith, N. C. (2003). Corporate social responsibility: Whether or how? *California Management Review*, 45(4):52–76.

- Statista (2023). Value of green bonds issued worldwide from 2014 to 2022. Technical report. <https://www.statista.com/statistics/1289406/green-bonds-issued-worldwide/>.
- Tang, D. Y. and Zhang, Y. (2020). Do shareholders benefit from green bonds? *Journal of Corporate Finance*, 61:101427.
- Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188.
- Wren-Lewis, L. (2013). Commitment in utility regulation: A model of reputation and policy applications. *Journal of Economic Behavior & Organization*, 89:210–231.

## A State Price Density in Vasicek model

*Lemma A.1. In the Vasicek (1977) term structure model*

$$dr(t) = (b - ar(t)) dt + \sigma dW_1(t),$$

*it holds that*

$$\begin{aligned} r(t) &= e^{-a(t-s)} r(s) + b \int_s^t e^{-a(t-u)} du + \sigma \int_s^t e^{-a(t-u)} dW_1(u), \\ \int_s^t r(u) du &= \mathcal{D}(s, t) r(s) + b \int_s^t \mathcal{D}(u, t) du + \sigma \int_s^t \mathcal{D}(u, t) dW_1(u), \end{aligned}$$

*where  $\mathcal{D}(s, t) := e^{as} \int_s^t e^{-au} du$ . Furthermore,*

$$\begin{aligned} \mathbb{E} \left[ \int_0^t r(u) du \middle| \mathcal{F}_s \right] &= \int_0^s r(u) du + \frac{1}{a} (1 - e^{-a(t-s)}) \left( r(s) - \frac{b}{a} \right) + \frac{b}{a} (t - s), \\ \text{Var} \left[ \int_0^t r(u) du \middle| \mathcal{F}_s \right] &= \frac{\sigma^2}{a^2} \left[ (t - s) - \frac{2}{a} (1 - e^{-a(t-s)}) + \frac{1}{2a} (1 - e^{-2a(t-s)}) \right] \end{aligned}$$

*and an  $\mathcal{F}_t$ -conditional variable  $\ln \frac{\delta_T}{\delta_t}$  has a normal distribution with the mean and variance given by*

$$\mu_{\delta_t} = \mathbb{E} \left[ \ln \frac{\delta_T}{\delta_t} \middle| \mathcal{F}_t \right] = -\mathbb{E} \left[ \int_t^T r(u) du \middle| \mathcal{F}_t \right] - \frac{1}{2} \kappa^2 (T - t) - \frac{1}{2} \tilde{\kappa}^2 (T - t),$$

$$\sigma_{\delta_t}^2 = \text{Var} \left[ \ln \frac{\delta_T}{\delta_t} \middle| \mathcal{F}_t \right] = \text{Var} \left[ \int_t^T r(u) du \middle| \mathcal{F}_t \right] + \kappa^2(T-t) + \tilde{\kappa}^2(T-t) + 2\sigma\kappa \int_t^T \mathcal{D}(u, T) du.$$

PROOF:

Equation (5) leads to

$$\ln \delta_T = - \int_0^T r(u) du - \int_0^T \kappa dW_1(u) - \int_0^T \tilde{\kappa} dW_2(u) - \frac{1}{2} \int_0^T \kappa^2 du - \frac{1}{2} \int_0^T \tilde{\kappa}^2 du.$$

State differently,  $\ln \delta_T$  is normally distributed with

$$\begin{aligned} \mu_{\delta_T} &= \mathbb{E}[\ln \delta_T] = -\mathbb{E} \left[ \int_0^T r(u) du \middle| \mathcal{F}_0 \right] - \frac{1}{2} \int_0^T \kappa^2 du - \frac{1}{2} \int_0^T \tilde{\kappa}^2 du, \\ \sigma_{\delta_T}^2 &= \text{Var}[\ln \delta_T] = \text{Var} \left[ \int_0^T r(u) du \middle| \mathcal{F}_0 \right] + \kappa^2 T + \tilde{\kappa}^2 T + 2 \int_0^T \sigma \kappa \mathcal{D}(u, T) du. \end{aligned}$$

□

The following result can be deduced directly from the normal distribution of  $\ln \frac{\delta_T}{\delta_t} \middle| \mathcal{F}_t$ .

*Lemma A.2. Under the hypothesis of Lemma A.1, it holds that*

$$\begin{aligned} \mathbb{E} \left[ \left( \frac{\delta_T}{\delta_t} \right)^{1-\frac{1}{\gamma}} \mathbb{1}_{\{\delta_T < \bar{\delta}^\epsilon\}} \middle| \mathcal{F}_t \right] &= \exp \left\{ \left(1 - \frac{1}{\gamma}\right) \mu_{\delta_t} + \frac{1}{2} \left(1 - \frac{1}{\gamma}\right)^2 \sigma_{\delta_t}^2 \right\} \Phi \left( \frac{\ln \frac{\bar{\delta}^\epsilon}{\delta_t} - \mu_{\delta_t} - \left(1 - \frac{1}{\gamma}\right) \sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right), \\ \mathbb{E} \left[ \frac{\delta_T}{\delta_t} \mathbb{1}_{\{\bar{\delta}^\epsilon \leq \delta_T < \bar{\delta}^\epsilon\}} \middle| \mathcal{F}_t \right] &= \exp \left\{ \mu_{\delta_t} + \frac{1}{2} \sigma_{\delta_t}^2 \right\} \left[ \Phi \left( \frac{\ln \frac{\bar{\delta}^\epsilon}{\delta_t} - \mu_{\delta_t} - \sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right) - \Phi \left( \frac{\ln \frac{\bar{\delta}^\epsilon}{\delta_t} - \mu_{\delta_t} - \sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right) \right], \\ \mathbb{E} \left[ \left( \frac{\delta_T}{\delta_t} \right)^{1-\frac{1}{\gamma}} \mathbb{1}_{\{\delta_T \geq \bar{\delta}^\epsilon\}} \middle| \mathcal{F}_t \right] &= \exp \left\{ \left(1 - \frac{1}{\gamma}\right) \mu_{\delta_t} + \frac{1}{2} \left(1 - \frac{1}{\gamma}\right)^2 \sigma_{\delta_t}^2 \right\} \left[ 1 - \Phi \left( \frac{\ln \frac{\bar{\delta}^\epsilon}{\delta_t} - \mu_{\delta_t} - \left(1 - \frac{1}{\gamma}\right) \sigma_{\delta_t}^2}{\sigma_{\delta_t}} \right) \right]. \end{aligned}$$



## B Proof of Theorem 3.1

Theorem 3.1 is proved by two following steps.

**Step 1:** First, we consider the following static optimization problem  $\max_{x \geq 0} \Psi_\lambda(x, \delta)$ , where the Lagrangian function  $\Psi_\lambda(x, \delta)$  is defined, for  $\lambda > 0$  and  $\delta > 0$  given, by

$$\Psi_\lambda(x, \delta) := \tilde{U}(x) - \lambda \delta x = \begin{cases} (1 + \epsilon(m)^{1-\gamma})U(x) - \lambda \delta x - \epsilon U(\underline{x}^G), & mx < \underline{x}^G \\ (1 + \iota(m)^{1-\gamma})U(x) - \lambda \delta x - \iota U(\underline{x}^G), & mx \geq \underline{x}^G. \end{cases}$$

The proof is then relied on the following Lemma.

*Lemma B.1. For  $\lambda > 0$  given, define*

$$(1 + \epsilon(m)^{1-\gamma})U'(\underline{x}^G/m)/\lambda := \bar{\delta}^\epsilon \quad \text{and} \quad (1 + \iota(m)^{1-\gamma})U'(\underline{x}^G/m)/\lambda := \bar{\delta}^\iota.$$

*It then holds that*

$$\operatorname{argmax}_{x \geq 0} \Psi_\lambda(x, \delta) = I^\iota(\lambda \delta) \mathbb{1}_{\delta < \bar{\delta}^\iota} + \frac{\underline{x}^G}{m} \mathbb{1}_{\bar{\delta}^\iota \leq \delta < \bar{\delta}^\epsilon} + I^\epsilon(\lambda \delta) \mathbb{1}_{\delta \geq \bar{\delta}^\epsilon}.$$

**PROOF:** The optimum of  $\Psi_\lambda(x, \delta)$  can be determined via a careful comparison of the local maximizers  $I^\iota(\lambda \delta)$  and  $I^\epsilon(\lambda \delta)$  with  $\underline{x}^G/m$ . Recall that from definition,  $\bar{\delta}^\iota < \bar{\delta}^\epsilon$ . Suppose first that  $\delta < \bar{\delta}^\iota$ , i.e.  $I^\epsilon(\lambda \delta) > I^\iota(\lambda \delta) > \frac{\underline{x}^G}{m}$ , we observe that  $\Psi_\lambda$  is increasing in  $[0, I^\iota(\lambda \delta)]$  and decreasing in  $[I^\iota(\lambda \delta), \infty)$ , which implies that  $I^\iota(\lambda \delta)$  is the global optimizer. Assume now  $\delta \geq \bar{\delta}^\epsilon$ , i.e.  $I^\iota(\lambda \delta) < I^\epsilon(\lambda \delta) \leq \frac{\underline{x}^G}{m}$ . In this case the Lagrangian  $\Psi_\lambda$  is increasing in  $[0, I^\epsilon(\lambda \delta)]$  and decreasing in  $[I^\epsilon(\lambda \delta), \infty)$ , which implies that  $I^\epsilon(\lambda \delta)$  is the maximizer. Finally, for  $\bar{\delta}^\iota \leq \delta < \bar{\delta}^\epsilon$ , it is straightforward to see that  $\frac{\underline{x}^G}{m}$  is the optimizer as  $\Psi_\lambda$  is increasing in  $[0, \frac{\underline{x}^G}{m}]$  and decreasing in  $[\frac{\underline{x}^G}{m}, \infty)$ .  $\square$

Applying Lemma B.1 for each  $\omega \in \Omega$ , we can conclude that the optimal terminal wealth of a green bond investor is given by

$$X_T^* = \operatorname{argmax}_{X_T} \Psi_\lambda(X_T, \delta_T) = I^\iota(\lambda \delta_T) \mathbb{1}_{\delta_T < \bar{\delta}^\iota} + \frac{\underline{x}^G}{m} \mathbb{1}_{\bar{\delta}^\iota \leq \delta_T < \bar{\delta}^\epsilon} + I^\epsilon(\lambda \delta_T) \mathbb{1}_{\delta_T \geq \bar{\delta}^\epsilon}. \quad (19)$$

**Step 2.** To finish, we prove the existence of the Lagrangian multiplier  $\lambda$ . Observe first that the function  $\lambda \mapsto \varphi(\lambda) := \mathbb{E}[\delta_T X_T^*]$  is well-defined by the integrability condition (12) and is continuous and strictly decreasing in  $(0, \infty)$  as since  $I$  is strictly decreasing. Moreover, it can be observed from Inada's condition that  $\lim_{\lambda \rightarrow 0} \varphi(\lambda) = \infty$  and

$\lim_{\lambda \rightarrow \infty} \varphi(\lambda) = 0$ . By intermediate value theorem, this implies that the existence of a unique  $\lambda > 0$  for any  $X_0 > 0$  such that the budget constraint  $\varphi(\lambda) = X_0$  is fulfilled.

## C Proof of Theorem 3.2

By Theorem 3.1, the optimal final wealth  $X_T^*$  expressed in Equation (19), where the optimal Lagrangian multiplier  $\lambda$  is determined by the binding budget constraint

$$\begin{aligned} X_0 &= \mathbb{E} \left[ \delta_T I'(\lambda \delta_T) \mathbb{1}_{\delta_T < \bar{\delta}^\iota} + \delta_T \frac{x^G}{m} \mathbb{1}_{\bar{\delta}^\iota \leq \delta_T < \bar{\delta}^\epsilon} + \delta_T I^\epsilon(\lambda \delta_T) \mathbb{1}_{\delta_T \geq \bar{\delta}^\epsilon} \right] \\ &= \mathbb{E} \left[ \frac{(\lambda)^{-\frac{1}{\gamma}} (\delta_T)^{1-\frac{1}{\gamma}} \mathbb{1}_{\delta_T < \bar{\delta}^\iota}}{(1 + \iota(m)^{1-\gamma})^{-1/\gamma}} + \delta_T \frac{x^G}{m} \mathbb{1}_{\bar{\delta}^\iota \leq \delta_T < \bar{\delta}^\epsilon} + \frac{(\lambda)^{-\frac{1}{\gamma}} (\delta_T)^{1-\frac{1}{\gamma}} \mathbb{1}_{\delta_T \geq \bar{\delta}^\epsilon}}{(1 + \epsilon(m)^{1-\gamma})^{-1/\gamma}} \right]. \end{aligned}$$

Since  $\delta_T$  is essentially  $\frac{\delta_T}{\delta_0}|_{\mathcal{F}_0}$ , using Lemma A.2, the above expectations can be computed explicitly. Correspondingly, the multiplier  $\lambda$  is given by

$$\lambda = \left( \frac{X_0 - \frac{x^G}{m} \exp \left\{ \mu_{\delta_T} + \frac{1}{2} \sigma_{\delta_T}^2 \right\} \left[ \Phi \left( \frac{\ln \bar{\delta}^\epsilon(\lambda) - \mu_{\delta_T} - \sigma_{\delta_T}^2}{\sigma_{\delta_T}} \right) - \Phi \left( \frac{\ln \bar{\delta}^\iota(\lambda) - \mu_{\delta_T} - \sigma_{\delta_T}^2}{\sigma_{\delta_T}} \right) \right]}{\exp \left\{ (1 - \frac{1}{\gamma}) \mu_{\delta_T} + \frac{1}{2} (1 - \frac{1}{\gamma})^2 \sigma_{\delta_T}^2 \right\} \cdot \left( \frac{\Phi \left( \frac{\ln \bar{\delta}^\iota(\lambda) - \mu_{\delta_T} - (1 - \frac{1}{\gamma}) \sigma_{\delta_T}^2}{\sigma_{\delta_T}} \right)}{(1 + \iota(m)^{1-\gamma})^{-1/\gamma}} + \frac{1 - \Phi \left( \frac{\ln \bar{\delta}^\epsilon(\lambda) - \mu_{\delta_T} - (1 - \frac{1}{\gamma}) \sigma_{\delta_T}^2}{\sigma_{\delta_T}} \right)}{(1 + \epsilon(m)^{1-\gamma})^{-1/\gamma}} \right)} \right)^{-\gamma}. \quad (20)$$

Let us determine the optimal wealth at  $t \in (0, T)$ . To this end, note that  $(X_t^* \delta_t)_{0 \leq t \leq T}$  is the martingale under  $\mathbb{P}$ . Therefore, the optimal wealth at any time  $t$  between 0 and  $T$  is given by the following conditional expectation:

$$\begin{aligned} X_t^* &= \frac{1}{\delta_t} \mathbb{E} [\delta_T X_T^* | \mathcal{F}_t] \\ &= \frac{1}{\delta_t} \mathbb{E} \left[ \frac{(\lambda)^{-\frac{1}{\gamma}} (\delta_T)^{1-\frac{1}{\gamma}} \mathbb{1}_{\delta_T < \bar{\delta}^\iota}}{(1 + \iota(m)^{1-\gamma})^{-1/\gamma}} + \delta_T \frac{x^G}{m} \mathbb{1}_{\bar{\delta}^\iota \leq \delta_T < \bar{\delta}^\epsilon} + \frac{(\lambda)^{-\frac{1}{\gamma}} (\delta_T)^{1-\frac{1}{\gamma}} \mathbb{1}_{\delta_T \geq \bar{\delta}^\epsilon}}{(1 + \epsilon(m)^{1-\gamma})^{-1/\gamma}} \middle| \mathcal{F}_t \right] \\ &= \frac{(\lambda \delta_t)^{-\frac{1}{\gamma}}}{(1 + \iota(m)^{1-\gamma})^{-1/\gamma}} \mathbb{E} \left[ \left( \frac{\delta_T}{\delta_t} \right)^{1-\frac{1}{\gamma}} \mathbb{1}_{\{\delta_T < \bar{\delta}^\iota\}} \middle| \mathcal{F}_t \right] + \frac{x^G}{m} \mathbb{E} \left[ \frac{\delta_T}{\delta_t} \mathbb{1}_{\{\bar{\delta}^\iota \leq \delta_T < \bar{\delta}^\epsilon\}} \middle| \mathcal{F}_t \right] \\ &\quad + \frac{(\lambda \delta_t)^{-\frac{1}{\gamma}}}{(1 + \epsilon(m)^{1-\gamma})^{-1/\gamma}} \mathbb{E} \left[ \left( \frac{\delta_T}{\delta_t} \right)^{1-\frac{1}{\gamma}} \mathbb{1}_{\{\delta_T \geq \bar{\delta}^\epsilon\}} \middle| \mathcal{F}_t \right]. \end{aligned}$$

## Curriculum Vitae

CV was removed for data privacy protection reasons.