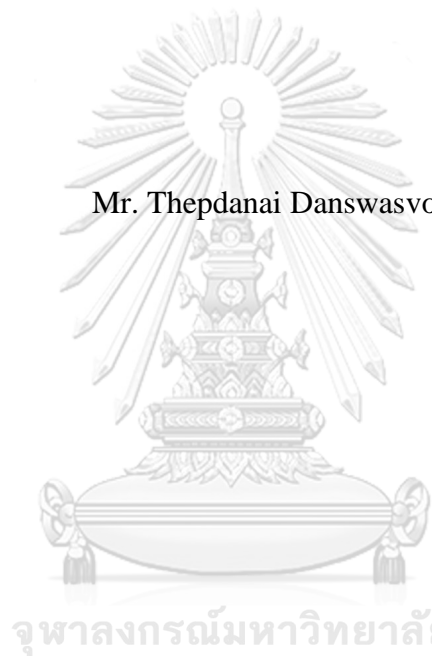Plan member's heterogeneity, economic regime effect and their implication on the management and sustainability of retirement funds

Mr. Thepdanai Danswasvong

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Quantitative Finance
Department of Banking and Finance
Faculty of Commerce and Accountancy
Chulalongkorn University
Academic Year 2017

ภาวะวิวิธพันธ์ของสมาชิกแผน, ผลกระทบสภาวะเศรษฐกิจ และ ความเกี่ยวพันต่อการ
บริหาร และ ความยั่งยืนของกองทุนเกษียณอายุ

นายเทพดนัย ด่านสวัสดิ์วงศ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาการเงินเชิงปริมาณ ภาควิชาการธนาคารและการเงิน
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2560
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Thesis Title	Plan member's heterogeneity, economic regime effect and their implication on the management and sustainability of retirement funds
By	Mr. Thepdanai Danswasvong
Field of Study	Quantitative Finance
Thesis Advisor	Assistant Professor Sira Suchintabandid, Ph.D.

Accepted by the Faculty of Commerce and Accountancy, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

.... Dean of the Faculty of Commerce and Accountancy
(Associate Professor Pasu Decharin, Ph.D.)

THESIS COMMITTEE

.... Chairman
(Associate Professor Manapol Ekkayokkaya, Ph.D.)

.... Thesis Advisor
(Assistant Professor Sira Suchintabandid, Ph.D.)

.... Examiner
(Assistant Professor Thaisiri Watewai, Ph.D.)

.... Examiner
(Tanawit Sae-Sue, Ph.D.)

.... External Examiner
(Anant Chiarawongse, Ph.D.)

จุฬาลงกรณ์มหาวิทยาลัย

CHULALONGKORN UNIVERSITY

เทพดนัย ด่านสวัสดิ์วงศ์ : ภาวะวิวิธพันธ์ของสมาชิกแผน, ผลกระทบสภาวะเศรษฐกิจ และ ความ เกี่ยวพันต่อการบริหาร และ ความยั่งยืนของกองทุนเกษียณอายุ (Plan member's heterogeneity, economic regime effect and their implication on the management and sustainability of retirement funds) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ. ดร. สิระ สุจินตะบัณฑิต, หน้า.

งานวิจัยชิ้นนี้ศึกษาความสำคัญของภาวะวิวิธพันธ์ของสมาชิกที่มีต่อการบริหารกองทุนบำนาญแบบ กำหนดประโยชน์ทดแทน เรานำเสนอแบบจำลองแบบใหม่ของกองทุนบำนาญแบบกำหนดประโยชน์ทดแทนที่มี สมาชิกหลายคนซึ่งอาจมีภาวะวิวิธพันธ์ในด้าน อายุ อัตราขึ้นเงินเดือน และ ด้านอื่นๆ เราเริ่มจากการแก้หาผลตอบ เชิงวิเคราะห์เพื่อให้ได้มาซึ่งวิธีการจัดการกองทุนที่เหมาะสมที่สุด เราแสดงให้เห็นว่าเงินสมทบของผู้อุปถัมภ์และ สัดส่วนการลงทุนในสินทรัพย์เสี่ยงที่เหมาะสมนั้นเป็นผลคูณไขว้ระหว่างมูลค่าคาดการณ์ของเงินผลประโยชน์ที่ ต้องจ่ายคืนให้สมาชิกหลังเกษียนกับสัมประสิทธิ์ส่วนลดตามภาวะวิวิธพันธ์ หลังจากนั้นเราแสดงให้เห็นว่าการมี อยู่ของภาวะวิวิธพันธ์อาจมีอิทธิพลอย่างมากต่อวิธีการบริหารกองทุนที่เหมาะสม และ การตัดสินใจของผู้บริหาร กองทุนโดยละเลยภาวะวิวิธพันธ์นั้นอาจก่อให้เกิดผลลัพธ์ที่แย่กว่าสิ่งที่ดีที่สุด

แต่อย่างไรก็ตาม การแก้หาวิธีการบริหารที่เหมาะสมสำหรับกองทุนที่มีสมาชิกหลายคนนั้นอาจต้องใช้ ทรัพยากรในการคำนวณที่สูงมากซึ่งจะเพิ่มขึ้นอย่างทวีคูณตามจำนวนของสมาชิกในแผน เรานำเสนอวิธีที่จะลด มิติของปัญหาลงโดยการทดแทนปัญหาของสมาชิกหลายคนด้วยอนุกรมของปัญหาที่มีสมาชิกหนึ่งเดียวซึ่งเป็น ตัวแทนของสมาชิกทุกคนในกองทุน (Adaptive Representative Agent หรือ ARA) ARA ของเรานั้นเป็น ค่าเฉลี่ยถ่วงน้ำหนักของสมาชิกกองทุนที่เหลืออยู่ โดยจะให้น้ำหนักที่มากกว่ากับสมาชิกที่มีมูลค่าคาดการณ์ที่ถูก ปรับปรุงโดยภาวะวิวิธพันธ์แล้วของเงินผลประโยชน์ที่ต้องจ่ายคืนให้กับสมาชิกนั้นๆสูง เราเปรียบเทียบ ประสิทธิภาพของตัวแทน ARA เมื่อเปรียบเทียบกับตัวแทนแบบค่าเฉลี่ย เราพบว่าวิธีใหม่ของเรานั้นดีกว่า โดยเฉพาะในสภาวะของตลาดลงทุนที่ไม่พึงประสงค์

ท้ายที่สุดเราศึกษาผลกระทบของภาวะวิวิธพันธ์ในระดับประเทศต่อความยั่งยืนของกองทุนเกษียณอายุ ประกันสังคม เราเสนอว่าภาวะวิวิธพันธ์ในประชากรนั้นเป็นกุญแจสำคัญในการควบรวมทฤษฎีต่างๆที่แข่งขันกัน อยู่ของนโยบายปฏิรูปเชิงกลยุทธ์และผลกระทบต่อความยั่งยืนของกองทุนประกันสังคม ภาวะวิวิธพันธ์ในด้านที่ ต่างกันจะส่งผลที่ไม่เหมือนกันต่ออัตราเงินที่สมทบเข้ากองทุนของลูกจ้างและต่อผลตอบแทนจากการลงทุนซึ่งจะ มีผลสืบเนื่องต่อความยั่งยืนของกองทุนต่อไป

ภาควิชา	การธนาคารและการเงิน	ลายมือชื่อนิสิต
สาขาวิชา	การเงินเชิงปริมาณ	ลายมือชื่อ อ.ที่ปรึกษาหลัก
ปีการศึกษา	2560	

5583053826 : MAJOR QUANTITATIVE FINANCE

THEPDANAI DANSWASVONG: Plan member's heterogeneity, economic regime effect and their implication on the management and sustainability of retirement funds. ADVISOR: ASST. PROF. SIRA SUCHINTABANDID, Ph.D., pp.

This research studies the importance of plan members' heterogeneity to the management of defined benefit (DB) pension fund. We propose a new multi-member model of DB pension fund that allows for heterogeneity in plan members' retirement ages, salary growths, and other characteristics. We first solve analytically for optimal management strategy and show that the sponsor's supplementary contribution and the fund's allocation in risky assets are determined by the cross-product between the fund's expected retirement liabilities and some heterogeneity-adjusted discount factors. We then demonstrate that the presence of heterogeneity can have a significant influence on the optimal management strategy and that a management decision made while ignoring heterogeneity will be suboptimal.

However, solving for the true optimality decision under multi-member setting requires high computational resources, which is exponentially increasing with the number of members. We suggest a way to reduce dimensionality by approximating the multi-member problem with a series of single-member (adaptive representative agent or ARA) problems. Our ARA is the weighted average of all the remaining plan members. Higher weight is given to plan member with higher heterogeneity-adjusted expected benefit liability. Compared to the simple average approach, our novel approach is shown to be better especially under unfavorable market conditions.

Lastly, we looked at the effect of heterogeneity on a country-level scale and try to understand its effects on the sustainability of a country's social security system. In our model, we argue that population heterogeneity is the key in unifying competing theories on strategic reforms and their effects on sustainability. Different aspects of heterogeneity have characteristically different effects to the system's aggregate taxation credit and investment return and thus the system's lifetime estimate.

Department:	Banking and Finance	Student's Signature	
Field of Study:	Quantitative Finance	Advisor's Signature	
Academic Year:	2017		

ACKNOWLEDGEMENTS

CONTENTS

Introduction

Types of Pension Funds

In the modern management of retirement funds, plan members are often grouped together and contributed towards a single fund. These retirement funds can be categorized as defined contribution (DC), defined benefit (DB) or social security based on the nature of the retirement benefit. In a defined contribution (DC) pension fund, contributions from the retirement plan investors are fixed and the benefits are solely dependent on the performance of the fund management.

As opposed to the DC pension fund, each DB plan participant is promised a pre-defined retirement income which typically corresponds to their final salary before retirement. In DC pension funds, plan members bear the investment risk while managers are not obligated to meet any target. In a DB fund, on the other hand, each plan participant is promised a pre-defined retirement benefit and the underfunding risk is borne by the plan sponsor. This mandate presents a challenge for the DB pension managers and raises many questions about the proper management of DB pension funds. These issues regarding DB pension funds will be the focus of this work.

In another type of retirement fund the social security system. The benefit and stakeholders structure are comparable to the defined benefit pension fund. Members of the social security plan contribute toward a centrally managed fund in return for promised benefits. Unlike in the DB fund, members of a social security plan are also compensated for sickness, maternity etc. in addition to the retirement income. Since the social security system is normally sponsored by the local government to support generations of the country's retired workforce.

Challenge of DB Management, and Importance of Estimating Projected Benefit Liability

The defined benefit pension plans all over the world are failing. More than 90% of DB schemes are now closed to new members. Most of the surviving schemes will pay out benefits at a far lower level than what members earned and paid for, over many years (*The Preventable Demise of Defined Benefit Pension Provision*, 2017).

To ensure the survivability of the plan, plan sponsors frequently fire investment managers for underperformance and hire them after they earn large positive excess returns as seen in Sialm, Starks, and Zhang (2015), Benartzi and Thaler (2001) and Huberman and Wei (2006). This is especially true for a DB plan since the underfunding risk is totally borne by the plan sponsors. Goyal and Wahal (2008) provide important evidence regarding the effects of DB pension plan sponsors on retirement money flows in which the plan sponsors often move from one investment manager to another when the performance is bad. Even though, this natural selection

process exists, the total market share of DB type of funds in the pension market has been decreasing throughout the 20th century (Farrell & Shoag, 2016).

The objective of this work is to address this problem and focus on the problem of DB pension fund management. Many reasons were attributed to the failure of defined benefit plans namely an increase in workers longevity, an increase in compliance costs, a decline in interest rates but in the end, it has always come back to the flaws in the fund's projected liability estimation (Farrell & Shoag, 2016). The issue of accurately determining the pension liability both in terms of magnitude and timing has been raised by many see Honda (2012), Ai, Brockett, and Jacobson (2015).

In a typical pension scheme, the projected liability is usually estimated collectively as if the plan members are homogeneous. Management decisions are then made under the same homogeneity assumptions. This has a direct impact on the performance of the fund which is the fund is the paramount concern for investment managers.

All in all, an accurate estimation of the fund's projected liability is extremely important for a DB fund. Plan sponsors needed to see the deficit to comprehend the current state of the plan. The liability is also important to the investment managers in which the fund's asset-liability structure is crucial in determining the course of action for the fund. In this work, we argue that the misestimation of the fund's liability arises from a cause that has thus far been overlooked by literature, namely, the heterogeneity of fund members.

The Existence of Plan Member's Characteristics Heterogeneity

In a pension plan of a big conglomerate or public pension plan, we observed plan members with a diverse personal profile such as age, salary, salary growth and its correlation with the capital market, health condition, the number of children, inherent risk tolerance and so on. Ayres and Nalebuff (2013) documented degrees of salary-stock market growth correlation across industries. For example, workers in financial, real estate and insurance industry can have more than fifty percent correlation with the stock market growth. On the other hand, workers in public relations, wholesale and retail can have as much as negative thirty percent correlation with the market.

Furthermore, an entry-level employee and an executive of a company are compensated differently but would exist in the same pension system. The salary growth of an executive is likely tied with the performance of the firm whereas the salary growth of an entry-level employee may as well be deterministic. The compensation package of two executives in a large conglomerate operating in multiple industries can differ significantly. This potential co-movement between income, retirement benefit liability and fund's investment opportunity has been shown to have a fundamental influence on the fund management decisions (see Cairns, Blake, and Dowd (2006) and Benzoni, Collin-Dufresne, and Goldstein (2007)).

Many researchers recognize the importance of plan members' heterogeneity and have studied its effects on the decisions of pension fund managers. Colombo and

Haberman (2005) study the impact of stochastic active-membership in which they allow age heterogeneity. Cremer, Gahvari, and Pestieau (2008) investigate pension funds whose members differ in fertility and ability to raise children. Honda (2012) studied pension funds whose members have different levels of risk aversion. Honda (2012) studied pension funds whose members have different levels of risk aversion. In the context of DC pension fund, Honda (2012) argues that participants in pension systems do not easily agree about the preferred portfolio strategy since the risk attitudes of pension participants may be heterogeneously distributed. Laun and Wallenius (2015) studied funds in which members differ in skills and susceptibility to illness throughout their life.

Literary Gap: Coupling Effects of Heterogeneity and Impact on Estimating Liability

Each of the research works mentioned above recognizes the importance of heterogeneity and focuses on one aspect of heterogeneity at a time (retirement date, risk aversion, etc.). Overlooked is the possibility that *multiple* aspects of heterogeneity may couple together to produce an even larger impact on the management decisions. Suppose, for example, that two pension funds have the same variability in members' age and salary growth rate. But if the young members in the first fund tend to have low salary growth, while the opposite is true in the second fund, the appropriate management of these two funds can be entirely different.

The coupling effects of multiple aspects of heterogeneity may be recast in the context of "wrong-way risks". In risk management, wrong-way risks refer to the phenomenon that multiple risk factors interact to amplify the risk. In the credit risk context, for instance, variability in counterparties' exposures and variability in counterparties' credit qualities can couple in an unfavorable way that increases the risk of the portfolio (see, for example, Hull and White (2012)). Here, we argue that wrong-way risks exist in the context of pension fund management in the form of the coupling effects of several heterogeneity aspects among the fund members.

In DB, we expect that this diversity in participants profile would have big influences to the portfolio decisions of a DB pension fund since they govern the very nature and timing of cash inflow via normal contribution and cash outflow via retirement benefit payouts. Since the underfunding risk is borne by the plan sponsor, it becomes imperative for managers to make decisions that minimize the sponsor's funding support while maintaining the ability to meet the pension obligation (see, e.g., Josa-Fombellida and Rincon-Zapatero (2004) and Cox, Lin, Tian, and Yu (2013)). The issue of accurately determining the pension liability both in terms of magnitude and timing, which has been raised previously by Honda (2012) in a DC setting, will become even more important in the DB setting. Since the nature of retirement benefit in a DB fund is closely related to plan member's final salary and years of service, plan members' heterogeneity in salary and in retirement date will have a significant impact on the projected liability structure and therefore on the management decisions as well.

The Objective of this Work

Each chapter in this dissertation attempts to address the issues above, as described as follows.

Chapter 1: Heterogeneity Effects on the Management of Retirement Fund

In the first chapter of our research, we want to show the fact that if we were to consider the even simple coupling of heterogeneity, they may have a substantial economic effect on the liability estimation and fund management strategy of a DB pension fund. This demonstration will be important to both the investment managers and DB plan sponsor. The findings will likely affect how the DB fund valued their projected benefit obligations which are important towards the investment decisions and understanding the current asset and the liability level of the plan.

The mathematical model that we will use to demonstrate the impact of heterogeneity shares many of the same features from the existing literature. We mainly follow the traditional discrete-time dynamic programming framework in which the fund manager makes investment decisions over time under the objective of minimizing variation in sponsor's funding. (Examples of similar set-ups can be found in Vigna and Haberman 2001, 2002.) But to be able to study the coupling effects of multiple heterogeneity features, this paper re-specifies the model in a way that allows for members' variability in age, salary growth, salary volatility, and salary's correlation with the investment asset. Each member's retirement benefit is assumed to be a constant multiple (possibly different for each member) of his/her final salary. With the objective of minimizing the expected variation in sponsor's funding, the fund manager must then make appropriate investment decisions so that the fund is able to meet its liabilities, amidst the influence of the said heterogeneity.

We found that the optimal investment strategy and sponsorship decisions are proportionate to the discounted value of expected benefit liability from each member. The extent to which we discount the liability depends on characteristics of each individual. With the different coupling of heterogeneity, the discount rates and ultimately how the liabilities will contribute toward the optimal decisions will be different.

We show that ignorance of some heterogeneity features will result in a welfare loss. Here, we use numerical demonstration to compare the decisions taken and performance of two investment managers with different perception towards heterogeneity managing the same fund. One manager is fully aware of heterogeneity. On the other hand, another manager believes that opposing features like salary growth correlation with investment asset will simply negate. We show that the latter will fail to anticipate the need to hedge against any salary growth and risky asset co-movement which will result in severe sub-optimality for the fund and the sponsor.

Chapter 2: Adapting to Heterogeneity in Multi-Member Pension Fund Management

The closed-form solution of management decisions in Chapter 1, while useful in the analytical investigation of the heterogeneity effects, exists only under very simplified assumptions. If we impose practical constraints such as short-selling restrictions, the analytical solution is unavailable, and we generally require numerical methods to solve for the optimal management decisions. The main objective of this Chapter deals with how to arrive at the optimal decisions in a manner that appropriately accounts for heterogeneity while being parsimonious at the same time. The issue here is that, in contrast to the relative ease with which one can solve a single-member problem, the computational effort required to solve a multi-member problem grows exponentially with the number of members. This is a challenge that one must face when managing a fund with a large number of heterogeneous members.

How to overcome this exponentially-increasing computational cost, in spite of the fact that the true optimal decisions that account for heterogeneity *cannot* be obtained by simply replacing the entire cohort with one representative agent over the entire period of the problem? To solve this dilemma, this paper presents a novel idea that we can arrive at the desired optimal solution by making decisions based on a special type of representative agent whose characteristic "update" at every time step to reflect the new environment of heterogeneity. If we can find such an agent at each time, then the problem of finding the optimal decisions for a fund of heterogeneous members amounts to solving a *series* of single-member problems along the duration of the fund. Thus, the computational effort becomes linear in the number of time steps rather than exponential in the number of members.

Our proposed method aims at providing a practical tool for the pension fund industry and investment managers, by offering an effective way to reduce problem dimensionality and thus computational resources that appropriately takes into account the coupling effect of heterogeneity.

Chapter 3: Adapting to Heterogeneity in Multi-Member Pension Fund Management

In the third chapter, we will turn our attention to another type of retirement fund that is the social security system, which, in essence, share many of the same features as the DB pension funds. We will be looking into the effect of heterogeneity but this time on a country-level scale and try to understand its effects on the sustainability of a country's social security system. The main objective of this chapter is to analyze and demonstrate the effects of heterogeneity on the sustainability of the social security system.

The long-run sustainability of the social security system is uncertain in the face of upcoming demographic shifts. The question of maintaining the sustainability of the system has been at the forefront of the public policy issues. In the current literature, less emphasis is given to the less apparent shift in economic policy and heterogeneity already existed within the workforce such as the change in labor force industry entry distribution, industry growth, and workforce productivity. Each of these aspects will

affect the system's aggregate taxation credit; social security benefit withdrawal and social security investment return which has a direct effect on the system's overall sustainability.

Academics and practitioners agreed that the long-run sustainability of the pension system must come with a reform of the existing system to close the budget gap. It is still debatable whether reformation policies like an increasing the social security taxation cap (Kitao (2014)) or investment capacity (Bohn (1999), Oshio (2004)) would yield superior sustainability benefit for the fund. Our work also tries to provide a unifying theory between choices of reformation schemes; increasing the social security taxation cap or investment capacity. We also establish that the optimal reformation choice is highly dependent on the country's demographic heterogeneity.

For our model to authentically reflect the impact of heterogeneity on the sustainability of the social security system, we have to collect various social security system and related data to be applied to our model. The country our interest is Thailand. One of the reasons we choose to study the sustainability of Thailand's social security system is because of the fact that relative to other countries, Thailand's population is aging fast. In a few years, approximately 1/6 of the country's population would be over 65 years of age and the proportion is rising even faster than in China (Economist, 2018). And despite such swift demographic shift which would place an extremely heavy burden on the country's social security system, there is a lack of urgency for reformation coming from the associated policymakers.

We contribute towards the social security sustainability literature by illustrating important implications of population heterogeneity towards the sustainability of social security fund and that the existence of multiple heterogeneity aspects may produce non-monotonic effect towards the sustainability of the system. Our work also laid out a unifying framework for competing theories in determining the optimal choice of reformation strategies through heterogeneity.

Organization of this paper

This dissertation is organized into the following sections and chapters. The immediately succeeding section is the literature review. The section offers an overview of literature in the context of DC and DB pension fund management, social security sustainability and heterogeneous agent models in economics and finance. Chapter 1, "Heterogeneity Effects on the Management of Retirement Fund" introduces our new heterogeneous agent model and illustrates the effects of heterogeneity on management decisions. Chapter 2 "Adjustment for Heterogeneity in Multi-Member Pension Fund Management" proposes an alternative novel approach "Adaptive Representative Agent" which reduces the dimensionality problem of multi-agent model. Chapter 3 "Social security model with population heterogeneity" demonstrates the effects of demographic heterogeneity on the sustainability of the social security system. The last section concludes the dissertation.

Literature Review

In the ancient past, portfolio decisions made by pension funds typically start with the actuarial calculation of the required return. A portfolio is then chosen such that this required expected return is attained with the smallest risk using the mean–variance approach in Markowitz (1952). Research on long-term investment decisions in the presence of a stochastic opportunity set has not been formally introduced until Merton (1971). It is the first demonstration that the presence of risk factors that impact the fund performance justifies the introduction of intertemporal hedging demands in an investor's optimal allocation. Merton framework revolutionized the researches on pension fund management. There have been many important studies that built upon the original work of Merton (1971) such as Deelstra, Grasselli, and Koehl (2000), Boulier, Huang, and Taillard (2001) and Cairns et al. (2006).

This literature review section offers an overview of literature in the context of DC and DB pension fund management, social security sustainability and heterogeneous agent models in economics and finance. It is necessary that we also understand the background and features of management in each type of pension funds. They share several common features such as they are a long-term investment fund and all trying to provide their plan members with sensible retirement pensions.

Defined Contribution (DC) Pension Fund

In a defined contribution (DC) pension fund, contributions from the retirement plan investors are fixed throughout the lifetime of investment and the benefits are solely dependent on the performance of the fund manager.

In the context of defined contribution pension funds, one of the most if not the most common asset allocation strategy is "Deterministic Lifestyling" or the "Glide Path". At the beginning of the plan, the fund invested mostly or entirely in equity. Gradually after a pre-specified target date, investments are switched over to a safer asset e.g. bonds. Overall, it is a simple strategy to explain to plan members and to implement. It is one of the default strategy offered by pensions providers.

Despite its widespread use, there are several counter-arguments directed towards this type of strategy. First, it cannot be identified objectively that the glide path strategy is an optimal strategy. Second, it is blind to the fact that the investor may have accumulated too little wealth in the initial years. Lastly, it can be counterproductive as by moving away from stocks to low return assets just when the size of their contributions (and accumulated fund) are growing larger in later years.
New works on DC fund investments are orientated towards the development of new dynamic allocation strategy and models. They are mostly driven by the inefficiency and limitation of the deterministic lifestyling strategy. Dated literature introduced several new stochastic components into the dynamic models in addition to the investment set. Deelstra, Grasselli, and Koehl (2003) investigate optimal asset allocation for DC pension funds with stochastic interest rates and a minimum

guarantee protection. Blake, Cairns, and Dowd (2001) developed a dynamic pension plan "Stochastic life-styling" taking into account the stochastic features of the plan member's lifetime salary progression and allow a possible correlation with the market. Hainaut and Devolder (2007) consider the optimal management policy and the asset allocation of a DB pension fund under the mortality risk. Martellini and Milhau (2010), proposed a dynamic allocation strategy considering a combination of an alternative "real-estate" investment option, interest rate and inflation. In our model, we also utilized the stochastic features of the plan member's lifetime salary progression from Blake et al. (2001).

Besides the new stochastic components, Dynamic Lifestyling models would reflect a member's attitude to risk via member's utility function. Numbers of research also introduced new controlling objectives for the pension fund. Vigna and Haberman (2001) minimize the risk measured by the quadratic mean deviation in a discrete time setting for DC pension fund asset allocation. A few others apply the Mean–Variance criterion to study the investment problem in the management of DC pension funds notably Vigna (2014), Yao, Yang, and Chen (2013) also utilize the Mean-Variance criterion but also take into account inflation. Succeeding their paper in 2001, Cairns et al. (2006) argue that at the time of retirement the plan member will be concerned about the preservation of his standard of living. So he will be interested in his retirement income relative to his pre-retirement (final) salary. They introduced salary-related numeraire as an argument in the plan member's utility function. Analytical suggest that under the presence of salary and asset return correlation, the optimal solution is to hedge against it. Basu, Byrne, and Drew (2011) proposed a simple dynamic strategy altering the allocation between growth and conservative assets based on cumulative portfolio performance relative to a set wealth target and demonstrated that the glide path produces inferior wealth outcomes for the investor. By simulation, Basu et al. (2011) show that the wealth outcome from using a dynamic allocation stochastically dominated the wealth from deterministic lifestyling. In our work, we employed a different interpretation but analogous objective function to Vigna and Haberman (2001).

Defined Benefit (DB) Pension Fund

Pension plan sponsors often have conflicting goals when designing the asset allocation of the plan (Kazemi, Black, & Chambers, 2016). The first goal is to earn a high return on pension assets, which will be used to reduce the employer's long-term contributions required to fund employee benefits. The second goal is to minimize the degree of underfunding or the amount of surplus risk incurred in the plan. The literature on DB pension fund aims at finding the optimal contribution and asset allocation strategies with varying degree of conflicting objectives and constraints.

A common framework of DB fund investment allocation often follows a strategy known as Liability-driven investing. Liability-driven investing (LDI) seeks to reduce surplus volatility by building a portfolio of assets that produce returns that are highly correlated with the change in the plan's liabilities. The liabilities, in this case, are the projected benefit obligation (PBO) which is the present value of the benefits assumed

to be paid to all future retirees of the firm. Estimation of PBO poses a significant challenge as the number of workers at the firm, employee turnover levels, employee final salary and years of service are unknown. Our optimal management strategies in our work also conform to the Liability-driven investing framework. It will also shed lights to how the investment manager should value their projected benefit obligation under heterogeneity.

An early work by Haberman and Sung (1994) and Haberman (1997) assume promised benefits and discount rate of liability and thus the PBOs are constant and deterministic. They consider contribution rate risk for defined benefit pension schemes with aim of minimizing the variability in the present value of future contributions. Josa-Fombellida and Rincon-Zapatero (2004) consider the benefits flow as a stochastic process (geometric Brownian motion) while investigating how to minimize both the contribution rate risk and the solvency risk. Josa-Fombellida and Rincon-Zapatero (2006) extend the above by considering three new different objectives for the fund i.e. trying to achieve some prescribed goal before some ruin point, minimizing the expected discounted cost of reaching a ruin point and maximization of utility function when the fund can suddenly terminate. The commonality among these works is the fact that they did not implicitly consider the magnitude and the actual timing of benefit payout in their models. These are addressed in the newer researches including ours where we consider and model the characteristics of each plan member directly.

Another strand of literature tries to model and study the demographic evolution of a pension plan's population. Colombo and Haberman (2005) study the impact of the stochastic evolution of the active membership population on the mismatch between assets and liabilities of a DB plan. Delong, Gerrard, and Haberman (2008) investigate the optimal investment strategy by minimizing funding variation including the supplementary contributions in a mean-variance problem setting. A recent work by Cox et al. (2013) minimizes total sponsor funding and its variation while employing random mortality evolution. Their model and analysis are based on one representative age cohort. Our work argued that referencing the fund with one representative member or cohort may lead to incorrect projected liability estimation. The possibility that *multiple* aspects of heterogeneity may couple together to produce an even larger impact on the projected liability and thus the management decisions.

Heterogeneous Agent Model in Finance and Economics

In terms of the literature on heterogeneous agent model in Economics and finance, typically, the heterogeneous agent approach attempts to explain market dynamics by means of the time-varying nature of expectations of investors. We are witnessing some shifts, from a representative, rational agent approach towards a behavioral, agent-based approach in which markets are populated with rational heterogeneous agents using the rule of thumb strategies. Hommes (2006) reviews how asset price dynamics can be explained by heterogeneous investors who apply time-varying investment strategies. Examples of Heterogeneous Agent Models (HAMs) trying to explain asset price dynamics reviewed by Hommes (2006) are fundamentalists versus

chartists and rational versus noise traders. According to a recent survey by Dieci and He (2018), newer Heterogeneous Agent Models applications are now dealing with multi-asset markets and financial market interlinkages, house price dynamics and Market Microstructure.

There is a limited availability of heterogeneous agent model in the context of pension funds. Colombo and Haberman (2005) study the impact of stochastic active-membership in which they allow age heterogeneity. Cremer et al. (2008) investigate pension funds whose members differ in fertility and ability to raise children. Laun and Wallenius (2015) studied funds in which members differ in skills and susceptibility to illness throughout their life. Honda (2012) formulated a dynamic pension fund portfolio management problem when risk preferences among the plan members are heterogenous. They consider an optimal portfolio problem of pension funds as a syndicate problem. Honda (2012) documented that the optimal portfolio is characterized as a weighted average of each member's optimal portfolio. Another important implication from this paper is the finding that the objective functions of syndicate pension funds do not tend to be constant absolute or relative risk aversion but tend to be decreasing relative risk aversion. Their finding strongly supports the claims that deterministic glide path cannot be optimal for pensioners. Utilizing Japanese public pension fund data, they demonstrate that the loss of social welfare from using inefficient risk sharing rules and suboptimal portfolio strategies is significant.

Each of the research works mentioned above focuses on one aspect of heterogeneity at a time (retirement date, risk aversion, etc.). Overlooked is the possibility that *multiple* aspects of heterogeneity may couple together to produce an even larger impact on the management decisions. Suppose, for example, that two pension funds have the same variability in members' age and salary growth rate. But if the young members in the first fund tend to have low salary growth, while the opposite is true in the second fund, the appropriate management of these two funds can be entirely different. Our research contributes towards the heterogeneous agent model literature in pension funds as we are one of the first to look at the coupling effects of multiple aspects of heterogeneity on the management decisions.

Social Security System and its Sustainability

Traditionally, economics literature attempted to compute the long-run stationary equilibrium to evaluate the effects of the demographic shifts and its policy reformation countermeasure. A vast literature has attacked this question using general equilibrium overlapping generations (OLG) models where individuals decide on consumption, savings and labor force participation. Static equilibrium in most models implies a balance between government's yearly spending, taxation incomes, benefit payout, clearance labor market and clearance capital market.

As for the literature on the sustainability of social security, there is a vast literature on equilibrium OLG models that evaluate quantitatively different scenarios for social security reform in a closed economy. The general equilibrium OLG life-cycle model

is pioneered by Auerbach and Kotlikoff (1987). Sustainability of social security is often analyzed through dynamic general-equilibrium models in a closed-economy framework. Using this OLG framework, De Nardi, Imrohoroglu, and Sargent (1999) and Kotlikoff, Smetters, and Walliser (2002) predict that 15% increase to the payroll tax in the next 100 years will keep the U.S. system solvent. Conesa and Krueger (1999) study the effect of social security reform with heterogeneous agents differing in the productivity.

Kitao (2014) utilizes this OLG framework and suggest that there are four options to make the social security sustainable under the coming demographic shift. Leveraging survival probabilities from Bell and Miller (2005) and Kitao (2014) drew a comparison between a benchmark economy with 2010 survival rates and economy with aging with 2100 survival rates. From the computation of stationary equilibrium, Kitao (2014) suggest that an increase payroll taxes by 6 percentage points, reduce replacement rates by one-third, raise the normal retirement age to 73, or means-test the benefits and reduce them in income would lead to the sustainability for the U.S. system. While achieving the sustainability goal, Kitao (2014) also showed that options have very different effects on labor force participation and retirement decisions of individuals.

Some authors argue that the closed-economy benchmark may not be the right one to study the implications of reforming the PAYG system and proposed an open economy framework. An open economy framework basically views the system as having multiple regions where capital flow between regions is allowed. Borsch-Supan (2003) proposed multi-region models of the developed world (i.e., a subset of OECD countries), where the focus is on the effects of pension reform in U.S. and Germany, respectively, in open economy. Attanasio, Kitao, and Violante (2007) proposed a two-region model (developed and developing) of the world economy where capital flows across regions equalize the rate of return on capital.

Moving on from the sustainability literature to the management of social security trust fund, literature suggested that social security investment can bring positive welfare in several ways. The US Social Security Trust Fund now amounts to over one trillion dollars, but the latest 2003 Social Security Trustee's Report released the official projection that the OASDI trust fund will be exhausted sometime around 2042 after the retirement of baby boom generation around 2010. Aaron and Reishauer (1998) insist on shifting the existing trust fund from (non-marketable) government bonds to private securities to raise the rate of return of the trust fund thus extending the period before its balance is exhausted. Bohn (1999) also find that stock returns are positively correlated with social security's wage-indexed benefit obligations, equity investments would also help to stabilize the payroll tax.

Oshio (2004) suggested that a trust fund in some form can offset the negative income effect of a PAYGO system. Social security trust fund is considered to be a "buffer," which is expected to absorb the adverse impact of a declining population and demographic shocks. He also discussed the policy implication of the use of the trust fund in the face of a rapidly aging population in Japan.

It is expected that the relationship between labor force entry rate, unemployment, and capital market performance is of a direct proportional. However, Cheng and French (2000), the Federal Reserve Bank of Chicago predicted that a run-up in the stock market should cause reductions in labor force participation rates, all else equal. Boyd, Hu, and Jagannathan (2005) documented that on average, an announcement of rising unemployment is good news for stocks during economic expansions and bad news during economic contractions.

1 Heterogeneity Effects on the Management of Retirement Fund

In a typical pension scheme, management decisions are usually made collectively as if the plan members are homogeneous, all of whom share a single set of parameters that are deemed representative of the entire cohort (Bovenberg, Koijen, Nijman, & Teulings, 2007). While this simplification tends to help the fund managers streamline the decision-making process, questions still remain whether the decisions made are truly optimal, since in reality, a fund would comprise members with different characteristics.

In reality, we observed plan members with differences in retirement date, salary growth, salary volatility, and salary's correlation with the capital market. These aspects of "heterogeneity" can have a substantial economic effect on the investment decisions of fund managers.

Indeed, plan participant's heterogeneity characteristics govern the nature of normal contribution (member's periodic contribution) cash flow to the fund and affect the liable amount to be paid at each of the retirement (Josa-Fombellida & Rincon-Zapatero, 2004). Many researchers recognize the importance of plan members' heterogeneity and have studied its effects on the decisions of pension fund managers. Colombo and Haberman (2005) study the impact of stochastic active-membership in which they allow age heterogeneity. Cremer et al. (2008) investigate pension funds whose members differ in fertility and ability to raise children. Honda (2012) studied pension funds whose members have different levels of risk aversion. Laun and Wallenius (2015) studied funds in which members differ in skills and susceptibility to illness throughout their life. Ayres and Nalebuff (2013) have documented variability in the salary-asset return correlation across industries, which has been shown to have a fundamental influence on investment decisions (see Cairns et al. (2006) and Benzoni et al. (2007)).

Each of the research works mentioned above focuses on one aspect of heterogeneity at a time (retirement date, risk aversion, etc.). Overlooked is the possibility that *multiple* aspects of heterogeneity may couple together to produce an even larger impact on the management decisions. Suppose, for example, that two pension funds have the same variability in members' age and salary growth rate. But if the young members in the first fund tend to have low salary growth, while the opposite is true in the second fund, the appropriate management of these two funds can be entirely different. The objective of this chapter is to demonstrate, both through analytical models and through numerical examples, the coupling effects of multiple aspects of heterogeneity on the management of pension funds.

The coupling effects of multiple aspects of heterogeneity may be recast in the context of "wrong-way risks". In risk management, wrong-way risks refer to the phenomenon that multiple risk factors interact to amplify the risk. In the credit risk context, for instance, variability in counterparties' exposures and variability in counterparties' credit qualities can couple in an unfavorable way that increases the risk of the

portfolio (see, for example, Hull and White (2012)). Here, we argue that wrong-way risks exist in the context of pension fund management in the form of the coupling effects of several heterogeneity aspects among the fund members.

In this work, the pension funds that we focus on are of the defined-benefit (DB) type. We believe that the effects of heterogeneity on management decisions are particularly interesting in the defined-benefit pension funds as opposed to the defined-contribution (DC) pension funds. In DC pension funds, plan members bear the investment risk while managers are not obligated to meet any target. In a DB fund, on the other hand, each plan participant is promised a pre-defined retirement benefit and the underfunding risk is borne by the plan sponsor. Thus, it becomes imperative for managers to make decisions that minimize the sponsor's funding support while maintaining the ability to meet the pension obligation (see, e.g., Josa-Fombellida and Rincon-Zapatero (2004) and Cox et al. (2013)). The issue of accurately determining the pension liability both in terms of magnitude and timing, which has been raised previously by Honda (2012) in a DC setting, will become even more important in the DB setting. Since the nature of retirement benefit in a DB fund is closely related to plan member's final salary and years of service, plan members' heterogeneity in salary and in retirement date will have a significant impact on the projected liability structure and therefore on the management decisions as well.

The mathematical model that we will use to demonstrate the impact of heterogeneity shares many of the same features from the existing literature. We mainly follow the traditional discrete-time dynamic programming framework in which the fund manager makes investment decisions over time with the objective of minimizing variation in sponsor's funding. (Examples of similar set-ups can be found in Vigna and Haberman 2001, 2002.) But to be able to study the coupling effects of multiple heterogeneity features, this paper re-specifies the model in a way that allows for members' variability in age, salary growth, salary volatility, and salary's correlation with the investment asset. Each member's retirement benefit is assumed to be a constant multiple (possibly different for each member) of his/her final salary. With the objective of minimizing the expected variation in sponsor's funding, the fund manager must then make appropriate investment decisions so that the fund is able to meet its liabilities, amidst the influence of the said heterogeneity.

In this chapter, the analysis of how several heterogeneity features interact and impact management decisions is presented in several steps. In the first part of the analysis, we derive a closed form for the optimal investment decisions under a simplified instance of our model and use it to analytically demonstrate the coupling effects of heterogeneity features on fund management. At the surface level, the outcome of our analysis in this part conforms to the existing asset-liability management concept, in which both the optimal risky investment allocation and supplementary funding decision are increasing in the fund's total projected pension obligation and decreasing in the current level of fund's wealth (see, for example, Cox et al. (2013)). However, to account for the interaction between heterogeneity features, our analysis introduces extra adjustments in the valuation of pension obligations. (The size of the adjustments turns out to be a function of several characteristics. For instance, we find that the

projected pension obligation should be adjusted upward for a member whose salary growth is negatively correlated with the investment asset return, and even more so if said member is still young.) We conclude the first part of our analysis by demonstrating that, if the various aspects of heterogeneity interact in different ways, the management decision will also differ significantly.

Next, for the more general case for which no closed forms are available, we use numerical examples to demonstrate that plan members' heterogeneity has significant economic impacts on the management strategy and that it leads to different levels of management difficulty. We showed that the investment managers for two funds that on the surface may seem conceivably indifferent may end up having to take different optimal strategies to achieve minimal variation in sponsor's funding. If your fund only has members whose salary grow quickly but with negative growth correlation with investment asset but does not have a younger member with positive correlation to help provide a natural hedge, it can be very difficult to manage without any sponsor's supplementary funding. In general, we find that the presence of a young member with positive correlation, low the average age and low liable retirement payables and wide retirement window will facilitate the management.

In the final part of our analysis, we show that ignorance of some heterogeneity features will result in a welfare loss. Here, we use numerical demonstration to compare the decisions taken and performance of two investment managers with different perception towards heterogeneity managing the same fund. One manager is fully aware of heterogeneity. On the other hand, another manager believes that opposing features like salary growth correlation with investment asset will simply negate. We show that the latter will fail to anticipate the need to hedge against any salary growth and risky asset co-movement which will result in severe suboptimality for the fund and the sponsor.

This chapter is organized into the following sections. Section 1.1, DB fund model with heterogeneity, we discuss the overview and the components of our model. In section 1.2, we present the analytical solution to our simplified model and investigate the effects of heterogeneity on the decisions. Section 1.3 demonstrates how the ignoring of heterogeneity can lead to sub-optimality and presents implications to members grouping. Section 1.4 concludes the chapter.

1.1 DB Model with Heterogeneous Members

In this section, we present the detail of our DB fund model with heterogeneous members in which multiple fund members contribute towards a collectively managed fund. Every plan member is guaranteed a lump sum which is a multiple of their final salary upon retirement by the plan sponsor. The role of the fund manager in our model is to decide on investment allocation (choices between a risk-free and one risky asset) and any request for supplementary funding from the sponsors if necessary. The objective of the fund manager in our model is to be able to settle all the promised benefit liabilities while requesting for minimal sponsorship from the plan sponsor. The model uses a discrete time setting. We describe each component in detail in the succeeding subsections.

1.1.1 Member's Salary and Retirement Benefit

Let N denote the number of members in the DB pension fund. At each time $t = 0, 1, 2, ...$, each member i earns a salary of $Y_{i,t}$, which varies with time according to the following stochastic process:

$$Y_{i,t+1} = Y_{i,t}\left(1 + \mu_i + \sigma_i Z_{i,t} + \sigma_{r,i} Z_{r,t}\right). \tag{1}$$

Here, μ_i denote the expected growth rate of member i's salary. Uncertainty in salary growth is introduced through the terms $\sigma_i Z_{i,t}$ and $\sigma_{r,i} Z_{r,t}$, where σ_i, $\sigma_{r,i}$ are constants and $Z_{i,t}$ $(i = 1, ..., N)$ and $Z_{r,t}$ are iid $\mathcal{N}(0,1)$'s for all t. Note that $Z_{i,t}$ is specific to member i, whereas $Z_{r,t}$ affects all members. This salary process setting is similar to Cairns et al. (2006).

Each member i retires at time τ_i, where τ_i is a constant integer and leaves the fund afterwards. Assume without loss of generality that $\tau_1 \leq \tau_2 \leq \cdots \leq \tau_N$, so that the model concludes at time τ_N when the last member retires. We assume for simplicity that there are no additional members introduced into the fund (see Delong et al. (2008) and Cox et al. (2013)). In every period leading up to his/her retirement, member i contributes a fixed fraction π_i of their salary to the fund. Let S_t be the cumulative contributions to the fund up to time t. Thus,

$$S_{t+1} = S_t + \sum_{i=1}^{N} 1_{\{\tau_i > t\}} \pi_i Y_{i,t}, \tag{2}$$

where $1_{\{\cdot\}}$ denote the indicator function.

We assume that each retirement pay-out to a retiring member is a lump sum which is a fixed multiplier K_i of the corresponding retiree final salary Y_{i,τ_i}. The benefit pay-outs $K_i Y_{i,\tau_i}$ from the fund will be made at each retirement until the last remaining member retires.

1.1.2 Investment and Fund's Wealth

The fund manager makes investment decisions by investing in a risk-free asset and a risky asset. Let r_f be a constant risk-free rate of return. The rate of return of the risky asset in period t is given by

$$r = r_f + \xi\sigma_r + \sigma_r Z_{r,t}. \tag{3}$$

The constant $\sigma_r > 0$ denotes the volatility of the return, and $\xi > 0$ denotes the risk premium per unit volatility. The presence of $Z_{r,t}$ in both (1) and (3) implies correlations between asset's return and members' salary growth. A member with positive (negative) $\sigma_{r,i}$ has salary growth that is positively (negatively) correlated with the investment asset's return.

Let W_t denote the total wealth of the pension fund at the beginning of time t, and let p_t denote the portion invested in the risky asset in period t. (For now, we allow short-selling both risk-free and risky assets, so p_t can be less than 0 or more than 1.) Therefore, the total wealth process satisfies:

$$W_{t+1} = W_t(r_f + p_t\xi\sigma_r + p_t\sigma_r Z_r) + S_{t+1} - S_t + Q_{t+1} - Q_t, \tag{4}$$

Where $S_{t+1} - S_t$ is the contribution from remaining members in period t, and $Q_{t+1} - Q_t$ is the amount of supplementary funding less withdrawals in period t, which we will formally define in the next section.

1.1.3 Supplementary Funding and Withdrawals

When member i retires at time $t = \tau_i$, the amount of $K_i Y_{i,\tau_i}$ is withdrawn from the fund to pay out his retirement benefit. The fund manager must ensure that the fund has sufficient capital to cover any immediate and future retirement pay-outs. We assume that, at each retirement date, the manager can request supplementary funding from the plan's sponsor if deemed necessary. Let X_i ($i = 1, \dots, N$) denote the supplementary funding at time τ_i. Let X_0 denote the initial funding provided at the beginning. Thus, the cumulative funding-less-withdrawal process can be written as:

$$Q_{t+1} = Q_t + \sum_{i=1}^{N} 1_{\{t=\tau_i\}}(X_i - K_i Y_{i,t}). \tag{5}$$

Here, our specification allows X_i's to be negative. In other words, we allow the sponsor to withdraw an amount from the fund if it is deemed to minimize overfunding in the future (see the manager's objective function in the next section). This assumption is for simplicity purpose and makes the model more tractable. It can also be argued that overfunding increase opportunity cost from other forgone profitable investment opportunities Vigna and Haberman (2001). If there are more than one member retires at the same time, we assume for simplicity that only one of them can have non-zero X_i. We impose the following constraint on X_N:

$$X_N = K_N Y_{N,\tau_N} - W_{\tau_N}. \tag{6}$$

In other words, at the time the last member retires, we either supply the extra funding necessary to satisfy the last retirement benefit or withdraw any over-funded surplus.

1.1.4 Objective Function and Optimization Problem

Define the expected *funding variation* as:

$$U = \mathrm{E}\left[\sum_{i=0}^{N} X_i{}^2\right] \tag{7}$$

We give the fund manager the objective of minimizing the expected funding variation by choosing the appropriate stochastic investment strategy $\{p_t\}_{t=1,\dots,\tau_N-1}$ and supplementary funding $\{X_i\}_{i=1,\dots,N-1}$. At each time t, the decision variables are allowed to depend on the state variables W_t and $Y_{i,t}$ $(i = 1,\dots,N)$. Thus, we have the following stochastic control problem:

$$\min_{p,X_0\dots X_{N-1}} \mathrm{E}\left[\sum_{i=0}^{N} X_i{}^2\right] \tag{8}$$
$$\text{Subject to } X_N = K_N Y_{N,\tau_N} - W_{\tau_N}$$

Note that our quadratic objective function penalizes both tails of the funding variation --- not only must the fund manager minimize the amount of sponsor's supplementary funding, he must also minimize instances of over-funding. The rationale is that an excessive surplus held in the pension fund is undesirable because it represents the sponsor's opportunity cost from forgone investment opportunities. The two-tailed quadratic from also makes the problem more tractable and allows us to solve it analytically in a later section. Examples of authors who also use this objective function include Vigna and Haberman (2001) and Haberman and Sung (2005). Note that some authors include discount factors in the funding variation, but we omit them here for the sake of simplicity.

1.1.5 Optimality Conditions

Now that we have formulated the manager's decision process as a stochastic control problem, we can follow the familiar dynamic programming framework and derive the optimality conditions for the decision variables $\{p_t\}$ and $\{X_i\}$.

Let $\mathbf{Y}_t = \left(Y_{1,t}, Y_{2,t}, \dots, Y_{N,t}\right)$ be the vector of members' salary at time t. Define the *value function* at time t, which is a function of the state variables $\mathbf{Y}_t$ and W_t, is defined as:

$$V_t(\mathbf{Y}_t, W_t) \quad = \quad \min_{\mathbf{p}_t, \mathbf{X}_t} E_t\left[\sum_{i:\,\tau_i\geq t} X_i^2\right], \tag{9}$$

where the minimum is taken over $\mathbf{p}_t = \{p_s\}_{s=t,t+1,\dots,\tau_N-1}$ and $\mathbf{X}_t = \{X_i\}_{i:\tau_i\geq t}$. From (6), it follows that

$$V_{\tau_N}\left(\mathbf{Y}_{\tau_N}, W_{\tau_N}\right) \quad = \quad \left(K_{\tau_N} Y_{N,\tau_N} - W_{\tau_N}\right)^2. \tag{10}$$

The value function $V_t(\mathbf{Y}_t, W_t)$ satisfies the following backward recursion: for all $t = \tau_N - 1,\ \tau_N - 2,\ \dots, 1, 0,$

$$V_t(\mathbf{Y}_t, W_t) \quad = \quad \min_{\mathbf{p},\mathbf{X}} E_t[V_{t+1}(\mathbf{Y}_{t+1}, W_{t+1})] \ + \sum_{i:\,\tau_i=t} X_i^2. \tag{11}$$

where Y_{t+1} and W_{t+1} are as given in (1) and (4). This equation provides the condition under which p_t and $\sum_{i:\,\tau_i=t} X_i$ are optimal at each time $t = \tau_N - 1, \tau_N - 2, \dots, 1, 0$.

The optimal solution, which depends on the state variables $\mathbf{Y}_t$ and W_t, shall be denoted by $\mathbf{p}^* = \left(p_0^*, p_1^*, \dots, p_{\tau_N}^*\right)$ and $\mathbf{X}^* = (X_0^*, X_1^*, \dots, X_{N-1}^*)$. In the following section, we will solve for $\mathbf{p}^*$ and $\mathbf{X}^*$ and analyze how they are affected by members' heterogeneity.

1.2 Heterogeneity Effects, Illustrated by Analytical Formulae

To begin our analysis of how heterogeneity affects management decisions, we first consider a simplified, unconstrained version of our model, for which we will derive closed-form expressions for the optimal investment decisions. The closed-form expressions will then allow us to analytically investigate the heterogeneity effects.

1.2.1 Closed-Form Solution to a Simplified Unconstrained Fund

Assuming that short-selling is allowed, the stochastic control problem in Section 1.1.5 can be solved in closed form, as summarized in the following proposition. (The proof is given in Appendix I.)

Proposition 1 *Assuming that the manager can short-sell the risk-free asset and the risky-asset, the solution p^* and X^* to the stochastic control problem in Section 1.5 can be characterized as follows. For $t = 0,1, \dots, \tau_N - 1$, let $\alpha_t = 1$ if no member retires at time t and $\alpha_t = 1 + x_t$ otherwise, where x_t is given by a backward recursion $x_t = x_{t+1}(1 + r)^2/(\alpha_{t+1} + \xi^2)$ with $x_{\tau_N - 1} = 1$. Then,*

$$p_t^* = \frac{\xi/\sigma_r}{\alpha_t + \xi^2}\left[\sum_{i=1}^{N} A_{i,t}\frac{K_i(1 + \mu_i)^{\tau_i - t}Y_{i,t}}{W_t} - 1 - r\right], \tag{12}$$

$$\sum_{i:\,\tau_i = t} X_i^* = \frac{\alpha_t - 1}{\alpha_t + \xi^2}\left[\sum_{i=1}^{N} B_{i,t}\,K_i(1 + \mu_i)^{\tau_i - t}Y_{i,t} - (1 + r)W_t\right], \tag{13}$$

Here, $A_{i,t}$ $B_{i,t}$ $(i = 1, \dots, N)$ are deterministic coefficients, independent of the state variables, satisfying the following backward recursion: $B_{i,\tau_N} = (1 + r)1_{\{\tau_i = \tau_N\}}$ and, for $t = \tau_N - 1, \tau_N - 2, \dots, 1, 0$,

$$A_{i,t} = B_{i,t+1}\frac{1 + \alpha_t\sigma_{r,i}\xi^{-1}/(1 + \mu_i)}{1 + r} - \frac{\pi'_{i,t}/K_i}{(1 + \mu_i)^{\tau_i - t}}, \tag{14}$$

$$B_{i,t} = B_{i,t+1}\frac{1 - \sigma_{r,i}\;\xi\,/\,(1 + \mu_i)}{1 + r} - \frac{\pi'_{i,t}/K_i}{(1 + \mu_i)^{\tau_i - t}}, \tag{15}$$

where $\pi'_{i,t} := \pi_i 1_{\{\tau_i > t\}} - K_i 1_{\{\tau_i = t\}}$. □

Proposition 1 gives explicit expressions for the investment strategies p_t^* and the funding decisions X_i^* for a DB fund with N heterogeneous members. Presently, we shall use the closed-form expressions to analyze how p_t^* and X_i^* are affected by members' heterogeneity.

1.2.2 Heterogeneity Effects on the Optimal Supplementary Contribution

In this section, we analyze how funding decisions X_i's are affected by heterogeneity in retirement date, salary growths and its correlations with the risky asset. To facilitate discussion, equation (13) is re-displayed here:

$$\sum_{i:\,\tau_i=t} X_i^* \;=\; \frac{\alpha_t - 1}{\alpha_t + \xi^2}\left[\sum_{i=1}^{N} B_{i,t}\,K_i(1+\mu_i)^{\tau_i-t}Y_{i,t} \;-\; (1+r)W_t\right], \tag{13}$$

Recall our assumption that the amounts $X_0, X_1, \ldots, X_N$ are added to or withdrawn from the fund only at times $t = 0, \tau_1, \tau_2, \ldots, \tau_N$. (At other times t, note that $\alpha_t = 1$ and therefore both sides of (13) are zero.) Since the expected final salary of member i is given by $E[Y_{i,\tau_i}|Y_{i,t}] = (1+\mu_i)^{\tau_i-t}Y_{i,t}$ (see (1)), the term $K_i(1+\mu_i)^{\tau_i-t}Y_{i,t}$ in the summation (13) represents his expected retirement benefit. Thus, (13) suggests that the supplementary funding at time t is proportional to the expected total retirement liability, where each member's expected future retirement benefit is "discounted" to the present time by a coefficient $B_{i,t}$ (more on this coefficient in the paragraph below). If this summation is greater than $(1+r)W_t$, then (13) suggests that a positive amount of extra funding should be supplemented to the fund. But if the summation is less than $(1+r)W_t$, then (13) suggests that the sponsor can withdraw some excess capital from the fund.

As seen from (15), the value of the discount factor $B_{i,t}$ is determined by member-specific parameters $\mu_i, \sigma_{r,i}, \pi_i, K_i, \tau_i$. Variability in the values of $B_{i,t}$ across heterogeneous members implies a different discount rate for each member's expected retirement benefits, which in turn determine the optimal funding decision under a said environment of heterogeneity. Thus, the coefficients $B_{i,t}$ can be viewed as a "heterogeneity-adjusted" discount factors for computing the projected fund liability. To get a broad idea of how the parameters $\mu_i, \sigma_{r,i}, \tau_i$ of heterogeneous members affect the values of $B_{i,t}$'s, consider (15) and assume that $\pi_i \ll K_i$, so that the last term in (15) vanishes for $t < \tau_i$ and we obtain the following approximation:

$$B_{i,t} \;\propto\; \left(\frac{1 - \xi\sigma_{r,i}/(1+\mu_i)}{1+r}\right)^{\tau_i-t} \tag{16}$$

First, note the factor $1+r$ in the denominator that signifies present-value calculation. However, not all members are given the same discount rate, as the numerator presents a member-specific adjustment based on the parameters μ_i and $\sigma_{r,i}$. The coupling effect of these two parameters on the values of $B_{i,t}$'s can be readily seen: Suppose two members, i and j, differ in that $0 < \sigma_{r,i} < \sigma_{r,j}$, then, holding everything else constant, we have $B_{i,t} > B_{j,t}$. But if we couple this with a further supposition that $\mu_i > \mu_j$, then $B_{i,t}$ will be even more dominant and $B_{j,t}$ even more subjugated. In fact, the salary growth μ_i not only affects the value of $B_{i,t}$, but also clearly raises the expected final retirement benefit $(1+\mu_i)^{\tau_i-t}Y_{i,t}$. In total, the effect of μ_i on each term in the summation (13) is approximately:

$$B_{i,t}(1+\mu_i)^{\tau_i-t} \;\propto\; \left(\frac{1 + \mu_i - \xi\sigma_{r,i}}{1+r}\right)^{\tau_i-t} \tag{17}$$

Note that the disparity between $B_{i,t}$ and $B_{j,t}$ will be even more accentuated when τ_i and τ_j are large. Generally, the summation in (13) will be dominated by members who are young, and has high salary growth with low or negative correlation with the investment asset. Under such condition of heterogeneity, equation (13) prompts manager to act very differently compared to when such condition is absent.

To explain the intuition behind the coupling effect of heterogeneity, let us give a concrete example of two hypothetical pension funds, each having only two members, as shown below.

					Fund Member's Parameters			
	τ_1	τ_2	$Y_{1,0}$	$Y_{2,0}$	μ_1	μ_2	$\sigma_{r,1}$	$\sigma_{r,2}$
Fund 1	25	35	1.5	1	$4.5\% + S$	$4.5\% - S$	$+\varsigma$	$-\varsigma$
Fund 2			1.5	1	$4.5\% + S$	$4.5\% - S$	$-\varsigma$	$+\varsigma$

Table 1, Plan members' parameters variation: Fund 1 and 2

Here, $\varsigma > 0$ and s are constants. Other parameters that are not shown in the table (such as π_i, K_i, etc.) are assumed to be the same for both members and for both funds. (For details of other parameters, please refer to Appendix II: Parameters for Analysis and Simulation.) Therefore, Fund 1 and Fund 2 seem identical in most respect. In both funds, we see the same variability in age, salary growth, and salary correlation with investment asset. However, these heterogeneous features couple in different manners in the two funds. In Fund 1, the older member has a positive salary growth correlation with the risky asset and the younger member has a negative salary growth correlation with the risky asset, whereas the converse is true in Fund 2.

For example, the following table shows $B_{i,0}$ for different values of ς, when fix $s = 1.5\%$ hence $\mu_1 = 6\%$ and $\mu_2 = 4.5\%$. This coupling of heterogeneity will induce different funding decisions for the two funds as shown in Table 2 below.

	Fund 1					Fund 2				
ς	$B_{1,0}$	$B_{2,0}$	$E_0[K_1 Y_{1,\tau_1}]$	$E_0[K_2 Y_{2,\tau_2}]$	X_0	$B_{1,0}$	$B_{2,0}$	$E_0[K_1 Y_{1,\tau_1}]$	$E_0[K_2 Y_{2,\tau_2}]$	X_0
0.012	0.28	0.29	45.06	19.70	3.89	0.44	0.12	45.06	19.70	4.38
0.018	0.25	0.35	45.06	19.70	3.80	0.50	0.09	45.06	19.70	4.59
0.024	0.22	0.43	45.06	19.70	3.74	0.55	0.07	45.06	19.70	4.84
0.030	0.19	0.51	45.06	19.70	3.71	0.62	0.05	45.06	19.70	5.11

Table 2, Comparison of discount factors $B_{1,0}$, $B_{2,0}$ and funding X_0: Varying ς

As seen from the Table 2, even though the expected benefit liability $\left(E_0[K_i Y_{i,\tau_i}] = K_i(1 + \mu_i)^{\tau_i} Y_i\right)$ of the two funds are equal, the coupling effect between variability in time-to-retirement and the correlation prompts the coefficients $B_{i,t}$ in the two funds to be very different. Coefficients $B_{i,t}$ tend to be higher for a member whose salary has a negative correlation with the investment asset. But because such member is younger in Fund 1 but older in Fund 2, the emphasis towards the member with negative

correlation is more pronounced in Fund 1. The higher the heterogeneity in ς, the more pronounced the difference.

This difference will be even more pronounced if we change the way in which heterogeneity in μ_i couple with other parameters. Consider the case where $s = -1.5$ versus the case where $s = 1.5$, as shown in the table below.

	Fund 1					Fund 2				
S	$B_{1,0}$	$B_{2,0}$	$E_0[K_1 Y_{1,\tau_1}]$	$E_0[K_2 Y_{2,\tau_2}]$	X_0	$B_{1,0}$	$B_{2,0}$	$E_0[K_1 Y_{1,\tau_1}]$	$E_0[K_2 Y_{2,\tau_2}]$	X_0
-0.015	0.20	0.43	21.99	53.80	2.09	0.48	0.14	21.99	53.80	0.10
0.015	0.24	0.38	45.07	19.70	0.10	0.51	0.09	45.07	19.70	1.51

Table 3, Comparison of discount factors $B_{1,0}$, $B_{2,0}$ and funding X_0: Varying S

In both of these two cases, we see the same variability in growth rates in both funds. But B_1 and B_2, which are already quite different in Fund 1 vs Fund 2 when $s = -1.5$, becomes even wider apart when $s = 1.5$ (because $s = 1.5$ implies that the young member in Fund 2 has both higher growth rate *and* negative correlation). This is to conclude that, even though the expected future liabilities are the same in both funds, the management decision turns out to be very different as a result of different present-value calculation of liability that takes into account the coupling effect of heterogeneity.

1.2.3 Heterogeneity Effects on Optimal Investment Decisions

Now, let us analyze how heterogeneity affects the optimal investment decision p_t^*. Recall (12) the analytical form of optimal investment allocation

$$p_t^* = \frac{\xi/\sigma_r}{\alpha_t + \xi^2} \left[\sum_{i=1}^{N} A_{i,t} \frac{K_i(1 + \mu_i)^{\tau_i - t} Y_{i,t}}{W_t} - 1 - r \right], \tag{12}$$

Analogously to the previous subsection, $K_i(1 + \mu_i)^{\tau_i - t} Y_{i,t}$ here represents the expected future retirement benefit of member i, which is discounted to the current time t with the coefficient $A_{i,t}$ (similar to how the expected future retirement benefit in (13) is discounted by $B_{i,t}$). Equation (12) implies that p_t^* is proportional to the ratio of the present-value of the total expected retirement benefit to the current level of fund value W_t. Thus, our optimal investment allocation in the risky asset is increasing in the total expected retirement benefit and decreasing in the fund wealth. This is consistent with the concept of "liability-driven" investment (see, for example, Farrell and Shoag (2016)).

Heterogeneity affects how the terms in the summation (12) are valued relative to one another, and this heterogeneity-dependent composition will, in turn, determine the optimal investment decision p_t^*. While the realized levels of the salary $Y_{i,t}$'s in the summation are implicitly affected by heterogeneity, let us first focus on how heterogeneity affects the discount factors $A_{i,t}$. Because of both $A_{i,t}$ and $B_{i,t}$ are computed from similar recursions (see (14)), it follows that $A_{i,t}$ depends on the member-specific parameters $\mu_i, \sigma_{r,i}, \tau_i$ in a similar way as $B_{i,t}$. *That is to say, the $A_{i,t}$*

terms in the summation (12) tend to be large for young members who have high salary growth, with low or negative correlation with the investment asset. In the presence of members who satisfy such condition of heterogeneity, equation (12) will prompt the manager to make a riskier investment than he would otherwise. Consider, for example, Fund 1 and Fund 2 from the previous section. Since we did not impose short sale constraint, for demonstrations in this section, we assume that $W_0 = 17$ for both funds to keep the range of p_0^* between zero and one. The values of $A_{i,t}$ and p_t^* at time zero are shown below.

	Fund 1					Fund 2				
ς	$A_{1,0}$	$A_{2,0}$	$E_0[K_1 Y_{1,\tau_1}]$	$E_0[K_2 Y_{2,\tau_2}]$	p_0	$A_{1,0}$	$A_{2,0}$	$E_0[K_1 Y_{1,\tau_1}]$	$E_0[K_2 Y_{2,\tau_2}]$	p_0
0.012	0.29	0.27	45.06	19.70	0.21	0.39	0.17	45.06	19.70	0.77
0.018	0.27	0.30	45.06	19.70	0.13	0.41	0.15	45.06	19.70	0.97
0.024	0.25	0.34	45.06	19.70	0.08	0.44	0.13	45.06	19.70	1.19
0.030	0.23	0.37	45.06	19.70	0.05	0.47	0.12	45.06	19.70	1.44

Table 4, Comparison of discount factors $A_{1,0}$, $A_{2,0}$ and asset allocation p_0: Varying ς

Here, Fund 1 has different $A_{i,0}$ than Fund 2 because the two funds exhibit different interactions between heterogeneity features: In Fund 1, the older member has positive $\sigma_{r,1}$ while the young member has negative $\sigma_{r,2}$. In Fund 2, the converse is true. As a result, even though Fund 1 and Fund 2 are the same in terms of the expected future retirement benefit, the investment strategy in the two funds can be different due to the difference in the coupling of heterogeneity features. Note that, in both funds, the coefficient $A_{i,0}$ is larger for the member whose salary growth is negatively correlated with investment return, which means that the retirement benefit of that member will greatly induce the manager to invest in the risky asset to accelerate returns early on, since it will be difficult later to rely on the investment return to hedge against an increase in that member's salary (and therefore increased liability).

It is worth noting that, unlike the simple statement (16) for $B_{i,t}$, the precise description of how $A_{i,t}$ depends on the parameters $\mu_i, \sigma_{r,i}, \tau_i$ is more complicated. From (14), if we assume that $\pi_i \ll K_i$ and ignore the last term for $t < \tau_i$, then $A_{i,t}$ is approximated as:

$$A_{i,t} \propto \frac{1 + \alpha_t \sigma_{r,i} \xi^{-1}/(1 + \mu_i)}{1 + r} \left(\frac{1 - \xi \sigma_{r,i}/(1 + \mu_i)}{1 + r} \right)^{\tau_i - t - 1} \tag{18}$$

Unlike (16), one can see that the right-hand side of (18) is not monotonic in μ_i and $\sigma_{r,i}$. Therefore, our discussion in the previous paragraph --- about the tendency for the member whose $\sigma_{i,r}$ is negative to have a larger value of $A_{i,t}$ --- might not hold for all t. Figure 1 below demonstrates how discount factors $A_{1,t}, A_{2,t}$ vary with time in Fund 1 and Fund 2.
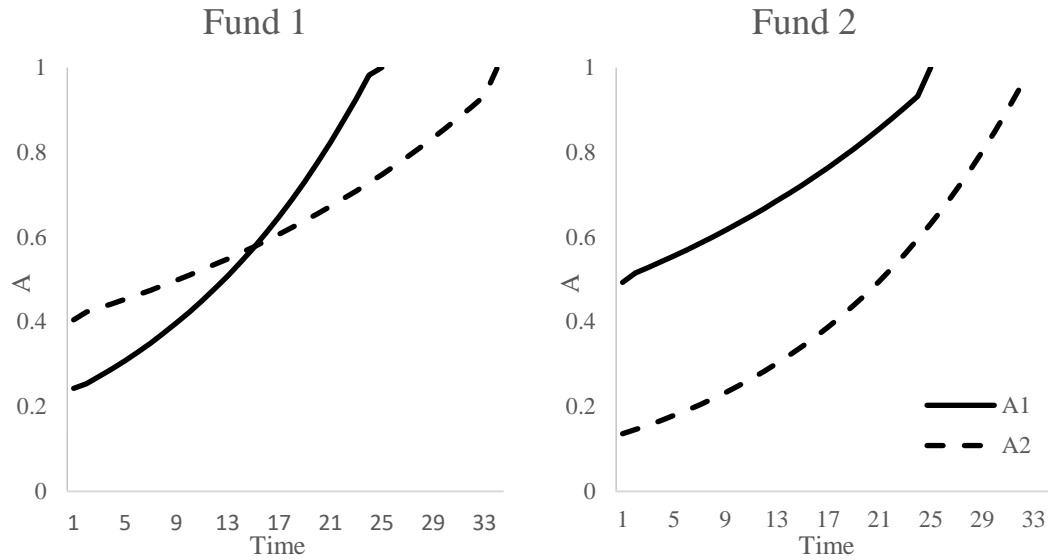
Figure 1, Comparison of discount factors $A_{1,t}$ and $A_{2,t}$ over time

Fund 1					Fund 2				
$A_{1,0}$	$A_{2,0}$	$E_0[K_1Y_{1,\tau_1}]$	$E_0[K_2Y_{2,\tau_2}]$	p_0	$A_{1,0}$	$A_{2,0}$	$E_0[K_1Y_{1,\tau_1}]$	$E_0[K_2Y_{2,\tau_2}]$	p_0
0.24	0.40	31.56	32.67	0.79	0.49	0.13	31.56	32.67	0.54

Table 5, Comparison of discount factors A_1, A_2 and asset allocation p_0

As seen from the Figure 1, $A_{i,t}$ tends to converge to 1 at the time τ_i. So, in Fund 1, we see $A_{1,t}$ takes a value larger than $A_{2,t}$ as t approaches τ_1, even though member 2 has a negative correlation. The low starting value of $A_{i,t}$ for a member with a positive correlation in Fund 1 implies that the manager can down-play this liability portion at the beginning and let the positive correlation hedge the liability movement towards the retirement time.

Lastly, similar to the previous subsection, the coupling effect between $\sigma_{r,i}$'s and τ_i's will be even more pronounced if we change how μ_i's interact with them. The following picture varies the parameter S in Table 1 from $S = -1.5$ to $S = +1.5$.

	Fund 1					Fund 2				
S	$A_{1,0}$	$A_{2,0}$	$E_0[K_1Y_{1,\tau_1}]$	$E_0[K_2Y_{2,\tau_2}]$	p_0	$A_{1,0}$	$A_{2,0}$	$E_0[K_1Y_{1,\tau_1}]$	$E_0[K_2Y_{2,\tau_2}]$	p_0
-0.015	0.23	0.34	45.06	19.70	-0.06	0.40	0.17	45.06	19.70	0.87
0.015	0.28	0.27	45.06	19.70	0.16	0.45	0.11	45.06	19.70	1.09

Table 6, Comparison of discount factors $A_{1,0}$, $A_{2,0}$ and asset allocation p_0: Varying S

Table 6 demonstrates numerically how "coupling" of heterogeneity affects the discount factors. The differences between discount factors $A_{1,t}$ and $A_{2,t}$ within the Fund 1 and Fund 2 themselves are even more pronounced. Comparing across the two funds, while the average growth of two members seemingly is still 4.5%, but because

of the coupling effect, it accentuate the differences of the discount factors between fund 1 and fund 2 especially when $S = 1.5$.

1.3 Economic Importance of Heterogeneity: Simulation Approach

In this section, firstly, we will show through numerical demonstration that the optimal management strategy differs for funds with different characteristics of plan members or heterogeneity. We will further show in the succeeding section, that if an investment manager ignores heterogeneity, it can lead to a severe sub-optimality for the fund. We impose no short-selling as an additional constraint to the numerical exercise. Furthermore, supplementary funding X_i must be such that the fund's wealth is always nonnegative. If the fund's wealth W_{τ_i} at the time that member i retires is less than the promised retirement benefit pay-out $K_i Y_{i,\tau_i}$, the supplementary contribution X_i must be enough to cover the immediate pay-out and to provide capital for future pay-outs. It was not possible to obtain a tractable analytical solution imposing the afore mentioned constraints.

			Fund Member's Parameters					
	τ_1	τ_2	$Y_{1,0}$	$Y_{2,0}$	μ_1	μ_2	$\sigma_{r,1}$	$\sigma_{r,2}$
Fund 1A (S=-1.5)			1.0	1.0	3.0%	6.0%	0.2	-0.2
	25	35						
Fund 2A (S=1.5)			1.0	1.0	6.0%	3.0%	-0.2	0.2

Table 7, Plan members' parameters variation: Fund 1A and 2A

We will use a designated version of two funds described previously for our analysis in this section. The parameters of the fund and its plan members are summarized in Table 7. We choose S and ς such that the two funds are moderately heterogenized. Fund 1A is Fund 1 with $S = -1.5$ whereas Fund 2A is essentially Fund 2 with $S = 1.5$. We fixed ς at 0.2. The salary growth rate and its correlation with the risky asset of plan members in both funds average to 4.5% and zero correlation.

1.3.1 Optimal Decisions under the Effects of Different Heterogeneities

As established previously, our optimal investment and funding decisions are dependent on members' characteristics as well as the current salary and wealth levels. The salary and wealth follow the stochastic processes as described in (1) and (4) for our simulation. Our purpose is to study the probabilistic distribution of decisions and compare across the three cases. We use Monte Carlo simulation to construct the distribution of the optimal investment and funding decisions which will enable us to observe how decisions are taken in each case. We further assume that investment leverage and short-selling is not allowed. The simulation base parameters are described in Appendix II: Parameters for Analysis and Simulation.
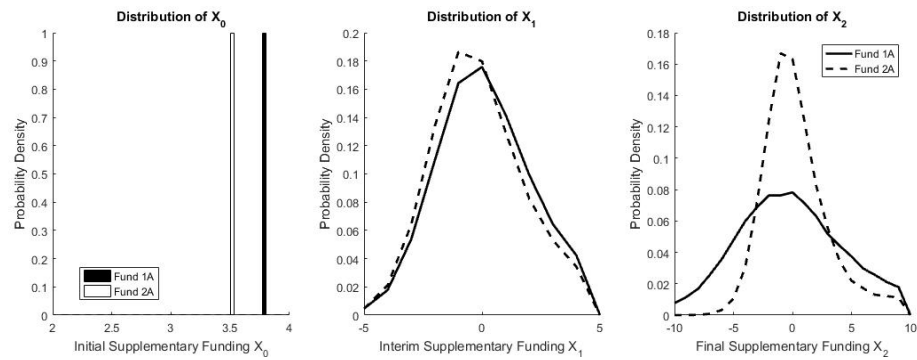
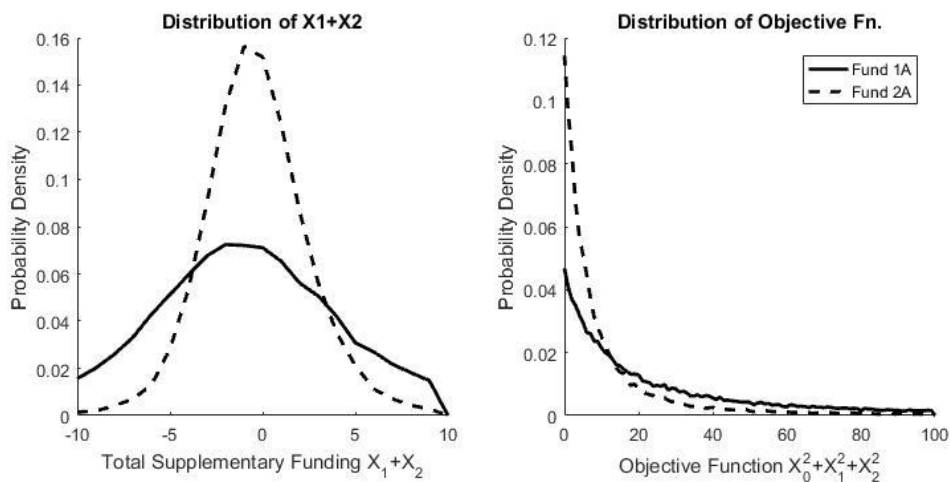*Figure 2, Distribution of optimal supplementary funding: Fund 1A vs. Fund 2A*

Figure 3, Distribution of objective function: Fund 1A vs. Fund 2A

Figure 2 shows the distributions of funding decisions X_0, X_1 and X_2 as well as the distribution of the objective function. Overall, one can see that different cases lead to different funding decisions. The presence of heterogeneity especially from correlation features has a significant impact to the overall supplementary contribution. As seen from Figure 2, the initial funding X_0 is slightly higher in Fund 1A compared to Fund 2A with the presence of negative correlation between salary growth and the risky asset coming from a young member rather than the older member. Terminally, the distribution of final sponsors contribution (X_2) is very different. The distribution of X_2 for Fund 1A has noticeably fatter tails than others. This is due to the inability to hedge the retirement income of the younger member. The management of Fund 1A is unlikely to meet his/her retirement target and results in the worst expected utility. On the contrary in Fund 2A, the management can consistently meets the fund's final wealth target as shown by the distribution peaking around zero.
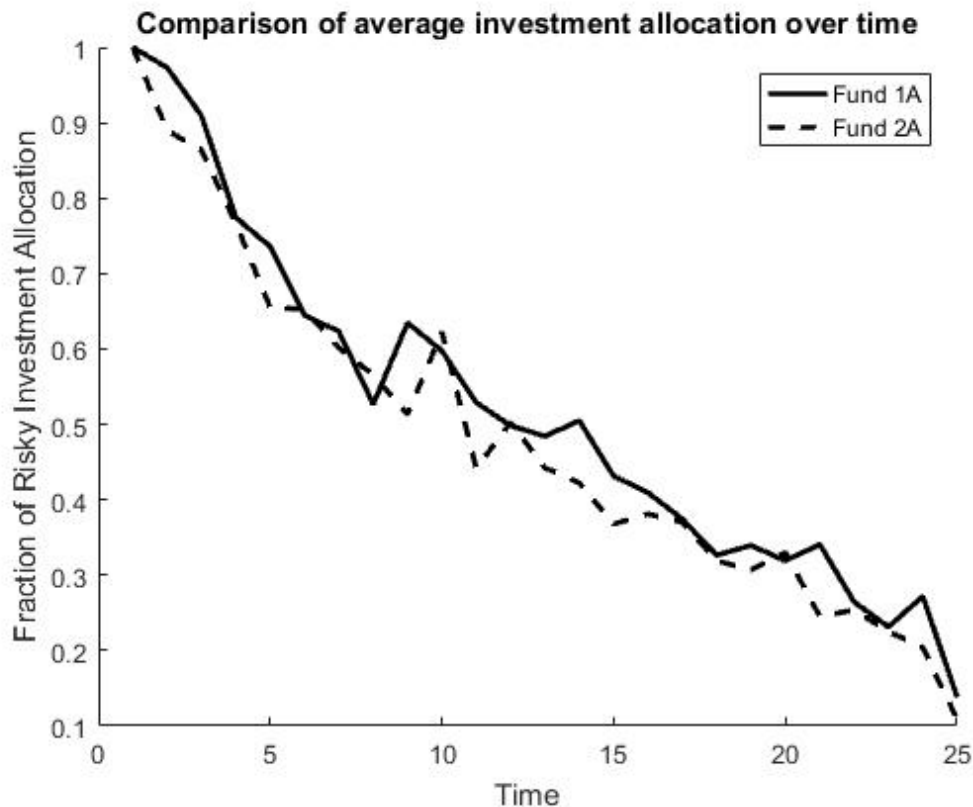
Figure 4, Comparison of average optimal allocation over time: Fund 1A vs Fund 2A

Figure 4 above shows the average risky investment allocation over the course of the fund from time zero to the first retirement where both members are still present in the fund. On average, the average optimal investment for Fund 1A puts more weight in the risky asset than Fund 2A. First and foremost, this figure confirms our understanding of the relationship drawn from our analytical solution. In Fund 1A where the fund is composed of an old member with low growth, positive correlation and a young member has high growth and negative correlation, the investment decisions are likely dominated by the unhedgeable nature of the younger's high retirement income. This forces the fund to try and generate wealth early and take a riskier investment allocation. On the other hand, the unhedgable component of the retirement income is smaller in Fund 2A where the old possess a negative correlation. It is easier for the management to achieve the hedgeable liability target from the younger member later on; we consequently observed that the investments are less risky. Apart from the mean, we have also compared the distributions over time and found that they are noticeably different.

This example shows that heterogeneity affects decisions: the fund manager and sponsors will be taking significantly different optimal actions depending on how the members' characteristics vary. Some combinations of heterogeneity can be more difficult to manage than the others (e.g. Fund 1A as indicated by higher objective score). A member with a negative correlation is easier for the manager when paired with a younger member with positive correlation. Although not shown explicitly in

the numerical example, we expect that a younger group with the possibility of a wider retirement window will be easier to manage.

1.3.2 What if we ignore Heterogeneity?

In contrast to the last section where the manager follows optimal decisions in each respective case, this section shows that if heterogeneity exists but is ignored by the manager, he will end up making a sub-optimal decision.

Suppose two managers are tasked with managing fund 1A. The first manager did not ignore heterogeneity and follow the true optimal funding and investment schedule. Another manager chooses to ignore heterogeneity and mistakenly assume that members are homogeneous with the growth rate $\mu_1 = \mu_2 = 4.5\%$ and $\sigma_{r,1} = \sigma_{r,2} = 0$ which is the average of the two plan members. This manager will be following a different optimal schedule founded on his ignorance. We will demonstrate the extent of severity and losses to the sponsor from such ignorance.

| | Probability Density | |
Objective Score	Fund 1A	Ignore Heterogeneity
0-25	0.5501	0.3116
26-50	0.157	0.0883
51-75	0.0756	0.0496
75-100	0.0429	0.033
100+	0.1744	0.5175

Table 8, Distribution of objective function: True optimal vs. Ignorance of heterogeneity

Table 8 displays the distribution of the fund's objective score or the total funding variation of each fund. The management that ignores heterogeneity has clearly led the fund into suboptimality as shown by the tail of the objective score distribution which is significantly fatter than its counterpart. This fatter tail suggested that the manager is unsuccessful in minimizing the amount and the variance of the sponsor's supplementary funding. It is almost three times (0.518 vs. 0.174) as likely that the ignorance of heterogeneity would result in an excessive funding request from the sponsors. To understand why the management that ignores of heterogeneity would end in such a downfall, we must analyze each decision taken in detail.
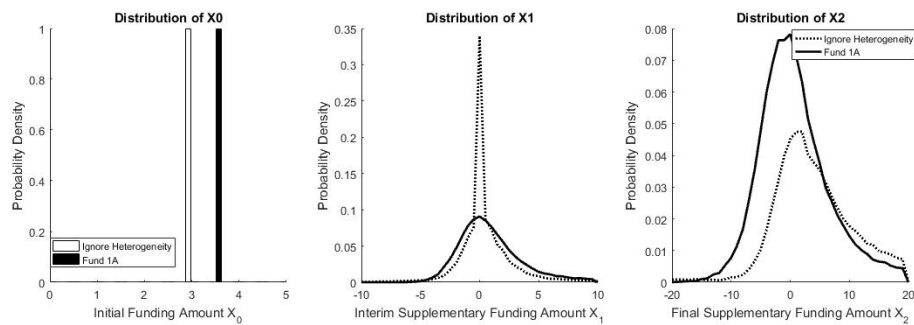
Figure 5, Distribution of supplementary funding: True optimal vs. Ignorance of heterogeneity
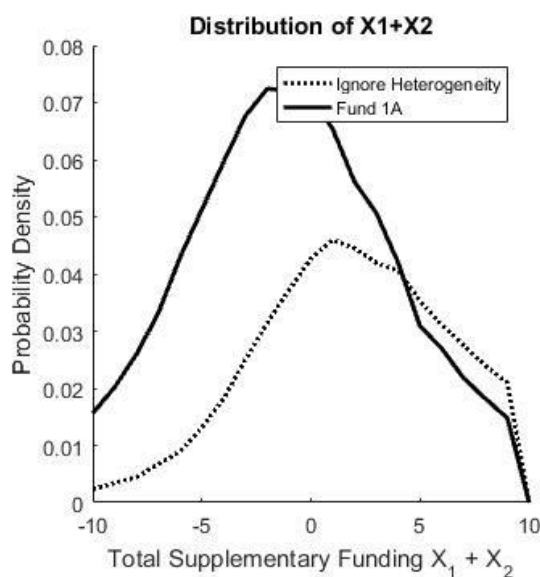

Figure 6, Distribution of total supplementary funding: True optimal vs. Ignorance of heterogeneity

Figure 5 and Figure 6 display the distribution of all the supplementary funding decisions. Let us first analyze the manager that ignore heterogeneity. The manager started out with a noticeably lower initial funding X_0 compared to its counterpart (Ignore: 3 vs. True: 3.8). Combined with investment return, this amount is deemed sufficient for the first retirement benefit payout by the manager. At the first retirement, the manager did not fully utilize this opportunity to adjust the fund's asset level. We observed a much less variability in the distribution of the interim funding X_1 from the manager that ignored heterogeneity. Note that the first retirement pay-out is for the older member with low salary growth and positive salary growth correlation with the risky asset. At this time, the harder target to hit is still yet to come (i.e. the liability owed to the younger, high growth and negatively correlated). The downfall of this manager occurs at the retirement of the younger member. We observed a significant increase in the amount and variance of the final funding X_2 which are the main source of sub-optimality to the objective score. The manager who ignores heterogeneity will fail to anticipate the need to hedge against any salary growth and risky asset co-movement which will result in severe suboptimality for the fund and the sponsor

The investment decisions taken by the manager told a similar story. As shown in Figure 7, the investment decisions taken show that no signs of anticipation for large unhedgable liability. Even with lower initial funding, there is no attempt to accumulate wealth early. The investment decisions taken by the manager is generally less risky than that of the true optimal.
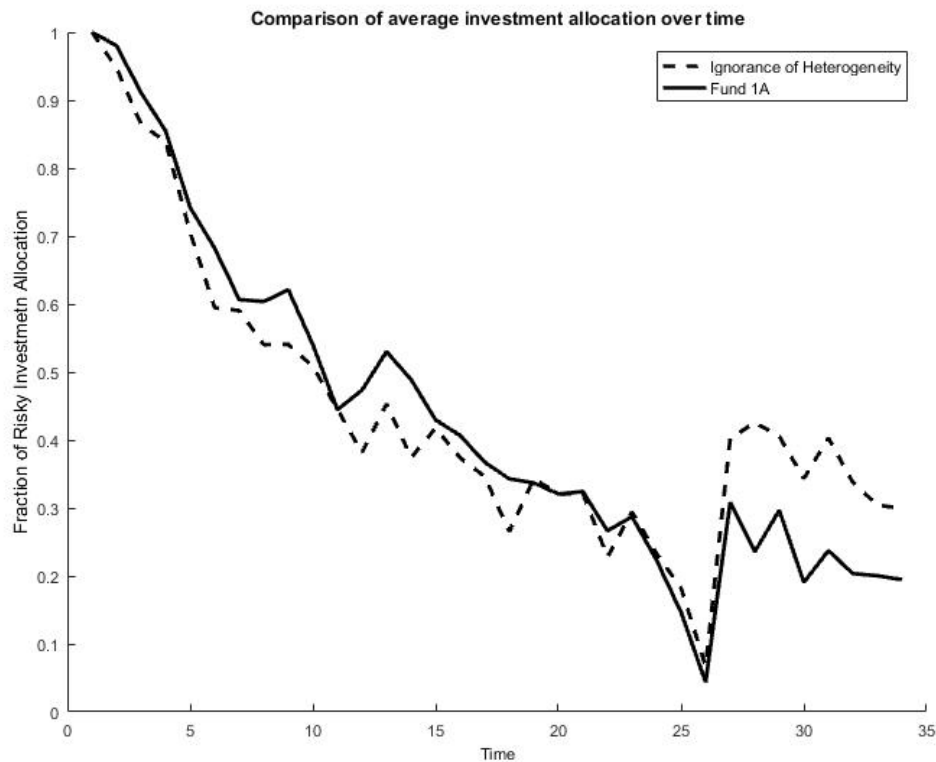

Figure 7, Comparison of average optimal allocation: Fund 1A vs. Ignorance of heterogeneity

Our analysis shows that on average, managers that ignore heterogeneity will end up requesting more funding from the plan sponsors. This can be accounted for the failure to anticipate the co-movement or reverse movement between the fund's investment return and the fund's liability. We expect that the sub-optimality can be even more severe under some adverse realized paths. In an adverse situation, the fund's investment return and net asset value are likely to go down, but your liability increases due to the negative correlation between the salary growth of plan members and the market. Incapability to anticipate and prepare for this realization will induce unnecessary cost to the plan sponsor.

1.4 Chapter Conclusion

In this chapter, we propose a multi-member defined benefit pension fund model that allows for heterogeneity among plan participant. We have demonstrated how demographic differences such as the retirement date and salary of plan members can affect the ongoing management decisions.

We derive an analytical solution to a simplified setting of our model where only two members are present. The optimal asset allocation turns out to be a proportion to total discounted future liabilities over the current wealth of the fund. Additionally, the optimal funding support depends on the sum of the discounted liabilities less any accumulated wealth. Our model conforms to the traditional liability-driven investment framework.

Heterogeneity features prompt different discount rates on the expected liabilities from each fund member. As shown in Section 1.2, it is entirely possible for an expected liability to be heavily discounted, depending on the period of time and the nature of heterogeneity. These discount factors, in turn, determine the optimal decisions over the course of the fund.

To disregard or misrepresent the heterogeneity can lead to suboptimal decisions. Optimal investment paths from cases with and without heterogeneity considered can differ by a large margin towards each of the retirement. The ignorance of seemingly unimportant correlational heterogeneity can increase the expected funding variation by a significant amount.

2 Adapting to Heterogeneity in Multi-Member Pension Fund Management

In the previous chapter, we have demonstrated how the coupling of demographic differences or heterogeneity of plan members can affect the management decisions over the course of the fund. We have also demonstrated that the simplistic method of replacing the entire cohort of heterogeneous members with one static representative agent, while computationally easy, would yield sub-optimal decisions that overlook the heterogeneity effects. Chapter 1 asserts the need for a model that allows for the interaction between multiple aspects of heterogeneity, which in turn raises the need for a practical way to solve for the corresponding optimal decisions.

In Chapter 1, we have employed a discrete-time dynamic programming framework and developed a model of manager's investment decisions that take into account plan members' heterogeneity in retirement age, salary growth, salary volatility, etc. Then, under a simplified assumption that allows short-selling, we have presented a closed-form solution of the optimal investment decisions, and use it to analytically demonstrate the coupling effect of heterogeneity features on fund management. We have found that both the optimal fraction of risky investment and funding decisions are proportional to the expected future liabilities of the fund, discounted by a heterogeneity-adjusted factor. Our analysis of the closed-form solution suggests that fund manager to treat each expected liability from heterogeneous plan participants differently in which they should give a higher valuation to the liability owed to members who are young and has high salary growth with low or negative correlation with the investment asset.

The closed-form solution of management decisions presented in Chapter 1, while useful in the analytical investigation of the heterogeneity effects, exists only under very simplified assumptions. If we impose practical constraints such as short-selling restrictions, the analytical solution is unavailable and we generally require numerical methods to solve for the optimal management decisions. The main objective of this Chapter deals with how to arrive at the optimal decisions in a manner that appropriately accounts for heterogeneity while being parsimonious at the same time. The issue here is that, in contrast to the relative ease with which one can solve a single-member problem, the computational effort required to solve a multi-member problem grows exponentially with the number of members. This is a challenge that one must face when managing a fund with a large number of heterogeneous members.

How to overcome this exponentially-increasing computational cost, in spite of the fact that the true optimal decisions that account for heterogeneity *cannot* be obtained by simply replacing the entire cohort with one representative agent over the entire period of the problem? To solve this dilemma, this paper presents a novel idea that we can arrive at the desired optimal solution by making decisions based on a special type of representative agent whose characteristic "update" at every time step to reflect the new environment of heterogeneity. (Hence, we name such an agent the *Adaptive*

Representative Agent or *ARA*). If we can find such an agent at each time, then the problem of finding the optimal decisions for a fund of heterogeneous members amounts to solving a *series* of single-member problems along the duration of the fund. Thus, the computational effort becomes linear in the number of time steps rather than exponential in the number of members.

Our proposed Adaptive Representative Agent (ARA) represents a time-specific adjustment for the effect of heterogeneity. The representative member at time t turns out to be the weighted average of the members. We give more weight to plan member with higher heterogeneity-adjusted expected liability. A young member with a negative salary growth correlation with the investment asset i.e. an unhedgable salary will contribute more towards the determination of both the optimal decisions and the representative agent. Even though model here is quite stylized (e.g., we assume no new members), the conclusion gives ideas how to weight and adjust that can be applied in practice.

We compare the characterization and performance of our ARA approach against two benchmarks in a multi-member heterogeneous fund set up. Our ARA is shown to be better at estimating the projected liability when there is a coupling of heterogeneity. The ARA approach is better prepared for an adverse situation in which the fund's net asset value is likely to go down, but your liability increases due to the negative correlation between the salary growth of plan members and the market. Our approach generally requests for higher initial funding than its counterparts but take slightly less risky investment choices over the course of the fund.

This chapter is organized into the following sections. Section 2.1 proposes the derivation and how to approximate a representative agent for our multi-member fund model. Section 2.2 shows a performance comparison between our ARA and a naïve agent. Section 2.3 concludes the chapter.

2.1 Proposed Methodology for Characterization of Adaptive Representative Agent

In this section, we aim to provide a procedure for finding the optimal investment decisions in a multi-member pension fund, with more practical constraints such as short-selling restrictions. Under such constraints, the analytical solution provided in Chapter 1 is no longer applicable, and we typically require numerical methods to solve for the optimal decisions. This is relatively easy if there are only a few members in the fund. But when there are many heterogeneous members in the fund, the numerical methods will require computational effort that grows exponentially in the number of members. In the following, we will propose a procedure, the Adaptive Representative Agent (ARA) approach, which reduces the computational burden and achieves the desired solution to the heterogeneity-laden problem.

Our approach recognizes that, in finding the true optimal solution for the multimember fund, it is ill-advised to over-simplify the computational effort by replacing the pool of heterogeneous members with one representative agent for the entire duration of the problem. Instead, the ARA approach attempts to arrive at the desired heterogeneity-adjusted solution by making decisions based on the so-called "adaptive" representative agent whose characteristics "update" at every time step to reflect the latest heterogeneity environment. Provided with such an agent at each time, finding the optimal decisions for a fund of heterogeneous members is tantamount to solving a *series* of single-member problems along the duration of the fund. In other words, we achieve the desired solution to the heterogeneous problem using a computational effort that is, not exponential in the number of members, but rather linear in the number of time steps. We now proceed by characterizing the adaptive representative agent.

2.1.1 Adaptive Representative Agent (ARA): Multi-member fund

At each time t, we let the adaptive representative agent be characterized by her salary $Y_{R,t}$, the salary growth $\mu_{R,t}$, the salary volatilities σ_t and $\sigma_{r,R,t}$, the contribution rate $\pi_{R,t}$, the retirement benefit $K_{R,t}$, and the effective retirement time $\tau_{R,t}$. (Note that these parameters are subscripted by t, since the ARA updates its parameters every time to reflect new heterogeneity environment.) To find the ARA's parameters, we consider, for the time being, the unconstrained case where short-selling is allowed. Using the special case of Proposition 1 where $N = 1$ and noting that $\alpha_t = 1$ for all $t > 0$ when there is only one member, the optimal investment decision based on ARA at time t is:

$$\hat{p}_{R,t} = \frac{\xi}{\sigma_r(1+\xi^2)}\left[\hat{A}_{R,t}\frac{K_{R,t}(1+\mu_{R,t})^{\tau_{R,t}-t}}{W_t}Y_{R,t} - 1 - r\right], \tag{19}$$

where $\hat{A}_{R,t}$ is computed by the recursion:

$$\hat{A}_{R,t} = \hat{B}_{R,t+1}\frac{1 + \sigma_{r,R,t}\,\xi^{-1}/(1+\mu_{R,t})}{1+r} - \frac{\pi'_{R,t}/K_{R,t}}{(1+\mu_{R,t})^{\tau_i-t}}, \tag{20}$$

$$\hat{B}_{R,t} = \hat{B}_{R,t+1}\frac{1 - \sigma_{r,R,t}\,\xi\,/(1+\mu_{R,t})}{1+r} - \frac{\pi'_{R,t}/K_{R,t}}{(1+\mu_{R,t})^{\tau_i-t}}, \tag{21}$$

for $t = \tau_{R,t} - 1, \tau_{R,t} - 2, \dots, 1, 0$, with $B_{R,\tau_{R,t}} = 1 + r$, and $\pi'_{R,t} := \pi_{R,t}1_{\{\tau_{R,t}<t\}} - K_{R,t}1_{\{\tau_{R,t}=t\}}$.

Because the decisions based on ARA should match the true optimal decisions of the multi-member fund that the ARA represents, we choose the parameters $\mu_{R,t}, \sigma_{R,t}, \pi_{R,t}, K_{R,t}, \tau_{R,t}$ and $Y_{R,t}$ in (19) such that $\hat{p}_{R,t} = p_t^*$, where p_t^* is the true optimal investment decision (12) for the underlying heterogeneous-member fund. (Since we focus on representing only the remaining members at each time t, we can assume that $\alpha_t = 1$ in the expression (12) for p_t^*.) The existence of such an adaptive representative agent is established in the following proposition (see Appendix III for detailed proof).

Proposition 2: *At each time t, let $Y'_{i,t} = Y_{i,t}\,1_{\{t<\tau_i\}}$, $i = 1, \dots, N$. Let $Y_{R,t}$ and $\pi_{R,t}$ be given by*

$$Y_{R,t} = Y'_{1,t} + Y'_{2,t} + \cdots + Y'_{N,t}, \tag{22A}$$

$$\pi_{R,t} = \frac{Y'_{1,t}}{Y_{R,t}}\pi_1 + \frac{Y'_{2,t}}{Y_{R,t}}\pi_2 + \cdots + \frac{Y'_{N,t}}{Y_{R,t}}\pi_N, \tag{22B}$$

Also, let $\mu_{R,t}, \sigma_{r,R,t}, K_{R,t}$ and $\tau_{R,t}$ be such that:

$$\mu_{R,t} = w_{1,t}\mu_1 + w_{2,t}\mu_2 + \cdots + w_{N,t}\mu_N, \tag{22C}$$

$$\sigma_{r,R,t} = w_{1,t}\sigma_{r,1} + w_{2,t}\sigma_{r,2} + \cdots + w_{N,t}\sigma_{r,N}, \tag{22D}$$

$$1/K_{R,t} = w_{1,t}/K_1 + w_{2,t}/K_2 + \cdots + w_{N,t}/K_N, \tag{22E}$$

$$1 = w_{1,t} + w_{2,t} + \cdots + w_{N,t}, \tag{22F}$$

and

$$w_{i,t} := \frac{B_{i,t+1}K_i(1+\mu_i)^{\tau_i-t-1}Y'_{i,t}}{B_{R,t+1}K_{R,t}(1+\mu_{R,t})^{\tau_{R,t}-t-1}Y_{R,t}}. \tag{22G}$$

where $B_{i,t+1}$ and $B_{R,t+1}$ are computed from (15) and (21), respectively. Then, where $\hat{p}_{R,t}$ is as given in (19) and p_t^* is as given in (12), we have that $\hat{p}_{R,t} = p_t^*$. □

Proposition 2 provides a recipe to compute the parameters $\mu_{R,t}, \sigma_{r,R,t}, \pi_{R,t}, K_{R,t}, \tau_{R,t}$ and $Y_{R,t}$ of the adaptive representative agent. We first compute the values of $B_{i,t+1}K_i(1+\mu_i)^{\tau_i-t-1}Y_{i,t}$ for every $i = 1, \dots, N$, then scale them appropriately to obtain $w_{1,t}, \dots, w_{N,t}$ that sum to 1. Then we use $w_{i,t}$'s to compute the parameters

$\mu_{R,t}, \tau_{R,t}$ and $K_{R,t}$ from (22C – 22E). Lastly, we calibrate τ_R so that the value of $B_{R,t+1} K_R (1 + \mu_R)^{\tau_R - t - 1} Y_{R,t}$ is indeed consistent with $w_{1,t} + \cdots + w_{N,t} = 1$ (In terms of our numerical implementation in the next section, we round τ_R to the nearest integer).

We now discuss the intuition behind Proposition 2. As seen from (22B – 22E), the adaptive representative agent comes from taking an average of the heterogeneous members. However, each member is not given equal weight. (Doing so would, in fact, distort the heterogeneity effects and lead to sub-optimal decisions, as we will demonstrate in a later example.) For computing $\mu_{R,t}, \sigma_{r,R,t}$, and $K_{R,t}$, the weight $w_{i,t}$ depends primarily on $K_i (1 + \mu_i)^{\tau_i - t - 1} Y_{i,t}$, which roughly reflects member i's expected future retirement benefit. This is consistent with the concept of liability-driven management mentioned in previous sections. But here, in (22G), the future liabilities are also discounted with the heterogeneity-adjusted factors $B_{i,t+1}$'s, as a way to account for the coupling effect of heterogeneity.

Since the weight $w_{i,t}$ is proportional to $B_{i,t+1}(1 + \mu_i)^{\tau_i - t - 1}$, it follows from (16) that the ARA characteristics will be dominated by members with high salary growths and negative correlation. Members whose salary growths are low and/or strongly positively correlated with the investment return are not given as much weight because their retirement benefits are deemed "easy to hedge", in the sense that their salaries, which determine their final retirement benefits, have low growth rate (and/or, even if their salaries do grow, the increase in retirement benefits will presumably be hedged by the accompanying positive investment returns).

In the general case where we have extra constraints like short-selling, the optimal decision cannot be written in closed form as in (19). However, we will still use this ARA to solve the constrained case numerically. We summarize the procedure in the next sub-section.

2.1.2 ARA Approach: Procedure for Implementation

The adaptive representative agent gives us the idea to solve for the solution of the heterogeneous problem. The procedure is summarized as follows:

1. For every t in 0 to τ_2 (the last retirement), an agent is created based on (22)

2. At every t, solve the original minimization problem for the ARA to obtain the optimal asset allocation and the optimal initial funding decision (at t=0)

3. Use the attained optimal decisions and observe the realization of state variables for this time period. Repeat 1 after each state realization.

Let us compare this method with the traditional way. To solve an N-member problem straightforwardly, we could have up to N+1 state variables (the salary of each member plus the fund's net worth). This is computationally expensive for large N. However, the procedure above only requires repeated calls to a one-member sub-routine. So we have reduced the computer time.

The proposed procedure is in line with the current representative agent practice. Instead of replacing the entire cohort with a salary-weighted or simple average representative agent, our representative agent's characteristics "update" at every time step to reflect the new environment of heterogeneity. Our procedures already accommodate the situation where new members are joining or leaving the fund. The investment manager can simply form a new ARA based on the current members of the fund. We illustrate the implementation and performance of our ARA approach in the upcoming section.

2.2 Performance of the Proposed ARA Approach: Numerical

Examples

In this section, we provide a numerical example to demonstrate our method for finding the investment decisions in a DB pension fund with heterogeneous members, under practical constraints such as restrictions on short-selling. Since a closed-form solution is unavailable under such constraints, we will solve this stochastic control problem numerically. But instead of using a straightforward numerical method that assigns state variables to every member in the fund, which would require computational efforts that grow exponentially with the number of members, we will use our proposed ARA approach.

2.2.1 Parameters and Benchmark Settings

Recall that, under the ARA approach, the manager makes decisions based on an adaptive representative agent whose characteristics $\mu_{R,t}, \sigma_{r,R,t}$, etc. at each time t come from taking a weighted-average of the heterogeneous members' characteristics. The weight $w_{i,t}$ given to each member i is proportional to his/her expected future retirement benefit discounted by a heterogeneity-adjusted factor, as prescribed in (15).

To demonstrate the advantage of this ARA approach, we shall compare it to a benchmark "simple-average" approach, in which the representative agent at each time t comes from taking a simple average (equally-weighted) of the members who remain in the fund at that time. We will show that the simple-average approach fails to account for the coupling effects of heterogeneity, and leads to sub-optimal decisions. In comparison, we shall see that the ARA manager tends to make better decisions (in the sense that it requires less funding, and more moderate amount of risky investment), especially when market conditions are unfavorable. (For completeness, Appendix IV also compares the performance of the ARA approach against a second benchmark which is the salary weighted approach. For further detail or mathematical definition and performance comparison against the salary weighted agent, please refer to Appendix IV and Appendix V.)

The fund in the problem set up contains ten plan members with varying degree of salary growths and correlations as shown in Table 9.

N=10	1	2	3	4	5	6	7	8	9	10
Time to Retirement τ_i	20	22	24	26	28	30	32	34	36	38
Salary Growth μ_i	0.06	0.03	0.06	0.03	0.06	0.03	0.06	0.03	0.06	0.03
Correlation $\sigma_{r,i}$	-0.2	0.2	-0.2	0.2	-0.2	0.2	-0.2	0.2	-0.2	0.2

Table 9, Plan members' parameters for the multimember setting

The oldest member will retire in twenty years whereas the youngest member shall retire after thirty-eight years (note that in our set up, heterogeneity exists in such a way that plan members have high growth and negative correlation at the same time). This manner of interaction, as we explain in the previous section, has a significant impact on the projected liability and therefore on decisions. We will test whether ARA can capture this effect.

To numerically compare the characteristics and performance of the competing approaches, we use simulation methodology to generate random paths for the agents' salary progression and the correlated risky asset return over the course of the fund. At each time, we compute the manager's investment decisions under two approaches, one using our proposed adaptive representative agent, and the other using the simple-average approach. Using 60 simulated scenarios, we compare how the characteristics of the ARA vary, and how the ARA manager performs versus the benchmark simple-average approach.

2.2.2 Characteristics of ARA

Before we discuss the management decisions under the ARA approach versus the simple-average approach, let us first understand the characteristics of the representative agents on which those decisions are based. In the ARA approach, the representative agent's characteristics such as $\mu_{R,t}$ and $\sigma_{R,t}$ are stochastic and path-dependent, based on the realized salary $Y_{i,t}$'s and the prevailing condition of heterogeneity at each time. The solid lines in Figure 8-9 show the mean process of $\mu_{R,t}$ and $\sigma_{R,t}$. (The shaded band represents the range of ±1SD around the mean.) In the simple-average approach, on the other hand, the representative agent's characteristics are deterministic processes, shown as the dashed lines in Figure 8-9. In our numerical example, the representative agent under the simple-average approach has, at most times, a salary growth of 4.5% and zero correlation with the investment return. Note that the interaction between multiple aspects of heterogeneity (members'

age, salary growths, and salary-investment correlations) is not taken into account in the simple-average approach.
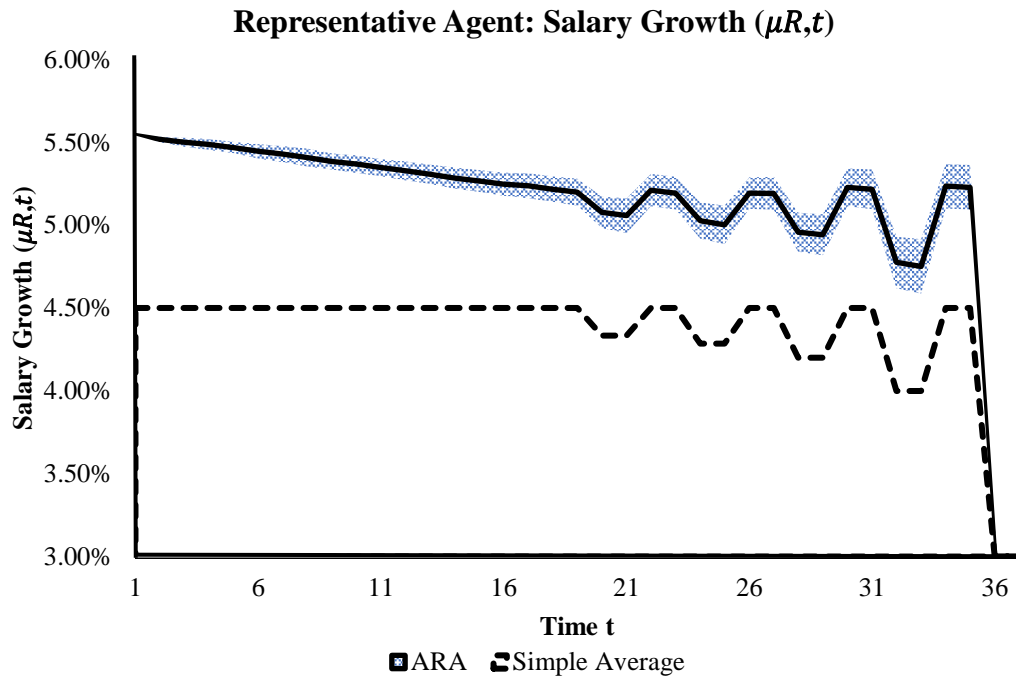

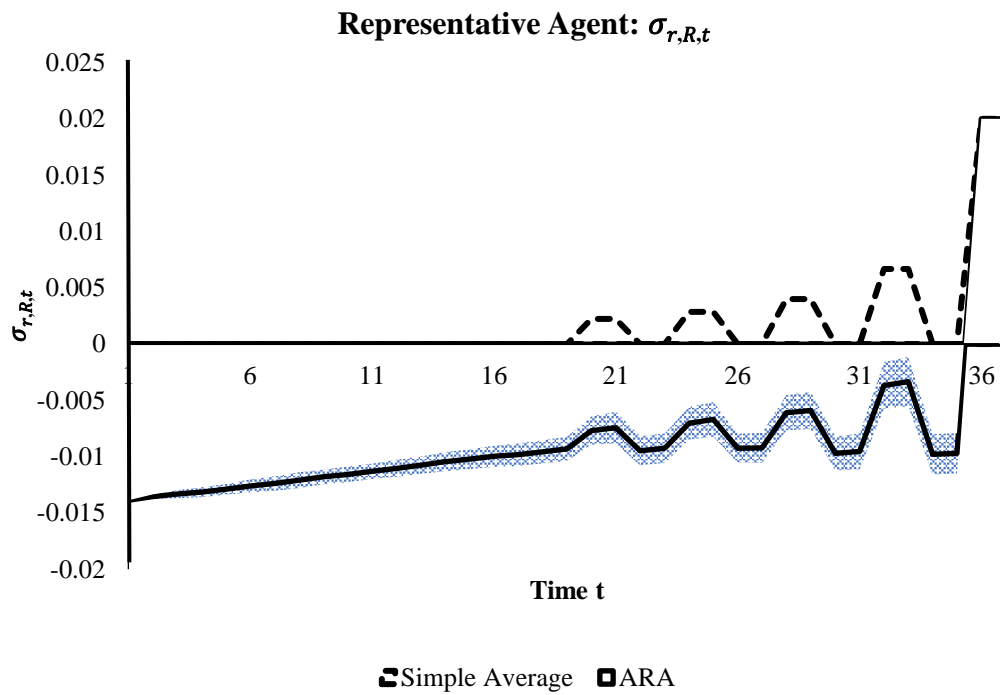*Figure 8, Evolution of Representative Agent: Salary Growth ($\mu_{R,t}$)*

Figure 9, Evolution of Representative Agent: $\sigma_{r,R,t}$

Previously in Section 2.1.2, we intuit that the ARA characteristics should be dominated by members with high salary growths and negative correlation since their retirement benefits are not "easily hedgeable". This is especially true in the early accumulation phase when such members are still young (recall the compounding effect of time-to-retirement on the magnitude of the heterogeneity-adjusted discount factor explained in (15)). As expected, the intuition that we have stated is confirmed by Figure 8-9: While the simple (equally-weighted) average of members' salary growths is only 4.5%, the ARA's salary growth $\mu_{R,t}$ starts at 5.5%, which is biased towards the higher salary growth in the underlying pool of members. And while the simple average of $\sigma_{1,r}, \ldots, \sigma_{10,r}$ is equal to zero, the ARA parameter $\sigma_{R,r,t}$ starts at a negative level -0.14. Note in Figure 9 that the mean value of $\sigma_{r,R,t}$ becomes less negative as t approaches members' retirement time. In other words, although the members whose salary correlates positively with the risky investment are subjugated during the early period, they become more relevant in our ARA as their retirement becomes imminent. (To see this, recall that, for a member with positive $\sigma_{r,i}$, the heterogeneity-adjusted discount factor (15) tends to increase as $t \to \tau_i$.) Intuitively, the fund manager will give more weight to a member with positive $\sigma_{r,i}$ as his retirement approaches, to drive the investment decisions towards the risky asset to mimic the co-movement between salary growth and the risky asset. Note that the ARA characteristics fluctuate noticeably when a member retires and left the fund.

Figure 10 below shows the time to retirement $\tau_{R,t} - t$ of the adaptive representative agent at each time.
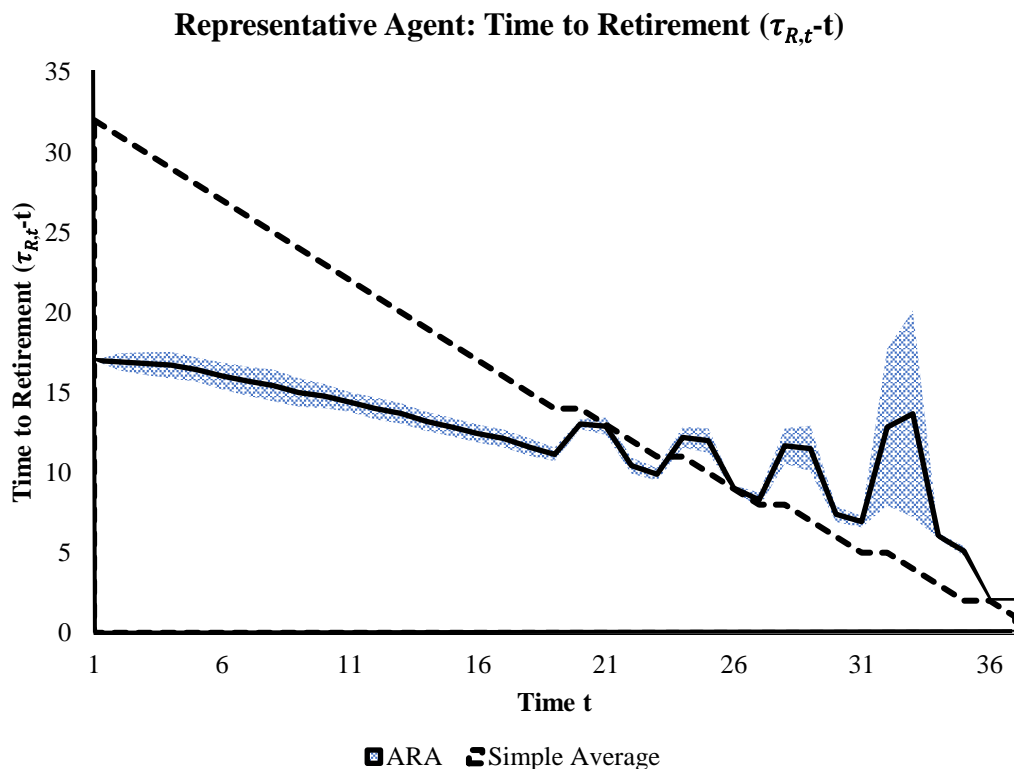

Representative Agent: Time to Retirement ($\tau_{R,t}$-t)

□ ARA   ⊇ Simple Average

Figure 10, Evolution of Representative Agent: Time to Retirement ($\tau_{R,t}$-t)

One can see that the time-to-retirement of the ARA tends to be nearer than that of the simple-average approach. In our example, in which some members have high salary growth that negatively correlates with the investment return, the ARA's retirement turns out to be even earlier than the actual first retirement. That is to say, the coupling of heterogeneity in our setting prompts the ARA to over-value the liabilities by shortening the time-to-retirement. Consequently, as we will see in the next section, the ARA will request for higher initial funding at first, but the overall supplementary funding will be lower and the investment decision will be less risky over the course of the fund.

2.2.3 ARA's Performance vs. Benchmark

In this section, we compared the performance of ARA against the simple average representative agent. We followed the procedure outlined in Section 2.1.2. At any time t, we created our representative agent based on our ARA characterization in (22A – 22G). We then solve the stochastic control problem defined in 1.14 numerically for optimal asset allocation and supplementary decisions. In the numerical method, we prohibit short-selling of both the risk-free and risky assets i.e. $p \in [0,1]$. Furthermore, the each supplementary funding added at retirement must always maintain the fund's asset level above or equal to zero.

In our scenarios analyses, we shall examine the performance based on the total supplementary funding requested by the investment manager. Figures below compare the performance of our novel approach with the two benchmarks.
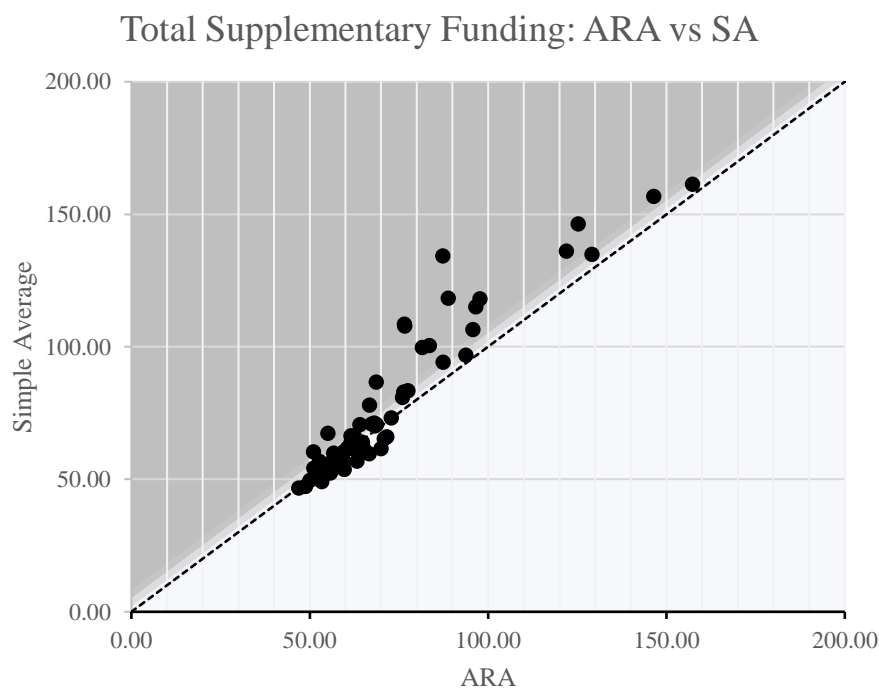

Figure 11, Total supplementary funding: ARA vs. Simple Average

The scatter plots in Figure 11 shows the comparison of the total supplementary funding between the ARA and the simple average agent benchmarks. Each point in the scatter plot compares the total supplementary funding needed for the two different approaches under the same realized scenario. For ease of viewing, we have drawn a diagonal dashed line to partition the graph into two regions: Points above the diagonal line are the scenarios that the simple average approach requires more funding than the ARA and vice versa.

Under favorable conditions i.e. good market returns, both the ARA and simple average managers can consistently meet the retirement target as shown in the lower left quadrant of the scatter plots. The variability in funding in this instance is mainly a consequence of the initial funding choices. The ARA always opted to start with a noticeably higher starting wealth compared to other approaches (ARA: 41 vs. SA: 24). The performance of the ARA approach is slightly worse than its counterparts in favorable conditions due to higher initial funding.

On the other hand, consider the upper quadrant of the scatter plot, it can be said that these scenario realizations are not favorable for the fund in the sense that they are unlikely to be able to pay all the liable retirement benefit without requesting for additional supplementary funding from the plan sponsor. In such adverse situation, the fund's investment return and net asset value are likely to go down, but your liability increases due to the negative correlation between the salary growth of plan members and the market. The simple average approach does not prepare for this eventuality and suffers by having to request for additional supplementary funding. On the other hand, the ARA approach has prepared for this eventuality by requesting higher initial funding ended up requesting much less money. The ARA approach performed better than the simple average approach in these adverse realizations as shown in Figure 11. The biggest funding support discrepancies between the two approaches in our scenarios can be as much as 60% (the simple average would need a total of 135 funding versus 85 for the ARA approach).

Figure 12 below shows the average supplementary funding choices for both approaches. On average, the amount of funding requested by the ARA approach is lower than those of the simple average across most retirements except the final one. Recall that our quadratic objective function minimizes the total *funding variation*, we are also interested in the minimization of the variance of the supplementary funding.

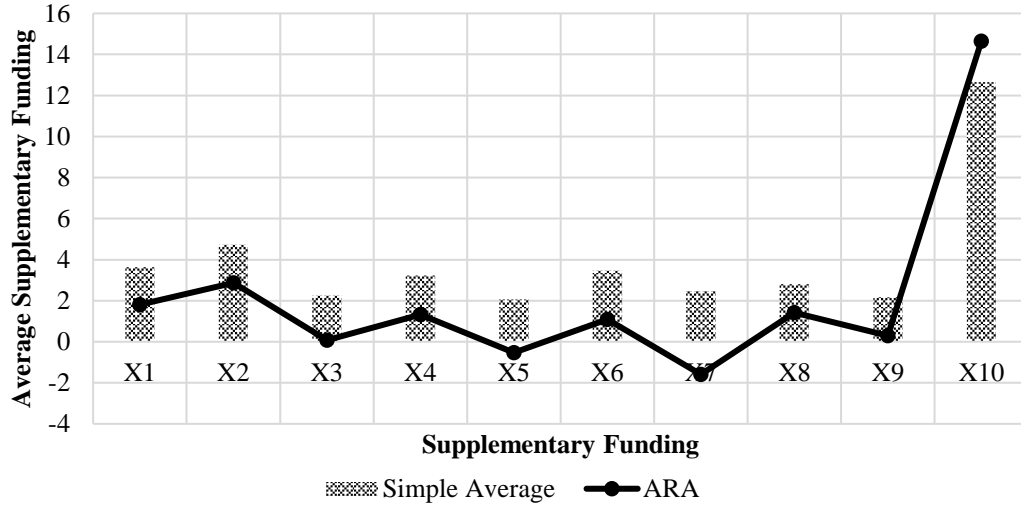**Comparison of Average Supplementary Funding Choices: ARA vs. SA**

Figure 12, Comparison of Average Supplementary Funding Choices: ARA vs. Simple Average

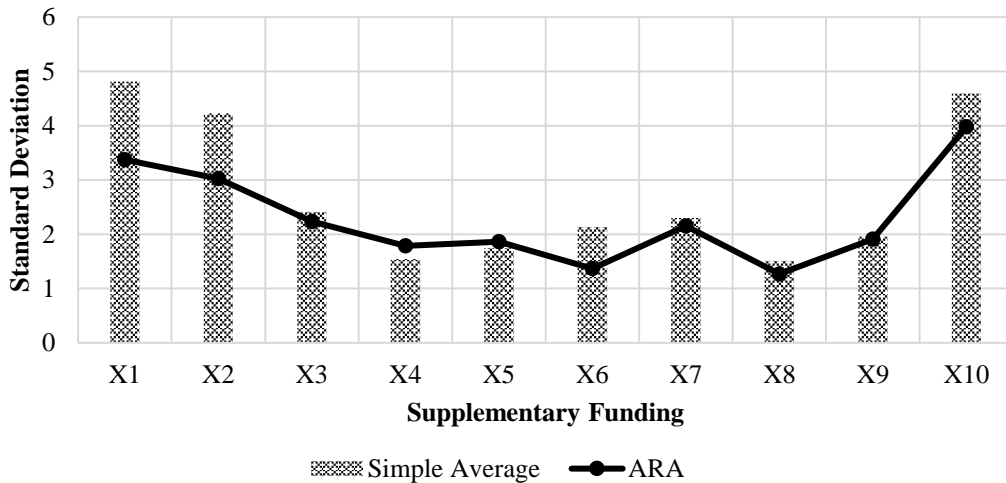**Comparison of Standard Deviation of Supplementary Funding Choices: ARA vs. SA**

Figure 13, Comparison of Deviation in Supplementary Funding Choices: ARA vs. Simple Average

Figure 13 display the comparison of variability in supplementary funding choices over our sample scenarios. These interim funding decisions represent the opportunity for the fund to re-adjust its asset value to be able to sufficiently satisfy all the future liable payouts. In both approaches, the variance of the first available funding request is naturally the highest since it has been the longest time without any funding adjustment. In general, we can see that the variance of these interim fund choices is lower for the ARA approach which can be attributed to better anticipation of the

eventuality by the ARA manager. We observe some fluctuations in the variance of the funding decision at each retirement. The variation is generally lower when a member with a positive salary growth correlation with the market retires due to some co-movement from the fund's investment and the liable payout. Next, we will look at how the investment choices differ between the two approaches.
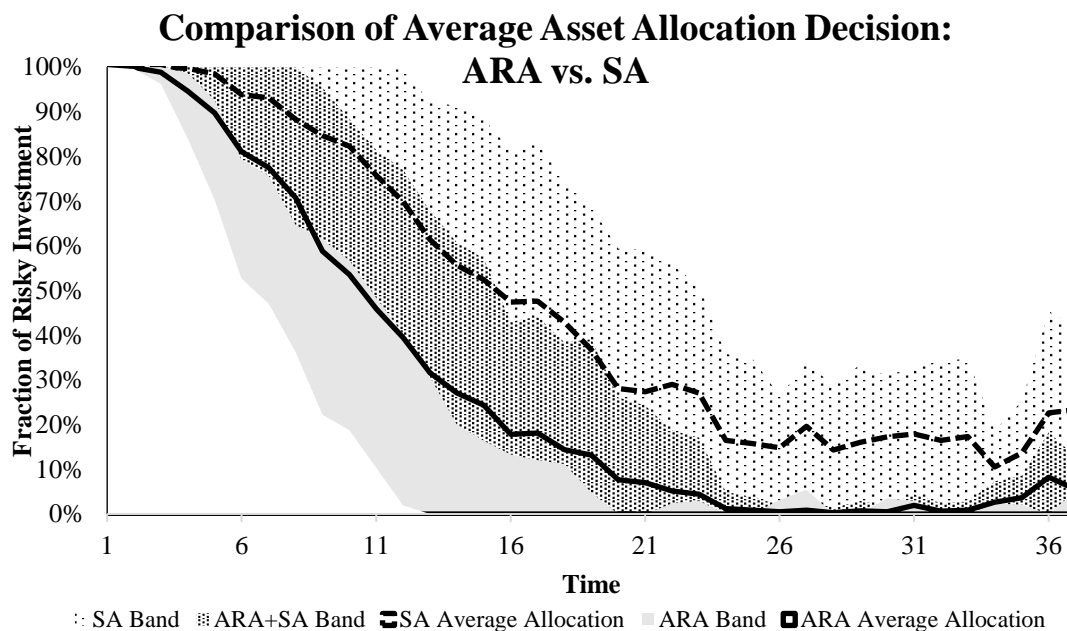

Figure 14, Comparison of Average Investment Decision: ARA vs. Simple Average

From Figure 14, it can be seen that on average the investment of ARA approach is less risky than the simple-average benchmark. This is the result of correctly estimating the fund's future liability at the beginning. As a result, even though both methods starts with 100% in the risky asset, as time progresses the ARA can gradually reduce the proportion of risky asset because it has higher NAV than SA. The simple average approach clearly made much risky investment decisions and even though the investment is riskier, the fund still faced a slight shortage at the last retirement. On the contrary, with higher initial funding for the ARA fund, it allows the investment to generate higher return early on. After sufficient accumulation, the ARA approach reduces risky investment and attempt to reduce the variability in the subsequent supplementary funding decisions. Since this is a scenario analysis, the outcome may not necessary reflects the true variance of the supplementary funding. It is arguable that the naïve manager's risky investment may not produce the same result and the variability in supplementary funding requests may even be bigger. It is not in the best interest of the fund to generate excessive wealth but rather to be able to consistently pay out the retirement income without requiring sponsorship. These decisions differences can be attributed to the differences in their representative agents and the simple average manager's incorrect estimation of the expected retirement liabilities.

In addition to the comparison of ARA with the simple average agent, we also compared the characterization and performance to another representative agent

benchmark: salary weighted approach. The salary weighted approach does not yield any significant advantages over the simple average approach. The key implication and messages are complete with the above analysis. For the detail of comparison between ARA and salary weighted approach, please refer to Appendix V.

2.3 Chapter Conclusion

In this chapter, we propose an adaptive representative agent approach for our problem such that the distinctive features of members are taken into account. Our derivation suggested that each member contribution towards the formation of the representative agent should correspond to their respective fraction of heterogeneity-adjusted retirement liability to the total liability of the fund. For instance, a member with longer retirement time and negative salary-asset correlation may weigh more or less in the construction of the representative agent.

Our results also show that both salary growth- risky asset correlation and age have a strong influence on our solutions. A young member with negative correlation i.e. an unhedgable salary will contribute more towards the determination of both the optimal decisions and the representative agent.

The performance comparison with the naive approach is promising especially in an unfavorable market condition. Our results show that our ARA can anticipate and handle the fund significantly better in unfavorable conditions.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

3 Heterogeneity and the Sustainability of Social Security Fund

In this third chapter, we will turn our attention to another type of retirement fund that is the social security system. Comparable to the defined benefit pension fund studied in the previous chapters, members of the social security plan contribute toward a centrally managed fund in return for promised benefits. Unlike in the DB fund, members of a social security plan are also compensated for sickness, maternity etc. in addition to the retirement income. Since the social security system is normally sponsored by the local government to support generations of the country's retired workforce, the main issue of concerns here is not the sponsorship funding but rather to maximize the sustainability of the system in the face of upcoming country's demographic change.

Previously, we show that characteristics heterogeneity such as the salary growth and its correlation reflected by a small group of defined benefit plan members can have a big impact on the fund's management decisions. But on the other hand, on a country's demographic level, we instead observe differences in labor productivity, its growth, and the labor force distribution across the different industry. Our primary objective of this chapter is to demonstrate that heterogeneity on a demographic scale can also affect the sustainability of the social security system. And that heterogeneity induces by certain economic policy reformation can bring about positive effects to the system's sustainability.

The long-run sustainability of the social security system is uncertain in the face of upcoming demographic shifts. The question of maintaining the sustainability of the system has been at the forefront of the public policy issues. Academics and practitioners agreed that the long-run sustainability of the pension system must come with a reform of the existing system to close the budget gap.

Currently, three of the most important projected demographic shifts which are expected to impact PAYGO system in the 21st century is an increase in life expectancy at 65 by 1.5 years per decade, a decline in fertility, negative rates of population growth and the retirement of the baby-boom generations, born in the 1950s. As a result of these demographic shifts, it is expected that the ratio of retirees to active workers will increase throughout the 21st century and would put the PAYGO system under severe strain. Without any fiscal reformation, the US social security will exhaust the trust fund entirely by 2042. Situations in Japan and Europe are even more intimidating. A vast amount of literature on PAYGO social security systems such as De Nardi et al. (1999), Kotlikoff et al. (2002) and Kitao (2014) focused on suggesting possible social security reformation under these three principal demographic shifts. It is still debatable whether reformation policies like an increasing the social security taxation cap or investment capacity would yield superior sustainability benefit for the fund. The secondary objective of this chapter is to establish that the optimal reformation choice is highly dependent on the country's demographic heterogeneity.

In the current literature, less emphasis is given to the less apparent shift in economic policy such as the change in labor force industry entry distribution, industry growth, and workforce productivity. Conesa and Krueger (1999) study the effect of social security reform with heterogeneous agents differing in productivity and documented significant differences in the sustainability of the social security fund in a closed economy. In addition to heterogeneity in workforce productivity, we suspect that a major shift to the country's economic regime such as the shift from agricultural based economy to a more industrialized economy may disturb the sustainability of the social security system. The shift would affect the social security system in several ways from industry-specific unemployment and productivity to the investment return of the fund as reflected through financial market performance. Collectively with the demographic shifts, the effect these have on the social security fund can be beyond devastation.

In concordance to our previous chapters, we will still be looking into the effect of heterogeneity but this time on a country-level scale and try to understand its effects on the sustainability of a country's social security system. As partially introduced previously, aspects of heterogeneity that we will be investigating this time are the labor force industry distribution, industry-specific labor productivity and its growth and the industry-specific unemployment. Each of these aspects will affect the system's aggregate taxation credit; social security benefit withdrawal and social security investment return which has a direct effect on the system's overall sustainability.

Traditionally, economics literature attempted to compute the long-run stationary equilibrium to evaluate the effects of the demographic shift and its policy reformation countermeasure. A number of researches utilized different formations of General Equilibrium Overlapping Generations models (OLG) which were pioneered by Auerbach and Kotlikoff (1987). In the models from De Nardi et al. (1999) and Kotlikoff et al. (2002), individuals decide on consumption, savings and labor force participation. Static equilibrium implies a balance between government's yearly spending, taxation incomes, benefit payout, clearance labor market and clearance capital market.

In oppose to evaluating the stationary equilibrium, the foundation of this research is to analyze and demonstrate the effects of heterogeneity on the sustainability of the social security system over a provisional time period. We extend the model proposed in Chapter I for this specific purpose.

We found that marginal modification to the growth of an appropriate industry could yield substantial results and deliver good improvement to the sustainability of the social security fund. The combined benefit of promoting growth and workforce entry into an industry with high GDP per capita that is also part of the market index yields the biggest sustainability benefit to the social security fund. On the other hand, clustering workforce into a few correlated industries could make the social security system susceptible to economic shocks.

This chapter is organized into the following sections. Section 3.1, social security model with population heterogeneity, we discuss the overview and the components of our model. In section 3.2, we present the dataset used in our analysis. Section 3.3 shows the comparison of the social security system with different type of heterogeneity. Section 3.4 concludes the paper.

3.1 Social Security Model with Population Heterogeneity

In this section, we present the detail of our social security system and describe each individual model's component.

3.1.1 Labor Force Demographics and Labor Entry

The Population is distributed into twenty-six industry-age cohorts relative to their industry and proposed labor productivity. The age cohorts ($i = 1 \dots 6$) are under 15, 15-29, 30-39, 40-49, 50-59 and retired. Within the working age group ($i = 2 - 5$), the age cohort are sub segmented by industries or sectors. We have six industry cohorts ($j = 1 \dots 6$) which are agricultural, manufacturing & transportation, construction & real estate, wholesale & retail, financial services and others. Let t be a non-negative integer representing time in year. At any time t, we have the following population I

1, Under 15				$I_{1,t}$			
Labor Force	1	2	3	4	5	6
2, 15-29							
3, 30-39							
4, 40-49				$I_{ij,t}$			
5, 50-49							
6, Retired				$I_{6,t}$			

where i_{ij} is the number of population in each cohort

Anyone with the age of 60 or over is considered retired. We assume that the population that is under 15 or retired will not be contributing taxation towards the social security system. The fraction of population in each age cohort j migrate to the next age cohort $j + 1$ according to a constant deterministic age migration matrix M as follows

$$M = [M_i]_{6*1} \text{ where } 1 \geq m \geq 0$$

Populations under 15 enter the workforce at a specified rate M_1 and retired populations leave the system at a specified rate M_6 which correspond to the population's death rate. New born are assumed to enter the under 15 cohort according to a certain birth rate M_0 times the total population.

To incorporate different industry distribution of new labor force entry, we proposed that fractions of the under 15s will join each of the specific industry according to a constant deterministic industry entry distribution L i.e.

$$L = [L_j]_{1*6} \text{ where } \Sigma L_j = 1$$

This will allow us to model different labor force entry distribution heterogeneity. With the age migration and labor force entry rates, we have the following dynamics for population movement from time t to $t + 1$

1, Under 15	$I_{1,t+1} = (1 - M_1)I_{1,t} + M_0\sum I$

Labor Force	1	2	3	4	5	6
2, 15-29	$I_{2j,t+1} = (1 - M_2)I_{2j,t} + M_1 L_j I_{1,t}$					
3, 30-39						
4, 40-49	$I_{ij,t+1} = (1 - M_i)I_{ij,t} + M_{i-1}I_{i-1\,j}$					
5, 50-49						

6, Retired	$I_{6,t+1} = (1 - M_6)I_{6,t} + M_5\sum I_{5j,t}$

3.1.2 Labor Force Productivity, Growth, Taxation, and Expenses

In our model, the productivity of adults in the working age is in the form of the yearly GDP per capita. We allow the productivity to vary across industry and age group. The productivity of population in each working cohort is represented by the following matrix

$$P_t = [P_{ij,t}]_{4*6}$$

Furthermore, yearly industry or sector growth will be reflected through the change in productivity (GDP per capita) of the workforce. The productivity growth rates for each industry are represented by matrix g

$$g = [g_j]_{1*6}$$

The dynamic of productivity in each industry category is modeled as follows

$$P_{j,t+1} = P_{j,t}(1 + g_j)$$

In our model, we assume that a fraction of the workforce is unemployed and that the unemployment rates u_j are industry specific and are constant over time. We have the following unemployment rate matrix u which represents the unemployment in each industry

$$u = [u_j]_{1*6}$$

The active workforce contributes a fraction π of their productivity as taxation toward the social security fund. The taxed amount is capped by P_{cap} which represents the amount taxed from a certain maximum taxable earning. The total social security contribution of the entire workforce is as follow

$$C_t = \sum_{i=2}^{5} \sum_{j=1}^{6} \min(\pi P_{j,t}, \pi P_{cap}) I_{ij,t}(1 - U_j)$$

On the other hand, the unemployed workforce will receive fixed financial aid z from the social security fund. The financial aid to the unemployed workforce will be an expense for the social security fund. The total expense from unemployment financial aid at any time t can be written as follow

$$Z_{1,t} = (z) \sum_{i=2}^{5} \sum_{j=1}^{6} I_{ij,t} U_j$$

In addition to the unemployment financial- aid, we also incorporate two other types of benefit commonly in place for social security contributors. The first type of benefit covers sickness, maternity, invalidity, and death and it's assumed to be a fixed expense Z_2 for the fund. The second type of social security support in our model is the retirement benefits. Annually, the fund pays a fixed benefit to each of the retired population. The total expense incurred to the social security fund from this retirement benefit at time t is

$$Z_{3,t} = K I_{6,t}$$

The total expense incurred to the fund from three categories of social security benefit is given as

$$Z_t = Z_{1,t} + Z_2 + Z_{3,t}$$

3.1.3 Social Security Total Asset, Investment Return

Before we describe the dynamic of the fund's net asset, let us first establish the association between labor force productivity and investment return. For social security fund's investment options, we will follow our formulation in Chapter 1. The management only allocates wealth in a risk-free asset and a market asset.

In this chapter, we suppose that the changes to the industry contribution to the country's overall GDP are directly proportional to the return of the market asset. Each industry's total GDP is equivalent to the number of workers working in that industry times the industry's GDP per capita.

$$G_{j,t} = \sum_{i=2}^{5} P_{j,t} I_{ij,t} (1 - U_j)$$

The fraction of each industry's contribution towards the overall GDP is

$$F_{j,t} = \frac{G_{j,t}}{\sum_{k=1}^{6} G_{k,t}} = \frac{\sum_{i=2}^{5} P_{j,t} I_{ij,t} (1 - U_j)}{\sum_{k=1}^{6} \sum_{i=2}^{5} P_{k,t} I_{ik,t} (1 - U_k)}$$

And the change in fractions contributed towards the overall GDP is

$$\Delta F_{j,t+1} = \frac{F_{j,t+1}}{F_{j,t}}$$

The industry's fraction of GDP contribution towards the market index at any time t is given by the matrix S_t

$$S_t = [S_{j,t}]_{1*6}$$

We suppose that change in fractions contributed towards the GDP will be directly translated to the contribution to the market asset. The dynamic of the industry contribution is given as follow

$$S_{j,t+1} = S_{j,t} (\Delta F_{j,t+1})$$

We assume that a portion of GDP generated from each industry is generated from companies in the market index and that an increase or decrease in this portion will directly affect the market asset return. The industry and total GDP generated by companies in the market index is given by

$$Q_{j,t} = S_{j,t} G_{j,t}$$
$$Q_t = \sum_{j=1}^{6} Q_{j,t}$$

The evolution of market asset return is given as follow

$$R_{t+1} = \frac{Q_{t+1}}{Q_t} R_t$$

The dynamic of the fund's net asset value is then given by

$$W_{t+1} = W_t[(1 - p_t)(1 + r_f) + (p_t)(1 + R_t)] + C_t - Z_t$$

With the current model set up, we can investigate the effects of workforce industry distribution entry and industry growths on the sustainability of a country's social security system. The social security system's sustainability measure here is the survival time until the fund is exhausted without external funding sources. The next few sections describe our datasets and demonstrate the effects of heterogeneity by using a numerical example.

3.2 Data

For our model to authentically reflect the impact of heterogeneity on the sustainability of the social security system, we have to collect various social security system and related data to be applied to our model. The country of interest is Thailand. One of the reasons we choose to study the sustainability of Thailand's social security system is because of the fact that relative to other countries, Thailand is aging fast. In a few years, approximately 1/6 of the country's population would be over 65 years of age and the proportion is rising even faster than in China (Economist, 2018). And despite such swift demographic shift which would place an extremely heavy burden on the country's social security system, there is a lack of urgency for reformation coming from the associated policymakers.

Firstly, let's look at the current demographic composition of Thailand's social security system. According to the latest Thailand Social Security Office (2016) annual report, we have the following age group composition (Table 10)

Age Cohort (i)	Insured			% of Workforce	Population			% of Population
	Male	Female	Total		Male	Female	Total	
6 Over 60	261,662	303,344	565,006	6%	4,340,137	5,461,943	9,802,080	15%
5 50-59	951,083	1,076,280	2,027,363	23%	4,247,100	4,733,659	8,980,759	14%
4 40-49	1,599,401	1,732,791	3,332,192	32%	5,089,037	5,387,628	10,476,665	16%
3 30-39	2,044,778	2,251,651	4,296,429	43%	4,998,559	4,999,388	9,997,947	15%
2 15-29	1,806,505	2,014,004	3,820,509	33%	6,984,227	4,614,136	11,598,363	18%
1 Under 15							15,076,186	23%
Grand Total			14,041,499	28%			65,932,000	

Table 10, Age group composition of Thailand's SSO system, 2016

Furthermore, we assume the following migration rate from one age cohort to another (Table 11). Quantities in the matrix represent the fraction of the population starting in each cohort moving to another. Thailand's birth and death rates are obtained from Thailand Factbook (2018) at 1.0% and 0.8% respectively. However, in our model, we only allow plan members to leave the system through death only after they retired. We assume that 0.8% death rate translates to approximately 3.5% death rate for the retired population.

Age Cohort (*i*)		1	2	3	4	5	6	
	Start/End	Under 15	15-29	30-39	40-49	50-59	Retired	Death
	Birth	1.0% (of Total Population)						
1	Under 15	93.3%	6.7%					
2	15-29		93.3%	6.7%				
3	30-39			90.0%	10.0%			
4	40-49				90.0%	10.0%		
5	50-59					90.0%	10.0%	
6	Retired						96.5%	3.5%
	Death							100.0%

Table 11, Age cohort migration rate

The 2016 annual report (Thailand Social Security Office, 2016) also documented the yearly expense of Thailand's social security office categorized based on the type of compensation (Table 12). From the information the table below, compensation for sickness is at the forefront of the current social security expense followed by the retirement compensation. Currently, only 6% of the populations over 60 of age are insured by the social security office but the retirement expense accounts for almost 13% of the total expenses. We expect this number to rise sharply in the future relative to the increase in the number of retired population. The total benefit expense for 2016 is 73 million THB.

Type	Fund	Fund (% Total)	Million THB	THB (% Total)	THB per case
Sickness	34,939,990	94.22	40,635	55.64	1,163
Maternity	294,169	0.79	7,060	9.67	24,000
Invalidity	12,139	0.03	773	1.06	63,679
Death	25,905	0.07	1,990	2.72	76,819
Child Allowance	1,326,518	3.58	6,653	9.11	5,015
Old-Age	344,937	0.93	9,384	12.85	27,205
Unemployment	141,267	0.38	6,541	8.64	46,302
Total	37,084,925		73,036		

Table 12, Compensation from Thailand's social security system, 2016

As for the current asset value, Thailand's social security investment fund amounts to 1.5 trillion THB in which 83% are invested in government bonds, Bank of Thailand bonds, state-enterprise bonds guaranteed by the Ministry of Finance and investment-grade corporate bonds. The remaining 17% is invested in various unit trusts and equities. In our model, we will assume that the risky asset is the SET50 index fund. These are summarized in Table 13 below. According to the latest (2016) Social Security Committee's regulation on the investment policy, the fund can invest in risky asset up to forty percent of the portfolio.

Current Wealth W_0	1,570,302	Million THB
Highly secured asset	83%	
Risky asset	17%	
Actual Investment return (2016)	3.36%	
Baseline risky return, R_0	1.50%	
Baseline risk-free return r_f	1.50%	

Table 13, Total asset, investment allocation and the return of SS fund, 2016

We also look to the national statistical office to find additional data on labor force industry distribution. According to the summary of the labor force survey in Thailand – January 2017, the current labor force industry distribution is as follow (Table 14). We assume that this distribution applies to all age group of the current workforce.

Industry (j)	Labor Force Industry Distribution L_j
1 Agriculture	28.6%
2 Manufacturing	16.7%
3 Transportation, Storage and Communications	3.5%
4 Wholesale and Retail Trade	18.1%
5 Construction & Mining	6.5%
6 Others	26.7%

Table 14, Thailand labor force distribution, Jan 2017

In addition to the workforce distribution, we acquired Thailand's GDP and GDP by sector information from the Bank of Thailand. Although almost 60% of our workforce is employed in the agriculture and wholesale-retail sectors, sector that contributed the most to the country's GDP is manufacturing at 40%. All the relevant GDP figures used in our model are summarized in Table 15. We assume that the insured workforce in each sector contributes an equal amount to the overall GDP as those who are uninsured.

	%GDP	GDP (Million THB)	Workforce	Base GDP per Capita (THB) $P_{j,0}$	Taxed Contribution per person (THB) $\min(\pi P_{j,t}, \pi P_{cap})$
GDP		12,204,000	41,053,734	297,269	
GDP from the Insured (33% of total workforce)		4,006,143	13,476,495	297,269	
By Industry (j)					
1 Agriculture	8.4	336,516	3,849,909	87,409	4,370
2 Manufacturing	39.2	1,570,408	2,249,101	698,238	9,000
3 Transportation, Storage and Communications	9.8	392,602	470,827	833,856	9,000
4 Wholesale and Retail Trade	13.4	536,823	2,433,810	220,569	9,000
5 Construction & Mining	4.3	172,264	876,461	196,545	9,000
6 Others	24.9	997,530	3,596,387	277,370	9,000

*Table 15, Thailand GDP and GDP by sector, 2017*

Last but not least, in our model, sector performance and fund's investment return are directly linked. As discussed in the model detail section, we proxy each sector performance using the amount of GDP its workforce is producing. We then tie the

yearly changes in the industry's GDP to the composition of the market index, in this case, the SET50. The current sector contributions to the SET50 index are shown in the following Table 16.

	Industry (j)	Contribution to SET50 Index $S_{j,0}$
1	Agriculture	0.00%
2	Manufacturing	17.82%
3	Transportation, Storage and Communications	2.68%
4	Wholesale and Retail Trade	23.82%
5	Construction & Mining	12.18%
6	Others	43.49%

Table 16, Contribution by sector to SET50, 2017

The next section will utilize our demographic, market, assumptions and other data described here to demonstrate the effects of heterogeneity to the sustainability of the social security fund using a numerical example.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY
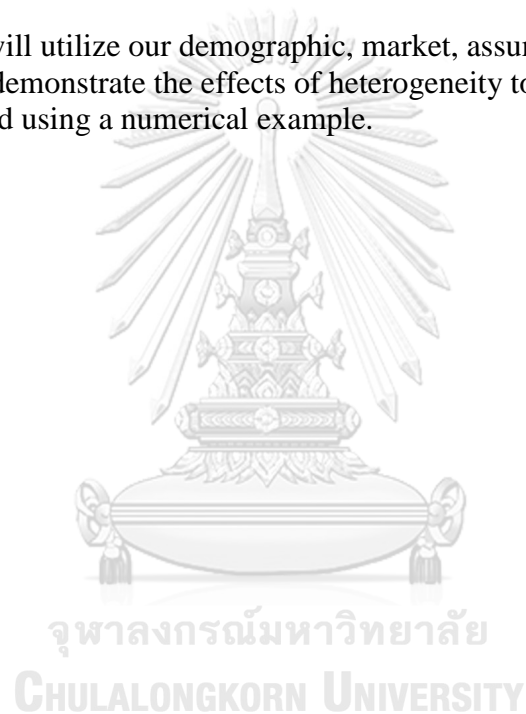
3.3 Heterogeneity Effects on Social Security System Sustainability:

Numerical Demonstrations

In this section, we will use numerical examples to demonstrate how economic policy on industry stimulation and its workforce distribution can affect the sustainability of a country's social security system. As presented in the previous section, our social security system of interest here is Thailand. We will utilize scenario analysis to illustrate our key findings regarding this matter. Most of our model inputs are defined in Section 3.2. For full detail of all the inputs please refer to Appendix VI.

In our analysis, we assume that the birth, death and migration rates (M) to and from each age cohort are constants and are defined by the matrix in Table 11 in section 3.2 Data. With such an assumption, the number of total, workforce and the retired population will be the same for every scenario. The number of insured worker, retirees and the ratio of retired to the workforce for the current demographic condition is shown in the Figure 15 below. If the there is no significant shift in birth and death rates, we will witness a steep rise in the retired-to-workforce ratio over the next twenty years which will result in a significant burden of the country's social security system. This figure would be even worse if all the current population aged over the 60s is to be insured and also if the population has a higher life expectancy.
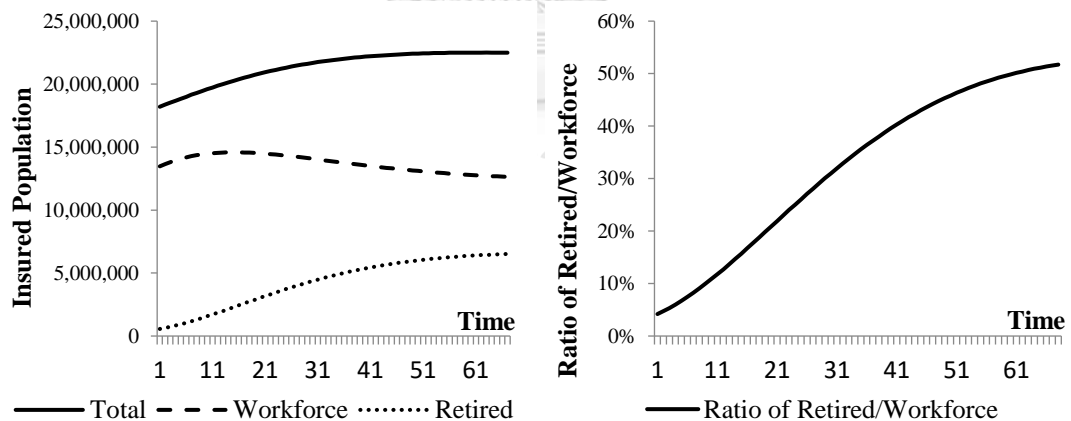
Figure 15, Number of the workforce, retirees and the ratio of retired to the workforce

3.3.1 Sustainability and Productivity Growth and Unemployment Heterogeneity

In the following demonstrations, we will demonstrate the effects of productivity growth and unemployment heterogeneity on the sustainability of the fund. The proportion of current and future workforce industry distribution L is assumed to be constant in this part of the analysis. Suppose there are two hypothetical countries whose demographics are similar to the current state of Thailand as in Section 3.2 Data. However, the industry or sector productivity that contributes toward to the

country's GDP growth and the unemployment are different. It can be postulate that these two countries took different paths in their growth and development policy. In the first country (Country 1), the source of GDP growth comes from growth in productivity from the manufacturing industry specifically. On the other hand, the growth comes from the agricultural industry in Country 2. In terms of the overall GDP growth and the unemployment rate, the two countries are identical. The parameters of interest of each country are shown in Table 17 below.

Industry (j)	Taxed Contribution per person (THB) $\min(\pi P_{j,t}, \pi P_{cap})$	Industry Distribution L_j	Country 1 Growth g_j	Country 2 Growth g_j
Country		100.00%	2.00%	2.00%
1 Agriculture	4,370	28.57%	0.00%	8.20%
2 Manufacturing	9,000	16.69%	1.75%	0.00%
3 Transportation, Storage and Communications	9,000	3.49%	0.00%	0.00%
4 Wholesale and Retail Trade	9,000	18.06%	0.00%	0.00%
5 Construction & Mining	9,000	6.50%	0.00%	0.00%
6 Others	9,000	26.69%	0.00%	0.00%

Table 17, Country 1 vs. 2, productivity growth and unemployment heterogeneity

The economic development policy adopted in Country 1 and Country 2 are very dissimilar. Country 1 focused on enhancing one of the industries with the highest productivity whereas Country 2 focused on the industry with the lowest GDP per capita. These policies are expected to have different implications to the social security fund. In Country 1, growth in the manufacturing industry will not increase any taxation contribution from the workforce since their GDP per capita is already well over the maximum taxable amount. However, the growth will instead be reflected in the social security fund's investment return since the manufacturing industry is a significant part of the market index as shown in Table 16. Country 1 here would represent a reformation where the policymaker tries to increase the investment return of the fund by promoting a specific industry.

On the other hand, the opposite will be observed from the policy adopted in Country 2. The agricultural industry has an extremely low GDP per capita which is a distance away from the taxation cap. Growth in the agricultural industry will increase taxation contribution from the agricultural workforce. However, the agricultural industry is not a constituent of the market index. We do not have any increase in the investment return for Country 2. As opposed to the previous instance, Country 2 would represent a reformation where the policymaker tries to increase the taxation contribution to the social security fund.
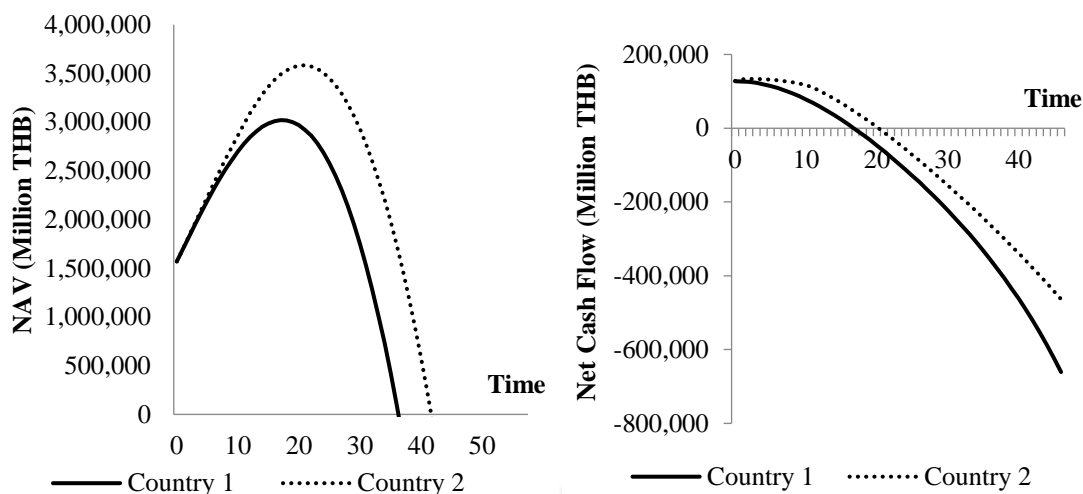
Figure 16, Country 1 vs. 2, social security fund net asset value and cash flow

The left plot in Figure 16 shows the progression of a fund's net asset value over time (years) for each country. The lifetime of the social security system in Country 1 is 37 years whereas the lifetime estimate (time until the fund's NAV hit zero) for Country 2 is considerably longer at 42 years (+5 years or +15%). It appears that the benefit of promoting growth in an industry with low GDP per capita outweighs the growth in high productivity industry even though the industry does not constitute toward the market index. The increase in contribution taxation from the workforce in the agricultural sector outweighs the increase in the fund's investment returns. It eventually took 28 years for the agricultural workers to reach the social security taxation cap. Productivity growth in an underdeveloped industry has a long-term effect on the fund. It increases the future contribution of the new workforce.

The right-hand side plot of Figure 16 displays the net cash in or outflow of each country's fund at a given time. The difference in time to the first deficit between Country 1 and 2 is six years. The funds reach deficit after the ratios of retired to working population are 17% and 21% respectively for Country 1 and 2. The gap or differences between net cash flow widens significantly over time and can be accounted for the differences in the investment return and contribution funding. The next figure shows the respective fund's income for Country 1 and 2.
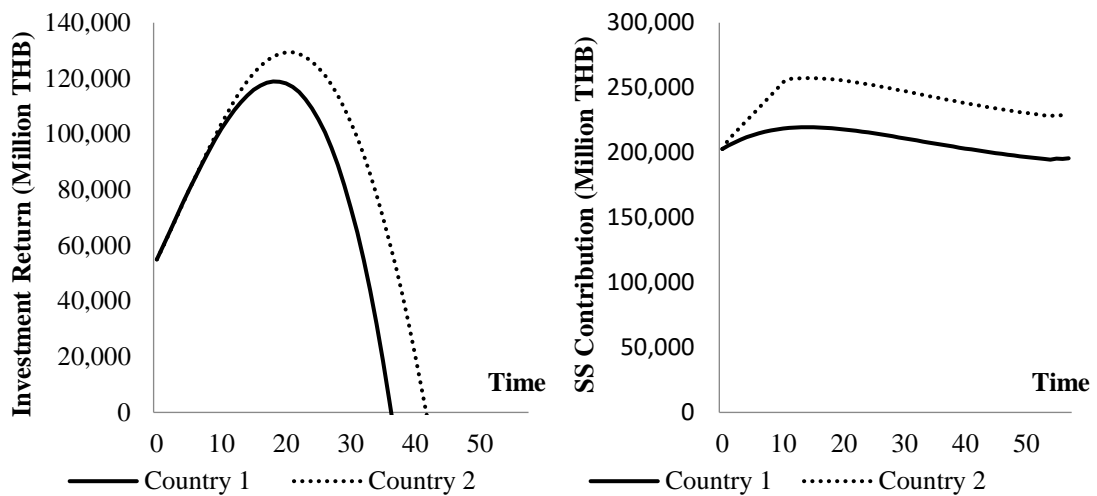
Figure 17, Country 1 vs. 2, investment return and SS taxed contribution over time

The left-hand panel of Figure 17 shows the progression of investment return and taxed contribution over the course of the fund. In the first ten years, the investment return of Country 1 is marginally more than that of Country 2. However, after ten years, the Country 2's fund has accumulated extra wealth from the taxed contribution and naturally generated more investment return. Furthermore, the right-hand panel of Figure 17 shows the social security taxation income. The policy adopted in Country 2 increases this income significantly (about 17% more at steady state). It takes about 10 years for the agricultural industry in Country 2 to reach taxation cap.

We have shown that marginal modification to the growth of an appropriate industry could yield substantial results and deliver good improvement to the sustainability of the social security fund. In our setting, the agricultural sector has relatively low GDP per capita compared to the others. Since initially a large portion of the workforce is already employed in the agricultural industry, the agricultural growth policy seems to yield healthier effects to the social security fund compared to the policy adopted in Country 1. We provide a baseline unifying model for both schemes, and it turns out that heterogeneity in workforce distribution is the key in this model that plays an important role in determining appropriate policy/course-of-actions. To further demonstrate this, based on a fixed 2-percent country's overall GDP growth and demographic conditions as per the previous exercise, the table below shows the social security fund lifetime estimates for different industry-focused growth policies in each of the interested industries (zero productivity growth in other industries similar to Country 1 & 2).

Policy	Industry of Focus (j)	Country's Overall GDP Growth	Industry Productivity Growth g_j	SS Lifetime Estimate
1	Agriculture	2.00%	8.20%	42
2	Manufacturing	2.00%	1.75%	37
3	Transportation, Storage and Communications	2.00%	6.95%	36
4	Wholesale and Retail Trade	2.00%	5.10%	38
5	Construction & Mining	2.00%	15.80%	>60
6	Others	2.00%	2.75%	39

Table 18, Comparison of industry-focused growth policies in different industries and their social security lifetime estimates

The lifetime estimates in Table 18 display substantial disparity among the adopted policy. In all industries except the agriculture, social security taxations are already at the capped amount. Productivity growths are thus reflected in the investment return of the SS fund. In these industries, the lifetime estimates are close to each other between 37 to 39 years which is lower than Country 2 (agricultural) except in the construction & mining industry. Under our model setting and Thailand's demographic conditions, a 15.8% productivity growth in the construction & mining will make the fund self-sustainable. Since the number of the workforce working in construction & mining is relatively low, it required a hefty amount of productivity growth to achieve a 2-percent increase in the country's overall GDP. This high productivity growth has a direct effect on the industry's GDP contribution to the market index and thus the fund's investment return. Although the productivity growth here can be unrealistic, this exercise established the importance of workforce distribution and industry contribution to the market index to the choice of reformation policy and its implication to the sustainability of the social security system. It is also in our interest to find the productivity growth thresholds which make the fund self-sustainable. Country's profile and aspects of heterogeneity will affect this threshold.

Policy	Industry of Focus (j)	SS Lifetime Estimate	Productivity Growth Threshold g_j	Country's Overall GDP Growth
1	Agriculture		Not Achievable	
2	Manufacturing		4.70%	3.17%
3	Transportation, Storage and Communications	Self-Sustainability	10.70%	2.37%
4	Wholesale and Retail Trade		7.30%	2.30%
5	Construction & Mining		11.60%	1.82%
6	Others		4.80%	2.50%

Table 19, Comparison of productivity growth threshold for SS fund self-sustainability across different industry-focused growth policies

Table 19 shows the productivity growth thresholds which make the social security fund self-sustainable (assuming zero growth in other industries). Interestingly, it is impossible to achieve self-sustainability from promoting productivity growth in the agricultural industry under the current demographic conditions. The increase in taxation income will eventually reach a certain limit which is still not enough to support the increasing expense. We can conclude that with equilibrium ratio of the

retiree to the workforce of approximately 50% (Figure 15), the increased in cash flow from social security taxation alone cannot be expected to support an increasing number of retired population. In a country's demography like Thailand, investment return is a crucial element to attain self-sustainability. Furthermore, it can be seen in Table 19 that promotion of productivity growths in several industries that contribute toward the market index can lead to self-sustainability of the SS fund. Manufacturing and others industry required the lowest industry-specific productivity growth to make the SS fund sustainable. The two industries differ in their base GDP per capita (700k THB vs. 277k THB), contribution to the market index (17.82% vs. 43.49%) and the proportion of workforce working in the respective industry (16.7% vs. 26.7%). Even though the contribution to the market index and proportion of workforce working in the manufacturing industry are low, growth on high base GDP per capita can adequately compensate for such numbers. Heterogeneity again plays an important role in determining appropriate policy/course-of-actions.

We did not demonstrate the heterogeneity in unemployment in our numerical example. The effect of unemployment is apparent. In our model, unemployment compensation is a fixed sum for the unemployed worker across all industry. Unemployment in high productivity industry will shorten the lifetime of the fund more than unemployment in low productivity industry like agriculture. This is due to lower taxed contribution from high productivity industry.

3.3.2 Sustainability and the Coupling of Productivity Growth and Workforce Distribution Heterogeneity

In this subsection, we will demonstrate the effects of workforce distribution heterogeneity on the sustainability of the fund. Again, consider two hypothetical countries that are similar to Country 1 and 2 in the previous demonstration. We will refer to these new countries as Country 1A and 2A. The proportion of the initial workforce industry distribution L is assumed to be equivalent to the current state of Thailand as in Section 3.2 Data or as in Country 1&2. We suppose that in Country 1A and 2A, the future labor force entry rate into each industry or workforce entry distribution differs. In Country 1A we will assume that 60% of the new workforce will join the manufacturing industry which correspond to the growth promoted in the previous section. The same principle is also applies to Country 2A, 60% of the future workforce will be joining the agricultural industry in this case. We use this extreme number for workforce entry rate (60%) to demonstrate how workforce distribution and productivity growth heterogeneity coupling can further affect the lifetime estimate of the social security system. Unless stated, we employed the same country's parameter settings as in the previous section. Each country's parameters for this example are summarized in Table 20.

	Industry (j)	Country 1A	Country 2A	Country 1A	Country 2A
		INITIAL ($t = 0$) Workforce Industry Distribution L_j		FUTURE ($t > 0$) new Workforce Industry Entry Distribution L_j	
1	Agriculture	28.57%	28.57%	5.0%	60.0%
2	Manufacturing	16.69%	16.69%	60.0%	5.0%
3	Transportation, Storage and Communications	3.49%	3.49%	5.0%	5.0%
4	Wholesale and Retail Trade	18.06%	18.06%	15.0%	15.0%
5	Construction & Mining	6.50%	6.50%	5.0%	5.0%
6	Others	26.69%	26.69%	10.0%	10.0%

Table 20, Country 1A vs. 2A, future workforce distribution heterogeneity

In terms of the policy, Country 1A's new workforce will be joining one of the industries with the highest productivity. However, in Country 2A, the new workforce will be joining the industry with the lowest GDP per capita but also has a very high growth. In the previous example, we observed that the effects of productivity growth translate into either an increase in contribution or investment return to the social security fund. The increase in contribution yields slightly longer lifetime for the fund. The combined effects of productivity growth and workforce distribution to sustainability can be complex to gauge. The gradual shift in the workforce to and from low and high productivity industry may exemplify or diminish the effects of policies employed in the previous demonstration.
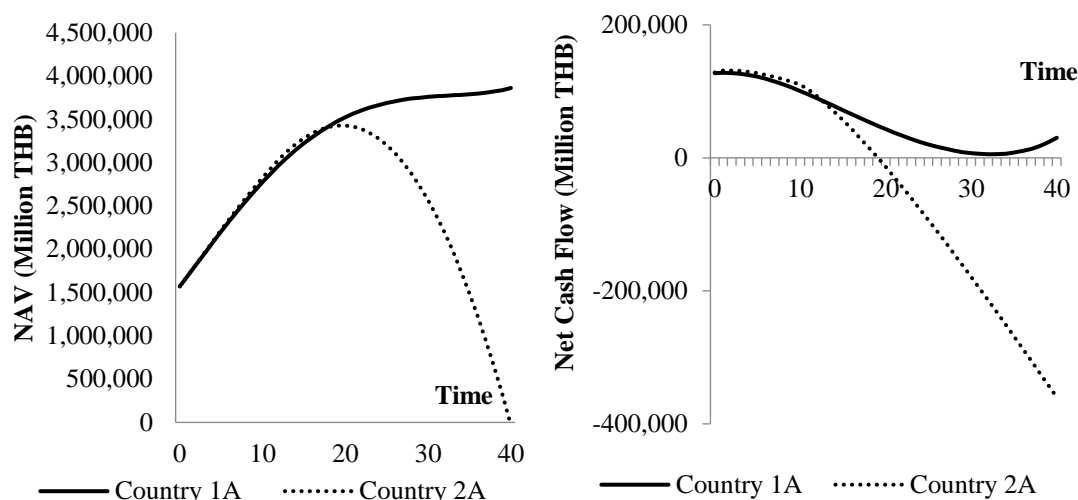
Figure 18, Country 1A vs. 2A, social security system sustainability and cashflow

The left plot in Figure 18 shows the progression of a fund's net asset value over time (years) for each country. The lifetime estimate of the social security system in Country 1A is over 60 years which is significantly longer than any of our hypothetical countries. As shown previously in Figure 15, the country's population and workforce structure reach a certain steady state after 60 years. In another word, we can say that the fund in Country 1A is self-sustained and will not perish unless big demographic shifts or economic recessions were to happen. On the other hand, the lifetime estimate for Country 2A is slightly lower than that of Country 2 at 40 years. The fund in Country 2A actually accumulated more wealth than the fund in Country 1A in the first 20 years but the end outcome is not as expected.

On a superficial level, one might think that increasing growth in the low-contribution sector will better improve sustainability as shown in the last section. However, the combined benefit of promoting growth and workforce entry into an industry with high GDP per capita that is also part of the market index yields the bigger sustainability benefit to the social security fund which is contradictory to the previous numerical example. This outcome resembles our discovery in Chapter 1. Many aspects of heterogeneity, when coupled, may produce non-monotonic effects. In order to understand the reasoning behind such phenomena, we need to look at the investment return and taxed contribution of each country.
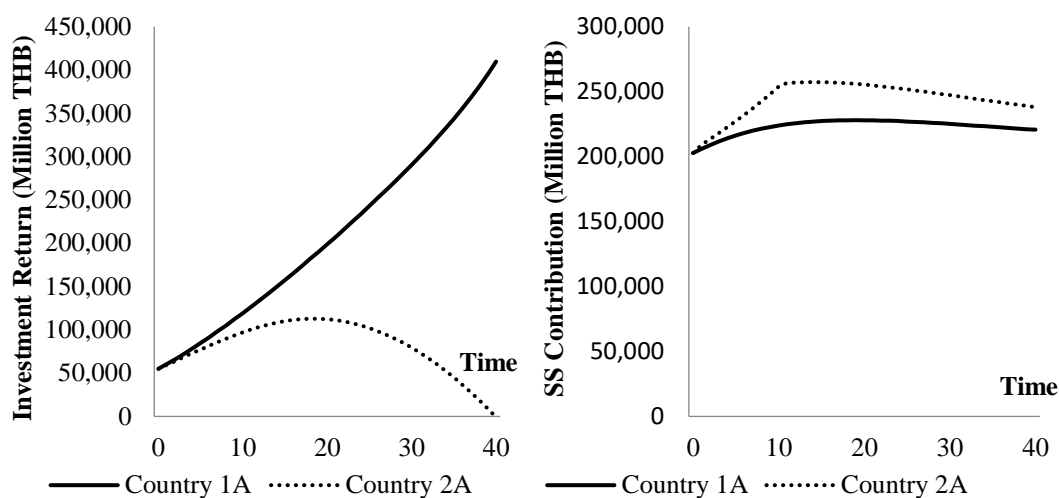
Figure 19, Country 1A vs. 2A, investment return and taxed contribution

Figure 19 shows the progression of investment return and taxed contribution over the course of the fund. Even though the fund in Country 2A accumulated more wealth in the first 20 years, the investment return in Country 1A is always more than that of Country 2A. The investment return in 1A is even superior to that in Country 1 in the previous analysis. As more and more workers move into the manufacturing industry, it generates higher GDP for the country which is in turn reflected in the investment return of the market asset. While investment return in Country 2A dropped as the fund depleted its asset base, the fund in Country 1A is able to generate the necessary return to offset the social security expenses.

On the other hand, consider the social security taxed contribution. The estimates are consistent with the previous example where the taxed contribution in Country 2A is always larger than its counterpart. But the difference is indifferent compared to the widening gap in investment returns, especially after 15 years. Country 2A has a shorter lifetime estimate than that of Country 2. This is because it takes a long time until the agricultural industry reaches the taxation cap. Besides, as the population birth and death are balanced, employing the workforce entry distribution as in 2A essentially move workers from high productivity industry to a lower one in agricultural. While the premature shift of workforce to a low productivity industry can shorten the lifetime estimate of the fund, shifting to the right industry can yield significant positivity to the sustainability of the fund. We will now look at the coupling effect of workforce distribution and growth in detail using a slight variation of set up of Country 1 and Country 2 in the previous section.

Industry (j)	Industry Distribution L_j	
	Country 1	Country 2
1 Agriculture	32.5−d%	32.5+d%
2 Manufacturing	32.5+d%	32.5−d%
3 Transportation, Storage and Communications	5.0%	5.0%
4 Wholesale and Retail Trade	15.0%	15.0%
5 Construction & Mining	5.0%	5.0%
6 Others	10.0%	10.0%

Table 21, Country 1 and 2 with allowance for variations in future workforce distribution

To understand and illustrate the coupling effects, let us consider a social security system in Country 1 and Country 2 with a choice of adjustment d to the future workforce industry distribution. This country specification is shown in Table 21 (Note that L_j cannot be negative, we will only consider $d \in [0, 0.325]$).

Figure 20, Social Security lifetime estimates of Country 1 vs. Country 2 SS: varying new labor force industry distribution d

Figure 20 display the variation of d and the lifetime estimates of Country 1 and Country 2 social security systems. Firstly, we see an opposing lifetime estimates trend from each country as more workers move into the focused-industry (Manufacturing in Country 1 and Agricultural in Country 2). In Country 1, the lifetime estimate increases as more workers move into the focused-industry. After a certain threshold ($d > 15\%$), we witness a sharp increase in the SS lifetime estimate. With the manufacturing productivity growth rate of 1.75% (2% Overall GDP growth) in Country 1, the country required 56.50% (32.50+24%) of its new workforce to join the manufacturing industry to make to social security system sustainable. On the other hand, for Country 2, the lifetime of the SS system declined as more workers move

into the agricultural industry. This can be attributed to several factors. Firstly, the premature shift of workforce to a low productivity industry before the productivity growth comes into full effect can shorten the lifetime. Furthermore, this shift can greatly deteriorate the investment return which as demonstrated previously is an essential component to achieve sustainability. The coupling of industry growth and workforce distribution is shown extensively in Figure 21. The figure display the lifetime estimates with variations of d and g_2 in Country 1. It can also be seen that the effects of both the workforce distribution and productivity growth on the lifetime estimate is exponential and non-linear.

Lifetime Estimate of Country 1 SS: Varying d and g
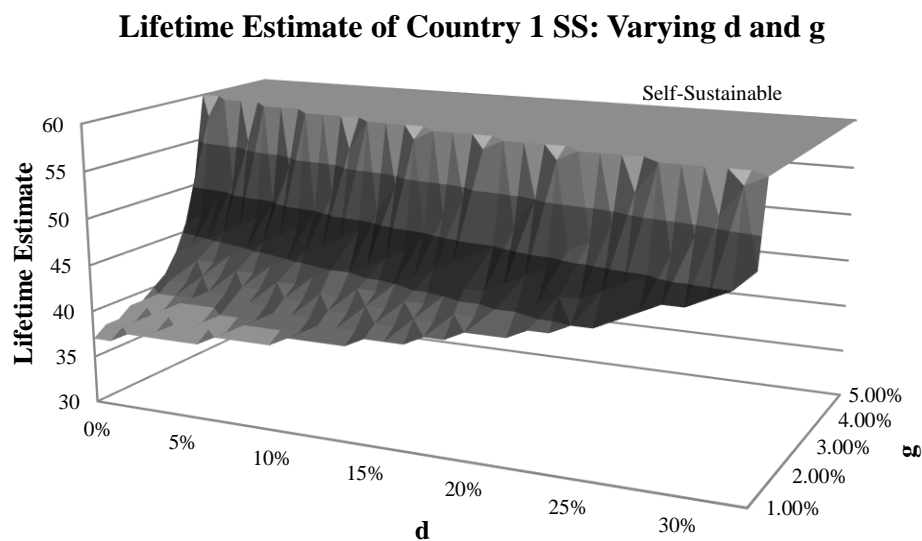
Figure 21, Lifetime Estimate of Country 1 SS: Varying d and g

These demonstrations are conditioned on constant productivity growth, unemployment, and workforce distribution which are all the desirables. Disturbances on the economic growths and unemployment will decay the sustainability of the social security fund. It is to be expected that shocks would affect any country unsymmetrically depending on its current heterogeneity demographic, productivity and labor distribution. Utilizing extreme reformation strategies requires thorough consideration since economic uncertainty can correspondingly lead to a shortfall. In order to achieve the desirable sustainability outcome, combinations of economic policy and reformation strategies must be carefully tailored subjected to the current demographic conditions of the country.

3.4 Chapter Conclusion

In this chapter, we looked at the effect of heterogeneity on a country-level scale and try to understand its effects on the sustainability of a country's social security system. Aspects of heterogeneity that were investigated are labor force entry industry distribution, industry-specific labor productivity growth, and unemployment.
Each of these aspects will affect the system's aggregate taxation credit, social security benefit withdrawal and social security investment return which has a direct effect on the system's overall sustainability. Our results show that combination between aspects of heterogeneity may produce non-monotonic outcome as observed in Chapter 1.

Our work tries to provide a unifying theory between choices of reformation schemes; increasing the social security taxation cap or investment capacity. It turns out that heterogeneity in workforce distribution plays an important role in determining appropriate policy/course-of-actions as illustrated in our numerical example.

Combination of promoting growth and workforce entry into an industry with high GDP per capita that is also part of the market index yields the biggest sustainability benefit to the social security fund. But on a side note, utilizing extreme reformation strategies requires thorough consideration. Clustering workforce in one or correlated industry can lead to a shortfall in the face of economic uncertainty.

Conclusion

In the first chapter, we propose a multi-member defined benefit pension fund model that allows for heterogeneity among plan participant. We have demonstrated how demographic differences such as the retirement date and salary of plan members can affect the ongoing management decisions.

We derive an analytical solution to a simplified setting of our model where only two members are present. The optimal asset allocation turns out to be a proportion to total discounted future liabilities over the current wealth of the fund. Additionally, the optimal funding support depends on the sum of the discounted liabilities less any accumulated wealth. Our model conforms to the traditional liability-driven investment framework.

Heterogeneity features prompt different discount rates on the expected liabilities from each fund member. As shown in Section 3, it is entirely possible for an expected liability to be heavily discounted, depending on the period of time and the nature of heterogeneity. These discount factors, in turn, determine the optimal decisions over the course of the fund.

To disregard or misrepresent the heterogeneity can lead to suboptimal decisions. Optimal investment paths from cases with and without heterogeneity considered can differ by a large margin towards each of the retirement. The ignorance of seemingly unimportant correlational heterogeneity can increase the expected funding variation by a significant amount.

In the second chapter, we propose an adaptive representative agent approach for our problem such that the distinctive features of members are taken into account. Our derivation suggested that each member contribution towards the formation of the representative agent should correspond to their respective fraction of heterogeneity-adjusted retirement liability to the total liability of the fund. For instance, a member with longer retirement time and negative salary-asset correlation may weigh more or less in the construction of the representative agent.

Our results also show that both salary growth- risky asset correlation and age have a strong influence on our solutions. A young member with negative correlation i.e. an unhedgable salary will contribute more towards the determination of both the optimal decisions and the representative agent. The performance comparison with the naive approach is promising. Our results show that our ARA can anticipate and handle the fund significantly better in unfavorable market conditions.

In the third chapter, we looked at the effect of heterogeneity on a country-level scale and try to understand its effects on the sustainability of a country's social security system. Aspects of heterogeneity that were investigated are labor force entry industry distribution, industry-specific labor productivity growth, and unemployment.

Each of these aspects will affect the system's aggregate taxation credit, social security benefit withdrawal and social security investment return which has a direct effect on the system's overall sustainability. Our results show that combination between aspects of heterogeneity may produce non-monotonic outcome as observed in Chapter 1.

We have shown that marginal modification to the growth of an appropriate industry could yield substantial results and deliver good improvement to the sustainability of the social security fund. Our work tries to provide a unifying theory between choices of reformation schemes; increasing the social security taxation cap or investment capacity. It turns out that heterogeneity in workforce distribution is the key in this model that plays an important role in determining appropriate policy/course-of-actions as illustrated in our numerical example.

Combination of promoting growth and workforce entry into an industry with high GDP per capita that is also part of the market index yields the biggest sustainability benefit to the social security fund. But on a side note, utilizing extreme reformation strategies requires thorough consideration. Clustering workforce in one or correlated industry can lead to a shortfall in the face of economic uncertainty.

Appendix I: Proof of Proposition 1

Proposition 1

Proof.

We will prove by induction that our optimal value function $V_t^*(\mathbf{Y}_t, W_t)$ as defined in (9) has a quadratic form in W_t and $\mathbf{Y}_t$ at all time t as follows

$$V_t^* = x_{t-1}\left(W_t^2 - \frac{2}{1+r}\sum_{i=1}^{N} B_{i,t}K_i(1+\mu_i)^{\tau_i-t}Y_{i,t}W_t + \sum_{i=1}^{N}\sum_{j=1}^{N} v_{ij,t-1}K_iK_jY_{i,t}Y_{j,t} \right). \qquad (23)$$

We will also find the coefficients x_{t-1} and $B_{i,t}$'s, which have direct effects on our optimal decisions. (As for $v's$, it will be shown later that it does not affect our optimal solution.) For ease of notation, we will not write the arguments $(\mathbf{Y}_t, W_t)$ to the value function V_t^*.

By definition, the value function at the last retirement is

$$V_{\tau_N}^* = \left(K_{\tau_N}Y_{N,\tau_N} - W_{\tau_N} \right)^2$$

which is quadratic in W_t and $\mathbf{Y}_t$ and correspond to the analytical form in (23) where $B_{N,\tau_N} = 1 + r$ and $x_{\tau_N-1} = 1$. To prove equation (23) by induction, we assume that V_{t+1}^* take the following form

$$V_{t+1}^* = x_t\left(W_{t+1}^2 - \frac{2}{1+r}\sum_{i=1}^{N} B_{i,t+1}K_i(1+\mu_i)^{\tau_i-t-1}Y_{i,t+1}W_{t+1} + \sum_{i=1}^{N}\sum_{j=1}^{N} v_{ij,t}K_iK_jY_{i,t+1}Y_{j,t+1} \right). \qquad (24)$$

We will use equation (11) to prove that V_t^* indeed has the quadratic form (23).

Define $\pi_{i,t}' := -K_i 1_{\{\tau_i=t\}} + \pi_i 1_{\{\tau_i>t\}}$. Let $\Gamma_t = \left[\Omega_t \quad \left(\sum_{i:\tau_i=t} X_i + \Upsilon_t\right)/x_t \right]^T$ where
$$\Omega_t := \sigma_r p_t W_t \quad \text{and} \quad \Upsilon_t := (1+r)W_t + \sum_{i=1}^{N} \pi_{i,t}' K_i Y_{i,t}.$$

Since $W_{t+1} = W_t\left[1 + r + p_t\sigma_r(\xi + Z_{r,t})\right] + \sum_{i=1}^{N} \pi_{i,t}'Y_{i,t} + \sum_{i:\tau_i=t} X_i$, we can easily re-write W_{t+1} in terms of Υ_t and Ω_t as $W_{t+1} = \left[\xi + Z_{r,t} \quad x_t\right]\Gamma_t$. This, combined with the fact that $Y_{i,t+1} = Y_{i,t}(1 + \mu_i + \sigma_i Z_{i,t} + \sigma_{r,i}Z_{r,t})$, yields

$$E_t[W_{t+1}^2] = \Gamma_t^T \begin{bmatrix} 1+\xi^2 & \xi x_t \\ \xi x_t & x_t^2 \end{bmatrix}\Gamma_t \quad , \quad E_t[Y_{i,t+1}W_{t+1}] = Y_{i,t}\Gamma_t^T \begin{bmatrix} \sigma_{r,i} + \xi(1+\mu_i) \\ x_t(1+\mu_i) \end{bmatrix}.$$

Let $\mathcal{M}_t := \frac{1}{1+r}\sum_{i=1}^{N} B_{i,t+1}K_i(1+\mu_i)^{\tau_i-t}Y_{i,t}$ and $\mathcal{S}_t := \frac{1}{1+r}\sum_{i=1}^{N} B_{i,t+1}K_i(1+\mu_i)^{\tau_i-t-1}Y_{i,t}\sigma_{r,i}$, it follows that

$$E_t\left[\frac{V_{t+1}^*}{x_t}\right] = \Psi_t - 2\Gamma_t^T \begin{bmatrix} \mathcal{S}_t + \xi\mathcal{M}_t \\ x_t\mathcal{M}_t \end{bmatrix} + \Gamma_t^T \begin{bmatrix} 1+\xi^2 & \xi x_t \\ \xi x_t & x_t^2 \end{bmatrix}\Gamma_t \qquad (25)$$

where $\Psi_t = \sum_{i=1}^{N}\sum_{j=1}^{N} v_{ij,t}K_iK_j E_t[Y_{i,t+1}Y_{j,t+1}]$. Notice that Ψ_t is in quadratic form of $Y_{i,t}$'s because $E_t[Y_{i,t+1}Y_{j,t+1}] = Y_{i,t}Y_{j,t}(1 + \mu_i + \mu_j + \mu_i\mu_j + \sigma_{r,i}^2 + \sigma_i^2 1_{\{i=j\}})$

From equation (11), we have

$$V_t^* = \min_{\Gamma_t} E_t[V_{t+1}^*] + \left(\sum_{i:\tau_i=t} X_i\right)^2. \qquad (26)$$

We will now solve the minimization problem for Γ_t^* (i.e. p_t^* and $\sum_{i:\tau_i=t} X_i^*$). The minimization problem can be divided into two instances. The first instance is where there is no retirement and the only decision variable is $\mathbf{p}_t$. The second instance is when there is a retirement in which there are two decisions variables in this case i.e. $\mathbf{p}_t$ and $\mathbf{X}$

Case 1: No retirement at time _t_ (i.e. $t \neq \tau_i \ \forall i, \ \sum_{i:\tau_i=t} X_i = 0$)

In this case the only decision variable is p_t. Hence, $\Gamma_t = [\ \Omega_t^* \quad \Upsilon_t / x_t\ ]^T$ and $\Omega_t^* = \sigma_r p_t^* W_t$

From equation (25), we have

$$E\left[\frac{V_{t+1}^*}{x_t}\right] = \Psi_t + \Upsilon_t^2 - 2\Upsilon_t \mathcal{M}_t - 2\xi(\mathcal{M}_t + \mathcal{S}_t/\xi - \Upsilon_t)\Omega_t + (1+\xi^2)\Omega_t^2.$$

Since the above expectation is quadratic in Ω_t, the closed form solution Ω_t^* that minimizes V_t^* can be easily obtained as follow

$$\Omega_t^* = \frac{\xi(\mathcal{M}_t + \mathcal{S}_t/\xi - \Upsilon_t)}{1+\xi^2} = \frac{\xi}{1+\xi^2}\left[\sum_{i=1}^{N} A_{i,t} K_i(1+\mu_i)^{\tau_i-t}Y_{i,t} - (1+r)W_t\right].$$

where $A_{i,t} = \frac{1}{1+r}B_{i,t+1}\big(1 + \sigma_{r,i}/\xi(1+\mu_i)\big) - \pi'_{i,t}/(1+\mu_i)^{\tau_i-t}$, which agrees with the equation (14) given in the proposition (because the α_t in (14) is one in this case). Substituting $\Omega_t^* = \sigma_r p_t^* W_t$, the optimal asset allocation p_t^* at any time t where there is no retirement is thus

$$p_t^* = \frac{\xi/\sigma_r}{1+\xi^2}\left[\sum_{i=1}^{N} A_{i,t}\frac{K_i(1+\mu_i)^{\tau_i-t}Y_{i,t}}{W_t} - 1 - r\right]. \tag{12A}$$

After we obtained Ω_t^*, then from (26) we have

$$\begin{aligned}
\frac{V_t^*}{x_t} &= \Psi_t - \mathcal{M}_t^2 + (\mathcal{M}_t - \Upsilon_t)^2 - \frac{\xi^2(\mathcal{M}_t + \mathcal{S}_t/\xi - \Upsilon_t)^2}{1+\xi^2} \\
&= \Psi_t - \mathcal{M}_t^2 - \mathcal{S}_t^2 + \frac{(\mathcal{M}_t - \xi\mathcal{S}_t - \Upsilon_t)^2 + 2\mathcal{S}_t^2}{1+\xi^2}
\end{aligned} \tag{27}$$

which is also quadratic in W_t and $\mathbf{Y}_t$. (Since $\mathcal{M}_t, \mathcal{S}_t$ and Υ_t are all linear in W_t and $\mathbf{Y}_t$, the product of these terms are quadratic in W_t and $\mathbf{Y}_t$.) We have shown that V_t^* has the same analytical form as per equation (23). The relationship between the coefficients $x_{t-1}, B_{i,t}$ and $x_t, B_{i,t+1}$ can be obtained from matching the coefficients of the interested terms in the above equation (27) and (23).

For instance, the only term with W_t^2 in V_t^* from equation (27) is Υ_t^2. From (27) and definition of Υ_t is $\frac{(1+r)^2}{1+\xi^2}$, The corresponding coefficient in front of W_t^2 is

$$x_{t-1} = x_t \frac{(1+r)^2}{1+\xi^2}.$$

Doing the same for $W_t Y_{i,t}$, we have

$$B_{i,t} = B_{i,t+1}\frac{1 - \xi\sigma_{r,i}/(1+\mu_i)}{1+r} - \frac{\pi'_{i,t}}{(1+\mu_i)^{\tau_i-t}}$$

which corresponds to equation (15),

Case 2: Retirement(s) at time t (i.e. $t = \tau_i \, \exists i$)

In this case the decision variables are p_t and X_is . Hence, $\Gamma_t^* = \begin{bmatrix} \Omega_t^* & \left(\sum_{i:\tau_i=t} X_i^* + \Upsilon_t\right)/x_t \end{bmatrix}^T$ and $\Omega_t^* = \sigma_r p_t^* W_t$.

Since $\left(\sum_{i:\tau_i=t} X_i\right)^2 = \Upsilon_t^2 - 2\Upsilon_t\left(\sum_{i:\tau_i=t} X_i + \Upsilon_t\right) + \left(\sum_{i:\tau_i=t} X_i + \Upsilon_t\right)^2$, at any time $t = \tau_i$ we will try to minimize the value function in equation (26) that is

$$\frac{\left(\sum_{i:\tau_i=t} X_i\right)^2 + E_t[V_{t+1}^*]}{x_t} = \Psi_t + \Upsilon_t^2/x_t - 2\Gamma_t^T \begin{bmatrix} \xi\mathcal{M}_t + \mathcal{S}_t \\ x_t\mathcal{M}_t + \Upsilon_t \end{bmatrix} + \Gamma_t^T \begin{bmatrix} 1 + \xi^2 & \xi x_t \\ \xi x_t & x_t + x_t^2 \end{bmatrix} \Gamma_t.$$

Since the above expectation is quadratic in Γ_t, the closed form solution that minimizes V_t^* can be obtained as follow

$$\Gamma_t^* = \begin{bmatrix} 1 + \xi^2 & \xi x_t \\ \xi x_t & x_t + x_t^2 \end{bmatrix}^{-1} \begin{bmatrix} \xi\mathcal{M}_t + \mathcal{S}_t \\ x_t\mathcal{M}_t + \Upsilon_t \end{bmatrix}$$

where

$$\begin{bmatrix} 1 + \xi^2 & \xi x_t \\ \xi x_t & x_t + x_t^2 \end{bmatrix}^{-1} = \frac{1}{1 + x_t + \xi^2} \begin{bmatrix} 1 + x_t & -\xi \\ -\xi & (1 + \xi^2)/x_t \end{bmatrix}$$

Therefore, the optimal solution to V_t^* is

$$(1 + x_t + \xi^2)\Gamma_t^* = \begin{bmatrix} 1 + x_t & -\xi \\ -\xi & (1 + \xi^2)/x_t \end{bmatrix}\begin{bmatrix} \xi\mathcal{M}_t + \mathcal{S}_t \\ x_t\mathcal{M}_t + \Upsilon_t \end{bmatrix}$$

$$(1 + x_t + \xi^2)\begin{bmatrix} \Omega_t^* \\ \left(\sum_{i:\tau_i=t} X_i^* + \Upsilon_t\right)/x_t \end{bmatrix} = \begin{bmatrix} \xi\mathcal{M}_t + (1 + x_t)\mathcal{S}_t - \xi\Upsilon_t \\ \mathcal{M}_t - \xi\mathcal{S}_t + (1 + \xi^2)\,\Upsilon_t/x_t \end{bmatrix}$$

$$(1 + x_t + \xi^2)\begin{bmatrix} \Omega_t^*/\xi \\ X_i^*/x_t \end{bmatrix} = \begin{bmatrix} \mathcal{M}_t + (1 + x_t)\,\mathcal{S}_t/\xi - \Upsilon_t \\ \mathcal{M}_t - \xi\mathcal{S}_t - \Upsilon_t \end{bmatrix} = \sum_{i=1}^{N} \begin{bmatrix} A_{i,t}K_i(1 + \mu_i)^{\tau_i-t}Y_{i,t} \\ B_{i,t}K_i(1 + \mu_i)^{\tau_i-t}Y_{i,t} \end{bmatrix} - W_t\begin{bmatrix} 1 + r \\ 1 + r \end{bmatrix}.$$

Where $B_{i,t}$ is the same as given in (15) and $A_{i,t} = \frac{1}{1+r}B_{i,t+1}\left(1 + \alpha_t\sigma_{r,i}/\xi(1 + \mu_i)\right) - \pi_{i,t}'/(1 + \mu_i)^{\tau_i-t}$ as shown in (14) (in this case $\alpha_t = 1 + x_t$). From the above equation, the optimal supplementary funding decisions X_i^* is

$$X_i^* = \frac{x_t}{1 + x_t + \xi^2}\left[\sum_{i=1}^{N} B_{i,t}K_i(1 + \mu_i)^{\tau_i-t}Y_{i,t} - (1 + r)W_t\right]$$

which proves equation (13).

Substituting $\Omega_t^* = \sigma_r p_t^* W_t$ to the above equation, we have

$$p_t^* = \frac{\xi/\sigma_r}{1 + x_t + \xi^2}\left[\sum_{i=1}^{N} A_{i,t}\frac{K_i(1 + \mu_i)^{\tau_i-t}Y_{i,t}}{W_t} - 1 - r\right] \tag{12B}$$

(12A) and (12B) prove equation (12)

The minimized value function is:

$$\frac{\sum_{i:\tau_i=t} X_i^* + E_t[V_{t+1}^*]}{x_t} = \Psi_t + \frac{\Upsilon_t^2}{x_t} - \frac{1}{1 + x_t + \xi^2}\begin{bmatrix} \xi\mathcal{M}_t + \mathcal{S}_t \\ x_t\mathcal{M}_t + \Upsilon_t \end{bmatrix} \cdot \begin{bmatrix} \xi\mathcal{M}_t + (1 + x_t)\mathcal{S}_t - \xi\Upsilon_t \\ \mathcal{M}_t - \xi\mathcal{S}_t + (1 + \xi^2)\,\Upsilon_t/x_t \end{bmatrix}$$

$$= \Psi_t - \mathcal{M}_t^2 - \mathcal{S}_t^2 + \frac{(\mathcal{M}_t - \xi\mathcal{S}_t - \Upsilon_t)^2}{1 + x_t + \xi^2}$$

which is again quadratic in W_t and $\mathbf{Y}_t$. We have proven by induction that $\forall t$, V_t^* is quadratic in W_t and $\mathbf{Y}_t$.

By grouping the terms with W_t^2 and $W_tY_{i,t}$ and compare the coefficient of each respective term with V_{t+1}^* defined previously. This implies that $B_{i,t}$ is indeed the same as given in (15), while

$$x_{t-1} = x_t\frac{(1 + r)^2}{1 + x_t + \xi^2}$$

We have thus proven, (12-15)

Appendix II: Parameters for Analysis and Simulation

Unless stated, we use the following set of base parameters in our analysis in Chapter 1 and Chapter 2.

Investment

$$W_0 = 0$$
$$r_f = 0.03$$
$$\sigma_r = 0.10$$
$$\xi = 0.5$$

Member 1

$$K_1 = 7$$
$$\sigma_1 = 0.10$$
$$\pi_1 = 0.10$$

Member 2

$$K_1 = 7$$
$$\sigma_2 = 0.10$$
$$\pi_2 = 0.10$$

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Appendix III: Proof of Proposition 2

Proposition 2
Proof.

Optimal investment decisions in a heterogeneity fund set up at $t \neq \tau_i$

$$p_t^* = \frac{\xi/\sigma_r}{1+\xi^2} \left[\sum_{i=1}^{N} A_{i,t} \frac{K_i (1+\mu_i)^{\tau_i-t} Y_{i,t}}{W_t} - 1 - r \right]$$

Since $A_{i,t} = B_{i,t+1} \frac{1+\sigma_{r,i}\xi^{-1}/(1+\mu_i)}{1+r} - \frac{\pi'_{i,t}/K_i}{(1+\mu_i)^{\tau_i-t}}$,

$$p_t^* = \frac{\xi/\sigma_r}{1+\xi^2} \left[\sum_{i=1}^{N} B_{i,t+1} \frac{(1+\mu_i+\sigma_{r,i}/\xi)(1+\mu_i)^{\tau_i-t-1}}{1+r} \frac{K_i Y_{i,t}}{W_t} - \sum_{i=1}^{N} \pi'_{i,t} \frac{Y_{i,t}}{W_t} - 1 - r \right]$$

We define $w_{i,t} := \frac{B_{i,t+1} K_i (1+\mu_i)^{\tau_i-t-1} Y_{i,t}}{B_{R,t+1} K_{R,t} (1+\mu_{R,t})^{\tau_{R,t}-t-1} Y_{R,t}}$ hence,

$$p_t^* = \frac{\xi/\sigma_r}{1+\xi^2} \left[\frac{B_{R,t+1}(1+\mu_{R,t})^{\tau_{R,t}-t-1}}{(1+r)} \frac{K_{R,t}Y_{R,t}}{W_t} \sum_{i=1}^{N}(1+\mu_i+\sigma_{r,i}\xi^{-1})w_{i,t} - \frac{Y_{R,t}}{W_t}\sum_{i=1}^{N}\pi'_{i,t}\frac{Y_{i,t}}{Y_{R,t}} - 1 - r \right]$$

Also by (22A – 22F),

$$p_t^* = \frac{\xi/\sigma_r}{1+\xi^2} \left[\frac{B_{R,t+1}(1+\mu_{R,t}+\sigma_{r,R,t}\xi^{-1})(1+\mu_{R,t})^{\tau_{R,t}-t-1}}{(1+r)} \frac{K_{R,t}Y_{R,t}}{W_t} - \pi'_{R,t}\frac{Y_{R,t}}{W_t} - 1 - r \right]$$

$$p_t^* = \frac{\xi/\sigma_r}{1+\xi^2} \left[A_{R,t} \frac{K_{R,t}(1+\mu_{R,t})^{\tau_{R,t}-t}Y_{R,t}}{W_t} - 1 - r \right]$$

Thus,

$$p_t^* = \hat{p}_{R,t}$$

Appendix IV: Parameters for Simple Average (SA) and Salary Weighted (WA) Agent

The characteristics of the representative agent under the simple average approach will simply be an average of all the remaining members of the fund as summarized as follow

$$Y_{R,t} = Y_{1,t} + Y_{2,t} + \cdots + Y_{N,t}$$
$$K_R = w_1 K_1 + w_2 K_2 + \cdots + w_N K_N$$
$$\mu_R = w_1 \mu_1 + w_2 \mu_2 + \cdots + w_N \mu_N$$
$$\sigma_{r,R} = w_1 \sigma_{r,1} + w_2 \sigma_{r,2} + \cdots + w_N \sigma_{r,N}$$
$$\pi_R = w_1 \pi_1 + w_2 \pi_2 + \cdots + w_N \pi_N$$
$$\tau_R = w_1 \tau_1 + w_2 \tau_2 + \cdots + w_N \tau_N$$

Where $w_{i,t} = \frac{1}{N}$

The characteristics of the representative agent under the salary weighted approach are summarized as follow

$$Y_{R,t} = Y_{1,t} + Y_{2,t} + \cdots + Y_{N,t}$$
$$K_R = w_1 K_1 + w_2 K_2 + \cdots + w_N K_N$$
$$\mu_R = w_1 \mu_1 + w_2 \mu_2 + \cdots + w_N \mu_N$$
$$\sigma_{r,R} = w_1 \sigma_{r,1} + w_2 \sigma_{r,2} + \cdots + w_N \sigma_{r,N}$$
$$\pi_R = w_1 \pi_1 + w_2 \pi_2 + \cdots + w_N \pi_N$$
$$\tau_R = w_1 \tau_1 + w_2 \tau_2 + \cdots + w_N \tau_N$$

Where $w_{i,t} = \frac{Y_{i,t}}{Y_{R,t}}$

Appendix V: Performance Comparison ARA vs. Salary Weighted (WA) Agent

Representative Agent Comparison

Representative Agent: Salary Growth ($\mu R,t$)

Representative Agent: $\sigma r,R,t$

Representative Agent: Time to Retirement ($\tau R,t-t$)

Performance Comparison
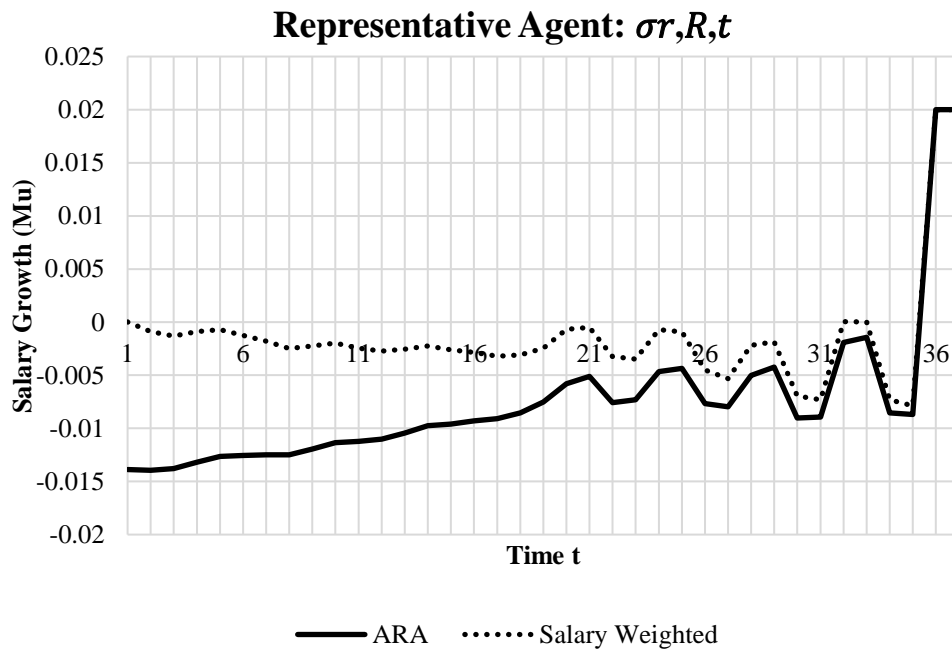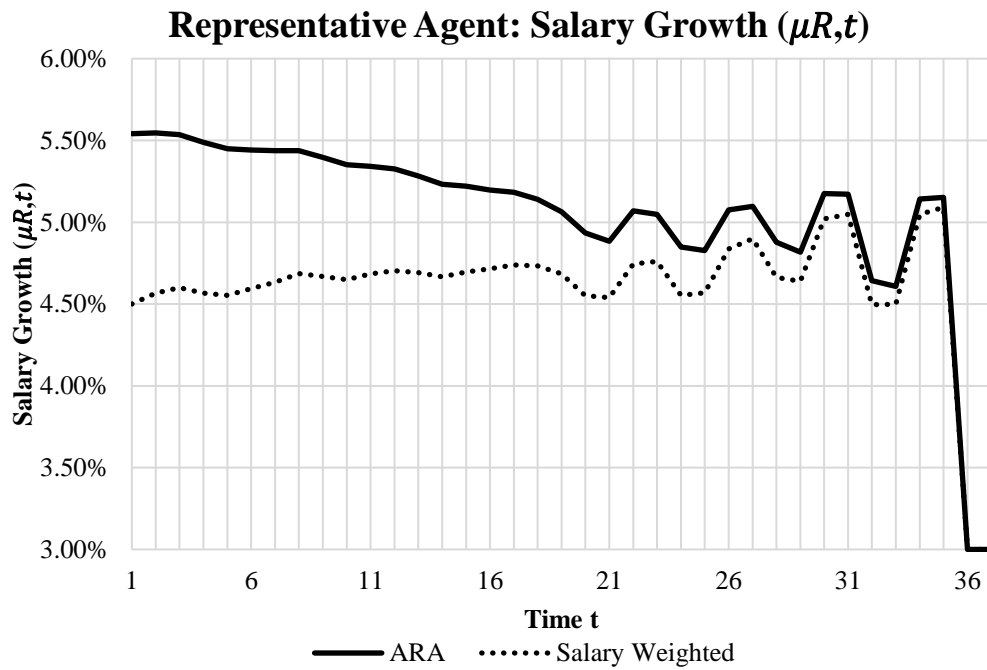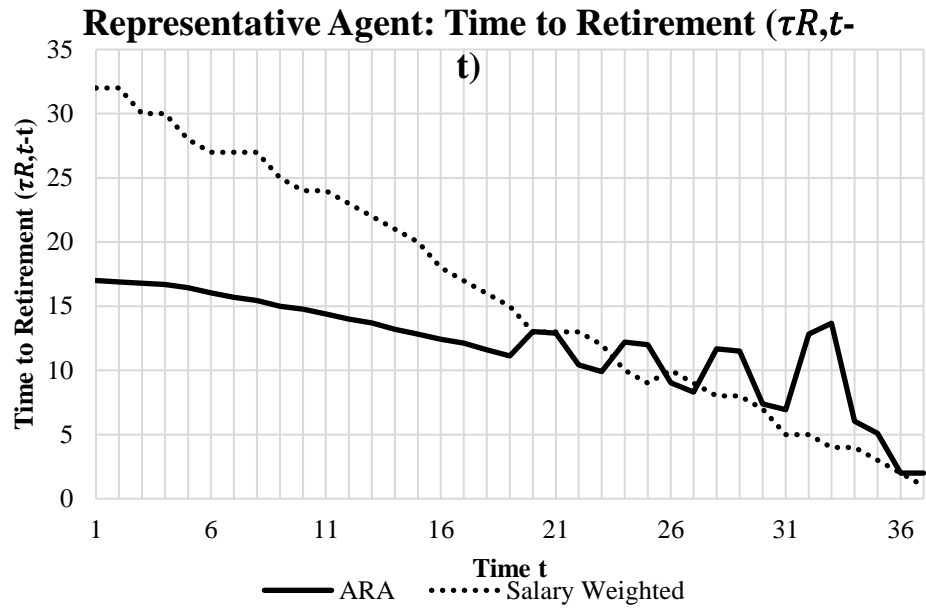
Total Supplementary Funding: ARA vs WA
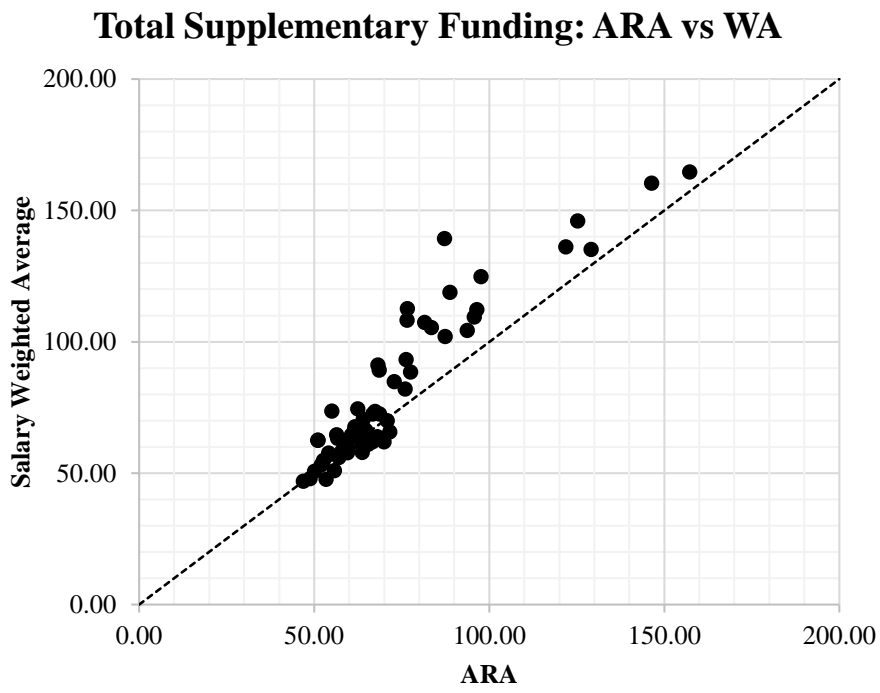
Decisions Comparison

Comparison of Average Asset Allocation Decision: ARA vs. WA

WA Band   ARA Band   WA Average Allocation   ARA Average Allocation
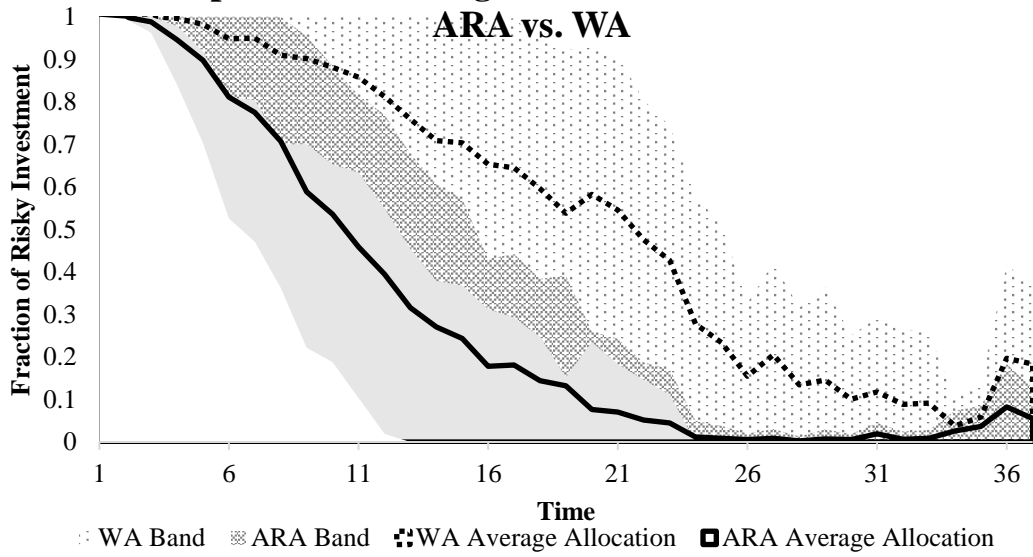
Comparison of Standard Deviation of Supplementary Funding Choices: ARA vs. WA

WA   ARA
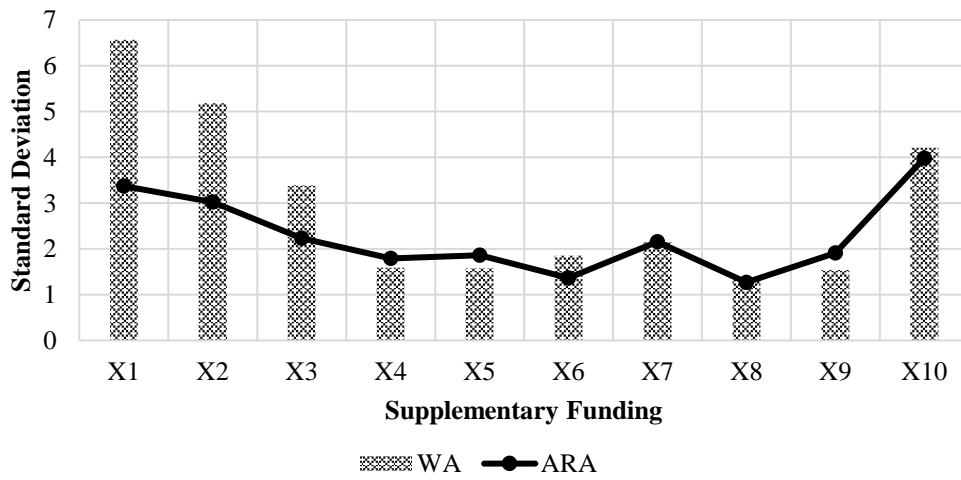
Appendix VI: Model Inputs for the Social Security Fund Model

Unless stated, we use the following set of base parameters in our analysis in Chapter 3
Initial Population

1, Under 15	$I_{1,0}$					

Labor Force	1	2	3	4	5	6
2, 15-29						
3, 30-39			$I_{ij,0}$			
4, 40-49						
5, 50-49						

6, Retired	$I_{6,0}$

1, Under 15	4,162,597					
Labor Force	1	2	3	4	5	6
2, 15-29	1,091,427	637,607	133,477	689,971	248,472	1,019,555
3, 30-39	1,227,386	717,034	150,104	775,921	279,424	1,146,561
4, 40-49	951,927	556,112	116,416	601,783	216,713	889,241
5, 50-49	579,169	338,348	70,830	366,135	131,852	541,030

6, Retired	565,006

Cohort Migration Matrix M

Age Cohort (i)	Start/End	1 Under 15	2 15-29	3 30-39	4 40-49	5 50-59	6 Retired	Death	M_i
0	Birth (of Total Population)	1.0%							1.0%
1	Under 15	93.3%	6.7%						6.7%
2	15-29		93.3%	6.7%					6.7%
3	30-39			90.0%	10.0%				10%
4	40-49				90.0%	10.0%			10%
5	50-59					90.0%	10.0%		10%
6	Retired						96.5%	3.5%	3.5%
	Death							100.0%	

Industry Entry Distribution L

	Industry (j)	Labor Force Industry Distribution L_j
1	Agriculture	28.6%
2	Manufacturing	16.7%
3	Transportation, Storage and Communications	3.5%
4	Wholesale and Retail Trade	18.1%
5	Construction & Mining	6.5%
6	Others	26.7%

Industry Productivity P

		Base GDP per Capita (THB) $P_{j,0}$	Taxed Contribution per person (THB) $\min(\pi P_{j,t}, \pi P_{cap})$
	GDP	297,269	
	GDP from the Insured (33% of total workforce)	297,269	
1	Agriculture	87,409	4,370
2	Manufacturing	698,238	9,000
3	Transportation, Storage and Communications	833,856	9,000
4	Wholesale and Retail Trade	220,569	9,000
5	Construction & Mining	196,545	9,000
6	Others	277,370	9,000

Unemployment $u = 2\%$

Social Security Expenses

$$\text{Unemployment } z = 90{,}000$$
$$\text{Retirement Benefit } K = 96{,}000$$
$$\text{Others } Z_2 = 50{,}000{,}000{,}000$$

Industry's fraction of GDP contribution to Market Index S

	Industry (j)	Contribution to SET50 Index $S_{j,0}$
1	Agriculture	0.00%
2	Manufacturing	17.82%
3	Transportation, Storage and Communications	2.68%
4	Wholesale and Retail Trade	23.82%
5	Construction & Mining	12.18%
6	Others	43.49%

Fund's starting NAV and Investment Parameters

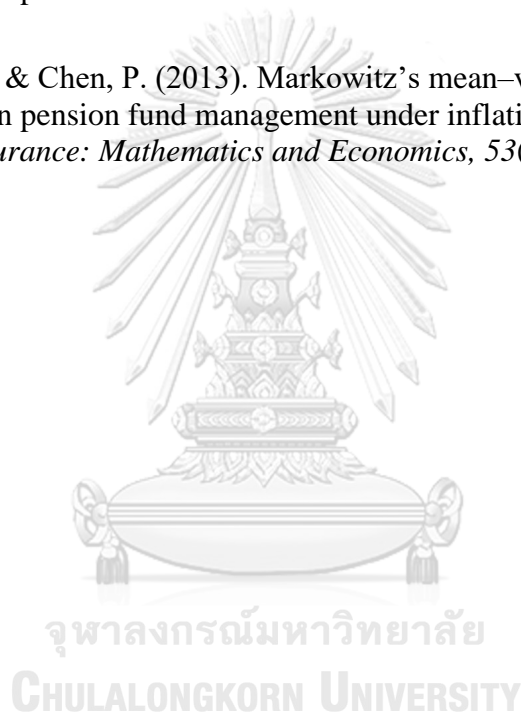|		
Current Wealth W_0	1,570,302	Million THB
Highly secured asset	83%	
Risky asset	17%	
Actual Investment return (2016)	3.36%	
Base line risky asset return, R_0	1.50%	
Base line risk free return r_f	1.50%	

Table of Figures

Table of Tables

REFERENCES

Aaron, H. J., & Reishauer, R. D. (1998). *Countdown to Reform*. New York: Century Foundation Press.

Ai, J., Brockett, P. L., & Jacobson, A. F. (2015). A new defined benefit pension risk measurement methodology. *Insurance: Mathematics and Economics, 63*, 40-51.

Attanasio, O. P., Kitao, S., & Violante, G. L. (2007). Global demographics trends and social security reform. *Journal of Monetary Economics, 54*(1), 144-198.

Auerbach, A. J., & Kotlikoff, L. J. (1987). *Dynamic Fiscal Policy*. Cambridge: Cambridge University Press.

Ayres, I., & Nalebuff, B. (2013). Diversification Across Time. *Journal of Portfolio Management, 39*(2), 73-86.

Basu, A., Byrne, A., & Drew, M. (2011). Dynamic Lifecycle Strategies for Target Date Retirement Funds. *The Journal of Portfolio Management, 37*(2), 83-96.

Bell, F. C., & Miller, M. L. (2005). *Life tables for the United States social security area 1900–2100*. Retrieved from Office of the Chief Actuary:

Benartzi, S., & Thaler, R. (2001). Na¨ıve diversification strategies in defined contribution savings plans. *American Economic Review, 91*, 79-98.

Benzoni, L., Collin-Dufresne, P., & Goldstein, R. (2007). Portfolio Choice over the Life-Cycle when the Stock and Labor Markets Are Cointegrated. *The Journal of Finance, LXII*(5), 2123-2168.

Blake, D., Cairns, A. J. G., & Dowd, K. (2001). Pensionmetrics I: stochastic pension plan design and value at risk during the accumulation phase. *Insurance: Mathematics and Economics, 29*(2), 187-215.

Bohn, H. (1999). Should the Social Security Trust Fund Hold Equities? An Intergenerational Welfare Analysis. *Review of Economic Dynamics, 2*, 666-697.

Borsch-Supan, A. (2003). Labor Market Effects of Population Aging. *Labour, Special Issue 17*, 5-44.

Boulier, J. F., Huang, S. J., & Taillard, G. (2001). Optimal management under stochastic interest rates: the case of a protected pension fund. *Insurance: Mathematics and Economics, 28*, 173-189.

Bovenberg, L., Koijen, R., Nijman, T., & Teulings, C. (2007). Savings and investing over the life cycle and the role of collective pension funds. *De Economist, 155*(4), 347-415.

Boyd, J. H., Hu, J., & Jagannathan, R. (2005). The Stock Market's Reaction to Unemployment News: Why Bad News Is Usually Good for Stocks. *The Journal of Finance, 60*(2), 649-672.

Cairns, A. J. G., Blake, D., & Dowd, K. (2006). Stochastic Lifestyling: Optimal Dynamic Asset Allocation for Defined Contribution Pension Plans. *Journal of Economic Dynamics and Control, 30*, 843-877.

Cheng, I. H., & French, E. (2000). The effect of the run-up in the stock market on labor supply. *ECONOMIC PERSPECTIVES-FEDERAL RESERVE BANK OF CHICAGO, 24*(4), 48-65.

Colombo, L., & Haberman, S. (2005). Optimal contributions in a defined benefit pension scheme with stochastic new entrants. *Insurance: Mathematics and Economics, 37*, 335-354.

Conesa, J. C., & Krueger, D. (1999). Social security with heterogeneous agents. *Review of Economic Dynamics, 2*(4), 757-795.

Cox, S. H., Lin, Y., Tian, T., & Yu, J. (2013). Managing Capital Market and Longevity Risks In A Defined Benefit Pension Plan. *The Journal of Risk and Insurance, 80*(3), 585-619.

Cremer, H., Gahvari, F., & Pestieau, P. (2008). Pensions with heterogenous individuals and endogenous fertility. *Journal of Population Economics, 21*, 961-981.

De Nardi, M., Imrohoroglu, S., & Sargent, T. J. (1999). Projected U.S. demographics and social security. *Review of Economic Dynamics, 2*, 575-615.

Deelstra, G., Grasselli, M., & Koehl, P. F. (2000). Optimal investment strategies in a CIR framework. *Journal of Applied Probability, 37*(936-946).

Deelstra, G., Grasselli, M., & Koehl, P. F. (2003). Optimal investment strategies in the presence of a minimum guarantee. *Insurance: Mathematics and Economics, 33*(1), 189-207.

Delong, L., Gerrard, R., & Haberman, S. (2008). Mean–variance optimization problems for an accumulation phase in a defined benefit plan. *Insurance: Mathematics and Economics, 42*, 107-118.

Dieci, R., & He, X. Z. (2018). Heterogeneous agent models in finance. *Handbook of Computational Economics, 4*.

Economist, T. (2018). The next Japan is not China but Thailand. *The Economist, Apr 5th 2018*.

Factbook, C. W. (2018). Thailand Demographics Profile.

Farrell, J., & Shoag, D. (2016). Asset management in public DB and non-DB Pension Plans. *Journal of Pension Economics and Finance, 15*(4), 379-406.

Goyal, A., & Wahal, W. (2008). The Selection and Termination of Investment Management Firms by Plan Sponsors. *The Journal of Finance, LXIII*(4), 1805-1846.

Haberman, S. (1997). Stochastic investment returns and contribution rate risk in a defined benefit pension scheme. *Insurance: Mathematics and Economics, 19*, 127-139.

Haberman, S., & Sung, J. H. (1994). Dynamics approaches to pension funding. *Insurance: Mathematics and Economics, 15*, 151-162.

Haberman, S., & Sung, J. H. (2005). Optimal Pension Funding Dynamics Over Infinite Control Horizon When Stochastic Rates of Return Are Stationary. *Insurance: Mathematics and Economics, 36*(1), 103-116.

Hainaut, D., & Devolder, P. (2007). Management of a pension fund under mortality and financial risks. *Insurance: Mathematics and Economics, 41*(1), 134-155.

Hommes, C. H. (2006). Heterogeneous Agent Models in Economics and Finance. In *Handbook of Computational Economics* (Vol. 2, pp. 1109-1186): Elsevier.

Honda, T. (2012). Dynamic Optimal Pension Fund Portfolios When Risk Preferences are Heterogeneous Among Pension Participants. *International Review of Finance, 12*(3), 329-355.

Huberman, G., & Wei, J. (2006). Offering versus choice in 401(k) plans: Equity exposure and number of funds. *Journal of Finance, 61*, 763-801.

Hull, J., & White, A. (2012). CVA and Wrong-Way Risk. *Financial Analysts Journal, 68*(5), 58-69.

Josa-Fombellida, R., & Rincon-Zapatero, J. P. (2004). Optimal Risk Management in Defined Benefit Stochastic Pension Funds. *Insurance: Mathematics and Economics, 34*(3), 489-503.

Josa-Fombellida, R., & Rincon-Zapatero, J. P. (2006). Optimal investment decisions with a liability: The case of defined benefit pension plans. *Insurance: Mathematics and Economics, 39*(1), 81-98.

Kazemi, H., Black, K. H., & Chambers, D. R. (2016). Chapter 4, Pension Fund Portfolio Management. In *Alternative Investments, CAIA Level II, 3rd Ed.*: CAIA Association.

Kitao, S. (2014). Sustainable social security: Four options. *Review of Economic Dynamics, 17*(4), 756-779.

Kotlikoff, L. J., Smetters, K. A., & Walliser, J. (2002). *Finding a way out of America's demographic dilemma*. The National Bureau of Economic Research.

Laun, T., & Wallenius, J. (2015). A life cycle model of health and retirement: The case of Swedish pension reform. *Journal of Public Economics, 127*, 127-136.

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance, 7*, 77-99.

Martellini, L., & Milhau, V. (2010). *From Deterministic to Stochastic Life-Cycle Investing: Implications for the Design of Improved Forms of Target Date Funds*. Retrieved from

Merton, R. C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory, 3*, 373-413.

Oshio, T. (2004). Social security and trust fund management. *Journal of the Japanese and International Economies, 18*(4), 528-550.

The Preventable Demise of Defined Benefit Pension Provision. (2017). Retrieved from

Sialm, C., Starks, L. T., & Zhang, H. (2015). Defined Contribution Pension Plans: Sticky or Discerning Money? *The Journal of Finance, 70*(2), 805-838. doi:10.1111/jofi.12232

Thailand Social Security Office, M. o. L. (2016). *Social Security Office Annual Report*. Retrieved from

Vigna, E. (2014). On efficiency of mean–variance based portfolio selection in denfined contribution pension schemes. *Quantitative Finance, 14*, 237-258.

Vigna, E., & Haberman, S. (2001). Optimal investment strategy for defined contribution pension schemes. *Insurance: Mathematics and Economics, 28*(2), 233-262.

Yao, H., Yang, Z., & Chen, P. (2013). Markowitz's mean–variance defined contribution pension fund management under inflation: A continuous-time model. *Insurance: Mathematics and Economics, 53*(3), 851-863.

APPENDIX

VITA

Thepdanai Danswasvong