

# **Risk Management with Basis Risk**

by

Jingong Zhang

A thesis  
presented to the University of Waterloo  
in fulfillment of the  
thesis requirement for the degree of  
Doctor of Philosophy  
in  
Actuarial Science

Waterloo, Ontario, Canada, 2018

© Jingong Zhang 2018

## **Examining Committee Membership**

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Rogemar S. Mamon  
Professor, Department of Statistical and Actuarial Sciences,  
Western University

Supervisor(s): Ken Seng Tan  
Professor, Department of Statistics and Actuarial Science,  
University of Waterloo  
Chengguo Weng  
Associate Professor, Department of Statistics and Actuarial Science,  
University of Waterloo

Internal Member(s): Adam Kolkiewicz  
Associate Professor, Department of Statistics and Actuarial Science,  
University of Waterloo  
Alexander Schied  
Professor, Department of Statistics and Actuarial Science,  
University of Waterloo

Internal-External Member: Peter A. Forsyth  
Distinguished Professor Emeritus, Cheriton School of Computer Science,  
University of Waterloo

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Jingong Zhang

## Abstract

Basis risk occurs naturally in a variety of financial and actuarial applications, and it introduces additional complexity to the risk management problems. Current literature on quantifying and managing basis risk is still quite limited, and one class of important questions that remains open is how to conduct effective risk mitigation when basis risk is involved and perfect hedging is either impossible or too expensive. The theme of this thesis is to study risk management problems in the presence of basis risk under three settings: 1) hedging equity-linked financial derivatives; 2) hedging longevity risk; and 3) index insurance design.

First we consider the problem of hedging a vanilla European option using a liquidly traded asset which is not the underlying asset but correlates to the underlying and we investigate an optimal construction of hedging portfolio involving such an asset. The mean-variance criterion is adopted to evaluate the hedging performance, and a subgame Nash equilibrium is used to define the optimal solution. The problem is solved by resorting to a dynamic programming procedure and a change-of-measure technique. A closed-form optimal control process is obtained under a general diffusion model. The solution we obtain is highly tractable and to the best of our knowledge, this is the first time the analytical solution exists for dynamic hedging of general vanilla European options with basis risk under the mean-variance criterion. Examples on hedging European call options are presented to foster the feasibility and importance of our optimal hedging strategy in the presence of basis risk.

We then explore the problem of optimal dynamic longevity hedge. From a pension plan sponsor's perspective, we study dynamic hedging strategies for longevity risk using standardized securities in a discrete-time setting. The hedging securities are linked to a population which may differ from the underlying population of the pension plan, and thus basis risk arises. Drawing from the technique of dynamic programming, we develop a framework which allows us to obtain analytical optimal dynamic hedging strategies to achieve the minimum variance of hedging error. For the first time in the literature, analytical optimal solutions are obtained for such a hedging problem. The most striking advantage of the method lies in its flexibility. While  $q$ -forwards are considered in the specific implementation in the paper, our method is readily applicable to other securities such as longevity swaps. Further, our method is implementable for a variety of

longevity models including Lee-Carter, Cairns-Blake-Dowd (CBD) and their variants. Extensive numerical experiments show that our hedging method significantly outperforms the standard “delta” hedging strategy which is commonly adopted in the literature.

Lastly we study the problem of optimal index insurance design under an expected utility maximization framework. For general utility functions, we formally prove the existence and uniqueness of optimal contract, and develop an effective numerical procedure to calculate the optimal solution. For exponential utility and quadratic utility functions, we obtain analytical expression of the optimal indemnity function. Our results show that the indemnity can be a highly non-linear and even non-monotonic function of the index variable in order to align with the actuarial loss variable so as to achieve the best reduction in basis risk. Due to the generality of model setup, our proposed method is readily applicable to a variety of insurance applications including index-linked mortality securities, weather index agriculture insurance and index-based catastrophe insurance. Our method is illustrated by a numerical example where weather index insurance is designed for protection against the adverse rice yield using temperature and precipitation as the underlying indices. Numerical results show that our optimal index insurance significantly outperforms linear-type index insurance contracts in terms of reducing basis risk.

## **Acknowledgements**

I would like to express my deepest gratitude to my supervisors, Professor Ken Seng Tan and Professor Chengguo Weng. Their profound knowledge, cautious attitude, inspiring guidance and continuous support turn me from a rookie who barely knew where to start into a self-proclaimed independent researcher. The freedom I am given to follow my research curiosity and interests makes the beginning of my academic career so colorful and enjoyable.

I would also like to thank all the committee members: Professor Adam Kolkiewicz, Professor Alexander Schied, Professor Peter Forsyth and Professor Rogemar Mamon for all those insightful suggestions and comments, which lead to significant improvements in both the content and the presentation of this thesis.

I acknowledge financial support from the Department of Statistics and Actuarial Science, University of Waterloo, and the Society of Actuaries Hickman Scholarship, which helps me a lot to relieve my financial stress and better focus on my research during my study life.

My friends and classmates also provide me lots of contributory ideas and opinions during my research process, and I truly appreciate all the help.

Finally, thank you to my parents and family for giving me unconditional support, without which this thesis would not even be possible in the first place. Special thanks to my wife Wenjun: meeting you is the most beautiful thing that happens in Waterloo and in my life.

## **Dedication**

This is dedicated to my family.

# Table of Contents

<b>List of Tables</b>	<b>xii</b>
<b>List of Figures</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background . . . . .	1
1.2 Objectives and outline . . . . .	5
<b>2 Optimal Hedging with Basis Risk under Mean-Variance Criterion</b>	<b>7</b>
2.1 Introduction . . . . .	7
2.2 Formulation of the optimal hedging problem . . . . .	11
2.3 Optimal time consistent hedging strategy . . . . .	16
2.3.1 The extended HJB equation . . . . .	17
2.3.2 Verification theorem . . . . .	18
2.3.3 Equilibrium solution . . . . .	21
2.4 Discussions . . . . .	29
2.4.1 The case with no basis risk . . . . .	30
2.4.2 The limiting case when $\gamma \rightarrow \infty$ . . . . .	31

2.4.3	The limiting case when $\gamma \rightarrow \infty$ with no basis risk	32
2.4.4	Solutions under geometric Brownian motions	34
2.5	Numerical examples	39
2.6	Conclusion	43
<b>3</b>	<b>Optimal Dynamic Longevity Hedge with Basis Risk</b>	<b>44</b>
3.1	Introduction	44
3.2	Problem setup	47
3.2.1	Stochastic mortality model	47
3.2.2	Pension liability	48
3.2.3	Dynamic hedge with q-forward	49
3.2.4	Hedging objective	52
3.3	Derivation of the optimal solution	55
3.3.1	The Bellman equation	55
3.3.2	Solution of the Bellman equation	57
3.4	Hedging Canadian mortality rates with UK mortality rates	60
3.4.1	Data and assumptions	61
3.4.2	Approximation to forward mortality rates	62
3.4.3	Benchmark method	63
3.4.4	Baseline result	64
3.4.5	Robustness to q-forwards' time to maturity	66
3.4.6	Robustness to q-forwards' reference age	68
3.4.7	Robustness to model risk	69
3.5	Conclusion	73

<b>4</b>	<b>Index Insurance Design</b>	<b>75</b>
4.1	Introduction . . . . .	75
4.2	Problem Setup . . . . .	79
4.3	Existence and Uniqueness of the Optimal Design . . . . .	81
4.3.1	Uniqueness of the optimal solution . . . . .	81
4.3.2	Existence of optimal solution . . . . .	82
4.4	Computing the Optimal Solution . . . . .	87
4.4.1	The ODE method . . . . .	88
4.4.2	Quadratic utility . . . . .	92
4.4.3	Exponential utility . . . . .	94
4.5	Applications in Weather Index Insurance Design . . . . .	95
4.5.1	Dependence modeling . . . . .	96
4.5.2	Quadratic utility . . . . .	97
4.5.3	Exponential utility . . . . .	101
4.5.4	Logarithmic utility . . . . .	102
4.5.5	Comparison with linear contracts . . . . .	103
4.5.6	An example of bivariate-index insurance . . . . .	105
4.6	Conclusion . . . . .	108
<b>5</b>	<b>Conclusion and Future Research</b>	<b>109</b>
	<b>References</b>	<b>112</b>
	<b>APPENDICES</b>	<b>120</b>

<b>A</b>	<b>Appendix for Chapter 2</b>	<b>121</b>
A.1	Proof of Proposition 2.1 . . . . .	121
A.2	Equilibrium value function for futures contract . . . . .	125
A.3	Equilibrium value function for European call options . . . . .	126
<b>B</b>	<b>Appendix for Chapter 3</b>	<b>127</b>
B.1	Approximations for forward survival rates in Section 3.4.2 . . . . .	127
B.2	Approximations for forward survival rates under CBD model in Section 3.4.7 . .	128
<b>C</b>	<b>Appendix for Chapter 4</b>	<b>131</b>
C.1	A Numerical Scheme for Solving ODE (4.15) . . . . .	131

# List of Tables

2.1	Parameter values for the hedging examples . . . . .	39
2.2	Comparison of different strategies on hedging European call option ( $\gamma = 1$ ) . . . . .	41
2.3	Hedging European call with $\rho = 1$ : “equilibrium solution” vs “naive delta hedge”. . . . .	42
2.4	Hedging European call with $\rho = 0.9$ : “equilibrium solution” vs “naive delta hedge”. . . . .	42
3.1	Results of Five Experiments . . . . .	65
3.2	HE results with different “true” and “assumption” models . . . . .	71

# List of Figures

3.1	HE with different time to maturity . . . . .	67
3.2	HE with different reference ages . . . . .	69
3.3	Change with time to maturity . . . . .	73
4.1	Joint density $f(x, y)$ of the actual loss and the average temperature. . . . .	96
4.2	Optimal indemnity functions for different premium levels. . . . .	98
4.3	Optimal indemnity for different maximum indemnity levels. . . . .	99
4.4	Basis risk (i.e., the standard deviation of residual risk) for different levels of premium and maximum indemnity. . . . .	100
4.5	Optimal indemnity under exponential utility. . . . .	101
4.6	Optimal indemnity under logarithmic utility. . . . .	103
4.7	Comparison of optimal indemnity functions under three different utilities. . . . .	104
4.8	Effectiveness: our optimal index insurance vs. the linear-type contract. . . . .	105
4.9	Indemnity function of a bivariate-index insurance: temperature and precipitation .	106
4.10	Effectiveness improvement of additional index. Basis risk is measured by standard deviation of residual risk. . . . .	107

# Chapter 1

## Introduction

### 1.1 Background

Broadly speaking, basis risk is defined as the non-hedgeable portion of risk as attributed to the imperfect correlation between the asset to be hedged and the asset used for hedging. It implies that the hedging will be imperfect and the hedged position still carries some residual risk. The existence of basis risk introduces additional complication for risk management, and it may have detrimental effects when it is overlooked. Therefore it is important for risk managers to pay attention to the identification, assessment, and management of basis risk.

Basis risk naturally occurs in a variety of financial and actuarial problems. A typical example in the financial market is that the hedging for a derivative written on a non-tradable asset is often conducted via trading over one liquidly traded asset which is closely correlated with the non-tradable underlying asset. A relevant concept in finance and economics is the incomplete market in which the number of Arrow-Debreu securities is less than the number of states of nature, and thus some payoffs in the market cannot be replicated by tradable securities in that market. In this case, the hedger would normally aim at designing a hedging scheme with other tradable assets to minimize the negative impact resulted from the mismatch between these assets and therefore achieve certain financial objectives.

The second example where basis risk plays a critical role is longevity risk management. In recent decades longevity risk has become one of the most dominant risks, particularly for pension plan sponsors and annuities providers. Longevity risk is any risk associated with the unpredicted increasing life expectancy, and it will eventually translate into higher than expected payout ratios for pension plans and insurance companies. Hedging of longevity risk is an important problem because the risk is non-diversifiable. As such some longevity securities such as longevity bonds and longevity swaps have been issued in an attempt to address longevity risk. However, these longevity securities have payoffs that are linked to some standardized populations while the pension plan sponsors and annuity providers are more concerned with the longevity experience underlying the pensioners and annuitants. These experiences are not perfectly related to the experience of the standardized population underlies the longevity securities and hence population basis risk arises. Li and Hardy (2011) give an empirical example to hedge a pension plan on the female Canadian population with a longevity index on the U.S. population. When measured by Value-at-Risk, the magnitude of reduction in basis risk can be as high as 10.14%.

Another area where basis risk is prevalent is the agricultural index insurance where the basis risk is cited as a primary concern in agricultural risk management (Brockett et al., 2005). Index insurance is an innovative approach to insurance provision which pays indemnity determined by certain relevant indices rather than the actual losses experienced by policyholders, and a typical example of using index insurance is in the area of agricultural insurance. Agricultural production is very vulnerable to weather risks and hence providing insurance protection to farmers is important. Traditionally agricultural insurance is indemnity based, i.e. farmers will be reimbursed based on the incurred losses. This product, however, is known to generate moral hazard with high underwriting cost, especially offering to regions with numerous small farms. Alternatively, weather index insurance is gaining popularity to hedge against agricultural production loss. Plausible indices include temperature, precipitation, sunshine as well as those remote sensing indices based on satellites images such as the Normalized Difference Vegetation Index (NDVI). In this context, farmers are similarly exposed to basis risk due to the imperfect correlation between the weather index and the crop yield.

This thesis focuses on three risk management problems, and the evolving theme in these applications is the basis risk that is present in derivative hedging, longevity risk hedging and

agricultural index insurance. It should be noted that there are other interesting risk management problems involving basis risk. Some notable examples include hedging of equity-linked insurance products such as variable annuities (Ng and Li, 2013) and catastrophic risk management (Doherty, 1997; Li and Yu, 2002; Cummins et al., 2004).

Basis risk generally comes from the mismatch between the hedging objective and the hedging instrument. More specifically, this mismatch may come in different forms and typical examples include: 1) hedging objective and hedging instrument are totally different type of assets, e.g., one is agricultural production and the other is a weather index; 2) hedging objective and hedging instrument are the same type but with different underlying assets, time to maturity and so on; 3) hedging objective and hedging instrument are based on exactly the same asset, but the way to construct the hedging portfolio makes the hedging imperfect due to budget constraint, trading frequency, and etc. Mathematically, we usually define this mismatch between the hedging objective and the hedging instrument by assuming that two random variables are not equal almost surely<sup>1</sup> in static settings, or two stochastic processes are not indistinguishable<sup>2</sup> in dynamic settings.

The study of basis risk contains two major aspects: 1) how to model the dependence structure between two random variables or stochastic processes and how to measure the difference between them in order to quantify basis risk; 2) based on a given dependence structure between the hedging objective and the hedging instrument, how to construct a hedging portfolio or a risk mitigation strategy in order to achieve the hedger's certain financial objective. The former, the modeling problem, itself is a very broad area, and under each specific setting the way to model the dependence structure can be very different. For example, in the literature of longevity risk, one common way to model the dependence between multiple populations is to introduce an additional common factor to represent their co-movement into the original model (such as Lee Carter model, Cairns-Blake-Dowd (CBD) model and etc.). This is quite different from those commonly used dependence modeling techniques such as copulas, and is not that commonly used in the other contexts of risk management. In this thesis we focus on the second aspect, and explicitly exploit the current existing models as our starting point. In other words, we study these

---

<sup>1</sup>Two random variables  $X$  and  $Y$  are defined as equal almost surely if  $\mathbb{P}(X = Y) = 1$ .

<sup>2</sup>Two stochastic processes  $\{X_t\}_{t \geq 0}$  and  $\{Y_t\}_{t \geq 0}$  are defined to be indistinguishable if  $\mathbb{P}(\omega : X(t, \omega) = Y(t, \omega), \forall t \geq 0) = 1$ .

risk management problems in the presence of basis risk by assuming the underlying models and dependence structures are given exogenously.

In the mathematical finance literature, there are many papers that discuss the problem of hedging contingent payoffs in an incomplete market setting. Since the contingent claim of interest cannot generally be replicated by any hedging strategies which are self-financing and perfectly match the target payoff function at the same time, there are two major approaches of constructing the hedge: first, the hedger may relax the self-financing constraint imposed on the hedging strategy by continuously injecting money into the hedging plan, and only requires it to perfectly match the payoff function of the contingent claim to be hedged; second, the hedger still sticks to self-financing hedging plans and strives to minimize the “distance” between the hedging objective and the hedging portfolio. The first method was studied by Föllmer and Schweizer (1991) and Schweizer (1999), and they introduced the concept of locally risk-minimizing strategy to minimize the expected cumulative cost under a quadratic criterion, and obtained the hedging strategy through the Föllmer-Schweizer decomposition. Some further research was conducted by Møller (1998, 2001) with applications for hedging insurance payoffs such as unit-linked life insurance contracts. The second method was discussed by Duffie and Richardson (1991) under the mean-variance criterion, and by Davis (2006) and Musiela and Zariphopoulou (2004) under an exponential utility maximization framework.

In the insurance and actuarial science literature, although there are articles addressing basis risk for various risk management problems, such as longevity hedge with standardized mortality-linked securities (Li and Hardy, 2011; Coughlan et al., 2011; Cairns, 2013; Zhou and Li, 2016), agricultural production protection with weather derivatives and weather insurance (Woodard and Garcia, 2008; Brockett et al., 2005; Chantarat et al., 2007; Jensen, et al., 2016) and catastrophe risk mitigation by catastrophe securities (Lee and Yu, 2002; Cummins et al., 2004), most of them are empirical studies concentrating on very specific applications.

## 1.2 Objectives and outline

The thesis aims at developing innovative methodologies by focusing on three risk management problems in the presence of basis risk, namely, financial derivative hedging, longevity risk hedging and index insurance design, in three separate chapters. Each chapter begins with formulating the problem in an optimization framework, proceeds with mathematical derivation to solve the optimization problem, and finishes with numerical examples showing the applicability and superiority of our proposed solutions. The rest of this thesis is organized as follows.

Chapter 2 studies the hedging problem for general European-style financial derivatives whose underlying assets are not traded in the market. Therefore we use another correlated and liquidly traded asset as the hedging instrument. We adopt the mean-variance criterion to evaluate the hedging performance, and use a subgame Nash equilibrium to define the optimal solution to overcome the “time-inconsistent” issue which arises inherently from the mean-variance criterion. The problem is solved by resorting to a dynamic programming procedure and a change-of-measure technique. Numerical examples for hedging futures and European call options are presented to showcase the performance of our proposed optimal strategy.

Chapter 3 investigates the problem of longevity hedge using standardized longevity securities. We study dynamic hedging strategies for longevity risk from a pension plan sponsor’s perspective in a discrete-time setting. Our assumptions about the underlying stochastic mortality models are quite general so that our results can be applied to those most popularly used longevity models such as Lee-Carter, Cairns-Blake-Dowd (CBD) and their variants. The hedging instruments are  $q$ -forward contracts that are linked to a population different from the pension plan’s underlying population, so basis risk arises due to such a population mismatch. We apply the dynamic programming technique to develop a framework which allows us to obtain analytical optimal dynamic hedging strategies to achieve the minimum variance of hedging error. Extensive numerical experiments show that our hedging method significantly outperforms the dynamic “delta” hedging strategy which is commonly adopted in the literature.

Chapter 4 studies the problem of index insurance design and its applications in agricultural weather index insurance. The goal is to design an optimal index insurance contract which max-

imizes the expected utility of policyholders. For a general strictly concave utility function, the optimal solution is characterized by an implicit ordinary differential equation (ODE) problem. Our results show that the optimal indemnity can be a highly non-linear and even non-monotonic function of the index variable in order to align with the actual loss variable so as to achieve the best reduction in basis risk. Our theoretical results are illustrated by a numerical example where weather index insurance is designed for protection against the adverse rice yield using temperature and precipitation as the underlying indices. Numerical results show that our optimal index insurance significantly outperforms those linear-type index insurance contracts, which are commonly adopted in the literature and insurance practice, in terms of reducing basis risk.

Chapter 5 concludes the thesis and discusses some potential future works. Some additional information complementing each chapter is collected in appendices.

## **Chapter 2**

# **Optimal Hedging with Basis Risk under Mean-Variance Criterion**

### **2.1 Introduction**

It is well-known in the financial theory that when an option is written on an asset that is tradable, it can be hedged by trading in the underlying asset. What if an option is written on an asset that is either illiquid or even non-tradable? In this case, a common hedging practice is to use another asset that is tradable, highly liquid, and also has the desirable property of being highly correlated to the underlying asset of the option. Because the hedged asset does not perfectly capture the behavior of the underlying asset, there is a mismatch between the risk exposure of the hedged portfolio and the option in question; this gives rise to the so-called basis risk. As shown in Davis (2006), the basis risk could be huge even though both assets have very high correlation. This implies that the basis risk can have a detrimental effect on the hedging performance and hence it needs to be prudently managed.

Basis risk does not just confine to hedging financial derivatives, it exists in many other settings, notably when an index-based security is used for hedging. For example, a pension plan sponsor may choose to hedge the plan's longevity risk by resorting to standard longevity instrument that is traded in the capital market. While such "standard" instrument provides liquidity

and transparency, its payoffs are typically determined by mortality indices based on one or more populations. As the longevity experience of the pension plan can deviate significantly from the reference populations, the basis risk, or more specifically, the population basis risk, is said to occur; see also Li and Hardy (2011), Coughlan et al. (2012). Another example is in the context of managing agricultural risk. In this application, using weather derivatives for hedging agricultural risk could give rise to variable basis risk and spatial basis risk (e.g., Brockett et al., 2005; Woodard and Garcia, 2008). Another situation for which basis risk arises is when a farmer purchases a crop insurance that is based on area yield, instead of individual yield. The area-yield crop insurance, which is known as the Group Risk Plan in the U.S., is an insurance scheme with indemnity depending on the aggregated county yields. The individual-yield crop insurance, which is known as the Annual Production History Insurance in the U.S., is another insurance scheme with payoff that is linked to individual farm yields. The discrepancy between yields at the county level and at the individual level gives rise to the basis risk; see for example Skees, et al. (1997) and Turvey and Islam (1995).

A typical example in the financial market is that the hedging for an option written on a non-tradable asset is often conducted via trading over one liquidly traded asset which is closely correlated with the non-tradable underlying asset. However, one should be very careful to use such a strategy since “close correlation” between the two underlying assets cannot guarantee the hedging performance to be as good as one may desire. Indeed, Davis (2006) showed that the unhedgeable noise, which is attributed to the mismatch between the two assets, may be huge even though the two underlying assets have very high correlation, and the “naive” hedging strategy may be ineffective.

In the existing literature, analytical results on optimal hedging in the presence of basis risk can broadly be classified into two streams. In order to ensure the model’s tractability, the first stream of investigation considers hedging general derivatives with basis risk under an exponential utility maximization framework. The pioneering closed-form optimal hedging strategies were obtained by Davis (2006).<sup>1</sup> The basic model of Davis (2006) was subsequently extended by Monoyios (2004) and Musiala and Zariphopoulou (2004) in a few interesting directions including indifference pricing, perturbation expansions, etc. All of these generalizations are restricted to

---

<sup>1</sup>Note that the work of Davis (2006) was done in 2000 but it was not formally published until 2006.

an exponential preference optimization framework. If we were to consider other optimization hedging frameworks such as under a mean-variance criterion, analytical optimal strategies with basis risk have been obtained but only for hedging futures. We classify this line of inquiry as the second stream. The main contribution is attributed to Duffie and Richardson (1991) who obtained the optimal continuous-time futures hedging policy under geometric Brownian motion assumptions. They demonstrated that the optimal hedging strategy can be derived from the normal equations for orthogonal projection in a Hilbert space. Their method, however, is not readily applicable to more general derivatives other than the futures contract. This is because their proposed method depends highly on the specific formulation of the problem and the trivial structure of the payoff function of the futures contract.

Motivated by the above two streams of investigation, this chapter attempts to address each of their limitations by studying the dynamic hedging of general European options with basis risk under a mean-variance criterion. Since the seminal work of Markowitz (1952), the mean-variance criterion has been widely applied in finance. A key advantage of the mean-variance criterion over a utility maximization objective is that in practice it is typically challenging to accurately evaluate a hedger's utility function while the mean-variance criterion provides a subjective measure. Furthermore, by comparing to the expected utility approach, MacLean et al. (2011) concluded that, for less volatile financial market, the mean-variance criterion yields a better investment portfolio return.

It is important to emphasize that the optimal portfolio model of Markowitz (1952) is a one-period model. If we are interested in a dynamic portfolio selection strategy, it is important to distinguish optimal strategy that is "pre-commitment" from "time-consistent planning" because of the added possibility of re-optimizing and re-balancing the portfolio at intertemporal times. After a decision maker obtained his/her optimal dynamic strategy at time  $t_1$ , the decision maker might find that the adopted strategy from  $t_1$  does not necessarily maximize his/her objective by the time he/she progresses to time  $t_2$ , where  $t_2 > t_1$ . In this situation, the decision maker can either continue to adopt the original plan or to devise a new plan that is "optimal" for him/her at time  $t_2$ . Strotz (1955) referred the former strategy as the "precommitment" strategy and the latter as the "consistent planing" strategy. Strotz (1955) also showed that the best investment strategy should be a plan for which the investor will actually follow, e.g., a consistent planing strategy.

The analytical solutions provided by Zhou and Li (2000) and Li and Ng (2000) for, respectively, the continuous-time and multiperiod analogs of Markowitz (1952) are examples of pre-commitment optimal strategies. To derive the optimal strategies that are time consistent under the mean-variance criterion is considerably more subtle. The complication is driven by the fact that the mean-variance function is not separable so that the Bellman optimality principle cannot be directly applied for deriving an optimal dynamic solution. This problem was not solved until another decade later by Basak and Chabakauri (2010) who provided a novel approach of obtaining a “consistent planning” solution to the portfolio selection problem involving mean-variance objective. They used the total variance formula to derive an extended Hamilton-Jacobi-Bellman (HJB) equation and ingeniously obtain the optimal hedging strategy without directly solving the extended HJB equation as a partial differential equation. Subsequently Björk and Murgoci (2010, 2014) developed a more rigorous theory for general time-inconsistent problems by providing a formal way of defining a “consistent planning” solution using game theoretic approach and providing the verification theorem. In recent years, the time consistent planning strategies have also been widely studied for decision problems in insurance, e.g., Li et al. (2012), Li et al. (2015a, 2015b), Liang and Song (2015), Wei et al. (2013), Wong et al. (2014), Wu and Zeng (2015), Zeng et al. (2013), Zhao et al. (2016), Zhou et al. (2016), just to name a few.

In this chapter, we aim to establish a “consistent planning” optimal hedging strategy in the sense of Björk and Murgoci (2010). The problem is solved by resorting to a dynamic programming procedure and solving an extended HJB equation using a certain change-of-measure technique. The solution we obtain is tractable and to the best of our knowledge, this is the first time the analytical solution exists for dynamic hedging of general European options with basis risk under the mean-variance criterion. The solution we obtained also reduces to the classical delta hedging strategy when the two involved assets are indistinguishable and the risk aversion coefficient in the mean-variance objective goes to infinity. For plain vanilla call options, the calculation of the optimal strategy requires only a minimum amount of numerical procedure. Examples based on hedging futures and European call options are presented to highlight the importance of our proposed optimal strategy, relative to other commonly adopted hedging strategies that do not take into consideration the basis risk.

The rest of the chapter proceeds as follows. The problem formulation is given in Section

2.2 and the consistent planning equilibrium solution is derived in Section 2.3. Discussions on some special cases are presented in Section 2.4. Some numerical examples are provided in Section 2.5 to highlight our theoretical results. Section 2.6 concludes the chapter. Finally, the appendix contains some technical proofs and semi closed-form expressions for the equilibrium value functions of both futures and European call options.

## 2.2 Formulation of the optimal hedging problem

Let us begin by first introducing the following notations. For a function  $F(t, s_1, s_2, x)$ , we use  $F_y$  to denote its first partial derivative with respect to (w.r.t.) variable  $y$  where  $y \in \{t, s_1, s_2, x\}$ . Analogously, we use  $F_{yz}$  to denote its second derivatives w.r.t. variables  $y$  and  $z$  where  $y, z \in \{t, s_1, s_2, x\}$ . Note that the function  $F$  and its derivatives can be time-dependent processes. In this case, each of the notation will be indexed by an argument  $t$ . Similarly, if the arguments  $s_1, s_2$  and  $x$  are also processes, then they will be denoted by  $S_1(t), S_2(t)$  and  $X(t)$ , respectively.

Consider a non-arbitrage market with two risky assets  $\{S_1(t), t \geq 0\}$  and  $\{S_2(t), t \geq 0\}$  as well as a risk-free asset earning at a constant rate of  $r > 0$ . The price processes of the two risk assets are defined over a common probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and they follow two general diffusion processes under the physical measure  $\mathbb{P}$  as below:

$$\begin{cases} \frac{dS_i(t)}{S_i(t)} = \mu_i(t, S_i(t))dt + \sigma_i(t, S_i(t))dW_i(t), & i = 1, 2, \\ dW_1(t)dW_2(t) = \rho(t)dt, \end{cases} \quad (2.1)$$

where  $W_1 := \{W_1(t), t \geq 0\}$  and  $W_2 := \{W_2(t), t \geq 0\}$  are two standard Brownian motions under  $\mathbb{P}$ . The coefficient  $\rho(t)$  is a deterministic function of  $t$ ,  $\mu_i(t, s) : \mathbb{R}_+ \times \mathbb{R} \mapsto \mathbb{R}$  and  $\sigma_i(t, s) : \mathbb{R}_+ \times \mathbb{R} \mapsto (0, \infty)$ ,  $i = 1, 2$ , where  $\mathbb{R}$  and  $\mathbb{R}_+$  respectively denote the real line and the set of nonnegative real numbers. When there is no confusion about their dependence on  $t$  and  $s$ , it is convenient to use the simplified notations  $\rho$ ,  $\mu_i$ , and  $\sigma_i$ , respectively. To ensure the existence of a unique strong solution to the stochastic differential equation (SDE) (2.1), we assume that the drift and diffusion coefficients for both  $S_i$  satisfy the global Lipschitz continuity condition, i.e.,

for  $i = 1, 2$ ,  $\exists K > 0$  s.t.  $\forall t \in \mathbb{R}_+$  and  $x, y \in \mathbb{R}$ ,

$$|x \cdot \mu_i(t, x) - y \cdot \mu_i(t, y)| + |x \cdot \sigma_i(t, x) - y \cdot \sigma_i(t, y)| \leq K|x - y|. \quad (2.2)$$

If we take  $y = 0$ , the above condition becomes

$$|\mu_i(t, x)| + |\sigma_i(t, x)| \leq K, \quad \forall x \in \mathbb{R}, \quad (2.3)$$

which means that both  $\mu_i(t, x)$  and  $\sigma_i(t, x)$  are bounded from above. Furthermore, we impose the non-degeneracy assumption on  $\sigma_i$ , i.e.,

$$\sigma_i(t) \geq \epsilon, \quad i = 1, 2, \quad \text{for some constant } \epsilon > 0. \quad (2.4)$$

The specification in equation (2.1) implies that the random sources between the two risky assets are correlated and the strength of correlation is governed by the coefficient function  $\rho(t)$ . Let  $G = G(S_2(T))$  be the payoff at maturity  $T > 0$  of a European option that is written on asset  $S_2$ , and write

$$\Pi(t, s_2) := \mathbb{E}_{t, s_2}[e^{-r(T-t)}G(S_2(T))]. \quad (2.5)$$

It is worth noting that  $\Pi(t, s_2)$  differs from the time- $t$  price of the European option  $G(S_2(T))$  since the expectation in (2.5) is taken under the physical measure  $\mathbb{P}$ , as opposed to a risk neutral probability measure.

For technical purposes, we assume the derivatives  $\Pi_t$  and  $\Pi_{s_2 s_2}$  exist and the condition

$$\mathbb{E} \left[ \int_0^T (S_2(t) \Pi_{s_2}(t, S_2(t)))^2 dt \right] < \infty \quad (2.6)$$

is satisfied. When the coefficients  $\mu_2$  and  $\sigma_2$  are constants, a sufficient and mild condition for the existence of the derivative  $\Pi_t$  and  $\Pi_{s_2 s_2}$  is given by

$$\exists a > 0 \text{ such that } \int_{-\infty}^{\infty} e^{-ax^2} |G(x)| dx < \infty; \quad (2.7)$$

see Musiela and Rutkowski (p.124, 1997) for detailed discussion. Both conditions (2.6) and (2.7) are quite mild and satisfied by most financial derivatives.

We assume that the hedging target is a short position of the contingent payoff  $G = G(S_2(T))$ , which could be either interpreted as a short position of a concrete financial derivative for a trader, or more generally, as a contingent payoff on the liability side of a product line. In both cases there is incentive to hedge against such a position for both internal risk management purposes and regulatory purposes. We assume that  $S_2$  is either a non-tradable asset or a thinly traded asset so that it lacks the necessary liquidity to be used for hedging the option that is written on it. Instead, we assume that  $S_1$  is a highly liquid and tradable asset so that together with the risk-free asset, a hedging portfolio can be constructed to hedge a short position of the above European option written on asset  $S_2$ . As  $S_1$  is related to  $S_2$  via the correlation parameter  $\rho(t)$ , using  $S_1$  to hedge  $G(S_2(T))$  gives rise to basis risk, unless in the special case  $\rho(t) = 1 \forall t \in [0, T]$ , and the coefficient functions  $\mu_1(\cdot, \cdot) = \mu_2(\cdot, \cdot)$  and  $\sigma_1(\cdot, \cdot) = \sigma_2(\cdot, \cdot)$ , where the two processes  $S_1$  and  $S_2$  are indistinguishable from each other as formally proved in Proposition 2.2 in Section 2.4.

At any time  $t \in [0, T]$ , the hedging portfolio is fully specified by the pair  $\{X^\theta(t), \theta(t)\}$ , where  $\theta(t)$  denotes the time- $t$  investment in the risky asset  $S_1$  and  $X^\theta(t)$  represents the time- $t$  hedging portfolio value resulting from a strategy  $\theta$ . This implies that the time- $t$  investment in the risk-free asset is given by  $X^\theta(t) - \theta(t)$ . At the inception of the option contract, i.e., at  $t = 0$ , the hedging cost is given by  $x_0 = X^\theta(0) > 0$ . This also corresponds to the initial value of the hedging portfolio. Then, the value process of the hedging portfolio is governed by the following SDE:

$$\begin{aligned} dX^\theta(t) &= \frac{\theta(t)}{S_1(t)} dS_1(t) + [X^\theta(t) - \theta(t)]r dt \\ &= [rX^\theta(t) + \theta(t)(\mu_1 - r)]dt + \theta(t)\sigma_1 dW_1(t), \quad t \in (0, T) \\ X^\theta(0) &= x_0, \end{aligned} \tag{2.8}$$

where  $\theta(t) = \theta(t, S_1(t), S_2(t), X^\theta(t))$ ,  $t \in [0, T]$ . Note that  $X^\theta(t)$  is a controlled Markovian process.

Let  $\mathcal{F} := \{\mathcal{F}_t, t \geq 0\}$  be the filtration generated by  $\{(S_1(t), S_2(t)), t \geq 0\}$  and write the con-

ditional expectation as  $E_{t,s_1,s_2,x}[\cdot] = E[\cdot | S_1(t) = s_1, S_2(t) = s_2, X^\theta(t) = x]$ ,  $\forall (t, s_1, s_2, x) \in [0, T] \times \mathbb{R}_+^2 \times \mathbb{R}$ . We are interested in the optimal hedging strategy among those admissible strategies in Definition 2.1 below.

**Definition 2.1.** *An admissible strategy  $\theta(t) = \theta(t, X^\theta(t), S_1(t), S_2(t))$ ,  $t \in [0, T]$  is defined as a progressively measurable process such that:*

$$(a) \ E \left[ \int_0^T \theta(u)^2 du \right] < \infty;$$

$$(b) \ E_{t,s_1,s_2} \left[ \int_t^T |\theta(u)| du \right] \leq K e^{K(s_1^2 + s_2^2)} \text{ for some constant } K > 0, \forall (t, s_1, s_2) \in [0, T] \times \mathbb{R}_+^2.$$

We use  $\Theta$  to denote the set of all admissible strategies.

Both conditions (a) and (b) in Definition 2.1 are quite mild. The square integrability condition in (a) is almost the minimum requirement to ensure that the SDE (2.8) for  $X^\theta$  is well defined, and the exponentially growth condition (b) allows a wide class of admissible strategies. Imposing condition (b) ensures the uniqueness of the solution to the partial differential equation (PDE) (2.37), which is critical to deriving an explicit optimal strategy as we will see in Section 2.3.3.

Because of the basis risk and the market incompleteness, the hedging strategy involving  $\theta$  can not perfectly replicate the maturity value of the European option. The hedging error at expiration of the option is given by  $G(S_2(T)) - X^\theta(T)$ . By defining  $V^\theta(T)$  as the terminal profit-and-loss random variable for the hedger, we have

$$V^\theta(T) = X^\theta(T) - G(S_2(T)). \quad (2.9)$$

For any time  $t < T$  and under mean-variance criterion, an optimal hedging strategy can be defined as the one that solves the following optimization problem:

$$\max_{\theta \in \Theta} \left\{ U(t, s_1, s_2, x; \theta) := E_{t,s_1,s_2,x}[V^\theta(T)] - \frac{\gamma}{2} \text{Var}_{t,s_1,s_2,x}[V^\theta(T)] \right\}. \quad (2.10)$$

where  $\text{Var}_{t,s_1,s_2,x}[\cdot] = \text{Var}[\cdot | S_1(t) = s_1, S_2(t) = s_2, X^\theta(t) = x]$ ,  $\text{Var}[\cdot]$  denotes the variance taken under the  $\mathbb{P}$  measure, and  $\gamma > 0$  is a constant parameter capturing the risk aversion of

the hedger. Note that the objective of the hedger is to choose an optimal hedging strategy  $\theta$  that maximizes hedger's (conditional) expected profit (i.e.  $E_{t,s_1,s_2,x}[V^\theta(T)]$ ) subject to the penalty attributed to the (conditional) variance of the profit (i.e.  $\text{Var}_{t,s_1,s_2,x}[V^\theta(T)]$ ). The degree of penalty is quantified by the parameter  $\gamma$ , which is called risk aversion coefficient. It is worthwhile noting that problem (2.10) is reduced to a portfolio selection problem if  $G(S_2(T)) = 0$ . In this chapter we use the term “hedging” as the terminal profit-and-loss variable can be interpreted as the difference between the hedging portfolio and the financial liability to be hedged and the variance term in the objective function represents the size of hedging error.

Let  $\{\tilde{\theta}_0(t), t \in [t_1, T]\}$  be a pre-commitment solution of problem (2.10) derived by sitting at time  $t_1$ . In other words,  $\{\tilde{\theta}_0(t), t \in [t_1, T]\}$  is the best strategy among the admissible set to maximize the mean-variance objective  $U(t_1, s_1, s_2, x; \theta)$ . For a mean-variance optimization problem, it is well-known that the truncated strategy  $\{\tilde{\theta}_0(t), t \in [t_2, T]\}$  is not generally optimal for the decision at a later time  $t_2 > t_1$  in the sense of maximizing the objective  $U(t_2, s_1, s_2, x; \theta)$ ; see, e.g., Basak and Chabakauri, (2010). Such a phenomenon is called the time-inconsistency issue associated with mean-variance analysis. Generally we call a problem time-inconsistent, if a strategy truncated from a strategy which optimizes the objective for an earlier time is not optimal for the objective at a later time; otherwise, the problem is called time-consistent.

To develop a time-consistent hedging strategy, we follow the game theoretic framework of Björk and Murgoci (2010) and Basak and Chabakauri (2010). Using the game theoretic formulation, the “optimality” is defined as a subgame perfect Nash equilibrium solution. The idea is to take the decision making process as a non-cooperative game among a continuum of players over the time horizon (who can be viewed as the future incarnations of the decision-maker), and each player can only influence the control process over an infinitesimal time interval. The formal mathematical definition is given as follows.

**Definition 2.2.** Consider a control process  $\theta^* \in \Theta$ . For any arbitrary constant  $q \in \mathbb{R}$ ,  $\tau \in \mathbb{R}_+$ , and initial point  $(t, s_1, s_2, x)$  for  $(t, S_1(t), S_2(t), X^\theta(t))$ , define the control process  $\hat{\theta}$  as

$$\hat{\theta}(v, s_1, s_2, x) = \begin{cases} q, & \text{for } t \leq v < t + \tau; \\ \theta^*(v, s_1, s_2, x), & \text{for } t + \tau \leq v \leq T. \end{cases} \quad (2.11)$$

Then  $\theta^*$  is an equilibrium control process, if

$$\liminf_{\tau \rightarrow 0^+} \frac{U(t, s_1, s_2, x; \theta^*) - U(t, s_1, s_2, x; \widehat{\theta})}{\tau} \geq 0 \quad (2.12)$$

holds for any  $q \in \mathbb{R}$  and  $t \geq 0$ , where  $U$  is the objective function in problem (2.10). Furthermore, the equilibrium value function is defined by

$$J(t, s_1, s_2, x) = U(t, s_1, s_2, x; \theta^*). \quad (2.13)$$

As defined in the above, for any time  $t \in (0, T)$ , the truncated strategy  $\{\theta_s^*, s \in [t, T]\}$  is still an equilibrium solution for the rest time horizon  $[t, T]$ . In other words, the decision at any later time under such a game theoretical framework sticks to the strategy  $\theta^*$ , and thus, such an equilibrium solution  $\theta^*$  is also called a time-consistent solution. For a time-inconsistent problem, the precommitment solution differs from the equilibrium solution in general. In contrast, for a time-consistent problem, the precommitment solution coincides with the equilibrium solution, since in this case, the truncation of a precommitment solution over a fractional period  $[t, T]$  is still an optimal solution for the objective at time  $t$  for any  $t \in (0, T)$ .

**Remark 2.1.** *If the physical measure  $\mathbb{P}$  is a martingale measure, i.e.,  $\mu_1(t, s) = r, \forall (t, s) \in \mathbb{R}_+ \times \mathbb{R}$ , problem (2.10) is in fact a time-consistent problem with  $E_{t, s_1, s_2, x}[V^\theta(T)]$  being a constant independent of  $\theta$ . Therefore in this chapter, when we use the term “time-inconsistent problem”, it might also include some time-consistent cases as its special cases.*

## 2.3 Optimal time consistent hedging strategy

When only a controlled Markovian process is involved, there exists a standard procedure to derive the extended HJB equation for mean-variance optimization, as one can see in some recent applications such as Björk and Murgoci (2010) and Li et al. (2012). For our problem (2.10), the profit-and-loss random variable  $V^\theta(T)$ , however, depends on not only the controlled Markovian process  $\{X^\theta(t), t \in [0, T]\}$  but also the price process  $\{S_2(t), t \in [0, T]\}$ . Explicit dependence of

$G(S_2(T))$  in  $V^\theta(T)$  distinguishes our model from other mean-variance based formulations and this complicates the derivation of an optimal solution. Subsection 2.3.1 will first establish an extended HJB equation for problem (2.10) in a heuristic way, subsection 2.3.2 will then formally justify our result by providing a verification theorem. Subsection 2.3.3 demonstrates that, with certain technical conditions, the proposed solution satisfies the conditions from the verification theorem in Subsection 2.3.2, and thus it is an equilibrium solution.

### 2.3.1 The extended HJB equation

We begin by obtaining an alternate expression for the objective function in problem (2.10). We achieve this via the following total variance decomposition for an admissible hedging strategy  $\theta \in \Theta$  and  $\tau \in \mathbb{R}_+$ :

$$\text{Var}_{t,s_1,s_2,x}(V^\theta(T)) = \mathbb{E}_{t,s_1,s_2,x}[\text{Var}_{t+\tau}(V^\theta(T))] + \text{Var}_{t,s_1,s_2,x}[\mathbb{E}_{t+\tau}(V^\theta(T))]. \quad (2.14)$$

The objective function in problem (2.10) therefore can be rewritten as

$$\begin{aligned} U(t, s_1, s_2, x; \theta) &= \mathbb{E}_{t,s_1,s_2,x}[V^\theta(T)] - \frac{\gamma}{2} \text{Var}_{t,s_1,s_2,x}[V^\theta(T)] \\ &= \mathbb{E}_{t,s_1,s_2,x}[V^\theta(T)] - \frac{\gamma}{2} \mathbb{E}_{t,s_1,s_2,x}[\text{Var}_{t+\tau}(V^\theta(T))] - \frac{\gamma}{2} \text{Var}_{t,s_1,s_2,x}[\mathbb{E}_{t+\tau}(V^\theta(T))] \\ &= \mathbb{E}_{t,s_1,s_2,x}[U^\theta(t+\tau)] - \frac{\gamma}{2} \text{Var}_{t,s_1,s_2,x}[\mathbb{E}_{t+\tau}(V^\theta(T))], \end{aligned} \quad (2.15)$$

where  $U^\theta(t+\tau) := U(t+\tau, S_1(t+\tau), S_2(t+\tau), X^\theta(t+\tau))$ .

Let  $m(t, s_1, s_2, x; \theta) := \mathbb{E}_{t,s_1,s_2,x}[V^\theta(T)]$ , and denote  $m^\theta(t) := m(t, S_1(t), S_2(t), X^\theta(t); \theta)$ . The, by definition we obtain

$$\begin{aligned} &\text{Var}_{t,s_1,s_2,x}[m^\theta(t+\tau)] \\ &= \mathbb{E}_{t,s_1,s_2,x} \left[ (m^\theta(t+\tau))^2 \right] - [\mathbb{E}_{t,s_1,s_2,x}(m^\theta(t+\tau))]^2 \end{aligned}$$

$$= \mathbb{E}_{t,s_1,s_2,x} \left[ (m^\theta(t+\tau))^2 - (m^\theta(t))^2 \right] - \left\{ [\mathbb{E}_{t,s_1,s_2,x} (m^\theta(t+\tau))]^2 - [\mathbb{E}_{t,s_1,s_2,x} (m^\theta(t))]^2 \right\}.$$

For any  $\tau \in \mathbb{R}_+$  and  $q \in \mathbb{R}$ , we let  $\widehat{\theta}$  to denote a hedging strategy with a generic admissible constant  $q \in \mathbb{R}$  applied over  $[t, t + \tau)$  and the equilibrium strategy  $\theta^*$  applied over  $[t + \tau, T)$ , i.e.,  $\widehat{\theta}$  is as defined in equation (2.11). Thus, dividing by  $\tau$  and letting  $\tau \rightarrow 0$  in (2.15) gives the following extended HJB equation:

$$0 = \max_{q \in \mathbb{R}} \left( \mathcal{A}^q F(t, s_1, s_2, x) - \xi^q(m(t, s_1, s_2, x; \theta^*)) \right), \quad (2.16)$$

where  $\mathcal{A}^q$  is the infinitesimal generator for processes  $\{S_1, S_2, X^q\}$  and is given by

$$\begin{aligned} \mathcal{A}^q F(t, s_1, s_2, x) = & F_t + F_x r x + q F_x (\mu_1 - r) + F_{s_1} s_1 \mu_1 + F_{s_2} s_2 \mu_2 \\ & + \frac{1}{2} F_{xx} (q \sigma_1)^2 + \frac{1}{2} F_{s_1 s_1} (s_1 \sigma_1)^2 + \frac{1}{2} F_{s_2 s_2} (s_2 \sigma_2)^2 \\ & + F_{x s_1} s_1 \sigma_1 q \sigma_1 + F_{x s_2} s_2 \sigma_2 q \sigma_1 \rho + F_{s_1 s_2} s_1 \sigma_1 s_2 \sigma_2 \rho, \end{aligned} \quad (2.17)$$

and

$$\xi^q(m(t, s_1, s_2, x)) = \frac{\gamma}{2} \left\{ \mathcal{A}^q [m(t, s_1, s_2, x)^2] - 2m(t, s_1, s_2, x) \mathcal{A}^q [m(t, s_1, s_2, x)] \right\}. \quad (2.18)$$

### 2.3.2 Verification theorem

Based on the extended HJB equation (2.16), the objective of this section is to develop a verification theorem, together with the required conditions, that guarantees a solution of the extended HJB equation, and to solve the mean-variance optimization problem (2.10).

**Theorem 2.1.** (*Verification Theorem*). *Suppose there exists a control process  $\theta^* \in \Theta$  and a function  $F$  such that*

$$\theta^*(t) = \arg \max_q \left\{ \mathcal{A}^q F(t, s_1, s_2, x) - \xi^q(g(t, s_1, s_2, x)) \right\}, \quad (2.19)$$

$$0 = \mathcal{A}^{\theta^*} F(t, s_1, s_2, x) - \xi^{\theta^*}(g(t, s_1, s_2, x)), \quad (2.20)$$

$$F(T, s_1, s_2, x) = x - G(s_2), \quad (2.21)$$

$$g(t, s_1, s_2, x) = \mathbb{E}_{t, s_1, s_2, x}[V^{\theta^*}(T)], \quad (2.22)$$

for any  $(t, s_1, s_2, x) \in [0, T] \times \mathbb{R}_+^2 \times \mathbb{R}$ . Then  $\{\theta^*(t)\}_{t \in [0, T]}$  is an equilibrium hedging strategy and  $F(t, s_1, s_2, x)$  is the equilibrium value function, i.e.,  $F(t, s_1, s_2, x) = U(t, s_1, s_2, x; \theta^*)$ ,  $\forall (t, s_1, s_2, x) \in [0, T] \times \mathbb{R}_+^2 \times \mathbb{R}$ .

*Proof.* We will first show that  $F(t, s_1, s_2, x) = U(t, s_1, s_2, x; \theta^*)$  given that  $\theta^*$  satisfies (2.19)-(2.22). Indeed, by Dynkin's formula,

$$\begin{aligned} & \mathbb{E}_{t, s_1, s_2, x} [F(T, S_1(T), S_2(T), X^{\theta^*}(T))] \\ &= F(t, s_1, s_2, x) + \mathbb{E}_{t, s_1, s_2, x} \left[ \int_t^T \mathcal{A}^{\theta^*} F(u, S_1(u), S_2(u), X^{\theta^*}(u)) du \right]. \end{aligned} \quad (2.23)$$

Then, using the shorthand  $g(u) := g(u, S_1(u), S_2(u), X^{\theta^*}(u))$  and equations (2.18) and (2.20) yields

$$\begin{aligned} & \mathbb{E}_{t, s_1, s_2, x} [F(T, S_1(T), S_2(T), X^{\theta^*}(T))] \\ &= F(t, s_1, s_2, x) + \mathbb{E}_{t, s_1, s_2, x} \left[ \int_t^T \xi^{\theta^*} (g(u, S_1(u), S_2(u), X^{\theta^*}(u))) du \right] \\ &= F(t, s_1, s_2, x) + \mathbb{E}_{t, s_1, s_2, x} \left\{ \int_t^T \left[ \frac{\gamma}{2} \mathcal{A}^{\theta^*} [g(u)^2] - \gamma g(u) \mathcal{A}^{\theta^*} [g(u)] \right] du \right\} \\ &= F(t, s_1, s_2, x) + \mathbb{E}_{t, s_1, s_2, x} \left\{ \int_t^T \frac{\gamma}{2} \mathcal{A}^{\theta^*} [g(u)^2] du \right\} \\ &= F(t, s_1, s_2, x) + \frac{\gamma}{2} \mathbb{E}_{t, s_1, s_2, x} [g^2(T, S_1(T), S_2(T), X^{\theta^*}(T))] - \frac{\gamma}{2} g^2(t, s_1, s_2, x). \end{aligned} \quad (2.24)$$

The third equality follows from the condition (2.22) and the last equality can be obtained by applying Dynkin's formula in conjunction with the definition of  $\xi^\theta$  given in (2.18). Moreover, by the boundary condition (2.21), we have

$$\begin{aligned} \mathbb{E}_{t, s_1, s_2, x} [F(T, S_1(T), S_2(T), X^{\theta^*}(T))] &= \mathbb{E}_{t, s_1, s_2, x} [X^{\theta^*}(T) - G(S_2(T))] \\ &= \mathbb{E}_{t, s_1, s_2, x} [V^{\theta^*}(T)]. \end{aligned} \quad (2.25)$$

Combining equations (2.24) and (2.25) yields

$$\begin{aligned}
& F(t, s_1, s_2, x) \\
&= \mathbb{E}_{t, s_1, s_2, x} [V^{\theta^*}(T)] - \frac{\gamma}{2} \mathbb{E}_{t, s_1, s_2, x} [g^2(T, S_1(T), S_2(T), X^{\theta^*}(T))] + \frac{\gamma}{2} g^2(t, s_1, s_2, x) \\
&= \mathbb{E}_{t, s_1, s_2, x} [V^{\theta^*}(T)] - \frac{\gamma}{2} \text{Var}_{t, s_1, s_2, x} (V^{\theta^*}(T)) \\
&= U(t, s_1, s_2, x; \theta^*), \tag{2.26}
\end{aligned}$$

where the second equality follows from the definition of  $g$  given in (2.22). What we have established in equation (2.26) is that a function  $F(t, s_1, s_2, x)$  that satisfies conditions (2.19)-(2.22) is the equilibrium value.

It remains to show that  $\theta^*$  is an equilibrium strategy. We start from equations (2.19) and (2.20) to obtain

$$\mathcal{A}^q F(t, s_1, s_2, x) - \xi^q(t, s_1, s_2, x) \leq 0, \quad \forall q \in \mathbb{R}.$$

Then, discretizing the left-hand-side of the above inequality leads to

$$\begin{aligned}
& \mathbb{E}_{t, s_1, s_2, x} [F(t + \tau, S_1(t + \tau), S_2(t + \tau), X^{\hat{\theta}}(t + \tau))] - F(t, s_1, s_2, x) \\
& - \frac{\gamma}{2} \left( \mathbb{E}_{t, s_1, s_2, x} [g(t + \tau, S_1(t + \tau), S_2(t + \tau), X^{\hat{\theta}}(t + \tau))^2] - g(t, s_1, s_2, x)^2 \right) \\
& + \frac{\gamma}{2} \left( \mathbb{E}_{t, s_1, s_2, x}^2 [g(t + \tau, S_1(t + \tau), S_2(t + \tau), X^{\hat{\theta}}(t + \tau))] - g(t, s_1, s_2, x)^2 \right) \leq o(\tau),
\end{aligned}$$

where  $o(\tau)/\tau \rightarrow 0$  as  $\tau \rightarrow 0^+$ . We further use the definition of  $g$  in equation (2.22) to obtain

$$\begin{aligned}
F(t, s_1, s_2, x) &\geq \mathbb{E}_{t, s_1, s_2, x} [F(t + \tau, S_1(t + \tau), S_2(t + \tau), X^{\hat{\theta}}(t + \tau))] \\
&\quad - \frac{\gamma}{2} \left( \mathbb{E}_{t, s_1, s_2, x} [g(t + \tau, S_1(t + \tau), S_2(t + \tau), X^{\hat{\theta}}(t + \tau))^2] \right. \\
&\quad \left. - \mathbb{E}_{t, s_1, s_2, x}^2 [g(t + \tau, S_1(t + \tau), S_2(t + \tau), X^{\hat{\theta}}(t + \tau))] \right) + o(\tau) \\
&= \mathbb{E}_{t, s_1, s_2, x} \left[ F(t + \tau, S_1(t + \tau), S_2(t + \tau), X^{\hat{\theta}}(t + \tau)) \right] \\
&\quad - \frac{\gamma}{2} \text{Var}_{t, s_1, s_2, x} \left[ \mathbb{E}_{t+\tau} (V^{\hat{\theta}}(T)) \right] + o(\tau).
\end{aligned}$$

Since we have proved that  $F(t, s_1, s_2, x) = U(t, s_1, s_2, x; \theta^*)$ , we combine equations (2.15) and the last display to obtain

$$\liminf_{\tau \rightarrow 0^+} \frac{U(t, s_1, s_2, x; \theta^*) - U(t, s_1, s_2, x; \hat{\theta})}{\tau} \geq 0, \quad \forall q \in \mathbb{R},$$

which implies that  $\theta^*$  is an equilibrium strategy.  $\square$

## 2.3.3 Equilibrium solution

### 2.3.3.1 Candidate solution and technical conditions

In the next subsection, we will show that, under some technical conditions, the following  $\theta^*$  is an equilibrium solution:

$$\theta^*(t, s_1, s_2) = e^{-r(T-t)} \left[ \frac{\mu_1 - r}{\gamma \sigma_1^2} - \eta_{s_1}(t) s_1 - \frac{s_2 \sigma_2 \rho}{\sigma_1} (\eta_{s_2}(t) - e^{r(T-t)} \Pi_{s_2}(t)) \right], \quad (2.27)$$

where

$$\eta(t, s_1, s_2) = \mathbb{E}_{t, s_1, s_2}^* \left\{ \int_t^T \left[ \frac{1}{\gamma} \left( \frac{\mu_1 - r}{\sigma_1} \right)^2 + (\mu_1 - r) \frac{\rho \sigma_2}{\sigma_1} S_2(u) e^{r(T-u)} \Pi_{s_2}(u) \right] du \right\}. \quad (2.28)$$

In the above,  $\eta_{s_1} = \frac{\partial}{\partial s_1} \eta(t, s_1, s_2)$ ,  $\Pi_{s_2} = \frac{\partial}{\partial s_2} \Pi(t, s_2)$ .  $\mathbb{E}_{t, s_1, s_2}^*[\cdot]$  denotes conditional expectation under probability measure  $\mathbb{P}^*$ , which is defined by the Radon-Nikodym derivative given as follows:

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = \exp \left( -\frac{1}{2} \int_0^t \left( \frac{\mu_1 - r}{\sigma_1} \right)^2 du - \int_0^t \frac{\mu_1 - r}{\sigma_1} dW_1(u) \right). \quad (2.29)$$

By the conditions on the boundedness of coefficients  $\mu_i$  and  $\sigma_i$  respectively given in equations (2.3) and (2.4), the right hand side of equation (2.29) is well defined and Novikov's condition

$$\mathbb{E} \left[ \exp \left\{ \frac{1}{2} \int_0^T \left( \frac{\mu_1 - r}{\sigma_1} \right)^2 du \right\} \right] < \infty$$

is satisfied. Consequently, by the Girsanov's Theorem (Karatzas and Shreve, 1998), under  $\mathbb{P}^*$ , the two risky assets  $S_1$  and  $S_2$  follow the following dynamics:

$$\begin{cases} \frac{dS_1(t)}{S_1(t)} = rdt + \sigma_1 dW_1^*(t), \\ \frac{dS_2(t)}{S_2(t)} = \left( \mu_2 - (\mu_1 - r) \frac{\rho\sigma_2}{\sigma_1} \right) dt + \sigma_2 dW_2^*(t), \end{cases} \quad (2.30)$$

where  $W_1^*$  and  $W_2^*$  are two standard Brownian motions with  $dW_1^* dW_2^* = \rho dt$  under  $\mathbb{P}^*$ .

It is notable that the introduction of measure  $\mathbb{P}^*$  allows to express the equilibrium solution  $\theta^*$  using (2.27) and (2.28) explicitly. In general, it is not clear whether  $\theta^*$  given in (2.27) is admissible, i.e.,  $\theta^* \in \Theta$  where  $\Theta$  is defined in Definition 2.1. To ensure this, we need certain technical conditions, and to facilitate the development, we define

$$A(u, s) := \frac{\mu_1(u, s) - r}{\sigma_1(u, s)} \quad \text{and} \quad B(u, s) := \sigma_2(u, s) s \Pi_{s_2}(u, s), \quad \forall (u, s) \in \mathbb{R}_+^2. \quad (2.31)$$

A sufficient condition for the admissibility of  $\theta$  is that the following conditions **C1** and **C2** hold (see Proposition 2.1 in the sequel for formal justification):

- C1.**  $\forall (t, s_1, s_2) \in [0, T] \times \mathbb{R}_+^2$ ,  $\eta_{s_1}(t, s_1, s_2)$  and  $\eta_{s_2}(t, s_1, s_2)$  exist, and there exists a constant  $K > 0$  such that  $|\eta(t, s_1, s_2)| \leq K \cdot e^{K(s_1^2 + s_2^2)}$ .
- C2.**  $\forall (t, s_1, s_2) \in [0, T] \times \mathbb{R}_+^2$  and  $\forall u \in [t, T]$ , there exists positive constants  $K$ ,  $K_1$  and  $K_2$  such that the following three partial derivatives exist and satisfy:

$$\begin{cases} \frac{\partial}{\partial s_1} \mathbb{E}_{t, s_1}^* [A(u, S_1(u))^2] \leq K(1 + |s_1|^{K_1}), \\ \frac{\partial}{\partial s_1} \mathbb{E}_{t, s_1, s_2}^* [A(u, S_1(u))B(u, S_2(u))] \leq K(1 + |s_1|^{K_1} + |s_2|^{K_2}), \\ \frac{\partial}{\partial s_2} \mathbb{E}_{t, s_1, s_2}^* [A(u, S_1(u))B(u, S_2(u))] \leq K(1 + |s_1|^{K_1} + |s_2|^{K_2}). \end{cases} \quad (2.32)$$

**Remark 2.2.** When coefficients  $\rho$ ,  $\mu_i$  and  $\sigma_i$ ,  $i = 1, 2$ , are all constants,  $\eta(t, s_1, s_2)$  is independent of  $s_1$  and  $A(u, s)$  is a constant. Thus, in this case,  $\eta_{s_1} = 0$  and the first condition in (2.32) is trivially true. As a consequence, having conditions **C1** and **C2** is equivalent to the following condition:

**H.**  $\forall (t, s_2) \in [0, T] \times \mathbb{R}_+^2$  and  $\forall u \in [t, T]$ ,  $|\tilde{\eta}(t, s_2)| \leq K \cdot e^{K \cdot s_2^2}$  holds for some constant  $K > 0$ , and the derivative  $\tilde{\eta}_{s_2}(t, s_2)$  exist, where

$$\tilde{\eta}(t, s_2) := \mathbb{E}_{t, s_2}^* \left[ \int_t^T S_2(u) e^{r(T-u)} \Pi_{s_2}(u) du \right], \quad u > t; \quad (2.33)$$

moreover, the following two derivatives exist and satisfy

$$\begin{cases} \Pi_{s_2}(t, s_2) \leq K(1 + |s_2|^{K_2}), \\ \Pi_{s_2 s_2}(t, s_2) \leq K(1 + |s_2|^{K_2}). \end{cases} \quad (2.34)$$

Note that, given constant coefficients for both processes  $S_1$  and  $S_2$ , condition H is fully determined by the payoff function of the European option, and it is easily satisfied for common derivatives; see Section 2.4.4 for more details w.r.t. futures contracts and European call options.

For a general model, conditions **C1** and **C2** are not transparent and they have to be checked based on the specific dynamics of asset prices  $S_1$  and  $S_2$  as well as the function  $\Pi$ , which further depends on the payoff of the European option. However, a sufficient condition which is a little stronger but more transparent to verify than condition **C2** is given as follows:

**C2'**. There exist positive constants  $K$ ,  $K_1$  and  $K_2$  such that,  $\forall (t, s_1, s_2) \in [0, T] \times \mathbb{R}_+^2$ ,

$$\begin{cases} \frac{\partial}{\partial s_1} A(t, s_1) \leq K(1 + |s_1|^{K_1}), \\ \Pi_{s_2}(t, s_2) \leq K(1 + |s_2|^{K_2}), \\ \frac{\partial}{\partial s_2} \sigma_2(t, s_2) \leq K(1 + |s_2|^{K_2}), \\ \Pi_{s_2 s_2}(t, s_2) \leq K(1 + |s_2|^{K_2}). \end{cases} \quad (2.35)$$

Proposition 2.1 below rigorously clarifies the relationship between **C2** and **C2'** and their sufficiency (together with **C1**) for  $\theta^* \in \Theta$ .

**Proposition 2.1.** *Given condition **C1** and the existence of the three derivatives in equation (2.32), condition **C2'** implies condition **C2**, which together with **C1** further implies  $\theta^* \in \Theta$ .*

*Proof.* See Appendix A.1. □

### 2.3.3.2 Verification of equilibrium solution

The formal justification of the optimality of  $\theta^*$  is given in Theorem 2.2 in the sequel and its proof depends on the following technical lemma.

**Lemma 2.1.** *Given that conditions **C1** and **C2** hold,  $\eta$  satisfies the following recursion:*

$$\eta(t, s_1, s_2) = \mathbb{E}_{t, s_1, s_2} \left[ \int_t^T e^{r(T-u)} \theta^*(u) (\mu_1 - r) du \right], \quad (2.36)$$

where  $\theta^*$  is given by equation (2.27).

*Proof.* By applying Feynman-Kac Theorem to function  $\eta(t, s_1, s_2)$  along with equation (2.30), we obtain the following partial differential equation (PDE):

$$\begin{aligned} \eta_t + s_1 \mu_1 \eta_{s_1} + s_2 \mu_2 \eta_{s_2} + \frac{1}{2} s_1^2 \sigma_1^2 \eta_{s_1 s_1} + \frac{1}{2} s_2^2 \sigma_2^2 \eta_{s_2 s_2} + \rho s_1 s_2 \sigma_1 \sigma_2 \eta_{s_1 s_2} \\ + \frac{1}{\gamma} \left( \frac{\mu_1 - r}{\sigma_1} \right)^2 - (\mu_1 - r) s_1 \eta_{s_1} - (\mu_1 - r) \frac{\rho \sigma_2}{\sigma_1} s_2 (\eta_{s_2} - e^{r(T-t)} \Pi_{s_2}) = 0. \end{aligned} \quad (2.37)$$

The given exponential growth condition in condition **C1** guarantees the uniqueness of solution to the second-order linear parabolic PDE (2.37) (see Chen, 2003; Lieberman, 1996). Thus, applying Feynman-Kac Theorem again, its solution is given by

$$\begin{aligned} \eta(t, s_1, s_2) &= \mathbb{E}_t \left[ \int_t^T \frac{1}{\gamma} \left( \frac{\mu_1 - r}{\sigma_1} \right)^2 du \right] \\ &\quad - \mathbb{E}_t \left[ \int_t^T (\mu_1 - r) \left( \eta_{s_1}(u) S_1(u) + \frac{\rho \sigma_2 S_2(u)}{\sigma_1} [\eta_{s_2}(u) - e^{r(T-s)} \Pi_{s_2}(u)] \right) du \right] \\ &= \mathbb{E}_{t, s_1, s_2} \left[ \int_t^T e^{r(T-u)} \theta^*(u) (\mu_1 - r) du \right], \end{aligned} \quad (2.38)$$

in view of equation (2.27). □

Recall from equation (2.9) that the random variable  $V^\theta(T) = X^\theta(T) - G(S_2(T))$  represents the profit-and-loss of the hedging strategy at maturity of the European option. For  $t < T$ , we

define  $V^\theta(t)$  as

$$V^\theta(t) \equiv V(t, S_2(t), X^\theta(t)) := X^\theta(t) - \Pi(t, s_2), \quad t \in [0, T], \quad (2.39)$$

where  $\Pi(t, s_2)$  is defined in equation (2.5).

To facilitate further development, we introduce the following functions:

$$\begin{cases} m(t, s_1, s_2, x; \theta) := \mathbb{E}_{t, s_1, s_2, x}[V^\theta(T)], \\ n(t, s_2, x) := [x - \Pi(t, s_2)] e^{r(T-t)}, \\ l(t, s_1, s_2, x; \theta) := \mathbb{E}_{t, s_1, s_2, x} \left[ \int_t^T e^{r(T-u)} \theta(u) (\mu_1 - r) du \right], \end{cases} \quad (2.40)$$

for  $(t, s_1, s_2, x) \in [0, T] \times \mathbb{R}_+^2 \times \mathbb{R}$ , and adopt the following short-hand notations:

$$\begin{cases} m^\theta(t) = m(t, S_1(t), S_2(t), X^\theta(t); \theta), \\ n^\theta(t) = n(t, S_2(t), X^\theta(t)), \\ l^\theta(t) = l(t, S_1(t), S_2(t), X^\theta(t); \theta), \end{cases}$$

for  $t \in [0, T]$ . By Feynman-Kac formula we obtain

$$r\Pi(t, s_2) = \Pi_t(t, s_2) + \Pi_{s_2}(t, s_2) s_2 \mu_2 + \frac{1}{2} \Pi_{s_2 s_2}(t, s_2) s_2^2 \sigma_2^2. \quad (2.41)$$

Consequently, for  $t \in (0, T)$ , we apply Itô's formula to  $n^\theta(t)$  in conjunction with equations (2.1), (2.8) and (2.41) to obtain

$$\begin{aligned} dn^\theta(t) &= d[X^\theta(t) e^{r(T-t)}] - d[\Pi(t) e^{r(T-t)}] \\ &= e^{r(T-t)} dX^\theta(t) - X^\theta(t) e^{r(T-t)} r dt + \Pi(t) e^{r(T-t)} r dt - e^{r(T-t)} d\Pi(t) \\ &= e^{r(T-t)} \theta(t) (\mu_1 - r) dt + e^{r(T-t)} [\theta(t) \sigma_1 dW_1(t) - \Pi_{s_2}(t) S_2(t) \sigma_2 dW_2(t)]. \end{aligned} \quad (2.42)$$

By Proposition 2.1,  $\theta^* \in \Theta$  and thus, the condition (a) in Definition 2.1 ensures

$$\mathbb{E}_{t, s_1, s_2, x} \left[ \int_0^t e^{r(T-u)} \theta(u) \sigma_1 dW_1(u) \right] = 0.$$

Moreover, from the conditions given in equations (2.3) and (2.6), it follows

$$\mathbb{E}_{t,s_1,s_2,x} \left[ \int_0^t e^{r(T-u)} \Pi_{s_2}(u) S_2(u) \sigma_2 dW_2(u) \right] = 0.$$

Thus, equation (2.42), in conjunction with the fact that  $n^\theta(T) = V^\theta(T)$  and  $n^\theta(t) = V^\theta(t)e^{r(T-t)}$ , yields

$$\begin{aligned} m(t, s_1, s_2, x; \theta) &\equiv \mathbb{E}_{t,s_1,s_2,x}[V^\theta(T)] \\ &= n(t, s_2, x) + \mathbb{E}_{t,s_1,s_2,x} \left( \int_t^T e^{r(T-u)} \theta(u) (\mu_1 - r) du \right) \\ &= n(t, s_2, x) + l(t, s_1, s_2, x; \theta), \end{aligned} \quad (2.43)$$

and

$$\begin{aligned} V^\theta(T) &= e^{r(T-t)} V^\theta(t) + \int_t^T d e^{r(T-u)} V^\theta(u) \\ &= n(t, s_2, x) + \int_t^T e^{r(T-u)} \theta(u) (\mu_1 - r) du \\ &\quad + \int_t^T e^{r(T-u)} [\theta(u) \sigma_1 dW_1(u) - \Pi_{s_2}(u) S_2(u) \sigma_2 dW_2(u)]. \end{aligned} \quad (2.44)$$

Consequently, we can rewrite the objective function in equation (2.10) as follows:

$$\begin{aligned} U(t, s_1, s_2, x; \theta) &= \mathbb{E}_{t,s_1,s_2,x}[V^\theta(T)] - \frac{\gamma}{2} \text{Var}_{t,s_1,s_2,x}[V^\theta(T)] \\ &= n(t, s_2, x) + l(t, s_1, s_2, x; \theta) \\ &\quad - \frac{\gamma}{2} \text{Var}_{t,s_1,s_2,x} \left[ e^{r(T-t)} V^\theta(t) + \int_t^T d(e^{r(T-u)} V^\theta(u)) \right] \\ &= n(t, s_2, x) + \tilde{U}(t, s_1, s_2, x; \theta), \end{aligned} \quad (2.45)$$

where

$$\tilde{U}(t, s_1, s_2, x; \theta)$$

$$\begin{aligned}
&= l(t, s_1, s_2, x; \theta) - \frac{\gamma}{2} \text{Var}_{t, s_1, s_2, x} \left( \int_t^T e^{r(T-u)} \theta(u) (\mu_1 - r) du \right. \\
&\quad \left. + \int_t^T e^{r(T-u)} [\theta(u) \sigma_1 dW_1(u) - \Pi_{s_2}(u) S_2(u) \sigma_2 dW_2(u)] \right). \quad (2.46)
\end{aligned}$$

Equation (2.45) means that our objective function  $U(t, s_1, s_2, x; \theta)$  can be separated into two parts, which is quite essential to solve for equilibrium control  $\theta^*$ . Let us denote  $C^{1,2,2,2} = \{f(t, s_1, s_2, x) : [0, T] \times \mathbb{R}_+^2 \times \mathbb{R} \mapsto \mathbb{R} \text{ s.t. } f \text{ is continuously differentiable in } t \text{ and twice continuously differentiable in } s_1, s_2 \text{ and } x\}$ .

**Theorem 2.2.** *With conditions C1 and C2 (or C2'),  $\theta^*$  defined in equation (2.27) is an equilibrium solution and the equilibrium value is given by  $U(t, s_1, s_2, x; \theta^*)$ , provided that  $U(t, s_1, s_2, x; \theta^*) \in C^{1,2,2,2}$ .*

*Proof.* We need to show  $(\theta^*, F)$  with  $F(\cdot) = U(\cdot, \cdot, \cdot, \cdot; \theta^*)$  solves the equation system (2.19)-(2.22). From equation (2.45), we apply Lemma 2.1 to obtain

$$F(t, s_1, s_2, x; \theta^*) = U(t, s_1, s_2, x; \theta^*) = n(t, x, s_2) + \tilde{U}(t, s_1, s_2; \theta^*), \quad (2.47)$$

with

$$\begin{aligned}
\tilde{U}(t, s_1, s_2; \theta^*) &= \eta(t, s_1, s_2) - \frac{\gamma}{2} \text{Var}_{t, s_1, s_2, x} \left( \int_t^T e^{r(T-u)} \theta^*(u) (\mu_1 - r) du \right. \\
&\quad \left. + \int_t^T e^{r(T-u)} [\theta^*(u) \sigma_1 dW_1(u) - \Pi_{s_2}(u) S_2(u) \sigma_2 dW_2(u)] \right),
\end{aligned}$$

which is independent of  $x$  because  $\theta^*$  does not depend on  $x$ . Therefore, equation (2.21) holds, since

$$F(T, s_1, s_2, x) = n(T, s_2, x) + \tilde{U}(T, s_1, s_2; \theta^*) = (x - \Pi(T, s_2)) + \eta(T, s_1, s_2) = x - G(s_2).$$

Furthermore, equation (2.22), in conjunction with equation (2.43), implies

$$\begin{aligned} g(t, s_1, s_2, x) = \mathbb{E}_{t, s_1, s_2, x}[V^{\theta^*}(T)] &= n(t, s_2, x) + \mathbb{E}_{t, s_1, s_2, x}\left(\int_t^T e^{r(T-u)}\theta^*(u)(\mu_1 - r)du\right) \\ &= n(t, s_2, x) + \eta(t, s_1, s_2), \end{aligned} \quad (2.48)$$

where the last equality follows from Lemma 2.1.

It remains to check equations (2.19) and (2.20). Regarding (2.20), we notice that  $g(t, s_1, s_2, x) = \mathbb{E}_{t, s_1, s_2, x}[V^{\theta^*}(T)]$  to obtain

$$\begin{aligned} &\mathcal{A}^{\theta^*} F(t, s_1, s_2, x) - \xi^{\theta^*}(g(t, s_1, s_2, x)) \\ &= \mathcal{A}^{\theta^*} U(t, s_1, s_2, x; \theta^*) - \xi^{\theta^*}(g(t, s_1, s_2, x)) \\ &= \mathcal{A}^{\theta^*} \left( g(t, s_1, s_2, x) - \frac{\gamma}{2} \mathbb{E}_{t, s_1, s_2, x}[V^{\theta^*}(T)^2] + \frac{\gamma}{2} g^2(t, s_1, s_2, x) \right) - \xi^{\theta^*}(g(t, s_1, s_2, x)), \end{aligned}$$

where the last equality is due to equation (2.15). Note that  $g(t, S_1(t), S_2(t), X^{\theta^*}(t))$  is a martingale. Also, if we denote  $h(t, s_1, s_2, x) = \mathbb{E}[V^{\theta^*}(T)^2]$ , then  $h(t, S_1(t), S_2(t), X^{\theta^*}(t))$  is also a martingale. Therefore,  $\mathcal{A}^{\theta^*} g(t, s_1, s_2, x) = \mathcal{A}^{\theta^*} h(t, s_1, s_2, x) = 0$ , and it follows from the definition of  $\xi^{\theta^*}$  that

$$\begin{aligned} &\mathcal{A}^{\theta^*} F(t, s_1, s_2, x) - \xi^{\theta^*}(g(t, s_1, s_2, x)) \\ &= \frac{\gamma}{2} \mathcal{A}^{\theta^*} (g^2(t, s_1, s_2, x)) - \gamma g(t, s_1, s_2, x) \mathcal{A}^{\theta^*} (g(t, s_1, s_2, x)) - \xi^{\theta^*}(g(t, s_1, s_2, x)) \\ &= 0, \end{aligned}$$

which implies condition (2.20).

Finally, we verify equation (2.19). By the definition of  $\xi^q$ , we obtain

$$\begin{aligned} \xi^q(g(t, s_1, s_2, x)) &= \frac{\gamma}{2} \left[ g_x^2(q\sigma_1)^2 + g_{s_1}^2(s_1\sigma_1)^2 + g_{s_2}^2(s_2\sigma_2)^2 \right. \\ &\quad \left. + 2g_{s_1}g_x s_1 q \sigma_1^2 + 2g_{s_2}g_x s_2 q \sigma_1 \sigma_2 \rho + 2g_{s_1}g_{s_2} s_1 s_2 \sigma_1 \sigma_2 \rho \right]. \end{aligned}$$

Therefore, we can use equation (2.47) to rewrite the right-hand-side of equation (2.19) as follows:

$$\mathcal{A}^q F(t, s_1, s_2, x) - \xi^q(g(t, s_1, s_2, x)) = \mathcal{A}^q \tilde{U}(t, s_1, s_2) - \frac{\gamma}{2}(a_0 q^2 + a_1 q + a_2), \quad (2.49)$$

where

$$\begin{cases} a_0 = g_x^2 \sigma_1^2, \\ a_1 = 2g_{s_1} g_x s_1 \sigma_1^2 + 2g_{s_2} g_x s_2 \sigma_1 \sigma_2 \rho - \frac{2}{\gamma} e^{r(T-t)} (\mu_1 - r), \\ a_2 = g_{s_1}^2 (s_1 \sigma_1)^2 + g_{s_2}^2 (s_2 \sigma_2)^2 + 2g_{s_1} g_{s_2} s_1 s_2 \sigma_1 \sigma_2 \rho \\ \quad - \frac{2}{\gamma} [e^{r(T-t)} r x + n_{s_2} \mu_2 s_2 + \frac{1}{2} n_{s_2 s_2} (s_2 \sigma_2)^2]. \end{cases}$$

Maximizing (2.49) with respect to  $q$  and using equations (2.48) and (2.27) yield

$$\begin{aligned} & \arg \max_q \{ \mathcal{A}^q F(t, s_1, s_2, x) - \xi^q(g(t, s_1, s_2, x)) \} \\ &= -\frac{a_1}{2a_0} \\ &= \frac{e^{r(T-t)} (\mu_1 - r) - \gamma g_{s_1} g_x s_1 \sigma_1^2 - \gamma g_{s_2} g_x s_2 \sigma_1 \sigma_2 \rho}{\gamma g_x^2 \sigma_1^2} \\ &= e^{-r(T-t)} \left[ \frac{\mu_1 - r}{\gamma \sigma_1^2} - \eta_{s_1}(t) s_1 - \frac{s_2 \sigma_2 \rho}{\sigma_1} (\eta_{s_2}(t) - e^{r(T-t)} \Pi_{s_2}(t)) \right] \\ &= \theta^*(t, s_1, s_2). \end{aligned}$$

This confirms equation (2.19) and the proof is complete.  $\square$

## 2.4 Discussions

In this section, we provide additional analysis on some special cases for the general results derived in the preceding section. In particular, optimal trading strategies for variance minimization and/or the absence of basis risk are shown to be special cases of the general results. We shall use the notation  $Y \stackrel{\text{ind.}}{=} Z$  for two processes  $Y$  and  $Z$  to denote that they are indistinguishable, i.e.,  $\mathbb{P}(Y(t) = Z(t), t \in [0, T]) = 1$ .

### 2.4.1 The case with no basis risk

A natural question one may ask is that under what conditions our problem degenerates to the case where the two stocks are perfectly correlated and there is no basis risk. It turns out that the case with no basis risk is indeed one special case of our problem (2.10), as shown in Proposition 2.2 below.

**Proposition 2.2.** *In equation (2.1), if we let  $\rho = 1$ ,  $\mu_1 \stackrel{\text{ind.}}{=} \mu_2 \stackrel{\text{ind.}}{=} \mu$  and  $\sigma_1 \stackrel{\text{ind.}}{=} \sigma_2 \stackrel{\text{ind.}}{=} \sigma$  for some progressively measurable processes  $\mu$  and  $\sigma$ , then the two stochastic processes of stock price  $S_1(t)$  and  $S_2(t)$  are indistinguishable and therefore they can be viewed as the same stock.*

*Further, the equilibrium solution in this case is given by:*

$$\theta^*(t, s) = s \cdot \Pi_s(t) + e^{-r(T-t)} \left[ \frac{\mu - r}{\gamma \sigma^2} - s \eta_s(t) \right], \quad (2.50)$$

and

$$\eta(t, s) = \mathbb{E}_{t,s}^* \left\{ \int_t^T \left[ \frac{1}{\gamma} \left( \frac{\mu - r}{\sigma} \right)^2 + (\mu - r) S(u) e^{r(T-u)} \Pi_s(u) \right] du \right\}, \quad (2.51)$$

where  $S(\cdot)$  is a progressively measurable process such that  $S \stackrel{\text{ind.}}{=} S_1 \stackrel{\text{ind.}}{=} S_2$ .

*Proof.* When  $\rho = 1$ ,  $\forall t_1, t_2 \in (0, T)$ ,

$$(W_1(t_2) - W_1(t_1), W_2(t_2) - W_2(t_1)) \sim N \left( (0, 0), \begin{bmatrix} t_2 - t_1 & t_2 - t_1 \\ t_2 - t_1 & t_2 - t_1 \end{bmatrix} \right).$$

Therefore  $(W_1(t_2) - W_1(t_1), W_2(t_2) - W_2(t_1))$  is a degenerate bivariate normal random variable with  $W_1(t_2) - W_1(t_1) = W_2(t_2) - W_2(t_1)$  a.s., and thus  $\forall t \in (0, T)$ ,  $W_1(t) = W_2(t)$  a.s., which along with the fact that both  $W_1$  and  $W_2$  have continuous paths almost surely, implies that  $W_1(t)$  and  $W_2(t)$  are indistinguishable, and so are  $S_1(t)$  and  $S_2(t)$ . Consequently, equation (2.50) and equation (2.51) are obtained trivially from equations (2.27) and (2.28).  $\square$

## 2.4.2 The limiting case when $\gamma \rightarrow \infty$

When the risk aversion coefficient  $\gamma$  in problem (2.10) becomes larger, this implies that the hedger is more risk averse and is more concerned with the variability of his/her hedging strategy. In the limiting case of  $\gamma \rightarrow \infty$ , the hedger can be perceived as one who is pre-dominantly concerned with the variability of the adopted hedging strategy and hence his/her objective boils down to minimizing  $\text{Var}_{t,s_1,s_2,x}[V^\theta(T)]$ . In this special case, it is of interest to investigate if the equilibrium solution given by equation (2.27) reduces to that of the variance minimization problem. The answer is affirmative as justified by the proposition below.

**Proposition 2.3.** *Denote  $\theta_\gamma^*$  as the equilibrium solution given in equation (2.27) to emphasize its dependence on the risk aversion coefficient  $\gamma$  in problem (2.10), and let*

$$\theta_0^*(t, s_1, s_2) = \lim_{\gamma \rightarrow \infty} \theta_\gamma^*(t, s_1, s_2) = e^{-r(T-t)} \left[ -\eta_{s_1}(t) s_1 - \frac{s_2 \sigma_2 \rho}{\sigma_1} (\eta_{s_2}(t) - e^{r(T-t)} \Pi_{s_2}(t)) \right] \quad (2.52)$$

and

$$\eta(t, s_1, s_2) = \mathbb{E}_{t,s_1,s_2}^* \left\{ \int_t^T \left[ (\mu_1 - r) \frac{\rho \sigma_2}{\sigma_1} S_2(u) e^{r(T-u)} \Pi_{s_2}(u) \right] du \right\}. \quad (2.53)$$

Then,  $\theta_0^*$  defined in (2.52) is an equilibrium solution to the following variance minimization problem:

$$\max_{\theta \in \Theta} \{ U(t, s_1, s_2, x; \theta) := -\text{Var}_{t,s_1,s_2,x}[V^\theta(T)] \}. \quad (2.54)$$

*Proof.* The technique used to derive an equilibrium solution of problem (2.54) parallels to that of problem (2.10) in Section 2.3, hence we only highlight some key steps of the proof. First, by the total variance formula, the objective function of problem (2.54) satisfies the recursion:

$$\begin{aligned} U(t, s_1, s_2, x; \theta) &= -\text{Var}_{t,s_1,s_2,x}[V^\theta(T)] \\ &= -\mathbb{E}_{t,s_1,s_2,x}[\text{Var}_{t+\tau}(V^\theta(T))] - \text{Var}_{t,s_1,s_2,x}[\mathbb{E}_{t+\tau}(V^\theta(T))] \\ &= \mathbb{E}_{t,s_1,s_2,x}[U^\theta(t+\tau)] - \text{Var}_{t,s_1,s_2,x}[\mathbb{E}_{t+\tau}(V^\theta(T))]. \end{aligned} \quad (2.55)$$

Second, in parallel to equations (2.45) and (2.46),

$$\begin{aligned}
& U(t, s_1, s_2, x; \theta) \\
= & -\text{Var}_{t, s_1, s_2, x} \left( \int_t^T e^{r(T-u)} \theta(u) (\mu_1 - r) du \right. \\
& \left. + \int_t^T e^{r(T-u)} [\theta(u) \sigma_1 dW_1(u) - \Pi_{s_2}(u) S_2(u) \sigma_2 dW_2(u)] \right).
\end{aligned}$$

Third, by a similar argument as in equation (2.16), we can use the recursion (2.55) to establish the following extended HJB equation for equilibrium solution  $\theta^*$  to satisfy:

$$0 = \max_{q \in \mathbb{R}} \left( \mathcal{A}^q F(t, s_1, s_2, x) - \xi^q(m(t, s_1, s_2, x; \theta^*)) \right), \quad (2.56)$$

where the generator  $\mathcal{A}^q$  is defined in equation (2.17),  $m(t, s_1, s_2, x; \theta) := \mathbb{E}_{t, s_1, s_2, x}[V^\theta(T)]$  as defined in equation (2.40), and in parallel to equation (2.18),

$$\xi^q(m(t, s_1, s_2, x)) = \mathcal{A}^q [m(t, s_1, s_2, x)^2] - 2m(t, s_1, s_2, x) \mathcal{A}^q [m(t, s_1, s_2, x)].$$

Finally it is straightforward to verify that Theorem 2.1 is still valid with  $\xi^q$  replaced by the above definition. Then following a procedure similar to Theorem 2.2, we can prove  $\theta_0^*$  is an equilibrium solution to problem (2.54).

**Remark 2.3.** *We note that while the equilibrium strategy  $\theta^*$  in Proposition 2.3 is derived from the extended HJB equation (2.56), the expressions of  $\theta^*$  and  $\eta$  given in Proposition 2.3 can be obtained from equations (2.50) and (2.51) by taking  $\gamma \rightarrow \infty$ .*

□

### 2.4.3 The limiting case when $\gamma \rightarrow \infty$ with no basis risk

We consider the case with no basis risk, i.e.,  $S_1$  and  $S_2$  are indistinguishable, for which sufficient conditions are  $\rho = 1$ ,  $\mu_1 \stackrel{\text{ind.}}{=} \mu_2$  and  $\sigma_1 \stackrel{\text{ind.}}{=} \sigma_2$ , as formally proved in Proposition 2.2. Let  $S$  be a

process which is indistinguishable from  $S_1$  and  $S_2$ . So, we can equivalently view the European option as if it is written on  $S$ . By further letting  $\sigma$  be a process indistinguishable from  $\sigma_1$  and  $\sigma_2$ , we obtain from equation (2.30) that, under the probability measure  $\mathbb{P}^*$ ,  $S$  follows a dynamic  $\frac{dS(t)}{S(t)} = rdt + \sigma dW^*(t)$ , where  $W^*$  is a standard Brownian motion under  $\mathbb{P}^*$ . Thus, the delta of the European option at time  $t$  is given by  $\Delta(t, s) = \frac{\partial}{\partial s}\Pi^*(t, s)$ , where

$$\Pi^*(t, s) = \mathbb{E}_{t,s}^* [e^{-r(T-t)}G(S(T))] . \quad (2.57)$$

In this special case with no basis risk, it is well known that a dynamic delta hedging can fully replicate the payoff  $G(S(T))$  of the European option. Therefore, a trading strategy  $\tilde{\theta} = \{\tilde{\theta}(t, s), t \in [0, T] \text{ and } s \in \mathbb{R}_+\}$  with  $\tilde{\theta}(t, s) = s\Delta(t, s)$  is an equilibrium solution to the variance minimization problem (2.54), because the variance attains its minimum value, zero, with such a trading strategy at any time  $t \in [0, T]$ .

Recall that (2.50) of Proposition 2.2 gives an equilibrium solution when there is no basis risk. With  $\gamma \rightarrow \infty$ ,  $\theta^*(t, s)$  in (2.50) reduces to

$$\theta^*(t, s) = e^{-r(T-t)} [-s (\eta_s(t, s) - e^{r(T-t)}\Pi_s(t, s))] , \quad t \in [0, T], \quad s \in \mathbb{R}_+ . \quad (2.58)$$

Therefore, in view of the above analysis, one may expect that  $\theta^*$  given in equation (2.58) is the same as the delta hedging strategy  $\tilde{\theta}$ . Proposition 2.4 below confirms such a conjecture.

**Proposition 2.4.**  $\tilde{\theta}(t, s) = \theta^*(t, s)$  for any  $t \in [0, T]$  and  $s \in \mathbb{R}_+$ , where  $\theta^*$  is given by equation (2.58).

*Proof.* Let us denote  $y(t, s) := \mathbb{E}_{t,s}[G(S(T))] = e^{-r(T-t)}\Pi(t, s)$ ,  $t \in [0, T]$  and  $s \in \mathbb{R}_+$ . Then, by Feynmann-Kac Theorem,

$$y_t + \mu s y_s + \frac{1}{2}\sigma^2 s^2 y_{ss} = 0. \quad (2.59)$$

Since  $\Pi(t, s) = \mathbb{E}_{t,s} [e^{-r(T-t)}G(S(T))]$ , it follows from (2.58) and Lemma 2.1 that

$$\theta^*(t, s) = s \cdot e^{-r(T-t)} [y_s(t, s) - \eta_s(t, s)] ,$$

with

$$\begin{aligned}\eta(t, s) &= \mathbb{E}_{t,s} \left[ \int_t^T e^{r(T-u)} \theta^*(u, S(u)) (\mu - r) du \right] \\ &= \mathbb{E}_{t,s} \left[ \int_t^T s(\mu - r) [y_s(u, S(u)) - \eta_s(u, S(u))] du \right].\end{aligned}$$

Applying Feynman-Kac Theorem again yields

$$\eta_t + \mu s \eta_s + \frac{1}{2} \sigma^2 s^2 \eta_{ss} + s(y_s - \eta_s)(\mu - r) = 0. \quad (2.60)$$

Combining equations (2.59) and (2.60) yields

$$(y - \eta)_t + r s (y - \eta)_s + \frac{1}{2} \sigma^2 s^2 (y - \eta)_{ss} = 0. \quad (2.61)$$

We further define  $y^*(t, s) := \mathbb{E}_{t,s}^* [G(S(T))]$ ,  $t \in [0, T]$  and  $s \in \mathbb{R}_+$ , and use Feynman-Kac Theorem to obtain

$$y_t^* + r s y_s^* + \frac{1}{2} \sigma^2 s^2 y_{ss}^* = 0. \quad (2.62)$$

Note that  $y^*(t, s)$  and  $y(t, s) - \eta(t, s)$  satisfy the same PDE with the same boundary condition, i.e.,  $y(T, s) - \eta(T, s) = G(S(T)) = y^*(T, s)$ . Thus,  $y^*(t, s) = y(t, s) - \eta(t, s)$  and  $y_s^*(t, s) = y_s(t, s) - \eta_s(t, s)$ . We further note  $\hat{\theta}(t, s) = s \Delta(t, s) = e^{-r(T-t)} y_s^*(t, s)$  and  $\theta^* = e^{-r(T-t)} [y_s(t, s) - \eta_s(t, s)]$  to conclude  $\hat{\theta}(t, s) = \theta^*(t, s)$ .  $\square$

## 2.4.4 Solutions under geometric Brownian motions

In this subsection, we investigate the hedging problem of two specific contingent claims – futures contract and the European call option – under the assumption that both asset price processes  $S_1$  and  $S_2$  are geometric Brownian motions. Under the assumption that  $S_1$  is a geometric Brownian motion, the coefficients  $\mu_1$  and  $\sigma_1$  are constant and the function  $\eta$  is independent of  $s_1$ . In this special case, we suppress the notation  $\eta(t, s_1, s_2)$  into  $\eta(t, s_2)$  throughout the section. Similarly

for the optimal hedging strategy  $\theta^*$  that only depends on time  $t$  and the price of asset  $S_2$ , and we use the simplified notation  $\theta^*(t, s_2)$ . Further, we denote

$$\mu_2^* = \mu_2 - (\mu_1 - r) \frac{\rho\sigma_2}{\sigma_1}, \quad (2.63)$$

which is the drift coefficient of  $S_2$  under the measure  $P^*$  as given in equation (2.30).

#### 2.4.4.1 Futures contract

In this section, we assume the hedging target is a future contract. The price of a future contract may depend on a variety of factors, such as the dividend payments of the underlying, the storage cost or the cost to carry for commodity-based futures contracts, and it may not converge to the spot underlying asset price at expiration. In this section, we do not consider those factors and assume for simplicity that the payoff of a futures contract is the same as the spot asset price at expiration so that  $G(S_2(T)) = S_2(T) - K$ . Hence, we have

$$\Pi(t, s_2) = E_{t, s_2} [e^{-r(T-t)} (S_2(T) - K)] = s_2 e^{(\mu_2 - r)(T-t)} - K e^{-r(T-t)},$$

and

$$E_{t, s_2}^* [S_2(u) \Pi_{s_2}(u)] = E_{t, s_2}^* [S_2(t) e^{(\mu_2 - r)(T-t)}] = s_2 e^{\mu_2^*(u-t)} \cdot e^{(\mu_2 - r)(T-t)}.$$

Substituting these into equation (2.28) yields

$$\eta(t, s_2) = \int_t^T \left[ \frac{1}{\gamma} \left( \frac{\mu_1 - r}{\sigma_1} \right)^2 - (\mu_1 - r) \frac{\rho\sigma_2}{\sigma_1} s_2 e^{\mu_2^*(u-t) + \mu_2(T-u)} \right] du,$$

and

$$\eta_{s_2} = \int_t^T \left[ (\mu_1 - r) \frac{\rho\sigma_2}{\sigma_1} e^{\mu_2^*(u-t) + \mu_2(T-u)} \right] du = e^{\mu_2(T-t)} - e^{\mu_2^*(T-t)}.$$

With the expressions of the above quantities, it is easy to check that condition **H** in Remark 2.2 is satisfied. Therefore, from equation (2.27) we obtain the optimal hedging strategy as follows

$$\theta^*(t, s_2) = \frac{\mu_1 - r}{\gamma(\sigma_1)^2} e^{-r(T-t)} + \frac{\rho\sigma_2}{\sigma_1} s_2 e^{(\mu_2^* - r)(T-t)}, \quad 0 \leq t \leq T. \quad (2.64)$$

Equilibrium value function  $J(t, s_1, s_2, x)$  does not have an explicit form, and thus need to be calculated numerically. A semi closed-form expression is given in Appendix A.2.

#### 2.4.4.2 European call option

Recall that the key contribution of this chapter is to provide an analytical optimal strategy to hedge derivative security. In this section, we provide an in-depth analysis by considering hedging a European call option under the Black-Scholes framework. The payoff of a European call option at maturity  $T$  is given by  $G(S_2(T)) = (S_2(T) - K)^+$ , where  $K$  is the pre-determined strike price and  $(x)^+ = \max(x, 0)$ . From the Black-Scholes formula, it can be shown that

$$\begin{aligned} \Pi(t, s_2) &= \mathbb{E}_{t, s_2} [e^{-r(T-t)} (S_2(T) - K)^+] \\ &= e^{(\mu_2 - r)(T-t)} \mathbb{E}_{t, s_2} [e^{-\mu_2(T-t)} (S_2(T) - K)^+] \\ &= e^{(\mu_2 - r)(T-t)} [s_2 \Phi(d_{1,t}) - K e^{-\mu_2(T-t)} \Phi(d_{2,t})], \end{aligned}$$

where  $\Phi(\cdot)$  is the standard normal distribution function, and

$$\begin{cases} d_{1,t} = \frac{\ln(s_2/K) + (\mu_2 + \frac{1}{2}\sigma_2^2)(T-t)}{\sigma_2\sqrt{T-t}}, \\ d_{2,t} = d_{1,t} - \sigma_2\sqrt{T-t} = \frac{\ln(\frac{s_2}{K}) + (\mu_2 - \frac{1}{2}\sigma_2^2)(T-t)}{\sigma_2\sqrt{T-t}}. \end{cases}$$

Hence

$$\Pi_{s_2} = \frac{\partial}{\partial s_2} \left( e^{(\mu_2 - r)(T-t)} [s_2 \Phi(d_{1,t}) - e^{-\mu_2(T-t)} K \Phi(d_{2,t})] \right)$$

$$\begin{aligned}
&= e^{(\mu_2-r)(T-t)} \frac{\partial}{\partial s_2} [s_2 \Phi(d_{1,t}) - e^{-\mu_2(T-s)} K \Phi(d_{2,t})] \\
&= e^{(\mu_2-r)(T-t)} \Phi(d_{1,t}).
\end{aligned} \tag{2.65}$$

By denoting

$$c(u, s_2) := \frac{\ln s_2 + (\mu_2^* - \frac{1}{2}\sigma_2^2)(u-t) - \ln K + (\mu_2 + \frac{\sigma_2^2}{2})(T-u)}{\sigma_2\sqrt{T-u}},$$

we obtain, for  $0 \leq t < u$ ,

$$\begin{aligned}
&E_{t,s_2}^* [S_2(u)\Pi_{s_2}(u)] \\
&= E_{t,s_2}^* \left[ S_2(u) e^{(\mu_2-r)(T-u)} \Phi \left( \frac{\ln(S_2(u)/K) + (\mu_2 + \frac{\sigma_2^2}{2})(T-u)}{\sigma_2\sqrt{T-u}} \right) \right] \\
&= e^{(\mu_2-r)(T-u)} \int_{-\infty}^{\infty} s_2 e^{(\mu_2^* - \frac{1}{2}\sigma_2^2)(u-t) + \sqrt{u-t}x} \Phi \left( \frac{\sqrt{u-t}}{\sigma_2\sqrt{T-u}} x + c(u, s_2) \right) \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx \\
&= s_2 e^{(\mu_2-r)(T-u)} e^{(\mu_2^* - \frac{\sigma_2^2}{2})(u-t)} \int_{-\infty}^{\infty} \left( \frac{1}{\sqrt{2\pi}} \right)^2 e^{\sqrt{u-t}x} e^{-\frac{x^2}{2}} \int_{-\infty}^{\frac{\sqrt{u-t}}{\sigma_2\sqrt{T-u}}x + c(u, s_2)} e^{-\frac{z^2}{2}} dz dx \\
&= s_2 e^{(\mu_2-r)(T-u)} e^{(\mu_2^* - \frac{\sigma_2^2}{2})(u-t)} e^{\frac{1}{2}(u-t)} \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{\sqrt{u-t}}{\sigma_2\sqrt{T-u}}x + c(u, s_2)} \frac{1}{2\pi} e^{-\frac{(x-\sqrt{u-t})^2 + z^2}{2}} dz dx \\
&= s_2 e^{(\mu_2-r)(T-u)} e^{(\mu_2^* - \frac{\sigma_2^2}{2})(u-t)} e^{\frac{1}{2}(u-t)} \mathbb{P} \left[ Z \leq \frac{\sqrt{u-t}}{\sigma_2\sqrt{T-u}} (X + \sqrt{u-t}) + c(u, s_2) \right],
\end{aligned}$$

where, under the probability measure  $\mathbb{P}$ ,  $X$  and  $Z$  are two independent standard normal variables so that  $\tilde{Z} := Z - \frac{\sqrt{u-t}}{\sigma_2\sqrt{T-u}}X$  is also a normal variable with  $E[\tilde{Z}] = 0$  and  $\text{Var}[\tilde{Z}] = \frac{\sigma_2^2(T-u) + (u-t)}{\sigma_2^2(T-u)}$ . Consequently,

$$\begin{aligned}
&E_{t,s_2}^* [S_2(u)\Pi_{s_2}(u)] \\
&= s_2 e^{(-\mu_2+r+\mu_2^* - \frac{\sigma_2^2}{2} + \frac{1}{2})u} e^{(\mu_2-r)T} e^{(-\mu_2^* + \frac{\sigma_2^2}{2} - \frac{1}{2})t} \Phi \left( \frac{(u-t) + c(u, s_2)\sigma_2\sqrt{T-u}}{\sqrt{(u-t) + \sigma_2^2(T-u)}} \right). \tag{2.66}
\end{aligned}$$

Using equation (2.28), we have

$$\eta(t, s_2) = \frac{1}{\gamma} \left( \frac{\mu_1 - r}{\sigma_1} \right)^2 (T - t) + (\mu_1 - r) \frac{\rho \sigma_2}{\sigma_1} e^{\mu_2 T} e^{-(\mu_2^* - \frac{\sigma_2^2}{2} + \frac{1}{2})t} \cdot s_2 \cdot \int_t^T e^{(\mu_2^* - \frac{\sigma_2^2}{2} + \frac{1}{2} - \mu_2)u} \Phi \left( \frac{(u - t) + c(u, s_2) \sigma_2 \sqrt{T - u}}{\sqrt{(u - t) + \sigma^2(T - u)}} \right) du,$$

and

$$\begin{aligned} \eta_{s_2} &= (\mu_1 - r) \frac{\rho \sigma_2}{\sigma_1} e^{\mu_2 T} e^{-(\mu_2^* - \frac{\sigma_2^2}{2} + \frac{1}{2})t} \int_t^T e^{(\mu_2^* - \frac{\sigma_2^2}{2} + \frac{1}{2} - \mu_2)u} \Phi \left( \frac{(u - t) + c(u, s_2) \sigma_2 \sqrt{T - u}}{\sqrt{(u - t) + \sigma_2^2(T - u)}} \right) du \\ &\quad + (\mu_1 - r) \frac{\rho \sigma_2}{\sigma_1} e^{\mu_2 T} s_2 e^{-(\mu_2^* - \frac{\sigma_2^2}{2} + \frac{1}{2})t} \cdot \int_t^T e^{(\mu_2^* - \frac{\sigma_2^2}{2} + \frac{1}{2} - \mu_2)u} \phi \left( \frac{(u - t) + c(u, s_2) \sigma_2 \sqrt{T - u}}{\sqrt{(u - t) + \sigma_2^2(T - u)}} \right) \frac{\sigma_2 \sqrt{T - u}}{\sqrt{(u - t) + \sigma_2^2(T - u)}} \frac{\partial c(u, s_2)}{\partial s_2} du \\ &= (\mu_1 - r) \frac{\rho \sigma_2}{\sigma_1} e^{\mu_2 T} e^{-(\mu_2^* - \frac{\sigma_2^2}{2} + \frac{1}{2})t} \cdot \int_t^T e^{(\mu_2^* - \frac{\sigma_2^2}{2} + \frac{1}{2} - \mu_2)u} \cdot \left[ \Phi \left( \frac{(u - t) + c(u, s_2) \sigma_2 \sqrt{T - u}}{\sqrt{(u - t) + \sigma_2^2(T - u)}} \right) + \phi \left( \frac{(u - t) + c(u, s_2) \sigma_2 \sqrt{T - u}}{\sqrt{(u - t) + \sigma_2^2(T - u)}} \right) \frac{1}{\sqrt{(u - t) + \sigma_2^2(T - u)}} \right] du, \end{aligned}$$

where  $\phi$  is the standard normal density function.

Based on the expression of  $\Pi_{s_2}$  given in (2.65), we can further derive an expression for  $\Pi_{s_2 s_2}$  and show that the polynomial bounded condition in equation (2.34) is satisfied. Further, based on equation (2.66) and noticing that the normal distribution function  $\Pi(x) \leq 1$  for any  $x \in \mathbb{R}$ , we can easily verify the condition in equation (2.33). Combining these implies that condition **H** in Remark 2.2 is satisfied. Therefore, we plug the corresponding expressions developed in the above into equation (2.27) to get the equilibrium solution as follows

$$\theta^*(t, s_2) = e^{-r(T-t)} \frac{\mu_1 - r}{\gamma \sigma_1^2} - e^{-r(T-t)} \frac{\rho \sigma_2}{\sigma_1} s_2 \eta_{s_2} + \frac{\rho \sigma_2}{\sigma_1} s_2 \Pi_{s_2}$$

$$\begin{aligned}
&= e^{-r(T-t)} \frac{\mu_1 - r}{\gamma \sigma_1^2} + \frac{\rho \sigma_2}{\sigma_1} s_2 e^{(\mu_2 - r)(T-t)} \Phi(d_{1,t}) \\
&\quad - e^{-(\mu_2^* - \frac{\sigma_2^2}{2} + \frac{1}{2} - r)t} e^{(\mu_2 - r)T} \left( \frac{\rho \sigma_2}{\sigma_1} \right)^2 (\mu_1 - r) s_2 \\
&\quad \cdot \int_t^T e^{(\mu_2^* - \frac{\sigma_2^2}{2} + \frac{1}{2} - \mu_2)u} \left[ \Phi(d_{*,u}) + \frac{\phi(d_{*,u})}{\sqrt{(u-t) + \sigma_2^2(T-u)}} \right] du, \quad (2.67)
\end{aligned}$$

where

$$d_{*,u} = \frac{(u-t) + c(u, s_2) \sigma_2 \sqrt{T-u}}{\sqrt{(u-t) + \sigma_2^2(T-u)}}.$$

Similar to futures contracts, equilibrium value function  $J(t, s_1, s_2, x)$  does not have an explicit form, and a semi closed-form expression is given in Appendix A.3.

## 2.5 Numerical examples

Based on the consistent planning equilibrium strategy derived in Subsection 2.4.4.2 for hedging a short position in European call option, this section provides some numerical evidences to highlight the importance of the proposed equilibrium hedging strategy. Throughout the numerical examples, we assume that the strike price of the European call option is  $K = 100$  with  $T = 1$  year time to maturity, and that both asset prices follow geometric Brownian motions with parameter values given in Table 2.1. Furthermore, the initial hedging cost is consistently set at  $x_0 = 20$ .

Table 2.1: Parameter values for the hedging examples

$S_1(0)$	$S_2(0)$	$K$	$r$	$\mu_1$	$\sigma_1$	$\mu_2$	$\sigma_2$	$T$	$x_0$
100	100	100	0.05	0.1	0.25	0.12	0.3	1	20

For our first set of comparison, we numerically evaluate the performance of our proposed mean-variance equilibrium strategy against two other hedging strategies. The first strategy we

benchmark against is known as the “naive delta hedge”, which is defined by

$$\theta_{naive} = \frac{\sigma_2 S_2(t)}{\sigma_1} \frac{\partial}{\partial s} BS(S_2(t)),$$

where  $BS(\cdot)$  denotes the Black-Scholes price and  $\frac{\partial}{\partial s} BS(\cdot)$  denotes the Black-Scholes delta. We refer  $\theta_{naive}$  as a naive delta hedge because such a strategy ignores the basis risk and applies delta hedge dynamically based on  $S_1$  (i.e. assuming  $S_1$  is the underlying asset). The naive delta hedge is not time consistent in terms of optimizing the mean-variance objective. The second strategy we benchmark against is simply the “no hedging” strategy, which merely investing the initial hedging amount of  $x_0 = 20$  in the risk-free bond to earn the risk-free rate of  $r = 5\%$ .

For each of the above strategies, we compute the mean and variance of the terminal wealth, along with the mean-variance objective value  $E[V(T)] - \frac{\gamma}{2}\text{Var}[V(T)]$ . These are depicted in Table 2.2 by assuming the risk aversion coefficient  $\gamma = 1$ . For both equilibrium and naive delta hedging strategies, we further assume that the correlation parameter  $\rho$  increases from 0.5 to 1, at increment of 0.1. Based on these results, it is clear that simply investing in the risk-free bond is an inadequate strategy, as can be seen from the unacceptable large negative value of  $|E[V(T)] - \frac{\gamma}{2}\text{Var}[V(T)]|$ . On the other hand, in the extreme case with  $\rho = 1$ , both equilibrium strategy and naive delta strategy are competitively effective for hedging the European call option. This is supported by the negligible variance of the terminal wealth of the hedger. The advantage of the equilibrium strategy becomes more pronounced as the correlation increases so that basis risk becomes less prominent. As  $\rho$  decreases, the variances of the terminal wealth of both strategies increase sharply (from perfect correlation case) but nevertheless the mean-variance objective values of the equilibrium strategy are consistently higher than the corresponding values from the naive delta hedging strategy, hence indicating the importance of taking into account the basis risk and the superiority of the equilibrium hedging strategy.

In the second set of comparison, we investigate the effect of risk aversion coefficient  $\gamma$  on the performance of the equilibrium hedging strategy and the naive delta hedging strategy. We similarly use the preceding example except by considering

$$\rho \in \{0.9, 1\} \text{ and } \gamma \in \{1/8, 1/4, 1/2, 1, 2, 4, \infty\}.$$

Table 2.2: Comparison of different strategies on hedging European call option ( $\gamma = 1$ )

		$E[V(T)]$	$\text{Var}[V(T)]$	$E[V(T)] - \frac{\gamma}{2}\text{Var}[V(T)]$
$\rho = 1$	equilibrium	5.49	0.91	5.04
	naive delta	5.39	1.14	4.82
$\rho = 0.9$	equilibrium	4.72	131.99	-61.27
	naive delta	5.19	137.46	-63.54
$\rho = 0.8$	equilibrium	4.20	249.89	-120.74
	naive delta	5.21	270.24	-129.91
$\rho = 0.7$	equilibrium	3.70	356.14	-174.38
	naive delta	5.25	402.84	-196.17
$\rho = 0.6$	equilibrium	3.20	449.59	-221.59
	naive delta	5.31	534.65	-262.02
$\rho = 0.5$	equilibrium	2.71	529.94	-262.26
	naive delta	5.36	664.97	-327.13
	no hedge	0.08	799.77	-399.81

The numerical results are reported in Tables 2.3 and 2.4 for  $\rho = 1$  and  $\rho = 0.9$ , respectively. Note that as shown in Proposition 2.3, the equilibrium strategy reduces to a solution for variance minimization as  $\gamma \rightarrow \infty$ .

The advantage of the equilibrium hedging strategy is highlighted in this example in that it is capable of reflecting the risk aversion of the hedger. When the hedger has a higher risk tolerance, he/she seeks an optimal strategy that has a higher mean value of the terminal wealth, though at the expense of higher terminal wealth variability. When the hedger becomes more and more risk averse, a greater penalty is imposed on the variability of the terminal wealth, which in turn also dampens the expected value of the terminal value. In contrast, the naive delta hedging strategy is invariant to the risk aversion of the hedger and hence produces the same set of mean and variance of the terminal wealth, irrespective of  $\gamma$ . Regardless of the risk aversion of the hedger, the mean-variance trade-off value of the equilibrium hedging strategy is consistently higher than the corresponding value from the naive delta hedging strategy, indicating the superiority of the former strategy. In the special case with  $\rho = 1$  and irrespective of  $\gamma$ , both hedging strategies are very competitive and effective, as signaled by the small variance of the terminal wealth. This observation is consistent with the earlier example.

Table 2.3: Hedging European call with  $\rho = 1$ : “equilibrium solution” vs “naive delta hedge”.

		$E[V(T)]$	$\text{Var}[V(T)]$	$E[V(T)] - \frac{\rho}{2}\text{Var}[V(T)]$
naive delta	$\gamma = 1/8$	5.387	1.140	5.316
	$\gamma = 1/4$			5.245
	$\gamma = 1/2$			5.102
	$\gamma = 1$			4.817
	$\gamma = 2$			4.247
	$\gamma = 4$			3.107
	$\gamma \rightarrow \infty$			N/A
equilibrium	$\gamma = 1/8$	5.802	4.493	5.521
	$\gamma = 1/4$	5.626	1.947	5.383
	$\gamma = 1/2$	5.539	1.174	5.245
	$\gamma = 1$	5.495	0.913	5.038
	$\gamma = 2$	5.473	0.814	4.659
	$\gamma = 4$	5.462	0.772	3.918
	$\gamma \rightarrow \infty$	5.451	0.736	N/A

Table 2.4: Hedging European call with  $\rho = 0.9$ : “equilibrium solution” vs “naive delta hedge”.

		$E[V(T)]$	$\text{Var}[V(T)]$	$E[V(T)] - \frac{\rho}{2}\text{Var}[V(T)]$
naive delta	$\gamma = 1/8$	5.187	137.462	-3.404
	$\gamma = 1/4$			-11.995
	$\gamma = 1/2$			-29.178
	$\gamma = 1$			-63.544
	$\gamma = 2$			-132.274
	$\gamma = 4$			-269.736
	$\gamma \rightarrow \infty$			N/A
equilibrium	$\gamma = 1/8$	5.009	134.731	-3.412
	$\gamma = 1/4$	4.846	132.632	-11.733
	$\gamma = 1/2$	4.765	132.115	-28.264
	$\gamma = 1$	4.725	131.989	-61.270
	$\gamma = 2$	4.704	131.960	-127.256
	$\gamma = 4$	4.694	131.953	-259.213
	$\gamma \rightarrow \infty$	4.684	131.952	N/A

## 2.6 Conclusion

The optimal dynamic hedging for European-style derivatives is studied in this chapter in the presence of basis risk where the underlying asset of the option is not traded in the market and is hedged by a traded asset. Under a diffusion model setup, analytical hedging strategy is obtained from optimizing a mean-variance criterion and resorting to the Nash subgame equilibrium framework. The derivation is based on an extended HJB equation and change-measure techniques. The existent literature usually either focus on the hedging of futures contracts or follow an exponential preference optimization framework for mathematical convenience.

In contrast, in this chapter formal analysis is provided for the mean-variance optimal hedging strategy on hedging general European-style derivatives. The optimal hedging strategies in the absence of basis risk and/or for variance minimization can be recovered, as special cases, from the general results that were established in this chapter.

## **Chapter 3**

# **Optimal Dynamic Longevity Hedge with Basis Risk**

### **3.1 Introduction**

Over the last 100 years, life expectancies have increased at the rate of approximately 2.5 years per decade. What this means is that you are expected to live 2.5 years longer than someone who was born 10 years before you were. While the improvement of our life expectancy is one of the greatest achievements in mankind and is to be celebrated, the unanticipated mortality improvements can have an undesirable effect on the society. More specifically, because of longevity, that is, we are living longer, this creates additional (and significant) financial burden to individuals, corporations and governments, as attributed to the greater retirement cost and medical cost, among others. The challenges with managing longevity risk, that is, the uncertainty associated with future mortality improvements, stem from the fact that it is a systematic risk, which means it cannot be mitigated via the typical diversification strategies. As such longevity risk has become a high profile risk in recent years. For example, in the International Monetary Fund (IMF) Global Financial Stability Report, which has received considerable attention since it was published in April 2012, it was demonstrated that if individuals live three years longer than expected, “the already large costs of aging could increase by another 50 percent, representing an additional cost

of 50 percent of 2010 GDP in advanced economies and 25 percent of 2010 GDP in emerging economies.

The corporate pension plan sponsors and the annuity providers are similarly facing the adverse financial effect of the longevity risk and raise considerable concern on their sustainability and the viability. Hence these stakeholders (or hedgers) are constantly seeking for more effective longevity risk management solutions. Some of the traditional approaches include pension buy-in, pension buy-out, and reinsurance and these are often considered as customization strategies in that actual longevity risk is effectively transferred to a third party. Another solution that has been advocated is via the capital market whereby hedgers hedge their longevity risk using standardized longevity securities that are linked to some certain longevity indices. While the former solution based on customization is undoubtedly more effective, it tends to be more costly. The latter longevity index-based solution, on the other hand, can be more cost effective. The downside of this strategy is the presence of basis risk, that is, the mismatch between the mortality experience underlies the standard longevity securities and the hedger's own mortality experience. The basis risk, therefore, diminishes the effectiveness of the longevity index-based solutions.

In recent years there have been some advances in addressing the longevity hedge in the presence of basis risk. Most of the existing literature on longevity hedge focus on static hedge, as opposed to dynamic hedge (Li and Hardy, 2011; Li and Luo, 2012; Cairns, 2013). Hedging based on static strategy can be inefficient due to the following reasons. First, it is very vulnerable to future market changes due to its inflexibility, and thus in general the hedging performance is not as good as dynamic strategies. Second, in order to achieve certain financial objectives, the hedger usually has to pay a considerable initial cost (or reserve for future cost), which might be impractical. Third, as pointed out in Zhou and Li (2016), constructing an effective static hedging portfolio generally requires long-dated hedging instruments which are expensive, unappealing to investors and have high counterparty risks. In conclusion, dynamic hedge, if implemented properly, can be a more effective strategy.

It appears that the most commonly used dynamic hedging strategies for longevity risk in the presence of basis risk is the dynamic "delta" hedging method proposed by Cairns (2011) (see also Zhou and Li, 2016). The key principle of the delta hedging method is to construct a hedging

portfolio that matches the sensitivity of the hedger's (usually an annuity provider's or a pension plan sponsor's) future liability with respect to changes in some underlying mortality index. However, it should be emphasized that in the presence of basis risk, there is no guarantee on the effectiveness of the above delta hedging strategy due to the imperfect correlation between the hedging portfolio and the underlying mortality risk. Additionally, performance of the dynamic "delta" hedging method highly depends on choice of assumption about mortality models and in practice it may be very computationally costly for case-by-case analysis.

The dynamic hedging problem formulated as an optimization problem under some certain criterion is a good alternative as it ensures some optimality on the performance of the proposed hedging strategy. Wong et al. (2014, 2015) discussed the optimal dynamic longevity hedging strategy using longevity bonds for an insurer under the mean-variance criterion. They derived a closed-form hedging policy using the Hamilton-Jacobi-Bellman (HJB) framework. A major deficiency of their results is that they used diffusion models for underlying mortality rates and their derivation was conducted in a continuous time setting. However, in the context of longevity risk, discrete time models such as the Lee-Cater model and the Cairns-Blake-Dowd (CBD) model are more popularly adopted. In this chapter we will use discrete time models and aim at obtaining hedging strategies which are more readily applicable in practice.

In this chapter, we also adopt the mean-variance objective and aim to find the optimal dynamic hedging strategy for a pension plan liability with population basis risk involved. We will show that under our assumption we can tackle with the hedging problem using the stochastic optimal control framework. Therefore a Bellman equation is derived to obtain a semi closed-form solution to efficiently calculate the optimal hedging plan for the original hedging problem.

Our work contributes to the existing literature in the following three ways: first, it is the first time to discuss optimal dynamic longevity hedging problem using  $q$ -forward contracts when basis risk is involved, and the results obtained in this chapter provides an effective solution for longevity risk management involving standardized securities; second, by numerical examples we can show that the hedging strategy we propose in this chapter outperforms the benchmark "delta" method in terms of reducing the hedger's basis risk, while still retaining a practical computational effort; third, we provide a very general framework for dynamic longevity hedging problems,

and therefore generalization of this chapter to other longevity models, hedging instruments and hedging structures is highly probable in the future.

The rest of this chapter is organized as follows. Section 3.2 introduces notation, the longevity models and formulates the hedging problem as an optimization problem. Section 3.3 derives the optimal control process by adopting the dynamic programming principle and the Bellman's equation framework. Section 3.4 provides a numerical example to demonstrate the effectiveness and feasibility of our proposed strategy. Section 3.5 concludes the chapter.

## 3.2 Problem setup

### 3.2.1 Stochastic mortality model

In this section, we describe the set-up of our longevity risk management framework. The methodology to deriving an optimal hedging strategy using the dynamic programming framework is quite general and it is applicable to many popular longevity models, including the Lee-Carter model, the Cairns-Blake-Dowd (CBD) model and their extensions. For illustration purpose, we proceed our mathematical derivation with the assumption that the future mortality improvement follows the Augmented Common Factor (ACF) model proposed by Li and Lee (2005). We consider the following two-population model:

$$\ln(m_{x,t}^{(i)}) = a_x^{(i)} + B_x K_t + b_x^{(i)} k_t^{(i)} + \epsilon_{x,t}^{(i)},$$

where  $i \in \{H, R\}$ , and  $H$  and  $R$  are two different populations;  $m_{x,t}^{(i)}$  denotes population  $i$ 's central death rate at age  $x$  in year  $t$ ;  $a_x^{(i)}$  denotes population  $i$ 's average mortality level at age  $x$ ;  $K_t$  represents the mortality improvement to both populations,  $B_x$  is the corresponding coefficient for age  $x$ ;  $k_t^{(i)}$  represents the mortality improvement specific to population  $i$ , and  $b_x^{(i)}$  is the corresponding coefficients for age  $x$  specific to population  $i$ ;  $\epsilon_{x,t}^{(i)}$  denotes the residual term which are modeled by independent and identically distributed (i.i.d.) normal random variables. In addition, the parameters  $B_x$ ,  $b_x^{(i)}$ ,  $K_t$  and  $k_t^{(i)}$  are subject to  $\sum_x B_x = \sum_x b_x^{(i)} = 1$  and  $\sum_t K_t = \sum_t k_t^{(i)} = 0$  to ensure identifiability of the model.

The time-varying indices  $\{K_t\}_{t \geq 0}$  and  $\{k_t^{(i)}\}_{t \geq 0}$ ,  $i \in \{H, R\}$  are further modeled by time series models, e.g., autoregressive integrated moving average (ARIMA) models. In this chapter, we follow Li and Lee (2005) and assume that  $\{K_t\}_{t \geq 0}$  follows a random walk with drift, while each of  $\{k_t^{(i)}\}_{t \geq 0}$ ,  $i \in \{H, R\}$  follows an AR(1) model:  $\forall t = 0, 1, 2, 3, \dots$

$$\begin{aligned} K_t &= C + K_{t-1} + \xi_t, \\ k_t^{(i)} &= \phi_0^{(i)} + \phi_1^{(i)} k_{t-1}^{(i)} + \zeta_t^{(i)}, \end{aligned}$$

where  $i \in \{H, R\}$ ,  $C$ ,  $\phi_0^{(i)}$  and  $\phi_1^{(i)}$  are constants, and  $\{\xi_t\}$  and  $\{\zeta_t^{(i)}\}$  are two mutually independent sequences of i.i.d. normal random variables with zero mean and constant variance:

$$\begin{aligned} \text{var}(\xi_t) &= \sigma_K^2, \\ \text{var}(\zeta_t^{(i)}) &= \sigma_{k,i}^2 \quad i \in \{H, R\}. \end{aligned}$$

Additionally, we assume  $|\phi_1^{(i)}| < 1$  to ensure that the two time series  $\{k_t^{(i)}\}_{t \geq 0}$  are stationary. Finally, we emphasize that the assumptions on the time-varying indices  $\{K_t\}_{t \geq 0}$  and  $\{k_t^{(i)}\}_{t \geq 0}$ ,  $i \in \{H, R\}$  are for illustration purpose only, and our derivation also applies to general ARIMA  $(p, d, q)$  models.

### 3.2.2 Pension liability

In order to define the pension plan liability and formulate our hedging problem, it is useful to introduce the following notations:

- $q_{x,t}^{(i)}$  denotes the probability that an individual aged  $x$  at time  $t - 1$  (alive) from population  $i$  dies between time  $t - 1$  and  $t$ ;
- $S_{x,t}^{(i)}(T) := \prod_{s=1}^T (1 - q_{x+s-1,t+s}^{(i)})$  denotes the probability that an individual from population  $i$  aged  $x$  at time  $t$  (alive) will survive to time  $t + T$ ;
- $p_{x,u}^{(i)}(T, K_t, k_t^{(i)}) := E(S_{x,u}^{(i)}(T) | \mathcal{F}_t) = E(S_{x,u}^{(i)}(T) | K_t, k_t^{(i)})$ ,  $u \geq t$ , is the expected survival probability given information up to time  $t$ , where  $\{\mathcal{F}_t\}_{t \geq 0}$  denotes the filtration generated

by  $\{K_u, k_u^{(H)}, k_u^{(R)}\}_{\{0 \leq u \leq t\}}$ . We call it a spot survival probability if  $u = t$  and a forward survival probability if  $u > t$ .

Now we consider a pension plan involving a single cohort of  $n$  pensioners all age  $x_0$  at time 0 from population  $H$ , and the sponsor of the pension plan would like to hedge against the unexpected future mortality improvement associated with the plan liability. Without any loss of generality we assume that the notional amount is  $\frac{1}{n}$  per pensioner, i.e., the plan pays each pensioner  $\frac{1}{n}$  at the end of each year until death, so that the total notional amount of the whole pension plan at time 0 is \$1. We also assume that the size of the plan,  $n$ , is large enough such that there is no sample risk and thus mortality experience of the underlying cohort perfectly matches the mortality rate of population  $H$ . As a result, the time- $t$  present value of future plan liability, denoted by  $FL_t$ , can be expressed by

$$FL_t = \sum_{s=t+1}^{\infty} (1+r)^{-s} S_{x_0+t,t}^{(H)}(s),$$

where  $r$  is the risk-free rate and is assumed to be a constant in this chapter. In particular, we are interested in designing a hedging strategy for the time-0 future liability variable, given by

$$FL_0 = \sum_{s=1}^{\infty} (1+r)^{-s} S_{x_0,0}^{(H)}(s). \quad (3.1)$$

### 3.2.3 Dynamic hedge with q-forward

Any unexpected changes in the mortality experience of population  $H$  can have an adverse effect to the pension plan sponsor. This can be seen from the time-0 value of the future pension plan liability given by (3.1). The pension plan sponsor, or the hedger, is therefore interested in hedging its longevity risk exposure. Here we assume that the hedger hedges its risk by trading index-linked q-forward contracts, instead of customized longevity securities. We further assume that the payoff of the q-forward depends on some other referenced population denoted by  $R$ . Because of the imperfect correlation between populations  $H$  and  $R$ , basis risk arises in our hedging

strategy. The magnitude of basis risk, or equivalently, the effectiveness of our hedging strategy depends on how the hedging portfolio is constructed. This chapter is concerned with an optimal construction of hedging portfolio such that the basis risk is minimized under some chosen criterion to be formally defined later.

To proceed, we now describe the q-forward. A q-forward contract is basically a zero-coupon swap with its floating leg proportional to the realized death probability for a certain reference age during the very last year before maturity, and its fixed leg proportional to a predetermined forward mortality rate. As swap type contracts, q-forwards do not generate any cash flows at the inception of the contract, and all payments are settled at the maturity date. In our setting, we assume that the q-forward contract written on population  $H$  is not liquidly traded, or even does not exist in the market. As an alternative, the hedger may choose the q-forward contract that is written on population  $R$  because it is sufficiently liquid as the hedging instrument. To study the effect of mismatch between populations on basis risk, in the rest of the chapter we only consider the longevity security linked to population  $R$  as the hedging instrument.

Whenever mortality rates and future improvements for populations  $H$  and  $R$  are not perfectly correlated, we would expect that basis risk to arise and a perfect hedge is not possible. Additionally, in reality it is usually the case that, besides population basis risk, other types of basis risk such as age basis risk and gender basis risk may also exist (Cairns et al., 2014). In this section we focus on basis risk associated with the population differences, and will have some discussion on factors related to the structure of the hedge in Sections 3.4.5 and 3.4.6.

We assume that the fixed leg of the q-forward is determined by the pure risk premium, i.e., there is no additional risk premium at the inception time of the contract,  $t_0$ . Therefore, by denoting the forward mortality rate for the fixed leg as  $q_{t_0}^f := q^f(t_0, K_{t_0}, k_{t_0}^{(R)}, x_f, T^*)$ , we have

$$q_{t_0}^f = E(q_{x_f, t_0 + T^*}^{(R)} | \mathcal{F}_{t_0}) = E(q_{x_f, t_0 + T^*}^{(R)} | K_{t_0}, k_{t_0}^{(R)}),$$

where  $x_f$  is a pre-specified reference age,  $t_0$  is the issuing time of the contract,  $T^*$  is the time to maturity at inception and  $t_0 + T^*$  is the maturity time.

As the hedger of the pension plan liability defined by equation (3.1), we want to be the fixed rate receiver between the two counterparties of the q-forward contracts, as this hedging instru-

ment is linked to the death probability but not the survival probability. Subsequently, a dynamic hedging scheme can be constructed via a so-called “rolling” strategy: at time  $t$ ,  $t = 0, 1, 2, \dots$ , we write a q-forward contract as the fixed leg receiver with a notional amount of  $h_t$  (with no initial cost); a year later at time  $t + 1$ , we close out the position written at time  $t$  (with profit or loss) and then write a new contract with notional amount  $h_{t+1}$ ; we repeat this process until the end of our hedging horizon. Based on the above description, the hedging strategy can be represented by the process  $\{h_t\}_{t \geq 0}$ , which is adapted to  $\{\mathcal{F}_t\}_{t \geq 0}$ . Via such a yearly rebalancing hedging scheme we are able to utilize the latest information obtained from the market and to adjust our position in our hedging instruments that best matches our future liability. Mathematically, we denote  $Q_{t_2}(t_1)$ ,  $t_2 \geq t_1$ , as the time- $t_2$  value of q-forward contract that was written at time  $t_1$  (to the fixed rate receiver), i.e.,

$$Q_{t_2}(t_1) = (1 + r)^{-(T^* - (t_2 - t_1))} \left[ q_{t_1}^f - E(q_{x_f, t_1 + T^*}^{(R)} | \mathcal{F}_{t_2}) \right]. \quad (3.2)$$

Therefore at time  $t = 0, 1, 2, \dots$ , the value of our position in the newly written q-forward contract is represented as:

$$h_t \cdot Q_t(t) = 0.$$

At time  $t + 1$ , we may experience a profit or loss from the actual realization of mortality improvement during the one-year time period, and the value of our position in q-forwards written at time  $t$  becomes:

$$\begin{aligned} & h_t \cdot Q_{t+1}(t) \\ = & h_t \cdot (1 + r)^{-(T^* - 1)} \left[ q_t^f - E(q_{x_f, t + T^*}^{(R)} | \mathcal{F}_{t+1}) \right] \\ = & h_t \cdot (1 + r)^{-(T^* - 1)} \left[ E(q_{x_f, t + T^*}^{(R)} | \mathcal{F}_t) - E(q_{x_f, t + T^*}^{(R)} | \mathcal{F}_{t+1}) \right] \\ = & h_t \cdot (1 + r)^{-(T^* - 1)} \left[ p_{x_f, t + T^* - 1}^{(R)}(1, K_{t+1}, k_{t+1}^{(R)}) - p_{x_f, t + T^* - 1}^{(R)}(1, K_t, k_t^{(R)}) \right]. \quad (3.3) \end{aligned}$$

An immediate conclusion we can draw about this hedging strategy is that, it does not affect

the hedger's expected rate of return. Mathematically we have the following proposition.

**Proposition 3.1.**  $\forall t = 0, 1, 2, \dots$ , we have

$$\mathbb{E}_t[Q_{t+1}(t)] = Q_t(t) = 0, \quad (3.4)$$

where  $\mathbb{E}_t[\cdot]$  denotes  $\mathbb{E}[\cdot | \mathcal{F}_t]$ .

*Proof.* By definition,

$$\begin{aligned} \mathbb{E}_t[Q_{t+1}(t)] &= (1+r)^{-(T^*-1)} \left\{ \mathbb{E}_t \left[ \mathbb{E}(q_{x_f, t+T^*}^{(R)} | \mathcal{F}_t) \right] - \mathbb{E}_t \left[ \mathbb{E}(q_{x_f, t+T^*}^{(R)} | \mathcal{F}_{t+1}) \right] \right\} \\ &= (1+r)^{-(T^*-1)} \left\{ \mathbb{E}_t \left[ q_{x_f, t+T^*}^{(R)} \right] - \mathbb{E}_t \left[ q_{x_f, t+T^*}^{(R)} \right] \right\} \\ &= 0. \end{aligned}$$

□

So far the hedging problem boils down to the problem of determining  $\{h_t\}_{t \geq 0}$  such that the hedged position will achieve certain desirable objectives of the hedger. In the rest of this chapter, we write  $\mathbb{E}_t[\cdot] := \mathbb{E}[\cdot | \mathcal{F}_t]$  and  $\text{Var}_t[\cdot] := \text{Var}[\cdot | \mathcal{F}_t]$ ,  $\forall t \geq 0$ . Furthermore, we also use  $\mathbb{E}[\cdot]$  and  $\text{Var}[\cdot]$  to denote  $\mathbb{E}_0[\cdot]$  and  $\text{Var}_0[\cdot]$  respectively.

### 3.2.4 Hedging objective

Without loss of generality, we assume that the hedger's initial net wealth is 0. We adopt the hedging strategy described in Section 3.2.3 and construct a yearly adjusted hedging portfolio starting at time 0 by trading q-forward contracts, in order to hedge the pension plan liability defined in Section 3.2.2. At each time point  $t$ ,  $t = 1, 2, 3, \dots$ , the net position of the hedger can be positive or negative (i.e. profit or loss) depending on relative values of the hedging portfolio and the realized pension liability. Define

$$F(t) := h_{t-1} \cdot Q_t(t-1) - S_{x_0, 0}^{(H)}(t), \quad t = 1, 2, 3, \dots \quad (3.5)$$

Then  $F(t)$  is the time- $t$  present value of net cash flow that adds up to the hedger's net wealth process. Also we define  $X(s)$  as the time-0 value of all future net cash flows truncated at time  $s$ ,  $\forall s > 0$ ,

$$X(s) := \sum_{t=1}^s (1+r)^{-t} F(t). \quad (3.6)$$

We denote  $\omega$  as the limiting age such that no human being survives beyond it under current medical conditions, and  $Y$  as a finite hedging horizon such that  $Y \leq \omega$ . Because  $\omega$  is a finite number and we know that  $S_{x_0,0}^{(H)}(t) = 0$  for large  $t \geq \omega$ , we will immediately stop constructing the hedging portfolio as our remaining plan liability reduces to 0. The reason why we further consider a hedging horizon  $Y \leq \omega$  is that the number of survivors in the pension plan during sufficiently late years will be very small and hence it is quite unnecessary to hedge such small plan liabilities. For the time-0 present value of all future net cash flows, we adopt the truncation and apply the following approximation,

$$\sum_{t=1}^{\infty} (1+r)^{-t} F(t) \approx X(Y).$$

We can always choose a large enough  $Y$  such that the above approximation yields a certain level of desirable accuracy.

Since basis risk exists and a perfect hedging strategy is in general impossible, we choose to optimize a mean-variance objective function of the time-0 present value of the truncated future net cash flows  $X(Y)$ . In other words, our proposed longevity hedging strategy is defined as a solution to the following optimization problem:

$$\max_{\{h_t\}_{t=0,1,2,\dots,Y-1}} \left\{ E_0[X(Y)] - \frac{\gamma}{2} \text{Var}_0[X(Y)] \right\}, \quad (3.7)$$

where  $\gamma > 0$  is the risk aversion coefficient.

Problem (3.7) can be recast into a quadratic hedging problem as demonstrated in the following proposition.

**Proposition 3.2.** *Problem (3.7) is equivalent to the following problem (3.8), in the sense that they yield the same optimal control process.*

$$\min_{\{h_t\}_{t=0,1,2,\dots,Y-1}} \mathbb{E}_0[X(Y)^2]. \quad (3.8)$$

*Proof.* The objective function (3.7) can be rewritten as

$$\begin{aligned} & \mathbb{E}_0[X(Y)] - \frac{\gamma}{2} \text{Var}_0[X(Y)] \\ = & \mathbb{E}_0[X(Y)] - \frac{\gamma}{2} \mathbb{E}_0[X(Y)^2] + \frac{\gamma}{2} (\mathbb{E}_0[X(Y)]^2) \\ = & \mathbb{E}_0 \left[ \sum_{t=1}^Y (1+r)^{-t} \left( h_{t-1} \cdot Q_t(t-1) - S_{x_0,0}^{(H)}(t) \right) \right] - \frac{\gamma}{2} \mathbb{E}_0[X(Y)^2] \\ & + \frac{\gamma}{2} \left( \mathbb{E}_0 \left[ \sum_{t=1}^Y (1+r)^{-t} \left( h_{t-1} \cdot Q_t(t-1) - S_{x_0,0}^{(H)}(t) \right) \right] \right)^2 \\ = & -\mathbb{E}_0 \left[ \sum_{t=1}^Y (1+r)^{-t} S_{x_0,0}^{(H)}(t) \right] - \frac{\gamma}{2} \mathbb{E}_0[X(Y)^2] + \frac{\gamma}{2} \left( \mathbb{E}_0 \left[ \sum_{t=1}^Y (1+r)^{-t} S_{x_0,0}^{(H)}(t) \right] \right)^2. \end{aligned}$$

The last equality is due to the fact that  $\mathbb{E}_0[Q_t(t-1)] = 0, \forall t > 0$ , by Proposition 3.1. The first term and the last term in the last line of the last display do not depend on control process  $\{h_t\}_{t=0,1,2,\dots,Y-1}$ , so maximizing  $\mathbb{E}_0[X(Y)] - \frac{\gamma}{2} \text{Var}_0[X(Y)]$  is equivalent to minimizing  $\mathbb{E}_0[X(Y)^2]$ .  $\square$

The objective function in (3.8) has a standard form of conditional expectation, and therefore it admits the Bellman's Principle of Optimality. In the rest of this chapter we will work on problem (3.8) instead of the original optimization problem (3.7). We denote  $\{h_t^*\}_{t=0,1,2,\dots,Y-1}$  as the minimizer of problem (3.8), i.e.,

$$\{h_t^*\}_{t=0,1,2,\dots,Y-1} = \underset{\{h_t\}_{t=0,1,2,\dots,Y-1}}{\text{argmin}} \mathbb{E}_0[X(Y)^2],$$

where  $\{h_t^*\}_{t=0,1,2,\dots,Y-1}$  is adapted to  $\{\mathcal{F}_t\}_{t=0,1,2,\dots,Y-1}$ . Our goal is to solve for  $\{h_t^*\}_{t=0,1,2,\dots,Y-1}$  such that hedging objective (3.8) is achieved.

### 3.3 Derivation of the optimal solution

In this section we solve problem (3.8) for the optimal hedging strategy  $\{h_t^*\}_{t=0,1,2,\dots,Y-1}$ , by resorting to the dynamic programming principle. Section 3.3.1 derives the Bellman equation and Section 3.3.2 solves for  $\{h_t^*\}_{t=0,1,2,\dots,Y-1}$  using the Bellman equation.

#### 3.3.1 The Bellman equation

Define

$$J_t := \mathbb{E}_t \left[ \left( \sum_{s=t+1}^Y (1+r)^{-s} F(s) \right)^2 \right], \quad t = 0, 1, \dots, Y,$$

and

$$V_t := \min_{\{h_s\}_{s=t,t+1,\dots,Y-1}} J_t, \quad t = 0, 1, \dots, Y. \quad (3.9)$$

Applying the dynamic programming principle, we obtain the Bellman equation, as shown in Proposition 3.3 below, to recursively solve for the optimization problem (3.8).

**Proposition 3.3.** *For problem (3.8), the value process  $\{V_t\}_{t=0,1,2,\dots,Y-1}$  satisfies the following Bellman equation,  $\forall t = 0, 1, 2, \dots, Y - 1$ ,*

$$\begin{aligned} V_t = \min_{h_t} \mathbb{E}_t \left\{ & V_{t+1} + (1+r)^{-2(t+1)} h_t^2 Q_{t+1}^2(t) + (1+r)^{-2(t+1)} S_{x_0,0}^{(H)}(t+1)^2 \right. \\ & - 2(1+r)^{-2(t+1)} h_t Q_{t+1}(t) S_{x_0,0}^{(H)}(t+1) \\ & + 2(1+r)^{-(t+1)} h_t Q_{t+1}(t) \mathbb{E}_{t+1} \left[ \sum_{s=t+2}^Y (1+r)^{-s} F^*(s) \right] \\ & \left. - 2(1+r)^{-(t+1)} S_{x_0,0}^{(H)}(t+1) \mathbb{E}_{t+1} \left[ \sum_{s=t+2}^Y (1+r)^{-s} F^*(s) \right] \right\}, \quad (3.10) \end{aligned}$$

with the boundary condition given by  $V_Y = X(Y)^2$ . The optimal control process  $\{h_t^*\}_{t=0,1,2,\dots,Y-1}$  is the corresponding minimizer.

*Proof.* From equation (3.8), the boundary condition is trivially true. In order to prove equation (3.10), we first rewrite the objective function  $J_t$  by splitting it into two terms:

$$\begin{aligned}
J_t &= \mathbb{E}_t \left\{ \left( (1+r)^{-(t+1)} F(t+1) + \sum_{s=t+2}^Y (1+r)^{-s} F(s) \right)^2 \right\} \\
&= \mathbb{E}_t \left\{ \mathbb{E}_{t+1} \left[ \left( (1+r)^{-(t+1)} F(t+1) + \sum_{s=t+2}^Y (1+r)^{-s} F(s) \right)^2 \right] \right\} \\
&= \mathbb{E}_t \left\{ \mathbb{E}_{t+1} \left[ (1+r)^{-2(t+1)} F(t+1)^2 + 2(1+r)^{-(t+1)} F(t+1) \sum_{s=t+2}^Y (1+r)^{-s} F(s) \right. \right. \\
&\quad \left. \left. + \left( \sum_{s=t+2}^Y (1+r)^{-s} F(s) \right)^2 \right] \right\} \\
&= \mathbb{E}_t \left\{ (1+r)^{-2(t+1)} F(t+1)^2 + 2(1+r)^{-(t+1)} F(t+1) \mathbb{E}_{t+1} \left[ \sum_{s=t+2}^Y (1+r)^{-s} F(s) \right] + J_{t+1} \right\}
\end{aligned}$$

We assume future cash flows,  $F(s)$ ,  $s = t+2, t+3, \dots, Y$ , are controlled by the optimal control process  $h_s^*$ ,  $s = t+1, t+2, \dots, Y-1$ , and use the notation  $F^*(s)$ ,  $s = t+2, t+3, \dots, Y$ , for  $F(s)$  controlled by the optimal control  $h_s^*$ . Therefore we obtain the Bellman equation:

$$\begin{aligned}
V_t &= \min_{\{h_s^*\}_{s=t,t+1,\dots,Y-1}} J_t \\
&= \min_{h_t} \mathbb{E}_t \left\{ (1+r)^{-2(t+1)} F(t+1)^2 + 2(1+r)^{-(t+1)} F(t+1) \mathbb{E}_{t+1} \left[ \sum_{s=t+2}^Y (1+r)^{-s} F^*(s) \right] \right. \\
&\quad \left. + V_{t+1} \right\} \\
&= \min_{h_t} \mathbb{E}_t \left\{ V_{t+1} + (1+r)^{-2(t+1)} h_t^2 Q_{t+1}^2(t) + (1+r)^{-2(t+1)} S_{x_0,0}^{(H)}(t+1)^2 \right\}
\end{aligned}$$

$$\begin{aligned}
& -2(1+r)^{-2(t+1)}h_tQ_{t+1}(t)S_{x_0,0}^{(H)}(t+1) \\
& +2(1+r)^{-(t+1)}h_tQ_{t+1}(t)E_{t+1}\left[\sum_{s=t+2}^Y(1+r)^{-s}F^*(s)\right] \\
& -2(1+r)^{-(t+1)}S_{x_0,0}^{(H)}(t+1)E_{t+1}\left[\sum_{s=t+2}^Y(1+r)^{-s}F^*(s)\right]\Bigg\},
\end{aligned}$$

and this completes the proof.  $\square$

### 3.3.2 Solution of the Bellman equation

We observe from (3.6) that  $F(t)$  depends on not only the current state  $\{t, K_t, k_t^{(H)}, k_t^{(R)}\}$  but also their realizations before time  $t$ . Thus, the term  $\left(\sum_{s=t+1}^Y(1+r)^{-s}F(s)\right)^2$  appears in the objective of problem (3.9) is non-Markovian. This implies that the value function  $V_t$  defined in (3.9) depends not only on the time- $t$  state  $\{t, K_t, k_t^{(H)}, k_t^{(R)}\}$  but also on the states before time  $t$ . In general, this non-Markovian feature may make the problem intractable. Fortunately, we are able to show that the value function  $V_t$  does not depend on the control policies adopted before time  $t$ , as shown in the proposition below.

**Proposition 3.4.**  $\forall 0 \leq t \leq Y - 2$ , the value function process  $V_{t+1}$  defined by equation (3.9) is independent of the control process up to time  $t$ , i.e.,  $\{h_s\}_{0 \leq s \leq t}$ .

*Proof.* We prove by backward induction. For  $t = Y - 2$ , from the Bellman equation (3.10),

$$\begin{aligned}
V_{Y-1} &= E_{Y-1} \left\{ [(1+r)^{-Y}F^*(Y)]^2 \right\} \\
&= (1+r)^{-2Y} E_{Y-1} \left\{ [h_{Y-1}^*Q_Y(Y-1) - S_{x_0,0}^{(H)}(Y)]^2 \right\},
\end{aligned}$$

which obviously depends on only  $h_{Y-1}^*$  but not  $\{h_s\}_{0 \leq s \leq Y-2}$ .

Suppose that the conclusion holds for  $t = k$ , i.e.,  $V_{k+1}$  is independent of  $\{h_s\}_{0 \leq s \leq k}$ . Then for

$t = k - 1$ , it follows from the Bellman equation (3.10) that

$$\begin{aligned}
V_k &= \min_{h_k} \mathbb{E}_k \left\{ V_{k+1} + (1+r)^{-2(k+1)} h_k^2 Q_{k+1}^2(k) + (1+r)^{-2(k+1)} S_{x_0,0}^{(H)}(k+1)^2 \right. \\
&\quad - 2(1+r)^{-2(k+1)} h_k Q_{k+1}(k) S_{x_0,0}^{(H)}(k+1) \\
&\quad \left. + 2(1+r)^{-(k+1)} h_k Q_{k+1}(k) \mathbb{E}_{k+1} \left[ \sum_{s=k+2}^Y (1+r)^{(-s)} F^*(s) \right] \right. \\
&\quad \left. - 2(1+r)^{-(k+1)} S_{x_0,0}^{(H)}(k+1) \mathbb{E}_{k+1} \left[ \sum_{s=k+2}^Y (1+r)^{(-s)} F^*(s) \right] \right\} \\
&= \mathbb{E}_k \{ V_{k+1} \} + \min_{h_k} \mathbb{E}_k \left\{ (1+r)^{-2(k+1)} h_k^2 Q_{k+1}^2(k) + (1+r)^{-2(k+1)} S_{x_0,0}^{(H)}(k+1)^2 \right. \\
&\quad - 2(1+r)^{-2(k+1)} h_k Q_{k+1}(k) S_{x_0,0}^{(H)}(k+1) \\
&\quad \left. + 2(1+r)^{-(k+1)} h_k Q_{k+1}(k) \mathbb{E}_{k+1} \left[ \sum_{s=k+2}^Y (1+r)^{(-s)} F^*(s) \right] \right. \\
&\quad \left. - 2(1+r)^{-(k+1)} S_{x_0,0}^{(H)}(k+1) \mathbb{E}_{k+1} \left[ \sum_{s=k+2}^Y (1+r)^{(-s)} F^*(s) \right] \right\}.
\end{aligned}$$

In the end of the last display, the first expectation term is independent of the control process  $\{h_s\}_{0 \leq s \leq k}$ , and the second expectation term only involves  $h_k$ . This implies  $V_k$  is independent of  $\{h_s\}_{0 \leq s \leq k-1}$  and by a backward induction, the proof is complete.  $\square$

By Proposition 3.4,  $\forall t = 0, 1, \dots, Y - 1$ , the optimal hedging strategy  $h_t^*$  can be solved as the maximizer of a quadratic form in equation (3.10). Subsequently we can obtain the optimal hedging strategy as described in the following proposition.

**Proposition 3.5.** *The optimal control process  $h_t^*$ ,  $\forall t = 0, 1, \dots, Y - 1$ , to the Bellman equation (3.10) is given by*

$$h_t^* = \frac{\mathbb{E}_t \left\{ Q_{t+1}(t) \sum_{s=t+1}^Y (1+r)^{-[s-(t+1)]} S_{x_0,0}^{(H)}(s) \right\}}{\mathbb{E}_t [Q_{t+1}^2(t)]}, \quad (3.11)$$

where

$$Q_{t+1}(t) = (1+r)^{-(T^*-1)} \left( p_{x_f, t+T^*-1}^{(R)}(1, K_{t+1}, k_{t+1}^{(R)}) - p_{x_f, t+T^*-1}^{(R)}(1, K_t, k_t^{(R)}) \right), \quad (3.12)$$

as defined in equation (3.2) and demonstrated in equation (3.3).

*Proof.* In equation (3.10), the optimization term has a quadratic form which opens upwards, and therefore the minimizer  $h_t^*$  always exists and is given by

$$h_t^* = \frac{\left\{ \mathbb{E}_t \left[ Q_{t+1}(t) S_{x_0,0}^{(H)}(t+1) \right] - (1+r)^{t+1} \mathbb{E}_t \left[ Q_{t+1}(t) \mathbb{E}_{t+1} \left( \sum_{s=t+2}^Y (1+r)^{-s} F^*(s) \right) \right] \right\}}{\mathbb{E}_t \left[ Q_{t+1}^2(t) \right]}, \quad (3.13)$$

for any  $t = 0, 1, 2, \dots, Y-1$ . In particular, for  $t = Y-1$ , obviously  $\sum_{s=t+2}^Y (1+r)^{-s} F^*(s) = 0$ , and thus,

$$h_{Y-1}^* = \frac{(1+r)^Y \mathbb{E}_{Y-1} \left[ Q_Y(Y-1) S_{x_0,0}^{(H)}(Y) \right]}{\mathbb{E}_{Y-1} \left[ Q_Y^2(Y-1) \right]}.$$

Next we consider the cases when  $t = 0, 1, \dots, Y-2$ . We note that  $h_{s-1}$  is  $\mathcal{F}_{s-1}$  measurable, and thus we apply Proposition 3.1 to conclude  $\mathbb{E}_{s-1} [h_{s-1} Q_s(s-1)] = 0$ . This implies  $\mathbb{E}_{t+1} \left( \sum_{s=t+2}^Y (1+r)^{-s} h_{s-1} \cdot Q_s(s-1) \right) = 0$ . Further from equation (3.5), we have

$$\mathbb{E}_t \left[ Q_{t+1}(t) \mathbb{E}_{t+1} \left( \sum_{s=t+2}^Y (1+r)^{-s} F^*(s) \right) \right] = -\mathbb{E}_t \left[ Q_{t+1}(t) \mathbb{E}_{t+1} \left( \sum_{s=t+2}^Y (1+r)^{-s} S_{x_0,0}^{(H)}(s) \right) \right].$$

Plugging the equation above into equation (3.13), we can simplify the expression of  $h_t^*$  into the desired expression

$$h_t^* = \frac{\left\{ \mathbb{E}_t \left[ Q_{t+1}(t) S_{x_0,0}^{(H)}(t+1) \right] + (1+r)^{t+1} \mathbb{E}_t \left[ Q_{t+1}(t) \mathbb{E}_{t+1} \left( \sum_{s=t+2}^Y (1+r)^{-s} S_{x_0,0}^{(H)}(s) \right) \right] \right\}}{\mathbb{E}_t \left[ Q_{t+1}^2(t) \right]}$$

$$\begin{aligned}
&= \frac{\left\{ (1+r)^{t+1} \mathbb{E}_t \left[ Q_{t+1}(t) \mathbb{E}_{t+1} \left( \sum_{s=t+1}^Y (1+r)^{-s} S_{x_0,0}^{(H)}(s) \right) \right] \right\}}{\mathbb{E}_t [Q_{t+1}^2(t)]} \\
&= \frac{\mathbb{E}_t \left\{ Q_{t+1}(t) \sum_{s=t+1}^Y (1+r)^{-[s-(t+1)]} S_{x_0,0}^{(H)}(s) \right\}}{\mathbb{E}_t [Q_{t+1}^2(t)]}.
\end{aligned}$$

□

**Remark 3.1.** Proposition 3.5 shows that the optimal control  $h_t^*$ ,  $\forall t = 0, 1, \dots, Y - 1$ , only depends on the current state variable  $\{t, K_t, k_t^{(H)}, k_t^{(R)}\}$  and this feature significantly simplifies the computation of the optimal hedging strategy. We also note that the value function  $V_t$ ,  $\forall t = 0, 1, \dots, Y$ , cannot be written as a function of the current state variable  $\{t, K_t, k_t^{(H)}, k_t^{(R)}\}$  only. Fortunately sitting at each time  $t$ , we are only interested in the particular value  $V_t$  and the optimal hedging strategy  $h_t^*$  determined by the state we observe then. In our numerical examples, we will resort to Monte Carlo simulation method for their computation.

**Remark 3.2.** From equation (3.12), we can see that the calculation of  $Q_{t+1}(t)$  involves two forward mortality rates which themselves are defined as conditional expectations. Then by equation (3.11), the numerical calculation of the optimal hedging strategy  $\{h_t^*\}_{t=0,1,\dots,Y-1}$  involves a two-step “nested” Monte Carlo simulation which can be computationally intensive. Due to its two-step nature, a direct Monte Carlo may still be feasible, but is highly ineffective in general. In the next section, in addition to the crude Monte Carlo method, we will adopt certain approximation to facilitate the Monte Carlo simulation procedure in implementing Proposition 3.5.

### 3.4 Hedging Canadian mortality rates with UK mortality rates

By using an example which involves hedging Canadian mortality rates based on UK mortality rates, this section compares and evaluates the effectiveness of various hedging strategies, including our strategy proposed in Section 3.3. Section 3.4.1 describes the data source, summarizes

our assumptions and introduces the criterion to evaluate hedging performance. Section 3.4.2 provides an approximation method which significantly reduces the computational time in calculating the hedging strategy, in order to overcome the “nested” computational issue as we mentioned in Remark 3.2. Section 3.4.3 introduces the “delta” strategy developed by Cairns (2011), and Zhou and Li (2016), and this is the benchmark in the literature. Section 3.4.4 compares the hedging performance between our optimal strategy and the “delta” strategy, and the results show that our proposed hedging scheme consistently outperforms the “delta” strategy in terms of variance reduction. Sections 3.4.5, 3.4.6 and 3.4.7 conduct several sensitivity tests to demonstrate that the performance of our proposed method is generally stable to the changes in the hedging instrument’s reference age, the hedging instrument’s time to maturity and the underlying mortality model, respectively.

### 3.4.1 Data and assumptions

In this section we describe the data and model assumptions pertaining to our numerical illustration. In Sections 3.4.1 to 3.4.6, we assume that future mortality improvements follow the ACF model introduced in Section 3.2.1, and model parameters are calibrated from the mortality data of Canadian unisex population and UK unisex population aged 60 to 89 over the period of 1966 to 2005<sup>1</sup>. The calibration process follows the procedures described in Zhou and Li (2016), which is based on a first-order singular value decomposition (SVD) procedure. In this example, Canadian unisex population is referred as population  $H$ , and the UK unisex population is referred as population  $R$ . Other main assumptions are listed below:

- The pension liability we hedge against is from a single cohort of individuals aged  $x_0 = 60$  at time 0 from population  $H$ . The total notional amount is \$1.
- The scale of the pension plan is large enough so that we do not consider sample risk.
- The hedging horizon is  $Y = 30$  years.

---

<sup>1</sup>Data source: <http://www.mortality.org>

- The q-forwards used as hedging instruments are linked to death rates of population  $R$ . The reference age for all the contracts is fixed at  $x_f = 75$  and the time to maturity at inception is  $T^* = 10$  years.
- The risk free rate is  $r = 4\%$ .

We use Hedge Effectiveness (HE), which measures the proportion of variance reduction of a hedging strategy relative to the unhedged position, as a measure of quantifying the performance of a hedging strategy. It is defined as

$$\begin{aligned}
 \text{HE} &= 1 - \frac{\text{Var}[\text{time-0 value of hedged portfolio}]}{\text{Var}[\text{time-0 value of unhedged portfolio}]} \\
 &= 1 - \frac{\text{Var}[X(30)]}{\text{Var}[\sum_{t=1}^{30} (1+r)^{-t} S_{x_0,0}^{(H)}(t)]}, \tag{3.14}
 \end{aligned}$$

where  $X(30)$  is defined by equation (3.6). By definition, whenever the hedging strategy has a non-negative impact on the hedger's position, HE is a real number between 0 and 1. A larger HE implies that the hedging strategy is more effective in reducing the variance of hedger's position at time 0, and vice versa.  $\text{HE} = 1$  will be achieved only if the hedge is perfect, and we know it is in general impossible due to the mismatch in mortality rates between the Canadian population and the UK population. Once we obtain the optimal hedging strategy using equations (3.11) and (3.12) based on a numerical procedure, we can use Monte Carlo simulation to estimate variance of the time-0 value of both hedged and unhedged cash flows, and thus the Hedge Effectiveness to show its performance.

### 3.4.2 Approximation to forward mortality rates

A direct simulation to calculate the optimal strategy  $\{h_t^*\}_{t=0,1,2,\dots,29}$  based on equations (3.11) and (3.12) is possible, however, numerical calculation of those double conditional expectations requires the so-called "nested Monte Carlo" which is generally computationally expensive. An alternative way is to adopt the first-order approximation formula for forward mortality rates proposed by Cairns (2011) and Zhou and Li (2016). This approximation method is introduced to es-

timate forward survival rates by the first-order Taylor expansion, instead of calculating the mean of their true distributions. Subsequently it can be used to analytically evaluate equation (3.12), and thus we only need a regular simulation procedure to evaluate equation (3.11). In this way we can avoid the “nested Monte Carlo” issue, and therefore have a much lower computation burden to calculate the hedging strategies. The approximation formula for the ACF model we use for this longevity hedging problem is provided in Appendix B.1 (see Zhou and Li (2016) for proof). Additionally, a similar approximation formula for the CBD model, which will be used later in Section 3.4.7, is given in Appendix B.2.

### 3.4.3 Benchmark method

Since we will calculate the optimal hedging strategy  $\{h_t^*\}_{t=0,1,2,\dots,29}$  using a Monte Carlo procedure combined with the approximation formula to forward mortality rates discussed in subsection 3.4.2, there is no guarantee that the strategy calculated using equation (3.11) is exactly the actual objective function minimizer to problem (3.8). Therefore in order to show that the error from approximation procedure is acceptable and the approximated hedging strategy is still very effective in reducing the hedger’s basis risk, we introduce the “delta” hedging strategy developed in Cairns (2011), and Zhou and Li (2016) as the benchmark to compare our results with. According to their method, the notional amount of q-forwards the hedger chooses to keep at time  $t$ , denoted by  $h_t^{**}$  for  $t = 0, 1, \dots, 29$ , is determined such that the hedging portfolio and the pension plan’s future liability have the same sensitivity to the mortality index  $K_t$ , i.e.,

$$\frac{\partial \left[ \sum_{s=1}^Y (1+r)^{-s} p_{x_0+t,t}^{(H)}(s, K_t, k_t^{(H)}) \right]}{\partial K_t} = h_t^{**} \cdot \frac{\partial Q_t(t-1)}{\partial K_t}. \quad (3.15)$$

The “delta” method is relatively easy to understand and implement, and generally computationally less intensive. However, as a trade-off, the solution is based on a heuristic idea rather than rigorous mathematical formulation and therefore neglects some useful information from the market, and can hardly capture the full dependence structure between different populations. In the next section, we will show that our proposed optimal strategy  $\{h_t^*\}_{t=0,1,2,\dots,29}$  defined by

equation (3.11) generally outperforms the “delta”  $\{h_t^{**}\}_{t=0,1,2,\dots,29}$  defined by equation (3.15) in terms of reducing the hedger’s basis risk.

### 3.4.4 Baseline result

In this section we will present simulation results of five experiments based on the optimal strategy  $\{h_t^*\}_{t=0,1,2,\dots,29}$  and the “delta” strategy  $\{h_t^{**}\}_{t=0,1,2,\dots,29}$ , and use these results to assess their relative efficiency. In each experiment, we generate  $N = 2,000$  random sample paths using the ACF model, and then for each generated path we calculate our optimal hedging strategy  $h_t^*$  or the “delta” hedging strategy  $h_t^{**}$  for each time  $t, t = 0, 1, \dots, 29$ , based on five different numerical procedures listed below.

- Method 1a: First analytically calculate the approximation formula (3.12) using equation (B.1), then simulate  $M = 1,000$  paths to estimate equation (3.11).
- Method 1b: Same as Method 1a but increase the number of paths generated in the second step to  $M = 10,000$ .
- Method 1c: Same as Method 1a but further increase the number of paths generated in the second step to  $M = 100,000$ .
- Method 2: “Delta” method based on equation (3.15).
- Method 3: A direct “nested” simulation based on numerical evaluation for equation (3.11) with  $M = 10,000$  simulated sample paths and numerical evaluation for equation (3.12) with  $M_1 = 10,000$  sample paths.

For each of the five methods, we calculate the variance of time-0 value for both unhedged and hedged portfolios from the 2,000 generated samples, to produce a single point estimate of the Hedge Effectiveness defined in equation (3.14). Furthermore, a bootstrapping of  $N_b = 100,000$  is implemented using these 2,000 simulated samples to yield estimates for variance and percentiles of the Hedging Effectiveness. Results of the Hedge Effectiveness from these five experiments are shown in Table 3.1.

HE	Method 1a	Method 1b	Method 1c	Method 2	Method 3
mean	0.9149	0.9147	0.9213	0.8938	0.9153
variance	1.5812e-005	1.3876e-005	1.2142e-005	2.3656e-005	1.4084e-005
Min	0.8962	0.8979	0.9035	0.8699	0.8968
Q1	0.9123	0.9122	0.9190	0.8906	0.9128
Q2	0.9150	0.9148	0.9214	0.8939	0.9154
Q3	0.9177	0.9172	0.9237	0.8971	0.9179
Max	0.9299	0.9301	0.9341	0.9127	0.9300
95% C.I.	(0.9069,0.9225)	(0.9072,0.9218)	(0.9143,0.9279)	(0.8839,0.9030)	(0.9077,0.9224)
time (hrs)	0.27	1.56	12.18	0.16	117.51

Table 3.1: Results of Five Experiments

The results from Methods 1a, 1b and 1c in Table 3.1 are based on the optimal strategy with a Monte Carlo procedure combined with the approximation formula to forward mortality rates. The only difference across these three columns is the sample size adopted in the Monte Carlo procedure. A comparison among the three columns shows that the Hedge Effectiveness increases as we increase the number of sample paths  $M$  in calculating equation (3.11). Note that the HE improves significantly from  $M = 10,000$  to  $M = 100,000$ . Also, the variance decreases as  $M$  increases, which means that we can get a better estimation for the HE as we increase  $M$ . The cost is, of course, at the expense of more computational efforts.

Comparing results from Method 1 (i.e. Methods 1a, 1b and 1c) based on our proposed method with Method 2 based on the benchmark method, we can see that our hedging strategy consistently outperforms the “delta” hedge, as indicated by the higher HE. Even with the smallest number of simulated paths  $M = 1,000$  (i.e. Method 1a), the average HE is consistently and significantly larger than that of Method 2, and the two 95% Confidence Intervals obtained by Method 1a and 2 do not share any overlap. In order to better demonstrate the huge difference in HE between Methods 1 and 2, we conduct a two-sample t-test with unequal variances on the hedge effectiveness between Methods 1a and 2. The t-test result rejects the null hypothesis with a p-value less than  $10^{-8}$ , and therefore strongly suggests that the advantage of our proposed strategy over “delta” strategy is significant. Additionally, all the HE from Method 1 have smaller variance than Method 2, which means that the results calculated from our proposed method are more sta-

ble. Finally, it should be noted that the computational time of Method 1a is less than double the computational time of Method 2 to ensure its feasibility in practice.

Comparing the HE obtained using Method 1 (which uses the approximation formula) with Method 3 (which is based on a direct “nested” Monte Carlo), we can see that the average HE from Methods 1a and 1b are lower than that from Method 3, while Method 1c yields a higher HE than Method 3 does. Although in theory Method 3 gives us the real objective optimizer, in practice due to the computational time limit we may not always have large enough number of sample paths to estimate equation (3.11) accurately. This explains why the mean and quantiles of HE from Method 1c are higher than those of Method 3 in their estimates. Comparison between the variance also shows that the HE obtained in Methods 1b and 1c are more accurate than that in Method 3. As a result, Method 1c should be considered as a preferred choice among the five methods in term of balancing between hedging performance and computational cost.

It is worth mentioning that although Method 3 seems to be unacceptably slow compared to other four methods and the HE we obtain from Method 3 is not as good as Method 1c, it is still the best strategy in theory if we have very powerful computation resources so that computational time is not a major issue. In this example, in order to illustrate our results we need to calculate the optimal hedging strategy  $\{h_t^*\}_{t=0,1,\dots,29}$  using equation (3.11) by a total of 60,000 ( $= 30 \times 2,000$ ) times for each experiment, so the total computation time gets quite long even though one single hedging strategy only takes about 7 seconds. But in practice it is usually the case that we only need to numerically calculate the optimal strategy for one particular period each time instead of the whole path, and thus we can always increase the number of simulation paths to approach the theoretically optimal hedging performance.

### 3.4.5 Robustness to q-forwards’ time to maturity

In this section we examine the robustness of the Hedge Effectiveness of our proposed hedging strategy with respect to the maturity of the q-forwards. In our previous baseline example considered in section 3.4.4, the q-forwards’ time to maturity  $T^*$  was set at 10 years, which is one third of our hedging horizon  $Y = 30$  years; and the q-forwards’ reference age was set at  $x_f = 75$  which is approximately the average age of the underlying population during the hedging period.

However, in practice, a certain type of q-forward contract may not be always liquidly traded in the market. Therefore it is of interest to evaluate the Hedge Effectiveness of our proposed strategy if the hedging instruments are based on different maturities and reference ages.

In this section we repeat our experiments using Methods 1c and 2 described in section 3.4.4, by keeping reference age fixed as  $x_f = 75$  and varying the q-forwards' time to maturity with  $T^* = 3, 4, \dots, 20$  years. Simulation results on the HE with different time to maturity are presented in figure 3.1.

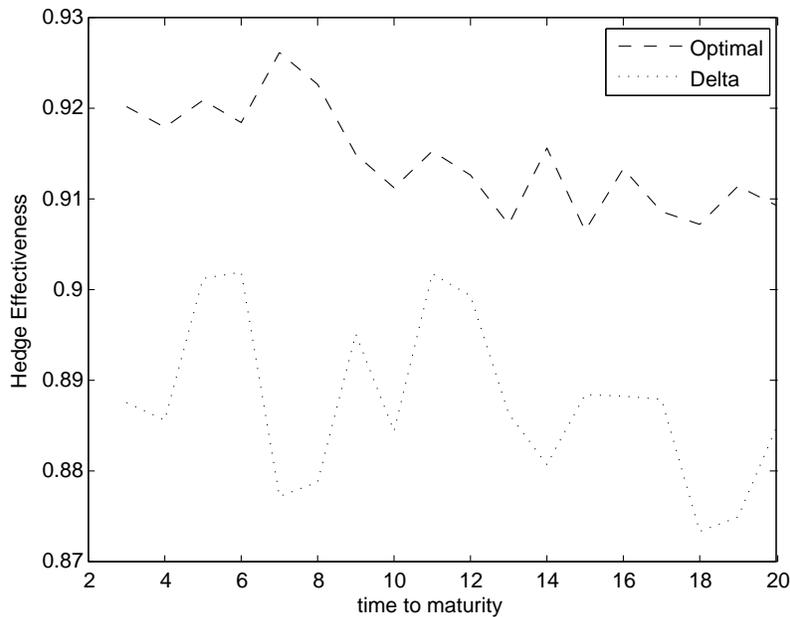


Figure 3.1: HE with different time to maturity

From figure 3.1 we can see that the obtained Hedge Effectiveness is very stable with respect to the maturity of q-forwards,  $T^*$ . For our proposed optimal hedging strategy, the HE is distributed within the interval [91%, 92%] while the dynamic “delta” method is consistently less effective. Since we adjust our hedging position in the hedging instrument every year based on the latest information from the market, we would not be very concerned if the amount of hedging instrument used for next year does not reflect mortality change in the far future. Additionally,

when the q-forwards' time to maturity is too long, for example,  $T^* = 20$  years, it may become less effective as time goes by because the mortality rate, which it is linked to, will be out of our hedging horizon.

This example reveals one of the major advantages of dynamic hedging strategy. Because longevity instruments with shorter time to maturity are generally much more appealing to investors and speculators than those longer term ones, longevity hedges will have more favorable expense and flexibility if the hedger adopts dynamic plans instead of static ones.

### 3.4.6 Robustness to q-forwards' reference age

Next we investigate the impact of q-forwards' reference age on the hedging performance. We fix time to maturity to be  $T^* = 10$  years as in our baseline example, and respectively consider reference ages  $x_f = 60, 61, \dots, 84$ . Since the underlying cohort of the pension plan liability are aged at 60 at inception, it does not make sense if we use q-forwards with reference ages that are lower than 60; for very high ages such as 85, it is unlikely to reflect our pension liability as well. This explains why we consider those values as the range for the reference age  $x_f$ . Results of the Hedge Effectiveness for different reference ages are shown in figure 3.2.

Figure 3.2 shows that the HE indeed heavily depends on q-forwards' reference age. Hedge effectiveness remains above 85% within the interval of  $71 \leq x_f \leq 81$ , while it fluctuates drastically for  $x_f < 71$  or  $x_f > 81$ . This is consistent with our conjecture that the best hedging performance should be obtained by using q-forwards that are linked to approximately the average age of the pension plan cohort during our hedging horizon. The curves shown in Figure 3.2 are not smooth due to the fact that age effects captured by  $B_x$  and  $b_x^{(i)}$ ,  $i = 1, 2$  in the ACF model are not necessarily smooth functions of  $x$ . Last but not least, we can also see that our hedging method generally outperforms the dynamic "delta" method, despite the unstable hedging performance when the reference age  $x_f$  is too large or too small.

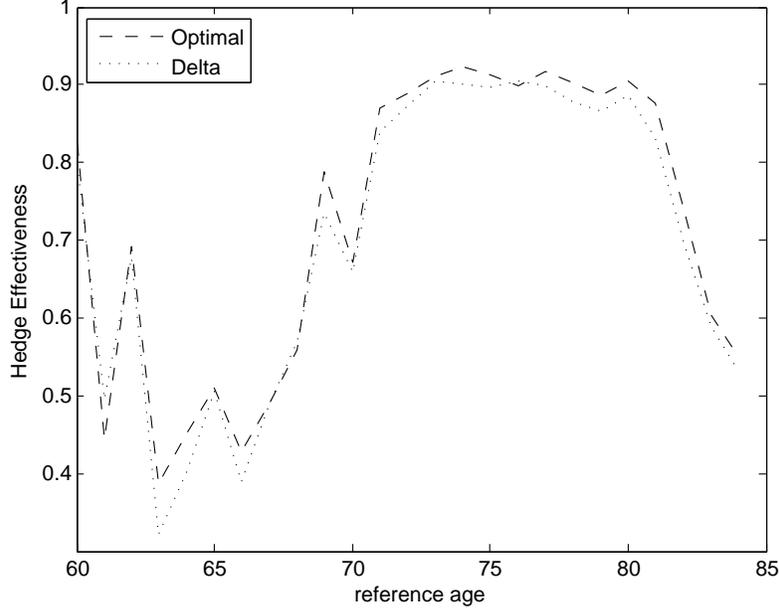


Figure 3.2: HE with different reference ages

### 3.4.7 Robustness to model risk

#### 3.4.7.1 The Cairns-Blake-Dowd (CBD) model

In this section, we study the impact of model risk, i.e., the risk coming from misidentifying the actual stochastic longevity model. In order to do so, we introduce the two-population Cairns-Blake-Dowd (CBD) model defined as follows:

$$\text{logit}(q_{x,t}^{(i)}) := \ln \left( \frac{q_{x,t}^{(i)}}{1 - q_{x,t}^{(i)}} \right) = \kappa_{1,t}^c + \kappa_{2,t}^c(x - \bar{x}) + \kappa_{1,t}^{(i)} + \kappa_{2,t}^{(i)}(x - \bar{x}) + \epsilon_{x,t}^{(i)}, \quad (3.16)$$

where  $i = H, R$ , are referred as two populations;  $\bar{x}$  denotes the average age over the sample age range  $[x_a, x_b]$  ( $=[60,89]$  in this example);  $\kappa_{1,t}^c$  and  $\kappa_{2,t}^c$  are time-varying factors which are common to both two populations;  $\kappa_{1,t}^{(i)}$  and  $\kappa_{2,t}^{(i)}$  are time-varying factors applied to the specific population  $i$ ;  $\epsilon_{x,t}^{(i)}$  captures all the remaining variations and the error terms for different populations, ages and

times are assumed to be independent.

Further the common factors  $\kappa_{1,t}^c$  and  $\kappa_{2,t}^c$  are modeled by a bivariate random walk with drift, i.e.,  $\forall t = 1, 2, 3, \dots$ ,

$$\begin{cases} \kappa_{1,t}^c = \mu_1^c + \kappa_{1,t-1}^c + \eta_{1,t}^c, \\ \kappa_{2,t}^c = \mu_2^c + \kappa_{2,t-1}^c + \eta_{2,t}^c, \end{cases}$$

where  $\eta_{1,t}^c$  and  $\eta_{2,t}^c$  are constantly correlated, i.e., the correlation coefficient  $\rho(\eta_{1,t}^c, \eta_{2,t}^c)$  is a constant number that does not change with time  $t$ . Moreover, the population specific factors  $\kappa_{1,t}^{(i)}$  and  $\kappa_{2,t}^{(i)}$  are modeled by two correlated AR(1) model,  $\forall t = 1, 2, 3, \dots$

$$\begin{cases} \kappa_{1,t}^{(i)} = \mu_1^{(i)} + \phi_1^{(i)} \kappa_{1,t-1}^{(i)} + \eta_{1,t}^{(i)}, \\ \kappa_{2,t}^{(i)} = \mu_2^{(i)} + \phi_2^{(i)} \kappa_{2,t-1}^{(i)} + \eta_{2,t}^{(i)}, \end{cases}$$

where  $\eta_{1,t}^{(i)}$  and  $\eta_{2,t}^{(i)}$  are also constantly correlated.

### 3.4.7.2 Hedging results with model risk

With model risk involved, we would expect a lower HE when applying our proposed hedging scheme due to the additional error from the model misspecification. To demonstrate this we investigate model risk by comparing results from the four following cases.

- Case 1: the “true” model is CBD model, but calculation of hedging strategy is based on ACF model (referred as “assumption” model in the rest of this section).
- Case 2: the “true” model is ACF model, but the “assumption” model is CBD model.
- Case 3: the “true” model is CBD model, and the “assumption” model is CBD model.
- Case 4: the “true” model is ACF model, and the “assumption” model is ACF model. This is the same as baseline results.

“True”/“assumption” model	mean	variance	min	max	95% C.I.
Case 1 (CBD/ACF)	0.6437	$1.7231 \times 10^{-4}$	0.5828	0.6942	(0.6174,0.6687)
Case 2 (ACF/CBD)	0.8712	$3.2003 \times 10^{-5}$	0.8416	0.8952	(0.8597,0.8819)
Case 3 (CBD/CBD)	0.6492	$1.7890 \times 10^{-4}$	0.5839	0.7029	(0.6222,0.6747)
Case 4 (ACF/ACF, baseline)	0.9213	$1.2142 \times 10^{-5}$	0.9035	0.9341	(0.9143,0.9279)

Table 3.2: HE results with different “true” and “assumption” models

In order to mimic this situation, we take the following numerical procedure for each of the four cases listed above:

- Step 1. Simulate  $N = 2,000$  mortality sample paths using the “true” model.
- Step 2. Based on each simulated path, calibrate model parameters for the “assumption” model. If the “true” model and the “assumption” model are identical, skip this step.
- Step 3. Calculate optimal hedging strategy for each time point on each simulated path based on the calibrated model parameters of the “assumption” model.
- Step 4. Calculate time-0 value of both hedged and unhedged position based on simulated paths and hedging strategy obtained in step 3.
- Step 5. Bootstrap of  $N_b = 100,000$  to obtain an estimation for the distribution of the Hedge Effectiveness.

To focus on model risk, in this example we fix the q-forward contract with reference age  $x_f = 75$  and time to maturity  $T^* = 10$ , to be consistent with the parameter values in our baseline experiments. Results of the Hedge Effectiveness under these four cases are shown in Table 3.2.

Table 3.2 shows that the Hedge Effectiveness by adopting our proposed hedging strategy is quite stable even when the “assumption” model deviates from the “true” model. Comparing Case 2 with baseline result Case 4, we can see that the average HE decreases from 92.13% to 87.12% if the “assumption” model is misidentified as the CBD model. All the percentiles also decrease slightly and variance of the HE increases substantially, from  $1.2142 \times 10^{-5}$  to  $3.2003 \times 10^{-5}$ .

When the model is misspecified, additional hedging error occurs due to the discrepancy between the “true” model and the “assumption” model, and the hedging strategy becomes less accurate, however, the HE we obtain in this example should still be considered as satisfactory in general.

When the “true” model is the CBD model (Cases 1 and 3), we observe a similar pattern. If the underlying model is correctly specified, the HE is slightly better in terms of average and percentiles, but the difference is quite small. It further implies that our hedging strategy is not sensitive to model risk.

It is worthwhile mentioning that the difference in HE achieved between “true” model of CBD cases (Cases 1 and 3) and “true” model of ACF cases (Cases 2 and 4) is not caused by the hedging strategy, but the implied population basis risk of different models. In this chapter correlations between mortality rates of different populations are modeled by common factors, and those model parameters we use in this example are estimated by a least square procedure. As a result, correlation between these two populations under different mortality model assumptions are not necessarily the same or even close, even though model parameters are calibrated from the same data set. Table 3.2 shows that in this example there is much larger population basis risk if future mortality rates follow a two-population CBD model.

### 3.4.7.3 Correlation coefficients under different models

In what follows, we provide some explanation about the seemingly large difference in basis risk we observed under different “true” model assumptions in section 3.4.7.2, by comparing the correlation coefficients between some representative death rates of two longevity models: the ACF model and the CBD model. To be consistent with our original hedging problem assumptions, we choose  $q_{59+t,t}^{(H)}$  and  $q_{75,t+9}^{(R)}$ ,  $t = 1, 2, 3, \dots, 30$ , as the representative death rates. Based on each of the ACF model and the CBD model calibrated from the same data set, we simulate  $N = 10,000$  future mortality paths and calculate correlation coefficient  $\rho(q_{59+t,t}^{(H)}, q_{75,t+9}^{(R)})$  for  $t = 1, 2, 3, \dots, 30$ , and numerical results are shown in Figure 3.3.

Figure 3.3 shows that the ACF model and the CBD model have different patterns to model the correlation between death rates  $q_{59+t,t}^{(H)}$  and  $q_{75,t+9}^{(R)}$  of two populations. When  $t$  is small,

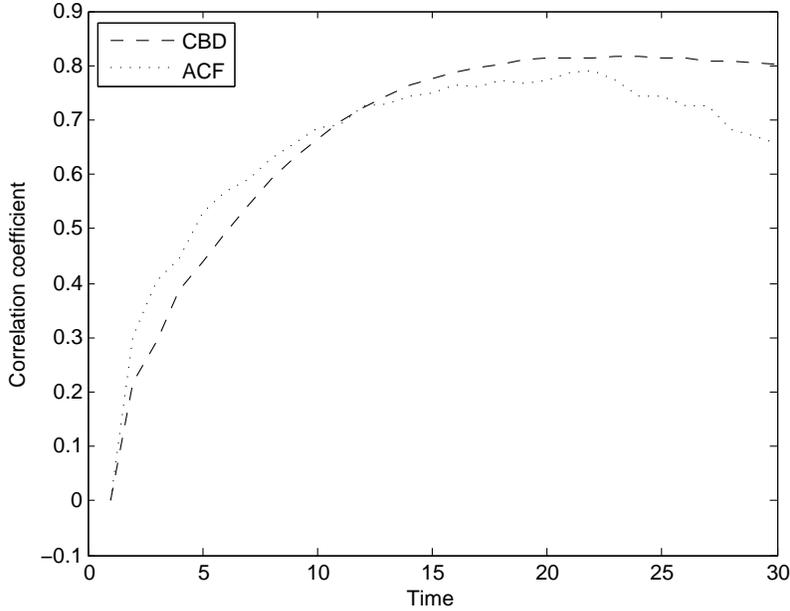


Figure 3.3: Change with time to maturity

correlation coefficients  $\rho(q_{59+t,t}^{(H)}, q_{75,t+9}^{(R)})$  for ACF are generally larger, however, as  $t$  increases,  $\rho(q_{59+t,t}^{(H)}, q_{75,t+9}^{(R)})$  for CBD model surpasses that from ACF model till the end of our hedging horizon  $T = 30$  years. Because the pension plan liability is more dependent on mortality rates in early years of the hedging horizon, intuitively it explains why there is much higher basis risk when the underlying model is the CBD model. A more comprehensive and rigorous analysis on the modeling issue of multi-population longevity models is beyond the scope of this chapter and worth more thorough investigations in the future.

### 3.5 Conclusion

In this chapter we study the optimal dynamic hedging strategy for pension plans that are exposed to longevity risk and basis risk. Under the commonly used longevity models such as the Augmented Common Factor model and the Cairns-Blake-Dowd model, we derive a semi closed-form

hedging plan for the pension plan sponsor which minimizes the variance of time-0 value of the total future liabilities. Further with a Monte Carlo simulation procedure and an approximation formula for forward mortality rates, we show that numerical calculation of optimal hedging strategy only requires moderate computational time and therefore it offers an effective solution for practical applications.

As the theoretically best strategy under the variance criterion, our proposed hedging scheme is also shown to outperform the “delta” strategy in terms of both effectiveness and stability. Additionally, extensive numerical results show that the hedging performance achieved by our hedging plan is robust to the hedging instrument we use and the underlying longevity model we choose. One useful discovery is that for dynamic longevity hedging plans, it is unnecessary to choose those long-maturity longevity instruments that tend to be more costly and less liquid, in order to achieve a satisfactory hedging performance.

# Chapter 4

## Index Insurance Design

### 4.1 Introduction

The purpose of this chapter is to provide an in-depth analysis on a class of insurance known as the index-indemnifying insurance, or simply the index insurance. As opposed to the traditional loss-indemnifying insurance for which its payout (indemnity payment) is a function of the actual loss incurred by the policyholder, the payout of an index insurance depends exclusively on a pre-determined index or some appropriately chosen indicators.

Prominent applications of index insurance can be found in insurance coverage provided to agricultural producers. In fact in recent years there is a surge of interest in piloting index insurance for agricultural households in developing economics. In these applications, an index may be an average county crop yield, the number of heating days, the amount of rainfall received by a particular area during the growing season, or based on remote sensing satellite vegetation data. For example Barnett and Mahul (2007) discuss the use of weather index insurance for agriculture in rural areas of lower-income countries. Chantararat et al. (2007) demonstrate that an index insurance with payout linked to some weather variables can be effective in improving drought response for famine prevention. Chantararat et al. (2013) describe an index-based livestock insurance by exploiting remote sensing vegetation data. Bokusheva et al. (2016) analyze the effectiveness of the indices constructed based on the satellite-based vegetation health indices

for insuring against drought-related yield losses. See also International Fund for Agricultural Development World Food Program (2010), Conradt et al. (2015), Carter et al. (2016) for recent advances in agricultural index insurance.

Other than hedging agricultural and livestock risks, index-based securities that are issued in the capital market have been effective in securitizing the catastrophic risks. See for example catastrophic-loss index options for hedging hurricane risk (Cummins et al., 2004) and the Swiss-Re mortality bonds for hedging mortality risk.

The popularity of index insurance stems from a number of reasons. The first and foremost reason lies in its potential of reducing or even eliminating completely the moral hazard and adverse selections since the indemnity payments are based on an index that is transparent, well-defined, and cannot be manipulated by either the insured or the insurer. The second reason is its low operational cost (such as the cost associated with the underwriting, administration, loss assessment). Because the indemnity payments are completely determined by an index, there is no need to assess the losses actually incurred by the agricultural producers. The loss assessment can be expensive and prohibitive, especially in rural areas where accessibility can be problematic. The number of small agricultural households further aggravates the cost if insurer needs to assess loss for all households. As a result, the claim settlement can also be processed more efficiently and more timely whenever there is a claim from an index insurance.

Despite all the aforementioned advantages, the challenge with the index insurance is the basis risk, which arises due to the discrepancy between the indemnity payments dictated by the index and the actual losses incurred by the insured. The imperfect correlation between the adopted index and the loss random variable casts doubt on the effectiveness of index insurance in hedging agricultural production risk and as such leads to low demand in some pilot index insurance programs. See, for example, Miranda and Farrin (2012) for a review of recent theoretical and empirical research on index insurance for developing countries and a summary of lessons learned from index insurance projects implemented in the developing world since 2000. See also Elabed, et al. (2013) and Jensen, et al. (2016) for additional discussion on basis risk associated with agricultural and livestock productions, respectively.

The presence of basis risk implies that the index must be chosen meticulously. A logical line

of inquiry is the determination of an index that optimally minimizes the basis risk. This basically relates to the optimal design of index insurance. The optimal design of loss-indemnifying contract is a widely studied problem in the actuarial literature. It is, however, important to point out some subtle differences between the formulation of optimal loss-indemnifying contract and optimal index-indemnifying contract. More specifically, the indemnity function in a loss-indemnifying insurance contract needs to be non-decreasing, bounded from above by the actual loss, and has a non-zero deductible, in order to avoid moral hazard. In contrast, the indemnity function of an index insurance can have very flexible structure. The indemnity is not necessarily increasing in the underlying indices. The indemnity payment can even exceed the loss incurred by the insured.

While index insurance is prevalent in agricultural production, its indemnity function in most cases, is relatively simple and is of linear type (e.g., Giné et al., 2007; Okhrin et al., 2013). While linear-type indemnity functions may work well in certain contexts, the basis risk is generally high in most cases. For example, in the context of agricultural insurance, the dependence structure between crop yields and weather indices such as temperature and precipitation are so complex that it cannot be accurately captured by a linear function. Thus, innovative weather index insurance products need to be developed for farmers to better protect against the decline in crop yields due to adverse weather conditions.

In this chapter, we adopt a utility maximization framework for the design of index insurance (Raviv, 1979) and define the optimal index insurance as the one that maximizes the insureds' expected utility. The variance minimization problem can be viewed as a special case in our general utility maximization framework when a quadratic utility function is adopted. Mahul (2000) and Vercaemmen (2001) are the two relevant references. The mathematical models in these two references share a similar structure as the present chapter, but they considered the problem of optimal loss-indemnifying (instead of index-indemnifying) insurance design in the presence of background risk under a utility maximization framework. They derived a characterization equation for the optimal solution, and presented certain interpretations on the shape of the optimal indemnity function. However, neither of these two articles studied the existence and uniqueness of the optimal contract, or offered a feasible procedure for the derivation of the optimal indemnity function. Therefore, their results are not sufficient for insurers to design effective index based

insurance products.

We contribute to the literature in the following aspects. First, we provide a rigorous mathematical examination on the existence and uniqueness of the optimal index insurance arrangement. Second, explicit form of the optimal index insurance is derived for utility functions commonly adopted in insurance economics including quadratic and exponential utility functions. For a general strictly concave utility function, the optimal solution is characterized by an implicit ordinary differential equation (ODE), for which the solution can be easily obtained numerically, for example, by the Runge-Kutta method (Burden and Faires, 2001). Third, an empirical agricultural index insurance is conducted and it shows that the index based contract from our results significantly outperforms those existing index contracts from the literature. Choosing the average temperature as the underlying index, we find that the optimal indemnity function generally follows a “first decreasing and then increasing” pattern and its specific shape relies on the premium level charged by the insurance contract, the maximum indemnity paid and the form of utility function. For quadratic utility function, the design is equivalent to minimizing the variance of insured’s resulting position, and our numerical results show that the effectiveness in terms of variance deduction does not continue to improve with the premium level after the premium exceeds certain threshold. This observation provides important and useful insights for government agency in making agricultural insurance premium subsidiaries. Further, our results also show that the proposed optimal contract generally outperforms the linear-type insurance contracts, and that the multi-index contracts can further reduce basis risk, when compared to the single-index ones.

The rest of this chapter is organized as follows. Section 4.2 describes the problem formulation of index insurance we will study in the chapter. Section 4.3 discusses the existence and uniqueness of the optimal index insurance contract for our formulation. Section 4.4 provides an ODE based method for the computation of the optimal solution, and applies this method to derive explicit optimal solution for quadratic and exponential utility functions, respectively. Section 4.5 provides an empirical study on the viability of our proposed optimal index insurance to weather index insurance. Section 4.6 concludes the chapter. A numerical procedure for solving the ODE arising from Section 4.4 is described in Appendix C.

## 4.2 Problem Setup

Suppose that a potential loss, which can hardly be insured or well hedged by any existing insurance of financial program on the market, is modeled by a random variable  $Y$ . Throughout this chapter, all the random variables are defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Our objective is to design an index-based insurance which is linked to an index  $X$  to protect an insured from such a risk. Let  $[c, d]$  and  $[a, b]$  with  $c < d$  and  $a < b$  be the supports of  $X$  and  $Y$  respectively. Further we assume that  $X$  and  $Y$  have a joint probability density function  $f(x, y)$  for  $(x, y) \in [c, d] \times [a, b]$  so that  $\int_c^d \int_a^b f(x, y) dy dx = 1$ . In this chapter, we assume that  $f(x, y)$  is continuous on  $[c, d] \times [a, b]$ , and we write the marginal density functions for  $X$  and  $Y$  respectively as follows

$$h(x) := \int_a^b f(x, y) dy \text{ for } x \in [c, d], \text{ and } g(y) := \int_c^d f(x, y) dx \text{ for } y \in [a, b].$$

Obviously  $g(y)$  and  $h(x)$  are continuous on  $[a, b]$  and  $[c, d]$  respectively. Additionally, we assume that  $f(x, y) > 0$  on  $[c, d] \times [a, b]$  a.e., and thus  $h(x) > 0$  a.e. on  $[c, d]$  and  $g(y) > 0$  a.e. on  $[a, b]$ .

Let  $I(X)$  be the indemnity function of the index insurance. This means that the actual payoff of the insurance is completely determined by the realization of the index  $X$ . We further assume that  $0 \leq I(X) \leq M$  for a constant  $M > 0$  which represents the maximum amount paid by the insurer. The maximum amount paid,  $M$ , is assumed exogenously in our discussion. It is possible to have  $M \geq b$  because the insured may want to over-insure its underlying for large losses in an incomplete market (Doherty and Schlesinger, 1983). Mathematically, we consider the following feasible set for the indemnity function in the design of index insurance:

$$\mathcal{I} := \{I \mid I : [c, d] \rightarrow [0, M] \text{ is measurable}\}.$$

For loss-indemnifying insurance where the payoff of the insurance contract depends on the actual loss occurred on the insured, the indemnity function is typically non-decreasing and bounded from above by the actual loss, and has a non-zero deductible, in order to preclude severe moral hazard from the insurance contract (e.g., Chi and Weng, 2013; Zhuang et al., 2016). For the

design of index insurance, however, we do not need to impose these restrictions on the indemnity function because the index can hardly be manipulated by either the insured or the insurer and thus no moral hazard is involved.

In this chapter we assume that the price of this insurance product is determined by the expected value premium principle:

$$P = \gamma \mathbb{E}[I(X)] = \gamma \int_c^d I(x)h(x) dx,$$

where  $\gamma - 1 \geq 0$  is the safety loading factor. For a given insurance premium level  $P \in (0, \gamma M)$ , the insurer aims to design an optimal insurance that maximizes its clients' expected utility. In other words, we are interested in solving the following optimization problem:

$$\begin{cases} \sup_{I \in \mathcal{I}} J(I) := \mathbb{E}\{U(w + I(X) - Y - (1 - \theta)P)\} \\ \text{s.t. } P = \gamma \int_c^d I(x)h(x) dx \end{cases} \quad (4.1)$$

where  $U$  is a strictly concave and non-decreasing utility function for the insured with  $U'(x) \geq 0$  and  $U''(x) < 0$  for  $x$  in the domain of the utility function  $U$ ,  $U''(x)$  is a continuous function,  $0 \leq \theta \leq 1$  denotes any possible subsidy to the insured by a third party (which is usually a government agency in practice),  $w$  is the initial wealth of the insured, and thus,  $w + I(X) - Y - (1 - \theta)P$  denotes the terminal wealth of the insured in the presence of an index insurance. The constraint  $P = \gamma \int_c^d I(x)h(x) dx$  may also be interpreted as the participation constraint for risk-neutral insurers when the insurance costs are proportional to the insurance payments (Raviv, 1979). We assume  $0 < P < \gamma M$  to ensure that the problem is well defined, and to exclude the trivial cases of  $P = 0$  or  $P = \gamma M$ , where the optimal indemnity is either zero or the upper bound  $M$ . We note that it is very common among most countries for a government to subsidize farmers for purchasing agricultural insurance. The inclusion of  $\theta$  in model (4.1) is to reflect such a practice. In the special case of  $\theta = 0$ , no subsidy is assumed for the insured in the above model.

## 4.3 Existence and Uniqueness of the Optimal Design

### 4.3.1 Uniqueness of the optimal solution

Due to the strict convexity of the utility function, we have the following proposition regarding the uniqueness of optimal solution to the insurance design problem (4.1).

**Proposition 4.1** (Uniqueness of optimal solution). *The optimal solution to problem (4.1) is unique up to the equality almost everywhere if it exists.*

*Proof.* Let  $I_1$  and  $I_2$  be two optimal solutions to problem (4.1) with  $\mu(D) > 0$  where  $D := \{x \in [c, d] \mid I_1(x) \neq I_2(x)\}$  and  $\mu(D)$  denotes the Lebesgue measure of the set  $D$ . Denote  $I_\lambda(x) := \lambda I_1(x) + (1 - \lambda)I_2(x)$ ,  $x \in [c, d]$  for a constant  $\lambda \in (0, 1)$ . Obviously,  $I_\lambda$  is a feasible indemnity function for problem (4.1) because  $I_\lambda \in \mathcal{I}$  and it satisfies the constraint in problem (4.1). We also note that  $X$  has a positive density function on the interval  $[c, d]$ . Consequently,  $\mu(D) > 0$  implies  $\mathbb{P}(A) > 0$  where  $A := \{\omega \in \Omega \mid I_1(X) \neq I_2(X)\}$ .

Let  $v(P)$  denote the supremum value for problem (4.1). We must have  $v(P) < \infty$  because both  $I(X)$  and  $Y$  are bounded random variables. Thus, using the strict concavity, we obtain

$$\begin{aligned} J(I_\lambda) &= \mathbb{E} \left[ U \left( w + \lambda I_1(X) + (1 - \lambda)I_2(X) - Y - (1 - \theta)P \right) \right] \\ &> \lambda \mathbb{E} \left[ U \left( w + I_1(X) - Y - (1 - \theta)P \right) \right] \\ &\quad + (1 - \lambda) \mathbb{E} \left[ U \left( w + I_2(X) - Y - (1 - \theta)P \right) \right] \\ &= \lambda v(P) + (1 - \lambda)v(P) \\ &= v(P), \end{aligned}$$

which contradicts to the optimality of  $I_1$  and  $I_2$ . Thus, the optimal solution to problem (4.1) is unique up to the equality almost everywhere if it exists.  $\square$

### 4.3.2 Existence of optimal solution

In order to solve problem (4.1), we introduce the Lagrange multiplier  $\lambda$  and define:

$$\begin{aligned}
K(I, \lambda) &:= J(I) + \lambda \left( P - \gamma \int_c^d I(x)h(x) dx \right) \\
&= \mathbb{E}[U(w + I(X) - Y - (1 - \theta)P)] + \lambda \int_c^d [P - \gamma I(x)]h(x) dx \\
&= \int_c^d \int_a^b U(w + I(x) - y - (1 - \theta)P) f(x, y) dy dx + \int_c^d \lambda(P - \gamma I(x))h(x) dx \\
&= \int_c^d \left\{ \int_a^b U(w + I(x) - y - (1 - \theta)P) f(y|x) dy + \lambda(P - \gamma I(x)) \right\} h(x) dx,
\end{aligned} \tag{4.2}$$

where  $f(y|x) = f(x, y)/h(x)$  is the conditional density function of  $Y$  given  $X = x$ . By the continuity and positiveness of  $f(x, y)$  and  $h(x)$ ,  $f(y|x)$  is also continuous and positive for  $x \in [c, d]$  and  $y \in [a, b]$ . The optimal solution to problem (4.1) can be recovered by the maximizer of  $K(I, \lambda)$  defined in (4.2), as stated in the following Lemma.

**Lemma 4.1.** *Let  $I_\lambda$  denote the maximizer of  $K(I, \lambda)$  defined by equation (4.2) for every  $\lambda \in \mathbb{R}$ . If there exists  $\lambda^*$  such that  $\mathbb{E}[I_{\lambda^*}] = P/\gamma$ , then  $I^* := I_{\lambda^*}$  solves problem (4.1).*

*Proof.* Recall that  $v(P)$  denotes the supremum value of problem (4.1). Therefore,

$$\begin{aligned}
v(P) &= \sup_{I \in \mathcal{I} \text{ s.t. } \gamma \mathbb{E}[I] = P} \mathbb{E}[J(I)] \\
&= \sup_{I \in \mathcal{I} \text{ s.t. } \gamma \mathbb{E}[I] = P} \{ \mathbb{E}[J(I)] + \lambda^*(P - \gamma \mathbb{E}[I]) \} \\
&\leq \sup_{I \in \mathcal{I}} \{ \mathbb{E}[J(I)] + \lambda^*(P - \gamma \mathbb{E}[I]) \} \\
&= \mathbb{E}[J(I^*)] + \lambda^*(P - \gamma \mathbb{E}[I^*]) \\
&= \mathbb{E}[J(I^*)] \\
&\leq \sup_{I \in \mathcal{I} \text{ s.t. } \gamma \mathbb{E}[I] = P} \mathbb{E}[J(I)] \\
&= v(P),
\end{aligned}$$

which implies that  $I^*$  is the solution of problem (4.1).  $\square$

By virtue of Lemma 4.1, we first investigate the maximizer of the function  $K(I, \lambda)$  with respect to  $I$  for a given  $\lambda \in \mathbb{R}$ . In view of equation (4.2), a sufficient condition is to pointwise maximize its integrand:

$$H(I(x), x, \lambda) := \int_a^b U(w + I(x) - y - (1 - \theta)P) f(y|x) dy + \lambda(P - \gamma I(x)), \quad x \in [c, d]. \quad (4.3)$$

The derivative of  $H(I(x), x, \lambda)$  with respect to  $I(x)$  is given by

$$\dot{H}(I(x), x, \lambda) := G(I(x), x) - \lambda\gamma, \quad (4.4)$$

where

$$\begin{aligned} G(\xi, x) &:= \int_a^b U'(w + \xi - y - (1 - \theta)P) f(y|x) dy \\ &= \mathbb{E}[U'(w + \xi - Y - (1 - \theta)P) | X = x]. \end{aligned} \quad (4.5)$$

We note that  $G(\xi, x)$  is strictly decreasing in  $\xi$  for any fixed  $x$ , since  $U$  is strictly concave. Accordingly,  $G(\xi, x)$  attains its maximum value at  $\xi = 0$  and its minimum value at  $\xi = M$  for a given  $x$ . Based on this fact, we define the following three sets:

$$S_1^\lambda := \left\{ x \in [c, d] \mid G(0, x) < \lambda\gamma \right\}, \quad (4.6)$$

$$S_2^\lambda := \left\{ x \in [c, d] \mid G(M, x) > \lambda\gamma \right\}, \quad (4.7)$$

$$S_3^\lambda := \left\{ x \in [c, d] \mid G(M, x) \leq \lambda\gamma \leq G(0, x) \right\}. \quad (4.8)$$

Since  $G(\xi, x)$  is strictly decreasing in  $\xi$  for a fixed  $x$ , we must have  $S_1^\lambda \cap S_2^\lambda = \emptyset$ , and thus  $S_1^\lambda$ ,

$S_2^\lambda$  and  $S_3^\lambda$  constitute a partition of the interval  $[c, d]$ . Consequently, it is obvious to have

$$I_\lambda(x) := \operatorname{argmin}_{I(x) \in [0, M]} H(I(x), x, \lambda) = \begin{cases} 0, & \text{for } x \in S_1^\lambda, \\ M, & \text{for } x \in S_2^\lambda, \\ \widehat{I}_\lambda(x), & \text{for } x \in S_3^\lambda, \end{cases} \quad (4.9)$$

where  $\widehat{I}_\lambda(x)$  satisfies  $\dot{H}(\widehat{I}_\lambda(x), x, \lambda) = 0$ , i.e.,

$$G(\widehat{I}_\lambda(x), x) = \lambda\gamma. \quad (4.10)$$

Obviously  $\dot{H}(0, x, \lambda) \geq 0$  and  $\dot{H}(M, x, \lambda) \leq 0$  for  $x \in S_3^\lambda$ . Thus, the continuity and strictly increasing property of  $\dot{H}(I(x), x, \lambda)$  as a function of  $I(x)$  implies that there exists a unique solution  $\widehat{I}_\lambda(x) \in [0, M]$  to equation (4.10) for every  $x \in S_3^\lambda$ .

**Remark 4.1.** *The partition by the three sets  $S_1^\lambda$ ,  $S_2^\lambda$  and  $S_3^\lambda$  for the index  $X$  represents different levels of insurance coverage for the insured. The expression of  $H(I(x), x, \lambda)$  given in (4.3) implies that its maximizer strives to keep a balance between the marginal utility gained and the marginal expense on insurance premium from an increase of insurance coverage. When the index value lies in the set  $S_1^\lambda$ , the marginal utility gained from each unit of insurance coverage is less than the marginal cost of premium, and thus a zero insurance coverage is optimal. For the index value on  $S_2^\lambda$ , the marginal utility for each unit of insurance coverage is larger than the marginal cost of insurance premium, and thus the maximum coverage is optimal. On  $S_3^\lambda$ , the optimal coverage makes the marginal benefit of utility equal to the marginal cost of insurance premium. The insurance coverage for index on the set  $S_3^\lambda$  is between 0 and  $M$ , and thus  $S_3^\lambda$  represents the relatively medium coverage region.*

In the rest of the section, we use Lemma 4.1 to show the existence of a solution to problem (4.1). We need to verify the existence of  $\lambda^*$  such that  $E[I_{\lambda^*}] = P/\gamma$  for  $I_\lambda$  in equation (4.9). To this end, we impose the following technical conditions:

**H1:**  $\mu(\{x \in [c, d] | G(0, x) = k_1\}) = \mu(\{x \in [c, d] | G(M, x) = k_2\}) = 0$  for any  $k_1, k_2 \in \mathbb{R}$ .

The above condition means that the level sets have a zero Lebesgue measure at any level for both functions  $G(0, x)$  and  $G(M, x)$ . This condition is quite mild from a practical point of view. For example, when the function  $G(M, x)$  is piecewise strictly monotonic over  $[c, d]$ , then condition **H1** is satisfied. In the context of optimal insurance or risk sharing with background risk, the concept of stochastic monotonicity is commonly used to describe the dependence structure between two random variables (e.g., Dana and Scarsini, 2007). A random variable  $Z_1$  is (strictly) stochastically monotonic in  $Z_2$ , if the map  $z \mapsto \mathbb{E}[f(Z_1)|Z_2 = z]$  is (strictly) monotonic for every (strictly) monotonic function  $f$ . Obviously, condition **H1** is satisfied when the actual loss variable  $Y$  is strictly stochastically monotonic in the index variable  $X$ . For index insurance design, the stochastic monotonicity is generally too strong to apply, but condition **H1** is general enough for most applications.

**Proposition 4.2** (Existence of optimal solution). *Assume that condition **H1** holds and  $P \in (0, \gamma M)$ . Then, there exists  $\lambda^*$  to satisfy  $\mathbb{E}[I_{\lambda^*}] = P/\gamma$  for  $I_\lambda$  defined by equations (4.9) and (4.10). In this case,  $I_{\lambda^*}$  is the optimal solution to problem (4.1).*

*Proof.* We only need to show the existence of  $\lambda^*$  to satisfy  $\mathbb{E}[I_{\lambda^*}] = P/\gamma$ , because this combined with Lemma 4.1 implies the optimality of  $I_{\lambda^*}$  for problem (4.1).

For  $x \in [c, d]$ , define  $\lambda_U := \max_{x \in [c, d]} \frac{1}{\gamma} G(0, x)$  and  $\lambda_L := \min_{x \in [c, d]} \frac{1}{\gamma} G(M, x)$ . Condition **H1** implies that both  $S_1^{\lambda_U}$  and  $S_2^{\lambda_L}$  differ from the set  $[c, d]$  by only a  $\mu$ -null set. Thus, by (4.9),  $\mathbb{E}[I_{\lambda_U}(X)] = 0$  and  $\mathbb{E}[I_{\lambda_L}(X)] = M$ . As a result, it is sufficient to show that  $\mathbb{E}[I_\lambda(X)]$  is continuous on  $[\lambda_L, \lambda_U]$ .

Below we only show the right continuity of  $\mathbb{E}[I_\lambda(X)]$  on  $[\lambda_L, \lambda_U]$ , as its left continuity follows in the same fashion. Define  $\Delta_\epsilon^\lambda := |\mathbb{E}[I_{\lambda+\epsilon}(X)] - \mathbb{E}[I_\lambda(X)]|$  for  $\lambda \in [\lambda_L, \lambda_U - \epsilon]$  and  $\epsilon > 0$ . Then,

$$\begin{aligned} \Delta_\epsilon^\lambda &= \left| M \cdot \mathbb{P}(X \in S_2^{\lambda+\epsilon}) + \int_{S_3^{\lambda+\epsilon}} I_{\lambda+\epsilon}(x) h(x) dx - M \cdot \mathbb{P}(X \in S_2^\lambda) - \int_{S_3^\lambda} I_\lambda(x) h(x) dx \right| \\ &\leq M \cdot |\mathbb{P}(X \in S_2^{\lambda+\epsilon}) - \mathbb{P}(X \in S_2^\lambda)| + \left| \int_{S_3^{\lambda+\epsilon}} I_{\lambda+\epsilon}(x) h(x) dx - \int_{S_3^\lambda} I_\lambda(x) h(x) dx \right| \\ &=: M |J_2^\epsilon| + |J_3^\epsilon|. \end{aligned}$$

By definition of  $S_2^\lambda$  in (4.7),  $J_2^\epsilon = -\mathbb{P}(\lambda\gamma < G(M, X) \leq (\lambda + \epsilon)\gamma) = \mathbb{P}(G(M, X) \leq \lambda\gamma) - \mathbb{P}(G(M, X) \leq (\lambda + \epsilon)\gamma) \rightarrow 0$  as  $\epsilon \rightarrow 0^+$ .

It remains to show  $J_3^\epsilon \rightarrow 0$  as  $\epsilon \rightarrow 0^+$ . Indeed,

$$\begin{aligned} J_3^\epsilon &= \int_{S_3^{\lambda+\epsilon} \cap \overline{S_3^\lambda}} I_{\lambda+\epsilon}(x)h(x) dx - \int_{S_3^\lambda \cap \overline{S_3^{\lambda+\epsilon}}} I_\lambda(x)h(x) dx \\ &\quad + \int_{S_3^{\lambda+\epsilon} \cap S_3^\lambda} (I_{\lambda+\epsilon}(x) - I_\lambda(x)) h(x) dx, \end{aligned} \quad (4.11)$$

where  $\bar{A}$  denotes the complement of a set  $A$ . It is easy to verify

$$S_3^{\lambda+\epsilon} \cap \overline{S_3^\lambda} = \{x \in [c, d] \mid \lambda\gamma < G(M, x) \leq (\lambda + \epsilon)\gamma \leq G(0, x)\},$$

and

$$S_3^\lambda \cap \overline{S_3^{\lambda+\epsilon}} = \{x \in [c, d] \mid G(M, x) \leq \lambda\gamma \leq G(0, x) < (\lambda + \epsilon)\gamma\}.$$

Condition **H1** implies both  $\mu\left(S_3^{\lambda+\epsilon} \cap \overline{S_3^\lambda}\right) \rightarrow 0$  and  $\mu\left(S_3^\lambda \cap \overline{S_3^{\lambda+\epsilon}}\right) \rightarrow 0$  as  $\epsilon \rightarrow 0^+$ . Further noting  $I_{\lambda+\eta}(x) \in [0, M]$  for any  $\eta \geq 0$  and  $x \in [c, d]$ , the first two items in (4.11) converge to 0 as  $\epsilon \rightarrow 0^+$ .

For  $x \in S_3^{\lambda+\epsilon} \cap S_3^\lambda$ , we have

$$G(I_{\lambda+\epsilon}(x), x) = (\lambda + \epsilon)\gamma \text{ and } G(I_\lambda(x), x) = \lambda\gamma.$$

Further,  $G(\xi, x)$  is a continuous and differentiable function of  $\xi$  for any  $x \in [c, d]$ , and thus, we apply the mean value theorem to obtain

$$\epsilon\gamma = G(I_{\lambda+\epsilon}(x), x) - G(I_\lambda(x), x) = G'_\xi(\xi, x) (I_{\lambda+\epsilon}(x) - I_\lambda(x))$$

for some constant  $\xi := \xi_{x,\lambda}$  valued between  $I_{\lambda+\epsilon}(x)$  and  $I_\lambda(x)$ , where  $G'_\xi(\xi, x) := \frac{\partial}{\partial \xi} G(\xi, x)$ .

This implies

$$|I_{\lambda+\epsilon}(x) - I_\lambda(x)| = \frac{\epsilon\gamma}{|G'_\xi(\xi, x)|} \leq \frac{\epsilon\gamma}{\inf_{\xi \in [0, M], x \in [c, d]} |G'_\xi(\xi, x)|}. \quad (4.12)$$

From (4.5), we have  $G'_\xi(\xi, x) = \int_a^b U''(w + \xi - y - (1 - \theta)P) f(y|x) dy$ . Since  $U''(\cdot)$  is a continuous function over its domain, there exists a constant  $\delta > 0$  such that

$$|G'_\xi(\xi, x)| \geq \delta \int_a^b f(y|x) dy = \delta, \quad \forall \xi \in [0, M] \text{ and } x \in [c, d].$$

Consequently, it follows from (4.12) that

$$\int_{S_3^{\lambda+\epsilon} \cap S_3^\lambda} (I_{\lambda+\epsilon}(x) - I_\lambda(x)) h(x) dx \leq \int_{S_3^{\lambda+\epsilon} \cap S_3^\lambda} \frac{\epsilon\gamma}{\delta} h(x) dx \leq \epsilon \frac{\gamma(d-c)}{\delta} \rightarrow 0, \quad \text{as } \epsilon \rightarrow 0^+.$$

Therefore, from (4.11) we have  $J_3^\epsilon \rightarrow 0$  as  $\epsilon \rightarrow 0^+$ , by which we complete the proof.  $\square$

## 4.4 Computing the Optimal Solution

In the previous section we have demonstrated the existence and uniqueness of the optimal insurance contract for problem (4.1). In order to derive a closed-form expression for the optimal solution, one may invoke Proposition 4.2. However, this involves the determination of the specific forms of the sets  $S_2^{\lambda^*}$  and  $S_3^{\lambda^*}$ , as well as solving the equation (4.10) for  $\widehat{I}_{\lambda^*}(x)$ , where  $\lambda^*$  is given in Proposition 4.2. Recall that  $\widehat{I}_{\lambda^*}(x)$  is defined on the set  $S_3^{\lambda^*}$  only and it is solved from equation (4.10) as the unique solution. In this section we consider the case where the analytical form of  $\widehat{I}_{\lambda^*}(x)$  derived from equation (4.10) can be extended to the whole interval  $[c, d]$ , i.e.,  $\widehat{I}_{\lambda^*}(x)$  is well defined for  $x \in [c, d]$ . We develop an ordinary differential equation (ODE) method which is more convenient for the derivation of the optimal solution. This ODE method will be demonstrated for quadratic and exponential utility functions in this section. For other strictly concave utility functions, a numerical procedure is attached in Appendix C for the derivation of optimal contract.

### 4.4.1 The ODE method

Lemma 4.2 below provides an equivalent but more computationally friendly way for deriving  $I_\lambda$  to maximize  $K(I, \lambda)$  in (4.2) when  $\widehat{I}_\lambda(x)$  derived from the equation (4.10) can be extended to the interval  $[c, d]$ . Note that  $\widehat{I}_\lambda(x)$  may no longer confine to the interval  $[0, M]$  for  $x$  outside the set  $S_3^\lambda$ .

**Lemma 4.2.** *Let  $\lambda$  be a constant such that  $S_3^\lambda \neq \emptyset$ , and assume that  $\widehat{I}_\lambda(x)$  solved from (4.10) exists on  $[c, d]$ . Then, the optimal solution to maximize  $K(I, \lambda)$  in (4.2) is given by*

$$I_\lambda(x) = \left[ \left( \widehat{I}_\lambda(x) \right) \vee 0 \right] \wedge M. \quad (4.13)$$

*Proof.* Without loss of generality, we assume that both  $S_1^\lambda$  and  $S_2^\lambda$  are non-empty. One can see from the rest of the proof that it is actually easier to show the desired result when either or both of the two sets are empty.

By equation (4.9), it is sufficient to show  $\widehat{I}_\lambda(x) \leq 0$  for  $x \in S_1^\lambda$  and  $\widehat{I}_\lambda(x) \geq M$  for  $x \in S_2^\lambda$ . In fact, if  $\widehat{I}_\lambda(x_1) > 0$  for some  $x_1 \in S_1$ , then

$$\begin{aligned} 0 &= G\left(\widehat{I}_\lambda(x_1), x_1\right) - \lambda\gamma \\ &= \int_a^b U'\left(w + \widehat{I}_\lambda(x_1) - y - (1 - \theta)P\right) f(y|x_1) dy - \lambda\gamma \\ &< \int_a^b U'(w - y - (1 - \theta)P) f(y|x_1) dy - \lambda\gamma \\ &= G(0, x_1) - \lambda\gamma \\ &< 0, \end{aligned}$$

where the first inequality is due to the strict convexity of  $U$  and the second one follows from the fact  $x_1 \in S_1^\lambda$ . The last display means  $0 < 0$ , a contradiction. Thus, we must have  $\widehat{I}_\lambda(x) \leq 0$  for  $x \in S_1^\lambda$ . We can use the same contradiction argument to show  $\widehat{I}_\lambda(x) \geq M$  for  $x \in S_2^\lambda$ .  $\square$

The advantage of Lemma 4.2 lies in the fact that we do not need to determine the sets  $S_2^\lambda$  and  $S_3^\lambda$  for the determination of the optimal solution  $I_\lambda(x)$ . Once we derive an analytical form

of  $\widehat{I}_\lambda(x)$  by solving equation (4.10) for  $x \in [c, d]$ , the optimal solution  $I_\lambda(x)$  can be derived via equation (4.13).

We can apply Lemma 4.2 to transform problem (4.1) into an ODE problem under certain smoothness condition for  $f(x, y)$  as shown in Proposition 4.4 in the sequel. The ODE method relies on the analytical continuation of  $\widehat{I}_{\lambda^*}(x)$  from  $S_3^{\lambda^*}$  to  $[c, d]$ , and thus, we need to make sure  $S_3^{\lambda^*}$  is non-empty to make the ODE method valid, where  $\lambda^*$  is given in Proposition 4.2. The following proposition confirms the non-emptiness of  $S_3^{\lambda^*}$ .

**Proposition 4.3.** *Assume that condition **H1** is satisfied. Then,  $S_3^{\lambda^*}$  is a non-empty subset of  $[c, d]$ , where  $\lambda^*$  is any constant such that  $E[I_{\lambda^*}(X)] = P/\gamma$  with the existence guaranteed by Proposition 4.2.*

*Proof.* We prove the proposition by contradiction. Suppose  $S_3^{\lambda^*} = \emptyset$ . Then, it must be one of the following three scenarios:

Case 1:  $S_1^{\lambda^*} = [c, d]$ ,

Case 2:  $S_2^{\lambda^*} = [c, d]$ ,

Case 3:  $S_1^{\lambda^*} \cup S_2^{\lambda^*} = [c, d]$ ,  $S_1^{\lambda^*} \neq \emptyset$  and  $S_2^{\lambda^*} \neq \emptyset$ .

For Case 1, it follows from (4.9) that  $I^*(x) = 0, \forall x \in [c, d]$  and thus  $P = \gamma E[I_{\lambda^*}(X)] = 0$ . Similarly, for Case 2  $I^*(x) = M, \forall x \in [c, d]$  and thus  $P = \gamma E[I_{\lambda^*}(X)] = \gamma M$ . Since the insurance premium budget  $P \in (0, \gamma M)$ , both Cases 1 and 2 are impossible.

Consider Case 3 and take  $x_1 \in S_1^{\lambda^*}$  and  $x_2 \in S_2^{\lambda^*}$ . By equations (4.6) and (4.7), and the fact that  $G(\xi, x)$  is strictly decreasing in  $\xi$ , we have

$$G(0, x_1) < \lambda\gamma < G(M, x_2) < G(0, x_2). \quad (4.14)$$

Further, since  $f(y|x)$  is continuous on  $(x, y) \in [c, d] \times [a, b]$ , it must be uniformly continuous on  $[a, b]$ . Therefore, for  $\epsilon > 0$ ,

$$|G(0, x + \epsilon) - G(0, x)|$$

$$\begin{aligned}
&= \left| \int_a^b U'(w - y - (1 - \theta)P) f(y|x + \epsilon) dy - \int_a^b U'(w - y - (1 - \theta)P) f(y|x) dy \right| \\
&\leq \max_{y \in [a, b]} |U'(w - y - (1 - \theta)P)| \cdot \int_a^b |f(y|x + \epsilon) - f(y|x)| dy \\
&\rightarrow 0, \text{ as } \epsilon \rightarrow 0^+, \forall x \in [c, d],
\end{aligned}$$

which implies the continuity of  $G(0, x)$  as a function of  $x$  on  $[c, d]$ . Thus, those inequalities in (4.14) imply the existence of a constant  $x_3$  between  $x_1$  and  $x_2$  to satisfy  $G(0, x_3) = \lambda\gamma$ . Again, by the strictly increasing property of  $G(\xi, x)$  in  $\xi$ , we have  $G(M, x_3) < G(0, x_3) = \lambda\gamma$ , which means that  $x_3 \in S_3^{\lambda*}$ . This contradicts to the assumption of  $S_3^{\lambda*} = \emptyset$ , and thus the proof is complete.  $\square$

The following proposition states that the index insurance design problem (4.1) can be solved by the ODE approach:

**Proposition 4.4.** *Suppose that the derivative  $\frac{\partial}{\partial x} f(y|x)$  exists and is continuous on  $[c, d] \times [a, b]$ , and a function  $\widehat{L} : [c, d] \mapsto \mathbb{R}$  solves the following ODE problem:*

$$\begin{cases} \frac{dL}{dx} = F(x, L), \\ P = \gamma E[(L(X) \vee 0) \wedge M], \end{cases} \quad (4.15)$$

where the function  $F : [c, d] \times \mathbb{R} \mapsto \mathbb{R}$  is defined by

$$F(x, L) := - \frac{\int_a^b U'(w + L - y - (1 - \theta)P) \frac{\partial}{\partial x} f(y|x) dy}{\int_a^b U''(w + L - y - (1 - \theta)P) f(y|x) dy}.$$

Then,  $L^* := (\widehat{L}(x) \vee 0) \wedge M$  is the optimal solution to problem (4.1).

*Proof.* By our assumption of  $U''(x) < 0$  and  $f(y|x) \equiv \frac{f(x,y)}{h(x)} > 0$ , a.e., for  $(x, y) \in [c, d] \times [a, b]$ , we have

$$\int_a^b U''(w + \widehat{L}(x) - y - (1 - \theta)P) f(y|x) dy < 0, \forall x \in [c, d],$$

and thus  $F(x, L)$  is well defined.

By Proposition 4.3,  $S_3^{\lambda^*} \neq \emptyset$  for any constant  $\lambda^*$  such that  $E[I_{\lambda^*}(X)] = P/\gamma$ , where  $I_{\lambda^*}(x)$  is defined in (4.9). If we could find a constant  $\lambda^*$  to satisfy  $E[I_{\lambda^*}(X)] = P/\gamma$  and show that  $\widehat{L}(x) = \widehat{I}_{\lambda^*}(x)$ , a.e., on  $S_3^{\lambda^*}$ , then Lemma 4.2, along with the fact that  $\widehat{L}(x)$  is well defined on  $[c, d]$ , implies that  $L^* = (\widehat{L}(x) \vee 0) \wedge M$  is the optimal solution to problem (4.1).

Since  $\widehat{L}(x)$  satisfies equation (4.15), we have

$$\int_a^b \left\{ U'' \left( w + \widehat{L}(x) - y - (1 - \theta)P \right) f(y|x) \frac{d\widehat{L}(x)}{dx} + U' \left( w + \widehat{L}(x) - y - (1 - \theta)P \right) \frac{\partial}{\partial x} f(y|x) \right\} dy = 0,$$

i.e.,

$$\frac{d}{dx} \int_a^b U' \left( w + \widehat{L}(x) - y - (1 - \theta)P \right) f(y|x) dy = 0, \quad x \in [c, d].$$

This implies

$$G(\widehat{L}(x), x) = \int_a^b U' \left( w + \widehat{L}(x) - y - (1 - \theta)P \right) f(y|x) dy = \lambda_0 \gamma, \quad x \in [c, d], \quad (4.16)$$

where the constant  $\lambda_0$  is defined as

$$\lambda_0 := \frac{1}{\gamma} \int_a^b U' \left( w + \widehat{L}(c) - y - (1 - \theta)P \right) f(y|c) dy.$$

The last two displays, together with the fact that the equation (4.10) has a unique solution  $\widehat{I}_\lambda(x)$  for every  $x \in S_3^{\lambda_0}$ , imply that  $\widehat{L}(x) = \widehat{I}_{\lambda^*}(x)$  on  $S_3^{\lambda^*}$  for  $\lambda^* = \lambda_0$ . Comparing (4.10) and (4.16), we see  $\widehat{L}$  and  $\widehat{I}_{\lambda^*}$  satisfy the same equation. Thus, from the proof of Lemma 4.2,  $\widehat{L}(x) \leq 0$  for  $x \in S_1^{\lambda^*}$  and  $\widehat{L}(x) \geq M$  for  $x \in S_2^{\lambda^*}$ . Further, the second equation in (4.15) obviously implies  $E[I_{\lambda^*}(X)] = P/\gamma$ , and thus, the proof is complete.  $\square$

In the next two sections we will demonstrate the applications of Proposition 4.4 for the derivation of optimal index insurance solutions for quadratic and exponential utility functions, respec-

tively.

#### 4.4.2 Quadratic utility

We suppose that the insured's utility function has a quadratic form, i.e.,  $U(x) = \alpha x - \beta x^2$ ,  $x \leq \frac{\alpha}{2\beta}$ , where the parameters  $\alpha > 0$  and  $\beta > 0$ . We also assume that  $w + M - a - (1 - \theta)P \leq \frac{\alpha}{2\beta}$  so that the insured's maximum possible wealth will not go beyond the domain of the utility function, and  $U'(x) \geq 0$ ,  $U''(x) < 0$  and continuity of  $U''(x)$  hold for every  $x$  in its domain. By invoking Proposition 4.4, we can derive a closed-form solution of optimal index insurance as shown in the following proposition.

**Proposition 4.5.** *Suppose that  $\frac{\partial}{\partial x} f(y|x)$  exists and is continuous on  $[c, d] \times [a, b]$ . If the policyholder's utility function  $U(x) = \alpha x - \beta x^2$ ,  $x \leq \frac{\alpha}{2\beta}$ , where the parameters  $\alpha > 0$  and  $\beta > 0$ , then the optimal index insurance is given by*

$$I^*(x) = [(E[Y|X = x] + \eta^*) \vee 0] \wedge M, \quad (4.17)$$

where  $\eta^*$  is determined by the equation

$$E[I^*(X)] = E\{[(E[Y|X] + \eta^*) \vee 0] \wedge M\} = \frac{P}{\gamma}.$$

*Proof.* With the given utility, the function  $F$  in equation (4.15) becomes,

$$\begin{aligned} F(x, L) &= -\frac{\int_a^b U'(w + L(x) - y - (1 - \theta)P) \frac{\partial}{\partial x} f(y|x) dy}{\int_a^b U''(w + L(x) - y - (1 - \theta)P) f(y|x) dy} \\ &= -\frac{\int_a^b (\alpha - 2\beta w - 2\beta L(x) + 2\beta(1 - \theta)P) \frac{\partial}{\partial x} f(y|x) dy + \int_a^b 2\beta y \frac{\partial}{\partial x} f(y|x) dy}{-2\beta} \\ &= \frac{(\alpha - 2\beta w - 2\beta L(x) + 2\beta(1 - \theta)P) \frac{\partial}{\partial x} \int_a^b f(y|x) dy + 2\beta \frac{\partial}{\partial x} \int_a^b y f(y|x) dy}{2\beta} \\ &= \frac{0 + 2\beta \frac{\partial}{\partial x} E[Y|X = x]}{2\beta} \end{aligned}$$

$$= \frac{\partial}{\partial x} \mathbb{E}[Y|X = x],$$

where we apply the fact that  $\int_a^b f(y|x) \, dy = 1$  and thus  $\frac{\partial}{\partial x} \int_a^b f(y|x) \, dy = 0$ .

Due to the existence and continuity of  $\frac{\partial}{\partial x} f(y|x)$ ,  $\frac{\partial}{\partial x} \mathbb{E}[Y|X = x] = \int_a^b y \frac{\partial}{\partial x} f(y|x) \, dy$  exists for every  $x \in [c, d]$ . Therefore, a direct application of Proposition 4.4 implies the following optimal index insurance

$$I^*(x) = [(\mathbb{E}[Y|X = x] + \eta^*) \vee 0] \wedge M,$$

given that a constant  $\eta^*$  exists to satisfy

$$\mathbb{E}[I^*(X)] = \mathbb{E}\{[(\mathbb{E}[Y|X] + \eta^*) \vee 0] \wedge M\} = \frac{P}{\gamma}.$$

In fact,  $\mathbb{E}[I^*(X)]$  is apparently continuous and non-decreasing in  $\eta^*$ , and thus, a solution  $\eta^*$  must exist for the above equation. Therefore, the proof is complete.  $\square$

**Remark 4.2.** *It is well known that the one-period quadratic utility maximization problem is equivalent to the one-period mean-variance problem. It is trivial to show that the optimal indemnity function is still given by Proposition 4.5 if the insurer aims at minimizing the variance of the insured' terminal wealth.*

**Remark 4.3.** (a) *As one can infer from Proposition 4.5, the optimal contract is independent of the parameters  $\alpha$  and  $\beta$  under the quadratic utility.*

(b) *Proposition 4.5 also shows that, under quadratic utility function, the optimal indemnity function is irrelevant to both the insured's initial wealth  $w$  and subsidy level  $\theta$ .*

**Remark 4.4.** *When the quadratic utility is adopted as the criterion, computation of the optimal index insurance is substantially simplified because it does not involve estimating the joint density function  $f(x, y)$ ,  $(x, y) \in [c, d] \times [a, b]$ , but only the conditional expectation function  $\mathbb{E}[Y|X = x]$ ,  $x \in [c, d]$ . In practice, it is sometimes possible to obtain a much quicker and more convenient estimation on  $\mathbb{E}[Y|X = x]$  directly without estimating  $f(x, y)$ .*

### 4.4.3 Exponential utility

In this section we consider the case when  $U$  is an exponential utility function, i.e.,  $U(x) = -\frac{1}{\alpha}e^{-\alpha x}$ , where the parameter  $\alpha > 0$ . It is easy to verify that  $U'(x) \geq 0$  and  $U''(x) < 0$  for all  $x \in \mathbb{R}$ . We invoke Proposition 4.4 to derive a closed-form solution of optimal index insurance as shown in the following proposition.

**Proposition 4.6.** *Suppose that  $\frac{\partial}{\partial x}f(y|x)$  exists and is continuous on  $[c, d] \times [a, b]$ . If  $U(x) = -\frac{1}{\alpha}e^{-\alpha x}$ ,  $x \in \mathbb{R}$ , with the utility parameter  $\alpha > 0$ , then the optimal indemnity function is given by*

$$I^*(x) = \left[ \left( \frac{1}{\alpha} \ln (\mathbb{E} [e^{\alpha Y} | X = x]) + \eta^* \right) \vee 0 \right] \wedge M,$$

where  $\eta^*$  is a constant determined by  $P = \gamma \mathbb{E} [I^*(X)]$ .

*Proof.* From the given utility, the function  $F$  in equation (4.15) becomes,

$$\begin{aligned} F(x, L) &= -\frac{\int_a^b U'(w + L(x) - y - (1 - \theta)P) \frac{\partial}{\partial x}f(y|x) dy}{\int_a^b U''(w + L(x) - y - (1 - \theta)P) f(y|x) dy} \\ &= -\frac{\int_a^b e^{-\alpha(w+L(x)-y-(1-\theta)P)} \frac{\partial}{\partial x}f(y|x) dy}{\int_a^b (-\alpha)e^{-\alpha(w+L(x)-y-(1-\theta)P)} f(y|x) dy} \\ &= \frac{1}{\alpha} \cdot \frac{e^{-\alpha(w+L(x)-(1-\theta)P)} \int_a^b e^{\alpha y} \frac{\partial}{\partial x}f(y|x) dy}{e^{-\alpha(w+L(x)-(1-\theta)P)} \int_a^b e^{\alpha y} f(y|x) dy} \\ &= \frac{1}{\alpha} \frac{\frac{\partial}{\partial x} \mathbb{E} [e^{\alpha Y} | X = x]}{\mathbb{E} [e^{\alpha Y} | X = x]} \\ &= \frac{\partial}{\partial x} \left\{ \frac{1}{\alpha} \ln (\mathbb{E} [e^{\alpha Y} | X = x]) \right\}. \end{aligned}$$

Similar to the proof of Proposition 4.5, the continuity of  $\frac{\partial}{\partial x}f(y|x)$  and  $f(y|x)$  implies that both  $\frac{\partial}{\partial x} \mathbb{E} [e^{\alpha Y} | X = x]$  and  $\mathbb{E} [e^{\alpha Y} | X = x]$  exist for every  $x \in [c, d]$ , and thus

$$\frac{\partial}{\partial x} \left\{ \frac{1}{\alpha} \ln (\mathbb{E} [e^{\alpha Y} | X = x]) \right\}$$

also exists for every  $x \in [c, d]$ . Then, by invoking Proposition 4.4, we derive the optimal index insurance as follows

$$I^*(x) = \left[ \left( \frac{1}{\alpha} \ln (\mathbb{E} [e^{\alpha Y} | X = x]) + \eta^* \right) \vee 0 \right] \wedge M,$$

given the existence of a constant  $\eta^*$  to satisfy  $P = \gamma \mathbb{E} [I^*(X)]$ . Apparently,  $\mathbb{E} [I^*(X)]$  is continuous and non-decreasing in  $\eta^*$ , which indicates the existence of  $\eta^*$ . Hence, the proof is complete.  $\square$

**Remark 4.5.** *Similar to the quadratic utility case as we commented in Remark 4.4, we only need to estimate the conditional expectation  $\mathbb{E} [e^{\alpha Y} | X = x]$ ,  $x \in [c, d]$ , in order to determine the optimal index insurance under the exponential utility function. We do not have to estimate the joint density function  $f(x, y)$  in real data applications.*

**Remark 4.6.** *Proposition 4.6 also indicates that the optimal indemnity function under the exponential utility is irrelevant to both the insured's initial wealth  $w$  and the subsidy level  $\theta$ . This is a similar fact to those comments we made in Remark 4.3 for optimal index insurance under the quadratic utility function.*

## 4.5 Applications in Weather Index Insurance Design

In this section we apply our theoretical results to an example of weather index insurance contract design, where basis risk is a primary concern for policyholders. For sections 4.5.1-4.5.5, we choose the temperature as the underlying index to protect insured's position from adverse weather conditions. We investigate the optimal index insurance under the quadratic, exponential and logarithmic utility functions, respectively. We also conduct certain comparison study between our optimal index design and linear-type contracts. Finally, we extend the optimal index design to the bivariate case where the indemnity depends on two indices, the temperature and the precipitation.

### 4.5.1 Dependence modeling

We choose the average temperature of the whole product growing cycle as the underlying index to protect insured's position from adverse weather conditions. For a certain kind of agricultural product, both too high and too low temperatures would normally have an adverse impact on the product yield. As a result, the indemnity function of a well designed contract should take larger values at both ends of the interval at temperature axis but smaller values in the middle area. In our study we use county-level data of rice yield and temperature data in Jiangsu Province, China during the period from 1992 to 2011. The same data set has been studied by Shi and Jiang (2016). We first remove the trend in the historical yield which represents production improvement factors over time such as technology, and then define the actual loss variable to the insured  $Y$  as the highest detrended yield during the last 20 years less the detrended yield. We apply kernel smoothing method to calibrate the joint density between the average temperature and the actual loss variables. The graph of the joint density function  $f(x, y)$  is illustrated in Figure 4.1.

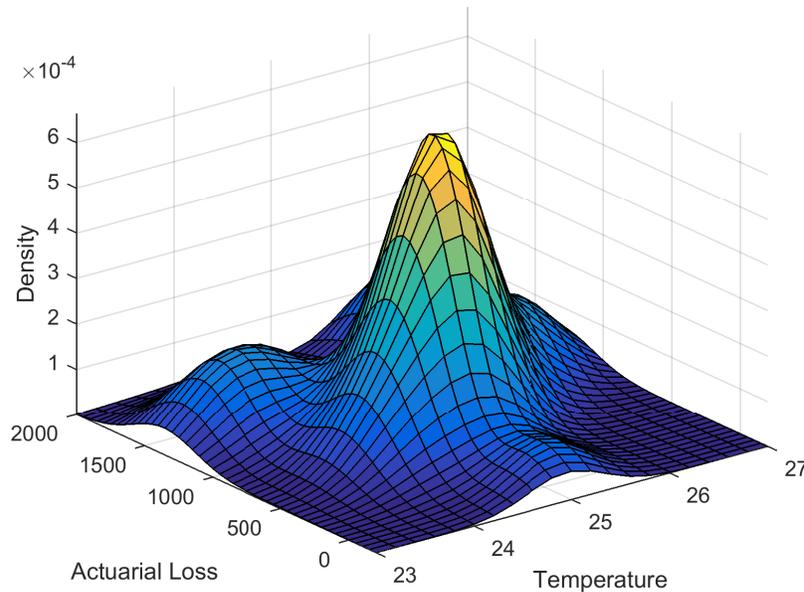


Figure 4.1: Joint density  $f(x, y)$  of the actual loss and the average temperature.

## 4.5.2 Quadratic utility

We begin with investigation into the shape of the optimal index insurance contract when the insured's risk preference is represented by a quadratic utility function. As we discussed in Remark 4.2, the optimal indemnity function is independent of the parameter values under a quadratic utility function, and therefore it is unnecessary to specify the parameter values for the investigation of the optimal index insurance. There are two major exogenous factors determining the shape and the scale of the optimal contract: the premium level  $P$  and the maximum indemnity level  $M$ . Weather index agricultural insurance is often subsidized by the government, and the premium level  $P$  also reflects the subsidy level because the latter is usually proportional to  $P$ . The maximum indemnity level  $M$  is also very important factor for index insurance design, because the insurer can prevent itself from extreme large losses by imposing such an upper limit.

### 4.5.2.1 Optimal contracts with different premium levels

We fix the maximum indemnity level at  $M = 300$ , and invoke Proposition 4.5 to construct the optimal indemnity functions for four different premium levels  $P$  varying from 20 to 80 in multiples of 20. The shapes of the optimal indemnity functions are illustrated in Figure 4.2 for each premium level.

Figure 4.2 shows the change in shape of the optimal indemnity function when we increase the premium level  $P$  from 20 to 80. The optimal indemnity function follows a “first decreasing and then increasing” pattern, which is in good accordance with our conjecture that a large portion of the premium paid by the insured goes to both ends of the weather index axis which indicates higher potential losses. Further, the slope of right half of the indemnity function is steeper than the left half, which means an asymmetric impact of temperature on the product yield: high temperatures have a more severely adverse impact on the crop yield than low temperatures do.

Recall from equation (4.9) that the indemnity from the optimal index insurance is zero over the set  $S_1^{\lambda^*}$ , attains the maximum amount  $M$  over  $S_2^{\lambda^*}$ , and lies between 0 and  $M$  over  $S_3^{\lambda^*}$ . Figure 4.2 indicates that, as the premium level  $P$  goes larger, the set  $S_1^{\lambda^*}$  diminishes and  $S_2^{\lambda^*}$  expands. When  $P = 20$ ,  $S_2^{\lambda^*} = \emptyset$ . This means that the premium level is too small to cover the

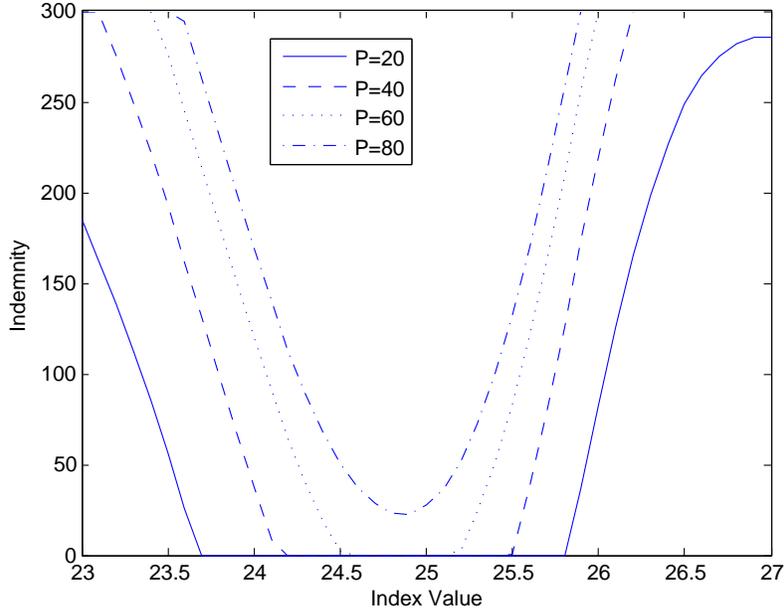


Figure 4.2: Optimal indemnity functions for different premium levels.

maximum indemnity level  $M = 300$  over any region. When  $P = 80$ ,  $S_1^{\lambda^*} = \emptyset$  which means that the premium level is large enough in this case to cover the whole range of the index variable.

#### 4.5.2.2 Optimal contracts with different maximum indemnity levels

In this section we fix the premium level  $P = 50$ , and investigate the shape of the optimal indemnity function for different maximum indemnity levels  $M$ . Numerical results are displayed in Figure 4.3.

Figure 4.3 shows that the optimal indemnity function also shows the “first decreasing and then increasing” pattern for all the different maximum indemnity levels we considered. As the maximum indemnity  $M$  increases, more premium transfers from the middle area to both tails, and thus the set  $S_1^{\lambda^*}$  becomes larger. In the meanwhile, the set  $S_2^{\lambda^*}$  becomes smaller as the indemnity payment on this set has been increased. From the perspective of the insured, when  $M$  is too small, the indemnity function does not sufficiently reflect the impact of the weather index

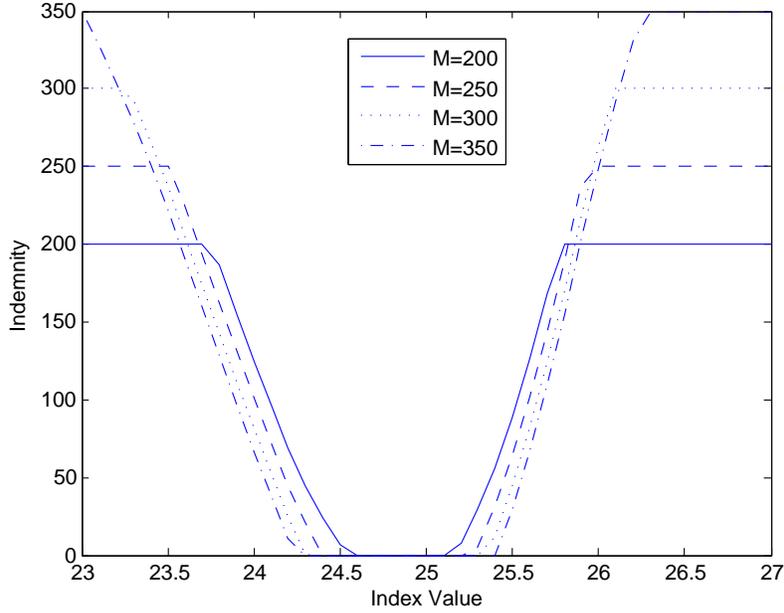


Figure 4.3: Optimal indemnity for different maximum indemnity levels.

on the insured's actual loss, and thus the index insurance contract is ineffective in this case; on the other hand, from the perspective of the insurance company, an increase in  $M$  also increases its own tail risk and thus potentially high capital cost. Therefore in practice the choice of  $M$  should be determined by the bargaining power between these two parties.

#### 4.5.2.3 Risk mitigation performance

In this section we are interested in the effectiveness of our proposed index insurance in reducing basis risk, which is measured by the standard deviation of the residual risk after the indemnity payment, i.e., the variance of the residual risk  $[Y - I^*(X)]$ . If the standard deviation of the residual risk is large, it means that policyholders' risk is not effectively mitigated and the basis risk is high, and vice versa. Figure 4.4 reports the standard deviation of the residual risk for  $P$  and  $M$  vary over intervals  $[0, 100]$  and  $[200, 600]$ , respectively.

The straight line intersected by the surface and plane  $P = 0$  in Figure 4.4 represents the

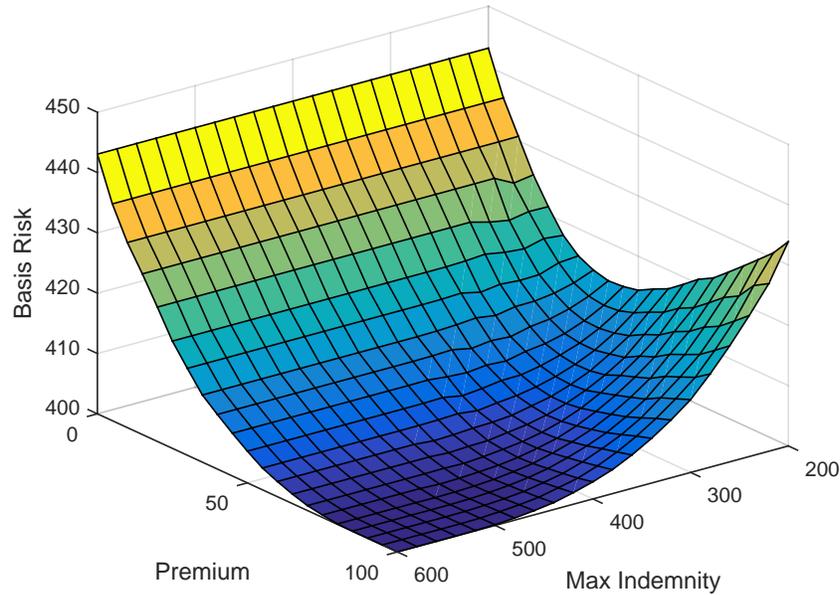


Figure 4.4: Basis risk (i.e., the standard deviation of residual risk) for different levels of premium and maximum indemnity.

uninsured position of the policyholder, and the standard deviation of the actual loss variable is about 443 in this case. The whole surface in the figure is lower than 443, which means a positive impact from our proposed index insurance on reducing the basis risk. The shape of the surface also indicates that the basis risk can be reduced by increasing premium level  $P$  and maximum indemnity level  $M$ . The exception occurs when  $P$  becomes too large relatively to  $M$ . This phenomenon can be explained by the curve on the top (corresponding to  $P = 80$ ) in Figure 4.2, which says that, when  $P$  is large relatively to  $M$ , any additional premium goes to cover the actual losses in the middle area, in other words, the coverage for small losses increases while coverage for large ones remains unchanged; as a result, the basis risk increases rather than decreases.

Since the premium level also indicates how much subsidy the government is paying for the policyholders, Figure 4.4 also provides some suggestions for the government in determining the subsidy amount according to  $M$ . This example shows that a wise choice for the subsidy amount needs to comply with the maximum indemnity level  $M$ .

### 4.5.3 Exponential utility

Risk preference of the insured is essential in index insurance contract design, and one advantage of our method is its capability to take into account the insured's utility function. In the present and the next sections we will investigate the optimal index insurance designs under the exponential and the logarithmic utility functions.

First we consider the shape of the optimal indemnity function under an exponential utility in the form of  $U(x) = -\frac{1}{\alpha}e^{-\alpha x}$ , where the parameter  $\alpha > 0$ . By equation (4.18), the optimal indemnity function depends on the parameter  $\alpha$ , which measures the degree of risk preference of the policyholder. Risk averse policyholders always have  $\alpha > 0$ , and a higher  $\alpha$  means a higher degree of risk aversion. In order to see how  $\alpha$  affects the shape of the optimal index insurance contract, we illustrate the optimal indemnity functions for four different values of  $\alpha$  in Figure 4.5. In this example we fix  $P = 50$  and  $M = 300$ .

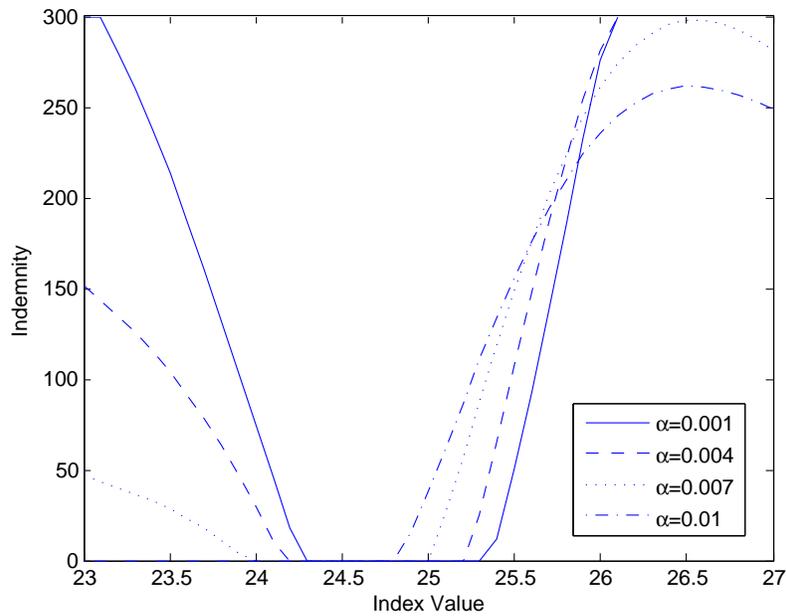


Figure 4.5: Optimal indemnity under exponential utility.

Figure 4.5 shows that the shape of the optimal indemnity function depends heavily on the choice of  $\alpha$ . For an  $\alpha$  as small as 0.001, the indemnity function also shows the “first decreasing and then increasing” pattern which has been previously observed under the quadratic utility. As  $\alpha$  becomes larger, i.e., the insured becomes more risk averse, the coverage on the low temperature region decreases and more premium is spent on the high temperature region. When the parameter  $\alpha$  is as large as 0.01, the optimal index insurance contract only indemnifies losses occurred in the high temperature region but not those in the low temperature region. This phenomenon can be explained from two perspectives. First, policyholders with a larger  $\alpha$  is generally more risk averse than those using “quadratic utility” as their risk preferences. As a result, they would like to have more coverage on the most severe losses, which occur in the high temperature region. Second, it can be explained by the asymmetric effects of temperature on the loss in rice yield. The adverse effect is more severe from the high temperatures than the low temperatures.

#### 4.5.4 Logarithmic utility

The logarithmic utility function takes a form of  $U(x) = \ln x$  for  $x > 0$ . There is no closed-form for the optimal index insurance contract under the logarithmic utility. We use Proposition 4.4 and apply a numerical scheme to solve the ODE in (4.15) for the optimal solution. The numerical scheme is specified in Appendix C. We fix the maximum indemnity level  $M = 300$ , and compute the optimal indemnity functions for a set of different premium levels  $P$ . The resulting optimal indemnity functions are illustrated in Figure 4.6. The figure shows that the optimal indemnity function takes a similar shape as the one under the quadratic utility function. It also generally follows the “first decreasing and then increasing” pattern and the coverage increases throughout the whole region as the premium increases.

In order to make a close comparison of the optimal index insurance among the three utility functions (i.e., quadratic, exponential and logarithmic), we fix  $P = 50$  and  $M = 300$  and demonstrate the resulting optimal indemnity functions in Figure 4.7, where  $\alpha = 0.005$  is set for the exponential utility function. Clearly, the contract under the logarithmic utility is quite similar to the one under the quadratic utility. The left parts of the two curves almost coincide with each other, and the right part of optimal indemnity function under the logarithmic utility is

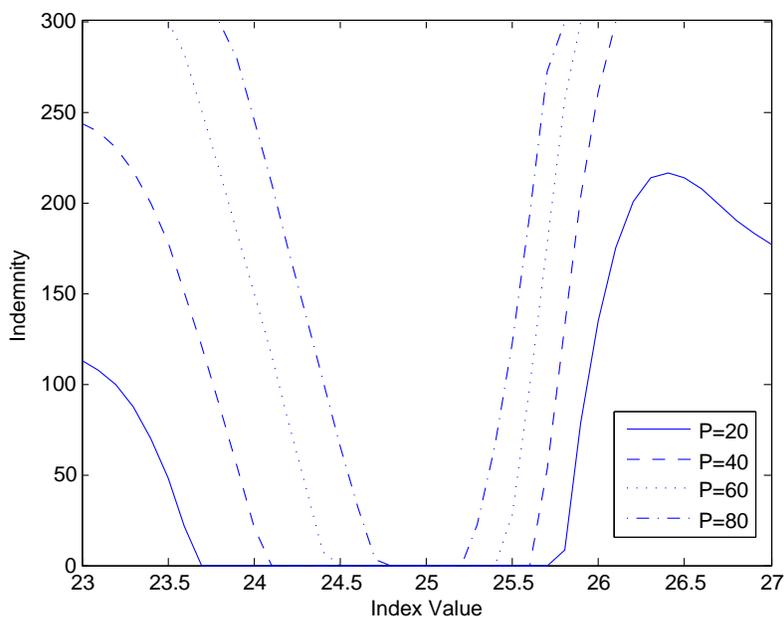


Figure 4.6: Optimal indemnity under logarithmic utility.

slightly steeper. In contrast, the optimal indemnity function under the exponential utility function is quite different from the other two, with less coverage on the left half but more coverage on the right half.

#### 4.5.5 Comparison with linear contracts

In this section we compare the effectiveness of our optimally designed index insurance contracts with the linear-type contract (e.g., Giné et al., 2007; Okhrin et al., 2013), which is based on a linear regression procedure and widely applied in both practice and academia as a benchmark. The effectiveness is measured by the standard deviation of the residual risk after the indemnity payment for the policyholders, which we also call basis risk. The comparison is conducted for a set of premium levels and two maximum indemnity payments at  $M = 250$  and  $M = 300$ , respectively. The results are demonstrated in Figure 4.8.

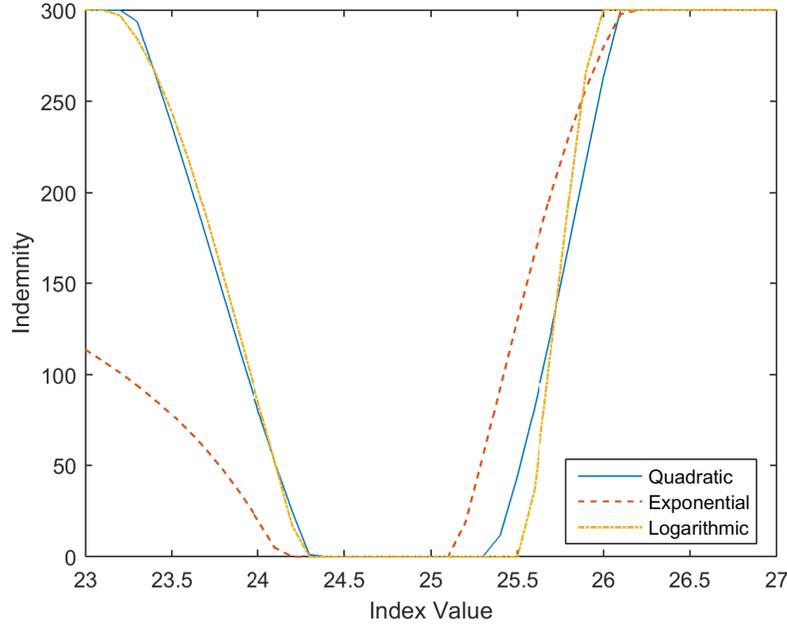


Figure 4.7: Comparison of optimal indemnity functions under three different utilities.

Maximizing the expected quadratic utility is the same as minimizing the variance of the residual risk for the insured in our index insurance design model. In theory our proposed indemnity function achieves the smallest standard deviation reduction of the residual risk for the insured, as guaranteed by Proposition 4.5. Figure 4.8 shows the superiority of our proposed insurance contract compared with the linear-type contract. The basis risk measured by the standard deviation of the insured's residual risk is smaller under our optimally design index insurance than the linear-type contract. Our index insurance performs equally well as the linear-type contract for small premium level (say,  $P < 40$ ). It substantially outperforms the linear-type contract for larger premium levels, and the advantage becomes more obvious as  $P$  increases. While the linear-type contract suffers from not being able to benefit from an increase in  $P$  and  $M$ , we can generally enhance the performance of our proposed contract by increasing the premium level  $P$  and enlarging the maximum coverage level  $M$ .

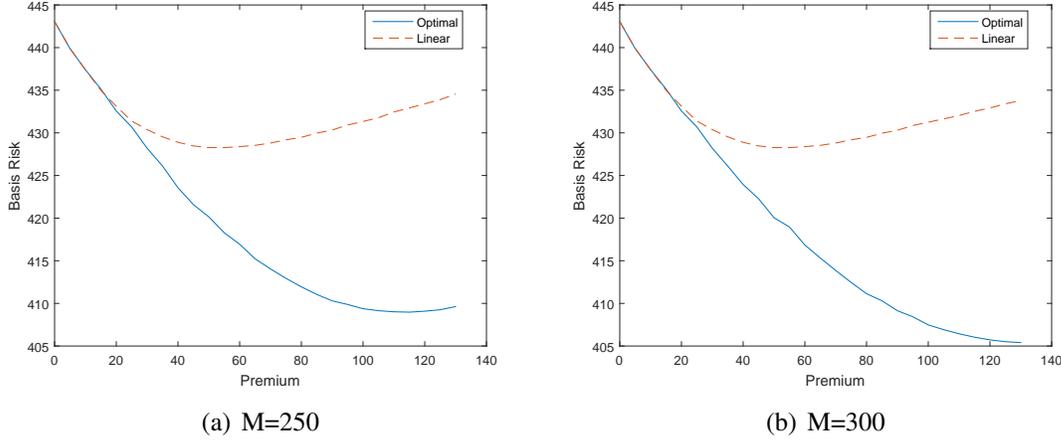


Figure 4.8: Effectiveness: our optimal index insurance vs. the linear-type contract.

#### 4.5.6 An example of bivariate-index insurance

In the previous sections, we focus on the single-index case, where the insurance indemnity is determined by a single index. In this section, we incorporate a second index into the insurance contract design and consider the optimal bivariate-index insurance. In principle introducing more indices into the insurance payoff function is always helpful to reduce basis risk, and we will illustrate this by an example of index insurance linked to two weather indices: temperature and precipitation. We use the average temperature and the total amount of precipitation during the whole growing period as the underlying indices to construct the index insurance contract. In this example we assume the policyholders' utility function has a quadratic form, and results under other utility functions can be analyzed similarly. Mathematically, it is straightforward to show that Proposition 4.5 still holds for the multi-dimensional case and the optimal indemnity function takes a similar form as  $I^*(x)$  in (4.17). Let  $X_1$  and  $X_2$  respectively denote the two indices under our consideration. Then the optimal indemnity function is given by

$$I^*(x_1, x_2) = [(E[Y|X_1 = x_1, X_2 = x_2] + \eta^*) \vee 0] \wedge M,$$

where where  $\eta^*$  is determined by  $E[I^*(X)] = P/\gamma$ . The optimal indemnity function of this bivariate-index insurance contract is illustrated in Figure 4.9.

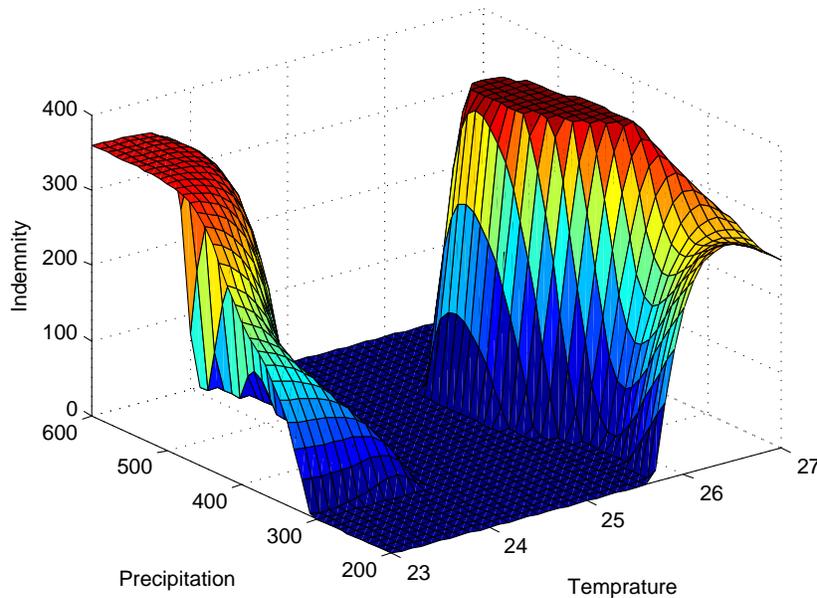


Figure 4.9: Indemnity function of a bivariate-index insurance: temperature and precipitation

The indemnity function also shows the “first decreasing and then increasing” pattern with respect to the increase in the temperature index. Further, the precipitation index also plays an important role in the optimal indemnity function. Indemnity amount is higher in the “low temperature & high precipitation” region than in the “low temperature & low precipitation” region. Within the “high temperature” region, a “medium precipitation” corresponds to the largest amount of indemnity, and a “low precipitation” leads to a larger indemnity amount than a “high precipitation”. Further, there is no indemnity payment when both temperature and precipitation are moderate, which corresponds to a good harvest and negligible actual loss.

To demonstrate the benefit of including the precipitation variable as the additional index in the optimal insurance contract, we compare the basis risk (measured by the standard deviation of residual risk) between the bivariate-index contract and the single-index contract under a set of

different premium levels. The maximum indemnity amount payment is fixed at  $M = 400$ . The comparison results are illustrated in Figure 4.10. Obviously, the inclusion of the precipitation index significantly reduces basis risk. Residual risk is constantly lower from the bivariate-index contract than the single-index contract. As the premium level  $P$  increases, the gap in residual risk between the two contracts becomes larger. In particular, the basis risk is reduced from 370 down to 340, which means a reduction rate of 8.1%, when the premium level  $P = 120$ . This suggests that more relevant indices should be included into the optimal insurance design if estimation of the joint distribution between the actual loss variable and indices is not an issue.

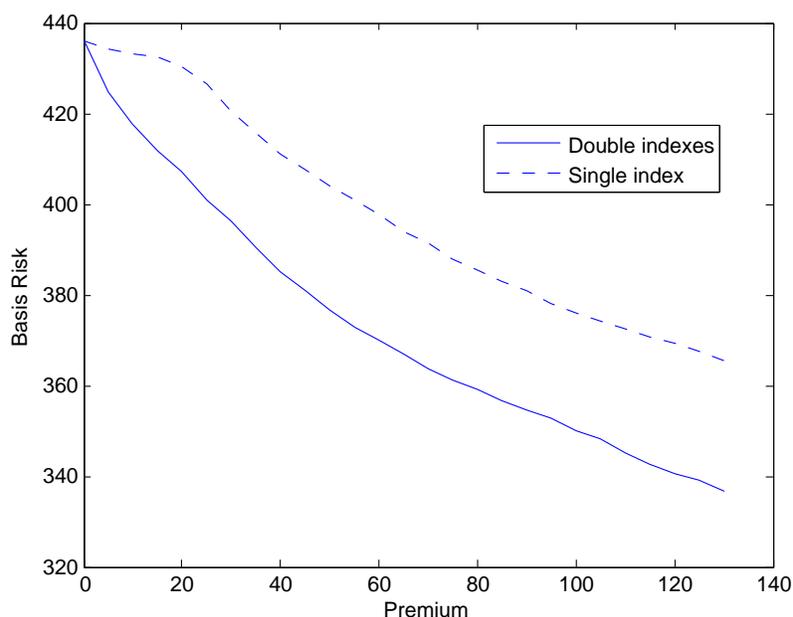


Figure 4.10: Effectiveness improvement of additional index. Basis risk is measured by standard deviation of residual risk.

## 4.6 Conclusion

In this chapter, we investigate the optimal index insurance design problem under a utility maximization framework. Under quite general and practical assumptions, we show that the optimal index insurance contract exists and is uniquely determined by the policyholder's utility function, the premium level and the maximum indemnity covered by the insurance contract. The optimal index insurance contract is obtained by solving an implicit ODE problem. Additionally, when the insured has a quadratic utility or an exponential utility, the optimal indemnity functions have explicit forms which are computationally friendly for real applications.

Our theoretical results are applied to a real data example, in which the temperature and precipitation variables are used as the underlying indices of the insurance contract to protect rice yield in Jiangsu, China. The shape of the optimal indemnity functions under different utility functions, premium levels and maximum indemnity amounts generally follow the “first decreasing and then increasing” pattern. The risk mitigation performance measured by the standard deviation reduction of the insured's residual risk is also discussed. Our results confirm that our optimally designed index insurance significantly outperforms the linear-type contract, which is a popular solution applied both in practice and in the literature for reducing farmers' basis risk. Finally, an example of a bivariate-index insurance contract based on the temperature and precipitation variables is introduced to show the benefit of incorporating multiple indices into the insurance contract design for mitigating basis risk.

# Chapter 5

## Conclusion and Future Research

This thesis studies the topic of *risk management with basis risk* by specializing in three problems from insurance and finance: financial derivative hedging, longevity risk hedging, and index insurance design. While these three problems are formulated from different economic settings, mathematically they all aim at solving a certain optimization problem for the best risk mitigation strategy from certain perspective. They all target to derive the best functional on the hedging instrument which minimizes the “distance” between the hedging objective and the functional representing the risk mitigation plan.

These three problems are investigated in three independent chapters, and they differ from each other in the following aspects:

1. By their nature, the financial derivative hedging and longevity risk hedging problems are studied in dynamic settings, whereas the index insurance design is analyzed in a static setting.
2. The financial derivative hedging problem is studied in a continuous-time setting, while the longevity risk hedging problem is tackled in a discrete-time setting for practicality reason.
3. The mean-variance criterion is adopted in our studies for the financial derivative hedging and longevity risk hedging problems. The index insurance design problem is studied under a utility maximization framework instead.

In the long-run, I plan to explore further interesting risk management problems with basis risk from the fields of insurance and finance. I am interested in developing methodologically innovative and practically effective risk management strategies for dealing with problems with basis risk. In the literature, basis risk is generally described as the non-hedgeable portion of risk as attributed to the mismatch between the hedging objective and the hedging instrument, and so far there has been no mathematically concrete definition for basis risk. As an important future research project, I plan to explore the potential intrinsic connections between the meanings of basis risk from different contexts, and to develop certain axiomatic scheme for basis risk.

In the near future, I plan to further my research on basis risk for the three problems: financial derivative hedging, longevity risk management and index insurance design, which have been studied in Chapters 2-4 of the thesis. Below I summarize the remaining important and relevant questions for each chapter.

Chapter 2 studies the problem of hedging general European vanilla options in a subgame Nash equilibrium framework. Two questions are natural to be further considered. First, what is the optimal hedging strategy for exotic or path-dependent options? This is indeed a very important but technically challenging problem. Most literature on dynamic hedging of path-dependent payoffs is based on the assumption of a complete market, and the problem of hedging general path-dependent payoffs with basis risk remains open in the literature. Second, what is the “pre-commitment” strategy, and how do the subgame Nash equilibrium strategy and the “pre-commitment” strategy differ from each other? Defined as the optimizer of the hedger’s objective function at time 0, the “pre-commitment” strategy has been widely studied by many researchers, however, the comparison between these two strategies currently focuses on the mathematical side, and further empirical evidence on pros and cons of each strategy is essential for people to better understand the nature of time inconsistency.

Chapter 3 provides a general framework to investigate the problem of optimal dynamic longevity hedge, and the framework can be used to conduct more comprehensive analysis on hedging problems based on different longevity models (e.g., extensions of ACF and CBD models) and hedging instruments (e.g., longevity swaps and longevity bonds). As one of my future research projects, I am interested in solving a multi-dimensional hedging problem and comparing

hedging effectiveness between different model assumptions and hedging portfolio constructions. Another problem worthwhile a more closely investigation is related to the calibration procedure of multi-population longevity models: what technique can one use to make the estimated correlation between populations to be more consistent across different underlying mortality models?

Chapter 4 discusses the problem of optimal index insurance design under an expected utility maximization framework. In the context of agricultural weather index insurance, the contract is designed to minimize the mismatch between the weather variable and the crop yield variable, which is known as the variable basis risk. A remaining yet critical problem is to take into consideration the spatial basis risk, which represents differences in weather and crop yield across locations, and offer a more comprehensive and handy solution for agricultural insurers.

# References

- [1] Barry J Barnett and Olivier Mahul. Weather index insurance for agriculture and rural areas in lower-income countries. *American Journal of Agricultural Economics*, 89(5):1241–1247, 2007.
- [2] Suleyman Basak and Georgy Chabakauri. Dynamic mean-variance asset allocation. *The Review of Financial Studies*, 23(8):2970–3016, 2010.
- [3] Suleyman Basak and Georgy Chabakauri. Dynamic hedging in incomplete markets: a simple solution. *The review of financial studies*, 25(6):1845–1896, 2012.
- [4] Tomas Björk and Agatha Murgoci. A general theory of markovian time inconsistent stochastic control problems, 2010. Working paper. Available at: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1694759](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1694759); accessed May-1-2018.
- [5] Tomas Björk and Agatha Murgoci. A theory of markovian time-inconsistent stochastic control in discrete time. *Finance and Stochastics*, 18(3):545–592, 2014.
- [6] Raushan Bokusheva, F Kogan, I Vitkovskaya, S Conradt, and M Batyrbayeva. Satellite-based vegetation health indices as a criteria for insuring against drought-related yield losses. *Agricultural and forest meteorology*, 220:200–206, 2016.
- [7] Patrick L Brockett, Mulong Wang, and Chuanhou Yang. Weather derivatives and weather risk management. *Risk Management and Insurance Review*, 8(1):127–140, 2005.

- [8] Richard L. Burden and J. Douglas Faires. Numerical analysis 9th edn (boston: Brooks/cole). 2011.
- [9] Andrew JG Cairns. Modelling and management of longevity risk: Approximations to survivor functions and dynamic hedging. *Insurance: Mathematics and Economics*, 49(3):438–453, 2011.
- [10] Andrew JG Cairns. Robust hedging of longevity risk. *Journal of Risk and Insurance*, 80(3):621–648, 2013.
- [11] Andrew JG Cairns, Kevin Dowd, David Blake, and Guy D Coughlan. Longevity hedge effectiveness: A decomposition. *Quantitative Finance*, 14(2):217–235, 2014.
- [12] Michael Carter, Alain de Janvry, Elisabeth Sadoulet, Alexander Sarris, et al. Index-based weather insurance for developing countries: A review of evidence and a set of propositions for up-scaling. *Development Policies Working Paper*, 111, 2014.
- [13] Michael R Carter, Lan Cheng, and Alexandros Sarris. Where and how index insurance can boost the adoption of improved agricultural technologies. *Journal of Development Economics*, 118:59–71, 2016.
- [14] Sommarat Chantarat, Christopher B Barrett, Andrew G Mude, and Calum G Turvey. Using weather index insurance to improve drought response for famine prevention. *American Journal of Agricultural Economics*, 89(5):1262–1268, 2007.
- [15] Sommarat Chantarat, Andrew G Mude, Christopher B Barrett, and Michael R Carter. Designing index-based livestock insurance for managing asset risk in northern kenya. *Journal of Risk and Insurance*, 80(1):205–237, 2013.
- [16] Ya-Zhe Chen. Parabolic partial differential equations of second order. *Series in Graduate Textbooks in Mathematics*, Beijing University Press, Beijing, 2003.
- [17] Yichun Chi and Chengguo Weng. Optimal reinsurance subject to vajda condition. *Insurance: Mathematics and Economics*, 53(1):179–189, 2013.

- [18] Sarah Conradt, Robert Finger, and Martina Spörri. Flexible weather index-based insurance design. *Climate Risk Management*, 10:106–117, 2015.
- [19] Guy D Coughlan, Marwa Khalaf-Allah, Yijing Ye, Sumit Kumar, Andrew JG Cairns, David Blake, and Kevin Dowd. Longevity hedging 101: A framework for longevity basis risk analysis and hedge effectiveness. *North American Actuarial Journal*, 15(2):150–176, 2011.
- [20] J David Cummins, David Lalonde, and Richard D Phillips. The basis risk of catastrophic-loss index securities. *Journal of Financial Economics*, 71(1):77–111, 2004.
- [21] Rose-Anne Dana and Marco Scarsini. Optimal risk sharing with background risk. *Journal of Economic Theory*, 133(1):152–176, 2007.
- [22] Mark HA Davis. Optimal hedging with basis risk. In *From stochastic calculus to mathematical finance*, pages 169–187. Springer, 2006.
- [23] Clemente De Rosa, Elisa Luciano, and Luca Regis. Basis risk in static versus dynamic longevity-risk hedging. *Scandinavian Actuarial Journal*, 2017(4):343–365, 2017.
- [24] Neil A Doherty. Innovations in managing catastrophe risk. *The Journal of risk and insurance*, 64(4):713–718, 1997.
- [25] Neil A Doherty and Harris Schlesinger. Optimal insurance in incomplete markets. *journal of political economy*, 91(6):1045–1054, 1983.
- [26] Darrell Duffie and Henry R Richardson. Mean-variance hedging in continuous time. *The Annals of Applied Probability*, pages 1–15, 1991.
- [27] Ghada Elabed, Marc F Bellemare, Michael R Carter, and Catherine Guirkinger. Managing basis risk with multiscale index insurance. *Agricultural Economics*, 44(4-5):419–431, 2013.
- [28] H Föllmer and Martin Schweizer. Hedging of contingent claims under incomplete information. *Applied stochastic analysis*, 5:389–414, 1991.

- [29] Xavier Giné, Robert Townsend, and James Vickery. Statistical analysis of rainfall insurance payouts in southern india. *American Journal of Agricultural Economics*, 89(5):1248–1254, 2007.
- [30] Peter Hazell, Jamie Anderson, Niels Balzer, A Hastrup Clemmensen, Ulrich Hess, and Francesco Rispoli. The potential for scale and sustainability in weather index insurance for agriculture and rural livelihoods. Technical report, World Food Programme (WFP), 2010.
- [31] Nathaniel D Jensen, Christopher B Barrett, and Andrew G Mude. Index insurance quality and basis risk: evidence from northern kenya. *American Journal of Agricultural Economics*, 98(5):1450–1469, 2016.
- [32] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.
- [33] Jin-Ping Lee and Min-Teh Yu. Pricing default-risky cat bonds with moral hazard and basis risk. *Journal of Risk and Insurance*, pages 25–44, 2002.
- [34] Ronald D Lee and Lawrence R Carter. Modeling and forecasting us mortality. *Journal of the American statistical association*, 87(419):659–671, 1992.
- [35] Danping Li, Ximin Rong, and Hui Zhao. Time-consistent reinsurance–investment strategy for a mean–variance insurer under stochastic interest rate model and inflation risk. *Insurance: Mathematics and Economics*, 64:28–44, 2015.
- [36] Duan Li and Wan-Lung Ng. Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. *Mathematical Finance*, 10(3):387–406, 2000.
- [37] Johnny Siu-Hang Li and Mary R Hardy. Measuring basis risk in longevity hedges. *North American Actuarial Journal*, 15(2):177–200, 2011.
- [38] Johnny Siu-Hang Li and Ancheng Luo. Key q-duration: A framework for hedging longevity risk. *ASTIN Bulletin: The Journal of the IAA*, 42(2):413–452, 2012.

- [39] Nan Li and Ronald Lee. Coherent mortality forecasts for a group of populations: An extension of the lee-carter method. *Demography*, 42(3):575–594, 2005.
- [40] Yongwu Li, Han Qiao, Shouyang Wang, and Ling Zhang. Time-consistent investment strategy under partial information. *Insurance: Mathematics and Economics*, 65:187–197, 2015.
- [41] Zhongfei Li, Yan Zeng, and Yongzeng Lai. Optimal time-consistent investment and reinsurance strategies for insurers under hestons sv model. *Insurance: Mathematics and Economics*, 51(1):191–203, 2012.
- [42] Zongxia Liang and Min Song. Time-consistent reinsurance and investment strategies for mean–variance insurer under partial information. *Insurance: Mathematics and Economics*, 65:66–76, 2015.
- [43] Gary M Lieberman. *Second order parabolic differential equations*. World scientific, 1996.
- [44] Yijia Lin and Samuel H Cox. Securitization of catastrophe mortality risks. *Insurance: Mathematics and Economics*, 42(2):628–637, 2008.
- [45] Yijia Lin, Ken Seng Tan, Ruilin Tian, and Jifeng Yu. Downside risk management of a defined benefit plan considering longevity basis risk. *North American Actuarial Journal*, 18(1):68–86, 2014.
- [46] Leonard C MacLean, Yonggan Zhao, and William T Ziemba. Mean-variance versus expected utility in dynamic investment analysis. *Computational Management Science*, 8(1-2):3–22, 2011.
- [47] Olivier Mahul. Optimum crop insurance under joint yield and price risk. *Journal of Risk and Insurance*, 67(1):109–122, 2000.
- [48] Harry Markowitz. Portfolio selection. *The journal of finance*, 7(1):77–91, 1952.
- [49] Mario J Miranda and Katie Farrin. Index insurance for developing countries. *Applied Economic Perspectives and Policy*, 34(3):391–427, 2012.

- [50] Thomas Møller. Risk-minimizing hedging strategies for unit-linked life insurance contracts. *ASTIN Bulletin: The Journal of the IAA*, 28(1):17–47, 1998.
- [51] Thomas Møller. Risk-minimizing hedging strategies for insurance payment processes. *Finance and Stochastics*, 5(4):419–446, 2001.
- [52] International Monetary Fund Monetary and Capital Markets Department. Global financial stability report, April 2012: The quest for lasting stability, 2012. Available at: <https://www.imf.org/en/Publications/GFSR/Issues/2016/12/31/Global-Financial-Stability-Report-April-2012-The-Quest-for-Lasting-Stability-25343>; accessed May-1-2018.
- [53] Michael Monoyios. Performance of utility-based strategies for hedging basis risk. *Quantitative Finance*, 4(3):245–255, 2004.
- [54] Marek Musiela and Marek Rutkowski. Martingale methods in financial modelling, volume 36 of applications of mathematics: Stochastic modelling and applied probability, 1997.
- [55] Marek Musiela and Thaleia Zariphopoulou. An example of indifference prices under exponential preferences. *Finance and Stochastics*, 8(2):229–239, 2004.
- [56] Andrew Cheuk-Yin Ng and Johnny Siu-Hang Li. Pricing and hedging variable annuity guarantees with multiasset stochastic investment models. *North American Actuarial Journal*, 17(1):41–62, 2013.
- [57] Ostap Okhrin, Martin Odening, and Wei Xu. Systemic weather risk and crop insurance: the case of china. *Journal of Risk and Insurance*, 80(2):351–372, 2013.
- [58] Huyên Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media, 2009.
- [59] Artur Raviv. The design of an optimal insurance policy. *The American Economic Review*, 69(1):84–96, 1979.
- [60] Martin Schweizer. A guided tour through quadratic hedging approaches. Technical report, Discussion Papers, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes, 1999.

- [61] Hong Shi and Zhihui Jiang. The efficiency of composite weather index insurance in hedging rice yield risk: evidence from china. *Agricultural Economics*, 47(3):319–328, 2016.
- [62] Jerry R Skees, J Roy Black, and Barry J Barnett. Designing and rating an area yield crop insurance contract. *American journal of agricultural economics*, 79(2):430–438, 1997.
- [63] Robert Henry Strotz. Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3):165–180, 1955.
- [64] Calum G Turvey and Zahirul Islam. Equity and efficiency considerations in area versus individual yield insurance. *Agricultural economics*, 12(1):23–35, 1995.
- [65] James Vercaemmen. Optimal insurance with nonseparable background risk. *Journal of Risk and Insurance*, 68(3):437–447, 2001.
- [66] J Wei, KC Wong, SCP Yam, and SP Yung. Markowitzs mean–variance asset–liability management with regime switching: A time-consistent approach. *Insurance: Mathematics and Economics*, 53(1):281–291, 2013.
- [67] Tat Wing Wong, Mei Choi Chiu, and Hoi Ying Wong. Time-consistent mean–variance hedging of longevity risk: Effect of cointegration. *Insurance: Mathematics and Economics*, 56:56–67, 2014.
- [68] Tat Wing Wong, Mei Choi Chiu, and Hoi Ying Wong. Managing mortality risk with longevity bonds when mortality rates are cointegrated. *Journal of Risk and Insurance*, 84(3):987–1023, 2017.
- [69] Joshua D Woodard and Philip Garcia. Basis risk and weather hedging effectiveness. *Agricultural Finance Review*, 68(1):99–117, 2008.
- [70] Huiling Wu and Yan Zeng. Equilibrium investment strategy for defined-contribution pension schemes with generalized mean–variance criterion and mortality risk. *Insurance: Mathematics and Economics*, 64:396–408, 2015.

- [71] Jiongmin Yong and Xun Yu Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Science & Business Media, 1999.
- [72] Yan Zeng, Zhongfei Li, and Yongzeng Lai. Time-consistent investment and reinsurance strategies for mean–variance insurers with jumps. *Insurance: Mathematics and Economics*, 52(3):498–507, 2013.
- [73] Hui Zhao, Yang Shen, and Yan Zeng. Time-consistent investment-reinsurance strategy for mean-variance insurers with a defaultable security. *Journal of Mathematical Analysis and Applications*, 437(2):1036–1057, 2016.
- [74] Kenneth Q Zhou and Johnny Siu-Hang Li. Dynamic longevity hedging in the presence of population basis risk: A feasibility analysis from technical and economic perspectives. *Journal of Risk and Insurance*, 84(S1):417–437, 2017.
- [75] Xun Yu Zhou and Duan Li. Continuous-time mean-variance portfolio selection: A stochastic lq framework. *Applied Mathematics and Optimization*, 42(1):19–33, 2000.
- [76] Zhongbao Zhou, Helu Xiao, Jialing Yin, Ximei Zeng, and Ling Lin. Pre-commitment vs. time-consistent strategies for the generalized multi-period portfolio optimization with stochastic cash flows. *Insurance: Mathematics and Economics*, 68:187–202, 2016.
- [77] Sheng Chao Zhuang, Chengguo Weng, Ken Seng Tan, and Hirbod Assa. Marginal indemnification function formulation for optimal reinsurance. *Insurance: Mathematics and Economics*, 67:65–76, 2016.

# **APPENDICES**

# Appendix A

## Appendix for Chapter 2

### A.1 Proof of Proposition 2.1

The proof depends on the following lemma.

**Lemma A.1.** *Under the Lipschitz continuity condition given in equation (2.2), there exists a unique strong solution to the SDE (2.1). Moreover, for all  $p > 1$  and  $T > 0$ , there exist positive constants  $C_{1,T}$  and  $C_{2,T}$  such that:*

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |S_i(t)|^p \right] \leq C_{i,T}(1 + s_i^p), \quad i = 1, 2,$$

given that  $S_i(0) = s_i$  for constant  $s_i > 0$ ,  $i = 1, 2$ . Moreover, for the strong solution  $\hat{S}_i(t)$ ,  $i = 1, 2$  to the SDE (2.1) with initial values  $\hat{S}_i(0) = \hat{s}_i$ , there exist positive constants  $D_{1,T}$  and  $D_{2,T}$  such that

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |S_i(t) - \hat{S}_i(t)|^p \right] \leq D_{i,T}(s_i - \hat{s}_i)^p, \quad i = 1, 2.$$

*Proof.* See Theorem 6.3 in Yong and Zhou (1999). □

**Proof of Proposition 2.1.** We first prove the implication of condition **C2'** to condition **C2**. We

begin with showing that the first two inequalities in (2.35) together imply the second inequality in (2.32).

Recall that we have defined  $A(u, s) = \frac{\mu_1(u, s) - r}{\sigma_1(u, s)}$  and  $B(u, s) = \sigma_2(u, s) s \Pi_{s_2}(u, s)$ ,  $\forall (u, s) \in \mathbb{R}_+ \times \mathbb{R}$ , in equation (2.31). Let  $S_i^{t,s}(u)$  be the unique strong solution to the SDE (2.1) with initial value  $S_i(t) = s$ ,  $i = 1, 2$ . We claim that there exist positive constants  $K_0$  and  $K_{01}$  to satisfy

$$\limsup_{h \rightarrow 0} \mathbb{E}^* \left\{ \left[ \frac{A(u, S_1^{t, s_1+h}(u)) - A(u, S_1^{t, s_1}(u))}{S_1^{t, s_1+h}(u) - S_1^{t, s_1}(u)} \right]^4 \right\} \leq K_0(1 + |s_1|^{K_{01}}), \quad \forall s_1 > 0. \quad (\text{A.1})$$

Indeed, by the Mean Value Theorem, there exists a  $\mathcal{F}_u$  measurable random variable  $\xi \in [S_1^{t, s_1+h}(u) \wedge S_1^{t, s_1}(u), S_1^{t, s_1+h}(u) \vee S_1^{t, s_1}(u)]$  such that

$$\frac{A(u, S_1^{t, s_1+h}(u)) - A(u, S_1^{t, s_1}(u))}{S_1^{t, s_1+h}(u) - S_1^{t, s_1}(u)} = A_s(u, \xi),$$

where  $A_s(u, \xi) := \frac{\partial}{\partial s_1} A(u, s_1) \Big|_{s_1=\xi}$ ,  $a \wedge b = \min(a, b)$ , and  $a \vee b = \max(a, b)$ . Therefore, by the first equality in display (2.35), there exist positive constants  $c_1$ ,  $c_2$  and  $c \geq 1$  such that

$$\mathbb{E}^* \left[ \left( \frac{A(u, S_1^{t, s_1+h}(u)) - A(u, S_1^{t, s_1}(u))}{S_1^{t, s_1+h}(u) - S_1^{t, s_1}(u)} \right)^4 \right] = \mathbb{E}^* [A_s^4(u, \xi)] \leq \mathbb{E}^* [c_1 |\xi|^c + c_2]. \quad (\text{A.2})$$

Noticing that  $\xi \in [S_1^{t, s_1+h}(u) \wedge S_1^{t, s_1}(u), S_1^{t, s_1+h}(u) \vee S_1^{t, s_1}(u)]$ , we obtain

$$\begin{aligned} \mathbb{E}^* [c_1 |\xi|^c + c_2] &\leq \mathbb{E}^* \left[ c_1 \left( \left| S_1^{t, s_1+h}(u) \wedge S_1^{t, s_1}(u) \right| \vee \left| S_1^{t, s_1+h}(u) \vee S_1^{t, s_1}(u) \right| \right)^c + c_2 \right] \\ &= \mathbb{E}^* \left[ c_1 \left( \left| S_1^{t, s_1+h}(u) \wedge S_1^{t, s_1}(u) \right|^c \vee \left| S_1^{t, s_1+h}(u) \vee S_1^{t, s_1}(u) \right|^c \right) + c_2 \right] \\ &\leq \mathbb{E}^* \left[ c_1 \left| S_1^{t, s_1+h}(u) \wedge S_1^{t, s_1}(u) \right|^c + c_1 \left| S_1^{t, s_1+h}(u) \vee S_1^{t, s_1}(u) \right|^c + c_2 \right]. \end{aligned}$$

Using the notation  $\|X\|_p = (\mathbb{E}^* [|X|^p])^{1/p}$  for all  $p > 1$  and applying Minkowski inequality, we

can continue the above display as follows

$$\begin{aligned}
\mathbb{E}^* [c_1|\xi|^c + c_2] &\leq c_2 + c_1 \left( \left\| S_1^{t,s_1+h}(u) \wedge S_1^{t,s_1}(u) \right\|_c \right)^c + c_1 \left( \left\| S_1^{t,s_1+h}(u) \vee S_1^{t,s_1}(u) \right\|_c \right)^c \\
&\leq c_2 + c_1 \left( \left\| S_1^{t,s_1+h}(u) \wedge S_1^{t,s_1}(u) - S_1^{t,s_1}(u) \right\|_c + \left\| S_1^{t,s_1}(u) \right\|_c \right)^c \\
&\quad + c_1 \left( \left\| S_1^{t,s_1+h}(u) \vee S_1^{t,s_1}(u) - S_1^{t,s_1}(u) \right\|_c + \left\| S_1^{t,s_1}(u) \right\|_c \right)^c \\
&\leq c_2 + c_1 \left( \left\| S_1^{t,s_1+h}(u) - S_1^{t,s_1}(u) \right\|_c + \left\| S_1^{t,s_1}(u) \right\|_c \right)^c \\
&\quad + c_1 \left( \left\| S_1^{t,s_1+h}(u) - S_1^{t,s_1}(u) \right\|_c + \left\| S_1^{t,s_1}(u) \right\|_c \right)^c.
\end{aligned}$$

Applying Lemma A.1 to the right hand side of the above inequality, we obtain

$$\mathbb{E}^* [c_1|\xi|^c + c_2] \leq c_2 + c_1 [(K_l|h|^c)^{1/c} + K_l(1 + s_1^c)^{1/c}]^c \quad (\text{A.3})$$

for some positive constant  $K_l$ . Combining (A.2) and (A.3) and letting  $h \rightarrow 0$ , we obtain (A.1).

We are ready to prove the second inequality in (2.32). Indeed, by repeatedly using Hölder's inequality, we obtain

$$\begin{aligned}
&\frac{1}{h} \mathbb{E}^* \left[ \left( A(u, S_1^{t,s_1+h}(u)) - A(u, S_1^{t,s_1}(u)) \right) B(u, S_2^{t,s_2}(u)) \right] \\
&\leq \frac{1}{h} \left\| \left( A(u, S_1^{t,s_1+h}(u)) - A(u, S_1^{t,s_1}(u)) \right) \right\|_2 \cdot \|B(u, S_2^{t,s_2}(u))\|_2 \\
&\leq \frac{1}{h} \left\| \frac{A(u, S_1^{t,s_1+h}(u)) - A(u, S_1^{t,s_1}(u))}{S_1^{t,s_1+h}(u) - S_1^{t,s_1}(u)} \right\|_4 \cdot \|S_1^{t,s_1+h}(u) - S_1^{t,s_1}(u)\|_4 \cdot \|B(u, S_2^{t,s_2}(u))\|_2.
\end{aligned}$$

Consequently, letting  $h \rightarrow 0$  in the above display and applying Lemma A.1, we obtain

$$\begin{aligned}
&\frac{\partial}{\partial s_1} \mathbb{E}_{t,s_1,s_2}^* [A(u, S_1(u))B(u, S_2(u))] \\
&\leq (K_1(1 + |s_1|^{K_{11}}))^{1/4} \cdot K_2^{1/4} \cdot [K_3(1 + |s_2|^{K_{32}})]^{1/2} \\
&\leq K(1 + |s_1|^{K_1} + |s_2|^{K_2}).
\end{aligned}$$

for some positive constants  $K$ ,  $K_1$  and  $K_2$ .

In a similar manner, we can show that the third and fourth inequalities in (2.35) together imply the third inequality in (2.32). Moreover, it is easy to check that the first inequality from equation (2.35) ensures the first inequality from (2.32).

It remains to show  $\theta^* \in \Theta$ , i.e., there exists some constant  $K > 0$  such that  $E[\int_0^T \theta(u)^2 du] < \infty$  and  $E_{t,s_1,s_2} \left[ \int_t^T |\theta(u)| du \right] \leq K e^{K(s_1^2 + s_2^2)}$ ,  $\forall (t, s_1, s_2) \in [0, T] \times \mathbb{R}_+^2$ . From the definition of  $\eta(\cdot, \cdot, \cdot)$  given in (2.28) and the first two inequalities in (2.32), we can obtain that,  $\forall (t, s_1, s_2) \in [0, T] \times \mathbb{R}_+^2$ ,

$$\begin{aligned} & \frac{\partial}{\partial s_1} \eta(t, s_1, s_2) \\ = & \int_t^T \frac{1}{\gamma} \frac{\partial}{\partial s_1} E_{t,s_1,s_2}^* \left[ \left( \frac{\mu_1 - r}{\sigma_1} \right)^2 \right] du + \int_t^T \frac{\partial}{\partial s_2} E_{t,s_1,s_2}^* \left[ (\mu_1 - r) \frac{\rho \sigma_2}{\sigma_1} S_2(u) e^{r(T-u)} \Pi_{s_2}(u) \right] du \\ \leq & K(1 + |s_1|^{K_1} + |s_2|^{K_2}). \end{aligned}$$

Similarly, by using the last inequality in (2.32), we can obtain

$$\frac{\partial}{\partial s_2} \eta(t, s_1, s_2) \leq K(1 + |s_1|^{K_1} + |s_2|^{K_2}).$$

Therefore, by equation (2.27),

$$\theta^*(t, s_1, s_2) \leq K_0(1 + |s_1|^{K_{01}} + |s_2|^{K_{02}}),$$

which, in conjecture of Lemma A.1, implies  $E[\int_0^T \theta(u)^2 du] < \infty$  and  $E_{t,s_1,s_2} \left[ \int_t^T |\theta(u)| du \right] \leq K e^{K(s_1^2 + s_2^2)}$  for some constant  $K > 0$ .

## A.2 Equilibrium value function for futures contract

A semi closed-form expression for the equilibrium value function is obtained by plugging equation (2.64) into equation (2.47) as follows:

$$\begin{aligned}
J(t, s_1, s_2, x) &= (x - \Pi(t, s_2))e^{r(T-t)} + E_{t, s_1, s_2, x} \left( \int_t^T e^{r(T-u)} [\theta^*(u)(\mu_1 - r)] du \right) \\
&\quad - \frac{\gamma}{2} E_{t, s_1, s_2, x} \left[ \left( \int_t^T e^{r(T-u)} [\theta^*(u)(\mu_1 - r)] du \right. \right. \\
&\quad \left. \left. + \int_t^T e^{r(T-u)} \theta^*(u) \sigma_1 dW_1(u) - \int_t^T e^{\mu_2(T-u)} S_2(u) \sigma_2 dW_2(u) \right)^2 \right] \\
&\quad + \frac{\gamma}{2} \left[ E_{t, s_1, s_2, x} \left( \int_t^T e^{r(T-u)} [\theta^*(u)(\mu_1 - r)] du \right) \right]^2 \\
&= xe^{r(T-t)} + xK - s_2 e^{\mu_2(T-t)} + \frac{(\mu_1 - r)^2 (T - t)}{2\gamma\sigma_1^2} \\
&\quad + \frac{1}{2} s_2 [e^{\mu_2(T-t)} - e^{\mu_2^*(T-t)}] - \frac{\gamma\rho^2\sigma_2^2 s_2^2}{4\mu_2 + 2\sigma_2^2 - 4\mu_2^*} [e^{(2\mu_2 + \sigma_2^2)(T-t)} - e^{2\mu_2^*(T-t)}] \\
&\quad - \frac{\gamma s_2^2}{2} e^{2\mu_2 T - 2\mu_2 t - \sigma_2^2 t} [e^{\sigma_2^2 T} - e^{\sigma_2^2 t}] + \frac{\rho(\mu_1 - r)\sigma_2}{2\sigma_1^2} e^{\mu_2(T-t)} s_2 (T - t) \\
&\quad + \frac{\gamma\rho^2\sigma_2^2 s_2^2}{2\sigma_1(\mu_2 + \sigma_2^2 - \mu_2^*)} [e^{(2\mu_2 + \sigma_2^2)(T-t)} - e^{(\mu_2^* + \mu_2)(T-t)}] \\
&\quad + \frac{\gamma}{2} \left\{ \frac{(\mu_1 - r)^2 (T - t)}{\gamma\sigma_1^2} + s_2 [e^{\mu_2(T-t)} - e^{\mu_2^*(T-t)}] \right\}^2 \\
&\quad - \frac{\gamma}{2} E_{t, s_1, s_2, x} \left[ \left( \int_t^T e^{r(T-u)} [\theta^*(u)(\mu_1 - r)] du \right)^2 \right] \\
&\quad - \gamma E_{t, s_1, s_2, x} \left[ \left( \int_t^T e^{r(T-u)} [\theta^*(u)(\mu_1 - r)] du \right) \right. \\
&\quad \left. \left( \int_t^T e^{r(T-u)} \theta^*(u) \sigma_1 dW_1(u) - \int_t^T e^{\mu_2(T-u)} S_2(u) \sigma_2 dW_2(u) \right) \right].
\end{aligned}$$

### A.3 Equilibrium value function for European call options

A semi closed-form expression for the equilibrium value function is obtained by plugging equation (2.67) into equation (2.47) as follows:

$$\begin{aligned}
 J(t, s_1, s_2, x) &= (x - \Pi(t, s_2)) e^{r(T-t)} + \mathbb{E}_{t, s_1, s_2, x} \left( \int_t^T e^{r(T-u)} [\theta^*(u)(\mu_1 - r)] du \right) \\
 &\quad - \frac{\gamma}{2} \mathbb{E}_{t, s_1, s_2, x} \left[ \left( \int_t^T e^{r(T-u)} [\theta^*(u)(\mu_1 - r)] du \right. \right. \\
 &\quad \left. \left. + \int_t^T e^{r(T-u)} \theta^*(u) \sigma_1 dW_1(u) - \int_t^T e^{\mu_2(T-u)} \Phi(d_1) S_2(u) \sigma_2 dW_2(u) \right)^2 \right] \\
 &\quad + \frac{\gamma}{2} \left[ \mathbb{E}_{t, s_1, s_2, x} \left( \int_t^T e^{r(T-u)} [\theta^*(u)(\mu_1 - r)] du \right) \right]^2.
 \end{aligned}$$

# Appendix B

## Appendix for Chapter 3

### B.1 Approximations for forward survival rates in Section 3.4.2

By Zhou and Li (2016), forward survival rates are approximated by,  $\forall u > t$ ,

$$p_{x,u}^{(R)}(T, k_t^{(R)}) \approx \Phi \left( \frac{-\mathbb{E}(V_u^{(R)} | \mathcal{F}_t)}{\sqrt{\text{Var}(V_u^{(R)} | \mathcal{F}_t)}} \right) \quad (\text{B.1})$$

where  $\Phi(\cdot)$  denotes the CDF of a standard normal distribution, and other components of the formula are defined as below:

$$\begin{aligned} \mathbb{E}(V_u^{(R)} | \mathcal{F}_t) &= -D_{x,u,0}^{(R)}(T) - D_{x,u,1}^{(R)}(T)(\mathbb{E}(K_u | \mathcal{F}_t) - \mathbb{E}(K_u | \mathcal{F}_0)) \\ &\quad - D_{x,u,2}^{(R)}(T)(\mathbb{E}(k_u^{(R)} | \mathcal{F}_t) - \mathbb{E}(k_u^{(R)} | \mathcal{F}_0)) \\ \text{Var}(V_u^{(R)} | \mathcal{F}_t) &= 1 + (D_{x,u,1}^{(R)}(T))^2 \text{Var}(K_u | \mathcal{F}_t) + (D_{x,u,2}^{(R)}(T))^2 \text{Var}(k_u^{(R)} | \mathcal{F}_t) \\ &\quad + 2D_{x,u,1}^{(R)}(T)D_{x,u,2}^{(R)}(T)\text{Cov}(K_u, k_u^{(R)} | \mathcal{F}_t) \\ \mathbb{E}(K_u | \mathcal{F}_t) &= K_t - K_0 - Ct \\ \mathbb{E}(k_u^{(R)} | \mathcal{F}_t) &= (\phi_1^{(R)})^u ((\phi_1^{(R)})^{-t} k_t^{(R)} - k_0^{(R)}) + \frac{(\phi_1^{(R)})^u (1 - \phi_1^{(R)})^{-t}}{1 - \phi_1^{(R)}} \phi_0^{(R)} \end{aligned}$$

$$\begin{aligned}
\text{Var}(K_u|\mathcal{F}_t) &= \sigma_K^2(u-t) \\
\text{Var}(k_u^{(R)}|\mathcal{F}_t) &= \frac{1 - (\phi_1^{(i)})^{2(u-t)}}{1 - (\phi_1^{(R)})^2} \sigma_{k,i}^2 \\
\text{Cov}(K_u, k_u^{(R)}|\mathcal{F}_t) &= 0 \\
f_{x,t}^{(R)}(T, k_t^{(R)}) &= \Phi^{-1}(p_{x,t}^{(R)}(T, k_t^{(R)})) \\
D_{x,t,0}^{(R)}(T) &= f_{x,t}^{(R)}(T, \mathbb{E}(K_t|\mathcal{F}_0), \mathbb{E}(k_t^{(R)}|\mathcal{F}_0)) \\
D_{x,t,1}^{(R)}(T) &= \left. \frac{\partial f_{x,t}^{(R)}(T, K_t, \mathbb{E}(k_t^{(R)}|\mathcal{F}_0))}{\partial K_t} \right|_{K_t=\mathbb{E}(K_t|\mathcal{F}_0)} \\
D_{x,t,2}^{(R)}(T) &= \left. \frac{\partial f_{x,t}^{(R)}(T, \mathbb{E}(K_t|\mathcal{F}_0), k_t^{(R)})}{\partial k_t^{(R)}} \right|_{k_t^{(R)}=\mathbb{E}(k_t^{(R)}|\mathcal{F}_0)}
\end{aligned}$$

For proof, see Zhou and Li (2016).

## B.2 Approximations for forward survival rates under CBD model in Section 3.4.7

In order to show the following derivation more concisely, we look at one specific population  $i$  and denote time-varying factors  $\kappa_{1,t}^c, \kappa_{2,t}^c, \kappa_{1,t}^{(i)}$  and  $\kappa_{2,t}^{(i)}$  in the CBD model as  $\kappa_{1,t}, \kappa_{2,t}, \kappa_{3,t}$  and  $\kappa_{4,t}$ , respectively. We also denote  $\sigma_1 = \sigma(\eta_{1,t}^c), \sigma_2 = \sigma(\eta_{2,t}^c), \rho_1 = \rho(\eta_{1,t}^c, \eta_{2,t}^c), \sigma_3 = \sigma(\eta_{1,t}^{(i)})$ ,  $\sigma_4 = \sigma(\eta_{2,t}^{(i)})$ ,  $\rho_2 = \rho(\eta_{1,t}^{(i)}, \eta_{2,t}^{(i)})$  and denote the vector  $(\kappa_{1,t}, \kappa_{2,t}, \kappa_{3,t}, \kappa_{4,t})^\top$  as  $\kappa_t$ . Then we apply the probit transformation and define  $f_{x,t}^{(i)}(T, \kappa_t) := \Phi^{-1}(p_{x,t}^{(i)}(T, \kappa_t))$ . Therefore an approximation based on Taylor's theorem at  $\kappa_t = \hat{\kappa}_t := \mathbb{E}[\kappa_t]$  is given by:

$$\begin{aligned}
f_{x,t}^{(i)}(T, \kappa_t) &= \tilde{f}_{x,t}^{(i)}(T, \kappa_t) \\
&= D_{x,t,0}^{(i)}(T) + D_{x,t,1}^{(i)}(T)^\top (\kappa_t - \hat{\kappa}_t) + \frac{1}{2} (\kappa_t - \hat{\kappa}_t)^\top D_{x,t,2}^{(i)}(T) (\kappa_t - \hat{\kappa}_t),
\end{aligned}$$

where

$$\begin{aligned}
D_{x,t,0}^{(i)}(T) &= f_{x,t}^{(i)}(T, \hat{\kappa}_t), \\
D_{x,t,1}^{(i)}(T)_j &= \left. \frac{\partial}{\partial \kappa_{j,t}} f_{x,t}^{(i)}(T, \kappa_t) \right|_{\kappa_t = \hat{\kappa}_t}, \quad j = 1, 2, 3, 4, \\
D_{x,t,0}^{(i)}(T) &\text{ is a } 4 \times 1 \text{ vector with } j\text{-th element defined by } D_{x,t,1}^{(i)}(T)_j, \\
D_{x,t,2}^{(i)}(T)_{jk} &= \left. \frac{\partial^2}{\partial \kappa_{j,t} \partial \kappa_{k,t}} f_{x,t}^{(i)}(T, \kappa_t) \right|_{\kappa_t = \hat{\kappa}_t}, \quad j, k = 1, 2, 3, 4, \\
D_{x,t,2}^{(i)}(T) &\text{ is a } 4 \times 4 \text{ matrix with } jk\text{-th element defined by } D_{x,t,2}^{(i)}(T)_{jk}.
\end{aligned}$$

Therefore, by definition, when  $u > t$ ,

$$p_{x,u}^{(i)}(T, \kappa_t) = \mathbb{E}[p_{x,u}^{(i)}(T, \kappa_u) | \mathcal{F}_t].$$

Applying a first-order approximation and letting  $A_u := D_{x,u,0}^{(i)}(T) + D_{x,u,1}^{(i)}(T)^\top(\kappa_u - \hat{\kappa}_u)$ , we have:

$$\begin{aligned}
p_{x,u}^{(i)}(T, \kappa_u) &\approx \Phi[D_{x,u,0}^{(i)}(T) + D_{x,u,1}^{(i)}(T)^\top(\kappa_u - \hat{\kappa}_u)] \\
&= \mathbb{P}[Z \leq A_u | \mathcal{F}_u] \quad \text{where } Z \text{ is standard normal and independent of } A_u \\
&= \mathbb{E}[\mathbb{1}_{\{Z \leq A_u\}} | \mathcal{F}_u].
\end{aligned}$$

Hence, letting  $B_u := Z - A_u = Z - D_{x,u,0}^{(i)}(T) - D_{x,u,1}^{(i)}(T)^\top(\kappa_u - \hat{\kappa}_u)$ ,

$$\begin{aligned}
p_{x,u}^{(i)}(T, \kappa_t) &\approx \mathbb{E}\{\mathbb{E}[\mathbb{1}_{\{Z \leq A_u\}} | \mathcal{F}_u] | \mathcal{F}_t\} \\
&= \mathbb{E}[\mathbb{1}_{\{Z \leq A_u\}} | \mathcal{F}_t] \\
&= \mathbb{P}[Z \leq A_u | \mathcal{F}_t] \\
&= \mathbb{P}[B_u \leq 0 | \mathcal{F}_t] \\
&= \Phi \left[ \frac{-\mathbb{E}(B_u | \mathcal{F}_t)}{\sqrt{\text{Var}(B_u | \mathcal{F}_t)}} \right],
\end{aligned}$$

where

$$\begin{aligned} \mathbb{E}(B_u|\mathcal{F}_t) &= -D_{x,u,0}^{(i)}(T) - D_{x,u,1}^{(i)}(T)^\top [\mathbb{E}(\kappa_u)|\mathcal{F}_t - \hat{\kappa}_u], \\ \text{Var}(B_u|\mathcal{F}_t) &= 1 + [\mathbb{E}(\kappa_u)|\mathcal{F}_t - \hat{\kappa}_u]^\top V_u [\mathbb{E}(\kappa_u)|\mathcal{F}_t - \hat{\kappa}_u], \end{aligned}$$

where  $V_u$  is a  $4 \times 4$  covariance matrix with  $jk$ -th element defined by:

$$V_{u,jk} = \text{Cov}(\kappa_{j,u}, \kappa_{k,u}|\mathcal{F}_t),$$

and

$$\begin{aligned} \mathbb{E}(\kappa_{j,u}|\mathcal{F}_t) - \hat{\kappa}_{j,u} &= \kappa_{j,t} - \kappa_{j,0} - \mu_1^c t, \quad j = 1, 2, \\ \mathbb{E}(\kappa_{j,u}|\mathcal{F}_t) - \hat{\kappa}_{j,u} &= (\phi_{j-2}^{(i)})^u ((\phi_{j-2}^{(i)})^{-t} k_{j-2,t}^{(i)} - k_{j-2,0}^{(i)}) + \frac{(\phi_{j-2}^{(i)})^u (1 - \phi_{j-2}^{(i)})^{-t}}{1 - \phi_{j-2}^{(i)}} \mu_{j-2}^{(i)}, \quad j = 3, 4, \\ V_{u,jj} &= \text{Var}(\kappa_{j,u}|\mathcal{F}_t) = \sigma_j^2 (u - t), \quad j = 1, 2, \\ V_{u,12} = V_{u,21} &= \text{Cov}(\kappa_{1,u}, \kappa_{2,u}|\mathcal{F}_t) = \rho \sigma_1 \sigma_2 (u - t), \\ V_{u,jj} &= \text{Var}(\kappa_{j,u}|\mathcal{F}_t) = \frac{1 - (\phi_{j-2}^{(i)})^{2(u-t)}}{1 - (\phi_{j-2}^{(i)})^2} \sigma_j^2, \quad j = 3, 4, \\ V_{u,34} = V_{u,43} &= \text{Cov}(\kappa_{3,u}, \kappa_{4,u}|\mathcal{F}_t) = \frac{1 - (\phi_1^{(i)} \phi_2^{(i)})^{(u-t)}}{1 - \phi_1^{(i)} \phi_2^{(i)}} \rho \sigma_3 \sigma_4, \\ V_{u,13} = V_{u,14} &= V_{u,23} = V_{u,24} = V_{u,31} = V_{u,32} = V_{u,41} = V_{u,42} = 0. \end{aligned}$$

# Appendix C

## Appendix for Chapter 4

### C.1 A Numerical Scheme for Solving ODE (4.15)

For utility functions other than the quadratic and exponential ones, closed-form solutions for the ODE (4.15) are generally unavailable and thus a numerical scheme is needed. The general boundary ODE problem (4.15) can be viewed as an initial value problem (C.1) along with an algebraic equation (C.2):

$$\begin{cases} \frac{dI}{dx} = F(x, I), & x \in [c, d], \\ I(c) = I_c, \end{cases} \quad (\text{C.1})$$

with initial value  $I_c$  determined by

$$P = \gamma E [I^*(X)] = \gamma E \left[ \left( \widehat{I}_{\lambda^*}(X) \vee 0 \right) \wedge M \right]. \quad (\text{C.2})$$

For any fixed  $I_c$ , the initial value problem (C.1) is a standard ODE problem. If equation (C.1) yields a unique solution for a given  $I_c$ , then equation (C.2) becomes an algebraic equation of  $I_c$ . By Theorem 5.4 and Theorem 5.6 in Burden and Faires (2001), a sufficient condition for existence and uniqueness of the solution, and the well-posedness of problem (C.1) is given by:

$$\mathbf{H2:} \begin{cases} \left| \frac{\partial F(x, I)}{\partial I} \right| \leq L, \forall (x, I) \in [c, d] \times \mathbb{R} \text{ for some constant } L > 0, \\ F(x, I) \text{ is continuous on } [c, d] \times \mathbb{R}. \end{cases}$$

If condition **H2** holds, then the implicit ODE (C.1)-(C.2) are well-posed and it can be solved using a numerical procedure. We recommend the 4th order Runge-Kutta (RK4) method combined with a binary search to numerically compute  $I^*(x)$ ,  $x \in [c, d]$ . The specific numerical scheme is summarized in six steps below.

Step 1 : Find a large enough interval  $[L_c, U_c]$  such that  $I(c) \in [L_c, U_c]$ . Check that  $(P_{L_c} - P)(P_{U_c} - P) < 0$ , where  $P_{I_c}$  denotes the premium calculated by equation (C.2) for the contract starting at  $I_c$ . Suppose  $P_{L_c} - P < 0$ ,  $P_{U_c} - P > 0$ , and define  $I_0(c) = \frac{1}{2}(L_c + U_c)$ .

Step 2 : Apply RK4 with a step-size  $\delta > 0$  to the initial value problem  $\frac{dI}{dx} = F(x, I)$ ,  $x \in [c, d]$ , with  $I(c) = I_0(c)$  : For  $n = 0, 1, 2, \dots, \lfloor \frac{d-c}{\delta} - 1 \rfloor$ , define

1.  $k_1 = F(x_n, I_n)$ ,
2.  $k_2 = F(x_n + \frac{\delta}{2}, I_n + \frac{\delta}{2}k_1)$ ,
3.  $k_3 = F(x_n + \frac{\delta}{2}, I_n + \frac{\delta}{2}k_2)$ ,
4.  $k_4 = F(x_n + \delta, I_n + \delta k_3)$ ,
5.  $x_{n+1} = x + \delta$ ,
6.  $I_{n+1} = x + \frac{\delta}{6}(k_1 + 2k_2 + 2k_3 + k_4)$ .

Step 3 : Define  $I_n^* = (I_n \vee 0) \wedge M$ ,  $n = 0, 1, 2, \dots, \frac{d-c}{\delta}$ , where  $I_n$  is obtained from the previous step.

Step 4 : Approximate the premium constraint  $P = \gamma E[I^*]$  numerically using

$$P_0 := \frac{\delta\gamma}{2} \left[ 2 \left( \sum_{n=0}^{\frac{d-c}{\delta}} I_n^* h(c + n\delta) \right) - I_0^* h(c) - I_{\frac{d-c}{\delta}}^* h(d) \right].$$

Step 5 : Verify whether  $|P_0 - P| < \epsilon$  is satisfied by the given tolerance  $\epsilon$ . If yes,  $I^*$  is already an accurate approximation to the solution of ODE (C.1)-(C.2), and we stop the algorithm; otherwise, we go to Step 6.

Step 6 : If  $P_0 < P$ , then define  $I_1(c) = \frac{1}{2}(I_0(c) + U_c)$ ; if  $P_0 > P$ , then define  $I_1(c) = \frac{1}{2}(L_c + I_0(c))$ . Go back to Step 2, replace the initial condition with  $I(c) = I_1(c)$ , and repeat Steps 2-6.