Robust Asset Allocation in Incomplete Markets

Sally Muling Shen
Robust Asset Allocation in Incomplete Markets

DISSERTATION

to obtain the degree of Doctor at Maastricht University,
on the authority of Prof. dr. L.L.G. Soete, Rector Magnificus,
in accordance with the decision of the Board of Deans,
to be defended in public
on Monday, 22 June 2015 at 12:00

by

Sally Muling Shen 沈牧龄
Promotors:

Prof. dr. Antoon Pelsser
Prof. dr. Peter Schotman

Evaluation committee:

Prof. dr. Jaap Bos, Chairman
Prof. dr. Hans Schumacher, (Tilburg University)
Prof. dr. Jean-Pierre Urbain
Prof. dr. ir. Michel Vellekoop, (University of Amsterdam)
To my papa Kai Shen and mama Qingyang Sun
致爸爸沈凯、妈妈孙清扬
Acknowledgements

As a curious person, I enjoy questioning and being questioned. I guess that is one of the reasons why I decided to pursue a PhD. Doing a PhD in a cross field of Quantitative Finance and Actuarial Science is one of the most wonderful adventure in my life. The best part of being a doctoral student is that I can get paid to learn and to develop new things (or models in my case). Another cool thing about being a PhD student is that I can travel all over the world to present my ideas in international conferences with a student discount. Research is a long, brain-cell killing, risky and not always rewarding journey that requires lots of energy, creativity, persistency, patience, inspirations and hard work. However, that is still far from enough without the support and the encouragement from an amazing group of people whom I am truly grateful for.

This dissertation would not have been possible without the support and guidance of my supervisors, Peter Schotman and Antoon Pelsser.

Peter, I feel so lucky to have you as my supervisor. In any phase of my doctoral journey, whenever I am stuck or scrambled, the first thing that jumped into my mind was: “talk to Peter”. You taught me many tricks of various numerical methods, challenged me with many interesting philosophical questions, criticized my ideas, walked me through the job market and helped me with many important life decisions. You have already been more than a supervisor for me. Thank you for tolerating my sassy personality.

Antoon, you are so smart that any little hints and inspirations can already help me out. Thank you for always pushing me further to a closed-form solution, challenging the limit of my maths skills and always being critical of my models. Your very humorous and inspiring presentation style surely has influenced my way of talking in public. Your slogan: “More accurately calculating a wrong answer.” is still one of my favourite.

A next word of gratitude goes to a very special group of people who are the primers of my research career. Hans Schumacher, Edward Ponds, Jiajia Cui and all my former colleagues from APG (CAR) research group, thank you for bringing me to this fascinating pension modelling world. I feel so lucky and truly humbled by the opportunities you have given to me. The 10-month research internship in APG - Amsterdam was one of the most exciting, intense and definitely rewarding experiences of my life. Hans, you surely is one of those guardian angels of mine. You witnessed my Bachelor graduation which I skipped because of your very important course “Financial Models”; my Master graduation and
my (almost) Doctoral graduation. Being detail-oriented is always your spirit and is also the most important lesson I learnt from you. I will surely bring it with me in my future academic career. I would also like to thank Henk Norde from Tilburg University who helped me accomplishing my dream of becoming an Econometrics students during the time when the undergraduate program of Econometrics was still taught in Dutch.

I have continuing appreciation to the members of my assessment committee Jaap Bos, Jean-Pierre Urbain, Hans Schumacher and Michel Vellekoop. Thank you for your kind words and valuable suggestions on my dissertation manuscript. Your comments cover the fields of game theory, microeconomics, macroeconomics, mathematics and engineering. Those feedbacks push me to give a deeper thought about my models in practice.

Experiencing the academic job market is like doing a roller coaster or bungee jumping. Nervousness, excitement and anxiety go up and down following a random tempo. It is an exhausting, brutal and highly competitive war. However, I have never been alone. I am so lucky to have so many friends, colleagues and family supporting me to fight against this war. You make my job market journey joyful and blessed.

Judy, my BFF. I can list hundreds of things that we are in common: Zara, UGGs, Steve Madden, sketching, painting, stuffed animal army, etc. I am very very lucky to have you (so is Rogier Q) and I firmly believe that you are the sexiest girl ever. If you are in Hollywood, Kim Kardashian would have never existed. Thank you so much for helping me with my job market preparations. Without you, I can never make it. I always enjoyed our coffee moment and I will surely miss that when I move to Beijing. Jaap, you are the coolest supervisor and I am envious of your students. You simply have the magic to make people smooth away the puzzles and feel relaxed especially when there is a cup of good coffee and of course I am not talking about the caffeine effect. You are a fascinating problem solver. Thank you for lighting my mind and guiding me the way to the next academic destination.

Further, I would like to thank Aleksandar, Leonard, Hong, Micheal, Matt, Jaap, Thomas, Kaj, Win King for all the supports and helps during the job market period. Aleks and Leonard, I learnt a lot of strategies from your job market experience. Thank you for sharing your stories and inside information with me. I wish you both a very successful career in academia and I hope you will become tenured in a very soon future. Matt, Thomas, Jaap and Kaj, thank you for giving me a hard time during the mock interview and the job market similar. Your suggestions, questions and comments on my show turned to the guts and the confidence that helped me survive the job market. Hong, JMC No.27 and Micheal, we had lots fun conversation in Nashville. Thank you for your encouragement during our “dark” days. Hong, I wish you a very successful tenure-track journey in Hong Kong, your dream place.

I also owe a debt of gratitude to the four dearest secretarial ladies of the Finance Department: Carina, Cécile, Els and Francien. You keep my life running in this coun-
try. Tax return, garbage recycling, fitness card, reimbursement, visa application, moving boxes, etc, you are like superwomen. Thank you for taking very good care of me, and making me feel home. I would also like to thank my dearest colleagues of the Finance Department. My roommates at B1.04: Aleks, Anna, Rogier Q, Wiebke, we shared many enjoyable moments. Aleks, you simply know everything about the pension funds. Your name will be surely on those top four finance journals. Anna, I always like your work hard play hard life style. Rogier, firstly, I have to apologize for calling you Angry Birdies for so many years. It must be weird for an Asian to make fun of the Dutch last name. You are the smartest and most efficient researcher I have ever seen. I am sure you will always be a legend in your future career. Wiebke, you are such a cute pretty girl with a beautiful personality. I am so amazed by your negotiation ability. Marketing is indeed your cup of tea. Omar, my officemate at TS51. You are a very cool roommate, sometime very quiet but sometimes super funny. Congratulations on your Dutch citizenship. I enjoyed your Egyptian style singing very much, although I thought you were praying at first. Mukul and Paul, we had a lot of fun time together, the Netpsar conference dinner, quiz nights or many other parties. I shared many secrets with you. I will miss the moments hanging out with you guys. Andrea, you are an amazing girl with a warm big heart and a super colorful lifestyle. I wish you a very successful career in Real Estate. Paulo, you are a good listener, a very easy-going PhD helper, a sophisticated thinker and a great scientist. Your door is always open for every one. Thank you for all your advice on my research plans. Matteo, Gildas and Oana (and Rogier H), we had a wonderful trip in Chicago 2012 and I was so lucky to celebrate my birthday in the downtown of Chicago with you guys. We went through the budget cut of the department, but eventually, we all survived and found the next destination of our life. I wish you all a great success in your careers and I will surely miss you all. I would also like to express my gratitude to my senior colleagues: Arian, Arvid, Dennis, Jeroen, Paulo, Piet, Rachel, Rob, Stefanie and Stef. My pleasant doctoral life in Maastricht also thanks to Alessandro, Carl, Clarissa, Ehsan, Katie, Lei, Rogier, Gaby, Georges, Hang, Iman, Inka, Patrick, Vincent, Pomme, Sanne, Sjoke, and Tobias. I will miss you all. I am also deeply indebted to my new boss, Li Qi. You are such an amazing Econometrician. Thank you for choosing me as a part of your team at CUEB. I feel so lucky and blessed. My future colleagues, Jing Shi, Wenchuan, Su Li, Lin Wei, Hou Jie and the rest of the ISEM group, thank you for your hospitality during my stay in Beijing. I am very looking forward to working with you guys.

Besides research, I also developed many other hobbies during my PhD life. Paul, your birthday party turned me to an amateur belly dancer. Andrea, thank you for your thanks giving dinners. Now, I know how to make michelin-level butternut squash and thanks giving stuffing. Gaby, your Zumba career makes me falling in love with Zumba fitness.

It seems that it was only yesterday when my parents brought me to international airport of Beijing and sent me to a small country that I had barely heard of before. From
18 to 28, the Netherlands witnessed the mostly important ten years of this funky Asian girl. There is a very special woman who is always standing on my side, supporting me, encouraging me and giving me love unconditionally. She is a high school English teacher. Her students call her “Sister QingYang” and I call her mum. Mama, I cannot find a good word in my dictionary to express how grateful I am for all you have done for me throughout my life. Sending me abroad right after the high school was surely a risky and a painful decision. After all, I hope you find it a good investment.

亲爱爸爸妈妈谢谢你们一路的支持，我这个人，缺点很多其中之一就是优点太少，在我仅存的优点当中，其一是坚持；其二是大爱，人因梦想而伟大，我会坚持我的梦想，同样我也会用一颗博爱的心温暖世间，这就是你们眼中永远长不大的，偶尔脾气很差的，有时候很固执但是大多数时候还算可爱的女儿。妈妈，谢谢你在我极度叛逆的青春期送我出国留学，漂泊异乡十年，学会最多的就是珍惜拥有的和感恩，爸爸，虽然我们性格不合经常斗嘴，但是我知道，你是这个世界上最爱我的男人，父女没有隔夜仇，更何况这一隔便是十年。小木林永远最爱你们。

Maastricht, April 2015
# Contents

Acknowledgements iii

1 Introduction 1

1.1 A Brief Review of Uncertainty 2

1.1.1 Why does uncertainty matter? 2

1.1.2 How to deal with uncertainty? 3

1.2 Property of Robustness 5

1.3 Highlights and Contributions of this Dissertation 5

1.3.1 Parameter Uncertainty 7

1.3.2 Individual Investors and Model Uncertainty 7

1.3.3 What does an Investor Have Doubts About? 8

1.3.4 How Large is Ambiguity Aversion? 8

1.4 Topics and Organization 10

1.4.1 Robust Asset Allocation with Downside Risk in an Incomplete Market 10

1.4.2 Pricing and Hedging an Ultra Long-Dated Liability 12

1.4.3 Life-Cycle Investment Under Model Uncertainty 14

2 Robust Hedging in Incomplete Markets 17

2.1 Introduction 17

2.2 Literature Review 19

2.3 Model 20

2.3.1 Asset and Liability Model 21

2.3.2 Robust Asset and Liability Model 21

2.3.3 Robust Optimization Problem 24

2.4 Static Robust Optimization 25

2.4.1 Static Solution 26

2.4.2 Static Robust Portfolio Choice 28

2.4.3 Cost of Preference for Robustness 31

2.4.4 Policy Evaluation 32

2.4.5 Sensitivity Analysis 34

2.5 Dynamic Robust Optimization 36
2.5.1 Dynamic Programming ........................................ 37
2.5.2 Numerical Solution ....................................... 39
2.5.3 Dynamic Policy Evaluation ............................... 41
2.6 Conclusion ..................................................... 42
2.7 Chapter 2 Appendix ........................................... 45
   2.7.1 Change of Measure .................................... 45
   2.7.2 Static Optimal Solution ............................... 47
   2.7.3 Numerical Methodology ............................... 48
       2.7.3.1 Simplification Stage 1 ....................... 49
       2.7.3.2 Simplification Stage 2 ....................... 50
       2.7.3.3 Explicit Finite Difference .................... 51
3 Robust Hedging of Long-Dated Liabilities .......................... 53
   3.1 Introduction ............................................. 53
   3.2 Term Structure Model ................................... 56
       3.2.1 Model Calibration ............................... 57
   3.3 Robust Optimal Portfolio Choice ........................ 59
       3.3.1 Robust Hedging .................................. 59
   3.4 Numerical Technique .................................... 61
   3.5 Long-term Investors and Bond Premia Uncertainty ......... 63
       3.5.1 Optimal Portfolio Choice ....................... 63
       3.5.1.1 Delta Hedging ................................ 63
       3.5.2 Robust Term Structure ......................... 65
   3.6 Conclusion ................................................. 66
   3.7 Chapter 3 Appendix ...................................... 67
       3.7.1 Robust LSMC algorithm .......................... 67
       3.7.2 Naive LSMC ....................................... 67
       3.7.2.1 Grid Generation .............................. 67
       3.7.2.2 LSMC .......................................... 67
       3.7.3 Closed-Form Nature’s Choice .................... 69
       3.7.4 Robust LSMC ....................................... 70
       3.7.5 Goodness of Fit .................................. 71
4 How Much Should Investors Adapt? .................................. 79
   4.1 Introduction ............................................. 79
   4.2 The Model ................................................ 83
       4.2.1 Financial Market ................................. 83
       4.2.2 Investor’s Labor Income ......................... 84
   4.3 Basic Optimal Portfolio Model ........................... 84
       4.3.1 Monetary Policy Uncertainty and Income Uncertainty .......................... 86
4.3.2 Robust Optimal Portfolio Model ........................................ 86
4.4 Optimal Portfolio Choice ..................................................... 87
  4.4.1 Model 1: No Labor Income, No Model Misspecification ............ 87
  4.4.2 Model 2: No Labor Income, Under Model Misspecification ....... 89
  4.4.3 Model 3: With Labor Income, No Model Misspecification ........ 90
  4.4.4 Model 4: With Labor Income and Model Misspecification ....... 92
4.5 Model Calibration ............................................................... 93
  4.5.1 Data ........................................................................ 93
  4.5.2 Estimation ................................................................. 94
  4.5.3 Hybrid Estimation Method ............................................. 94
4.6 How Large is Ambiguity Aversion Parameter ............................ 97
4.7 Numerical Solution ............................................................. 100
  4.7.1 Robust Optimal Life-Cycle Decision without Labor Income .... 100
4.8 Robust Optimal Life-Cycle Decision with Labor Income ............ 103
4.9 Conclusion .................................................................. 105
4.10 Chapter 4 Appendix ........................................................... 107
  4.10.1 Nominal bond pricing ................................................... 107
  4.10.2 Estimation Procedure .................................................. 107
  4.10.3 Proofs ...................................................................... 112
    4.10.3.1 Proof of Proposition 4.4.1 ....................................... 112
    4.10.3.2 Proof of Proposition 4.4.2 ....................................... 113
    4.10.3.3 Proof of Proposition 4.4.3 ....................................... 114
    4.10.3.4 Proof of Proposition 4.4.4 ....................................... 117
    4.10.3.5 Proof of Proposition 4.4.5 ....................................... 118
    4.10.3.6 Proof of Proposition 4.6.1 ....................................... 119
5 Conclusion ................................................................. 129
  5.1 Summary of Findings ....................................................... 129
    5.1.1 Robustness is not Conservativeness .............................. 130
  5.2 Future Research ............................................................. 130
    5.2.1 Mean Reversion Parameter Uncertainty ......................... 130
    5.2.2 Consequence of Model Uncertainty with Investment and Borrowing
         Constraint ................................................................. 131
6 Who Needs Robustness? ........................................................ 133
  6.1 Pension Funds Regulators .................................................. 133
  6.2 Portfolio Managers .......................................................... 135
Nederlanse Samenvatting ...................................................... 143
## List of Figures

1.1 Greek Myth of Fragility, Resilience and Antifragility. . . . . . . . . . . . . 6  
1.2 Equity Premium Puzzle Remains a Puzzle. . . . . . . . . . . . . . . . . . . 10  
1.3 Projection of future funding ratio. . . . . . . . . . . . . . . . . . . . . . . 11  
1.4 Extrapolated term structure curves. . . . . . . . . . . . . . . . . . . . . . 13  
1.5 Inflation Survey Data Histogram. . . . . . . . . . . . . . . . . . . . . . . . 15  
1.6 Dispersion of income growth. . . . . . . . . . . . . . . . . . . . . . . . . . 16  
2.1 Static optimal portfolio choice. . . . . . . . . . . . . . . . . . . . . . . . . 29  
2.2 Static optimal perturbations $\lambda_1$ and $\lambda_2$. . . . . . . . . . . . . 30  
2.3 Mean rate of stock and liability return with and without the preference for robustness. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31  
2.4 Cost of Preference for Robustness $T$. . . . . . . . . . . . . . . . . . . . . 32  
2.5 Loss function equivalent curves. . . . . . . . . . . . . . . . . . . . . . . . . 34  
2.6 Sensitivity analysis with $\rho = 1$. . . . . . . . . . . . . . . . . . . . . . . 35  
2.7 Sensitivity analysis with respect to $\rho$. . . . . . . . . . . . . . . . . . . . 36  
2.8 Dynamic robust optimal hedging strategy. . . . . . . . . . . . . . . . . . . 39  
2.9 Dynamic robust and naive optimal hedging strategy. . . . . . . . . . . . . 40  
2.10 Dynamic optimal perturbation processes. . . . . . . . . . . . . . . . . . . 41  
2.11 Dynamic perturbation effect on drift terms. . . . . . . . . . . . . . . . . . 42  
2.12 Dynamic loss function equivalent curve. . . . . . . . . . . . . . . . . . . . 43  
3.1 Bond Premia . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60  
3.2 Nature’s Optimal Decision. . . . . . . . . . . . . . . . . . . . . . . . . . . . 62  
3.3 Optimal Bond Portfolio for Different Horizons . . . . . . . . . . . . . . . 72  
3.4 Optimal Portfolio under Different Funding Ratio (1). . . . . . . . . . . . . 73  
3.5 Optimal Portfolio under Different Funding Ratio (2). . . . . . . . . . . . . 74  
3.6 Robust Policy under Different Spot Rate. . . . . . . . . . . . . . . . . . . 75  
3.7 Optimal Expected Shortfall (1). . . . . . . . . . . . . . . . . . . . . . . . . 76  
3.8 Optimal Expected Shortfall (2). . . . . . . . . . . . . . . . . . . . . . . . . 77  
3.9 Robust Yield Curve . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 78
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Summary Statistics</td>
<td>57</td>
</tr>
<tr>
<td>3.2</td>
<td>GMM Estimates</td>
<td>58</td>
</tr>
<tr>
<td>4.1</td>
<td>Estimation of Model Parameters.</td>
<td>120</td>
</tr>
<tr>
<td>4.2</td>
<td>Comparison of Sample Moments with Model Implied Distribution</td>
<td>121</td>
</tr>
<tr>
<td>4.3</td>
<td>Bond Risk Premia and Volatilities</td>
<td>121</td>
</tr>
<tr>
<td>4.4</td>
<td>Correlation between returns on the stock, CPI and bonds with maturities of 1, 3, 6 and 10 years.</td>
<td>121</td>
</tr>
<tr>
<td>4.5</td>
<td>Optimal Portfolio Choice and Consumption without Labor Income</td>
<td>122</td>
</tr>
<tr>
<td>4.6</td>
<td>Optimal Portfolio Choice and Consumption with Labor Income.</td>
<td>124</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

How to make an investment decision when an economic agent does not trust his or her model? The goal of this dissertation is to introduce a powerful decision-making rule that remains stable in the presence of model uncertainty. A decision-making rule of this type is called “a robust strategy”. According to Hansen and Sargent (2007), a model is a probability distribution over a sequence of random variables. The economic decision-maker builds up her own subjective mental models of the economy based on her understanding of the financial market and economic policies. However, if a mental model fails to capture the dynamics of the economy, then the model is misspecified. Life, or nature is full of uncertainties, financial crises, policy changes, climate changes, natural disasters, etc. These uncertainties are unpredictable and unmeasurable, but can influence the economy as much as the economic decision-making. Doubts about the accuracy of the model makes an agent treat it as an approximation of an unknown true model. She wants her decision rules to work well over a set of models in the neighborhood of the approximating model. In this dissertation, I use a robust decision theory to formalize model uncertainty and I argue that in particular, institutions under severe financial distress as well as long-term investors should consider employing a robust decision rule, for these agents are especially sensitive to model uncertainty and a robust policy can protect them from the unknown crises.

This chapter consists of four sections. Section 1.1 explains why it is important to consider model uncertainty when making economic decisions. Given that model uncertainty matters, then how to measure it? Armed with max-min expected utility decision criterion, I employ robust control theory to quantify model uncertainty. Section 1.2 presents more details about the preference for robustness as well as the philosophy behind this. Section 1.3 narrows down into the scope of my thesis. First, I consider a heterogeneous model where an individual agent’s economic decision does not alter the equilibrium of the economy. Second, the particular type of model uncertainty that I deal with is also
CHAPTER 1. INTRODUCTION

called parameter uncertainty, meaning that some dimensions of the approximate model are measured incorrectly. The misspecification of function forms is beyond the scope of this dissertation. Third, I present two situations in which the model is more likely to be confronted with parameter uncertainty. One situation is when parameters are difficult to estimate. A typical example is the mean return parameter. The other case is when facing incomplete financial markets. Fourth, I introduce a new method for measuring the magnitude of the uncertainty aversion parameter: this is one of the contributions of this thesis to the robust asset allocation literature. Section 1.4 summarizes three studies included in this dissertation that deal with robust asset allocation problems in an incomplete market.

The first research question is: how should pension funds with large downside risks allocate their wealth when confronted with non-tradable liability risks and misspecified expected asset and liability returns. The second research project introduces a robust policy for hedging and pricing ultra-long-dated liabilities where bond markets do not exist. The third project quantifies the impact on optimal consumption and investment strategies when a robust life-cycle investor\footnote{A robust investor is characterized by a max-min expected utility decision criterion. She looks for the optimal solution that maximize her utility under an endogenous worst case scenario.} is afraid of inflation and income model uncertainty.

1.1 A Brief Review of Uncertainty

1.1.1 Why does uncertainty matter?

Uncertainty is a state of having limited knowledge. There is a meaningful distinction between risk and uncertainty. The necessity of distinguishing risk from uncertainty has been posited by Knight (1921).\footnote{This kind of uncertainty is also called Knightian uncertainty or ambiguity.} Risk is a set of uncertain outcomes but with well-defined probabilities. Risk can be distinguished from uncertainty in the sense that risk is measurable, but uncertainty is the unknown. The Ellsberg (1961) paradox provides experimental evidence arguing that people tend to prefer situation with given rather than unknown probabilities. Ellsberg’s experiment can be summarized as follows. The decision maker can choose between two urns. The first urn has 50 black balls and 50 white balls. The second urn also contains 100 balls, both black and white, but there is no information about the proportion of white and black balls. Decision-makers will win a prize if they pick a black ball. The majority of the participants in the experiment choose the first urn, indicating that people are ambiguity\footnote{Guidolin and Rinaldi (2013) state on p.184 “ambiguity and uncertainty are not always clearly distinguished”. Camerer and Weber (1992) state on p.330 “Ambiguity is uncertainty about probability, created by missing information that is relevant and could be known.” In this dissertation, I treat “ambiguity” and “uncertainty” as interchangeable.} averse. The Ellsberg paradox
1.1. A BRIEF REVIEW OF UNCERTAINTY

shows that decision-makers care about uncertainty, and the experiment makes the axiom of Savage (1954) dubious since Savage only considers risk in the decision-making process. The experiment’s result reveals the phenomenon of uncertainty or ambiguity aversion.

Model uncertainty means having fragile beliefs about the probability distribution over a sequence due to limited knowledge or information incompleteness. However, traditional economic decision theory restricts the freedom to hold alternative beliefs about the likelihoods of states of the world. For example, before the mid-20th century, the Bayesian paradigm dominated the economic decision literature. According to the Bayesian approach, economic agents should always have their own subjective beliefs about the true probability measures even in the absence of objective probabilities.\(^4\) In other words, the Bayesian paradigm does not distinguish risk from uncertainty as all unknowns are represented by probabilities. Similarly, since the 1970s, the rational expectations hypothesis presented by Robert Lucas\(^6\) has become the mainstay of modern macroeconomics and finance literature. Armed with the rational expectations hypothesis, the decision-makers are assumed to have knowledge of the objective data generation process of the state variables, or their subjective beliefs coincide with the objective state of the economy.

Unfortunately, the standard economic decision theories (both Bayesian approach and rational expectations) are challenged by many empirical facts. For instance, Ellsberg’s experiment contradicts the Bayesian paradigm but is in line with the argument of Knight. In addition, the rational expectations hypothesis is challenged by a large number of asset pricing puzzles. For example, Mehra and Prescott (1985) conclude that a suspiciously high degree of risk aversion is needed to explain the huge positive difference between the equity premium observed from the data and from the model, which results in the equity premium puzzle.

1.1.2 How to deal with uncertainty?

The Ellsberg paradox shows that agents dislike facing uncertainties. But how to deal with uncertainties when making economic decisions? In practice, the amount of uncertainty that agents face depends on their knowledge of the true economy. The more useful information they can observe, the less uncertainty they are facing. Uncertainties arise both from problems in observing the relevant current economic variables as well as forecasting.

\(^4\)Hansen and Sargent (2010) define “fragile beliefs” as the sensitivity of pessimistic probabilities to the arrival news.

\(^5\)Cyert and DeGroot (1974) state on p.524 “To the Bayesian, all uncertainty can be represented by probability distribution.” Gilboa and Marinacci (2011) write “Even if you don’t know what the probabilities are, you should better adopt some probabilities and make decisions in accordance with them.”

\(^6\)Muth (1961) first introduced the idea of rational expectations which was further developed and included in macroeconomics by Lucas (1976).
CHAPTER 1. INTRODUCTION

future developments. However, much economic data, such as the GDP growth rate or the unemployment rate, is measured on a quarterly or an annual basis and the methods used to measure such economic data changes over time. Therefore the useful information an agent can obtain is very limited. Financial data have higher frequencies than economic data. Many stock prices are even available at minute-by-minute frequency, but forecasting future stock prices or stock returns has little success because the agent does not know the future distributions of the stock prices.

In the literature on decision-making under ambiguous beliefs in asset pricing and portfolio choice, many approaches have been developed in the past few decades to represent and to deal with ambiguity.\textsuperscript{7} For instance, Schmeidler (1989) proposes the Choquet expected utility (CEU) which allows for non-additive probabilities to represent uncertain events so as to compensate for the information incompleteness. Gilboa and Schmeidler (1989) provide the static multiple-priors utility model in which uncertain prospects are represented by a set of probability measures. The multiplicity of priors justifies decisions under a max-min expected utility criterion of Wald (1949). Departing from Bayesian or rational expectations in which decision makers believe only in one single model, the Gilboa and Schmeidler model allows decision makers to choose between and rank a set of subjective models, with the one that minimizes their utility bringing aversion to uncertainty. Inspired by Gilboa and Schmeidler max-min expected utility theory, Lars Hansen and Thomas Sargent develop a new school of ambiguity representation, namely model misspecification and robustness. A series of studies, such as Anderson et al. (2003), Hansen and Sargent (2001) and Hansen and Sargent (2007) employ robust control theory from the operational research literature such as Ben-Tal et al. (2009) to deal with ambiguity. Robust control theory shares the same rationale as Gilboa and Schmeidler axioms, for in both cases, rational decision makers maximize their expected utility under some worst-case model decided by ‘nature’.\textsuperscript{8}

In this dissertation, I adopt the Hansen and Sargent framework to represent the fear of uncertainty. An economic decision maker who worries about model misspecification looks for a prudential policy that is resilient to fragile beliefs about the likelihood of the state variables. Such decision rules are called robust policies. A preference for robustness protects the decision maker from model uncertainty. In order to obtain a robust decision, one has to follow a max-min expected utility criterion.

\textsuperscript{7}See Gilboa and Marinacci (2011) and Guidolin and Rinaldi (2013) for more details about the review of those methods.

\textsuperscript{8}In his Risk Inaugural Lecture “The Beauty of Uncertainty” at Bocconi University in 2012, Thomas Sargent defines “nature” as “an artificial agent who helps an economic decision maker to construct bounds of value functions which allow the decision maker to analyze the fragility of the decision rule under a set of models that she think may prevail.”
1.2 Property of Robustness

What is robustness? Robustness is a property allowing toleration of distortions. Systems that are robust do not break down in volatile circumstances. To achieve vivid understanding of a robust system, I borrow an idea from Taleb (2012). In his recent book, *Antifragile: Things That Gain from Disorder*, he uses the Phoenix from Greek mythology to represent robustness. The Phoenix is a symbol of rebirth and self-curing and remains the same through recurrent demise and rebirth. Correspondingly, a robust model remains steady in disordered circumstances. A robust policy adopts the max-min expected utility principle to make sure the system not only works when the underlying model is correctly specified but also provides a healthy outcome in adversity.

In contrast, rational expectations and Bayesian approaches are both fragile, for decision makers under both paradigms rely wholly on one single model in which they believe. If the economic model is misspecified, then the corresponding economic decisions will not work in the real world. Taleb (2012) uses the Sword of Damocles as a powerful metaphor for a fragile system, since by adding in a bit of chaos, the sword hanging by a thin horse-hair will kill Damocles instantly.

Taleb argues that antifragility is the most ideal system for it grows and strengthens itself from volatility, like the mythical Hydra. In my opinion, nature herself is an antifragile system because she is unpredictable, filled with noise. Nature grows from complexity. It is impossible to capture nature’s movement. Voltaire says “Doubt is not a pleasant condition, but certainty is absurd.” Wilde (1891) says in “The Picture of Dorian Gray” that “knowledge would be fatal, it is the uncertainty that charms one.” As an economic decision maker, staying robust and resilient while enjoying the beauty of uncertainty is my idea of rationality.

1.3 Highlights and Contributions of this Dissertation

In their work on robust economic decision-making under model uncertainty, Hansen and Sargent (2007) conjecture that economic decision makers’ sluggish responses towards the recent recession or the changes in fiscal and monetary policy can be explained by the uncertainty they face. The applications of the Hansen and Sargent model are mostly in...
CHAPTER 1. INTRODUCTION

Figure 1.1: Greek Myth of Fragility, Resilience and Antifragility.

macroeconomics. How does model uncertainty and the preference for robustness influence the decision-making of an individual investor (such as a life-cycle investor, a pension fund, an insurance company or an entrepreneur)? More importantly, how much should a robust investor adapt her behavior when faced with model fragility? Which factors determine the preference for robustness? How can ambiguity aversion be measured?

My dissertation searches for answers to these questions while attempting to fill gaps in the literature on financial econometrics and robust asset allocation. I use econometric theory to show that investors’ uncertainty aversion levels are highly dependent on the revealingness of the underlying information. This property not only provides an exogenous boundary for the uncertainty aversion parameter but also allows for some potential insights into various asset pricing puzzles such as the equity home bias.

For instance, after the 2008 financial crisis, most pension funds were faced with severe solvency problems. The classic asset allocation theory argues that the optimal hedging
1.3. HIGHLIGHTS AND CONTRIBUTIONS OF THIS DISSERTATION

Portfolio is independent of the funding ratio. However, empirical evidence\textsuperscript{11} shows that pension funds with low funding ratios are more likely to invest in equities. How to hedge a long-dated and stochastic cash flow commitment with an insufficient amount of wealth has become a challenge to most pension funds. My primary focus of interest is therefore on pension funds with funding shortfalls.

1.3.1 Parameter Uncertainty

According to Hansen (2014), model uncertainty can be interpreted in two ways. First, external uncertainty, which refers to uncertainty about different functional forms. Model builders, such as econometricians and government policy makers are confronted with this type of model uncertainty. Second, internal uncertainty, which means some dimensions of one particular approximating model are misspecified. Model users, such as fund manager and individual investors, face internal uncertainty. My research mainly focuses on the second type of uncertainty. In my work, I rename “internal uncertainty” as “parameter uncertainty”.

1.3.2 Individual Investors and Model Uncertainty

Individual decision makers such as pension fund asset managers or life cycle investors differ from economists, government policy makers or econometricians in two main ways. First, as individual investors, their investment decisions are heterogenous and will not influence the market price or policy making. Heterogeneity can be expressed in terms of an investor’s financial status, life expectancy, education level, etc. Therefore I only deal with partial equilibrium problems in this dissertation, and a general equilibrium cannot be obtained due to the heterogeneity design. Consequently, Lucas (1978) style equilibrium asset pricing model is beyond the scope of this dissertation.\textsuperscript{12}

Second, an individual investor is a model user rather than a model builder. As a result, an ambiguity averse investor who worries about model uncertainty restricts her doubts to parameter misspecification. For instance, a life cycle investor may fear that some dimension of her labor income model, such as the income growth rate parameter or the income volatility parameter, is misspecified.

\textsuperscript{11}Rauh (2009) and Andonov et al. (2014) show that U.S. public pension funds have increased their allocation to riskier assets in the past two decades.

\textsuperscript{12}For example, Maenhout (2004) and Ju and Miao (2012) manage to adopt the Hansen and Sargent framework in a Lucas-style equilibrium asset pricing model to explain various asset pricing puzzles, such as the equity premium puzzle and the risk-free rate puzzle.
1.3.3 What does an Investor Have Doubts About?

An economic model may contain multiple parameters, but it is costly to consider misspecification of all dimensions. In this dissertation, I argue that parameters with the following features are most likely to suffer from specification errors. First, parameters that are notoriously difficult to estimate. A typical example is the expected equity return. Merton (1980) argues that it is difficult to estimate expected returns from time series of realized stock return data. Cochrane (1998) provides a statistical explanation. The standard deviation of the expected return is \( \sigma / \sqrt{T} \) where \( \sigma \) is the standard deviation of annual returns and \( T \) is the data size (number of years). The formula indicates that stock returns are too volatile to measure statistically. However, the estimates of second moments can be much more accurate (see Andersen and Bollerslev (1998)).

Second, pessimism also stems from the incomplete information. If a decision maker shares the same amount of information as nature, then she is obviously uncertainty neutral. However, in practice, this is unlikely to be the case. In general, there are two types of information incompleteness. One is a lack of data at hand. Small sample problems challenge economic research. The other type of information incompleteness is an incomplete market. In financial literature, an incomplete market refers to the impossibility of achieving a perfect hedge, meaning that not every payoff can be replicated by constructing a self-financing portfolio. For instance, a pension liability is exposed to longevity risk, long-term interest rate risk, mortality risk, etc. However, these risks cannot be fully hedged, since financial instruments exposed to these risks are either illiquid or do not even exist.

1.3.4 How Large is Ambiguity Aversion?

How much should investors adapt their behavior when facing model (parameter) ambiguity? In other words, how to measure the ambiguity aversion parameter? In the financial literature, there are two approaches. First, one can use an Ellsberg-style experiment to estimate the ambiguity aversion parameter (see Camerer (1995)). Chen et al. (2013) use the experimental method to determine the magnitude of ambiguity and they find that a reasonable range of the ambiguity aversion parameter is between 50 to 90 when the risk aversion parameter takes the range of 0 \( \sim \) 10. However, Maenhout (2004) argues that the experimental method is not suitable for portfolio choice and asset pricing problems because financial markets are much more complicated than the stylized Ellsberg urns.

Second, one may derive an equilibrium from the robust model, then use some econo-

---

\(^{13}\)For example, if \( T = 100 \) and \( \sigma = 16\% \), then the standard error of the equity premium is \( 1.6\% \left( \frac{16\%}{\sqrt{100}} \right) \), which leads to an approximate 95% confidence interval span of \( 6.3\% \pm (1.96 \times 1.6\%) \). Although the interval shrinks with the square root of time period for estimation, it is difficult to maintain the same data generating process throughout the entire period.
1.3. HIGHLIGHTS AND CONTRIBUTIONS OF THIS DISSERTATION

metric methods such as likelihood ratio tests or matching moments (e.g., Hansen and Singleton (1982) and Anderson et al. (2003)) to measure the ambiguity aversion parameters. For example, Maenhout (2004) uses the gap between the observed equity premium and the pessimistic equity premium derived from the equilibrium of the robust asset pricing model to measure the ambiguity aversion parameter, in order to claim that model uncertainty can explain the equity premium puzzle (see Figure 1.2). Ju and Miao (2012) parameterize the gap between the market interest rate and the equilibrium interest rate as a function of the degree of ambiguity. The estimated ambiguity level helps to interpret the interest rate puzzle. However, the logic behind this method is questionable. It is due to the manipulated hypothesis that model uncertainty is the only explanation of the asset pricing puzzles that one can quantify the ambiguity parameters in this way. One can also argue that model uncertainty might not be the only explanation of the puzzles. Further, the resulting ambiguity aversion level is heterogeneous, depending on the sample observations.

In this dissertation, I develop a new approach to measuring the ambiguity aversion parameter systematically. The new approach consists of three steps. First, I employ standard estimation theory to construct a joint statistical boundary for misspecified parameters. The boundary provides a confidence region for all ambiguous estimates. Second, I use robust control theory to obtain optimal perturbations which quantify the distances between point estimates and the true value of parameters. The optimal distortions are functions of ambiguity aversion parameters. Third, by simulating a large number of optimal perturbations under a series of trial preference parameters and mapping to the statistical boundary space, one can obtain a viable value for the ambiguity aversion parameter. Simulations outside the statistical boundary are highly unlikely to occur and the corresponding ambiguity levels are too high to be accepted.

The statistical boundary systematically constrains investors’ pessimism towards the reference model. Further, this boundary also rules out a large part of the numerical solutions from the previous robust asset pricing studies such as Maenhout (2004) and Branger et al. (2013) because these studies set the uncertainty aversion level too high, leading to a worst-case scenario that is statistically implausible. Figure 1.2 demonstrates that only a very small part of the equity premium puzzle can be explained by applying my approach (Sally’s Model) to Maenhout’s equilibrium asset pricing model. Therefore, the equity premium puzzle remains a puzzle!
CHAPTER 1. INTRODUCTION

Figure 1.2: Equity Premium Puzzle Remains a Puzzle. 2.1% refers to equity premium obtained from the reference model. The upward sloping line plots the expected equity premium observed from the data (1891 ∼ 1994) under different detection-error probabilities as a function of the ambiguity aversion parameter. See Maenhout (2004), Table 4.

1.4 Topics and Organization

The remaining part of the introduction provides a summary of main findings of each chapter.

1.4.1 Robust Asset Allocation with Downside Risk in an Incomplete Market

Following the 2008 global financial crisis, the performance of U.S. pension funds has remained depressed. The dismal solvency situation has been driven by a declining discount rate and also a fall in equity prices. As shown in Figure 1.3, in the past half decade, funding ratios (asset values divided by projected benefit obligations) of the top 100 largest U.S. corporate defined-benefit pension plans have not rebounded. At the end of 2014, the average funding ratio was 83.6%. The downside risks are so large that even an optimistic projection can barely meet liabilities by the end of 2016. In addition, Figure 1.3 also presents an increasingly wide range of funding ratio projections over the next two years, resulting in tremendous uncertainty. Figure 1.3 raises the question of how to price and
1.4. TOPICS AND ORGANIZATION

Figure 1.3: Projection of future funding ratio. This figure presents the Milliman 100 pension funding index average funding ratio from 2010 to 2014 and projections from 2015 to 2016. See Milliman Pension Fund Index us.milliman.com.

hedge downside risks when confronted with fragile beliefs about the likelihood of different funding ratio scenario. Chapter 2 provides a robust solution.

Pricing and hedging pension or insurance liabilities faces two problems. First, the market is incomplete. Liability risks are typically not (actively) traded in the financial market. For instance, a pension liability is exposed to long-term interest rate risk and longevity risk. Due to the missing market, these risks cannot be hedged by constructing a replicating portfolio. Mitigating such kinds of risks involves a tradeoff between risk and expected return. Second, on the other side of the balance sheet, expected asset return is very difficult to estimate from historical data (see Merton (1980), Cochrane (1998)). A poorly estimated expected return will lead to a suboptimal portfolio choice.

To illustrate the challenges discussed above, I present in Chapter 2 a robust optimal hedging strategy that takes both downside risks and market incompleteness into account for agents who fear model misspecification. Robust agents are assumed to minimize the shortfall between the assets and liabilities under the subjective worst case scenario by means of solving a min-max robust optimization problem. Without a doubt, the robust model includes three crucial elements, namely (1) downside risks; (2) an incomplete market and (3) model misspecification.

- Downside risks, the failure to fulfill pension commitments, are calculated by expected shortfall measurement. In a static model where control variables are stable over time, the shortfall between the assets and liabilities can be linked with the payoff on an exchange option introduced by Margrabe (1978). Ang et al. (2012) adopt such an option-type payoff function as an additional penalty term upon the
classic mean-variance hedge criterion.

- An investor who follows a liability-driven investment strategy faces undiversifiable idiosyncratic risks. Market prices of these unhedgeable risks are uncertain and cannot be spanned. The market incompleteness is expressed by an unhedgeable risk factor which is orthogonal to financial risks. Cochrane and Saa-Requejo (1996) introduce a Good-Deal-Bound technique to quantify the unspecified aggregate market price of risk. The insight of Good Deal Bound is also employed in this Chapter, but for a different purpose.

- A preference for a min-max expected shortfall objection function is used to deal with parameter uncertainty. I design a stylized robust hedging problem that integrates the Good-Deal-Bound rationale into the Hansen-Sargent approach. Inspired by Anderson et al. (2003) and Maenhout (2004), I introduce drift distortions to represent parameter misspecification. These drift distortions perturb the true data generation process from approximate models. The statistical measure for discriminating between models is determined by a relative entropy parameter used in Hansen and Sargent studies. The relative entropy is used to penalize specification errors. Meanwhile, drift distortions also generate additional market prices of risks. I use the Good-Deal-Bound method to limit the size of uncertainty, for the bound gives some insight into the investor’s ambiguity aversion level.

I provide both static and dynamic solutions to this robust optimization problem. First, I find that ambiguity-averse investors are more conservative with their investment decisions compared with the naive investor who is ambiguity neutral. Cutting down the portfolio’s stock market weighting simultaneously protects investors from misspecified parameters. Second, the investment decision also depends on the planning horizon. The longer the hedging horizon, the lower the risk exposure becomes. Third, both robust and naive models induce investors to gamble their way out of trouble when the instantaneous funding ratio is very low. Fourth, I also find that below a given level of downside risk, a robust policy requires less initial wealth when the expected stock return is overestimated.

1.4.2 Pricing and Hedging an Ultra Long-Dated Liability

Pension funds are obliged to face an ultra long-dated payoff commitment. Pricing and hedging long dated liability faces two challenges. On the one hand, there is no market for long maturity financial instruments. The longest government bonds even in developed financial markets (such as the US and Germany) have maturities no longer than 30 year. However, pension liabilities are well-defined future cash flow obligations with maturities
1.4. TOPICS AND ORGANIZATION

Various term structure models imply various yield curves when maturity is beyond 20 years. Recourse of the figure: Bragt and Slagter (2013), Figure 1, “Ultimate forward rate: The way forward?”

of at least over 50 years. Solvency II regulations require that pension and insurance companies should follow a market-consistent principle in valuing their liabilities. Here comes the puzzle: how could we implement market-based valuation of those commitments with maturities of longer than 30 years; or how to discount the far future cash flows in the absence of a complete bond market?

On the other hand, there is model uncertainty. Any valuation of a long dated liability with a partially missing bond market must be model based. Extrapolating the term structure curve can be dated by the 1960s. To date, mountains of work has been done on improving the specification of the term structure model. Conditional on a suitable model, we could derive a term structure for all maturities. However, as shown in Figure 1.4, there are a large number of different models that perfectly fit bond prices up to maturities of 20 years but yet imply different prices for the non-traded-period bonds. Therefore, given the incompleteness of the bond market, it is impossible to avoid model misspecification no matter how sophisticated the underlying model is.

In Chapter 3, I introduce a robust investment strategy to hedge long dated liabilities under model misspecification and incomplete bond markets. A robust agent who worries about misspecified bond premia follows a min-max expected shortfall criterion to protect against model uncertainty. I employ a backward least squares Monte Carlo method to
CHAPTER 1. INTRODUCTION

solve this dynamic robust optimization problem.

I find that the preference for robustness induces a strong demand for long-term bonds when the solvency ratio is low. Robust investors worry about model uncertainty in the sense that they are afraid that nature would choose lower bond premia than they expected, hence they need a risker portfolio in order to gamble their way out of trouble. Both naive and robust optimal portfolios depend on the hedging horizon. The longer the horizon, the more risk exposure on long-term bond markets. I also find that a robust policy requires more initial wealth than the naive policy, in order to meet the shortfall fall target. In other words, the robust yield curve is always lower than the naive yield curve. However, when the spot rate is low, both policy-based yield curves are higher than the reference yield curve.

1.4.3 Life-Cycle Investment Under Model Uncertainty

Departing from pension funds risk management, in Chapter 4, I investigate a life-cycle investment and consumption problem where a life cycle investor is ambiguity averse due to the fear of misspecified inflation and income models. Although the rate of price inflation is less volatile than equity returns, Stock and Watson (2007) argue that US inflation has become harder to forecast. Ang et al. (2007) claim that survey forecasts beat model-based inflation prediction performances. As shown in Figure 1.5, there is a huge diversity of beliefs among economic agents about the expected future inflation rate. For long-term individual investors (young cohorts), inflation dynamics play an important role in optimal bond portfolios (see Brennan and Xia (2002), Munk et al. (2004)). However, inflation risk cannot be fully hedged for the long run, because the real-bond market is incomplete. Consequently, a misspecified inflation model can jeopardize optimal investment and consumption.

Labor income drives the optimal long-term investment decision of an individual investor dramatically (Cocco et al. (2005)). Guvenen (2007) argues that the income growth rate should be individual specific. Haider (2001) and Guvenen and Kuruscu (2010) show empirically that income-growth dispersion has been broadening since the 1970s. Heterogeneity, however, adds huge uncertainty into the estimation of the expected income growth rate (Wang (2009)). As demonstrated in Figure 1.6, ignoring the heterogeneity of income growth results in a huge bias in estimated persistence.

14Stock and Watson (1999) claims that Phillips curve-based forecasts outperform other models’ performance. However, Stock and Watson (2007) states that “after 1984 it has been harder to be an inflation forecaster, in the sense that it is more difficult to improve upon simple univariate models, at least using activity-based backward-looking Phillips curves.”

15Inspired by Giordani and Söderlind (2003), I use the Survey of Professional Forecasters as a proxy for inflation model uncertainty. The survey was carried out in 2012, and 30 economists filled in the questionnaire. See www.philadelphiafed.org/.
The research question of this chapter is as follows. First, where should life cycle investors put their money when they have fragile beliefs in their income process and future inflation dynamics. Second, if a robust policy is presented to protect the investor from ambiguities, then how much should their wealth be reallocated?

To answer these questions, I first set up a dynamic robust life cycle consumption and investment model under which investors with a max-min expected utility preference can allocate their wealth to various financial instruments including stocks, nominal bonds. I provide a tractable solution with the help of “homothetic robustness” which is developed by Maenhout (2004). Robustness increases the demand for long-term bonds while reducing exposure to stocks. Second, I present a feasible boundary for the ambiguity aversion parameter using econometric theory. This parameter may help the investor to quantify systematically how much to adapt her investment and consumption behavior. The ambiguity aversion parameter ranges from 0 to 2 based on my calibration result.\textsuperscript{16} Choosing

\textsuperscript{16}This range is much narrower than any other calibration results such as Ju and Miao (2012) and Maenhout (2004).
Figure 1.6: The figure presents the projections of future expected income growth rates of 25-year old cohorts under different education levels. See Guvenen (2009), Figure 1.

values beyond 2 is over-pessimistic as those ultra bad scenarios are statistically unlikely to happen.
Chapter 2
Robust Hedging in Incomplete Markets*

2.1 Introduction

Pricing and hedging pension or insurance liabilities faces two problems. First, the market is incomplete. Liability risks are typically not - or not actively - traded in the financial market. For instance, a pension liability is exposed to long-term interest rate risk and longevity risk. Due to the lack of a market, these risks cannot be hedged by constructing a replicating portfolio. Mitigating such kinds of risks involves a tradeoff between risk and expected return.

Second, on the other side of the balance sheet, the expected asset return is very difficult to estimate from historical data. Suppose we have a series of stock price for $T$ years. The expected log return can be estimated by \( \frac{\ln(S_T) - \ln(S_0)}{T} \). If we have 100 years of historical data and the market volatility is 16\%, then the standard error of the average equity return is 1.6\% (\( \frac{16\%}{\sqrt{100}} \)), leading to an approximate 95\% confidence interval span of 6.3\% (±1.96 × 1.6\%). Although the interval shrinks by the square root of the length of the sample period for estimation, it is difficult to maintain the same data generating process during the entire period. The investor is therefore exposed to estimation error in the expected asset returns. A poorly estimated expected return will lead to a suboptimal portfolio choice.

Despite many studies on pricing and hedging in incomplete markets, very few studies consider the effect of parameter uncertainty. The literature on portfolio choice under estimation risk can date back to 1970s. One of the pioneering works by Bawa et al.

---

*This chapter is based on my working paper Shen et al. (2013) co-authored with Antoon Pelsser and Peter Schotman (Maastricht University and Netspar).

2Kandel and Stambaugh (1996) and Barberis (2000) show that regressing stock returns on a set of “predictive” variables such as the lagged dividend yield has almost zero explanatory power (\( R^2 < 4\% \)).
(1979) had already noted that the optimal hedge differs from the classic Morgenstern and Von Neumann (1953) axioms if estimation risk is taken into consideration.

In this paper, we analyse a robust hedging strategy in incomplete markets. The main goal is to design a hedging strategy that not only works well when the underlying model is correct but also performs reasonably well under model misspecification. To be more specific, we assume that the agent makes an investment decision to hedge the downside liability risk, the payoff of which cannot be fully replicated due to market incompleteness. Meanwhile, the agent is pessimistic towards the underlying model. In order to neutralize the effect of model ambiguity, the agent follows a robust policy that insures him against a worst case scenario. The robust policy is less sensitive to the estimation error, but this additional guarantee also makes the robust policy more expensive.

There are two ways to understand the preference for robustness. First, market incompleteness creates an unknown market price of risk. This unobservable market price of risk leads to the rationale of Cochrane and Saa-Requejo (1996)’s Good-Deal-Bound that limits the maximum Sharpe ratio of the market. The true market price of risk could be anywhere within this Good-Deal-Bound. Second, investors who fear parameter misspecification are ambiguity averse. They believe that the true model parameters differ from the estimated parameters. To formulate model misspecification, Hansen and Sargent (2001) and Hansen and Sargent (2007) employ a relative entropy factor. This relative entropy factor captures the perturbation between the estimated model and the unobservable true model. Although the economic interpretations of the two are different, both interpretations can be understood as requiring an additional premium to represent the estimation error.

Our robust optimization problem involves four elements. First is an incomplete market. We introduce two uncorrelated risk drivers in our model, one hedgeable and the other not hedgeable. The unhedgeable risk captures the incompleteness of the market. The asset market is exposed to hedgeable risk only, but the liability side is exposed to both types of risk. The second element is the parameter estimation error. We introduce two perturbations: one only works on the asset side, whereas the other operates on both the asset and the liability side of the balance sheet. Each of the two is defined as an additional drift term on the Brownian motion. Economically, an additional drift on the Brownian motion can be understood as the unobservable market price of risk, which is in line with Cochrane and Saa-Requejo (1996)’s explanation. Technically, the two parameters measure the discrepancies between alternative probability distributions. This explanation is in line with Hansen and Sargent (2007)’s relative entropy approach. The third element

---

3Under the Von Neumann-Morgenstern paradigm, an individual maximizes the expected utility of returns to obtain the optimal portfolio where the expected utility function is unique either subjectively or objectively.
2.2 Literature Review

This section reviews the literature on hedging in incomplete markets. The past decade has witnessed a large number of studies on pricing and hedging in incomplete markets. Cochrane and Saa-Requejo (1996) weaken arbitrage pricing theory by using a “good deal bound” (GDB) as an acceptability criterion to rule out high Sharpe ratio portfolios. However, the size of the GDB is determined exogenously. In this chapter, we provide a statistical measure to quantify the interval of the GDB. Sangvinatsos and Wachter (2005) and He and Pearson (1991) introduce a min-max local martingale measure to solve the optimization problems in incomplete markets. They suggest completing the market by using an additional parameter on the pricing kernel so that the demand for the new claim is zero. A similar technique is also employed by Schachermayer (2004). However, unlike these studies, we only consider downside risks in an incomplete market. Hodges and Neuberger (1989) introduce an indifference pricing principle to value European options.
CHAPTER 2. ROBUST HEDGING IN INCOMPLETE MARKETS

in the presence of transaction cost. More recently, Young and Zariphopoulou (2002) have extended this technique to pricing of a dynamic insurance risk. Young (2004) calculates the price of an insurance claim which is exposed to catastrophe risk and stochastic interest rate risk. Most of these studies are surveyed by Henderson and Hobson (2004).

This study is also related to robust asset pricing literature with a max-min expected utility preference. The max-min expected utility paradigm is developed by Gilboa and Schmeidler (1989). Gilboa and Marinacci (2011) claim that Gilboa-Schmeidler's axiom is a neo-Bayesian paradigm because it allows decision makers to have a set of subjective priors. Agents aim to maximize their utility under the least preferred prior so as to display an aversion to uncertainty. As an extension, Hansen and Sargent (2001) managed to transform Gilboa-Schmeidler's static theory into a dynamic version through the techniques of robust control theory, which had already been widely applied in engineering and applied mathematics since the 1980s. Maenhout (2004) is one of the first studies to employ the Hansen-Sargent framework to revisit the Merton (1969) optimal portfolio problem. He concludes that uncertainty aversion causes investors to reduce their stock market risk exposure dramatically. Uppal and Wang (2003) extend Maenhout’s model in a multi-dimensional setting. This paper is closely related to the Hansen and Sargent framework, although instead of using relative entropy to measure the magnitude of uncertainty, we employ the GDB.

The rest of the chapter proceeds as follows. Section 2.3 introduces the general economic setup. In this section, we formulate the market incompleteness and model misspecification and construct the uncertainty set. We also define the robust optimization problem. In Section 2.4, we derive the value function analytically under a static environment, and numerically calculate the static robust portfolio. We also evaluate the robust policy via the loss function. Section 2.5, provides dynamic analysis of the robust policy. We solve the partial differential equation numerically using an explicit finite difference approach. Section 2.6 gives the summary and conclusions.

2.3 Model

We construct a continuous-time incomplete market with a finite trading horizon \([0, T]\). The risk is modeled by a filtered probability space \((\Omega, \mathcal{F}, \mathbb{P})\), on which are defined two uncorrelated risk factors, a hedgeable risk \(W_1\) and a unhedgeable risk \(W_2\). Both \(W_1\) and \(W_2\) are univariate standard Brownian motions and we consider \(\{\mathcal{F}_t : t \in [0, T]\}\) as the completion of the filtration generated by \(W_1\) and \(W_2\). A hedgeable risk means we can

---

4 In order to obtain a closed form solution, Maenhout (2004) modifies Hansen-Sargent’s framework by transforming the constant Lagrange multiplier into a function of state variables. However, Pathak (2002) argue that this modification breaks the link to the world of Gilboa-Schmeidler.
replicate the payoff of this kind of risk perfectly. The payoff for a unhedgeable risk is not replicable because it is not traded.

### 2.3.1 Asset and Liability Model

On the asset side, we have a risk-free money-market account $B_t$, which earns a deterministic risk-free rate of interest $r$, so $dB_t = rB_t dt$. We also have a stock market. The stock price follows a geometric Brownian Motion process $dS_t = \mu S_t dt + \sigma S_t dW_1$. The agent can only invest in the money-market account and the stock market. Denote the value of the assets at time $t$ by $A_t$. The investor puts $wA_t$ amount in the stock market at time $t$. Note, $w = w_t$ is a stochastic process but not constant. We ignore the time subscript $t$ for notation convenience. The remaining part of the assets $(1 - w)A_t$ is put into the money-market account. The asset diffusion process follows as

$$dA_t = \left(r + w(\mu - r)\right) A_t dt + w \sigma A_t dW_1,$$

where $w$ is the possibly time-varying hedging strategy. We do not set a constraint on $w$, therefore short positions are allowed.

The liability is exposed to both hedgeable risk $W_1$ and unhedgeable risk $W_2$. We assume that the diffusion process of the liability $L_t$ follows an exogenously given geometric Brownian motion with constant drift term and constant volatility,

$$dL_t = a L_t dt + b L_t \left( \rho dW_1 + \sqrt{1 - \rho^2} dW_2 \right),$$

where $a$ is the drift of the liability and $b$ is its volatility. The non-traded risk driver, $dW_2$, represents the incomplete part of the market. We introduce a correlation parameter $\rho \in [-1, 1]$ between asset risk and liability risk. It controls the risk exposure to $W_2$ of the liability. If $\rho = \pm 1$, then the non-traded risk $W_2$ disappears from the liability side. The liability in this case can be perfectly hedged by a replicating portfolio. We are interested in the case when $\rho$ is strictly between $-1$ and $1$.

### 2.3.2 Robust Asset and Liability Model

We use the Hansen and Sargent (2007) framework to integrate the preference for robustness to the asset-liability model (2.1) and (2.2). With a preference for robustness, the agent treats (2.1) and (2.2) as an approximate model for the unknown true state evolution of $A_t$ and $L_t$. We design a particular form of model misspecification by limiting the parameter uncertainty to the drift terms $\mu$ and $a$ only. We assume that the volatility term
σ is known. The approximate model only provides an estimated value of the drift terms, but the expected return is imprecisely estimated and is subject to estimation error. However, the constant volatility parameters σ and b can be estimated using high frequency observations and are therefore not subject to parameter estimation error. This reduces the misspecification to uncertainty about the drift terms of the state variables.

In the Hansen and Sargent framework, the robust model contains an unknown drift term on the Brownian motion, so in our case the robust models dW\textsubscript{1} and dW\textsubscript{2} in (2.1) and (2.2) are replaced by dW\textsubscript{1} + λ\textsubscript{1}dt and dW\textsubscript{2} + λ\textsubscript{2}dt. The two drift terms λ\textsubscript{1} and λ\textsubscript{2} are defined as two perturbation time series processes that quantify the misspecification of the underlying model. The values of λ\textsubscript{1} and λ\textsubscript{2} are constrained by an uncertainty set. We provide two interpretations of these two additional terms. First, they shift the mean distribution of the asset and the liability diffusion process by a unit of wσλ\textsubscript{1} and bρλ\textsubscript{1} + b√1−ρ\textsuperscript{2}λ\textsubscript{2}, respectively. Hence they specify a set of alternative measures referring to different specifications of the stochastic process known as a Girsanov kernel.

Second, the misspecified expected return also generates an error in the market price of risk. The two additional drifts correct the estimation error of the market price of risk. The perturbed evolution of the state variable is given by:

\begin{align*}
\text{dA}_t &= \left( r + w(\mu - r) \right) A_t \, dt + w\sigma A_t \left( dW_1 + \lambda_1 dt \right), \\
\text{dL}_t &= a L_t \, dt + b L_t \left( \rho \left( dW_1 + \lambda_1 dt \right) + \sqrt{1 - \rho^2} \left( dW_2 + \lambda_2 dt \right) \right).
\end{align*}

The perturbation of the model is bounded by an uncertainty set S. The larger the uncertainty set S, the more pessimistic the agent is about the accuracy of the underlying model. To describe the uncertainty set, we introduce some additional notations. Let δ be the vector of the estimated drift terms from the approximate model, and let δ\textsubscript{1} be the estimation error which is a vector of the additional drift terms created from the perturbation processes λ\textsubscript{1} and λ\textsubscript{2}. Hence the true expected return is δ\textsubscript{0} = δ + δ\textsubscript{1}. Let Σ be the covariance matrix with ΓΓ\textsuperscript{′} = Σ.

\[ \delta = \begin{pmatrix} \mu \\ a \end{pmatrix}, \quad \delta_1 = \begin{pmatrix} \sigma \lambda_1 \\ b\rho \lambda_1 + b\sqrt{1 - \rho^2} \lambda_2 \end{pmatrix}, \quad \Gamma = \begin{pmatrix} \sigma & 0 \\ b\rho & b\sqrt{1 - \rho^2} \end{pmatrix}. \]

Hence we have δ\textsubscript{1} = ΓΛ where Λ = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} is the perturbation vector. Suppose we have N observations. The difference between the true expected return and the estimated expected return δ\textsubscript{0} − δ is asymptotically normal with mean zero and variance \( \frac{\Sigma}{N} \) where

\[ \Sigma = \begin{pmatrix} \sigma^2 & b\rho \sigma \\ b\rho \sigma & b^2 \end{pmatrix}. \]

We also know that \( \delta_1^\prime \left( \frac{\Sigma}{N} \right)^{-1} \delta_1 \) is a Chi-square distribution with two
degrees of freedom. Denoting the critical value at $\alpha$ significance level as $CV_\alpha$, we then have a probability of $1 - \alpha$ that

$$\delta_i' \Sigma^{-1} \delta_1 \leq \kappa^2$$  \hspace{1cm} (2.4)

where $\kappa^2 = \frac{CV_\alpha}{N}$. Equation (2.4) provides a natural boundary of the perturbation parameters. Simplifying (2.4) further, we get

$$(\Gamma \lambda)' (\Gamma \Gamma')^{-1} (\Gamma \lambda) \leq \kappa^2$$

and it becomes $\lambda \lambda \leq \kappa^2$. Hence our uncertainty set is as follows,

$$S = \{ \lambda_1, \lambda_2 | \lambda_1^2 + \lambda_2^2 \leq \kappa^2 \}$$  \hspace{1cm} (2.5)

Our uncertainty set has a circular shape in $\lambda$ space centered by zero. Hence, we can write the confidence interval of $\delta_0$ as

$$\delta_0 \in \{ \delta + \Gamma \lambda | S \}$$  \hspace{1cm} (2.6)

The true drift term $\delta_0$ is constrained by an ellipsoid uncertainty set centered by $\delta$ and it can be at any point within this set. The size of the uncertainty set depends on two factors, one is the significance level $\alpha$ and the other is the sample size $N$. If the agent has infinite observations, then the uncertainty set shrinks to the point estimate $\delta$. However, the agent only obtains limited observations which means $\delta$ moves away from $\delta_0$. The fewer the observations, the higher $\kappa^2$ will be and the more parameter uncertainty the agent is exposed to.

Our stylized uncertainty set is related to the Good Deal Bounds rational proposed by Cochrane and Saa-Requejo (1996). The idea here is to add a confidence interval surrounding the observable market price of risk, and constrain the total market of risk to a reasonable value. Equation (2.4) could also be understood as the Good Deal Bounds constraint in which we only put a boundary on the unobservable part of the market price of risks, $\lambda_1$ and $\lambda_2$. If we completely follow the Good Deal Bounds method, our uncertainty set should be as follows $(\delta_0 - \iota r)' \left( \frac{\Sigma}{T} \right)^{-1} (\delta_0 - \iota r) \leq G$ where $\iota = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $G$ is the value of the Good Deal Bounds. Denote $\bar{\delta} = \delta - \iota r$ the demeaned point estimate. Then the Good Deal Bounds condition becomes $(\bar{\delta} + \delta_1)' \Sigma^{-1} (\bar{\delta} + \delta_1) \leq \frac{G}{T}$. The function can be further simplified to $\bar{\delta} \Sigma^{-1} \bar{\delta} + 2 \delta_1' \Sigma^{-1} \delta_1 + \lambda_1^2 + \lambda_2^2 \leq \frac{G}{T}$.

The uncertainty set we propose differs from the Good Deal Bounds in the way in which our uncertainty set is derived from the distribution theory. The uncertainty set

23
parameter $\kappa$ depends only on the statistical quantities, $\alpha$ and $T$. However, the Good Deal Bounds method is inspired by an economic belief that the total market price of risk in an incomplete market has to have limits.

2.3.3 Robust Optimization Problem

Utility is defined at time $t$ as a function of $A_t$ and $L_t$. The optimal hedging strategy maximizes the utility function $U(A_t, L_t, t)$. The optimization problem is given by

$$\max_w U(A_t, L_t, t)$$

(2.7)

We call the hedging strategy that does not consider model misspecification a naive policy, denoted $w_{na}$. The agent is completely confident about the estimated model and his hedging criterion is to maximize the utility function by making an instantaneous investment decision $w_{na}$.

If the agent is afraid of model misspecification, then how to deal with it? We assume that the uncertainty averse agent divorces from the naive framework by having a set of models which are all likely to be the true model to represent the dynamics of the asset and the liability. The agent looks for a hedging policy that works well over a set of models that may prevail. This policy is later on call a robust policy. To obtain the robust hedging policy, the agent employs a max-min expected utility framework which is defined as follows

$$\max_{w} \min_{\lambda_1, \lambda_2} U(A_t, L_t, t)$$

subject to state variable evolutions (2.3). The control variables $\lambda_1$ and $\lambda_2$ are subject to the uncertainty set $S$. This is a robust control problem. The max-min optimization problem is according to Anderson et al. (2003) a two-player zero-sum game. This is a sequential game between the decision maker and a malevolent nature. Player 1, or max-player (the robust agent) moves first by choosing instantaneous investment decisions to maximize the utility function at time $t$ a set of models, and then player 2, or the min-player (the imaginary nature) picks the worst state of nature for player 1 by making an instantaneous choice of $\lambda_1$ and $\lambda_2$, given player 1’s choice. Given the decision of player 2, player 1 can define her optimal portfolio at time $t$ denoting $w_{rob}$. In other words, the agent is maximizing while the nature is choosing the worst possible model. We call the optimal portfolio choice $w_{rob}$ from (2.8) a robust decision. Of course, we do not really believe in the evil nature. It is simply an artificial agent that help the robust agent to construct bounds of the value function that allow us to analyze the fragility of the decision
rules over a set of possible models.

## 2.4 Static Robust Optimization

Given the information at time $t$, our hedging strategy is defined over the hedging error $L_T - A_T$ at a predetermined time $T$. Our utility function takes the form of the shortfall risk $- [L_T - A_T]^+$, which specifies the downside risk on the liability shortfall. The lower the shortfall risk, the higher the agent’s utility will be. Hence, we employ an optimal hedging strategy that minimizes the expected shortfall at time $T$. The naive optimization problem is given by

$$
\min_w E[(L_T - A_T)^+ | \mathcal{F}_t] \quad (2.9)
$$

and the robust optimization is

$$
\min_w \max_{\lambda_1, \lambda_2} E[(L_T - A_T)^+ | \mathcal{F}_t] \quad (2.10)
$$

Note that, in this case the min-player (the robust agent) moves first by making a hedging decision to minimize the expected shortfall at time $T$ over a set of models. Then the max-player (the artificial nature) choose the worst state given the min-player’s choice. The reason that the economic agent is no longer a max-player as introduced in (2.8) is that the objective of the agent in this particular case is to minimize a function of the downside risk.

It is also interesting to notice that, the order of the two players is interchangeable in the static case. According to the saddle-point existence Theorem mentioned in Delbaen (2002) and Rockafellar (1976), the optimal solution of (2.10) is a saddle point, since both control variables are constrained by a convex set and the value function is bilinear. Hence (2.10) and its dual problem $\max_{\lambda_1, \lambda_2} \min_w E[(L_T - A_T)^+ | \mathcal{F}_t]$ has the same optimal solution.

In the following two sections, we will show how to solve the robust optimization problem and how the robust solution differs from the naive one, and also how we can benefit from the robust decision. We start with the relatively simple static case, where both agent and nature only make decisions now at $t = 0$ without rebalancing until the expiration period $T$. The static case is technically easy to solve, but still provides us with some intuition about the robust policy. However, the static solution is not optimal, if the agent is able to rebalance before $T$. We therefore also provide a dynamic solution in the next section in which $\lambda_1$ and $\lambda_2$ are time series processes.
2.4.1 Static Solution

If the control variables \( w, \lambda_1 \) and \( \lambda_2 \) are static, our criterion function \( E[(L_T - A_T)^+] \) is very similar to the value of an “exchange option” (see e.g. Margrabe (1978)) which changes one asset for another at time \( T \). We can also consider this payoff function as a European option with a floating strike price. Margrabe (1978) provides a formula for valuing this type of option. The problem in our case is more complicated, because we are in an incomplete market, which means the equivalent martingale is not unique, or in other words, the so-called risk-neutral \( Q \) measure is not unique, but depending on \( \lambda_1 \) and \( \lambda_2 \).

To facilitate calculation, let

\[
\begin{align*}
\mu_S &= \mu + \sigma \lambda_1, \\
\mu_A &= r + w(\mu_s - r), \\
\mu_L &= a + b \rho \lambda_1 + b \sqrt{1 - \rho^2} \lambda_2,
\end{align*}
\]

representing the drift terms of the stock market, the asset and the liability respectively.

There are many ways to solve this static criterion function. We use the change of probability measure technique. By multiplying and dividing \( E[L_T] \) inside the valuation function, we can create a Radon-Nikodym process \( \frac{L_T}{E[L_T]} \) that changes the probability measure from \( P \) to a new measure called \( L \). That is to say \( \frac{L_T}{E[L_T]} \) is a strictly positive random variable with expected value one, \( E\left( \frac{L_T}{E[L_T]} \right) = 1 \), hence the expected shortfall function can be rewritten as,

\[
E \left[ (L_T - A_T)^+ \right] = E[L_T] E \left[ \frac{L_T}{E[L_T]} \left( 1 - \frac{A_T}{L_T} \right)^+ \right] = E[L_T] E^k \left[ (1 - C_T)^+ \right],
\]

We denote the coverage ratio (also known as funding ratio) at time \( t \) as \( C_t = \frac{A_t}{L_t} \). It is a common criterion used to describe the performance of a financial institution. If the coverage ratio is less than one, the fund is facing a solvency risk. Hence, we reconstruct our “exchange option” as a product of the expected value of \( L_T \) under \( P \) and the value of a European put option under measure \( L \). The first term \( E[L_T] \) is equal to \( L_0 \exp(\mu_L T) \), and the second term is known from Margrabe (1978).

Applying Ito’s lemma on \( dC_t \), we can derive the diffusion process of \( C_t \). We show the calculation in Appendix 2.7.1. We also show how to obtain the process \( W^1_T \) and \( W^2_T \) under the new probability measure. The diffusion process of the coverage ratio \( C_t \) under
the new measure $L$ is given by

$$dC_t = C_t \left[ (-\mu_L + \mu_A) dt + (w\sigma - b\rho) dW_t^L - b\sqrt{1 - \rho^2} dW_t^L \right]$$  \hspace{1cm} (2.13)$$

Note that $dC_t$ process under $L$ has a drift term, containing $\mu_A$ and $\mu_L$ only. The variance of coverage ratio is $\sigma_C^2 = (w\sigma - b\rho)^2 + b^2 (1 - \rho^2)$. Therefore, the analytical solution of our objective function under the static case is given by $\bar{L} \left[ \Phi(-d_2) - \bar{C} \Phi(-d_1) \right]$, or

$$\mathbb{E} \left[ (L_T - A_T)^+ \right] = \bar{L} \Phi(-d_2) - \bar{A} \Phi(-d_1) = \bar{L} (\Phi(-d_2) - \bar{C} \Phi(-d_1))$$  \hspace{1cm} (2.14)$$

where

$$\bar{L} = L_0 \exp(\mu_LT)$$
$$\bar{A} = A_0 \exp(\mu_AT)$$
$$\bar{C} = C_0 \exp[(\mu_A - \mu_L)T]$$

and

$$d_1 = \frac{\ln \bar{C} + \sigma_C^2 T}{\sigma_C \sqrt{T}}$$
$$d_2 = d_1 - \sigma_C \sqrt{T}$$

The function $\Phi$ is the cumulative probability distribution function for a standard normal distribution. Therefore, for given $\lambda_1$ and $\lambda_2$, the optional hedge disregarding the preference for robustness is the solution of the first order condition for maximizing $\mathbb{E}^L \left[ (1 - C_T)^+ \right]$ with respect to $w$ (see Appendix (2.7.2)),

$$\frac{\partial}{\partial w} [\Phi(-d_2) - \bar{C} \Phi(-d_1)] = -\Phi(-d_1) \bar{C} (\mu - r) T + \bar{C} \phi(d_1) \sqrt{T} \frac{w\sigma^2 - b\rho\sigma}{\sigma_C} = 0$$  \hspace{1cm} (2.15)$$

where function $\phi$ denotes the standard normal probability density function. Note that $-\Phi(-d_1)$ is the delta of the BS put-option which is always less than zero, and $\bar{C} \phi(d_1) \sqrt{T}$ denotes the vega of the BS option which is always positive. Therefore, we see from (2.15) that the optimal $w$ strikes a balance between the “delta effect” that reduces the value of the option and the “vega effect” that increases the value of the option. There is a special case when $\mu = r$ where the “delta-effect” disappears and the optimal $w$ is then given by the minimum variance solution $w = \frac{b\rho}{\sigma}$. 

27
2.4.2 Static Robust Portfolio Choice

In this subsection, we consider the static robust optimization problem. We can rewrite the objective function (2.9) as

\[
\min_w \max_{\lambda_1, \lambda_2} \lambda_1 L_0 \exp(\mu_A T) \Phi(-d_2) - A_0 \exp(\mu_A T) \Phi(-d_1) = 0
\]

subject to \(\lambda_1 + \lambda_2 \leq \kappa^2\). Notice that \(\mu_A\) and \(\mu_L\) are functions of \(\lambda\).

This problem can only be solved numerically. As a benchmark scenario, we assume \(\mu = 0.04, \sigma = 0.16, r = 0, a = 0, b = 0.1, \rho = 0.5, \kappa = 0.25\). We assume that the stock return \(\mu\) is higher than the liability return \(a\) because the stock volatility \(\sigma\) is normally higher than the volatility of the liability \(b\).

As we have discussed in Section 2.3.2 the uncertainty set parameter \(\kappa\) depends on the significance level \(\alpha\) and the sample size \(N\). Hence it is fixed and state variable independent. We choose the significance level \(\alpha = 0.05\) at which the corresponding \(\chi^2\) value with 2 degrees of freedom is 5.99. The choice of \(\kappa\) is also based on an implicit assumption that the risk premium \(\mu_S - r\sigma\) is always positive, which means \(\kappa \leq \frac{\mu_S - r\sigma}{\sigma} = 0.25\). Hence, the sample size \(N\) has to be larger than 96 so as to satisfy this implicit assumption.

The robust hedge provides the optimal solution under the worst-case scenario. Under the benchmark scenario, the worst-case scenario occurs when the true expected asset return is lower than the estimated value, \(\mu_A < r + w (\mu - r)\), and the true liability return is higher than the estimated result \(\mu_L > a\). In this case, the true hedging error will be much higher than expected. A negative \(\lambda_1\) and a positive \(\lambda_2\) lead to the worst-case scenario. To see this, we first look at the asset side. If \(\lambda_1\) is negative, then the expected asset return is reduced. But a negative \(\lambda_1\) also reduces the liability return, since \(\lambda_1\) also plays a role on the liability side except when \(w\sigma - b\rho = 0\). In this case, the effect of \(\lambda_1\) on the liability side cancels out the effect of \(\lambda_1\) on the asset side completely. In this particular case, the decision of \(\lambda_2\) is independent of the choice of \(\lambda_1\). Besides, the effect of \(\lambda_2\) is partially canceled and it has to be sufficiently big to beat the negative effect of \(\lambda_1\) and push up the liability return. In short, the agent is afraid of a negative \(\lambda_1\) and a positive \(\lambda_2\).

In Figure 2.1, we show the static optimal portfolio choice at time \(t = 0\) as the function of the current funding ratio, \(C_0\). The solid-dot curve depicts the robust portfolio decision and the open-dot curve shows the naive optimal portfolio choice with \(\lambda_1 = \lambda_2 = 0\). When there is underfunding, which means \(C_0 < 1\), the robust and naive policies differs. Both take substantial risk, but the robust portfolio is more conservative than the naive one. For example, if the current funding ratio equals 80%, then the robust policy will reduce the risky asset exposure by approximately 6% based on the naive policy. A naive investor
2.4. STATIC ROBUST OPTIMIZATION

Figure 2.1: Static optimal portfolio choice. This figure compares the robust and naive static optimal hedging policies. The investor makes an investment decision at time $t = 0$ with given current funding ratio $C_0$ so as to minimize the expected shortfall at time period $T$. The naive policy relies completely on the estimation parameters. The robust policy takes the parameter uncertainty into consideration and insures against the worst-case scenario. The horizontal axis depicts the present funding ratio. The results are based on the benchmark estimation parameters $\mu = 0.04$, $\sigma = 0.16$, $r = 0$, $a = 0$, $b = 0.1$, $\rho = 0.5$, $\kappa = 0.25$, and $T = 5$.

relies completely on the underlying model and believes that there is a positive expected asset return of 4%. However, the robust investor is not so sure about the asset return and is afraid that the true expected asset return is not as high as the model estimated. Therefore the robust investor is more conservative with his portfolio choice.

Robust hedges are not always more conservative than naive hedges. If the fund is already balanced - or even overfunded with $C_0 \geq 1$ - the two policies are almost identical, which means the robustness effect diminishes if $C_0$ goes up. The two curves converge to a hedging ratio $\frac{b\rho}{\sigma} = 0.3125$ if $C_0$ is sufficiently large. The ratio $\frac{b\rho}{\sigma}$ is the result of a minimum-variance hedge that neutralizes the traded part of the liability risk. The resulting volatility becomes $b(1 - \rho^2)$, which is the unhedgeable part of the liability risk. Also, this position neutralizes the $\lambda_1$ effect such that the misspecification of asset return does not influence the performance of hedges.

The decision of nature is displayed in Figure 2.2. We show the movement of $\lambda_1$ and

\[ \sigma_C^2 = (w\sigma - b\rho)^2 + b^2(1 - \rho^2), \]  
with respect to $w$, the optimal hedging ratio $w$ equals to $\frac{b\rho}{\sigma}$.

---

5To minimize the volatility of the hedging error, $\sigma_C^2 = (w\sigma - b\rho)^2 + b^2(1 - \rho^2)$, with respect to $w$, the optimal hedging ratio $w$ equals to $\frac{b\rho}{\sigma}$. 29
CHAPTER 2. ROBUST HEDGING IN INCOMPLETE MARKETS

Figure 2.2: Static optimal perturbations $\lambda_1$ and $\lambda_2$. This figure depicts the optimal $\lambda_1$ and $\lambda_2$ as functions of the present coverage ratio $C_0$ under the benchmark scenario with $\mu = 0.04$, $\sigma = 0.16$, $r = 0$, $a = 0$, $b = 0.1$, $\rho = 0.5$, $\kappa = 0.25$, and $T = 5$. Nature makes decisions of $\lambda_1$ and $\lambda_2$ at time 0 under the constraint $\lambda_1^2 + \lambda_2^2 \leq \kappa^2$ so as to maximize the expected shortfall at period $T$.

$\lambda_2$ as a function of the present coverage ratio $C_0$. To facilitate the comparison, we put the two perturbations in one graph. With given $C_0$, each combination of $\lambda_1$ and $\lambda_2$ leads to the biggest perturbations between models. We find that $\lambda_1$ is always below the zero line given any coverage ratio level and is close to zero when $C_0$ is high, but $\lambda_2$ is always positive and converges to $\kappa$. We also find that the optimal choice of $\lambda_1$ and $\lambda_2$ is always on the circle $\lambda_1^2 + \lambda_2^2 = \kappa^2$, which means the worst-case scenario is always at the boundary of the uncertainty set.

The resulting negative $\lambda_1$ represents the fear of an over-estimated asset return. Hence, the absolute value of $\lambda_1$ is increasing with the exposure to the stock market, $w$. We know from Figure 2.1 that risk exposure and the coverage ratio are negatively related. The lower the coverage ratio is, the higher the risk exposure will be and therefore the more negative the value of $\lambda_1$ will be. However, if the coverage ratio is sufficiently high, the investor will put less wealth into the risky asset, as the penalty from $\lambda_1$ is smaller.

However, the penalty term $\lambda_1$ also plays a role in the liability return. A negative $\lambda_1$ can benefit the agent by reducing the expected liability return. However, the agent is afraid of an under-estimated liability return. To capitalize on the fear of an increase in the liability return, Nature chooses a positive $\lambda_2$ so as to compensate for the negative
2.4. STATIC ROBUST OPTIMIZATION

(a) Robust vs. Naive

(b) Robust vs. Naive

Figure 2.3: Mean rate of stock and liability return with and without the preference for robustness. This figure displays the expected stock and liability returns before and after considering parameter uncertainty as functions of the present coverage ratio. Panel 2.3a: comparing the robust stock drift $\mu_S = \mu + \sigma \lambda_1$ with the naive drift term $\mu_S = \mu$. Panel 2.3b: comparing the robust liability drift term $\mu_L = a + b \rho \lambda_1 + b \sqrt{1 - \rho^2} \lambda_2$ with the naive one $\mu_L = a$ under the benchmark scenario.

We further examine how the perturbation terms impact the expected returns. Figure 2.3 displays both the naive and robust mean rate of the stock return and the liability return as functions of $C_0$. Without the preference for robustness, both drift terms are constant at the benchmark level irrespective of the value of $C_0$. However, if the investor is aware of the model misspecification and aims to insure against the worst-case scenario, the perturbed expected stock return is dragged down by $|\sigma \lambda_1|$ and the worst-case liability drift is pushed up by $|b \rho \lambda_1 + b \sqrt{1 - \rho^2} \lambda_2|$.

The robust policy differs from the naive policy in the way in which the agent introduces an additional guarantee on top of the naive contract in order to neutralize the estimation error. This additional insurance makes the robust policy more expensive. In Appendix 2.4.3, we show the cost of hedging using the two different policies.

2.4.3 Cost of Preference for Robustness

In Figure 2.4, we show the additional expense required from the robust policy. The upper panel shows the expected shortfall of the two policies at period $T$ as a function of the present coverage ratio $C_0$. The solid-dot curve is the robust result and the dotted curve is the naive result. Remark, the two curves are not comparable. It is trivial that the robust
Figure 2.4: Cost of Preference for Robustness. The upper panel plots the expected shortfall at period $T$ as a function of the present coverage ratio by following different investment policies under the benchmark scenario. The solid-dotted line is the worst case scenario with robust optimal $w$, $\lambda_1$, and $\lambda_2$. The empty dotted curve is the expected shortfall following naive investment policy without perturbations. The bottom panel plots the gap of the two curves. It measures the cost of the preference for robustness.

Policy is more expensive because the agent buys an additional insurance to neutralize himself from the model misspecification. The bottom panel measures the cost of adding the preference for robustness on the hedging contract which is the gap between the two curves.

We assume the present liability value is $L_0 = 100$, if coverage ratio is 80% which means the current value of asset is $A_0 = 80$ and the current mismatch of the fund is $L_0 - A_0 = 20$, then after $T$ periods the naive expected shortfall is 13 but the worst case scenario ends up with 23, which means the agent has to pay additional 10 to guarantee himself against the estimation error. The gap between the two curves captures the robust agent’s pessimism in two directions: the penalty cost from an over estimated asset return and the penalty cost from an under estimated liability drift.

2.4.4 Policy Evaluation

The agent is motivated to follow the robust policy because it is less sensitive to the estimation error. In this section, we will show how and when the agent can benefit from
the robust policy. Let $Q(w, \delta)$ be the cost of hedging following a certain policy $w$, where $\delta$ is the assumed value of the drift parameters. In our case, the cost of hedging is defined by

$$Q(w, \delta) = \mathbb{E}_t [(L_T - A_T)^+ | w, \delta]$$

(2.17)

The optimal hedging policy has a cost $q(\delta) = \min_w Q(w, \delta)$ for given $\delta$. Let $\delta_0$ be the true value of $\delta$, with $\delta_0 = \left(\begin{array}{c} \mu_0 \\ a_0 \end{array}\right)$ and denote $q_0$ as the minimum hedging cost when the investor implements the associated optimal hedging policy $w_0$ under the true value $\delta_0$. Any other alternative hedging policies $w_a (w_a \neq w_0)$ have higher expected shortfall.

Defining the loss function $K(w_a | \delta_0)$ as the difference between the cost of hedging following a suboptimal policy $Q(w_a, \delta_0)$ and the true minimum cost,

$$K(w_a | \delta_0) = Q(w_a, \delta_0) - q_0$$

(2.18)

In case $\delta$ is misspecified with $\delta \neq \delta_0$, the agent is facing estimation error, therefore $w_a \neq w_0$ and $K(w_a | \delta_0) > 0$.

The agent does not know the true value of the drift terms $\delta_0$. Given the estimated drift terms $\delta$, he can choose between two alternative hedging policies, a robust policy $w_\text{rob}$ and a naive policy $w_\text{na}$. At the benchmark scenario when the present coverage ratio $C_0 = 80\%$, the solutions are $w_\text{rob} = 0.81$ and $w_\text{na} = 0.87$. When $C_0 = 90\%$, we find that $w_\text{rob} = 0.67$ and $w_\text{na} = 0.69$. Next, we will investigate in which circumstances the robust policy will perform better than the naive policy.

$$K(w_\text{rob} | \delta_0) < K(w_\text{na} | \delta_0)$$

(2.19)

We display the loss indifference curves in Figure 2.5 when the present coverage ratio is 80\% and 90\%. The x-axis and y-axis represent the true value of liability return $a_0$ and asset return $\mu_0$ respectively. The $[\mu = 0, \mu = 0.04]$ spot represents the estimated expected return $\delta$. We also display the ellipsoid uncertainty set of the true drift term $\delta_0$ in the figure. The area outside the ellipse is assumed to be improbable and irrelevant for decision making.

On the policy indifference curve, when $K(w_\text{rob} | \delta_0) = K(w_\text{na} | \delta_0)$, the two policies require the same amount of wealth to hedge. In the region below the indifference curve for both scenarios (when $C_0 = 80\%$ and 90\%), the robust policy requires less initial wealth.

\textsuperscript{6}“Cost of hedging” here is defined as the initial wealth required to obtain a particular level of expected shortfall.
CHAPTER 2. ROBUST HEDGING IN INCOMPLETE MARKETS

Figure 2.5: Loss function equivalent curves. The figure plots the indifference curve of the loss when $K(w_{rob} | \delta_0) = K(w_{na} | \delta_0)$. The $y$-axis is the true value of the expected stock return $\mu_0$ and the $x$-axis is the true value of the liability drift $a_0$. The estimated value is $\mu = 0.04$ and $a = 0$. The solid-dot indifference curve represents the case when $C_0 = 80\%$ and the open-dot curve is the when $C_0 = 90\%$. In the region below the curve, the robust policy outperforms the naive policy and in the region above, it is the other way around.

Then the naive policy to hedge a certain amount of expected shortfall. The value of $\delta_0$ in this region is lower than the estimated value $\delta$. Hence we can conclude that when the true drift term $\delta_0$ is over-estimated, the robust policy performs better.

We also find that this beneficial region is positively related to the present coverage ratio $C_0$. Figure 2.4 shows that the additional cost of hedging using the robust policy is increasing with $C_0$. Therefore a lower $C_0$ leads to a smaller beneficial region.

2.4.5 Sensitivity Analysis

The correlation parameter $\rho$, representing the completeness of the market, plays an important role in the model. If $\rho = \pm 1$, and $\lambda_1 = \lambda_2 = 0$, then the market becomes complete and the unhedgeable risk driver $W_2$ does not play a role. In this section, we investigate how sensitive the optimal hedges are with respect to a change of $\rho$.

In Figure 2.6, we show an extreme case when $\rho = 1$. The non-traded risk driver $W_2$ disappears from the liability diffusion process and the perturbation parameter $\lambda_2$ does
2.4. STATIC ROBUST OPTIMIZATION

Figure 2.6: Sensitivity analysis with $\rho = 1$. The figure depicts the optimal portfolio choice when $\rho = 1$. The remaining parameters stay at the benchmark level. The solid-dot line represents the robust policy and the empty dotted curve is the naive policy. The naive agent considers such an economy a complete market if $\lambda_1 = \lambda_2 = 0$, since the non-tradable risk driver $W_2$ is gone. However, the robust agent still stays in the incomplete market, because the model misspecification ($\lambda_1 \neq 0, \lambda_2 \neq 0$) is also considered, as another source of market incompleteness.

not play a role either. Nature can only control $\lambda_1$ to maximize the expected shortfall at period $T$. The naive agent considers this as a complete market. However, the robust agent still faces another source of incompleteness, caused by model misspecification.

If there is insufficient wealth in the fund, the robust policy deviates from the naive one much more severely compared with the benchmark case. When the asset risk and the liability risk are perfectly correlated, nature will choose a more negative $\lambda_1$ so as to maximize the expected shortfall. Although a negative $\lambda_1$ reduces the expected liability return as well, the liability drift term is less sensitive to the change of $\lambda_1$ than the expected asset return, since $\sigma > b$. As the result the robust investor’s fear of an over-estimated asset return is stronger than the benchmark level.

In the case of overfunding, the two policies are identical. The hedging error volatility becomes $\sigma^2 = (w\sigma - bp)^2$. The investor can fully replicate the liability by following a Delta-neutral strategy $w = \frac{b}{\sigma} = 62\%$ if she has sufficient assets. In that case robustness does not play a role because the Delta hedge neutralizes the $\lambda_1$ effect.

In Figure 2.7, we show the two hedging policies as a function of correlation parameter $\rho$. We display two scenarios, one when $C_0 = 80\%$ and the other when $C_0 = 90\%$. The
Figure 2.7: Sensitivity analysis with respect to $\rho$. The figure plots the optimal naive and robust hedging policies as a function of correlation parameter $\rho$. We show two pairs of comparison. The red pair with cube-dot is when the present coverage ratio $C_0$ is 80%, and the blue pair with circle-dot is when $C_0 = 90%$. The solid-dot curves represent the robust policy and the empty-dot curves are the naive policy.

The relation between the optimal portfolios and $\rho$ is not monotone but is hump shaped. This is because the volatility of the value function $\sigma_C$ is a quadratic function of $\rho$.

The optimal portfolio initially increasing with $\rho$ for either policies because the liability market is more exposed to the tradable risk driver $W_1$. Therefore the risky portfolio has to increase as well, in order to hedge the traded liability risk. Next, the optimal portfolio reaches the peak where $\rho$ maximizes the total volatility $\sigma_C$. After the peak, the risky portfolio goes down with $\rho$, because after the peak, any higher level of correlation will reduce $\sigma_C$. From Figure 2.7, we can also see that the difference between the two policies under the lower coverage ratio is wider than under the higher $C_0$.

### 2.5 Dynamic Robust Optimization

In this section, we will extend the problem to a dynamic world. The robust investor still aims to minimize the final-period expected shortfall under the worst case scenario, but instead of making a static portfolio choice, he is now considering a dynamic optimal portfolio. Nature also needs to rebalance her choice of $\lambda_1, \lambda_2$ instantaneously given
2.5. Dynamic Robust Optimization

the intertemporal decision of \(w\). We employ dynamic programming to solve this robust optimization problem. The structure of this section is as follows: we first formulate the dynamic robust optimization problem and discuss the analytical solution at the extreme case. Next, we employ numerical methods to solve the partial differential equation. Lastly, we will investigate the dynamic effect on the policy indifference curve analyzed in Section 2.4.4.

2.5.1 Dynamic Programming

Let the value function at time \(t\) be \(U(A, L, t)\). Then using Feynman-Kač we can derive the Hamilton-Jacobi-Bellman equation (henceforth HJB) or partial differential equation (pde) for the investor’s min-max problem:

\[
0 = \min_w \max_{\lambda_1, \lambda_2} \left[ U_t + U_{AA} (r + w(\mu - r) + w\sigma \lambda_1) + U_{LL} \left( a + b \rho \lambda_1 + b \sqrt{1 - \rho^2} \lambda_2 \right) + \frac{1}{2} U_{AA} \sigma^2 A^2 + \frac{1}{2} U_{LL} b^2 L^2 + U_{AL} \rho \sigma \lambda_1 \right] \tag{2.20}
\]

with the boundary condition: \(\lambda_1^2 + \lambda_2^2 \leq \kappa^2\). We employ the method of Lagrange by introducing the multiplier \(\nu\) and forming the Lagrangian function:

\[
0 = \min_w \max_{\lambda_1, \lambda_2} \left[ U_t + U_{AA} (r + w(\mu - r) + w\sigma \lambda_1) + U_{LL} \left( a + b \rho \lambda_1 + b \sqrt{1 - \rho^2} \lambda_2 \right) + \frac{1}{2} U_{AA} \sigma^2 A^2 + \frac{1}{2} U_{LL} b^2 L^2 + U_{AL} \rho \sigma \lambda_1 - \frac{1}{2} \nu (\lambda_1^2 + \lambda_2^2 - \kappa^2) \right] \tag{2.21}
\]

By solving a linear system of equations based on the first order condition of (2.21) with respect to \(w\), \(\lambda_1\) and \(\lambda_2\), we have

\[
w^* = -\frac{(\mu - r) U_{AA} \nu}{\sigma^2 (U_{AA} A^2 + U_{AA} A^2)} - \frac{U_{AL} b \rho \sigma \nu + U_{LL} U_{AA} A^2}{\sigma^2 (U_{AA} A^2 + U_{AA} A^2)} \tag{2.22a}
\]

\[
\lambda_1^* = -\frac{(\mu - r) U_{AA} \sigma^2}{\sigma^2 (U_{AA} A^2 + U_{AA} A^2)} - \frac{b \rho (U_{LL} + U_{AA} A^2)}{(U_{AA} A^2 + U_{AA} A^2)} \tag{2.22b}
\]

\[
\lambda_2^* = \frac{U_{LL} b \sqrt{1 - \rho^2}}{\nu} \tag{2.22c}
\]

Note: this is a partial solution. The Lagrange multiplier \(\nu\) still needs to be solved, but numerically.

The sign of optimal \(\lambda_2\) must be positive since it increases the expected liability return but does not influence the pension asset. The sign of \(\lambda_1\) is not however defined. A positive \(\lambda_1\) not only increases the liability but also the asset, but the net effect depends on the value of other input variables.
CHAPTER 2. ROBUST HEDGING IN INCOMPLETE MARKETS

The solution (2.22) has a very interesting structure. The investor’s dynamic portfolio choice \( w \) and lifetime discrepancy between the reference model and perturbations, \( \lambda \)'s, are dueling against each other. The dynamic optimal investment strategy \( w^* \) is a tradeoff between hedging and speculation. We can see this by considering the extreme case when \( \nu \to 0 \) and \( \nu \to \infty \).

**When \( \nu \) is 0.** For \( \nu \to 0 \), the discrepancy parameters \( \lambda \)'s have more freedom to choose an arbitrarily large aversion pair of drift for the Brownian Motions, or in other words, the agent is extremely pessimistic about the approximation model. When \( \nu \to 0 \), we have

\[
w_{\nu \to 0} = - \frac{U_L b \rho}{U_A \sigma},
\]

This is a pure hedging portfolio, where the agent invests an amount in risky assets such that the change in value \( U(A, L, t) \) due to \( L \) is (as much as possible) offset by a change in value due to \( A \). It is not possible to completely eliminate the volatility of \( L \). This is because the liabilities are exposed both to hedgeable risk \( W_1 \) and unhedgeable risk \( W_2 \), but only the hedgeable part \( W_1 \) can be eliminated.

The optimal value for \( \lambda^*_1 \) when \( \nu \to 0 \) is given by

\[
\lambda^*_{1, \nu \to 0} = -\frac{\mu - r}{\sigma} - \frac{b \rho (U_{AL} A - U_{LA} U_{AA}) L}{U_A^2}
\]

which contains two terms. The first term is the observable market-price of risk which we can see from the Black-Scholes setup. The second term in more interesting. Note that

\[
- \frac{b \rho (U_{AL} U_A - U_{LA} U_{AA}) L}{U_A^2} = \sigma \left( \frac{w^*_L}{\partial A} \right) = w^*_{\nu \to 0} \sigma + \sigma A \left( \frac{U_{AL}}{U_{AA}} \right),
\]

this reflects to what extent the agent’s best possible hedging strategy is influenced by the instantaneous wealth level \( A_t \).

**When \( \nu \) is infinity.** At the other extreme, when we consider the case \( \nu \to \infty \), then both \( \lambda_1 \) and \( \lambda_2 \) shrink to zero, so \( \kappa = 0 \). This corresponds to the case when the agent faces no model misspecification. Hence we recover the “classical” Merton’s solution for the optimal portfolio choice:

\[
w^*_{\nu \to \infty} = - \frac{\mu - r}{\sigma^2} \frac{U_A}{U_{AA} A} - \frac{U_{AL} L b \rho}{U_{AA} A \sigma}
\]

The first term is a speculative portfolio where the agent invests in the stock market to obtain the optimal trade-off between the observable market price of risk \( \frac{w^*_L}{\sigma} \) and the local risk aversion \( \frac{U_{AL}}{U_{AA}} \). The second term is the intertemporal hedging component, but the optimal amount to hedge is now measured in terms of the “CAPM-beta”. That is, the optimal hedge is the local covariance term \( b \rho \sigma \) divided by local variance term \( \sigma^2 \), i.e.
2.5. DYNAMIC ROBUST OPTIMIZATION

Figure 2.8: Dynamic robust optimal hedging strategy. This figure displays the robust optimal investment policy as a function of time and the instantaneous coverage ratio $C_t$ with benchmark input parameters. Panel 2.8a plots the robust portfolio choice as a function of the instantaneous coverage ratio and the time movement. Panel 2.8b abstracts one dimension from Panel 2.8a and plots the robust solution as a function of the instantaneous coverage ratio $C_t$ through period 0 to period $T$; therefore it is in three dimensions. Panel 2.8b only depicts the solutions at period $t = 0, 3, 5$. Due to technical limitations, our grid searching interval for the risky portfolio $w$ has to be smaller than 1.95, otherwise we will confront a negative probability problem in some trinomial trees.

the stock market investment that minimizes locally the (unhedgeable) variance in the portfolio.

2.5.2 Numerical Solution

Due to the complexity of our problem, we cannot solve the PDE analytically. We employ an explicit finite difference method to solve the PDE. In Appendix 2.7.3, we elaborate on the numerical procedure for solving our dynamic programming problem. In this section, we will show the numerical result of the dynamic optimization problem.

In Figure 2.8, we show the dynamic robust investment policy as a function of the instantaneous coverage ratio $C_t$ and time. We can conclude from the figure that on the one hand, if the coverage ratio is continuously low, the investor will increase the risk exposure over time. The reason is that given the poor performance of the fund, the investor is afraid of an even worse funding ratio in the subsequent period. We have learnt from the static case that the optimal risk exposure is supposed to be high when the coverage ratio is low. On the other hand, if the liability payoff is already fully covered with $C_t > 1$, the investor will reduce his risk exposure over time, and the optimal portfolio converges faster to the hedging ratio Delta ($\Delta \lambda$) when $t$ is approaching expiration.

Next, we would like to investigate the difference between the robust and the naive policy in the dynamic version. In Figure 2.9, we show the two investment policies as a
function of the instantaneous coverage ratio at two selected time periods, $t = 0$ and $t = 3$. We highlight two findings from the figure. First, the robust policy is always less risky than the naive one as long as the instantaneous coverage ratio is lower than 1. Second, we find that the difference between the two policies increases over time if the instantaneous coverage ratio is low. The growing difference between the two policies is caused by the accumulated $\lambda_1$ effect.

Next, we investigate the dynamic optimal perturbation process. Figure 2.10 shows the dynamic optimal $\lambda_1$, $\lambda_2$ as functions of the coverage ratio at three different time periods. It is no longer fresh that $\lambda_1$ is always negative and $\lambda_2$ is always positive. We now focus on the dynamic effect of the processes.

We first look at the low coverage ratio region ($C_t < 1$). If $C_t$ is low, $\lambda_1$ is decreasing over time. We provide two intuitive explanations of this finding. First, if a fund is continually underperforming, the agent would become more and more pessimistic about the underlying model and more likely to believe that the true expected asset return could be lower, hence the negative $\lambda_1$ effect is growing. Second, the agent is expecting a decreasing instantaneous coverage ratio, hence her risky portfolio increases over time. As a result, the exposure to the estimation error is increasing over time as well. An increasing $\lambda_1$ effect is always accompanied by a decreasing $\lambda_2$ effect due to our specific uncertainty set design.

If the instantaneous coverage ratio is high, the agent is approaching a Delta hedge in order to neutralize the $\lambda_1$ effect. Therefore, $\lambda_1$’s negative effect is diminishing over time as $t \to T$. By contrast, the $\lambda_2$ effect is growing over time and converging quickly to $\kappa$, such that nature can maximize the shortfall risk.
2.5. Dynamic Robust Optimization

In this figure we show the optimal perturbation processes $\lambda_1$ and $\lambda_2$ as functions of the instantaneous coverage ratio at time $t = 0, 3, 5$. The solid lines are the movement of $\lambda_1$ and $\lambda_2$ at time $t = 0$, the dashed curves are at time $t = 3$ and the dotted curves are at time $t = 5$. The upper panel with positive perturbations gives the optimal results of $\lambda_2$. The negative portion of the figure gives the optimal solutions of $\lambda_1$.

Now let us look at the movement of dynamic perturbed drift terms displayed in Figure 2.11. Panel 2.11a plots the perturbed expected stock return process $\mu_S$ as a function of $C_t$ and at time $t = 0, 3, 5$. Since $\mu_S = \mu + \sigma \lambda_1$ is a linear function of $\lambda_1$, it shares common characteristics with $\lambda_1$. In short, $\mu_s$ decreases over time given underfunding and vice versa if $C_t > 1$.

Panel 2.11b shows the movement of $\mu_L$. When $C_t$ is low, the perturbed expected liability return $\mu_L$ decreases over time because $\lambda_1$ and $\lambda_2$ is declining over time. For a large $C_t$, the negative effect of $\lambda_1$ diminishes over time and $\lambda_2$ converges to $\kappa$; therefore $\mu_L$ goes up over time and converges to $a + b\sqrt{1 - \rho^2}\kappa$.

2.5.3 Dynamic Policy Evaluation

We have established from the static case that the robust policy performs better when the drift term is over-estimated. In this section, we will investigate the policy indifference curve as a function of time.

Figure 2.12 displays the policy indifference curve in different time period. Panel 2.12a
CHAPTER 2. ROBUST HEDGING IN INCOMPLETE MARKETS

Figure 2.11: Dynamic perturbation effect on drift terms. In this figure, we plot the dynamic movement of the perturbed drift terms as functions of the instantaneous coverage ratio at time $t = 0, 3, 5$. Panel 2.11a depicts the movement of $\mu^S = \mu + \sigma \lambda_1$ and Panel 2.11b shows $\mu^L = a + b \rho \lambda_1 + b \sqrt{1 - \rho^2} \lambda_2$.

displays the scenario when $C_t = 80\%$ and Panel 2.12b shows the plots at $C_t = 90\%$. Each panel plots the indifference curve at time $t = 0, 3, 5$. In the area beneath the indifference curve, the robust policy is better in the sense that the cost of hedging with a given amount of estimation error is lower. In line with the finding from Figure 2.5, we also observe from Figure 2.12 that the robust policy’s beneficial region has a positive relationship with coverage ratio $C_t$. Additionally, we find that the beneficial region decreases over time. However this dynamic effect is relatively weaker in Panel 2.12b when the instantaneous coverage ratio is relatively high.

2.6 Conclusion

In this chapter, we have provided a robust hedging strategy under the condition that the market is incomplete and the underlying model is misspecified. In short, a robust policy requires an extra cost of capital to guarantee against model uncertainty. It can, however, tolerate noises of the model and provide more successful hedges.

From our analysis, we summarize two major characteristics of the robust policy. We first find that the robustness effect strongly depends on the instantaneous coverage ratio. The preference for robustness only influences the hedging policy when the coverage ratio is low; if the fund’s assets are large enough to cover the liability payoff, then the robust and the naive policies are identical. Therefore, underfunded pension funds should consider employing a robust asset allocation policy, as a robust policy can effectively protect them from uncertain future crises which may lead to even worse solvency positions and ruin their recovery plan.
2.6. CONCLUSION

Second, we find that the agent can benefit from the robust policy when the expected return is over-estimated. That means, with a given expected-shortfall hedging target, the robust policy requires less initial wealth to obtain a successful hedge than the naive policy if the true expected stock return is lower than the estimated value.

In practice, robust polices are difficult to implement, because the max-min expected utility framework is difficult to solve, especially in the real world with multiple dimensions. However, a crucial insight of the robust policy is that it provides a low-risk and relatively high-return portfolio. Empirical evidences, such as Miller and Scholes (1972) and Jensen et al. (1972) have shown that low-beta stocks can resemble the property of robust portfolios. However, are low-beta stocks safe enough to substitute the robust portfolio? The 2008 financial crisis has witnessed the failure of the low-beta stocks as most of them did not even beat the S&P 500 which dropped by 37% by the end of the year.

There are several limitations of the our robust model, which can be further improved in the future research. Firstly, we assume a fixed correlation between the asset and the liability. However, correlations between assets and liabilities can vary considerably over time. A time-varying correlation parameter on one hand alters the incompleteness of the financial market, and on the other hand adjusts nature’s freedom of perturbation. Secondly, we should widen the source of uncertainty. For instance, instead of considering additional drift terms in Geometric Brownian Motion only, we should also study the uncertainty of the drift term with mean reversion. Wachter (2002) has shown that mean-reverting risk premium plays an important role in life-cycle investment model. Maenhout
(2006) argues that a misspecified mean-reverting risk premium parameter may result in a suboptimal portfolio choice. However, there is no study considering robust hedging rules with liability constraint for a mean-reversion risk premium.
2.7 Chapter 2 Appendix

2.7.1 Change of Measure

We employ a change of probability measure approach to solve \( \mathbb{E}[(L_T - A_T)^+] \) using the value of the liabilities \( L_t \) as a numeraire. The model is given by

\[
\begin{align*}
    dA_t &= \mu_A A_t dt + \sigma_A A_t dW_1, \\
    dL_t &= \mu_L L_t dt + \beta \rho L_t dW_1 + b \sqrt{1 - \rho^2} L_t dW_2
\end{align*}
\]

First we need to construct a Radon-Nikodym derivative \( \theta \) that changes measure from \( P \) to a new measure \( L \). According to Pelsser (2000), with given measure \( P \) and \( L \), if there is a random variable \( \theta = \frac{dL}{dP} \) such that \( \mathbb{E}^L[X] = \mathbb{E}^P[\theta X] \) for all random variables \( X \), then we say \( \theta \) is the Radon-Nikodym derivative of \( L \) with respect to \( P \). The expectation under \( P \) of \( \theta \) must be equal to 1 because \( \mathbb{E}^P[\theta] = \mathbb{E}^L[1] = 1 \). By following the derivation below

\[
\begin{align*}
    \mathbb{E}[(L_T - A_T)^+] &= \mathbb{E}\left[L_T \left(1 - \frac{A_t}{L_t}\right)^+\right] \\
    &= \mathbb{E}\left[\mathbb{E}[L_T] \frac{L_T}{\mathbb{E}[L_T]} (1 - C_T)^+\right] \\
    &= \mathbb{E}[L_T] \mathbb{E}^L[(1 - C_T)^+] \\
    &= \mathbb{E}[L_T] \mathbb{E}^L[(1 - C_T)^+] \\
\end{align*}
\]

we construct a Radon-Nikodym process \( \{\theta_t\} = \frac{L_t}{\mathbb{E}[L_t]} \) that changes the probability measure from \( P \) to \( L \). Notice that \( \{\theta_t\} \) is strictly positive and its expectation under the \( P \) measure is \( \mathbb{E}(\theta_t) = 1 \).

According to Ito’s Lemma \( ^7 \)

\[
dC_t = d[A_t (L_t^{-1})] = A_t d(L_t^{-1}) + L_t^{-1} dA_t + d[A_t, L_t^{-1}] \tag{2.27}
\]

\( ^7 \)We apply multivariate Ito’s rule for computing quadratic covariation. Suppose \( x \) and \( y \) are functions with finite quadratic variation. Define \( \phi(X,Y) = XY \) a product function of the two variables. Then

\[
d(XY) = Y dX + X dY + d[X,Y] \tag{2.26}
\]

where \( d[X,Y] \) is the quadratic covariations between \( X \) and \( Y \).
where
\[
A_t d(L_t^{-1}) = A_t \left[ -\frac{dL_t}{L_t^2} + \frac{d[L_t, L_t]}{L_t^3} \right] \\
= C_t \left[ (-\mu_L + b^2) dt - b\sigma dW_1 - b\sqrt{1-\rho^2} dW_2 \right]
\]
\[
L_t^{-1} dA_t = C_t [\mu_A dt + \omega \sigma dW_1] \\
d[A_t, L_t^{-1}] = -C_t b\rho \omega \sigma dt
\]

Hence the diffusion process of the coverage ratio \( C_t \) under \( P \) measure is given by
\[
dC_t = C_t \left[ (-\mu_L + \mu_A + b^2 - b\rho \sigma) dt + (\omega \sigma - b\rho) dW_1 - b\sqrt{1-\rho^2} dW_2 \right]
\]
(2.28)

Next, Girsanov Theorem is applied to define the new measure. We show here how to find the scalar process \( \lambda_t \) in our model. Given that \( \theta_t = \frac{L_t}{E[L_t]} \) with the expectation under measure \( P \), we can derive that
\[
d\theta_t = \frac{dL_t}{E[L_t]} - \frac{dE[L_t]}{E[L_t]^2} \\
= \theta_t \left[ b\rho dW_1 + b\sqrt{1-\rho^2} dW_2 \right]
\]
Hence, according to Girsanov Theorem, the Radon-Nikodym process is given by
\[
d\theta_t = -\theta_t \left( -b\rho, -b\sqrt{1-\rho^2} \right) \begin{pmatrix} dW_1 \\ dW_2 \end{pmatrix}
\]
so \( \lambda_t = \begin{pmatrix} -b\rho \\ -b\sqrt{1-\rho^2} \end{pmatrix} \), and the new Brownian motion under the \( Q \) measure is given by
\[
dW^L_1 = -b\rho dt + dW_1 \\
dW^L_2 = -b\sqrt{1-\rho^2} dt + dW_2
\]

Then we can rewrite equation (2.28) under measure \( L \) as,
\[
dC_t = (\mu_A - \mu_L) C_t dt + (\omega \sigma - b\rho) C_t dW^L_1 - b\sqrt{1-\rho^2} C_t dW^L_2
\]

Next we show how to calculate \( E^L [(1 - C_T)^+] \),
\[
E^L [(1 - C_T)^+] = E^L [1_{12C_T}] - E^L [C_T 1_{12C_T}]
\]
where 1 is an indicator function.

\[ 1 \geq C_T \Rightarrow \ln \left( \frac{1}{C_0} \right) \geq \mu_A - \mu_L - \frac{1}{2} \left( (w\sigma - b\rho)^2 + (1 - \rho^2) b^2 \right) T + (w\sigma - b\rho) \sqrt{T} W_1^T - b\sqrt{1 - \rho^2} \sqrt{T} W_2^T \]

\[ \Rightarrow \ln \left( \frac{1}{C_0} \right) \geq \mu_A - \mu_L - \frac{1}{2} \sigma^2_T + \sigma_C \sqrt{T} Z \]

\[ \Rightarrow Z \leq -\ln \bar{C} + \frac{1}{2} \sigma^2_C \sqrt{T} \]

\[ \sigma^2_C = (w\sigma - b\rho)^2 + (1 - \rho^2) b^2, \quad Z \sim N(0, 1) \] and \( \bar{C} = C_0 \exp((\mu_A - \mu_L) t) \). Hence

\[ E_L[C_T \mathbb{1}_{1 \geq C_T}] = \Phi(-d_2). \]

\[ E[L_T] = E \left[ L_0 \exp \left( \left( \mu_L - \frac{1}{2} b^2 \right) T + b\rho \sqrt{T} W_1 + b\sqrt{1 - \rho^2} \sqrt{T} W_2 \right) \right] = L_0 \exp \left( \mu_L T - \frac{1}{2} b^2 + \frac{1}{2} \left( b\rho \right)^2 + \frac{1}{2} \left( b\sqrt{1 - \rho^2} \right)^2 \right) = L_0 \exp (\mu_L T) = L \]

To summarize

\[ E[L_T] E^L \left[ (1 - C_T)^+ \right] = \bar{L} \left( \Phi(-d_2) - \bar{C} \Phi(-d_1) \right) \]

### 2.7.2 Static Optimal Solution

We show in this section the first order condition of the static value function with respect to each control variable. We first introduce some properties that will be applied in the calculation.

**Property 1:** \( d_2 = d_1 - \sigma_C \sqrt{T} \).

**Property 2:** \( d_2^2 = d_1^2 - 2 \ln \bar{C} \).

**Property 3:** \( \Phi'(-d_1) = \Phi'(d_1) \)
CHAPTER 2. ROBUST HEDGING IN INCOMPLETE MARKETS

Property 4: \( \Phi'(-d_2) = \Phi'(-d_1) \bar{C} \)

Property 5: \( \frac{\partial (-d_2)}{\partial w} - \frac{\partial (-d_1)}{\partial w} = \frac{\partial (\sigma \sqrt{T})}{\partial \lambda_2} = 0 \)

Property 6: \( \frac{\partial (-d_2)}{\partial \lambda_1}, \frac{\partial (-d_1)}{\partial \lambda_1} = \frac{\partial (\sigma \sqrt{T})}{\partial \lambda_1} = 0 \)

where \( \Phi'(d_1) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2}d_1^2 \right) \)

- FOC w.r.t \( w \)

\[
\frac{\partial}{\partial w} \left[ \bar{L} \Phi (-d_2) - \bar{A} \Phi (-d_1) \right] = \bar{L} \frac{\partial}{\partial w} \left[ \Phi (-d_2) - \bar{C} \Phi (d_1) \right] = \bar{L} \left[ \Phi'(-d_1) \bar{C} \frac{\partial (-d_2)}{\partial w} - \bar{C} \frac{\partial (-d_1)}{\partial w} \right] = \bar{L} \left[ \Phi' (d_1) \bar{C} \sqrt{T} \frac{w \sigma^2 - b \rho \sigma}{\sigma_c} - \bar{C} \Phi (-d_1) (\mu - r + \sigma \lambda_1) T \right]
\]

- FOC w.r.t \( \lambda_1 \)

\[
\frac{\partial}{\partial \lambda_1} \left[ \bar{L} \Phi (-d_2) - \bar{A} \Phi (-d_1) \right] = \frac{\partial \bar{L}}{\partial \lambda_1} \Phi (-d_2) + \bar{L} \frac{\partial \Phi (-d_2)}{\partial \lambda_1} \Phi (-d_1) - \frac{\partial \bar{A}}{\partial \lambda_1} \Phi (-d_1) = \bar{L} \lambda_1 \left[ \Phi' (d_1) \bar{C} \sqrt{T} \frac{w \sigma^2 - b \rho \sigma}{\sigma_c} - \bar{C} \Phi (-d_1) (\mu - r + \sigma \lambda_1) T \right]
\]

- FOC w.r.t \( \lambda_2 \)

\[
\frac{\partial}{\partial \lambda_2} \left[ \bar{L} \Phi (-d_2) - \bar{A} \Phi (-d_1) \right] = \frac{\partial \bar{L}}{\partial \lambda_2} \Phi (-d_2) + \bar{A} \left[ \Phi' (d_1) \frac{\partial (-d_2)}{\partial \lambda_2} - \frac{\partial (-d_1)}{\partial \lambda_2} \right] = \frac{\partial \bar{L}}{\partial \lambda_2} \Phi (-d_2) = \bar{L} b \sqrt{1 - \rho^2} \Phi (-d_2)
\]

2.7.3 Numerical Methodology

In this appendix, we will elaborate the numerical procedure of dynamic programming problem. We simplify our optimization problem in two steps then we employ explicit finite difference method to show the problem.
2.7.3.1 Simplification Stage 1

We first simplify the optimization problem (2.10) to

$$\min_{w} \max_{\lambda_1, \lambda_2} L_t \mathbb{E}[(1 - C_T)^+ | \mathcal{F}_t]$$  \hspace{1cm} (2.32)

where $C_t = \frac{A_t}{L_t}$ and the control variables are functions of time. The simplification does not influence the final solution of the control variables since $L_t$ is not affected by the agent’s decision.

We now show the proof this property. Denoting the value function by $U(A_t, L_t, t)$ with boundary condition $U(A_T, L_T, T) = \mathbb{E}_T[(L_T - A_T)^+]$. We also define value function $V(C_t, t)$ with boundary condition $V(C_T, T) = \mathbb{E}_T [(1 - C_T)^+]$. We want to prove that $U(A_t, L_t, t) = L_t V(C_t, t)$. If this property holds, then the simplification of the optimization problem states in (2.32) is valid.

The partial differential equation for $U(A_t, L_t, t)$ can be written as

$$U_t + U_A r + w(\mu - r) + w \sigma \lambda_1 + U_L \left( a + b p \lambda_1 + b \sqrt{1 - \rho^2 \lambda_2} \right)$$

$$+ \frac{1}{2} U_{AA} \sigma^2 A^2 + \frac{1}{2} U_{LL} b^2 L^2 + U_{AL} b p w \sigma A L = 0$$ \hspace{1cm} (2.33)

Assume $U(A_t, L_t, t) = L_t V(C_t, t)$ is valid, then $L_t$ on the right side only has a scale effect and we would expect a univariate PDE for $V(C_t, t)$ that does not depend on $L_t$. The partial derivative of $U(A_t, L_t, t)$ can be expressed in terms of $V(C_t, t)$

$$U_A = V_C$$ \hspace{1cm} (2.34a)

$$U_L = V - V_C C$$ \hspace{1cm} (2.34b)

$$U_{AA} = V_{CC} \frac{1}{L}$$ \hspace{1cm} (2.34c)

$$U_{LL} = V_{CC} \frac{C^2}{L}$$ \hspace{1cm} (2.34d)

$$U_{AL} = -V_{CC} \frac{C}{L}$$ \hspace{1cm} (2.34e)

$$U_t = LV_t$$ \hspace{1cm} (2.34f)
CHAPTER 2. ROBUST HEDGING IN INCOMPLETE MARKETS

Replace (2.34) into (2.33) we get

\[ V_t + V_C C (r + w(\mu - r) + w\sigma\lambda_1) + (V - V_C C) \left( a + b\rho\lambda_1 + b\sqrt{1 - \rho^2}\lambda_2 \right) \\
+ \frac{1}{2} V_C C^2 w^2\sigma^2 + \frac{1}{2} V_C C^2 b^2 - V_C C^2 b\rho\sigma = 0 \quad (2.35) \]

PDE (2.35) is what we would have expected by using the new lognormal process \( dC \) see (2.28). Since (2.35) is derived from (2.33) and is \( L \) neutral, we should be able to get the same optimal \( w, \lambda \)'s from the two PDE's, but the latter one is easier to solve numerically since it is a univariate PDE.

2.7.3.2 Simplification Stage 2

The second simplification procedure is related to the numerical aspect. It is more efficient to use finite difference methods with \( \ln C \) instead of \( C \). Define \( Z = \ln C \), then we have

\[ dZ = \left( \mu_A - \mu_L - \frac{1}{2} b^2 - \frac{1}{2} w^2\sigma^2 \right) dt + (w\sigma - b\rho) dW_1 - b\sqrt{1 - \rho^2} dW_2 \quad (2.36) \]

and the corresponding simplified HJB equation (2.35) in terms of \( U(Z,t) \) is given by

\[ -U_t = \min_{w} \max_{\lambda_1, \lambda_2, \nu} U_Z \left( \mu_A - \mu_L - \frac{1}{2} \sigma^2 \right) + U(\mu_L) + U_{ZZ} \left( \frac{1}{2} \sigma^2 \right) \quad (2.37) \]

using the following relations,

\[ U_Z = V_C C \quad (2.38a) \]
\[ U_{ZZ} = V_C C^2 + V_C C \quad (2.38b) \]

If we replace the explicit optimal solution for \( \lambda_1 \) and \( \lambda_2 \) (see (2.22)) into the uncertainty set constraint \( \lambda_1^2 + \lambda_2^2 = \kappa^2 \), we will get a fourth order polynomial equation of \( \nu \)

\[ (s_0 t_1 + s_1 t_0)^2\nu^2 + s_2 (\nu t_0 + t_1^2)^2 = \kappa^2 \nu^2(\nu t_0 + t_1^2)^2, \quad (2.39) \]

where \( s_0 = (U_{ZZ} - U_Z)b\sigma - U_Z(\mu - r), s_1 = -(U - U_Z)b\rho, s_2 = -(U - U_Z)b\sqrt{1 - \rho^2}, t_0 = (U_{ZZ} - U_Z)\sigma \) and \( t_1 = U_Z\sigma \).

Equation (2.39) contains four roots, but only one root gives us the min-max solution.\(^8\)

\(^8\)The advantage of using \( Z \) instead of \( C \) (under the condition that \( C_t \) is a Geometric Brownian motion process) as the state variable of the value function is that we can turn the partial differential equation state-variable neutral, so that we can get a relatively simple version of PDE equation.

\(^9\)The remaining three solutions of equation (2.39) result in the min-min, max-max and max-min value
2.7. CHAPTER 2 APPENDIX

Also, this specific \( \nu \) has to be positive, this is because, on one hand the optimal \( \lambda_2^* \) (see equation (2.22)) is a function of \( \nu \), and we have discussed that \( \lambda_2 \) has to be positive; and on the other hand the numerator of the fraction \( \lambda_2^* \) is positive. Hence that results in a positive \( \nu \).

The partial differential equation for \( U(Z,t) \) is given by

\[
U_t + U_Z (r + w(\mu - r) + w\sigma\lambda_1) + (U - U_Z) \left( a + bp\lambda_1 + b\sqrt{1 - \rho^2}\lambda_2 \right) \\
+ (U_{ZZ} - U_Z) \left( \frac{1}{2} w^2\sigma^2 + \frac{1}{2} b^2 - bpw\sigma \right) = 0 \quad (2.40)
\]

as we can see, PDE (2.40) is state variable free, so compared with the PDE (2.35) from \( U(C,t) \), it is relatively more efficient to use log coverage ratio.

2.7.3.3 Explicit Finite Difference

We generate a stylized finite difference grid. We divide time \( T \) into a certain \( N \) (unknown) equally spaced interval with length \( \Delta t = \frac{T}{N} \). Let \( C_{\text{min}} \) and \( C_{\text{max}} \) be the two extremes of the coverage ratio with corresponding log extremes \( Z_{\text{min}} \) and \( Z_{\text{max}} \). We divide the interval \([Z_{\text{min}}, Z_{\text{max}}]\) into \( M \) particular spaced intervals. with length \( \Delta Z \). According to Hull (2009), it is more efficient to set \( \Delta Z = \sigma\sqrt{3\Delta t} \). Therefore, we set the length \( \Delta Z = \frac{3\sigma^2}{Z_{\text{max}} - Z_{\text{min}}} \)

\[^{10}\text{Given that } M = N \text{ and } M = \frac{Z_{\text{max}} - Z_{\text{min}}}{\Delta Z} \text{ and } \Delta Z = \sigma\sqrt{3\Delta t} = \sigma\sqrt{\frac{3T}{N}}, \text{ we have } \Delta Z^2 = \sigma^2 \frac{3T}{N}.\]
Chapter 3

Robust Long-Term Interest Rate Risk Hedging in Incomplete Bond Markets*

3.1 Introduction

It is important for insurance companies and pension funds to make a long-term strategic investment plan. Most people start working and participating in a pension scheme from age 25. The average life expectancy in most developed countries is above 80 years. Hence insurance companies and pension funds face commitments with maturities of more than 50 years.

Pricing and hedging long dated liabilities faces two challenges. On the one hand, the market for long maturity financial instruments is incomplete, since the market for fixed income securities with maturities of over 20 years is illiquid. However, regulations in many countries (especially in Europe under the Solvency II program) require pension and insurance companies to follow a “market-consistent” principle in valuing their liabilities, which means that insurance companies want to invest in very long-maturity financial instruments in order to replicate the long-dated cash flows.

The standard replication theory does not work because long-maturity financial contracts barely exist. Consols are among the rare examples of perpetual bonds, issued by the UK in 1917 in order to help pay for World War I. However, the consols market is not very liquid and is only a very small part of the UK government bond portfolio. The longest government bonds even in developed financial markets (such as US and Germany) have maturities no longer than 30 years. In developing countries (such as Asia, Eastern

*This chapter is based on my working paper Shen (2014) co-authored with Antoon Pelsser and Peter Schotman (Maastricht University and Netspar).
Europe and South America), the longest government bonds have maturities no longer than 10 years.

On the other hand, it is difficult to avoid parameter misspecification. Due to the incompleteness of the bond market, it is necessary to extrapolate the term structure of interest rates beyond the maturity of longest dated market-available instrument. In the absence of a bond market, any valuation of a long dated liability must therefore be model based. Conditional on a model, we could derive a term structure for all maturities. However, there are a large number of models that fit bond prices up to maturity of 20 years perfectly, but still indicate different prices for the longer maturity bonds.

Despite much work on term structure models, very few studies consider the impact of model uncertainty on long dated liability valuation. In this chapter, a robust optimal hedging policy is proposed that minimizes the shortfall of long dated liabilities in the presence of parameter uncertainty and missing bond markets. Our replicating portfolio is robust to model misspecification in the sense that the investment policy is less sensitive to the choice of model. A second aim is to propose an alternative term structure that is robust against long-term interest rate risk.

The challenge of pricing in incomplete markets is closely related to model misspecification. Market incompleteness creates an unknown market price of risk. This unobservable market price of risk has prompted the rationale of Cochrane and Saa-Requejo (1996)’s Good-Deal-Bound, which restricts the maximum Sharpe ratio of the market. The true market price of risk could be anywhere within this Good-Deal-Bound.

Investor who fear parameter misspecification are ambiguity averse. They believe that the true model parameters differ from the estimated ones. To formulate model misspecification, Hansen and Sargent (2001) and Hansen and Sargent (2007) employ a relative entropy factor. This relative entropy captures the perturbation between the estimated model and the unobservable true model. Although the economic interpretations of the two are different, technically the two have identical motives. Both interpretations can be understood as requiring an additional premium to represent the estimation error.

We solve a dynamic robust optimization problem by employing Hansen and Sargent (2007)’s framework. In the model, the agent allocates her wealth between a short-term and a medium-term bond in order to minimize the expected shortfall of a long maturity commitment. On the other hand, nature perturbs the parameters in order to maximize the expected shortfall given the decision of the agent. The robust optimal portfolio is therefore robust against model ambiguity.

We propose a feasible region of uncertainty set that gives nature a reasonably bounded freedom of decision-making. We use the generalized method of moments (GMM) approach to estimate the one-factor affine term structure model. By assuming the estimation errors
3.1. INTRODUCTION

to be asymptotically normal, one can use the property that standardized error term
square becomes a Chi-squared distribution in order to shape a joint confidence interval
for estimation parameters.

To solve our robust optimization problem, we propose a regression-based method. The
new method belongs to the family of the backward least squares Monte Carlo method.
It combines, in essence, the methods of Brandt et al. (2005) and Koijen et al. (2007),
proposed to approximate the conditional expectation encountered in solving the dynamic
program by polynomial expansions in the state variables. In addition, we add the prefer-
ence for robustness into the algorithm.

We find that the preference for robustness induces strong demand for long-term bonds
when the solvency ratio is low. Robust investors worry about model uncertainty in the
sense that they are afraid that nature would choose a lower bond premium than they
expected, hence they need a risker portfolio so as to gamble their way out of trouble.
Both naive and robust optimal portfolios depend on the hedging horizon. The longer the
horizon, the more risk exposure on long-term bond markets. We also find that the robust
policy requires more initial wealth than the naive policy in order to meet the shortfall
target. In other words, the robust yield curve is always lower than the naive yield curve.
However, when the spot rate is low, both policy-based yield curves are higher than the
Vasicek yield curve.

This chapter is related to three different literatures. The first concerns optimal portfo-
lio choice under model uncertainty. Early works are dominated by the Bayesian paradigm.
Klein and Bawa (1976) is one of the first studies to consider the effect of model uncertainty
on portfolio choice. They look at a two-period model and find that in the presence of es-
timation risk, the optimal hedging portfolio differs from the traditional analysis. Barberis
(2000) extends the study in a multi-period economy setting while keeping the hedging
decision static, and finds that ignoring estimation risk may result in an over-aggressive
portfolio. Brennan (1998) incorporates learning with parameter uncertainty and finds
that risk lovers put more wealth into risky assets after learning while the risk averse
investors are more conservative with their portfolio.

A more recent approach is the max-min expected utility paradigm developed by Gilboa
and Schmeidler (1989). Gilboa and Marinacci (2011) claim that Gilboa-Schmeidler’s
axiom is a neo-Bayesian paradigm because it allows decision makers to have a set of
subjective priors. The agent aims to maximize her utility under the least preferred prior
in order to demonstrate an aversion to uncertainty. As an extension, Hansen and Sargent
(2001) managed to transform Gilboa-Schmeidler’s static theory into a dynamic version
through the techniques of robust control theory.

Our work also relates to the long-term investment literature. Early work by Brennan
and Xia (2002) and Campbell and Viceira (2002) studies the optimal hedge demand for long-term bonds assuming constant and specified bond premia. Koijen et al. (2010) and Sangvinatsos and Wachter (2005) allow for time-varying bond risk premia and both find that time-varying risk premia induce a large hedging demand for long-term bonds. None of the papers above consider bond premia misspecification or the incompleteness of the very long-term bond markets.

Additionally, studies on dynamic hedging in incomplete markets are related to our work. Basak and Chabakauri (2012), Föllmer and Leukert (1999) and Föllmer and Leukert (2000) all provide a closed form solution of dynamic hedge in incomplete markets while assuming that the underlying models are correctly specified. Basak and Chabakauri (2012) introduce an unhedgeable risk driver to represent market incompleteness, and employ minimum-variance as hedging criterion; they conclude that a larger hedging demand is induced due to speculative uncertainty. Föllmer and Leukert (1999) maximize the success of hedge and Föllmer and Leukert (2000) minimize the probability of shortfall risk. They define the incomplete market as a system with multiple martingale measures. Both studies find that the optimal solution depends on the shortfall target, as is also found in our study.

The rest of the chapter proceeds as follows. Section 3.2 describes the one-factor affine term structure model employed in our economy. Section 3.2.1 uses the GMM method to estimate the structural parameters. The dynamic robust optimization problem is explained in Section 3.3. Section 3.4 elaborates the regression-based techniques used in solving our dynamic programming problem. Section 3.5 discusses the robust optimal solution and we provide policy evaluation of the robust policy. Section 3.6 gives conclusions.

3.2 Term Structure Model

Our term structure model follows the framework of Duffee (2002) and Duffie and Kan (1996). We assume that the spot rate $r$ follows the one-factor Vasicek model

$$dr = \kappa (\theta - r) dt + \sigma dW,$$  \hspace{1cm} (3.1)

where $\theta$ is the unconditional mean, $\kappa$ is the mean reversion, $\sigma$ is a volatility parameter and $dW$ a univariate Brownian motion. Let $P(t, T)$ be the time $t$ price of a discount bond maturing at time $T$. The Vasicek model implies that bond prices follow the diffusion

$$dP = rP dt + B \sigma P \left( \lambda dt + dW \right),$$  \hspace{1cm} (3.2)

where $B = \frac{1-e^{-\kappa T}}{\kappa}$ is the volatility of long-term bond returns relative to the spot rate volatility and where $\lambda$ is the price of risk. In the original Vasicek (1977) model, the price
3.2. TERM STRUCTURE MODEL

Table 3.1: Summary Statistics
Means, standard deviation and autocorrelations of daily yield curve spot rate with three different maturities and their difference. The variable \( r_t \) denotes three-month spot rate. ADF denotes the Augmented Dickey-Fuller unit root statistics with a 5% critical value of \(-3.43\).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std</th>
<th>Std ∆</th>
<th>ADF</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_t )</td>
<td>1.62%</td>
<td>1.45%</td>
<td>3.28bp</td>
<td>-0.61</td>
</tr>
<tr>
<td>( Y_t(5) )</td>
<td>2.61%</td>
<td>1.10%</td>
<td>4.27bp</td>
<td>-0.40</td>
</tr>
<tr>
<td>( Y_t(10) )</td>
<td>3.36%</td>
<td>0.80%</td>
<td>3.97bp</td>
<td>-0.94</td>
</tr>
</tbody>
</table>

of risk is constant. In the essential affine extension of Duffee (2002), the price of risk also depends on the spot rate,

\[
\lambda = \Lambda_0 + \Lambda_1 r \tag{3.3}
\]

In our model we will assume that the volatilities \( \sigma \), \( \kappa \) and \( \theta \) can be estimated very precisely, whereas the expected excess return parameters \( \Lambda_0 \) and \( \Lambda_1 \) are subject to model uncertainty.

3.2.1 Model Calibration

The model parameters are estimated using standard GMM. We use Euro area nominal government bonds with triple A issuing ratings, obtained from the ECB statistical data warehouse. We use daily annualized data on 3 constant maturity zero rates with maturities of 3 months, and 5 and 10 years for the period from September 6, 2004 to November 15, 2013.\(^2\)

Summary statistics are displayed in Table 3.1. The average three-month rate is 1.62% with a standard deviation of 1.45%. The 5-year rate has a mean of 2.61% with standard deviation of 1.1%. The 10-year bond has highest average rate but lowest volatility among the three. The ADF test shows that we cannot reject the unit root hypothesis for any of the three series.

In Table 3.2, we report the parameter estimation results for the full sample. The parameters are expressed in annual terms. The first two columns of Table 3.2 report the estimated value and standard deviation of each structural parameter. The unconditional mean of spot rate, \( \theta \), is about 162 basis points. The mean reversion coefficient \( \kappa \) implies half-life innovation of 4 years.

The rest of the table reports the correlation matrix of parameters. We find an extremely high correlation between \( \kappa \) and \( \Lambda_1 \), and between \( \theta \) and \( \Lambda_0 \). The low volatility of

\(^2\)The yield data is available at sdw.ecb.europa.eu. Sample size is 2360, and there are 9.2 (approximately) years in our sample, hence the average yearly number of trading days is \( \frac{2360}{9.2} \approx 255 \).
Table 3.2: GMM Estimates.
The term structure model is estimated by the GMM method using daily European data. The bond maturities used are three-month, five-year and ten-year maturities. The parameters are expressed in annual terms. We choose 20 as the number of lags applied in the Newey West estimator.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std error</th>
<th>Correlation Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \kappa )</td>
<td>17.16%</td>
<td>17.43%</td>
<td>1.0000</td>
</tr>
<tr>
<td>( \theta )</td>
<td>1.62%</td>
<td>1.24%</td>
<td>-0.5893</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>0.52%</td>
<td>0.14bp</td>
<td>0.8674</td>
</tr>
<tr>
<td>( \sigma \Lambda_0 )</td>
<td>-0.16%</td>
<td>0.36%</td>
<td>0.0380</td>
</tr>
<tr>
<td>( \sigma \Lambda_1 )</td>
<td>-14.19%</td>
<td>17.49%</td>
<td>-0.9987</td>
</tr>
</tbody>
</table>

\( \sigma \) implies an accurate estimation performance on spot rate volatility. Hence we ignore the parameter uncertainty on \( \sigma \). Checking the eigenvalues of the covariance matrix of the estimates, we find two positive eigenvalues 0.0609, 0.0001 whereas the remaining three are very close to zero. From this we conclude that two sources of parameters estimation error exist.

In this section, we consider the misspecification of the two risk-price factors. As a result, we fix the dynamics of the spot rate and only consider the effect of uncertainty in the risk premium of long-term bonds. In other words, only the drift of \( \frac{dP}{P} = rd\,t \) is effected by the parameter uncertainty and the volatility is fixed. The layout of the affine term structure model does not allow us to consider the misspecification of other parameters (neither \( \kappa \) nor \( \theta \)), because in this paper we only deal with the first-moment model uncertainty. The misspecification of \( \Lambda_0 \) and \( \Lambda_1 \) will only result in a drift distortion of bond premium, hence it can be related to a change of probability measure. However, the misspecification of \( \kappa \) and \( \theta \) will lead to a higher moment distortion, which is beyond the scope of our technical skills.

We assume prices of risk factors are jointly normal with mean \( \hat{\Lambda}_0 \) and \( \hat{\Lambda}_1 \) and covariance matrix \( \Omega \).

\[
\begin{pmatrix}
\hat{\Lambda}_0 \\
\hat{\Lambda}_1
\end{pmatrix} \sim N \left( \begin{pmatrix}
\hat{\Lambda}_0 \\
\hat{\Lambda}_1
\end{pmatrix}, \Omega \right)
\]

We denote \( c_0 \) and \( c_1 \) as the estimation errors of the two misspecified parameters. We know that the standardized error terms square is a Chi-squared distribution with two degrees of freedom. This leads to a confidence interval of \( \Lambda_0 \) and \( \Lambda_1 \), denoted \( S \)

\[
S := \left\{ \begin{pmatrix}
\hat{\Lambda}_0 \\
\hat{\Lambda}_1
\end{pmatrix} + \begin{pmatrix}
c_0 \\
c_1
\end{pmatrix} \right| \begin{pmatrix}
c_0 \\
c_1
\end{pmatrix} \Omega^{-1} \begin{pmatrix}
c_0 \\
c_1
\end{pmatrix} \leq \chi^2_2 \right\}
\]

58
where $\gamma^2$ represents the level of preference for robustness. When $\gamma^2 = 0$, the uncertainty set $S$ shrinks at the point estimates. The bigger $\gamma^2$ is, the more a robust investor worries about model uncertainty. Statistically, $\gamma^2$ is the critical value of a Chi-squared distribution with two degrees of freedom, which implies $\gamma^2$ should be equal to 5.99 at the 5% significance level.

We choose a value of $\gamma^2$ such that even in the worst-case scenario, we can still guarantee a non-negative bond premium $-B\lambda\sigma$. A feasible boundary of $\gamma^2$ is given by

$$0 \leq \gamma^2 \leq \frac{(\hat{\Lambda}_0^\prime \alpha)^2}{\alpha^\prime \Omega^{-1} \alpha} \quad (3.6)$$

where $\alpha = (1, r)'$ and $\hat{\Lambda} = (\hat{\Lambda}_0, \hat{\Lambda}_1)'$. Eq.(3.6) shows that the feasible value of $\gamma^2$ depends on the spot rate. The lower the spot rate, the lower the feasible $\gamma^2$ will be. To rule out negative bond premium scenarios for any plausible value of the spot rate, Eq.(3.6) is applied. The lowest upper bound occurs at the lowest spot rate level. Our sample data shows the lowest three-month spot is -0.017% and the corresponding value of $\gamma^2$ is 0.22. Therefore a feasible region is $[0, 0.22]$. As a benchmark, we select $\gamma^2 = 0.03$.

The impact of term structure factor $r$ on bond premia is governed by $\Lambda_1$, if $\Lambda_1$ is estimated with error, then the impact of the spot rate is also ambiguous. Figure 3.1 presents the 20-year nominal bond risk premia for a realistic range of spot rates. The risk premium is increasing with the spot rate. Misspecification of market price of risk results in a perturbation of the bond premium of around 200 basis points either downwards or upwards.

### 3.3 Robust Optimal Portfolio Choice

#### 3.3.1 Robust Hedging

The agent with wealth $X_t$ at time $t = 0$ aims to hedge a long dated liability with payoff equal to one at maturity time $T$ by investing in two financial instruments: a zero coupon bond with maturity $\tau_2 < T$; and a short term bond. Suppose the agent is only interested in eliminating the downside risk. Then our hedging criterion is defined over the expected shortfall $[1 - X_T]^+$ at time $T$. If the agent is not aware of the parameter uncertainty, then her hedging decision relies fully on the point estimator $\hat{\Lambda}$. The dynamic optimization

---

\[3^3\] Why do we choose such a small value? There are two explanations. First, for technical reasons $\gamma^2$ cannot be too high, otherwise the numerical results would be very unstable. Second, it is too conservative to let nature choose at a perturbation boundary of 95% confidence level at every decision time-point. However, the design of our optimization problem does not allow for a dynamic $\kappa^2$. An alternative solution to make the robust policy less conservative is to shrink the uncertainty set by reducing the value of $\kappa^2$. 

---

59
Figure 3.1: Bond Premia
The figure presents the 20-year bond risk premia for different spot rate under different choices of $\lambda_0$ and $\lambda_1$. The ambiguity of market price of risk constrained by set $S$ is shown in Eq.(3.5) with $\gamma^2 = 0.03$. Panel 3.1a presents the minimum and maximum values of bond premia within the feasible region for different $r$. The point estimate premia are located in between the two extremes. Panel 3.1b plots the feasible region of bond premia for different choices of $\lambda_1$ under the unconditional expectation of spot rate.

(a) Uncertainty constraint of bond premia
(b) Mother nature makes choices of $\lambda_0$ and $\lambda_1$, that makes the true value of bond premia deviate from the point estimate.

problem is defined as

$$\min_{w_t \in [0, T]} E \left[ (1 - X_T)^+ | \mathcal{F}_T \right]$$  \hspace{1cm} (3.7)$$

The fraction of wealth allocated to long-term bonds with maturity $\tau_2$ at time $t$ is indicated by $w_t$. The dynamics of wealth is given by

$$dX = (r - w B(\tau_2) \sigma \lambda) X dt - w B(\tau_2) \sigma X dW$$  \hspace{1cm} (3.8)$$

where we ignore the subscripts $t$. However, if the agent is uncertainty averse, then she fears misspecified bond premia. We find from Figure 3.1 that a different decision on $\Lambda$ would imply a different bond premium and hence will imply a different hedging position. In order to protect against parameter uncertainty, a robust investor seeks a robust policy that minimizes the hedging error under the worst-case decision of nature. The robust optimization problem is given by

$$\min_{w_t \in [0, T]} \max_{(\Lambda_0, \Lambda_1) \in S} E \left[ (1 - X_T)^+ | \mathcal{F}_T \right]$$  \hspace{1cm} (3.9)$$

We assume nature controls the price of risk parameters $\Lambda_0$ and $\Lambda_1$ and her action is centered around the point estimators $\hat{\Lambda}_0$ and $\hat{\Lambda}_1$ bounded by the uncertainty set $S$. The dynamic robust optimization problem can be interpreted as a two-player game. At each re-
balancing time $t$, player one, the robust investor, makes a hedging decision that minimizes the expected shortfall at the end of hedging horizon. Player two is nature. Conditional on player one’s choice, she makes a decision on bond premium parameters so as to maximize the expected shortfall at time $T$. The equilibrium of the game gives us an instantaneous robust hedge.

### 3.4 Numerical Technique

Robust hedging with an expected shortfall objective function does not allow for an analytical solution, so we use a numerical approach instead. Our method in essence combines the methods proposed by Brandt et al. (2005) and Koijen et al. (2007) to approximate the conditional expectation that we encounter in solving the dynamic program by polynomial expansions in the state variables. We also follow Diris (2011)’s approach to parameterize the coefficients of the approximation function by a quadratic function of portfolio weights, such that we can calculate the optimal portfolio under each path analytically. Furthermore, this method allows us to achieve an accurate result using a very small grid of testing portfolios. We integrate robustness with the standard simulation based algorithm.

We start by simulating a large number of $N$ sample paths with length $T$ years of bond returns using Euler discretisation. We also choose an $M$-dimensional grid for financial wealth values. The wealth grid points are indicated by $X_j, j = 1, \cdots, M$. In total, we have $(M \times N)$ grid points at each point in time.

The algorithm for the robust policy consists of two parts. We first solve nature’s decision analytically. Nature’s decision only influences the long-term bond premia. Therefore, under the assumption that $w_t \geq 0$, maximizing the expected shortfall boils down to minimizing bond premia. Therefore, we can analytically solve for nature’s decision.

The optimal $c_0, c_1$ are the solution of following quadratic programming problem.

$$\min_{\alpha_t} -B(\tau_2)\sigma \left( \hat{\Lambda} + c_t \right) \alpha_t,$$

s.t. $c_t^\prime A c_t = 1$

where $\hat{\Lambda} = (\hat{\Lambda}_0, \hat{\Lambda}_1)'$, $c_t = (c_0, c_1)'$, $\alpha_t = (1, r_t)$ $A = \frac{\Omega}{2}$. We can easily find the optimal solution for $c_t$:

$$c_t^\ast = \frac{A^{-1}\alpha_t}{\alpha_t^\prime A^{-1}\alpha_t}$$

Therefore, nature’s decision at each rebalancing time step $t$, depends solely on $r_t$. The min-max problem is now simplified to the minimization problem alone.
CHAPTER 3. ROBUST HEDGING OF LONG-DATED LIABILITIES

Figure 3.2: Nature’s Decision on Bond Premium. The figure presents nature’s choice of bond risk premium against the naive estimate $-B(\tau)\sigma\lambda$ under different bond maturities. The spot rate is equal to 2%.

Figure 3.2 displays the robust optimal bond premium as a function of bond maturity against the point estimate. Nature intends to minimize the bond premium so as to maximize the expected shortfall at time $T$, therefore, the robust bond premium at any maturity is always lower then the point estimate.

In the second part of the algorithm, we use the backward least square monte carlo (LSMC) method to solve for optimal portfolio. We summarize the LSMC algorithm in three steps.

**Step 1** At time step $t$, i.e. starting from period $T - \Delta t$ and iterating to period 0, construct realized loss $V_T = (1 - X_T)^+$ for all simulated path using the following information

- Robust optimal portfolio from previous time steps $w^*_s(s = t + \Delta t, \ldots, T - \Delta t)$
- Optimal natures decision from previous steps $\lambda^*_s$ (this is analytically solved)
- A small grid of testing portfolio $w_h, h = 1, \ldots, H$
- Spot rate $r_{i,s}$ with $i = 1, \ldots, N$
- Current wealth $X_{j,t}, j = 1, \ldots, M$

so we have a cross-section size of $N \times M \times H$.
3.5. LONG-TERM INVESTORS AND BOND PREMIA UNCERTAINTY

Step 2a Run a cross-sectional regression by approximating the realized objective function $V_{ijh}$ calculated from Step 1 on trajectories of state variables as well as the testing portfolio $w_h$ on a second-order polynomial expansion (including cross term) at a certain point in time.

$$V_{ijh,T}(F_t) = \beta' f(X_{j,t}, r_{i,t}, w_h) + \epsilon_{ijh,t} \quad (3.10)$$

Step 2b Now we parameterize the approximate conditional value function as a quadratic function of portfolio weights such that we can analytically calculate the optimal portfolio.

For a naive investor, the first part of the algorithm can be ignored since nature does not play a role in the naive hedging framework. Appendix 3.7.1 elaborates on the LSMC algorithm in more detail.

3.5 Long-term Investors and Bond Premia Uncertainty

In this section, we investigate the impact of model uncertainty on bond portfolio allocations. Section 3.5.1 analyzes the robust optimal portfolio. Section 3.5.2 discusses the property of the robust yield curve.

3.5.1 Optimal Portfolio Choice

Figure 3.3 plots the optimal portfolio of 20-year ($\tau_2 = 20$) nominal bonds as a function of the investment horizon based on three different hedging policies, namely naive policy, robust policy and delta hedging policy.

3.5.1.1 Delta Hedging

The delta of a long dated liability is defined as the rate of change of the price of liability with respect to the price of the underlying $\tau_2$-year bond. The delta of long-term liability is

$$\Delta = \frac{\partial P(0,T)}{\partial P(0,\tau_2)} = \frac{B(T)}{B(\tau_2)}$$

We use Delta hedge as a benchmark to investigate the optimal portfolio choice as a function of the solvency condition. We use the funding ratio, a fraction of the current wealth level $X_0$ and the hypothetical price of liability $P(0,T)$, to measure the solvency.
position of a fund. If the funding ratio is lower than one, then the fund has a higher expected value of liability than of its assets, which implies under-funding.

Suppose we have a funding ratio greater than one, and the underlying model is correctly specified, then the Delta position is indeed the optimal portfolio to hedge the long-dated liability. This is because when the current funding ratio is above one, downside risks fade away and the hedging problem boils down to a complete-market setting where the long-dated payoff can be fully replicated by traded bond portfolios. It is confirmed by Panel (b) that if the current funding ratio is above one, a delta neutral position is the optimal policy.

Figure 3.3 shows that the bond allocations are highly dependent on the hedging horizon and the present funding ratio for both the naive and the robust policies. Both hedging methods are identical when the horizon $T \leq 20$, because market completeness eliminates the model uncertainty. Also, both policies suggest a riskier position than the Delta hedge. When the funding ratio is low, following the Delta strategy may hedge the long-term interest rate risk but does not help to meet the long-dated commitment, due to insufficient current wealth. Therefore the Delta hedge is not optimal and we need a riskier portfolio in order to obtain higher expected excess return.

With a long-term investment horizon, a robust investor takes a more risky position than a naive investor when the present funding ratio is low. This is because a robust investor worries that nature might choose even lower bond premia than the model estimated, and therefore the robust method is even more aggressive. The result for the robust policy is the opposite of many results in the literature eg. Maenhout (2004). Usually, the robust policy is more conservative compared to the naive one. In Maenhout (2004)’s model, an investor aims to maximize her terminal wealth utility function. The preference for robustness suppresses the beliefs in risk premium. Without the liability constraint, a robust investor will therefore take less risk.

We now demonstrate how the optimal policy respond to changes in the funding ratio and the spot rate. Figure 3.4 displays the optimal asset allocation to 20-year bonds under the naive policy (panel (a)) and the robust policy (panel (b)) for a reasonable range of the funding ratio (FR). These figures are constructed by regressing the optimal policy along all trajectories of state variables used in the simulations at a certain point in time.

Figure 3.5 summarizes the insight of Figure 3.4. The lower the funding ratio, the riskier the position for hedging against the shortfall risk. This property holds true for both policies. However, the robust policy is riskier due to the fear of underestimated bond premia. Figure 3.3 also leads us to the same conclusion.

Figure 3.6 presents the optimal asset allocation under different spot rates when the funding ratio is 80%. For both policies, the optimal allocation increases with the spot
3.5. LONG-TERM INVESTORS AND BOND PREMIA UNCERTAINTY

rate, as the spot rate is positively related to bond premia. A higher bond premium leads to a higher risk exposure on long-term bonds. The robust policy (panel (b)) differs from the naive policy (panel (a)) in two ways. First, the robust policy is less sensitive to the spot rate. This is an important and a desired feature of the robust policy. Second, the difference between the two policies gets larger for a lower value of the spot rate.

Figure 3.7 presents the optimal expected shortfall value as a function of the funding ratio and hedging horizon. Figure 3.8 summarizes the insights of Figure 3.7. First, without considering model uncertainty, the expected shortfall value is dramatically underestimated when the funding ratio is low. Secondly, the mispricing error increases with the investment horizon. The increasing amount of mispricing is caused by an accumulative fear of the underestimated bond premia.

3.5.2 Robust Term Structure

An investor is interested in the optimal strategy that guarantees the liability with the lowest required initial wealth \( X^* \), which is called the super-hedging strategy. Once we have determined the minimum wealth \( X^* \), we can define the implicit period discount rate as

\[
y(T) = -\frac{1}{T} \ln X^* \tag{3.11}
\]

For most models of the term structure, the super-hedging strategy will be extremely costly. In our model, a super-hedging strategy will not exist, since there will always be a small probability of underfunding because our interest rate process is Gaussian. We define approximately risk-free rates through the minimum assets required as having an expected shortfall of less than \( S \),

\[
y(T) = -\frac{1}{T} \ln X^*_S, \quad \text{with} \quad S = E_0 [(1 - X_T)^+ | X_0 = X^*_S] \tag{3.12}
\]

Expected shortfall does not reward any upside. The lower \( S \) is, the more initial wealth is required to hedge against the targeting shortfall risk, hence the lower \( y(T) \) will be. The lowest level of \( y(T) \) is simply the underlying term structure model that by definition guarantees zero expected shortfall.

Figure 3.9 presents the implied yield curves under different hedging policies at different shortfall levels and initial spot rates. As a benchmark, we plot the two hedging policy based yields together against the Vasicek yield curve. We highlight two crucial insights from Figure 3.9. First, the robust yield is always lower than the naive yields regardless of the current spot rate or shortfall target. Due to the concern about an underestimated
CHAPTER 3. ROBUST HEDGING OF LONG-DATED LIABILITIES

bond premium, a robust investor requires more initial wealth to meet the shortfall target. Second, both policy-based yield curves are higher than the Vasicek curve when the spot rate is low (panel (a), (b)). When the spot rate is low, the present value of liability will be higher, hence the required initial wealth would be higher as well, in order to meet the long-term commitment. The Vasicek model is based on a zero shortfall requirement, and hence has a lower yield curve compared with the other two curves which allow for positive expected shortfall. We also find that the Vasicek yield curve overlaps the naive curve when the spot rate is high if $S$ is sufficiently small (panel (c)).

3.6 Conclusion

We have constructed an optimal bond portfolio to minimize the replication error of a long-dated cash flow in the presence of incomplete bond markets and model uncertainty. The robust policy is not always more cautious than a naive policy. With a fixed long-term commitment, we find that the robust portfolio suggests holding riskier positions.

The robust hedging strategy is a powerful method for pricing an ultra long-dated liability. Although the robust policy requires more initial wealth to guarantee a certain level of shortfall risk, it provides a more successful and resilient hedge, especially in a disordered environment where the bond premium is misspecified.

There are at least two directions in which future research along these lines would be interesting. The first is to consider mean-reversion parameter uncertainty. We find from calibration results that the mean reversion parameter is highly correlated with other parameters, such as the long-run mean parameter and the volatility term as well as the market-price-of-risk parameters. However, the layout of the optimization problem in this study does not allow for parameter uncertainty of other dimensions. Mean-reversion is especially relevant in interest rate models. The mean reversion parameter not only influences the expected value of the bond return, but also determines the volatility of the bond diffusion process for longer term bonds. Second, our robust model is applicable to any term structure model, and the Vasicek model used in this paper is simply a start. We can play our hedging problem under various reference models to see whether the model builder’s decision also matters.
3.7 Chapter 3 Appendix

3.7.1 Robust LSMC algorithm

In this section, we elaborate the numerical method used in this paper. We show how the regression-based method works in our robust hedging problem and we also analyze the accuracy of our algorithm. This section proceeds as follows, we first introduce the naive LSMC algorithm without considering model uncertainty. Second, we show how nature’s choice can be solved analytically. Third, we explain the robust LSMC algorithm. Last, we analyze the accuracy of our algorithm.

3.7.2 Naive LSMC

3.7.2.1 Grid Generation

The hedging period is from 0 to $T$. We partition $[0, T]$ into $m$ subintervals of length $\Delta t = \frac{T}{m}$. Hence the bond portfolio is rebalanced every $\Delta t$ unit of time. We start by simulating $N$ trajectories of $m$ time periods of spot rate under $\mathbb{P}$ measure using discrete Euler approximation.

$$r_{t+\Delta t} = r_t + \kappa (\theta - r_t) \Delta t + \sigma \sqrt{\Delta t} Z$$

with $Z$ a standard normal random variable. We indicate the spot rate at time $t$ in trajectory $i$ by $r_{i,t}$, $i = 1, \cdots, N$, $t = 0, \Delta t, \cdots, T - \Delta t, T$. At step 0, $r_0$ is also random with mean $\theta$ and volatility $\sigma \sqrt{\Delta t}$.

Next, we need to generate an $M$ dimensional grid of funding ratio. Funding ratio at time $t$ is defined as a fraction of instantaneous wealth against the hypothetical bond price maturing at $T$, $P(t, T)$. The funding ratio grids are indicated by $FR_j$, $j = 1, \cdots, M$. The reason for choosing funding ratio grid instead of wealth grid will be explained later in LSMC algorithm.

We also generate a small grid of testing portfolios, denoting $w_h$, $h = 1, \cdots, H$. We assume the portfolio grid is bounded between 0 and twice of delta hedge $\Delta = \frac{\partial P}{\partial \tau}$, $w_h \in [0, 2\Delta]$.

As a benchmark, we choose $\Delta t = 0.25$, $N = 10,000$, $M = 40$ and $H = 5$.

3.7.2.2 LSMC

The problem is solved by means of simulation based dynamic programming. We outline the general recursion. We first explain the naive case. In addition to the generated grids of the state variables, we also generate $N$ paths of gross bond returns based on point
estimates of bond premia $\hat{\Lambda}_0$ and $\hat{\Lambda}_1$,

$$R_{t+\Delta t} = 1 + \left( r(t_1) - B(t_2)\sigma \hat{\Lambda}_1 - B(t_2)\sigma \hat{\Lambda}_0 \right) \Delta t - B(t_2)\sigma \sqrt{\Delta t} Z$$

The gross bond return at time $t$ in trajectory $i$ is denoted by $R_{i,t}$.

**Time** $T - \Delta t$  The problem at time step $T - \Delta t$ can be summarized by

$$\min_{w_{T-\Delta t}} \mathbb{E} \left[ (1 - X_T)^+ \mid \mathcal{F}_{T-\Delta t} \right]$$

We first generate a time-dynamic wealth grid denoting $X_{j,T-\Delta t} = FR_j \times P(T - \Delta t, T)$, where $P(T - \Delta t, T)$ is the hypothetical bond price at time $T - \Delta$ maturing at $T$ depending on $r_{T-\Delta t}$. 

The reason we choose a fixed funding ratio grid instead of a fixed wealth grid is because our value function depends on terminal wealth $X_T$, keeping wealth grid fixed cannot guarantee us a reasonable range of terminal wealth, typically for long hedging horizon and when recursion approaches to time step 0.

Next we construct the realized terminal wealth under each simulated path $(i, j)$,

$$X_{ijh,T}(\mathcal{F}_{T-\Delta t}) = X_{j,T-\Delta} \left( (1 + r_{i,T-\Delta})(1 - w_h) + R_{i,T}w_h \right)$$

The realized objective function is $V_{ijh,T}(\mathcal{F}_{T-\Delta t}) = (1 - X_{ijh}(\mathcal{F}_{T-\Delta t}))^+$. Next, we regress the realized value function with a polynomial expansion of the state variables.

$$V_{ijh,T}(\mathcal{F}_{T-\Delta t}) = (a_0 + a_1 r_{i,T-\Delta} + a_2 FR_j + a_3 IFR_j + a_4 FR_j IFR_j)$$

$$+ (b_0 + b_1 r_{i,T-\Delta} + b_2 FR_j + b_3 IFR_j + b_4 FR_j IFR_j) w_h$$

$$+ (c_0 + c_1 r_{i,T-\Delta} + c_2 FR_j + c_3 IFR_j + c_4 FR_j IFR_j) w_h^2 + \epsilon_{T-\Delta t} \quad (3.13)$$

where $IFR$ is an indicator function, with $IFR_j = (1 - FR_j)^+$.

It is noticed that, along each path, the conditional variables are known hence the conditional expectation of the approximation regression is a quadratic function of the portfolio weight. Therefore, minimizing the conditional expectation boils down to minimizing a quadratic function of portfolio. We rewrite the conditional expectation of the value function as follows

$$\mathbb{E} [V_{ijh,T} \mid \mathcal{F}_{T-\Delta t}] = a_{ij} + b_{ij} w_{T-\Delta t} + c_{ij} w_{T-\Delta t}^2 \quad (3.14)$$
where

\[
\begin{align*}
    a_{ij}(F_{T-\Delta t}) &= a_0 + a_1 r_{i,T-\Delta t} + a_2 FR_j + a_3 IFR_j + a_4 FR_j IFR_j \\
    b_{ij}(F_{T-\Delta t}) &= b_0 + b_1 r_{i,T-\Delta t} + b_2 FR_j + b_3 IFR_j + b_4 FR_j IFR_j \\
    c_{ij}(F_{T-\Delta t}) &= c_0 + c_1 r_{i,T-\Delta t} + c_2 FR_j + c_3 IFR_j + c_4 FR_j IFR_j
\end{align*}
\]

The optimization problem boils down to solving for the root of Eq. (3.14),

\[
w_{ij,T-\Delta t}^* = -\frac{b_{ij}}{2c_{ij}}
\]

if \(c_{ij} > 0\).

**Time** \(t = T - 2\Delta t, \cdots, 0\) We now discuss the general recursion of all any other point in time. Suppose we have optimized the hedging policy as of time \(t + \Delta t\) onwards. The realized value function at time \(t\) is given by

\[
V_{ijh,T}(F_t) = \left[1 - X_{j,t}((1 + r_{i,t})(1 - w_h) + R_{i,t+\Delta t}w_h) \prod_{s=t+\Delta t}^{T-\Delta t} ((1 + r_{i,s})(1 - w_{h,s}) + R_{i,s+\Delta t}w_{h,s}) \right]^{-1}
\]

where \(X_{j,t} = FR_j \times P(t, T)\). Next, we approximate conditional expectations \([V_{ijh,T} \mid F_t]\) by functions of state variables and the testing portfolio

\[
E[V_{ijh,T} \mid F_t] = \beta'(X_{j,t}, r_{j,t}, w_h)
\]

Last, we rewrite the conditional expectation of each path as a quadratic function of portfolio \(w_t\) and solving for the root,

\[
\min_{w_t} E[V_{ijh,T} \mid F_t] = \min_{w_t} f(w_t)
\]

We need to re-calculate the dynamic allocation \(\frac{T}{\Delta t}\) times to retrieve the optimal decision now \(w_0\) that we are interested in. Along this way, we have also obtained all other optimal portfolios for different hedging horizons \(\tau < T\).

### 3.7.3 Closed-Form Nature’s Choice

Next, we consider the robust case when bond premia are misspecified. A straightforward but more complicated method is to repeat the naive algorithm over a set of testing bond premia taking into uncertainty set constraint. Then we obtain a set of naive optimal policies under each path conditional on a testing bond premium. Then we could find the optimal nature’s choice under each path either by grid search or using regression based method. However, neither methods are efficient nor accurate. If we use grid search method, we need to generate a fine grid of \(\lambda_0\) and \(\lambda_1\). This costs a huge computational
memory. Unlike the portfolio weight, the quadratic approximation does not work in the nature’s choice \((\lambda_0, \lambda_1)\) due to low \(R^2\).

A more efficient and accurate method we proposed is to approximate the expected shortfall by a function of bond return. Remark: this only works because we have simplified the problem so much that only \(\Lambda\) is uncertainty. We find that value function is monotonically decreasing with bond return. The approximation is sufficiently accurate since \(R^2\) is nearly one. Therefore, we transform a maximization problem to a minimizing bond return problem, which can be analytically solved by linear programming.

To test the validity of the new methods, we first generate \(L\) pairs of testing market price of risk (MPR) denoting \((\lambda_{0,l}, \lambda_{1,l})\), with \(l = 1, \cdots, L\) and \((\lambda_{0,l}, \lambda_{1,l}) \in S\). Under each pair of testing MPR, we simulate \(N\) paths of gross bond returns over \(m\) periods. The gross return at time \(t\) in path \(i\) under testing MPR \((\lambda_{0,l}, \lambda_{1,l})\) is denoted by \(R_{il,t}\).

At step \(T - \Delta t\), we construct realized value function \(V_{ijl,T}\) based on mul-specification of bond returns \(R_{il,T}\) using a random fixed portfolio weight, then we approximate the conditional value function as a function of \(R_{il,T}\) and fund ratio

\[
E \left[ V_{ijl,T} \left( w^{\tau + \Delta t}_l \right) \mid \mathcal{F}_{T - \Delta t} \right] = \beta_0 + \beta_1 R_{il,T} + \beta_2 R^2_{il,T} + \beta_3 FR_j + \beta_4 IFR_j
\]

The goodness of fit is larger than 0.99 regardless of the portfolio weight we choose. We find that the conditional expectation of value function is a strict downward sloping convex function of \(R_{il,T}\). This property holds for the entire hedging horizon if we recurse the algorithm backward till step 0.

Therefore, the global maximization problem boils down to minimizing bond return which is equivalent to minimizing bond premia since nature can only control over the drift term of bond diffusion process. Further, the ellipsoid uncertainty set \(S\) is convex, hence the minimum bond premium should locates on the ellipsoid.

The resulting nature’s optimal decision depends only on the instantaneous spot rate \(r_t\).

### 3.7.4 Robust LSMC

We can follow the naive LSMC algorithm to calculate the robust optimal portfolio expect that we first need to analytically solve for nature’s choice at each backward step in time. Hence at time \(t\), realized value function contains both optimal portfolios as well as nature’s choice of bond premia from steps onwards

\[
V_{ijh,T} = \left[ 1 - X_{jh} \left( (1 + r_{ih})(1 - w_h) + R^*_T \prod_{s=t+\Delta t}^{T-\Delta t} \left( (1 + r_{is})(1 - w^*_s) + R^*_s \right) \right) \right]^+ \]

70
where $R^*_{i,t+\Delta}$ indicates the optimal bond return at time $t$ on path $i$ with optimal MPR $\lambda^*_{i,t} + \lambda^*_1 r_{i,t}$

### 3.7.5 Goodness of Fit

We investigate the accuracy of our algorithm by means of $R^2$. The $R^2$s of parametrization regression at step $T - \Delta t$ are higher than 0.995 for both naive and robust case. The goodness of fit by construction, has to decay backwards of time, because we are accumulating cross sectional information over each time period onward. The quality of the global quadratic approximation depends on $R^2$ at first approximation step $T - \Delta t$ and the speed of decaying. The longer the hedging horizon is, the lower $R^2$ will be at time 0. i.e. If $T \leq 20$ years, $R^2$ at time 0 is higher than 95%. If we set investment horizon extremely long, $T = 80$ years, $R^2$ drops to 0.83 at time step 0. This is still reasonably high, since the last step cross-sectional regression contains 320 time steps of cross-section information. Grid sizes or rebalancing frequency do not influence the goodness of fit.
CHAPTER 3. ROBUST HEDGING OF LONG-DATED LIABILITIES

Figure 3.3: Optimal Bond Portfolio for Different Horizons
The figure plots 20-year bond weights against different investment horizons when current funding ratio is 80% (panel (a)) and 110% (panel (b)) for different hedging policies. The present spot rate is 2%. The weights are based on three different hedging policies: naive hedge, robust hedge and delta hedge. Results are based on 10,000 draws of predictive bond distribution.

(a) 80% Funding Ratio

(b) 110% Funding Ratio
Figure 3.4: Optimal Portfolio under Different Funding Ratio (1)
The figure plots the optimal portfolio of 20-year bond as a function of investment horizon and current funding ratio with spot rate equal to 2%.

(a) Naive Policy

(b) Robust Policy
Figure 3.5: Optimal Portfolio under Different Funding Ratio (2).
The figure summarizes the key insight of Figure 3.4 under three reasonable funding ratio levels.

(a) Naive Policy

(b) Robust Policy
Figure 3.6: Robust Policy under Different Spot Rate.
The figure plots naive (panel (a)) and robust (panel(b)) optimal bond portfolio as a function of spot rate and hedging horizon when current funding ratio is 80%.

(a) Naive Policy

(b) Robust Policy
Figure 3.7: Optimal Expected Shortfall (1).
The figure plots the expected shortfall as a function of hedging horizon and funding ratio under naive (panel (a)) and robust (panel(b)) policies when spot rate is 2%.

(a) Naive Expected Shortfall

(b) Robust Expected Shortfall
Figure 3.8: Optimal Expected Shortfall (2).
The figure plots the robust expected shortfall against the naive one as a function of
funding ratio when hedging horizon is 40 years (panel (a)) and when the horizon is 60
years (panel(b)).

(a) $T = 40$

(b) $T = 60$

The difference represents mispricing.

77
Figure 3.9: Robust Yield Curve

The figure plots policy-based yield curve against the benchmark Vasicek yield curve under different spot rate and shortfall target. The benchmark yield curve by definition has zero expected shortfall.

(a) $r_0 = 0.02$, $S = 0.01$

(b) $r_0 = 0.02$, $S = 0.05$

(c) $r_0 = 0.05$, $S = 0.01$

(d) $r_0 = 0.05$, $S = 0.05$
Chapter 4

How much should life-cycle investors adapt their behavior when confronted with model uncertainty?*

4.1 Introduction

How to make a rational lifetime investment decision in a world we do not understand? It is well documented in the optimal life-cycle investment literature that labor income and inflation play crucial roles in optimal asset allocation of individual investors. Several examples, e.g. Bodie et al. (1992) incorporate nontradeable labor income into the standard Merton (1969) model and find that extra demand for risky assets is seen when the investors are employed. Brennan and Xia (2002) study the impact of stochastic inflation on optimal asset allocation and find that index bonds should be part of the optimal hedging component in addition to the mean-variance tangency portfolio. In light of these studies, it is not surprising that investors have a strong wish to protect themselves against income and inflation shocks. However, as these studies are cast in a setting where the underlying model parameters are known, their ‘optimal’ portfolio choice can be suboptimal if the parameters are wrong.

In this chapter, I extend the classic life-cycle problem by allowing for model misspecification. I provide a robust optimal asset allocation strategy as well as a robust optimal lifetime consumption plan that can protect investors from misspecification of the income and inflation models. A robust investor considers alternative models that perturb the reference model. Following the Anderson et al. (2003) and Maenhout (2004) framework, an investor who is ambiguity averse looks for a robust optimal investment strategy that maximizes her working period utility function under a worst-case scenario.

*This is my job market paper.
Why inflation model uncertainty matters? There are at least three reasons. First, it is obvious that every consumer or any economic agent worries about the price of the commodity that keeps their life running and it is the inflation that drives the price of the supermarket goods. However, inflation rate is part of the government monetary policy which is filled with uncertainty. If an economic agent is afraid of the volatile and unpredictable stock market, she can simply stay away from it by keeping her money in cash or investing in the bond market. However, different from the noisy stock market, inflation uncertainty is not avoidable since everyone has to go to the supermarket. Second, although the rate of price inflation is less volatile than equity returns, Stock and Watson (2007) argue that US inflation has become harder to forecast. Third, for long-term individual investors (young cohorts), inflation risk cannot be fully hedged for the long run, because the real-bond market is incomplete. Consequently, a misspecified inflation model can jeopardize optimal investment and consumption.

The financial reference model in this chapter relates closely to Koijen et al. (2010). I follow Hamilton and Wu (2012)’s approach for estimating the two-factor affine term structure, model and use the generalized method of moments GMM method to estimate the stock return process as well as the inflation dynamics. Both methods belong to the class of minimum distance estimation. The reference financial model is based on monthly US return data from 1961 to 2013.

I follow the Hansen and Sargent framework for measuring aversion to uncertainty. The perturbation between the unobservable true model and the reference model is captured by a relative entropy parameter, which penalizes the failure of specification. The ambiguity aversion pertains to the first moment of return processes. In other words, the robust investor only worries about a misspecified expected inflation rate and / or a wrong expected income growth rate. Nature has limited freedom to distort the two expected returns. The two drift distortions jointly satisfy a chi-squared statistical constraint that reflects the investor’s confidence about the underlying model. The same constraint indirectly helps to derive the feasible values of the relative entropy parameter.

I solve a max-min robust dynamic programming problem analytically and derive the robust optimal investment and consumption decisions in closed form. To analyze the impact of macroeconomic uncertainty on optimal lifetime investment decisions, I proceed in two steps. First, I solve the robust life-cycle problem in the absence of labor income, which means that the investor is confronted only with inflation rate uncertainty. The optimal portfolio is independent of instantaneous wealth, but has a strong horizon effect. I find that the preference for robustness induces a stronger demand for long-term bonds but a more conservative position in stocks. Further, nature’s optimal decision on the inflation rate distortion reflects a robust investor’s fear of an underestimated inflation.
4.1. INTRODUCTION

rate.

Second, when I add labor income to the life-cycle problem, human capital becomes part of the optimal investment strategy. The labor income process is based on Viceira (2001)'s framework, but in a continuous time setting. In addition, I allow for hedging against income risk by assuming that income shocks are partially correlated with the financial market. The remaining portion of income risks is assumed to be idiosyncratic.

I find that, first, the optimal hedging strategy is highly dependent on the ratio of human capital to wealth, and exhibits stronger horizon effects. Therefore, a larger risk exposure is expected for a young investor, because the human capital for a long-term investor is much larger than her financial wealth. Secondly, I also find that in the presence of human capital, the risk-shifting between different asset classes induced by a preference for robustness depends on state variables and the investor's risk aversion. A risk-averse investor is more in favour of long-term bonds. Nature's distortions lead to a lower expected income growth rate and a higher inflation rate.

The contribution of this research is twofold. First of all, I integrate and extend the literature on both the life-cycle asset allocation problem and the robust hedging problem. Most robust optimal portfolio studies that are based on the Hansen and Sargent (2007) and Anderson et al. (1999) framework only consider one source of misspecification, namely stock return ambiguity. None of these studies (such as Maenhout (2004), Uppal and Wang (2003) and Branger et al. (2013)) consider monetary policy uncertainty and/or income process misspecification. Why does the specification of the inflation process matter, or why do macroeconomic models in general matter? Because robust investors can avoid stock return misspecification by reducing their equity exposure or in an extreme case simply stop participating the stock market, but it is impossible for them to avoid inflation rate estimation error or unexpected income shocks. Further, I consider two sources of ambiguities jointly.

The second contribution of this research is that I develop a new method for calibrating the ambiguity aversion parameter. The new approach fills the gap between robust asset pricing and financial econometrics. First, I solve for the optimal perturbations as a function of an endogenous relative entropy parameter, which is a linear function of the ambiguity aversion parameter. Second, under the assumption that the estimation errors are asymptotically normal, together with the standard probability theory that a standardized normal distribution square becomes a Chi-squared distribution, I derive a statistical constraint of the drift distortions at the second moment. The methodology resembles the Bayesian paradigm but in an opposite way, for I do not consider the reference model obtained from the historical data and the true model. Lastly, I simulate a large number of robust optimal solutions under a set of trial ambiguity aversion parameters. Among
these trial series, the largest, under which all simulations are located within the statistical boundary obtained from the second step, is the feasible upper bound of the ambiguity aversion parameter.

This study relates to the optimal asset allocation problem for long-term investors. It is demonstrated in Merton (1971) that time-independent myopic demand is not sufficient to hedge the investment opportunity for a long-term investor when the interest rate is stochastic. A large number of studies such as Brennan and Xia (2002), Sørensen (1999), Wachter (2003) and Wachter (2010) have argued that when investment opportunities are time-varying, long horizon investors tend to increase risk exposure to long-term bonds. Furthermore, these studies also demonstrate a strong horizon effect. The longer the hedging horizon, the larger the required position in long-term bonds. However, none of these studies take labor income into account, and none of them consider the problem of parameter uncertainty either.

Pioneer work by Bodie et al. (1992) emphasizes the importance of human capital for optimal asset allocation over the lifespan, and they argue that the young should borrow from their future human capital to invest in risky assets. More recent studies such as Viceira (2001) and Cocco et al. (2005) Koijen et al. (2010), Benzoni et al. (2007) and Munk and Sørensen (2010) consider the impact of the stochastic income process on the optimal investment decision. However, none of these studies take income model uncertainty into consideration.

This research also relates to studies of robust dynamic optimal portfolio choice. Maenhout (2004) extends Merton (1969)'s model by allowing for stock return misspecification, and he finds that ambiguity aversion reduces the investor's demand for risky asset. Maenhout (2006) again considers the expected equity return uncertainty on the basis of a mean-reversion market price of risk. Branger et al. (2013) extend Maenhout (2006)'s study in an incomplete equity market by introducing a non-tradable risk factor, and they also allow for learning. They find that both ambiguity aversion and learning matter for a robust investor. Learning is not considered in this paper, as I assume that the data generation process does not stay the same forever.

The rest of the chapter proceeds as follows. In Section 4.2, I first introduce the reference financial and macroeconomic models. Next, I show the drift distortions as well as their influence on the underlying model. The robust optimization problem is also introduced in this section. In Section 4.4, I derive the closed form solution for investors under different preferences. Section 4.5 estimates the reference model using US data. In Section 4.6, I develop a new method to quantify the ambiguity aversion parameter. Section 4.7 presents the numerical solution for the robust life-time investor and Section 4.9 gives the conclusion.
4.2 The Model

4.2.1 Financial Market

The financial model is closely related to Sangvinatsos and Wachter (2005) and Koijen et al. (2010). I assume that a life-cycle investor can invest in a stock (index), a long-term nominal bond and a nominal money market account. The instantaneous nominal risk free rate, $r_t$, is assumed to be affine in two state variables $X_{1t}$ and $X_{2t}$.

$$r_t = \delta_0 + \delta^\top X_t, \quad \delta_0 > 0. \quad (4.1)$$

where $\delta_0$ is scalar and $\delta \in \mathbb{R}^2$. The latent factors $X_t = \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix}^\top$, governing yields are characterized by an Ornstein-Uhlenbeck process with mean-reverting drift around zero under the physical measure,

$$dX_t = -\kappa X dt + \sigma_X^\top dZ, \quad (4.2)$$

where $\kappa \in \mathbb{R}^{2 \times 2}$. Following Dai and Singleton (2000), I assume $\sigma_X$ equal to an identity matrix with $\sigma_X = \begin{bmatrix} I_{2 \times 2} & 0_{2 \times 2} \end{bmatrix}^\top$. Let $dZ \in \mathbb{R}^4$ represent a vector of independent risk drivers following Brownian motions. The nominal pricing kernel is given by

$$d\phi = -r_t dt - \Lambda_t^\top dZ, \quad (4.3)$$

where market price of risk $\Lambda_t$ is affine in $X_t$

$$\Lambda_t = \lambda_0 + \lambda_1 X_t, \quad (4.4)$$

with $\lambda_0 \in \mathbb{R}^4$ and $\lambda_1 \in \mathbb{R}^{4 \times 2}$. Denote $P(t, T)$ as the nominal price at time $t$ of a zero-coupon bond maturing at time $T = t + \tau$ with a nominal payoff of 1. Following Duffie and Kan (1996), I conjecture that bond price is exponential affine in the state variables,

$$P(t, T) = \exp(B(\tau)^\top X_t + A(\tau)), \quad (4.5)$$

where $B(\tau) \in \mathbb{R}^2$ and $A(\tau)$ is a scalar. Therefore, the corresponding yield is

$$Y^\tau_t = a(\tau) + b(\tau)^\top X_t, \quad \text{with} \quad a(\tau) = -\frac{A(\tau)}{\tau}, \quad b(\tau) = -\frac{B(\tau)}{\tau}. \quad (4.6)$$
CHAPTER 4. HOW MUCH SHOULD INVESTORS ADAPT?

The expression of $B(\tau)$ and $A(\tau)$ can be solved in a closed form and is shown in Appendix 4.10.1. Applying Itô’s lemma, bond price dynamics follow

$$
\frac{dP}{P} = (r_t + B(\tau)^\top \sigma_X^t A_t) dt + B(\tau)^\top \sigma^t dZ
$$

(4.7)

The commodity price level $\Pi_t$ follows a diffusion process

$$
\frac{d\Pi_t}{\Pi_t} = \pi_t dt + \sigma_{\Pi}^t dZ,
$$

with $\sigma_{\Pi} \in \mathbb{R}^4$

(4.8)

where the expected inflation, $\pi_t$, is assumed to be affine in state variables as well

$$
\pi_t = \xi_0 + \xi^\top X_t,
$$

with $\xi \in \mathbb{R}^2$.

(4.9)

I assume the dynamics of stock price is given by

$$
\frac{dS_t}{S_t} = (r_t + \eta_S) dt + \sigma_S^t dZ
$$

(4.10)

where $\sigma_S \in \mathbb{R}^4$. The following constraint for expected excess stock return holds

$$
\eta_S = \sigma_S^t A_t
$$

(4.11)

4.2.2 Investor’s Labor Income

The nominal labor income model is based on Viceira (2001). In addition, I assume income shocks that are partially correlated with the financial market introduced in Section 4.2.1. The investor’s income dynamics at time $t$ are given by

$$
dY = \left( g + \frac{\sigma_y^2}{2} \right) Y dt + \sigma_y Y \rho_{\rho_y}^t dZ + \sigma_y Y \sqrt{1 - ||\rho_y||^2} dZ_Y
$$

(4.12)

where $Y$ is the instantaneous nominal income, $\rho_y \in \mathbb{R}^4$ and $dZ_Y$ represents idiosyncratic risk.

4.3 Basic Optimal Portfolio Model

I first solve the standard life cycle problem in the absence of model misspecification. Consider an investor who has constant relative risk aversion utility function (CRRA).
Her preferences before retirement can be represented as follows:

\[
\max_{C_t,x_t} \mathbb{E}_0 \left[ \int_0^T \frac{\exp(-\beta t)}{1-\gamma} \left( \frac{C_t}{\Pi_t} \right)^{1-\gamma} dt + \frac{\varphi \exp(-\beta T)}{1-\gamma} \left( \frac{W_T}{\Pi_T} \right)^{1-\gamma} \right]
\] (4.13)

where \( \beta > 0 \) is the subjective discount factor. The post-retirement consumption and investment decision is not included in agent’s planning horizon. Instead, I introduce a bequest function scaled by \( \varphi \) to capture the retirement consumption. In particular, I assume that the value of bequest function is equivalent to the utility value of annuitization at retirement, as such, the value of \( \varphi \) can be calibrated.\(^2\)

The asset menu includes one stock, one nominal long-term bond with maturity \( \tau \) and a short term nominal government bond. The nominal wealth follows the diffusion process

\[
dW = \left( x_t^\top (\mu_t - \iota r_t) + r_t \right) W dt - C dt + Y dt + W x_t^\top \sigma^\top dZ
\] (4.14)

where \( x_t \) denotes the fraction of nominal wealth invested in difference assets,

\[
\mu_t = \begin{pmatrix} B(\tau)^\top \sigma_{XA} \\ \sigma_{XA} \end{pmatrix} + \iota r_t
\] (4.15)

\[
\sigma^\top = \begin{pmatrix} B(\tau)^\top \sigma_X \\ \sigma_s \end{pmatrix}
\] (4.16)

with \( \mu_t \in \mathbb{R}^2 \) and \( \sigma \in \mathbb{R}^{4 \times 2} \).

Denote the value function at time \( t \) by \( J(W,Y,\Pi,X,t) \), I omit the time subscripts for notation convenience. The Hamilton-Jacobi-Bellman HJB equation of the dynamic optimization problem (4.13) is given by

\[
0 = \max_{C_t,x_t} \left[ \frac{1}{1-\gamma} \left( \frac{C_t}{\Pi_t} \right)^{1-\gamma} - \beta J + \mathcal{D}(C_t, x_t) J(W,Y,\Pi,X,t) \right]
\] (4.17)

where \( \mathcal{D}(C_t, x_t) J(W,Y,\Pi,X,t) \) represents the remaining body of the HJB equation besides the existing two terms. The full expression of the HJB equation is explained in Appendix 4.10.3.4.

---

\(^2\) Along the line with Koijen et al. (2010)’s idea, the scaling factor \( \varphi \) is defined as

\[
\varphi = A_T^{-1} \int_T^{85} \exp(-\beta t) dt
\]

where \( A_T \) represents the annuity payment at retirement, and I assume a survival probability of 1 up to age of 85.
CHAPTER 4.  HOW MUCH SHOULD INVESTORS ADAPT?

4.3.1 Monetary Policy Uncertainty and Income Uncertainty

Suppose the expected inflation rate $\pi_t$ and the expected income growth rate $g$ are mis-specified. A robust life cycle investor, who suspects the inaccuracy of her macroeconomic models, takes (4.8) and (4.12) as an approximation and considers alternative models. Adopting Anderson et al. (2003) framework, a perturbed model can be expressed with an added drift term on the Brownian motions $e^\top dZ$ and $dZ_Y$ of the reference model, with $e = (0, 0, 1, 0)^\top$. Therefore an alternative inflation diffusion process is given by

$$\frac{d\Pi}{\Pi} = \pi_t dt + \sigma_\Pi (dZ + e_{\gamma_1,t} dt)$$  \hspace{1cm} (4.18)

and a perturbed income diffusion process is

$$dY = \left( g + \frac{\sigma_y^2}{2} \right) Y dt + \sigma_y \rho_{\gamma_1} (dZ + e_{\gamma_1,t} dt) + \sigma_y \sqrt{1 - ||\rho_y||^2} (dZ_Y + \gamma_2 dt)$$  \hspace{1cm} (4.19)

Each alternative macroeconomic model is characterized by a joint stochastic process $(\gamma_1,t, \gamma_2,t)$ and is specified under a candidate measure which is equivalent to the reference measure $P$.

Nature is entitled to decide the sign and the size of the drift distortions. Consider an investor with preferences (4.13), who can not observe the true model and fears misspecification. The most detrimental nature’s decision to this investor, is when the reference model (4.9) underestimates the expected inflation while in contrast, (4.12) overestimates the expected income return. Both scenarios imply that the investor’s realized optimal preference is much lower than her expected level based on the reference model.

4.3.2 Robust Optimal Portfolio Model

The robust investor searches for a more resilient consumption and investment strategy to protect herself against model fragility. Her preferences are assumed to be

$$\min_{\{\gamma_1,t, \gamma_2,t\}} \max_{C_t,x_t} \left[ \int_0^T \exp(-\beta t) \left( \frac{C_t}{\Pi_t} \right)^{1-\gamma} dt + \frac{\varphi \exp(-\beta T)}{1-\gamma} \left( \frac{W_T}{\Pi_T} \right)^{1-\gamma} \right]$$  \hspace{1cm} (4.20)

The perturbation of the inflation process (4.18) triggers the drift distortion of the nominal wealth dynamics

$$dW = \left( x_t^\top (\mu_t - r_t) + r_t \right) W dt - C dt + Y dt + W x_t^\top \sigma^T (dZ + e_{\gamma_1,t} dt)$$  \hspace{1cm} (4.21)

\[^3\text{Define } e \text{ as a weighting vector controlling the fragility of each dimension of the state. As in this paper I focus on the inflation model uncertainty, I put a full weight on the inflation risk dimension. However, it is important to note that by taking } e = (0, 0, 1, 0)^\top, \text{ the drift distortions may also influence asset prices.}\]
Following Anderson et al. (2003), the robust HJB equation is given by

$$0 = \min_{(\gamma_1, \gamma_2)} \max_{C_t, x_t} \left\{ \gamma_1 \left( \frac{C}{\Pi} \right)^{1-\gamma_1} - \beta J + D(C_t, x_t) J(W, Y, \Pi, X, t) + J_{W} W x^\top \sigma e_{\gamma_1} + J_{Y} Y Y^\top (\sigma y \rho^\top y e_{\gamma_1} + J_{\Pi} \Pi \sigma_{\Pi} e_{\gamma_1} + \frac{1}{2 \Psi} (\gamma_1^2 + \gamma_2^2)) \right\}$$

(4.22)

In addition to the reference-model based HJB equation (4.17), there are four additional terms. The first three additional terms reflect the adjustment of the Bellman equation due to the additional drift components. The last term of (4.22) can be considered as a penalty function, to penalize those alternative models located too far away from the reference model. Hence, nature’s decision is constrained.

The penalty term is scaled by a non-negative function $\Psi = \Psi(W, Y, \Pi, X, t)$ which captures investor’s preference for robustness. If $\Psi \to \infty$, the penalty term vanishes. As a result, nature has tremendous freedom to choose drift distortion and the investor has almost no confidence over the underlying model. In contrast, when $\Psi \to 0$, the penalty function becomes infinitely large. Hence, it is too costly for nature to make any alternative decision deviating from the reference model. In this case, the investor is extremely optimistic towards the reference model.

In order to obtain an explicit solution, homotheticity of the HJB equation is usually required under CRRA preferences. Inspired by Maenhout (2004), I assume the penalty term $\Psi$ is a scaled function of the indirect utility $J(W, Y, \Pi, X, t)$, which takes the form:

$$\Psi(W, Y, \Pi, X, t) = \frac{\theta}{(1-\gamma) J(W, Y, \Pi, X, t)}$$

(4.23)

where $\theta$ represents investor’s ambiguity aversion level which is non-negative and is assume constant over time. The bigger the value of $\theta$ is, the more pessimistic the investor is towards the underlying model.

### 4.4 Optimal Portfolio Choice

#### 4.4.1 Model 1: No Labor Income, No Model Misspecification

I first consider a benchmark problem without labor income ($Y_t = 0$). This is a special case of the model of Sangvinatsos and Wachter (2005) and can be solved explicitly in two different ways. In this paper, I obtain the optimal investment and consumption by means of solving the Bellman equation. An alternative method would be to use martingale technique of Cox and Huang (1989). Studies, such as Brennan and Xia (2002), Wachter...
CHAPTER 4. HOW MUCH SHOULD INVESTORS ADAPT?

(2002) and Sangvinatsos and Wachter (2005) use this method to solve their dynamic portfolio choice problem. However, the martingale technique is not applicable when the market price of risk is not unique, hence cannot be used to solve the robust optimization problem (4.20).

**Proposition 4.4.1.** The indirect utility function for an investor with preferences represented by (4.13) without labor income and without borrowing constraint\(^4\) is given by

\[ J_{M1} (W, \Pi, X, t) = \frac{1}{1 - \gamma} b_{M1} (X, t)^{\gamma} \left( \frac{W}{\Pi} \right)^{1-\gamma} \]  

(4.24)

with

\[ b_{M1} (X, t) = \exp \left\{ \frac{1}{\gamma} \left( \frac{1}{2} X^\top \Gamma_{1M1} (\tau) X + \Gamma_{2M1} (\tau) X + \Gamma_{3M1} (\tau) \right) \right\} \]  

(4.25)

where \( \tau = T - t \) represents the remaining investment horizon before retirement and the deterministic functions \( \Gamma_{1M1} \in \mathbb{R}^{2x2}, \Gamma_{2M1} \in \mathbb{R}^{1x2}, \) and scalar \( \Gamma_{3M1} \) solve a system of ordinary differential equations stated in Appendix 4.10.3.1.

The optimal consumption is \( C_{M1} = \frac{W}{b_{M1}} \), and the optimal portfolio is given by

\[ x_{M1} = \frac{1}{\gamma} \left( \sigma^\top \sigma \right)^{-1} (\mu - \iota r) + \frac{\gamma - 1}{\gamma} \left( \sigma^\top \sigma \right)^{-1} \left( \sigma^\top \sigma_{\Pi} \right) + \left( \sigma^\top \sigma \right)^{-1} \left( \sigma^\top \sigma_{X} \right) b_{M1} \frac{b_{M1}^\top X}{b_{M1}} \]  

(4.26)

where \( \frac{b_{M1}^\top X}{b_{M1}} = \frac{1}{\gamma} \left( \Gamma_{1M1} (\tau) X + \Gamma_{2M1} (\tau) \right) \).

The optimal portfolio of **Model 1** consists of three components. The first component is simply Merton’s solution that solves for the instantaneous mean-variance efficiency scaled by a risk aversion function. The second component means to replicate the unexpected inflation so as to adjust the nominal mean-variance portfolio to real terms. The first two terms together are also known as “myopic demand”. Without labor income, the myopic demand is independent of both instantaneous wealth and hedging horizon. The last term represents the hedging demand for long-term nominal bond. It is demonstrated in Sangvinatsos and Wachter (2005) that a time-varying bond risk premium can generate hedging demand for long-term bond for those long-run investors. The third component depends on investor’s remaining investment horizon and the value of state variables.

\(^4\)I allow for short positions and borrowing from the future income.
4.4. OPTIMAL PORTFOLIO CHOICE

4.4.2 Model 2: No Labor Income, Under Model Misspecification

Next, I investigate the optimal solution for a robust investor. Consider a special case of (4.20) subject to (4.18) and (4.21), disregarding the income effect $Y = 0$. The absence of labor income implicitly wipes away the effect of distortion drift term $\gamma_2$. Nature can only adjust the expected inflation rate to minimize investor’s preferences. **Model 2** is the robust version of **Model 1**.

**Proposition 4.4.2.** Under the assumptions stated above, the optimal consumption is $C^{M2} = \frac{W}{b^{M2}}$, where function $b^{M2}$ takes the same structure as (4.25), while with the quadratic, linear and constant coefficients of the exponential function replaced by $\Gamma_1^{M2}$, $\Gamma_2^{M2}$ and $\Gamma_3^{M2}$ respectively.

The robust optimal portfolio for an investor with preference (4.20) is given by

$$x^{M2} = \left(\gamma \sigma^\top \sigma + \theta \sigma^\top \mathbf{e} \mathbf{e}^\top \sigma\right)^{-1} \left\{ (\mu - \nu r) + (\gamma - 1) \sigma^\top \sigma_{\Pi} + \theta \sigma^\top \mathbf{e} \mathbf{e}^\top \sigma_{\Pi} ight. \\
+ \left. \sigma^\top \sigma_X \left[ \Gamma_1^{M2}(\tau) X_\tau + \Gamma_2^{M2}(\tau) \right] \right\}$$  (4.27)

where $\Gamma_1^{M2}$ and $\Gamma_2^{M2}$ satisfying a system of ODE stated in Appendix 4.10.3.2.

The worst case distortion is given by

$$\gamma_1^{M2} = -\theta \left( x^{M2^\top} \mathbf{e} - \sigma_{\Pi} \right)$$  (4.28)

Without the preference for robustness $\theta = 0$, $x^{M2}$ is equivalent to $x^{M1}$. The robust optimal portfolio of **Model 2** also contains three components. The first component

$$\left(\gamma \sigma^\top \sigma + \theta \sigma^\top \mathbf{e} \mathbf{e}^\top \sigma\right)^{-1} (\mu - \nu r)$$

modifies Maenhout (2004)’s solution in two ways. First, instead of a constant investment opportunity, the investor faces a time-varying nominal myopic demand. Second, rather than fearing for a misspecified stock return, I assume the robust investor only worries about inflation rate misspecification, while the stock and bond return processes are correctly specified.

The adjustment term $\theta \sigma^\top \mathbf{e} \mathbf{e}^\top \sigma$, penalizes the misspecification of the inflation rate. A robust investor increases her risk-aversion level by a unit of $\theta$ in the inflation risk dimension. The penalty parameter $\theta$ reduces the demand for nominal mean-variance hedging. The second component

$$\left(\gamma \sigma^\top \sigma + \theta \sigma^\top \mathbf{e} \mathbf{e}^\top \sigma\right)^{-1} \left( (\gamma - 1) \sigma^\top \sigma_{\Pi} + \theta \sigma^\top \mathbf{e} \mathbf{e}^\top \sigma_{\Pi} \right)$$
CHAPTER 4. HOW MUCH SHOULD INVESTORS ADAPT?

is an ambiguity-adjusted inflation replication portfolio. When the ambiguity parameter goes to infinity $\theta \to \infty$, this term converges to $(\sigma^\top \mathbf{e} \mathbf{e}^\top \sigma)^{-1} \sigma^\top \mathbf{e} \mathbf{e}^\top \sigma \Pi$ and the first component vanishes. It indicates that when inflation rate is extremely hard to estimate, the most robust investment strategy is to become inflation risk neutral (in the absence of time-varying bond premia). As the penalty function occurs on both numerator and denominator of the fraction, the joint effect on inflation adjustment is unclear. The last part

$$(\gamma \sigma^\top \sigma + \theta \sigma^\top \mathbf{e} \mathbf{e}^\top \sigma)^{-1} \sigma^\top \sigma X \left[ \Gamma_{M1}^M (\tau) X_t + \Gamma_{M2}^M (\tau) \right]$$

shows the adjustment of hedging demand on long-term bond. As the penalty parameter $\theta$ also plays a role in function $\Gamma_{M1}^M$ and $\Gamma_{M2}^M$, the quantitative influence on hedging demand of the long-term interest rate risk is not clear at this stage.

It is also interesting to note that the optimal nature’s decision $\gamma_{M1}^M$ depends on the investor’s investment decision. As the impact of ambiguity aversion on the portfolio choice is not obvious, it is hard to detect nature’s intention.

4.4.3 Model 3: With Labor Income, No Model Misspecification

Classic optimal life-cycle investment studies, such as Merton (1969) and Samuelson (1969) show that the optimal asset allocation is independent of age. However, pioneer work by Bodie et al. (1992) show that young should go short and investor’s life-time risk exposure is a diminishing function of age if the non-tradeable human capital is taken into account.

In Model 3, I assume investor has preferences (4.13) subject to both financial wealth constraint (4.14) and human capital constraint (4.12). This problem is comparable to Munk and Sørensen (2010)’s work but with two small variations. First, the inflation risk is not considered in their model and the investor’s asset menu is assumed in real terms. Second, they employ one-factor Vasicek (1977) model instead of a multi-factor affine model.

Human capital is defined as the cumulative discounted remaining labor income stream under the risk neutral measure. Assuming that the income stream is locally risk free, $\sigma_y \equiv 0$,$^5$ the market value of the income stream at time $t$ over the working period $[t, T]$ is defined as

$$H_t = H(Y_t, X_t, t) = \mathbb{E}_t^Q \left[ \int_t^T Y_u \exp \left( - \int_t^u r_s ds \right) du \right] \quad (4.29)$$

$^5$This assumption is crucial and is also used in Munk and Sørensen (2010). If $\sigma_y$ is not spanned, then the market is incomplete, due to the non-traded risk driver $dZ_y$. As the result, the risk neutral measure is no longer unique and human capital cannot be priced.
4.4. OPTIMAL PORTFOLIO CHOICE

The explicit expression of \( H_t \) is shown in the following proposition. The proof is given in Appendix 4.10.3.3.

**Proposition 4.4.3.** Under the assumption that the idiosyncratic risk is zero, the nominal human capital is given by

\[
H(Y, X, t) = Y_t M(X, t) = Y_t \int_t^T \exp \left( M_1 (u - t) + M_2 (u - t) X_t \right) du \quad (4.30)
\]

The expression of \( M(X, t) \) is derived in Appendix 4.10.3.3. Proposition 4.4.3 indicates that human capital is a linear function of instantaneous income \( Y_t \). The value of human capital depends also on age. The younger the investor is, the longer her future expected labor income stream will be.

**Proposition 4.4.4.** The indirect utility function for an ambiguity-neutral investor with preference (4.13) subject to (4.12) and (4.14) is given by

\[
J^{M_3}(W, Y, \Pi, X, t) = \frac{1}{1 - \gamma} M_3(X, t)^\gamma \left( W + H \right)^{1-\gamma} \quad (4.31)
\]

where the human capital \( H \) follows Proposition 4.10.3.3, under a special case when state variables \( X \) take values of their unconditional means in function \( M(X, t) \). Function \( b^{M_3}(X, t) \) takes the same form as (4.25), but with the coefficients of the polynomial function replaced by \( \Gamma^{M_3}_1(\tau), \Gamma^{M_3}_2(\tau), \) and \( \Gamma^{M_3}_3(\tau) \). See Appendix 4.10.3.4 for full expression.

For an ambiguity neutral investor, the optimal portfolio weight is

\[
x^{M_3} = \frac{W + H}{\gamma W} \left( \sigma^\top \sigma \right)^{-1} (\mu - \nu \cdot \tau) - \frac{H}{W} \left( \sigma^\top \sigma \right)^{-1} \sigma^\top \rho_\sigma \sigma_\gamma + \frac{\gamma - 1}{\gamma} \frac{W + H}{W} \left( \sigma^\top \sigma \right)^{-1} (\sigma^\top \sigma_\Pi) + \frac{W + H}{W} \left( \sigma^\top \sigma \right)^{-1} (\sigma^\top \sigma_X) b^{M_3}_X b^{M_3}_X \quad (4.32)
\]

where \( b^{M_3}_X = \frac{1}{2} (\Gamma^{M_3}_1(\tau) X + \Gamma^{M_3}_2(\tau)) \). The remaining part of wealth, \( 1 - x^\top t \), is invested in nominal short-term bond. The optimal nominal consumption equals

\[
C^{M_3} = \frac{W + H}{b^{M_3}}.
\]

Note that in Proposition 4.4.5, human capital \( H \) takes an approximate form. I assume that \( M(X, t) \approx M(\bar{X}, t) \) with unconditional mean \( \bar{X} = 0 \), because numerical experiments reveals that \( M \) is not very sensitive to the state variable. Munk and Sørensen (2010) also use this assumption in the numerical analysis.

The human-capital adjusted optimal portfolio (4.32) differs from the benchmark case (Model 1) in two aspects. First, the solution is no longer instantaneously wealth neutral,
CHAPTER 4. HOW MUCH SHOULD INVESTORS ADAPT?

but instead each of the three hedging components discussed in Model 1 is scaled by a total wealth over financial wealth ratio, where total wealth is the sum of financial wealth \( W \) and human capital \( H \). It is demonstrated in Bovenberg et al. (2007) that human capital plays a crucial role in life-cycle financial planning. The ratio \( \frac{W}{W+H} \) is downward sloping as a function of age. For the human capital dominates the total wealth and it is a decreasing function over age. As a result, the optimal portfolio \( x^{M3} \) also decreases over age.

The second difference comes from the additional income hedging component

\[- \frac{H}{W} \left( \sigma^\top \sigma \right)^{-1} \sigma^\top \rho_y \sigma_y.\]

As income risk \( dZ_Y \) is assumed correlated to the financial shocks, this hedging portfolio, scaled by human capital over wealth ratio, reflects the hedging demand against non-idiosyncratic income risks.

4.4.4 Model 4: With Labor Income and Model Misspecification

**Proposition 4.4.5.** Under the assumption that the human capital \( H \) is ambiguity neutral and is state variable insensitive, the indirect utility function of a robust investor with preferences represented by (4.20) subject to (4.19), (4.18) and (4.21) is given by

\[
J^{M4} (W,Y,\Pi, X, t) = \frac{1}{1 - \gamma} b^{M4} (X, t)^\gamma \left( \frac{W + H}{\Pi} \right)^{1-\gamma} \tag{4.33}
\]

where \( b^{M4} \) is an exponential affine function of \( X \), taking the same form as (4.25) while with coefficients replaced by \( \Gamma^{M4}_1 (\tau), \Gamma^{M4}_2 (\tau) \) and \( \Gamma^{M4}_3 (\tau) \) respectively. The robust optimal portfolio is given by

\[
x^{M4} = \left( \frac{\gamma W}{W + H} \sigma^\top \sigma + \frac{\theta W}{W + H} \sigma^\top \epsilon \epsilon^\top \sigma \right)^{-1} \left\{ (\mu - \varepsilon r) + \left( (\gamma - 1) \sigma^\top \Pi \right) + \theta \sigma^\top \epsilon \epsilon^\top \sigma \right\}
- \left( \frac{\gamma H}{W + H} \sigma^\top \rho_y \sigma_y + \frac{\theta H}{W + H} \sigma^\top \epsilon \epsilon^\top \rho_y \sigma_y \right) + \gamma \sigma^\top \sigma X b^{M4} X \tag{4.34}
\]

where \( \frac{b^{M4}_1}{\sigma^\top \sigma} = \frac{1}{\gamma} \left( \Gamma^{M4}_1 (\tau) X + \Gamma^{M4}_2 (\tau) \right) \). The optimal consumption takes the same form as Model 3 while applying indirect utility function \( J^{M4} \). Nature’s optimal decision on the two drift distortions are

\[
\gamma_1^{M4} = -\theta \left( \frac{W}{W + H} x^{M4} \epsilon + \frac{H}{W + H} \sigma^\top \rho_y \epsilon - \sigma^\top \Pi \right) \tag{4.35}
\]

\[
\gamma_2^{M4} = -\theta \left[ \frac{H}{W + H} \sigma^\top \rho_y \sqrt{1 - ||\rho_y||^2} \right] \tag{4.36}
\]
Appendix 4.10.3.5 provides the proof.

The robust optimal portfolio contains four hedging components. The first two components of $x^{M4}$ are equivalent to the robust myopic demand of Model 2 but scaled by the labor-income adjusted scale $\frac{W+H}{W}$. Therefore, instead of age independent portfolios, these two hedging portfolios are aggregately decreasing with age. With age given, the presence of preference for robustness reduced the hedging demand for nominal mean-variance portfolio, however, the impact of inflation hedging portfolio is unclear but depends on the input parameters.

The simplified third component is

\[-H \frac{1}{W} \left( (\gamma\sigma^\top\rho + \theta\sigma^\top e\sigma) - 1 \right) \left( \gamma\sigma^\top\rho\rho + \theta\sigma^\top e\rho\rho \right)\]

It captures the hedging demand for tradeable income risk in the present of model mis-specification. As the entropy parameter $\theta$ brings a structure change on both sides of the income-hedge portfolio fraction, the hedging demand for this term is not clear.

Nature's decision on inflation rate drift distortion contains three mixed-sign components and is also portfolio decision dependent. Hence it is difficult to foretell the sign of $\gamma_1^{M4}$. However, the sign of the other drift distortion term $\gamma_2^{M4}$ is obvious. Nature would always prefer a non-positive drift distortion on income expected return. In other words, the robust investor is afraid of an overestimated expected income growth rate.

4.5 Model Calibration

4.5.1 Data

I use monthly US data from June 1961 to December 2013. The daily US government yield data are taken from Gürkaynak et al. (2007)\(^6\). I collect the last business day of the month to obtain monthly data. I use three yields with maturities of 1, 3 and 6 years to estimate the two-factor affine term structure model. Data on price index and stock market are as in Koiijen et al. (2010) and Sangvinatsos and Wachter (2005). Consumption price index data is obtained from the Bureau of Labor Statistics.\(^7\) The stock return data I use is from CRSP value-weighted NYSE/Ames/Nasdaq index.

---


\(^7\)See http://www.bls.gov/ for details.
The financial model introduced in Section 4.2.1 consists of four independent risk driver. I assume that the volatility matrix stacking $\sigma_X$, $\sigma_\Pi$ and $\sigma_S$ is lower triangular.

$$
\begin{pmatrix}
\sigma_X^T \\
\sigma_\Pi^T \\
\sigma_S^T
\end{pmatrix} =
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
\sigma_{\Pi_1} & \sigma_{\Pi_2} & \sigma_{\Pi_3} & 0 \\
\sigma_{S_1} & \sigma_{S_2} & \sigma_{S_3} & \sigma_{S_4}
\end{pmatrix}
$$

(4.37)

The market price of risk (4.4) can take the following form to satisfy the no arbitrage assumption

$$
\Lambda_t = \begin{pmatrix} \lambda_{01} \\ \lambda_{02} \\ 0 \\ \ast \end{pmatrix} + \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \\ 0 & 0 \end{pmatrix} X_t
$$

(4.38)

where the fourth row of $\Lambda_t$ satisfies the restriction that $\sigma_S^T \lambda_0 = \eta_S$ and $\sigma_S^T \lambda_1 = 0$. Under the assumption that price of inflation rate risk is only driven by the two state variables, $X_1$ and $X_2$, the third row of $\Lambda_0$ turns to zero.

### 4.5.3 Hybrid Estimation Method

Quasi maximum likelihood estimation (QMLE) is considered a standard method to calibrate the affine term structure models, e.g. Duffee (2002) and Sangvinatsos and Wachter (2005). However, practical experience has seen tremendous estimation difficulties. The numerical challenges come from two aspects. First, these latent factor models are characterized by a complex likelihood surface. Due to a large number of parameters to be estimated, the nonlinear optimization problem often does not converge. Second, as argued by Ang and Piazzesi (2003), it is hard to find good starting values to achieve convergence due to high degrees of freedom.

I use a hybrid estimation method that combines the method of Hamilton and Wu (2012) and the Generalized method of moments (GMM) method introduced by Hansen (1982). In particular, I use minimum chi-square estimation (MCSE) as an alternative to QMLE to estimate the affine term structure model as well as the stock and inflation diffusion processes. Both methods could be viewed as specially cases of minimum distance estimation (MDE), in which one minimizes a distance square between restricted and unrestricted statistics. Estimation details are described in Appendix 4.10.2.
Table 4.1 reports the parameter estimation results followed by their estimation errors. The results are displayed in annual terms. First, I summarize some properties of the short rate process and the inflation rate. The constant terms $\delta_0$ and $\xi_0$ are both comparable to Koijen et al. (2010) and Sangvinatsos and Wachter (2005). The short rate is increasing with the two latent factors and is more sensitive to $X_2$. One unit of increase on $X_2$ with the other factor $X_1$ fixed, results in a 188 basis points increase of the nominal interest rate. However, the two factors $X_1$ and $X_2$ have an opposite impact on the inflation rate.

Next I turn to the mean reversion parameters. The risk neutral mean-reversion matrix $\kappa^Q$ is a lower triangular matrix. The bond market price of risk excluding the unconditional part $\sigma^Q \lambda_1$ fills the difference between $\kappa^Q$ and $\kappa$. The eigenvalues of the mean-reversion matrix $\kappa$ are 0.57 and 0.10. It corresponds to a half-life innovation of 1.2 years for $X_1$ and approximately seven years for $X_2$. Therefore, $X_2$ is estimated to be more persistent than $X_1$. The half-life innovation for the nominal interest rate must be also 1.2 year, the same as innovation speed of $X_1$. 8

Last, I highlight some properties of the stock return parameters from the penultimate panel of Table 4.1. First, the value of $\eta_S$ represents the expected excess return. The standard deviation of stock return parameter is relatively higher than the volatility of other (unconditional) expected return parameters, such at $\delta_0$ and $\xi_0$. This result confirms that the expected stock return is harder to estimate than the other return parameters in the model. Second, negative values of $\sigma_{S_1}$ and $\sigma_{S_2}$ indicates that stock and bond returns must be positively correlated, since the first two rows of bond volatility vector elements $\sigma_X B(\tau)$ are also negative.

Table 4.2 provides some insights in the implied moments of the estimated model. I compare the summary statistics of the raw sample series with the estimated series from the discretized model using the estimation results from Table 4.1. Table 4.2 provides an overview of the first two moments of stock returns, inflation and the three yields on a monthly basis. The estimated latent factors $X_t$ are linear combinations of the one-year and six-year yields. Hence, by construction, the estimated sample moments of the one-year and six-year yields match the distribution of raw sample exactly. The remaining three-year yield, stock return and inflation process all fit the raw sample distribution reasonably well. The model underestimates the expected monthly stock return by approximately 10 basis points, but the high volatility of the stock return series make the difference between the two statistically insignificant. To conclude, Table 4.2 implies that the model for long-term bond yields also captures the dynamics of inflation and stock returns.

Figure 4.1 shows that the expected inflation rate implied from the model does not

---

8To calculate half-life of nominal rate under the multi factors affine model, one has to use the largest eigenvalue of mean reversion matrix $\kappa$ as a proxy of mean reversion parameter.
always manage to forecast the realized inflation rate despite the tightly fitted moments shown in Table 4.2. For instance, in 1973 during the oil crisis, the model underestimates the inflation rate. In contrast, the inflation rate is over estimated during the dot-com bubble, around the year 2001.

Next, I elaborate on some properties of the price of risk parameters. Table 4.3 reports the unconditional bond premia $B(\tau)^\top \sigma\chi\lambda_0$ as well as their volatilities $\sqrt{B(\tau)^\top \sigma\chi\sigma\chi B(\tau)}$ for various maturities with latent factors equal to their long-term means of zero. It is reported in Table 4.1 that both unconditional bond market price of risk factors $\lambda_{01}$ and $\lambda_{02}$ are negative. Since $B(\tau)$ is a negative vector, the unconditional bond premia $B(\tau)^\top \sigma\chi\lambda_0$ must be strictly positive and are increasing with bond maturity. Further, Table 4.3 also implies that both bond risk premia and volatilities are positively correlated with their maturity $\tau$, however the Sharpe ratio $^9$ is negatively correlated with the bond maturity. Therefore, the one-year nominal bond must have a higher Sharpe ratio than the ten-year nominal bond (0.36 versus 0.18).

Table 4.4 reports the correlations between the returns of different assets included in the asset menu and the inflation rate. These correlations can influence the hedging demand to

---

$^9$Sharpe ratio is a fraction of a bond risk premium and its corresponding volatility.
hedge future investment opportunities, hence are important to investors. I first summarize
the property of the correlation between inflation rate and other asset returns. The value
of $\sigma_\Pi$ shown in Table 4.1 implies that inflation rate risk is negatively correlated with
the bond return risk. Indeed, as Table 4.4 shows that correlation is negative. Further,
Table 4.4 also reports a weak negative correlation between stock return and inflation
rate, since the third elements of $\sigma_\Pi$ and $\sigma_S$ have opposite sign. Second, bond returns are
always positively correlated with the stock returns, but the correlation value is negatively
correlated with the bond maturity. This property is consistent with Koijen et al. (2010).

Figure 4.8 presents the conditional nominal and real interest rates as well as inflation
for a reasonable range of $X_1$ and $X_2$. Panel (a) fixes the latent factor $X_2$ at zero
and $X_1$ varies between -2 to 2. Panel (b) is the other way around. Figure 4.8 shows
that both nominal and real interest rates are positively correlated with the two state
variables. Inflation is negatively correlated with $X_1$, for $\xi_1$ (see Table 4.1) is negative.
The time-variation of bond premia is governed of $\lambda_1$. Figure 4.3 shows that bond premia
are increasing with $X_1$ but decreasing with $X_2$. The figure shows that the long-term bond
premium is more sensitive to the movement of both state variables than the short term
bond premium.

4.6 How Large is Ambiguity Aversion Parameter

How much should an investor adapt her behavior when facing model (parameter) ambi-
guity? In other words, how to calibrate the penalty parameter, $\theta$? One may derive an
optimal decision from the robust model, then use an econometric estimate to calibrate
the ambiguity aversion parameters. For example, Maenhout (2004) uses the gap between
the observed equity premium and the pessimistic equity premium derived from the equi-
librium of the robust asset pricing model to calibrate the ambiguity aversion parameter,
so as to claim that model uncertainty can explain the equity premium puzzle. However,
in this paper, income process is individual specific. It is difficult to obtain an economic
equilibrium from a life cycle model with heterogenous labor income.

In this section, I introduce a new approach to measure the ambiguity aversion param-

\[
X_T = \exp\left(-\kappa(T - t)\right) X_t + \int_t^T \exp\left(-\kappa(T - s)\right) \sigma_\lambda dZ_s
\]

therefore $X_t$ is a Gaussian with mean $E[X_T | F_t] = \exp\left(-\kappa(T - t)\right) X_t$ and variance $\text{Var}[X_T | F_t] = \int_t^T \exp\left(-\kappa(T - s)\sigma_\lambda^2 \sigma_\chi \exp\left(-\kappa(T - s)\right)\right) ds$. The volatilities for $X_1$ and $X_2$ over one year is 0.57 and
0.91 respectively based on the estimation results 4.1. Assume that the long-term mean of $X_1$ and $X_2$ are
close to zero, the 99% confidence interval for $X_1$ approximately between -1.5 to 1.5 and for $X_2$ is between
-2.4 to 2.4.
Figure 4.2: **State-Variable Dependent Interest Rate and Inflation** The figure presents the interest rate in both nominal and real terms and inflation rate (in percentage) as a function of the two state variables.

(a) $X_2 = 0$

(b) $X_1 = 0$

**Proposition 4.6.1.** Assume that the estimate $\hat{\psi}$ for parameter $\psi = \left( \begin{array}{c} \pi_t \\ \gamma_1 + \sigma_2^2 \end{array} \right)$ is asymptotically normal. The estimation error $\left( \hat{\psi} - \psi \right) = \Sigma \Gamma^T$ is also normally distributed with mean zero and variance $\Sigma \Sigma^T$ where $\Sigma$ is the covariance matrix of inflation and income processes (see Appendix 4.10.3.6, $\Gamma = (\gamma_1, \gamma_2)^T$ is the perturbation vector and $T$ is the number of sample observations. We also know that $T \left( \hat{\psi} - \psi \right)^T \left( \Sigma \Sigma^T \right)^{-1} \left( \hat{\psi} - \psi \right)$ is a Chi-square distribution with two degrees of freedom. Denote the critical value at $\alpha$ significance level as $CV_\alpha$, then there is a probability of $1 - \alpha$ that

$$T \left( \hat{\psi} - \psi \right)^T \left( \Sigma \Sigma^T \right)^{-1} \left( \hat{\psi} - \psi \right) \leq \kappa^2$$

where $\kappa^2 = \frac{CV_\alpha}{T}$. Therefore, the perturbation parameters satisfy the following constraint set

$$S = \{ \gamma_1, \gamma_2 \mid \gamma_1^2 + \gamma_2^2 \leq \kappa^2 \}$$

See Appendix 4.10.3.6 for the derivation of Proposition 4.6.1. The uncertainty set
4.6. HOW LARGE IS AMBIGUITY AVERSION PARAMETER

Figure 4.3: **State-Variable Dependent Bond Premia** The figure plots the 3-year and 10-year nominal bond premia as functions of state variable $X_1$ (Panel (a)) and state variable $X_2$ (Panel (b)).

(a) $X_2 = 0$

(b) $X_1 = 0$

S transfers nature’s freedom to a confidence interval of the estimation parameters. The small significance level triggers a bigger set $S$, hence will result in a larger perturbation.

- Second, I use robust control theory to obtain optimal perturbations which quantify the distance between point estimates and the true value of parameters. The optimal distortions are a function of the ambiguity aversion parameter. This step has been detailed in Propositions 4.4.2 and 4.4.5.

- Third, by simulating a large number of optimal perturbations under a series of trial preference parameters and mapping to the statistical confidence set, one can obtain the feasible value of the ambiguity aversion parameter. Simulations outside the statistical boundary are highly improbable to occur and the corresponding ambiguity levels are too high to be accepted.

Figure 4.4 plots the simulated robust optimal expected income returns as a function of the optimal perturbations $\gamma_1$ based on Proposition 4.4.5 under different trial $\theta$’s. If $\theta = 1$, all simulations drop inside the statistical boundary $S$. However, if one increases $\theta$ to 2, then none of the simulations are inside the boundary.

Figure 4.5 plots the maximum value of penalty $\theta$ under different statistical significance level $\alpha$ as a function of human capital over wealth ratio. The lower the confidence interval is, the low the required penalty term. In absence of labor income, when $H = 0$, the maximum value of $\theta$ equals 4 at the 95% confidence level. Younger cohorts, usually have a high human capital over wealth ratio and is less ambiguity aversion than the middle-age
cohort.

To conclude, the value of $\theta$ can be chosen between 0 to 4 for Model 2. For Model 4, however, a feasible region of $\theta$ depends on the $\frac{H}{W}$ ratio. For example, if an investor has a human capital over wealth ratio equal to 100, then her maximum preference for robustness even under the lowest significance level is barely above 1.

4.7 Numerical Solution

This section aims to discover the quantitative impact of parameter uncertainty on the optimal life-time asset allocation. I first determine the feasible region of the penalty parameter $\theta$ in Section 4.6. In Section 4.7 I demonstrate the optimal risk taking under various preferences for different age cohorts.

4.7.1 Robust Optimal Life-Cycle Decision without Labor Income

Table 4.5 reports the optimal strategies of Model 1 and Model 2 for different age cohorts without considering labor income. The risk aversion level is chosen at $\gamma = 5$ and the ambiguity aversion level for Model 2 is $\theta = 4$, which is a reasonable value according to Proposition 4.6.1. Investor’s age indirectly determines the remaining hedging horizon, as I assume a fixed retirement age. Therefore, the younger an investor is, the longer her hedging horizon will be.

For both models, the myopic demand contains a long position in 10-year nominal
bonds and a long position in stocks. The myopic demand depends on the current value of the state variable, but is independent of hedging horizon. Therefore, if $X$ is fixed, the myopic demand also stays the same for different age cohorts. When both state variables take values of their long-run mean of zero, the estimated risk premium for 10-year bond is 2.6% (see Table 4.3), which is higher than the stock premium. Not surprisingly, the myopic demand for long-term bonds is higher than that for stocks. The 10-year bond also enjoys a lower level of volatility and hence a higher marker price of risk than stocks.

The preference for robustness creates a small myopic demand for long-term bonds while in contrast, a reduction in the demand for stocks. What drives the risk shifting effect? On one hand, the inflation penalty term $\theta \sigma^\top \epsilon \epsilon^\top \sigma_{11}$ shown in (4.27) brings a negative shock on the expected stock return. However, it does not influence the expected bond premia. On the other hand, the additional volatility penalty term $\theta \sigma^\top \epsilon \epsilon^\top \sigma$ raises the volatility effect on the stock dimension while by construction keeps the bond volatility part unchanged. Because of the positive correlation between returns on the stock and the 10-year bond, a decrease of price of risk for the stock leads to a higher allocation on the long-term bond.

The interest rate hedging demand contains a long position in nominal bonds and a short position in stocks. There are strong horizon effects on the interest rate hedge, because functions $\Gamma_i^{Mi}(\tau)$, $\Gamma_i^{Mi}(\tau)$ with $i = 1, 2$ are increasing with the remaining hedging

Figure 4.5: A feasible value of penalty parameter as a function of human capital over financial wealth ratio under a confidence level of 90% and 95%.
horizon $\tau$. The horizon effects on long-term bond is much more dramatic. For the correlation between nominal interest rate and long-term bond return is much higher than the correlation between the nominal rate and the stock return. Ditto, the penalty term also brings an extra hedging demand for long-term bonds. However it has a very subtle effect on the allocation of stocks.

A large number of studies (e.g. Wachter (2010), Brennan and Xia (2002), Koijen et al. (2010) and Wachter (2003)) find that a time-varying short rate can generate a large amount of hedging demand for long-term bonds. The closed-form solutions Proposition 4.4.1 and 4.4.2 imply that the hedging opportunity for bond portfolio comes from two sources. One is from the time-varying risk bond premia and the other one comes from the investor’s desire for real interest rate hedge $r - \pi$ which can be found is PDE (4.57). The second source creates the well-known horizon effects for long-term bonds.

Nature’s decision influences the reference model in two aspects. On one hand, nature’s decision obviously would perturb the expected inflation rate to a different level. Figure 4.6 shows that a positive inflation shock is considered as the worst scenario, which also means that nature must be choosing a negative drift distortion term $\gamma_1 < 0$ for different age cohorts. Hence, a robust investor must worry about under estimation of the inflation level. On the other hand, a negative drift distortion would also brings a downside shock to the expected stock return. This explains why a robust investor wants to reduce her risk exposure on the stock market shown in Table 4.5. There are also horizon effects of the optimal perturbations, as the older investors are more worried about the inflation rate misspecification.

In order to understand how the robust decision varies over different state variables and different risk preferences, I compare the changes of the optimal hedge at each asset class from Model 1 to Model 2 for different scenarios of state variables as well for difference risk aversion level. Figure 4.7 demonstrates the risk shifting effects when an ambiguity neutral investor is aware of the model misspecification and becomes ambiguity aversion. First, with the preference for robustness, a robust investor reduces the hedge demand for stocks. For the inflation drift distortion would reduce the expected stock return. In other words, the realized hedging opportunity for stocks is over estimated (see also Figure 4.6). Secondly, Figure 4.7 also shows that a robust investor always required more hedging demand for long-term bond and this demand increases with $X_2$ while decreases with age. As shown in Figure 4.8 that the real interest rate is increasing with $X_2$, and a higher real interest rate triggers a larger demand for long-term bond. Not surprisingly, risk takers (e.g. $\gamma = 2$) enjoy a more aggressive risk shifting movement than the high risk aversion investors. For the risk takers are willing to put a higher weight on each investment opportunity.
4.8. ROBUST OPTIMAL LIFE-CYCLE DECISION WITH LABOR INCOME

Figure 4.6: Distortion (in bp) on inflation rate and expected stock return. The penalty parameter $\theta$ is assumed equal to 4. State variables are assumed equal to their long-run means zero.

4.8 Robust Optimal Life-Cycle Decision with Labor Income

Different from the first two models, when labor income is included and is assumed partially correlated with the financial market, an additional investment opportunity is created. Therefore, in addition to myopic demand and interest rate hedging components, a new hedging component occurs in Model 3 and Model 4 named income risk hedging component. Table 4.6 displays the hedging components for Model 3 and Model 4 for various age cohorts. When labor income is introduced, nature has freedom to control two drift distortions. As the result, the impact of parameter uncertainty becomes much more complicated than for Model 2.

The optimal portfolio choices are no longer wealth independent. However, as shown in (4.32) and (4.34) that the human capital over wealth ratio $\frac{H}{W}$ plays a crucial role in lifetime optimal asset allocation. Munk and Sørensen (2010) also emphasize the importance of $\frac{H}{W}$ ratio but not the two separately to the investment decision making. According to Proposition 4.4.3, human capital, the accumulative expected future income, is a function of the remaining working time horizon and the income growth rate. Figure 4.10.3.3 indicates that the human capital over wealth ratio $\frac{H}{W}$ is a decreasing function of age, for
human capital diminishing over time, while in the meantime financial wealth is growing over employment years.

Two important results are obtained from Table 4.6. First, the human capital parameter has a strong effect on the optimal decision for the young cohort. As demonstrated in Proposition 4.4.4 that, both myopic demand and the interest ratio hedging demand are scaled by the total wealth over fiance wealth ratio $\frac{W+H}{W}$. The scaling effect influences both hedging components but in different ways. For myopic demand, it is not surprising that the youth should enlarge their risk exposure on both long-term bonds and stocks by a noticeable amount, since for instance the mean-variance investment opportunity is enlarged for more than six times for investors younger than 30.

However, it is interesting to notice that the interest rate hedging component suggests to short the long-term bonds for young investors. What drives such a dramatic change? In the previous section the real interest rate $r - \pi$ also generates hedge demand for long-term bond and creates horizon effect in the first two models. Mathematically, when $X = 0$, the interest rate hedging component depends on $\Gamma_M^i$ for $i = 3, 4$ with other parameters given. The full expression of $\Gamma_M^i$ are shown in Appendix 4.10.3.4 and 4.10.3.5. It shows in function (4.68) that the last two terms $\frac{W}{W+H} \delta^T - \xi^T$ which represents the desire for interest rate change make the real rate hedging much smaller or even negative when $H$ is high compared to the previous two models. However, when an investor is about to retire, her human capital vanishes and her preferences converge to the cases without human capital. That also explains the similarity between the four preferences for the older cohorts for each hedging component.

Figure 4.9 shows more details about the asset allocation when human capital is involved. The allocation to each asset class depends on both investors’ risk aversion level and the state variables. First, the demand for long-term bond is decreasing with $X_2$, because of the negative relationship between bond premia and $X_2$, shown in Figure 4.3. Therefore, a lower $X_2$ with $X_1$ given implies a higher bond premia hence leads to a bigger risk exposure. Second, the optimal long-term bond allocation is very sensitive to the risk aversion level. An extreme long-position on long-term bond is suggested for a risk takers. However, a risk averse investor requires a short position. What makes the variation so dramatic? It is notable from Table 4.6 that both myopic demand and interest rate hedging component dominate the demand for long-term bond and have two opposite signs for the young cohorts. A high risk aversion level, say $\gamma = 10$ brings a bigger negative effect on function $\frac{\delta M^4}{\rho r}$ and also reduce the myopic demand although the myopic demand always stays positive. Therefore, the joint effect leads to a short position on the long-term bond.

The second message from Table 4.6 is that the perturbations bring a much bigger risk shifting effect especially for the young cohorts compared with Model 2. The investor
worries about ambiguity exposure of total wealth to the drift distortions, and not only to financial wealth. Figure 4.10 provides more details about the risk shifting effects over as functions of age. It is important to notice that the risk shifting unit used in Figure 4.10 is percentage, while in Figure 4.7, the risk shifting amount is in unit of basis point, because the risk shifting effect is very small without labor income uncertainty. I only focus on the dynamics for the young cohorts, because the human capital parameter can only influence the young cohorts’ investment behavior.

The asset allocation shifting is caused by the uncertainty of the expected returns. Nature’s decision represents a robust investor’s fear for the worst case scenario. Therefore, to sum up, as shown in Figure 4.11 that the perturbation will reduce expected equity return as well as the expected income growth rate, while in contract will increase the inflation rate, since all those three scenarios can make the value of investor’s utility lower than expected. The fear for the an overestimated equity return leads to a more conservative decision on the stock market, therefore in Figure 4.10, the shifting of stock allocation is always negative.

What makes the robust demand for long-term bond so sensitive to the state variables as well as the risk aversion level? The extra demand for long-term bond comes from the three hedging component. The myopic demand and income risk hedging demand both requires a longer position on the long-term bond when an investor is ambiguity averse. Therefore, it is the interest rate hedging component that controls the sign of the the demand change.

4.9 Conclusion

In this chapter, I present a robust optimal life cycle consumption and investment strategy for an uncertainty averse investor who cannot fully diversify her inflation and labor income risk. A new approach is developed to calibrate the ambiguity aversion parameter which combines insights from robust asset allocation and financial econometrics.

The impact of parameter uncertainty on the optimal asset allocation is heterogeneous, depending on an investor’s age, risk aversion level, income stability as well as the instantaneous state variables. Young investors are particularly sensitive to parameter uncertainty due to the longer planing horizon. The robust policy is particular in favor of long-term bond market especially when expected inflation is highly uncertainty.

There are a few limitations of my model. First, throughout the paper, I do not set any constraint on portfolio weights, hence borrowing from the future income and short selling are allowed. These assumptions are necessary in order to obtain closed-form solution but they can be unrealistic in reality. Therefore, a numerical solution with borrowing
constraint bearing in mind would be useful extension. Second, in order to provide a closed form solution for the human capital function $H$, I assume away idiosyncratic risks and neutralize the uncertainty effect. Again, numerical methods can help to relax these restrictions.

This chapter provides a robust life-cycle decision theory for a single agent. However, as an economist, it is also important to think about the multi-agent decision theory. In other words, how does income and inflation model uncertainty affect the equilibrium concept? The robust investment policy suggests that embedding this model into general equilibrium might help explain various life cycle puzzles. For instance, the risk shifting effect (see Figure 4.10) can to some extend support the fact that the share of risky assets increases with age and can also explains the low share of risky assets (equity premium puzzle).
4.10 Chapter 4 Appendix

4.10.1 Nominal bond pricing

In line with Cox et al. (1985), Sangvinatsos and Wachter (2005) and many other term structure studies, I assume that bond prices are continuous functions of state variables $X$ and of time $t$. I omit time subscription for notation convenience. No arbitrage condition implies that $P(t,T)$ satisfies

$$-\kappa XP_X + \frac{1}{2} \text{tr} \left( P_{XX} \sigma_X^T \sigma_X \right) + P_t - rP = P_X \sigma_X^T \Lambda$$

(4.40)

Assume that bond price is exponential affine in the $X$, (see (4.5)) and substitute this conjecture bond price function to (4.40),

$$-PB^T(\tau) \kappa X + \frac{1}{2} PB^T(\tau) \sigma_X^T \sigma_X B(\tau) + P (-B'(\tau)X - A') - rP = PB^T(\tau) \sigma_X^T \Lambda$$

and matching coefficients on $X$ and the constants produce the following system of ordinary differential equation for row vector $B(\tau)$, $1 \times m$ and the scalar $A(\tau)$

$$B'(\tau) = - \left( \kappa^T + \lambda_1^T \sigma_X \right) B(\tau) - \delta \quad (4.41)$$

$$A'(\tau) = \frac{1}{2} B(\tau)^T \sigma_X^T \sigma_X B(\tau) - \delta_0 - B(\tau)^T \sigma_X^T \lambda_0 \quad (4.42)$$

The boundary condition is that $B(0) = 0$ and $A(0) = 0$. The ODE can be solved analytically,

$$B(\tau) = (\kappa^T + \lambda_1^T \sigma_X)^{-1} \left[ \exp \left( - \left( \kappa^T + \lambda_1^T \sigma_X \right) \tau \right) - I_{2 \times 2} \right] \delta_1 \quad (4.43)$$

$$A(\tau) = \int_0^\tau A'(s) ds \quad (4.44)$$

I define $\kappa^Q = \kappa + \sigma_X^T \lambda_1$, where $\kappa^Q$ is a lower triangular matrix.

4.10.2 Estimation Procedure

The estimation procedure is as follows. First, I estimation the term structure model using MCSE method. Second, I use GMM method to estimation inflation and stock return processes.

The term structure model estimation method I use combines in essence the method of Hamilton and Wu (2012). I use three different yields to estimate the term structure, namely $\tau_1 = 1$ year, $\tau_2 = 6$ year and $\tau_3 = 3$ year maturities. In line with Duffee (2002), I assume that 1 year and 10 year yield bonds are measured without error, and the other
CHAPTER 4. HOW MUCH SHOULD INVESTORS ADAPT?

3-year to maturity bond is assumed to be measured with serially uncorrelated, mean-zero measurement error. Yield function with maturity \( \tau_i \), with \( i = 1, 2, 3 \) is given by

\[
Y_{t}^{\tau_i} = a(\tau_i) + b(\tau_i)^\top X_t
\]

Let \( Y_{1,t} \) denote the vector of perfectly observed yield at time \( t \) and \( Y_{2,t} \) denotes the yield which is observed imperfectly. More specifically

\[
\begin{pmatrix}
Y_{1,t}^{(2\times1)} \\
Y_{2,t}^{(1\times1)}
\end{pmatrix}
= \begin{pmatrix}
A_1^{(2\times1)} \\
A_2^{(1\times1)}
\end{pmatrix}
+ \begin{pmatrix}
B_1^{(2\times2)} \\
B_2^{(1\times1)}
\end{pmatrix} X_t
+ \begin{pmatrix}
0^{(2\times1)} \\
\Sigma_e^{(1\times1)}
\end{pmatrix} u_t^{(1\times1)}
\]

where \( \Sigma_e \) represents the variance of estimation error with \( u_t \sim N(0, \Delta t) \). There are 13 unknown parameters involved in the model, namely 1 in \( \delta_0 \), 2 in \( \delta \), 4 in \( \kappa \), 3 in \( \kappa^0 \), 2 in \( \lambda_0 \) and 1 in \( \Sigma_e \).

MCSE algorithm consists of three steps. First, I reconstruct the affine term structure model to a vector autoregression (VAR) model of \( Y_{t}^{\tau} \), then estimate VAR model using OLS. The coefficients obtained from VAR is called reduced-form parameters. Second step is to map between structural parameters and reduced-form paymasters. Last step is solve minimum chi-square criterion to obtain the structural parameters.

**Step 1** I estimate parameters of continuous time model using a discrete time econometrics specification. The implied state variable can be derived from (4.6)

\[
X_t = B_1^{-1}(Y_{1,t} - A_1)
\]  

(4.45)

A discretization of (4.2) yields the discrete time model of state variables

\[
X_{t+\Delta t} = \mu_X X_t + \sqrt{\Delta t} z_{t+\Delta t} \quad \text{with} \quad \mu_X = \exp(-\kappa \Delta t)
\]

where \( \Delta t = \frac{1}{12} \) and \( z_{t+\Delta t} \) represents bivariate standard normal distribution. Both sides of equation above multiply by \( B_1 \) and plus \( A_1 \), I get

\[
A_1 + B_1 X_{t+\Delta t} = A_1 + B_1 \mu_X B_1^{-1} B_1 X_t + B_1 \sqrt{\Delta t} z_{t+\Delta t}
\]

which can be rewritten as a VAR model of \( Y_{1,t} \)

\[
Y_{1,t+\Delta t} = A_1 + B_1 \mu_X B_1^{-1} A_1 + B_1 \mu_X B_1^{-1} Y_{1,t} + B_1 \sqrt{\Delta t} z_{t+\Delta t}
\]

(4.46)

\[
= A_1^* + \phi_{11}^* Y_{1,t} + u_{1,t+\Delta t}^*
\]
where \( A_1^* = A_1 - B_1 \mu X B_1^{-1} A_1 \), \( \phi_{11}^* = B_1 \mu X B_1^{-1} \) and \( \Omega_1^* = B_1 B_1^\top \Delta t = u_1^{\top} u_1 \Delta t \Delta t \).

Similarly

\[
Y_{2,t} = A_2^* + \phi_{21}^* + u_{2,t}^*
\]  

with \( A_2^* = A_2 - B_2 B_1^{-1} A_1 \), \( \phi_{21}^* = B_2 B_1^{-1} \) and \( \Omega_2^* = \Sigma_e \Sigma_e \Delta t = u_{2,t}^* u_{2,t}^\top \). Hence the reduced-form parameter vector is

\[
\pi = \{ \text{vec} \left( (A_1^*, \phi_{11}^*)' \right), \text{vech} (\Omega_1^*), \text{vec} \left( (A_2^*, \phi_{21}^*)' \right), \Omega_2^* \}
\]

where vec(\( X \)) collects the columns of the matrix \( X \) into a vector. If \( X \) is symmetric matrix, vech(\( X \)) does the same but only use the elements below and including diagonal.

The reduced-form parameters can easily be obtained via OLS estimation with \( \Omega_1^* \) and \( \Omega_2^* \) products of those OLS residuals:

\[
\Omega_1^* = N^{-1} \sum_{t=1}^{N \Delta t} (Y_{1,t+\Delta t} - A_1^* - \phi_{11}^* Y_{1,t}) (Y_{1,t+\Delta t} - A_1^* - \phi_{11}^* Y_{1,t})^\top
\]

\[
\Omega_2^* = N^{-1} \sum_{t=1}^{N \Delta t} (Y_{2,t} - A_2^* - \phi_{21}^* Y_{1,t})^2
\]

where \( T \) represents the number of observations from the historical sample data. Next I will map the structural parameters to the

**Step 2**  Mapping between structural and reduced-form parameters can be done in four steps:

1. Estimation of \( \Sigma_e \) is obtained analytically, \( \Omega_2^* = \Sigma_e^2 \Delta t \).

2. Estimates of the 5 unknowns in \( \kappa^Q \) and \( \delta \), are found by numerically solving 5 equations from following relation

\[
B_1 B_1^\top \Delta = \Omega_1^* \\
B_2 B_1^\top \Delta = \phi_{21}^* \Omega_1^*
\]

3. Estimate of 4 unknowns in \( \kappa \) can be obtained analytically via

\[
\mu_X = B_1^{-1} \phi_{11}^* B_1, \quad \text{with} \quad \mu_X = \exp(-\kappa \Delta t) = N \exp(-D \Delta t) N^{-1}
\]

under the assumption that \( \kappa \) can be diagonalized and \( \kappa = N D N^{-1} \).
CHAPTER 4. HOW MUCH SHOULD INVESTORS ADAPT?

4. Numerically solve the remaining 3 unknowns $\delta_0, \lambda_0$ from equation $A_1^*, A_2^*$:

$$A_1 - B_1 \mu X B_1^{-1} A_1 = A_1^*$$
$$A_2 - B_2 B_1^{-1} A_1 = A_2^*$$

**Step 3** Let $\theta_B$ stack all 13 structure parameters of the term structure model into a vector. Next one can test the hypothesis that $\pi = g(\theta_B)$ using Wald statistics $T[\hat{\pi} - g(\theta_B)]^T \hat{R}[\hat{\pi} - g(\theta_B)]$ which would have an asymptotic $\chi^2(13)$ distribution, in which $\hat{R}$ is a consistent estimate of the information matrix $R = -\frac{1}{T} E \left[ \frac{\partial^2 L(\pi; X)}{\partial \pi \partial \pi'} \right]$ with $L(\pi; X)$ the log likelihood for the entire sample.

The reduced-form equation (4.46) and (4.47) form 2 independent blocks. For $i = 1, 2$

$$Y_{i,t} = \Pi_i x_{it} + u_{i,t}^*$$

with $u_{i,t}^* \sim N(0, \Omega_i^*)$. The information matrix $R$ for the full system of reduced-from parameter is given by

$$\hat{R} = \begin{pmatrix} \hat{R}_1 & 0 \\ 0 & \hat{R}_2 \end{pmatrix}$$

where as shown in Neudecker and Magnus (1988),

$$\hat{R}_i = \left( (\Omega_i^*)^{-1} \otimes \sum_{a} x_{it} x_{at}' \right) \frac{1}{2} D_q((\Omega_i^*)^{-1} \otimes (\Omega_i^*)^{-1}) D_q$$

for $D_N$ the $N^2 \times N(N + 1)/2$ duplication matrix satisfying $D_N \text{vech} = \text{vec} \Omega$, and $q$ represents the dimension of $\Omega$ matrix.

Following Rothenberg (1973), the minimum-chi-square estimator $\hat{\theta}_B$ is the solution of

$$\min_{\theta_B} T [\hat{\pi} - g(\theta_B)]^T R [\hat{\pi} - g(\theta_B)]$$

(4.51)

According to Hamilton and Wu (2012), the asymptotic distribution of $\theta_B$ is

$$\sqrt{T} \left( \hat{\theta}_B - \theta_B \right) \rightarrow N \left( 0, \Gamma^T R \Gamma \right)^{-1}$$

(4.52)

where $\Gamma = \frac{\partial g(\theta_B)}{\partial \theta_B}$ is the Jacobian matrix of $g(\theta_B)$ as a function of $\theta_B$.

Next, I estimation inflation and stock return process using GMM method. The discrete
time inflation return process is as follows

\[ \Delta \pi_{t+\Delta t} = (\xi_0 + \xi^T X_t) \Delta t + \sqrt{\Delta t} \sigma_\pi \epsilon_{t+\Delta t} \]

where \( \Delta \pi_{t+\Delta t} = \frac{\Delta P_{t+\Delta t}}{P_t} \). Hence, there are 6 unknown parameters, 1 in \( \xi_0 \), 2 in \( \xi \), 3 in \( \sigma_\pi \). I can generate six moment conditions based on inflation dynamics and state variable diffusion process, namely

\[
E [\Delta \pi_{t+\Delta t} - (\xi_0 + \xi^T X_t) \Delta t] = 0
\]

\[
E [X_t (\Delta \pi_{t+\Delta t} - (\xi_0 + \xi^T X_t) \Delta t)] = 0
\]

\[
E [(\Delta \pi_{t+\Delta t} - (\xi_0 + \xi^T X_t) \Delta t)^2 - (\sigma_{\pi}^T \sigma_{\pi}) \Delta t] = 0
\]

\[
E [(\Delta X_{t+\Delta t} - (\mu_X - 1)X_t) (\Delta \pi_{t+\Delta t} - (\xi_0 + \xi^T X_t) \Delta t) - \sigma_X^T \sigma_{\pi} \Delta t] = 0
\]

Similarly, one generate five moment conditions based on the discrete time stock return model

\[ r_{s,t+\Delta t} = (\delta_0 + \delta^T X_t + \eta_s) \Delta t + \sqrt{\Delta t} \sigma_S \epsilon_{s,t+\Delta t} \]

Let \( r_{s,t+\Delta t} = \frac{\Delta S_{s,t+\Delta t}}{S_t} \). Five moment conditions are

\[
E [r_{s,t+\Delta t} - (\delta_0 + \delta^T X_t + \eta_s) \Delta t] = 0
\]

\[
E [(r_{s,t+\Delta t} - (\delta_0 + \delta^T X_t + \eta_s) \Delta t) (\Delta X_{t+\Delta t} - (\mu_X - 1)X_t) - \Delta t \sigma_X^T \sigma_S] = 0
\]

\[
E [(r_{s,t+\Delta t} - (\delta_0 + \delta^T X_t + \eta_s) \Delta t)^2 - \Delta t \sigma_S^T \sigma_S] = 0
\]

\[
E [(r_{s,t+\Delta t} - (\delta_0 + \delta^T X_t + \eta_s) \Delta t) (\Delta \pi_{t+\Delta t} - (\xi_0 + \xi^T X_t) \Delta t) - \Delta t \sigma_{\pi}^T \sigma_{\pi}] = 0
\]

Denote \( \theta_\pi \) a vector of inflation process parameters and \( \theta_S \) a vector of stock process parameters. The GMM estimates \( \hat{\theta}_{GMM} = (\hat{\theta}_\pi, \hat{\theta}_S)^T \) aim to minimize a quadratic form of sample mean of the moment conditions \( g(\theta_{GMM}) = \frac{1}{T} \sum f_t(\theta_{GMM}) \), where \( f_t(\theta_{GMM}) \) denotes a vector of errors of moment conditions at time \( t \). The long run covariance matrix is defined by

\[ S = \sum_{j=-\infty}^{\infty} E[f_t(\theta_{GMM}) f_{t-j}'(\theta_{GMM})] \] (4.53)

I use Newey West method to estimate \( S \) with \( j = 12 \). Hence under the efficient GMM, the estimate \( \theta \) is asymptotically normal

\[ \sqrt{T}(\hat{\theta}_{GMM} - \theta_{GMM}) \sim N(0, (d'S^{-1}d)^{-1}) \] (4.54)
where \( d \) is the Jacobian matric of the population moments vector estimated by \( \frac{\partial(g_{GMM})}{\partial(\theta_{GMM})} \).

4.10.3 Proofs

4.10.3.1 Proof of Proposition 4.4.1

**Model 1** is a standard life-cycle problem with consumption. Wachter (2010) provides a systematic overview of modern techniques on solving asset allocation problems. In this paper, I use standard dynamic programming method by means of solving a HJB equation to obtain the optimal decision. Define value function of **Model 1** at time \( t \) as \( J^{M1}(W, \Pi, X, t) \). The HJB equation of **Model 1** reads as

\[
\beta J^{M1} = \max_{C_t, \sigma_t} \left\{ \frac{1}{1 - \gamma} \left( \frac{C}{\Pi} \right)^{1-\gamma} + J^{M1}_t + J^{M1}_W \left[ (\mu - r) W - C \right] + \frac{1}{2} J^{M1}_W W \sigma^\top \sigma \right\} - \frac{1}{2} \text{tr} \left( J^{M1} X \sigma^\top \sigma X \right) + \frac{1}{2} J^{M1} \Pi \Pi \sigma^\top \sigma \sigma^\top \Pi - \frac{1}{2} \text{tr} \left( \gamma J^{M1} X \sigma X \right) + \frac{1}{2} \text{tr} \left( \gamma J^{M1} \Pi \sigma X \right) \]

(4.55)

The first order condition for consumption equals to

\[
C^{M1} = \left( J^{M1}_W \Pi \right)^{\frac{1}{1-\gamma}} \]

(4.56)

Then, substitute the conjecture function (4.24) as well the the optimal consumption (4.56) back to the equation. Function \( b^{M1} \) satisfies the following partial differential equation (PDE)

\[
0 = \left( -\frac{\beta}{1 - \gamma} + x^\top (\mu - r) + (\rho - \pi) - \frac{1}{2} \gamma x^\top \sigma^\top \sigma x - \frac{\gamma - 2}{2} \sigma_x \right) + \frac{1}{2} \left( \frac{\partial b^{M1}}{\partial \Pi} \right) \left( \gamma - 1 \right) x^\top \sigma^\top \sigma X + \frac{1}{2} \text{tr} \left( \frac{\partial b^{M1}}{\partial \Pi} \right) \sigma^\top \sigma \sigma^\top \Pi \]

(4.57)

The PDE above is comparable to equation (36) of Sangvinatsos and Wachter (2005) expect the additional term \( \frac{b^{M1}}{\partial \Pi (1 - \gamma)} \) which is generated from the optimal consumption decision \( C^{M1} \). Their work shows that as long as the optimal portfolio \( x^{M1} \) is affine in state variable \( X \), the PDE above can be solve explicitly in the absence of consumption, and the trial function is exponentially quadratic in the state variable.

Conjecture the form of \( b^{M1} \) as (4.25). Differ from Sangvinatsos and Wachter (2005) case, however, when consumption is considered, the PDE above is no longer homothetic.
hence cannot be solved explicitly by coefficients matching. Inspired by Wachter (2010) equation (49), the inverse of trial function is approximately affine in \( \log(b^{M1}) \)

\[
\frac{1}{b^{M1}} = b_0^{M1} - b_1^{M1} \log(b^{M1}) \tag{4.58}
\]

where \( b_1^{M1} = \exp\left(\mathbb{E}[-\log(b^{M1})]\right) \) and \( b_0^{M1} = b_1^{M1} (1 - \log b_1^{M1}) \). This approximation technique retrieves the PDE to a homothetic environment. The optimal portfolio obtained from the first order condition on HJB equation can be expressed as

\[
x^{M1} = a_0^{M1}(\tau) + a_1^{M1}(\tau)X \tag{4.59}
\]

where

\[
a_0^{M1}(\tau) = \frac{1}{\gamma} \left(\sigma^\top \sigma\right)^{-1} \sigma^\top \lambda_0 + \frac{\gamma - 1}{\gamma} \left(\sigma^\top \sigma\right)^{-1} \sigma^\top \sigma \Pi + \frac{1}{\gamma} \left(\sigma^\top \sigma\right)^{-1} \sigma^\top \sigma \chi \Gamma_2^{M1}(\tau)^\top
\]

\[
a_1^{M1}(\tau) = \frac{1}{\gamma} \left(\sigma^\top \sigma\right)^{-1} \sigma^\top \lambda_1 + \frac{1}{2\gamma} \left(\sigma^\top \sigma\right)^{-1} \sigma^\top \sigma \chi \left(\Gamma_1^{M1}(\tau) + \Gamma_2^{M1}(\tau)^\top\right)
\]

Substituting the trial solution (4.25) to the PDE (4.57) and match the coefficients on \( X^\top(\cdot)X, X \) and the constant term, the parameters \( \Gamma_1^{M1} \) and \( \Gamma_2^{M1} \) satisfy the following system of ODE

\[
\Gamma_1^{M1}(\tau)' = \left(\Gamma_1^{M1}(\tau) + \Gamma_2^{M1}(\tau)^\top\right) \left(-\kappa + (1 - \gamma) \sigma^\top \sigma \chi a_1^{M1}\right) + \frac{\Gamma_1^{M1}(\tau) + \Gamma_2^{M1}(\tau)^\top}{\sigma^\top \sigma \chi} \left(\Gamma_1^{M1}(\tau) + \Gamma_2^{M1}(\tau)^\top\right) + 2(1 - \gamma) a_1^{M1} \sigma^\top \sigma \lambda_1
\]

\[
\Gamma_2^{M1}(\tau)' = \Gamma_2^{M1}(\tau) \left[-b_1^{M1} - \kappa + (1 - \gamma) \sigma^\top \sigma \chi a_1^{M1} + \frac{\Gamma_1^{M1}(\tau) + \Gamma_2^{M1}(\tau)^\top}{\sigma^\top \sigma \chi}\right]
\]

\[
+ \left[(1 - \gamma) a_0^{M1} - (1 - \gamma) a_0^{M1} \sigma^\top \chi - (1 - \gamma) a_0^{M1} \sigma^\top \chi\right] \Gamma_1^{M1}(\tau) + \frac{\Gamma_1^{M1}(\tau) + \Gamma_2^{M1}(\tau)^\top}{\sigma^\top \sigma \chi}
\]

\[
- \gamma (1 - \gamma) a_0^{M1} \sigma^\top \chi a_1^{M1} + (1 - \gamma) \left[a_0^{M1} \sigma^\top \chi \lambda_1 + \lambda_0^{M1} \sigma^\top a_1^{M1} + \delta^\top - \xi^\top\right]
\]

with the boundary condition that \( \Gamma_1^{M1}(0) = 0, \Gamma_2^{M1}(0) = 0 \) and \( \Gamma_3^{M1}(0) = 0 \). The expression for \( \Gamma_3^{M1} \) is not required for the optimal portfolio and hence I do not provide.

4.10.3.2 Proof of Proposition 4.4.2

Denote \( J^{M2}(W, \Pi, X, t) \) an indirect utility function of Model 2. The robust HJB equation with only one source of parameter uncertainty is a special case of (4.22) stated in Section 4.3.2, with \( Y = 0 \) and \( \gamma_2 = 0 \). The solution technique is very similar to Model 1,
CHAPTER 4. HOW MUCH SHOULD INVESTORS ADAPT?

expect that robust HJB equation has an additional penalty function. Thanks to Maenhout (2004)'s trick (4.23), the robust HJB equation still stays homothetic hence is friendly to the closed-form solution.

The robust solution can be obtained in three steps.

- Step 1. Replace the optimal consumption, the optimal affine portfolio choice and the optimal distortion, obtained from the first order conditions respectively, back to the robust HJB equation. The optimal portfolio takes the form

\[ x^{M_2} = a_0^{M_2} + a_1^{M_2}X, \quad a_0^{M_2} \in \mathbb{R}^{2 \times 1}, \quad a_1^{M_2} \in \mathbb{R}^{2 \times 2} \]

- Step 2. Substitute the conjecture for \( J^{M_2} \) back to the PDE derived from Step 1. The trail function of \( J^{M_2} \) takes the same form as (4.24)

\[ J^{M_2} = \frac{1}{\gamma} \left( W \Pi \right)^{1-\gamma} \exp \left\{ \frac{1}{\gamma} \left( \frac{1}{2} \chi^{M_2}(\tau) X + \Gamma_1^{M_2}(\tau) X + \Gamma_2^{M_2}(\tau) X + \Gamma_3^{M_2}(\tau) \right) \right\} \]

(4.62)

- Step 3. Eventually, the simplified PDE is an affine function of \( X^\top (\cdot) X \) and \( X \). The last step is to match the coefficients of each polynomial order of the state variable.

\[
\Gamma_1^{M_2}(\tau) = \Gamma_1^{M_1}(\tau) \left( a_0^{M_2}, a_1^{M_2} \right) - (1 - \gamma) \theta a_1^{M_2} \sigma^\top \sigma a_1^{M_2}
\]

\[
\Gamma_2^{M_2}(\tau) = \Gamma_2^{M_1}(\tau) \left( a_0^{M_2}, a_1^{M_2} \right) + (1 - \gamma) \theta \left[ \sigma_1^\top \sigma a_1^{M_2} - a_0^{M_2} \sigma^\top \sigma a_1^{M_2} \right]
\]

The first part of ODE system repeats the corresponding ODE of Model 1 while replacing the deterministic functions function \( a_0^{M_1} \) and \( a_1^{M_1} \) to \( a_0^{M_2} \) and \( a_1^{M_2} \) respectively.

Then substituting the explicit solution of \( J^{M_2} \) back to the first-order conditions, one can arrive at Proposition 4.4.2. The expression of \( a_0^{M_2} \) and \( a_1^{M_2} \) can easily be obtained from (4.27) via coefficients matching.

4.10.3.3 Proof of Proposition 4.4.3

The state variable under \( Q \) measure follows the diffusion process

\[ dX_t = \left( -\kappa^Q X_t - \sigma^\top X_t \lambda_0 \right) dt + \sigma^\top dZ^Q_t \]

(4.63)

where \( \kappa^Q = \kappa + \sigma^\top \lambda_1 \). Assume that \( \kappa^Q \) is diagonalizable, then there exists a diagonal matrix \( D \) and a matrix \( N \) such that

\[ \kappa = NDN^{-1} \]
Hence the explicit solution of $X_s$, $s \geq t$ is given by

$$X_s = N \exp \left( -D (s - t) \right) N^{-1} X_t - \int_t^s N \exp \left( -D (s - v) \right) N^{-1} \sigma_N^T \lambda_0 dv$$

$$+ \int_t^s N \exp \left( -D (s - v) \right) N^{-1} \sigma_N^T dZ^Q_v$$

hence

$$\int_t^u X_s ds = \int_t^u N \exp \left( -D (s - t) \right) N^{-1} X_t ds - \int_t^u \int_t^s N \exp \left( -D (s - v) \right) N^{-1} \sigma_N^T \lambda_0 dv ds$$

$$+ \int_t^u \int_t^s N \exp \left( -D (s - v) \right) N^{-1} \sigma_N^T dZ^Q_v ds$$

The integration above contains three parts. The first two parts can be solved analytically.

One can reduce one level of integration of the third part by applying the Fubini rule.

Then I get

$$\int_t^u X_s ds = \int_t^u N \exp \left( -D (s - t) \right) N^{-1} X_t ds - \int_t^u \int_t^s N \exp \left( -D (s - v) \right) N^{-1} \sigma_N^T \lambda_0 dv ds$$

$$+ \int_t^u \int_t^s N \exp \left( -D (s - v) \right) N^{-1} \sigma_N^T dZ^Q_v ds$$

(4.64)

where $d_i$ represents the diagonal element of matrix $D$ with $\{i = 1, 2\}$, and

$$f (d_i, \tau) = -\frac{1}{d_i} \left[ \exp \left( -d_i (u - t) \right) - 1 \right].$$

The nominal income process under $Q$ measure is given by

$$dY_t = \left( \frac{\sigma_y^2}{2} - \sigma_y \rho \lambda_0 - \sigma_y \rho \lambda_1 X_t \right) Y_t dt + Y_t \sigma_y dZ^Q_t \tag{4.65}$$

Hence the explicit solution of nominal income stream at time $u$ is given by

$$Y_u = Y_t \exp \left[ \int_t^u \left( \frac{\sigma_y^2}{2} - \sigma_y \rho \lambda_0 - \frac{1}{2} \sigma_y^2 \rho \lambda_0 \right) ds \right.$$

$$- \int_t^u \sigma_y \rho \lambda_1 X_s ds + \int_t^u \sigma_y dZ^Q_s \left. \right]$$
CHAPTER 4. HOW MUCH SHOULD INVESTORS ADAPT?

The explicit expression of $Y_u \exp \left( - \int_t^u r_s ds \right)$ is log normally distributed

$$Y_u \exp \left( - \int_t^u r_s ds \right) = Y_t \exp \left\{ \int_t^u \left( g + \frac{\sigma_u^2}{2} - \sigma_u \rho_y \lambda_0 - \frac{1}{2} \sigma_u^2 \rho_y \rho_y - \delta_0 \right) ds - (\sigma_u \rho_y \lambda_1 + \delta^\top) \left[ N f (d_i, u - t)_i N^{-1} X_i - N \left( \frac{1}{d_i} (u - t) - \frac{1}{d_i} f (d_i, u - t)_i \right) N^{-1} \rho_y \lambda_0 \right. \\
\left. + \int_t^u f (d_i, u - s)_i \rho_y \sigma_y d\bar{Z} \right\} (4.66)$$

Taking the expectation of exponential of normal random variable

$$E^Q_t \left[ Y_u \exp \left( - \int_t^u r_s ds \right) \right] = Y_t \exp \left\{ \mu_{H_1} (u - t) + \mu_{H_2} N \left( \frac{1}{d_i} (u - t) - \frac{1}{d_i} f (d_i, u - t)_i \right) N^{-1} \rho_y \lambda_0 + \frac{1}{2} \int_t^u \Sigma_H(s) \Sigma_H(s)^\top ds + \mu_{H_2} N f (d_i, u - t)_i N^{-1} X_i \right\}$$

$$= Y_t \exp (M_1 (u - t) + M_2 (u - t) X_i)$$

where

$$\mu_{H_1} = g + \frac{\sigma_u^2}{2} - \sigma_u \rho_y \lambda_0 - \frac{1}{2} \sigma_u^2 \rho_y \rho_y - \delta_0$$

$$\mu_{H_2} = \sigma_u \rho_y \lambda_1 + \delta^\top$$

$$\Sigma_H(s) = f (d_i, u - s)_i \rho_y \sigma_y + \sigma_y \rho_y \sigma_y$$

The exponential term consists of four parts and are separated into two groups. The first three parts together named $M_1$ is a scalar. The remaining part is state variable dependent.

Integrating the expectation above over $u$ arrives Proposition 4.4.3.
4.10.3.4 Proof of Proposition 4.4.4

The complete version of the Hamilton-Jacobi-Bellman (HJB) equation (4.17) of the dynamic optimization problem (4.13) is given by

$$\beta J = \max_{C_t, x_t} \left\{ \frac{1}{1 - \gamma} \left( \frac{C_t}{\Pi} \right)^{1-\gamma} + J_t + J_W \left[ (x^\top (\mu - r) + r) W - C + Y \right] + \frac{1}{2} J_{WW} x^\top \sigma^\top \sigma x + J_{YY} \left( g + \sigma^2_y \right) + \frac{1}{2} J_{YX} X_{\sigma} \right\} \tag{4.67}$$

Hence the right side of (4.67) except the first term represents the term $D(C_t, x_t) J(W, Y, \Pi, X, t)$ of (4.17). The HJB equation of Model 1 is a special case of (4.67), under the condition that $Y = 0$. Solution technique is the same as previous two models, expected the form of indirect utility function. Under the assumption of Model 3, the indirect utility function denoting $J^{M3}$ must be a function of labor income. $J(W, Y, \Pi, X, t) = J^{M3}(W + H(Y, X, t), \Pi, X, t)$

where $H$ is the value of human capital for an investor at age $t$ (See Proposition 4.4.3 for details) under the condition of $X = \bar{X} = 0$. Then I borrow the idea of Bodie et al. (1992) by assuming a trial function of (4.31). The rest of the work repeats the solution procedure of Model 1.

The coefficients of an exponential affine function $b^{M3}$ (see (4.31)) has to satisfy follow-

\footnote{The assumption that $\frac{\partial H}{\partial X} = 0$ is crucial for the closed form solution. It is stated in Munk and Sørensen (2010) that, human capital is relatively insensitive to the latent factors, hence it is valid to assume that $H(Y, X, t) = H(Y, X, t)$. Without this assumption, the state variable $X$ appears in both function $H$ and $b^{M3}$. If $b^{M3}$ is still assumed exponential, the HJB equation of Model 3 will not be homothetic hence cannot be solved analytically.}
CHAPTER 4. HOW MUCH SHOULD INVESTORS ADAPT?

ing ODE system

\[ 
\Gamma_1^{M3}(\tau)' = \left( \Gamma_1^{M3}(\tau) + \Gamma_1^{M3}(\tau)\right) \left( -\kappa + (1 - \gamma) \frac{W}{W + H} \sigma_0^\top \sigma_0 a_1^{M3} \right) \\
+ \frac{\Gamma_1^{M3}(\tau) + \Gamma_1^{M3}(\tau)\sigma_0^\top \sigma_0 M3}{2} \left( \Gamma_1^{M3}(\tau) + \Gamma_1^{M3}(\tau)\right) + 2 (1 - \gamma) \frac{W}{W + H} a_1^{M3} \sigma_0^\top \lambda_1 \\
- \gamma (1 - \gamma) \frac{W^2}{(W + H)^2} \sigma_0^{M3} \sigma_0^\top \sigma_0 a_1^{M3} - b_1^{M3} (\tau) \right) \\
\left[ -b_1^{M3} - \kappa + (1 - \gamma) \frac{W}{W + H} \sigma_0^\top \sigma_0 M1 + \frac{\Gamma_1^{M3}(\tau) + \Gamma_1^{M3}(\tau)\sigma_0^\top \sigma_0 M1}{2} \right] \\
+ \left[ (1 - \gamma) \frac{W}{W + H} \sigma_0^{M3} \sigma_0^\top \sigma_0 M1 + (1 - \gamma) \frac{H}{W + H} \sigma_0^\top \sigma_0 M1 \right] \\
- (1 - \gamma) \sigma_0^\top \sigma_0 \left[ \Gamma_1^{M3}(\tau) + \Gamma_1^{M3}(\tau)\sigma_0^\top \sigma_0 M3 \right] - (1 - \gamma)^2 \frac{W}{W + H} \sigma_0^\top \sigma_0 M3 \\
- \gamma (1 - \gamma) \frac{W^2}{(W + H)^2} \sigma_0^{M3} \sigma_0^\top \sigma_0 a_1^{M3} - (1 - \gamma) \frac{H W}{(W + H)^2} \sigma_0^\top \sigma_0 M3 \sigma_0^\top \rho_0 \sigma_0 \\
+ (1 - \gamma) \frac{W}{W + H} \left[ \sigma_0^{M3} \sigma_0^\top \lambda_1 + \lambda_0^\top \sigma_0 a_1^{M3} + \frac{W}{W + H} \delta^\top - \xi^\top \right] \right) \\
\tag{4.68}
\right]

where \( a_0^{M3} \) and \( a_1^{M3} \) represents the coefficients of affine optimal portfolio \( x^{M3} \), and can easily obtained from (4.32).

4.10.3.5 Proof of Proposition 4.4.5

The solution procedure of Model 4 follows the same steps as solving for Model 2. Hence, I show directly the explicit expression of the deterministic coefficients appeared in the indirect utility function (4.33).

\[ 
\Gamma_1^{M4}(\tau)' = \Gamma_1^{M4}(\tau)' (a_0^{M4}, a_1^{M4}) - \theta (1 - \gamma) \frac{W^2}{(W + H)^2} a_1^{M4} \sigma_0^\top \sigma_0 a_1^{M4} \] \\
\tag{4.70}
\left[ \Gamma_2^{M4}(\tau)' (a_0^{M4}, a_1^{M4}) - \theta (1 - \gamma) \frac{W^2}{(W + H)^2} a_0^{M4} \sigma_0^\top \sigma_0 a_1^{M4} \\
- \theta (1 - \gamma) W \frac{(W + H)^2}{(W + H)^2} e^\top \rho_0 \sigma_0 (W + H) e^\top \sigma_0 a_1^{M4} \right] \\
\tag{4.71}

The expressions of each coefficient is comparable to the corresponding parameters in Model 3, but under a different affine function optimal portfolio. I ignore the scaler term \( \Gamma_3^{M4} \), as it is not required in the optimal function function \( x^{M4} \).
4.10.3.6 Proof of Proposition 4.6.1

Define $\psi = \begin{pmatrix} \pi_t \\ g + \sigma^2 y^2 \end{pmatrix}$. I write $\hat{\psi}$ to denote an estimate. Hence the estimation error is defined as

$$
\begin{pmatrix}
\sigma_{\hat{\psi}}^2 e_1 \\
\sigma_y \rho_y e_2 + \sigma_y \sqrt{1 - ||\rho_y||^2} \gamma_2
\end{pmatrix} = \Sigma \Gamma^T
$$

(4.72)

where $\Gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}$ and $\Sigma = \begin{pmatrix} \sigma_y \rho_y e_2 & \sigma_y \sqrt{1 - ||\rho_y||^2} \gamma_2 \\
0 & \sigma_y \rho_y e_2 + \sigma_y \sqrt{1 - ||\rho_y||^2} \gamma_2
\end{pmatrix}$ Assume the perturbation vector is jointly normal with mean zero and volatility $\Sigma \Sigma^T$

$$
\sqrt{T} (\hat{\psi} - \psi) \sim N(0, \Sigma \Sigma^T)
$$

(4.73)

In other words, standardized error term square is chi-square distributed

$$
T (\hat{\psi} - \psi)^T (\Sigma \Sigma^T)^{-1} (\hat{\psi} - \psi) \sim \chi^2(2)
$$

(4.74)

Suppose the critical value at $\alpha$ significance level as $CV_{\alpha}$, then I can derive that

$$
(\hat{\psi} - \psi)^T (\Sigma \Sigma^T)^{-1} (\hat{\psi} - \psi) \leq \frac{CV_{\alpha}}{T} = \kappa^2
$$

The simplification of the constraint reaches Proposition 4.6.1.
Table 4.1: Estimation of Model Parameters.
Estimation result of the two-factor financial model described in Section 4.2.1. The financial model is estimated using monthly U.S. data on three bond yields, inflation and stock return over the period from June 1961 to December 2013. The bond maturities used in estimation are one-year, three-year and six-year. The three-year bond yield is assumed to be measured with error.

<table>
<thead>
<tr>
<th>Parameter Description</th>
<th>Estimate</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal Interest Rate $r_t = \delta_0 + \delta' X_t$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_0$</td>
<td>4.62%</td>
<td>2.11%</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.31%</td>
<td>0.10%</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>1.88%</td>
<td>0.05%</td>
</tr>
<tr>
<td>Expected inflation rate $\pi_t = \xi_0 + \xi' X_t$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\xi_0$</td>
<td>3.67%</td>
<td>0.26%</td>
</tr>
<tr>
<td>$\xi_1$</td>
<td>-0.14%</td>
<td>0.18%</td>
</tr>
<tr>
<td>$\xi_2$</td>
<td>1.43%</td>
<td>0.30%</td>
</tr>
<tr>
<td>Market price of risk $\Lambda_t = \lambda_0 + \lambda_1 X_t$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_0$</td>
<td>-0.1083</td>
<td>0.1313</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>-0.3451</td>
<td>0.1313</td>
</tr>
<tr>
<td>$\lambda_{11}$</td>
<td>-0.2551</td>
<td>0.1023</td>
</tr>
<tr>
<td>$\lambda_{12}$</td>
<td>0.4189</td>
<td>0.1316</td>
</tr>
<tr>
<td>$\lambda_{21}$</td>
<td>-0.1883</td>
<td>0.1011</td>
</tr>
<tr>
<td>$\lambda_{22}$</td>
<td>0.0914</td>
<td>0.1316</td>
</tr>
<tr>
<td>Latent factors $dX_t = -\kappa X_t dt + \sigma' \tilde{Z}_t$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\kappa_{11}$</td>
<td>0.3473</td>
<td>0.1025</td>
</tr>
<tr>
<td>$\kappa_{12}$</td>
<td>-0.4189</td>
<td>0.1316</td>
</tr>
<tr>
<td>$\kappa_{21}$</td>
<td>-0.1310</td>
<td>0.1029</td>
</tr>
<tr>
<td>$\kappa_{22}$</td>
<td>0.3192</td>
<td>0.1341</td>
</tr>
<tr>
<td>Mean reversion under risk neutral measure $\kappa^Q = \kappa + \sigma \tilde{\lambda}_1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\kappa_{11}^Q$</td>
<td>0.0922</td>
<td>0.0076</td>
</tr>
<tr>
<td>$\kappa_{12}^Q$</td>
<td>-0.3193</td>
<td>0.0194</td>
</tr>
<tr>
<td>$\kappa_{22}^Q$</td>
<td>0.4107</td>
<td>0.0260</td>
</tr>
<tr>
<td>Realized inflation process $d\Pi_t = \pi_t dt + \sigma_{\Pi} \tilde{Z}_t$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_{\Pi_1}$</td>
<td>0.16%</td>
<td>0.06%</td>
</tr>
<tr>
<td>$\sigma_{\Pi_2}$</td>
<td>0.10%</td>
<td>0.04%</td>
</tr>
<tr>
<td>$\sigma_{\Pi_3}$</td>
<td>-1.09%</td>
<td>0.08%</td>
</tr>
<tr>
<td>Stock return process $dS_t = (r_t + \eta) dt + \sigma' \tilde{Z}_t$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\eta_s$</td>
<td>5.59%</td>
<td>2.25%</td>
</tr>
<tr>
<td>$\sigma_{S_1}$</td>
<td>-0.84%</td>
<td>0.86%</td>
</tr>
<tr>
<td>$\sigma_{S_2}$</td>
<td>-2.26%</td>
<td>0.96%</td>
</tr>
<tr>
<td>$\sigma_{S_3}$</td>
<td>1.01%</td>
<td>0.79%</td>
</tr>
<tr>
<td>$\sigma_{S_4}$</td>
<td>15.31%</td>
<td>0.88%</td>
</tr>
<tr>
<td>Standard errors of yield measurement error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Sigma_e$</td>
<td>0.48%</td>
<td>0.11bp</td>
</tr>
</tbody>
</table>
Table 4.2: **Comparison of Sample Moments with Model Implied Distribution.**

Compare mean and standard deviation of inflation rates, stock returns and yields of three maturities that follow from the data and from the model using the estimated parameters reported from Table 4.1. The two estimated latent factors are estimated based on one-year and six-year yields. The distribution from both historical data and model estimation are on a monthly basis.

<table>
<thead>
<tr>
<th>average</th>
<th>volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data</td>
</tr>
<tr>
<td>Stock Return</td>
<td>0.89%</td>
</tr>
<tr>
<td>Inflation</td>
<td>0.33%</td>
</tr>
<tr>
<td>1 Year Yield</td>
<td>5.45%</td>
</tr>
<tr>
<td>3 Year Yield</td>
<td>5.85%</td>
</tr>
<tr>
<td>6 Year Yield</td>
<td>6.24%</td>
</tr>
</tbody>
</table>

Table 4.3: **Bond Risk Premia and Volatilities.**

This table displays the implied bond risk premia and return volatility with various maturities using the estimation results of Table 4.1. The market price of risk $\lambda_t$ is at the unconditional expectation with $X_t = 0_{2 \times 1}$.

<table>
<thead>
<tr>
<th>Bond Maturity</th>
<th>1 year</th>
<th>3 year</th>
<th>6 year</th>
<th>10 year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bond Premia $B(\tau)\sigma_X\lambda_0$</td>
<td>0.59%</td>
<td>1.39%</td>
<td>2.09%</td>
<td>2.62%</td>
</tr>
<tr>
<td>Volatility $\sqrt{B(\tau)^\top \sigma_X\lambda_0 B(\tau)}$</td>
<td>1.64%</td>
<td>4.10%</td>
<td>7.26%</td>
<td>10.82%</td>
</tr>
</tbody>
</table>

Table 4.4: **Correlation between returns on the stock, CPI and bonds with maturities of 1, 3, 6 and 10 years.**

The table depicts the correlation between stock return, inflation rate and returns on nominal bonds with maturities 1 year, 3 year, 6 year and 10 year on the basis of estimation results shown in Table 4.1. For example, the correlation between ten-year bond return and inflation rate is equal to $B(10)^\top \sigma_X \sigma_\Pi \left( B(10)^\top \sigma_X B(10) \right)^{-1/2} \left( \sigma_\Pi \sigma_\Pi \right)^{-1/2}$.

<table>
<thead>
<tr>
<th></th>
<th>Stock</th>
<th>Inflation</th>
<th>1 Year</th>
<th>3 Year</th>
<th>6 Year</th>
<th>10 Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stock return</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inflation</td>
<td>-0.0930</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Year bond return</td>
<td>0.1574</td>
<td>-0.1330</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Year bond return</td>
<td>0.1561</td>
<td>-0.1582</td>
<td>0.9511</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Year bond return</td>
<td>0.1407</td>
<td>-0.1678</td>
<td>0.8184</td>
<td>0.9559</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10 Year bond return</td>
<td>0.1246</td>
<td>-0.1663</td>
<td>0.6975</td>
<td>0.8848</td>
<td>0.9826</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 4.5: Optimal Portfolio Choice and Consumption without Labor Income. The optimal portfolio choice between 10-year nominal bonds, stock market and a nominal money market (in percentage of total wealth) as well as the optimal consumption over wealth ratio for investors at age 30, 40, 50 and 60. The state variables $X_1$ and $X_2$ are set equal to their unconditional mean, namely zero. Panel A is on the basis of Proposition 4.4.1. Panel B is based on Model 2 with the portfolio composition shown in Proposition 4.4.2. The penalty term $\theta$ is set equal to 4, which is approximately equivalent to 5% significance level of the perturbation distribution discussed in Figure 4.5.

<table>
<thead>
<tr>
<th>Age</th>
<th>10Year</th>
<th>Stock</th>
<th>10Year</th>
<th>Stock</th>
<th>10Year</th>
<th>Stock</th>
<th>10Year</th>
<th>Stock</th>
<th>10Year</th>
<th>Stock</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>40.46</td>
<td>17.38</td>
<td>84.06</td>
<td>-2.60</td>
<td>124.51</td>
<td>14.79</td>
<td>40.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>40.46</td>
<td>17.38</td>
<td>78.14</td>
<td>-2.57</td>
<td>118.60</td>
<td>14.82</td>
<td>40.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>40.46</td>
<td>17.38</td>
<td>64.46</td>
<td>-2.50</td>
<td>104.92</td>
<td>14.88</td>
<td>40.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>40.46</td>
<td>17.38</td>
<td>34.61</td>
<td>-2.30</td>
<td>75.07</td>
<td>15.08</td>
<td>43.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel A: Model 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel B: Model 2 with $\theta = 4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>40.54</td>
<td>16.94</td>
<td>84.18</td>
<td>-2.59</td>
<td>124.71</td>
<td>14.35</td>
<td>40.35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>40.54</td>
<td>16.94</td>
<td>78.24</td>
<td>-2.56</td>
<td>118.77</td>
<td>14.38</td>
<td>40.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>40.54</td>
<td>16.94</td>
<td>64.52</td>
<td>-2.50</td>
<td>105.06</td>
<td>14.44</td>
<td>40.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>40.54</td>
<td>16.94</td>
<td>34.63</td>
<td>-2.30</td>
<td>75.17</td>
<td>14.64</td>
<td>43.11</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.7: **Risk Shifting** $x^{M2} - x^{M1}$. The figure shows the change of asset allocation (in basis point) among different asset classes from an ambiguity neutral investor to an ambiguity aversion investor under different risk aversion and different state variables scenarios. The difference between the two optimal strategies ($x^{M1}$ and $x^{M2}$) is named risk shifting. The figure shows the risk shifting effect as a function of age. The state variable $X_1$ is assumed equal to zero. The other state variable $X_2$ is zero at the first column and is set equal to 1 and -1 in the other two column. The relation between state variables and interest rate or inflation rate is described in Figure 4.8.

(a) $\gamma = 2 \ X_2 = 0$  
(b) $\gamma = 2 \ X_2 = 1$  
(c) $\gamma = 2 \ X_2 = -1$

(d) $\gamma = 5 \ X_2 = 0$  
(e) $\gamma = 5 \ X_2 = 1$  
(f) $\gamma = 5 \ X_2 = -1$

(g) $\gamma = 10 \ X_2 = 0$  
(h) $\gamma = 10 \ X_2 = 1$  
(i) $\gamma = 10 \ X_2 = -1$
CHAPTER 4. HOW MUCH SHOULD INVESTORS ADAPT?

Figure 4.8: Human capital, financial wealth and their ratio. Panel (a) plots the value human capital, financial wealth and total wealth as a function of age. Human capital is calculated on the basis of (4.29) with $X_1 = X_2 = 0$ and the initial nominal income $Y_0 = 1$. Income volatility is set equal to $\sigma_y = 0.08$ and the expected return is assumed equal to $g + \frac{\sigma_y^2}{2} = 0.1$. Initial wealth is set to $W_0 = 10$ and is increasing with accumulative income over life cycle. Panel (b) shows the human capital over financial wealth under different age cohorts.

![Graph showing human capital, financial wealth, and total wealth vs. age](image)

Table 4.6: Optimal Portfolio Choice and Consumption with Labor Income.
The optimal portfolio choice (in percentage) on the basis of Model 3 and Model 4 for investors at age 30, 40, 50 and 60. The risk aversion level is set at $\gamma = 5$. The state variables $X_1$ and $X_2$ are set equal to zero, namely zero. The penalty term $\theta$ is set equal to 1, which is approximately equivalent to 10% significance level of the perturbation distribution discussed in Figure 4.5. The first six columns show each the demand for each hedging components which contains myopic demand, interest hedging ratio and income risk hedging demand. The next two columns aggregate the total hedging demand for the long-term bonds and the stocks. The last column gives the optimal consumption over total wealth ratio, which is $\frac{1}{\theta}$ for Model 3 and is $\frac{1}{\theta}$ for Model 4.

<table>
<thead>
<tr>
<th>Age</th>
<th>10Year Stock</th>
<th>10Year Interest Rate Hedge</th>
<th>10Year Income Risk Hedge</th>
<th>Total Portfolio</th>
<th>$\frac{1}{\theta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>344.03</td>
<td>147.82</td>
<td>-111.53</td>
<td>-29.96</td>
<td>2.49</td>
</tr>
<tr>
<td>40</td>
<td>98.87</td>
<td>42.483</td>
<td>52.19</td>
<td>-7.83</td>
<td>0.48</td>
</tr>
<tr>
<td>50</td>
<td>50.05</td>
<td>21.506</td>
<td>66.13</td>
<td>-3.35</td>
<td>0.08</td>
</tr>
<tr>
<td>60</td>
<td>41.44</td>
<td>17.807</td>
<td>35.01</td>
<td>-2.38</td>
<td>0.81bp</td>
</tr>
</tbody>
</table>

Panel A: Model 3

<table>
<thead>
<tr>
<th>Age</th>
<th>10Year Stock</th>
<th>10Year Interest Rate Hedge</th>
<th>10Year Income Risk Hedge</th>
<th>Total Portfolio</th>
<th>$\frac{1}{\theta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>344.21</td>
<td>146.79</td>
<td>-112.68</td>
<td>-29.96</td>
<td>4.98</td>
</tr>
<tr>
<td>40</td>
<td>98.92</td>
<td>42.193</td>
<td>52.234</td>
<td>-7.83</td>
<td>0.96</td>
</tr>
<tr>
<td>50</td>
<td>50.08</td>
<td>21.36</td>
<td>66.145</td>
<td>-3.35</td>
<td>0.16</td>
</tr>
<tr>
<td>60</td>
<td>41.47</td>
<td>17.68</td>
<td>35.014</td>
<td>-2.38</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Panel B: Model 4 with $\theta = 1$
Figure 4.9: **Robust Optimal Portfolio with Human Capital.** The figure presents the optimal asset allocation on different classes for investors with preferences of Model when both labor and parameter uncertainty are concerned. State variable $X_1 = 0$ and state variable $X_2 = -1$ for the first column and is 1 for the second column. The penalty parameter is assumed equal to $\theta = 1$.

(a) $\gamma = 2 \ X_2 = -1$

(b) $\gamma = 2 \ X_2 = 1$

(c) $\gamma = 5 \ X_2 = -1$

(d) $\gamma = 5 \ X_2 = 1$

(e) $\gamma = 10 \ X_2 = -1$

(f) $\gamma = 10 \ X_2 = 1$
Figure 4.10: **Risk Shifting** $x^{M4} - x^{M3}$. The figure presents the risk shifting (in percentage) effect for an investor with preferences moving from **Model3** to **Model 4**. The state variable $X_1$ is assumed equal to zero. The other state variable $X_2$ is zero at the first column and is set equal to 1 and -1 in the other two column.

(a) $\gamma = 2 \ X_2 = 0$

(b) $\gamma = 2 \ X_2 = 1$

(c) $\gamma = 2 \ X_2 = -1$

(d) $\gamma = 5 \ X_2 = 0$

(e) $\gamma = 5 \ X_2 = 1$

(f) $\gamma = 5 \ X_2 = -1$

(g) $\gamma = 10 \ X_2 = 0$

(h) $\gamma = 10 \ X_2 = 1$

(i) $\gamma = 10 \ X_2 = -1$
Figure 4.11: **Nature’s Perturbation Effect.** The figure presents nature’s optimal decision on the basis on Model 4, as well as the perturbation impact on equity and income returns.
Chapter 5

Conclusion

5.1 Summary of Findings

Robust policies provide economic decision makers with protection against unexpected disorders. The cost of robust policies is higher than that of naive policies. That being so, then who needs this protection? In this dissertation, I identify two types of investors who should be particularly afraid of model uncertainty, and who should consider employing a robust decision-making framework. First, investors under severe financial distress benefit strongly from robust polices. For example, pension funds with large solvency issues should consider choosing a prudential risk management strategy. To hedge large downside risks, these investors face substantial uncertainty from the equity market. Empirical evidence barely supports the idea of equity return predictability. Hence the higher the exposure to the equity market, the more uncertainty one may face. Optimistic scenarios that predict a high equity return as well as a high interest rate are surely attractive, but it is extremely dangerous for an insolvent investor to make a bet on them. A robust policy can however provide a successful hedge that is resilient and has a very low probability of breaking down during a crisis.

Secondly, robust financial policies are also attractive to a long-term investor. There are two reasons for this. The first still concerns the debate on financial market predictability. Even if historical data can provide some explanatory power due to autocorrelation effects, its forecasting ability is only limited to the short run. Secondly, it is difficult to stick to a single data generation process for decades. Literature on successful forecasting of very long-horizon equity return is very rare. Taken together, long-term investors are more likely to be ambiguity averse.
CHAPTER 5. CONCLUSION

5.1.1 Robustness is not Conservativeness

Most robust asset allocation studies such as Maenhout (2004) and Ju and Miao (2012) show that robust policies produce more cautious investment decisions. However, this does not mean that a robust decision equates to cutting down the risky portfolio. As shown in Chapter 3 when the hedging criterion is to replicate a fixed future payoff, the robust policy sometimes suggests a more aggressive portfolio. Therefore it is not so obvious how a robust investor should allocate her wealth. It depends on the hedging criterion and hedging instruments as well as the correlation between state variables.

5.2 Future Research

5.2.1 Mean Reversion Parameter Uncertainty

Existing studies of robust asset allocation on the basis of the Anderson et al. (2003) framework only consider the parameter misspecification of the first moment, namely, the expected equity returns. Chapter 2 also considers parameter uncertainty on drift terms. I employ drift distortions to capture the perturbation between the reference econometric model and the true data generation process of the sequence. It is well known that the expected asset return is very difficult to estimate based on historical data. Therefore, considering parameter uncertainty on drift terms is a valid starting point.

However, it is natural to consider parameter uncertainty of higher moments, such as mean reversion and time-varying volatility parameters. In Chapter 3, I explore the impact of market-price-of-risk misspecification on hedging demand for long-term bonds, for a robust investor with a long-dated liability. The calibration result shows that the speed of reversion parameter for the bond price diffusion process is very difficult to estimate. Further, I also find that the mean reversion parameter is highly correlated with other parameters, such as the long-run mean parameter and the volatility term as well as the market-price-of-risk parameters. However, the layout of the optimization problem in this paper does not allow for parameter uncertainty of other dimensions. Therefore considering mean-reversion parameter uncertainty could be a follow up project to this research.

Mean reversion is especially relevant in interest rate models. The mean reversion parameter not only influences the expected value of the bond return, but also determines the volatility of the bond diffusion process for longer-term bonds. As a result, the Anderson et al. (2003) framework is not applicable in this case. Since the uncertainty cannot simply be represented by additional drift terms, more dimensions are involved. I plan to use numerical methods to analyze the impact of mean-reversion parameter uncertainty on decision-making for long-term investors.
5.2. FUTURE RESEARCH

5.2.2 Consequence of Model Uncertainty with Investment and Borrowing Constraint

Chapter 4 develops a model to explain how income and inflation model uncertainty influence an individual investor’s optimal investment and consumption decisions. The model provides an investor with a robust investment and consumption strategy over the lifetime. Further, I provide a feasible boundary for the uncertainty aversion parameter, which measures the investor’s preference for robustness using an econometric estimate of the parameter estimation error. I find that robust investors would modify their investment portfolio by reducing the risk exposure on the stock market while increasing the long-term bond portfolio. The econometric boundary for uncertainty aversion parameter rules out a large part of the numerical solutions from the previous robust asset pricing studies such as Maenhout (2004), Branger et al. (2013) as these studies set the uncertainty aversion level too high, leading to a worst-case scenario that is statistically implausible.

From a methodological viewpoint, I use a Dynamic Programming approach in my research projects. In Chapter 4, I derive the robust optimal solution explicitly based on two crucial assumptions. First, I assume that the labor income diffusion process under the risk neutral measure is idiosyncratically risk free and is also model uncertainty neutral. As a result, human capital, the expectation of the discounted future cumulative income stream, is model-ambiguity neutral. Second, I ignore the short selling constraint and allow for borrowing from future income. Both assumptions are not realistic and may influence the robust decision-making.

However, if both assumptions are relaxed, it is impossible to obtain an analytical solution. The evolutionary research methodology proposed in my projects is called Robust least squares Monte Carlo algorithm (RLSMC). This is a numerical approach to solving dynamic programming problems inspired by the regression-based methods of Brandt et al. (2005) and Koijen et al. (2007).

In Chapter 3, I use the RLSMC algorithm to solve for the robust hedging problem. However, in the current setting, the optimal perturbations (also called nature’s decisions) are independent of the hedging decisions. In other words, nature’s decision depends only on the state variables, regardless of an investor’s behavior. As a result, the RLSMC algorithm in this project boils down to the classic LSMC algorithm together with a predetermined instantaneous nature’s decision.

In Chapter 4 however, I find that Mother Nature’s decisions on distortion parameters are influenced by the optimal portfolio choice. This property makes RLSMC a desirable algorithm.
Chapter 6


This chapter aims to elaborate the implementation of robust policy in real society. In particular, I consider two groups of people who should consider robust policies seriously, namely pension policy regulators as well the portfolio managers.

6.1 Pension Funds Regulators

The introduction of the Ultimate Forward Rate (UFR) has created a big debate among the Dutch Central Bank (DNB) and the European pension and insurance regulators. The UFR is a liability discounting approach adopted by the Dutch pension funds and insurance companies since September 2012. According to the UFR policy, interest rates beyond the last liquid point are required to be adjusted such that the one-year forward converges towards a fixed UFR of 4.2%.

The essential motivation of introducing the UFR is to provide an economic extrapolation method to place the market-consistent value on non-traded commitments, namely, ultra-long-dated liabilities. The long-maturity bond market is not liquid enough to approve a stable market value of the long-term payoff. The adoption of the UFR helps pension funds to smooth the market volatility at the long-end of the yield curve and makes them less dependent on the long-term interest rates.

In addition, the application of UFR valuation method also reduces the present value of long-term liabilities meanwhile raises funding ratios, because the UFR triggers higher interest rates with a maturity beyond 20 years (see Figure 6.1). A higher funding ratio based on the UFR policy benefits the current pensioners in the sense that the present

\footnote{Solvency 2 (Article 76) requires that the liability valuation method must be based on financial market data.}
CHAPTER 6. WHO NEEDS ROBUSTNESS?

Figure 6.1: Spot yield curve as a function of years to maturity. The red curve plots the UFR-adjusted interest rates and the blue curve is the Euro swap curve as of September 30, 2013. Recourse of the figure: Stef van Wseesel (2014), ‘Ultimate Forward Rate: does it create more risk?’.

retirees do not need to worry about a cut of their pension income and the employees are less likely to suffer from a higher contribution rate. However, is the fixed UFR policy always appealing?

Unfortunately, the fixed UFR makes the valuation approach fragile. For instance, if the yield is lower than assumed then the true funding ratio would be also lower than expected and hence will create a measurement error. This measurement error will hurt the future pensioners as there will be insufficient funds to payout the long-term commitment.

One of the reasons that makes the UFR policy fragile is that the fixed UFR is independent of the economic reality. Instead, the fixed rate of 4.2% is simply the sum of a long term inflation projection of 2% and an expected short-term bond rate of 2.2%. If the long-end rate rises, then the UFR would over estimate the value of liabilities. For instance, in 1994 the Fed increased short rates by 25 basis points, which resulted in a raise of long-end rates by 200-300 basis points in the western countries. In this case, the UFR is too prudential.

In October 2013, the UFR committee suggested to abandon the current method with a fixed rate of 4.2%. Instead, the committee proposed a moving-target policy in which the level of the UFR is based on the average 20-year forward rates of the last 120 months. The new policy is monthly adjustable hence can reflect the real economy and will reduce the measurement error to a large extend. Although the market adjusted UFR is much more sustainable than the fixed UFR, it is not resilient enough to protect against shocks. For the updated valuation method is based on the average return of the previous 10-year historical data, rather than extreme scenarios. It cannot smooth out the impact of
unexpected monetary policy or fiscal policy changes in the coming decades.

A more shock-resilient and stable approach is proposed in Chapter 3. Regulators can consider using the dynamic robust hedging method to price the long-dated commitments. There are three advantages of the dynamic robust hedging method. First, this is a market-consistent method. The robust yield curve moves with the instantaneous short rate. When the short rate goes down, the robust curve will shift down simultaneously and vice versa. Second, dynamic hedging can efficiently hedge the long-term risks as much as possible. As shown in Chapter 3, even without the preference for robustness, dynamic hedging can help to reduce the initial wealth requirement in order to meet the long-term payoff. Third, the max-min preference makes the policy robust against the uncertain future economy.

To sum up, the robust policy can not only reflect the movement of the real economy but can also hedge the unmanageable political risks significantly.

6.2 Portfolio Managers

By the end of fiscal year 2014, the accumulated deficit of U.S. public pension funds had reached $2 trillion.\(^2\) In less than a decade, the shortfall tripled. Growing downside risks, more volatile markets and slow recovery of the economy make pension funds to reconsider their asset allocations.

On one hand, low funding ratio can result in an excessive risk-taking to add returns. According to Rauh (2009), around 60% of pension assets are allocated in equity securities. Sharpe (1976) and Sharpe and Treynor (1977) also argue that underfunded pensions funds are more likely to have moral hazard (or risk shifting) incentives which characterize the potential of increasing risk exposure by pension sponsors. The creation of the federal Pension Benefit Guaranty Corporation (PBGC)\(^3\) even exacerbated the risk shifting incentives. On the other hand, from a risk management perspective, Rauh (2009) suggests that underfunded funds should reduce risk exposure to equities to avoid poor asset performance.

In Chapter 2, I develop a robust investment strategy for pension funds with large shortfalls. Theoretically, the robust portfolio has lower risk exposure compared with the benchmark expected shortfall hedging portfolio. However, the risk-taking is still substantial, especially for underfunded funds. For instance, when funding ratio is 85%, the robust portfolio puts 75% of the fund assets in risky securities, which is 10% less than the benchmark (see Figure 2.1). A robust policy can effectively improve the funding ratio while avoid moral hazard incentives.


\(^3\)PBGC provides an ultimate pension-income guarantee to workers for those pension funds with insufficient assets to meet their liabilities.
In practice, robust portfolios are not commonly used because of the limited knowledge about the max-min framework. However, an important takeaway from Chapter 2 is that low-risk and high-return stocks are desirable for underfunded pension funds. In an efficient market, such kind of free lunch does not exist. However, empirical evidence such as Black (1972) has shown that a positive risk-return relationship is not always supported by the data. Pioneer works by Miller and Scholes (1972) and Jensen et al. (1972) showed that low-beta stocks in US outperforms the Capital Asset Pricing Model (CAPM) expectations during the period (1931-1965). In other words, low-beta stocks combines both low risks and high returns. Therefore a low-beta stock could be an ideal instrument for investors with the preference for robustness. However a low-beta past does not necessarily mean a low-beta in the future. For instance, before the 2008 financial crisis Bank of America, General Electric and Microsoft all had betas below one.\textsuperscript{4} During the crisis, however, these stocks performed even worse than the S&P 500 which dropped 37\% by the end of the year 2008. This story tells that low-beta stocks are not always resilient as least during the market downturn.

References


REFERENCES


138
REFERENCES


REFERENCES


Jensen, M. C., Black, F., and Scholes, M. S. (1972). The capital asset pricing model: Some empirical tests. 43, 136


Koijen, R., Nijman, T., and Werker, B. (2007). Appendix describing the numerical method used in ‘when can life-cycle investors benefit from time-varying bond risk premia?’. Available at SSRN 945720. 55, 61, 131


Maenhout, P. (2004). Robust portfolio rules and asset pricing. Review of Financial Studies, 17(4):951. 7, 8, 9, 10, 12, 15, 20, 64, 79, 81, 82, 87, 89, 97, 114, 130, 131


140
REFERENCES


Morgenstern, O. and Von Neumann, J. (1953). Theory of games and economic behavior. 18


141
REFERENCES


Wilde, O. (1891). The picture of dorian gray. 1891. The Portable Oscar Wilde. 5


142
Nederlanse Samenvatting

Mijn onderzoek is geïnspireerd door het werk van Hansen and Sargent (2007) over het nemen van robuste economische beslissingen met een onzeker model. In dat werk vermoe- den zij dat het langzame respons van economische beslissers tijdens de recente recessie of voor het maken van fiscaal en monetair beleid verklaard kan worden door de onzek- erheid waarin zij verkeren. Echter, het werk door Hansen en Sargent is gelimiteerd tot macro-economische perspectieven en de parameter waarbij robuustheid wordt geprefer- eerd endogeen is beslist. Hoe kan modelonzekerheid en een voorkeur voor robuustheid een individuele belegger (zoals een lange-termijn belegger of een pensioenfonds) het maken van beslissingen beïnvloeden? Hoe meet men model onzekerheid in een incomplete finan- cieele wereld? Hoeveel moeten conservatieve beleggers hun gedrag aanpassen wanneer zij te maken krijgen met model onzekerheid? Welke factoren beslissen de voorkeur voor robuustheid? Is het vermijden van onzekerheid een begrensde parameter? Zo ja, hoe kan dit gemeten worden?

Mijn onderzoek zoekt naar antwoorden op deze vragen terwijl tegelijkertijd geprobeerd wordt om de gaten in de literatuur te vullen met betrekking tot financiële econometry en robuuste asset allocatie. Ik gebruik econometrische theorieën om aan te tonen dat het gedrag van beleggers om onzekerheid te vermijden erg afhankelijk is van hoever de onderliggende informatie onthuld is. Hoe meer gegevens, des te kleiner is de afkeer tegen onzekerheid bij een belegger. Deze eigenschap levert niet alleen een exogene grens aan de afkeer tegen onzekerheid-parameter, maar levert ook potentiële inzichten in verschillende asset pricing puzzels zoals voorkeur voor aandelen van de thuismarkt.

Mijn onderzoek interesse in risicomanagement van pensioenfondsen is gemotiveerd door mijn professionele ervaring als een junior onderzoeker in Netspar, waar ik actief betrokken was bij discussies over de meest recente pensioen problemen. Een voorbeeld is dat sinds de financiële crisis in 2008, de meeste pensioenfondsen geconfronteerd zijn met ernstige solvabiliteitsproblemen. De klassieke theorie over beleggingskeuze beargumenteert dat het optimale afdekkingsportefeuille onafhankelijk is van de dekkingsgraad. Echter, empirisch bewijs toont aan dat pensioenfondsen met lage dekkingsgraden eerder geneigd zijn te investeren in aandelen. Het afdekken van een langdurige en stochastis-
che kasstroom verplichtingen met onvoldoende geld is een uitdaging geworden voor de meeste pensioenfondsen. Dus richt mijn primaire interesse zich op pensioenfondsen met solvabiliteitstekorten.

Mijn onderzoek interesse in het verleden, het heden en de toekomst kunnen worden samengevat in drie categorieën: (1) Beslissingen nemen met model onzekerheid; (2) Afdekken en prijzen in incomplete markten; (3) Risicomanagement van pensioenfondsen.
Curriculum Vitae

Sally Muling Shen 沈牧龄 was born on 8 January in Nanjing, China. She earned her Bachelor of Science (2005-2008) under the program of International Economics and Finance at Tilburg University. During her second and third year of Bachelor studies, Sally gave tutorials of Mathematics 1 and Statistics 1 for the first year Bachelor students at Tilburg University. In 2009, she earned her Master of Science in Econometrics under the track of Quantitative Finance and Actuarial Science also at Tilburg University. During the final phase of her Master studies, Sally did a research project on pension funds indexation policy design for All Pension Group (APG) in Amsterdam. The project was supervised by Prof. Eduard Ponds and Prof. Hans Schumacher. After graduation, Sally continued her research internship at APG for another half year.

In 2010, Sally joined the Finance Department of Maastricht University as a PhD Candidate. Meanwhile, she became a junior research fellow of NETSPAR (Network for Studies on Pensions, Aging and Retirement). In Maastricht University, she also taught master courses on Financial Research Methods. Her research aims to answer: How to make an investment decision when an economic agent does not trust his or her model? Her dissertation “Robust Asset Allocation in Incomplete Markets” provides a powerful decision-making rule that remains stable in the presence of model uncertainty. Her research ideas and findings had been presented at academic and industry conferences, such as China International Conference in Finance (CICF), European Economic Association (EEA) and Bachelier Finance.

In September 2015, Sally will start as an Assistant Professor in Finance at the International School of Economics and Management at Capital University of Economics and Business in Beijing, China. Her fields of interest include actuarial science, financial engineering, asset pricing, financial econometrics, portfolio choice and pension risk management. Her research focuses on dynamic asset allocation under model misspecification, pricing and hedging in incomplete financial markets and robust life-cycle investments.
About the Hand-Drawn Cover

The cover of this book is designed and drawn by Sally Banana
The Artwork on the cover is entitled "When the Black Swan Meets the Phoenix"