

Portfolio Choice with General Pricing Kernel

BY

ZHIFAN ZHANG

B.S. (Central University of Finance and Economics) 2007

M.S. (University of Chicago) 2008

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Chicago, 2014

Chicago, Illinois

Defense Committee:

Jie Yang, Chair and Advisor

Jing Wang

Ryan Martin

Cheng Ouyang

Fangfang Wang, Department of Information and Decision Sciences

Copyright by

Zhifan Zhang

2014

To my family and friends.

ACKNOWLEDGMENTS

I wish to express my grateful appreciation to my advisor Professor Jie Yang, for leading me through this long and arduous process. His extraordinary support and enlightening guidance have profoundly influenced my study, research and future career.

I would love to thank all my committee members, Professor Jing Wang, Professor Ryan Martin, Professor Cheng Ouyang and Professor Fangfang Wang, for their tremendous help and valuable opinions in my research work and dissertation.

I would love to thank Professor Qingshuo Song for his tremendous help and valuable opinions in my research work and dissertation.

I want to give my thanks for Professor Samad Hedayat, Professor T.E.S. Raghavan, Professor Jie Yang, Professor David Nicholls, Professor Jing Wang, Professor Fangfang Wang, Professor Junhui Wang, and Professor Irina Nenciu for their great help in my course work and research work throughout my years at UIC.

I would love to thank Professor Ramin Takloo-Bighash, Ms. Maureen Madden and other staff members in Department of Mathematics, Statistics, and Computer Science for their kind-

ACKNOWLEDGMENTS (Continued)

ness and help.

Last, but not the least, I want to thank my husband, my parents, my grandparents and my friends. They provided endless encouragement and support through the ups and downs of the entire process. They gave me space and time when I was busy, they listened when I wanted to complain, and congratulated me for the success both small and large.

TABLE OF CONTENTS

<u>CHAPTER</u>		<u>PAGE</u>
1	INTRODUCTION	1
1.1	Introduction to Portfolio Choice Problem	2
1.2	Neyman-Pearson Lemma and Optimal Tests	6
1.2.1	Neyman-Pearson Lemma in Hypothesis Testing	6
1.2.2	Alternative Description of Neyman-Pearson Lemma	9
1.2.3	Pure and Randomized Hypothesis Tests	10
1.2.4	Composite Hypothesis Tests	12
1.2.5	Generalized Neyman-Pearson Lemma	13
1.2.6	Optimal Tests Analysis	13
1.3	Portfolio Choice Optimization	15
1.3.1	Models for Portfolio Optimization	16
2	GOAL REACHING MODEL	19
2.1	Introduction and Model Setup	20
2.2	Modified Optimization Problem of Goal Reaching Model	21
2.3	Solution of Goal Reaching Model with Atomic Pricing Kernel	24
2.4	Example: Pricing Kernel Follows Geometric Brownian Motion	26
2.4.1	Explicit Solution with Positive Cumulative Sharpe Ratio	27
2.4.2	Explicit Solution with Zero Cumulative Sharpe Ratio	28
2.4.3	Back Tests and Numerical Results	29
3	YAARI'S DUAL MODEL	36
3.1	Introduction and Model Setup	37
3.2	Properties of an Optimal Solution	39
3.2.1	Properties of Optimal Solution on Non-atomic ρ	39
3.2.2	Properties of Optimal Solution on Atomic ρ	40
3.2.3	Example of Extreme Case: Cantor Function ρ	50
3.3	Solution of Modified Optimization Problem	51
3.3.1	Form of the Optimal Solution to a Modified Optimization Problem	51
3.3.2	Discussion on Quantile Functions	53
3.3.3	Algorithm: Search-and-Cut	64
3.3.4	Proof of Lemma 7	66
3.3.5	Example of Extreme Case: Infinite Number of Monotone Pieces	72
3.3.6	Derivation of Optimal Solution	76
3.3.7	Existence and Uniqueness of Optimal Solution	78
3.4	Optimal Solution to Yaari's Dual Model	87

TABLE OF CONTENTS (Continued)

<u>CHAPTER</u>		<u>PAGE</u>
4	CONCLUSION	91
	CITED LITERATURE	94
	VITA	100

LIST OF TABLES

<u>TABLE</u>		<u>PAGE</u>
I	COMPARISON BETWEEN THE GRID SEARCH SOLUTION AND EXPLICIT SOLUTION	30

LIST OF FIGURES

<u>FIGURE</u>		<u>PAGE</u>
1	Convergence of Two Optimal Values	32
2	Outperformance Probability versus Initial Capital with Different Sharpe Ratios	33
3	Outperformance Probability versus Initial Capital with Different Interest Rates	35
4	Graph of X Non-increasing on $A_i \cup A_j$	44
5	Graph of C_l , A_i , and C_r	47
6	Case I of Lemma 5	47
7	Case II of Lemma 5	48
8	Cantor Function ρ	52
9	Graph of F_X	54
10	Graph of Left-continuous Quantile Function $F_X^{\leftarrow}(z)$	55
11	Graph of Right-continuous Quantile Function $F_X^{\rightarrow}(z)$	56
12	Graph of an Artificial Distortion Function w	73
13	Graph of $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ with Infinite Number of Monotonic Pieces in Example 4	74

SUMMARY

The portfolio choice optimization problem we study in this thesis is to construct a continuous-time portfolio which maximizes the probability of outperformance. In the literature of mathematical finance, this type of problem is typically solved by the quantile approach, which requires a non-atom pricing kernel.

In real financial practice, the pricing kernel can be atomic, i.e., the probability that the pricing kernel equals to a constant can be positive. For example, an extreme case is that the pricing kernel equals to a constant with probability 1 (see Example 1 in Section 2.4). Another example is the scenario analysis in risk management. Risk analysis is done by setting the asset price to be certain extreme values. In this case, the pricing kernel is atomic at those extreme values.

In this thesis, we consider two portfolio choice optimization models, goal reaching model and Yaari's dual model, with more general pricing kernels which may allow the existence of atoms.

For goal reaching model, we discuss the properties of the solution, and derive a modified optimization problem, which has a similar mathematical format to the optimal hypothesis test problems. Therefore, a general solution scheme for both non-atomic and atomic pricing kernel is derived based on a generalized Neyman-Pearson Lemma, which is famous in classical statistical theory. We also provide an example with pricing kernel follows geometric Brownian

SUMMARY (Continued)

motion, to show the explicit solution based on our results. Our numerical experiments validate the optimal solution as well.

For Yaari's dual model, we discuss the properties of optimal solution that is an optimal terminal cash flow which is nonincreasing with respect to the pricing kernel. The pricing kernel here could contain atoms and thus is more general than non-atomic ones. Under the assumption that probability distortion/weighting is differentiable, we derive a modified optimization problem that contains left-continuous quantile function of the pricing kernel and terminal cash flow. A sub-optimization problem with Lagrange multiplier is studied. We propose an algorithm, called Search-and-Cut Algorithm to find the optimal solution, which is good for cases where the weighting/pricing-kernel ratio consists of a finite number of monotone pieces. We prove the existence and uniqueness of the optimal solution as well. Finally, we derive an optimal solution of Yaari's dual model for more general pricing kernels and probability distortions.

The approaches we propose in this thesis could be used for other portfolio choice models, as well as for problems solved by non-atomic quantile approaches.

CHAPTER 1

INTRODUCTION

1.1 Introduction to Portfolio Choice Problem

Given an initial wealth x_0 , a benchmark b and an investment time horizon $[0, T]$ are chosen by the investment agent. The benchmark b can be set up as a value (usually higher than initial wealth) or a market index (e.g., the S&P 500 index). The terminal value is the value of portfolio at the end of this time horizon, i.e., at time T . It is called outperformance, if the terminal value of the portfolio exceeds the benchmark. The portfolio choice optimization problem we study in this thesis is to construct a continuous-time portfolio which maximizes the probability of outperformance.

Continuous-time portfolio choice has been studied by Samuelson (1969) and Merton (1969) with discussion on expected utility maximization. Utility Functions can be used in portfolio allocation, taking into account investors' preferences to return and risk. Portfolios need to be developed to satisfy investors' goal on return, and at the same time, take the risk associated with market events and price changes under consideration.

There is no portfolio that could maximize the return while minimizing the risk. A utility function indicates, for a given level of return, how much satisfaction or utility we get. In most of the cases, people prefer higher level of return to lower level return, that is the utility function is increasing with respect to level of return. If the marginal utility decreases on level of return, then the investor is said to be risk averse. Therefore, the portfolio allocation problem can be transformed into an expected utility maximization problem (see Section 1.3.2 Models for Portfolio Optimization).

Essentially, there are two ways to solve this type problem. One is to employ "stochastic control

or dynamic programming approach”, which was initially proposed by Merton (1969, 1971). Under this approach, the problem is transformed into “solving a partial differential equation, named the Hamilton-Jacobi-Bellman (HJB) equation”. Based on Bellman (1957, 1962), dynamic programming is an optimal technique that finds the decision which maximizes the expected value given a dynamic model of state behavior and value. The optimization problem is broken down into subproblems, a sequence of decision steps over time $t_1, t_2, \dots, t_i, \dots, t_n$. At time t_i , we define s_{t_i} as the state, d_{t_i} as the decision, and w_{t_i} as the state uncertainty. The state transition is defined by an arbitrary function f :

$$s_{t_{i+1}} = f(s_{t_i}, d_{t_i}, w_{t_i})$$

where $s_{t_{i+1}}$ represents the new state at time t_{i+1} , which is influenced by the state, decision and state uncertainty at time t_i . We can write the value function as:

$$V_{t_i}(s_{t_i}) = E[G(s_{t_i}, d_{t_i}, w_{t_i}) + V_{t_{i+1}}(s_{t_{i+1}})]$$

where G is the value for current state, and V is the value function at a certain time and state. $V_{t_i}(s_{t_i})$ is the value function at time t_i at state s_{t_i} , is the expected value of the summation of current value and value at time t_{i+1} , and state $s_{t_{i+1}}$. At each time t_i , the optimal strategy is to choose d_{t_i} such that the value is maximized:

$$V_{t_i}^*(s_{t_i}) = \max_{d_{t_i}} E[G(s_{t_i}, d_{t_i}, w_{t_i}) + V_{t_{i+1}}(s_{t_{i+1}})]$$

This is also known as the discrete-time Bellman Equation, and $V_{t_i}^*(s)$ can be found by backward induction. $V_{t_i}^*(s)$ at any state s_i is calculated from $V_{t_{i+1}}^*(s)$ by maximizing a value function of gain at time t_{i+1} based on the decision made at time t_i . Since terminal time value $V_{t_n}^*(s)$ is given, the above operation yields $V_{t_i}^*(s)$, $i = 1, 2, \dots, n - 1$. $V_{t_1}^*(s)$ is the initial value of the optimal solution.

The other approach is to apply the martingale characterization, that is, the discounted value of a portfolio under the risk neutral probability measure is a martingale. In financial models, we set up a sample space as the set of possible scenarios for the future, which has an actual probability measure \mathbb{P} . However, for the purpose of pricing derivatives or portfolios, we use a risk-neutral measure $\tilde{\mathbb{P}}$. The two measures are equivalent, since they agree on possible events, even though they may not agree on the probability of these events. In the continuous-time market, if we set up hedges that work under probability measure $\tilde{\mathbb{P}}$, then these hedges should also work with \mathbb{P} probability measure. Based on Girsanov's Theorem, one could verify that the discounted value of a stock/portfolio under $\tilde{\mathbb{P}}$ is a martingale, that is,

$$d(D(t)X(t)) = \Delta(t)\sigma(t)D(t)S(t)d\tilde{W}(t)$$

where $D(t)$ is the discounted function on time t , $X(t)$ is the portfolio capital, $\Delta(t)$ is the shares held in the portfolio, $S(t)$ is the stock price, $\sigma(t)$ is the volatility of the stock, and $\tilde{W}(t)$ is a Brownian Motion under risk-neutral probability measure. There is no drift term, the expected value at time t should always equal to the initial capital $X(0)$. One can get $\tilde{E}[X(t)] = X(0)$.

Let $\rho = \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}$ be the Radon-Nikodým derivative of $\tilde{\mathbb{P}}$ with respect to \mathbb{P} , then $E[\rho X(t)] = X(0)$, we also consider ρ as the pricing kernel. This approach uses a budget constraint (initial capital constraint) for dynamic wealth and then achieves the optimal terminal cash flow by solving the optimization problem.

Since the expected utility measurement, as an objective function has a conflict while maximizing return and minimizing risk, many alternative risk measures have been discussed, for example, Yaari's dual theory of choice with a probability distortion by Yaari (1987), and goal reaching problem by Kulldorff (1993). Dynamic programming and HJB equation were employed to solve these problems, see Browne (1999, 2000) and Hamada, Sherris, and van der Hoek (2006). The distortion probability, which represents the subjective inflation/deflation of the true probability, is usually non-linear, and it rises difficulties for dynamic programming approach. To avoid this type of problem, He and Zhou (2011) provided solution scheme to the quantile function of the terminal cash flow, but it highly depends on the requirement that the pricing kernel is atomless.

In real financial practice, the pricing kernel can be atomic, i.e. the probability of price equals to a constant can be positive. For example, an extreme case is that the pricing kernel equals to a constant with probability 1 (see Example 1 in Section 2.4). Another example is the scenario analysis in risk management. Risk analysis is done by setting the asset price to be certain extreme values. In this case, the pricing kernel is atomic at those extreme values.

As far as we know, atomic pricing kernel were not considered in the current literature. In this thesis, we apply a generalized Neyman-Pearson Lemma, which is famous in classical statistical

theory, and a modified left-continuous quantile function to find the optimal terminal cash flow, for more general pricing kernels, which may allow the existence of atoms.

In Section 1.2, we introduce the statistical background of Neyman-Pearson Lemma and Optimal Tests. In Section 1.3, the financial background of the continuous-time market and models for portfolio optimization are discussed.

Chapter 2 describes a general solution scheme for both non-atomic and atomic pricing kernel for goal reaching model. In Chapter 3, we propose an algorithm for Yaari's model, called Search-and-Cut Algorithm to find the optimal solution, which is good for cases where the distortion/pricing-kernel ratio consists of a finite number of monotone pieces. Finally, Chapter 4 concludes.

1.2 Neyman-Pearson Lemma and Optimal Tests

1.2.1 Neyman-Pearson Lemma in Hypothesis Testing

Hypothesis testing is the use of statistics to choose between competing hypotheses about a probability distribution, based on observed data from the distribution. We are interested in a random variable X which has *p.d.f.* or *p.m.f.* $f(x; \theta)$, where $\theta \in \Omega$. We assume that $\theta \in \omega_0$ or $\theta \in \omega_1$, where ω_0 and ω_1 are disjoint subsets of Ω and $\omega_0 \cup \omega_1 = \Omega$. We label the hypotheses as

$$H_0 : \theta \in \omega_0 \text{ versus } H_1 : \theta \in \omega_1.$$

The Hypothesis H_0 is referred to as the null hypothesis, while H_1 is referred to as alternative hypothesis. The hypothesis test is based on a sample X_1, \dots, X_n from the distribution of X .

Let S denote the support of the random sample $X' = (X_1, \dots, X_n)$.

A test of H_0 versus H_1 is based on a subset C of S . This set C is called the critical region and its corresponding decision rule is:

Reject H_0 if $X \in C$

Retain H_0 if $X \in C^c$

To summarize the results of the hypothesis test in terms of the true state of nature, besides the correct decisions, two errors can occur. A Type I error occurs if H_0 is rejected when it is true, while a Type II error occurs if H_0 is accepted when H_1 is true. The size or significant level of the test is the probability of a Type I error:

$$\alpha = \max_{\theta \in \omega_0} P_{\theta}(X \in C)$$

The power of the test is given by:

$$\gamma_C(\theta) = P_{\theta}(X \in C); \theta \in \omega_1$$

For testing a simple hypothesis H_0 against a simple alternative H_1 . Let $f(x; \theta)$ denote the *p.d.f.* or *p.m.f.* of a random variable X , where $\theta \in \Omega = \{\theta', \theta''\}$. Let $\omega_0 = \theta'$ and $\omega_1 = \theta''$.

In classical statistical textbook (see Hogg, McKean, and Craig (2005)) Neyman-Pearson Theorem could be expressed as:

Let X_1, X_2, \dots, X_n denote a random sample from a distribution that has p.d.f. or p.m.f. $f(x; \theta)$.

Then the likelihood of X_1, X_2, \dots, X_n is

$$L(\theta; x) = \prod_{i=1}^n f(x_i; \theta), \text{ for } x' = (x_1, x_2, \dots, x_n).$$

Suppose $\theta' \neq \theta''$, and let k be a positive number. Let C be a subset of \mathbb{R}^n such that

- (a) $\frac{L(\theta'; x)}{L(\theta''; x)} \leq k$, for each point $x \in C$
- (b) $\frac{L(\theta'; x)}{L(\theta''; x)} \geq k$, for each point $x \in C^c$
- (c) $\alpha = P_{H_0}[X \in C]$

Then C is a best critical region of size α for testing the simple hypothesis $H_0 : \theta = \theta'$ against the alternative simple hypothesis $H_1 : \theta = \theta''$.

As stated in the theorem, conditions (a), (b), (c) are sufficient for the region C to be a best critical region of size α . They are also necessary.

1.2.2 Alternative Description of Neyman-Pearson Lemma

An alternative description of Neyman-Pearson Lemma, is first proposed by Lehmann (1986) and was further discussed by Cvitanic and Karatzas (2001).

On a measurable space (Ω, \mathcal{F}) , we want to discriminate between two probability measures/distributions:

$$P \longrightarrow \text{measure/distribution under } H_0$$

$$Q \longrightarrow \text{measure/distribution under } H_1$$

Neyman-Pearson Lemma assures that there exists a randomized test

$$X : \Omega \rightarrow [0, 1]$$

which maximizes the power of the test: $E^Q(X)$, while keeping the level of type I error: $E^P(X) = \alpha$. Furthermore, the optimal test X takes the form of

$$\hat{X} = \mathbf{1}_{\{\frac{dQ}{dP} > z\}} + B\mathbf{1}_{\{\frac{dQ}{dP} = z\}}$$

for some $z > 0$ with probability 1. In this thesis, the pricing kernel ρ is defined as $\frac{d\tilde{P}}{dP}$, where P (Q in this section) is the real probability measure, and \tilde{P} (P in this section) is the risk-neutral probability measure.

1.2.3 Pure and Randomized Hypothesis Tests

A pure test is a random variable $X : \Omega \rightarrow \{0, 1\}$. Consider the hypothesis testing problem

$$\begin{aligned} \max \quad & Q(X = 1) \\ \text{subject to} \quad & P(X = 1) \leq \alpha \end{aligned} \tag{1.1}$$

Suppose μ is another probability measure, which satisfy

$$P \ll \mu \text{ and } Q \ll \mu$$

Based on change of measure, let G and H be Radon-Nikodým derivatives

$$G := \frac{dQ}{d\mu} \text{ and } H := \frac{dP}{d\mu}$$

Neyman-Pearson Lemma assures that solution takes the form of

$$\begin{aligned} \hat{X} &= \mathbf{1}_{\{\frac{dQ}{dP} > \hat{z}\}} \\ &= \mathbf{1}_{\{\frac{\frac{dQ}{d\mu}}{\frac{dP}{d\mu}} > \hat{z}\}} \\ &= \mathbf{1}_{\{\frac{G}{H} > \hat{z}\}} \\ &= \mathbf{1}_{\{\hat{z}H < G\}} \end{aligned}$$

for some $\hat{z} \in (0, \infty)$. Note that it is possible that $P(\hat{z}H < G) < \alpha$.

For randomized hypothesis tests, in order to maximize the power of the test, we want to find a randomized hypothesis test $X : \Omega \rightarrow [0, 1]$, which solves

$$\begin{aligned} \max \quad & E^Q(X) \\ \text{subject to} \quad & E^P(X) \leq \alpha \end{aligned} \tag{1.2}$$

The optimal problem is equivalent to

$$\sup_{X \in \mathcal{X}_\alpha} E^Q(X)$$

where $\mathcal{X}_\alpha := \{X : \Omega \rightarrow [0, 1]; E^P(X) \leq \alpha\}$, for any given significant level $\alpha \in (0, 1)$.

Neyman-Pearson Lemma assures that the optimal solution for this randomized test is

$$\hat{X} = \mathbf{1}_{\{\hat{z}H < G\}} + b\mathbf{1}_{\{\hat{z}H = G\}}$$

with some $\hat{z} \in (0, \infty)$ and $b \in [0, 1]$. This is randomized test, since b is a random variable on $[0, 1]$ and $E[b] = \frac{\alpha - P[\hat{z}H < G]}{P[\hat{z}H = G]}$. Note that $E^P(\hat{X}) = \alpha$.

1.2.4 Composite Hypothesis Tests

Suppose on the measurable space (Ω, \mathcal{F}) , we have a family \mathbf{P} of probability measures (the composite null hypothesis), and another family \mathbf{Q} of probability measures (the composite alternative hypothesis). There is some other probability measure μ . We assume

$$\mathbf{Q} \cap \mathbf{P} = \emptyset$$

$$Q \ll \mu, P \ll \mu, \forall Q \in \mathbf{Q}, \forall P \in \mathbf{P}$$

Based on change of measure, let G_Q and H_P be two sets of Radon-Nikodým derivative. Define the random variables

$$G_Q := \frac{dQ}{d\mu} \quad (Q \in \mathbf{Q})$$

$$H_P := \frac{dP}{d\mu} \quad (P \in \mathbf{P})$$

and the set of feasible randomized hypothesis tests

$$\mathbf{X}_\alpha := \{X : \Omega \rightarrow [0, 1]; E^P(X) \leq \alpha, \forall P \in \mathbf{P}\}$$

1.2.5 Generalized Neyman-Pearson Lemma

Since we have a family \mathbf{Q} , which contains alternative hypotheses, we may use the *max-min criterion*, and let

$$\underline{V}(\alpha) := \sup_{X \in \mathbf{X}_\alpha} (\inf_{Q \in \mathbf{Q}} E^Q(X)) \quad (1.3)$$

Following Cvitanic, J. and Karatzas, I. (2001), there exists an optimal max-min randomized test taking the form of

$$\hat{X} = \mathbf{1}_{\{\hat{z}\hat{H} < \hat{G}\}} + B\mathbf{1}_{\{\hat{z}\hat{H} = \hat{G}\}}$$

where B is a random variable with values in $[0, 1]$; $\hat{G} = \frac{d\hat{Q}}{d\mu}$ is a random variable for some $\hat{Q} \in \mathbf{Q}$, given that \mathbf{Q} is convex and closed. While \hat{H} is a random variable chosen from the closure of the convex hull of $\{\frac{dP}{d\mu}, P \in \mathbf{P}\}$, and \hat{z} is a suitable positive number (see also Leung, Song, and Yang (2013)).

1.2.6 Optimal Tests Analysis

For a randomized composite hypothesis testing problem, let

$$\mathbb{X}_\alpha^0 := \{X : \Omega \rightarrow [0, 1], E[HX] \leq \alpha, \forall H \in \mathbf{H}\}$$

where $\mathbf{H} = \{\frac{dP}{d\mu}, P \in \mathbf{P}\}$. This is equivalent to solve:

$$V_0(\alpha) := \sup_{X \in \mathbb{X}_\alpha^0} (\inf_{G \in \mathbf{G}} E[GX]) \quad (1.4)$$

where $\mathbf{G} = \{\frac{dQ}{d\mu}, Q \in \mathbf{Q}\}$.

For a pure composite hypothesis testing problem, define

$$\mathbb{X}_\alpha^1 := \{X : \Omega \rightarrow \{0, 1\}, E[HX] \leq \alpha, \forall H \in \mathbf{H}\}$$

Hypothesis testing problem can also be equivalently expressed as:

$$V_1(\alpha) := \sup_{X \in \mathbb{X}_\alpha^1} (\inf_{G \in \mathbf{G}} E[GX]) \quad (1.5)$$

From the definitions, we could see that $V_0(\alpha) \geq V_1(\alpha)$. When $V_0(\alpha) = V_1(\alpha)$, the best pure hypothesis test performs as well as the best randomized test.

Leung, Song, and Yang (2013) discussed sufficient conditions for equivalent optimal values. Following (C1) in Theorem 2.10, suppose that G and H are singletons, and there exists an \mathcal{F} -measurable random variable with a continuous c.d.f. with respect to \mathbb{P} ; then $V_0(x) = V_1(x)$, and there exists an indicator function \hat{X} that maximizes $V_0(x)$ and $V_1(x)$ simultaneously. Furthermore, $x \mapsto V_1(x)$ is continuous, concave, and non-decreasing.

There are relevant literature in Financial Mathematics that used pure hypothesis test to solve optimization problems. For Example, Outperformance portfolio optimization: Browne (1999), Spivak and Cvitanić (1999), Bayraktar, Huang, and Song (2011), Quantile hedging: Föllmer and Leukert (1999) show that the quantile hedging problem can be formulated as a pure hypothesis testing problem, and Shortfall risk minimization: Cvitanić (2000), Föllmer and Schied (2002), Schied (2004), Rudloff (2007).

In this thesis, we will apply the generalized Neyman-Pearson Lemma to the Goal Reaching problem in Chapter 2.

1.3 Portfolio Choice Optimization

Continuous-time portfolio choice has been studied by Samuelson (1969) and Merton (1969) with discussion on expected utility maximization. Two typical ways of solving this type of problem are dynamic programming approach by Merton (1969, 1971) and the martingale approach. In a complete market, an optimal strategy can be derived by replicating the optimal terminal cash flow.

If the market is incomplete with possible portfolio constraints, the martingale approach is further studied by adding convex duality machinery, see Cvitanic and Karatzas (1992), Kramkov and Schachermayer (1999) and Goll and Ruschendorf (2001).

In this thesis, we discuss the problem under a continuous-time market.

Let $T > 0$ be a given terminal time point, and $(\Omega, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbf{P})$ is a filtered probability space, on which one could defined a standard \mathcal{F}_t measurable Brownian motion $W(t)$, with $W(0) = 0$. It is assumed that $\mathcal{F}_t = \sigma\{W(s) : 0 \leq s \leq t\}$.

1.3.1 Models for Portfolio Optimization

There are three models which are commonly used to solve portfolio optimization. Let X be the terminal cash flow and ρ be the pricing kernel at time T . These models are different in terms of the mathematical formulations and economical interpretations, but all three could employ the characteristic of martingale, that is, to solve the optimal problem with a static budget constraint, i.e., $E[\rho X] = x_0$.

Model 1: Expected Utility Maximization

$$\begin{aligned} \max_X \quad & Eu(X) \\ \text{subject to} \quad & E[\rho X] = x_0, X \geq 0, X \in \mathcal{F}_T \end{aligned}$$

where $u(\cdot)$ is a utility function. $E[\rho X] = x_0$ is the budget constraint. $X \geq 0$ is the no-bankruptcy assumption. This model was first proposed and studied by Samuelson (1969) and Merton (1969). The solution X^* to this optimization problem is the target cash flow at the terminal time T . The optimal portfolio will be the one that replicates X^* .

Since the expected utility measurement, as an objective function has a conflict while maximizing return and minimizing risk, many alternative risk measures have been discussed, for example,

Yaari's dual theory of choice with a probability distortion by Yaari (1987), and goal-reaching problem by Kulldorff (1993).

Model 2: Goal Reaching

$$\begin{aligned} \max_X \quad & P(X \geq b) \\ \text{subject to} \quad & E[\rho X] = x_0, X \geq 0, X \in \mathcal{F}_T \end{aligned}$$

where b is the goal or benchmark that is to be reached at terminal time T . This model is initialed by Kulldorff (1993) and further studied by Browne (1999, 2000). Browne (1999, 2000) solved this model by employing dynamic programming and HJB equation.

Model 3: Yaari's Dual Theory

$$\begin{aligned} \max_X \quad & \int_0^\infty w(P(X > x)) dx \\ \text{subject to} \quad & E[\rho X] = x_0, X \geq 0, X \in \mathcal{F}_T \end{aligned}$$

where $w : [0, 1] \rightarrow [0, 1]$ is a probability distortion or weighting function, which may represent a subjective inflation or deflation of the true probability. This model was first proposed by Yaari (1987), including the risk preference by $w(\cdot)$, for example, $w(\cdot)$ is convex, if the investor is risk-averse.

In this thesis, Model 2 and Model 3 are further discussed. Methods in the current literature require a non-atom pricing kernel. Here, new methods are derived to solve these two models for

more general pricing kernel, and new algorithm are designed and implemented to find optimal solutions.

CHAPTER 2

GOAL REACHING MODEL

In this chapter, we discuss the properties of the optimal solution, and derive a modified optimization problem which has similar mathematical format to the optimal hypothesis test problems. Therefore, a general solution scheme for both non-atomic and atomic pricing kernel is derived based on a generalized Neyman-Pearson Lemma. We also provide an example where the pricing kernel follows geometric Brownian motion, to show the explicit solution based on our results.

2.1 Introduction and Model Setup

The Goal Reaching Model was first proposed by Kulldorff(1993) and studied extensively by Browne(1999), He and Zhou (2011). The Goal Reaching problem is to solve

$$\begin{aligned} \max_X \quad & P[X \geq b] \\ \text{subject to} \quad & E[\rho X] = x_0, X \geq 0, X \in \mathcal{F}_T \end{aligned} \tag{2.1}$$

where $\mathcal{F}_T = \sigma\{W(s) : 0 \leq s \leq T\}$, $b > 0$ is the goal intended to reach at time T and $x_0 > 0$ is the budget constraint. The optimal solution X^* is the optimal terminal cash flow to be achieved. The optimal portfolio will then be the one replicating X^* . Here ρ is the pricing kernel at T . Thus, $\rho \in \mathcal{F}_T$, $\rho \geq 0$, $P(\rho > 0) > 0$ and $\rho < \infty$. Typically, we set $\frac{x_0}{b} < E[\rho] < \infty$. Note that if $E[\rho] \leq \frac{x_0}{b}$, then a trivial solution $X^* \equiv \frac{x_0}{E[\rho]}$ solving (Equation 2.1) with $P[X^* \geq b] = 1$.

2.2 Modified Optimization Problem of Goal Reaching Model

It can be verified that if X is the optimal solution of (Equation 2.1), then $P((X > b) \cap (\rho > 0)) = 0$. Otherwise, one could find $\underline{b} \in (0, b)$, such that $P((\underline{b} < X \leq b) \cap (\rho > 0)) > 0$, and

$$E[\rho X] \geq E[\rho(X \mathbf{1}_{\{(X \leq \underline{b}) \cup (\rho = 0)\}} + b \mathbf{1}_{\{(\underline{b} < X) \cap (\rho > 0)\}})]$$

Existence of \underline{b} can be proved as following:

As $\underline{b} \uparrow b$, set

$$(\underline{b} < X < b) \cap (\rho > 0) \downarrow \emptyset,$$

and

$$\begin{aligned} & E[\rho X \mathbf{1}_{\{(\underline{b} < X < b) \cap (\rho > 0)\}}] - E[\rho b \mathbf{1}_{\{(\underline{b} < X < b) \cap (\rho > 0)\}}] \\ &= E[\rho(X - b) \mathbf{1}_{\{(\underline{b} < X < b) \cap (\rho > 0)\}}] \uparrow 0 \end{aligned}$$

Therefore, $\underline{b} \in (0, b)$ always exists.

Define

$$X' = X \mathbf{1}_{\{(X \leq \underline{b}) \cup (\rho = 0)\}} + b \mathbf{1}_{\{(\underline{b} < X) \cap (\rho > 0)\}}$$

Thus,

$$E[\rho X] \geq E[\rho X'], \text{ and } P[X \geq b] < P[X' \geq b],$$

which leads to a contradiction.

Also, it can be verified that if X is an optimal solution of (Equation 2.1), then

$$P((0 < X < b) \cap (\rho > 0)) = 0.$$

Otherwise, one could find $\bar{b} \in (0, b)$, such that

$$P((\bar{b} < X < b) \cap (\rho > 0)) > 0,$$

and

$$E[\rho X] \geq E[\rho(X \mathbf{1}_{\{(X=b) \cup (X=0) \cup (\rho=0)\}} + b \mathbf{1}_{\{(\bar{b} < X < b) \cap (\rho > 0)\}})]$$

We can prove the existence of \bar{b} :

As $\bar{b} \uparrow b$, set

$$(\bar{b} < X < b) \cap (\rho > 0) \downarrow \emptyset,$$

and

$$\begin{aligned} & E[\rho X \mathbf{1}_{\{(\bar{b} < X < b) \cap (\rho > 0)\}}] - E[\rho b \mathbf{1}_{\{(\bar{b} < X < b) \cap (\rho > 0)\}}] \\ = & E[\rho(X - b) \mathbf{1}_{\{(\bar{b} < X < b) \cap (\rho > 0)\}}] \uparrow 0 \end{aligned}$$

Therefore, $\bar{b} \in (0, b)$ always exists.

Define

$$X' = X\mathbf{1}_{\{(X=b) \cup (X=0) \cup (\rho=0)\}} + b\mathbf{1}_{\{(\bar{b} < X < b) \cap (\rho > 0)\}}$$

Then, $E[\rho X] \geq E[\rho X']$, and $P[X \geq b] < P[X' \geq b]$, which leads to contradiction.

The previous arguments have justified the results as follows:

Theorem 1 *There exists an optimal solution of (Equation 2.1), such that*

Case 1: If $P(\rho > 0) = 1$, then any solution X of (Equation 2.1) must satisfy $P(X \in \{0, b\}) = 1$.

Case 2: If $P(\rho = 0) > 0$, then any solution X of (Equation 2.1) must satisfy $P(X \in \{0, b\} | \rho > 0) = 1$ and $P(X \geq b | \rho = 0) = 1$.

In order to find solution of (Equation 2.1) for more general pricing kernel ρ , especially ρ with atom, we need the following assumption.

Assumption 1 *There exists a random variable $Y \in \mathcal{F}_T$, such that Y has a continuous cumulative distribution function.*

In a continuous-time financial market, \mathcal{F}_T is typically associated with Brownian motion(s)(for example, see He and Zhou (2011, section 2.1)), and Assumption 1 is automatically satisfied.

According to Theorem 1, we can restrict the optimal problem (Equation 2.1) to $\{X|P(0 \leq X \leq b) = 1\}$ only. By replacing X with $\frac{X}{b}$, we can consider a modified optimization problem

$$\begin{aligned} \max_X \quad & E[X] \\ \text{subject to} \quad & E[\rho X] = \frac{x_0}{b}, P(0 \leq X \leq 1) = 1, X \in \mathcal{F}_T \end{aligned} \quad (2.2)$$

2.3 Solution of Goal Reaching Model with Atomic Pricing Kernel

As an application of a generalized Neyman-Pearson Lemma (Leung, Song, and Yang (2013), Thm 2.10(C1)), the optimal solution of (Equation 2.2) exists and takes the form of

$$X^* = \mathbf{1}_{\{[\rho < \frac{1}{\lambda^*}] \cup ([\rho = \frac{1}{\lambda^*}] \cap [Y \leq \mu^*])\}} \quad (2.3)$$

for some $\lambda^* > 0$ and $\mu^* \in \mathbb{R}$, such that

$$\frac{x_0}{b} = E[\rho X^*] = E[\rho \mathbf{1}_{\{\rho < \frac{1}{\lambda^*}\}}] + \frac{1}{\lambda^*} P(A(\mu^*)),$$

where $A(\mu) = [\rho = \frac{1}{\lambda^*}] \cap [Y \leq \mu]$.

The existence of λ^* and μ^* can be justified as follows:

First, we find a $\lambda^* > 0$, such that

$$E[\rho \mathbf{1}_{\{\rho < \frac{1}{\lambda^*}\}}] \leq \frac{x_0}{b} \leq E[\rho \mathbf{1}_{\{\rho \leq \frac{1}{\lambda^*}\}}] \quad (2.4)$$

The existence of λ^* is guaranteed by $\rho \geq 0$ and $E[\rho] > \frac{x_0}{b}$. Given $\lambda^* > 0$ satisfying (Equation 2.4), $P(A(\mu))$ is increasing and continuous on μ , and $P(A(\infty)) \geq 0$, which leads to the existence of $\mu^* \in \mathbb{R}$. Since $E[X^*] = P[X^* = 1]$, bX^* is a solution for (Equation 2.1). Note that we don't require the assumption that ρ contains no atom as in He and Zhou (2011).

The proceeding arguments prove the following theorem.

Theorem 2 *Under Assumption 1, there exist $\lambda^* > 0$ and $\mu^* \in \mathbb{R}$, such that*

$$X^* = b\mathbf{1}_{\{[\rho < \frac{1}{\lambda^*}] \cup ([\rho = \frac{1}{\lambda^*}] \cap [Y \leq \mu^*])\}}$$

solves optimization problem (Equation 2.1), where λ^ can be determined by*

$$E[\rho\mathbf{1}_{\{\rho < \frac{1}{\lambda^*}\}}] \leq \frac{x_0}{b} \leq E[\rho\mathbf{1}_{\{\rho \leq \frac{1}{\lambda^*}\}}],$$

μ^ is determined by*

$$P([\rho = \frac{1}{\lambda^*}] \cap [Y \leq \mu^*]) = \lambda^* (\frac{x_0}{b} - E[\rho\mathbf{1}_{\{\rho < \frac{1}{\lambda^*}\}}]),$$

and $Y \in \mathcal{F}_T$ has a continuous c.d.f.. Specially, if $E[\rho\mathbf{1}_{\{\rho < \frac{1}{\lambda^}\}}] = \frac{x_0}{b}$, then $\mu^* = -\infty$ and $X^* = b\mathbf{1}_{\{\rho < \frac{1}{\lambda^*}\}}$; if $E[\rho\mathbf{1}_{\{\rho \leq \frac{1}{\lambda^*}\}}] = \frac{x_0}{b}$, then $\mu^* = \infty$ and $X^* = b\mathbf{1}_{\{\rho \leq \frac{1}{\lambda^*}\}}$.*

Corollary 1 *If ρ contains no atom, that is, ρ has a continuous c.d.f., then X^* takes the form of $b\mathbf{1}_{\{\rho \leq \frac{1}{\lambda^*}\}}$.*

Note that Corollary 1 was first derived by He and Zhou (2011, Thm 3.2).

2.4 Example: Pricing Kernel Follows Geometric Brownian Motion

Example 1 *A continuous-time market consists of two assets, a bond and a stock. Let S_0 be the price process of the bond and S_1 be the process of another asset, they follow the stochastic differential equations:*

$$dS_0(t) = r(t)S_0(t)dt, t \in [0, T]; S_0(0) = s_0 > 0$$

$$dS_1(t) = S_1(t)[b_1(t)dt + \sigma(t)dW(t)], t \in [0, T]; S_1(0) = s_1 > 0$$

where $W(t)$, $t \in [0, T]$ is a Brownian motion, the interest rate $r(\cdot)$, appreciation rate $b_1(\cdot)$ and volatility rate $\sigma(\cdot)$ are all deterministic functions. We assume $\sigma(t) > 0$, since S_1 is a risky asset. Let θ_0 be the Sharpe ratio:

$$\theta_0(t) = \frac{b_1(t) - r(t)}{\sigma(t)} \tag{2.5}$$

Note that this market frame can be generalized to a portfolio with m assets and $W(t)$ as an n -dimensional Brownian motion.

Define the pricing kernel at t :

$$\begin{aligned}\rho(t) &:= \exp\left\{-\int_0^t [r(s) + \frac{1}{2}\theta_0(s)^2]ds + \int_0^t \theta_0(s)dW(s)\right\} \\ \rho &:= \rho(T)\end{aligned}$$

2.4.1 Explicit Solution with Positive Cumulative Sharpe Ratio

If $\int_0^T \theta_0^2(s)ds > 0$, then it has a positive cumulative Sharpe ratio.

Then ρ follows a lognormal distribution which contains no atom. Based on Corollary 1, the solution to Equation 2.1 takes the form of

$$X^* = b\mathbf{1}_{\{\rho \leq \frac{1}{\lambda^*}\}}$$

where λ^* is obtained by solving $E[\rho\mathbf{1}_{\{\rho \leq \frac{1}{\lambda^*}\}}] = \frac{x_0}{b}$

$$\lambda^* = \exp\left\{-\Phi^{-1}\left(\frac{x_0}{b}e^{\int_0^T r(s)ds}\right)\left(\int_0^T \theta_0^2(s)ds\right)^{\frac{1}{2}} + \int_0^T r(s)ds - \frac{1}{2}\int_0^T \theta_0^2(s)ds\right\}$$

Therefore, the optimal value of (Equation 2.1) is

$$P[X^* \geq b] = P\left[\rho \leq \frac{1}{\lambda^*}\right] = \Phi\left[\Phi^{-1}\left(\frac{x_0}{b}e^{\int_0^T r(s)ds}\right) + \left(\int_0^T \theta_0^2(s)ds\right)^{\frac{1}{2}}\right]$$

2.4.2 Explicit Solution with Zero Cumulative Sharpe Ratio

If $\int_0^T \theta_0^2(s) ds = 0$, then it has a zero cumulative Sharpe ratio.

For example, $b_1(t) \equiv r(t)$ which leads to $\theta_0(t) \equiv 0$ and $P[\rho = e^{-\int_0^T r(s) ds}] = 1$. Let $\lambda^* = e^{\int_0^T r(s) ds}$, then $P[\rho \leq \frac{1}{\lambda^*}] = 1$. Let $Y = S_1(T)$ as required in Theorem 2. One can find μ^* by solving $P[S_1(T) \leq \mu^*] = \frac{x_0}{b} e^{\int_0^T r(s) ds}$. An optimal solution solving (Equation 2.1) with $\mathcal{F}_T = \sigma\{W(t) : 0 \leq t \leq T\}$ is $X^* = b \mathbf{1}_{\{S_1(T) \leq \mu^*\}}$. Since

$$S_1(T) = s_1 e^{\int_0^T [b_1(s) - \frac{1}{2}\sigma^2(s)] ds + \int_0^T \sigma(s) dW(s)}$$

follows a lognormal distribution, then

$$\mu^* = s_1 e^{\Phi^{-1}\left(\frac{x_0}{b} e^{\int_0^T r(s) ds}\right) \left(\int_0^T \sigma^2(s) ds\right)^{\frac{1}{2}} + \int_0^T [b_1(s) - \frac{1}{2}\sigma^2(s)] ds}.$$

Therefore, the optimal value of (Equation 2.1) in this case is

$$\lambda^* \frac{x_0}{b} = \frac{x_0}{b} e^{\int_0^T r(s) ds}.$$

2.4.3 Back Tests and Numerical Results

For pricing kernel follows a Geometric Brownian Motion as discussed above, we could calculate the explicit solution and compare it with the numerical solution based on grid search. Let $x_0 = 1$, $b = 1.05$, $T = 1$, $r = 0.03$ and $\theta = 0.1$. Denote λ_S^* , V_S and V_E to be the optimal value of λ and optimal outperforming probability from grid search, and optimal outperforming probability from explicit solution respectively.

TABLE I
COMPARISON BETWEEN THE GRID SEARCH SOLUTION AND EXPLICIT
SOLUTION

Number of Grids	λ_S^*	V_S	$V_E - V_S$
100	0.8422365	0.98236	0.00500
200	0.8369955	0.98520	0.00216
400	0.8366255	0.98542	0.00194
800	0.8364413	0.98548	0.00188
1600	0.8363493	0.98550	0.00186
3200	0.8363034	0.98554	0.00182
6400	0.8362804	0.98556	0.00180
12800	0.8362689	0.98558	0.00178

Table 1 shows the comparison between the solution based on grid search and explicit solution. From the results, we can see that as the number of grids increases, the probability of outperformance increases, and converges to the explicit solution from our theoretical result.

The relationship between the number of simulation step and the optimal value from the simulation, and the convergence of two optimal values are shown in Figure 1.

The upper figure shows as the number of grid steps increases, the probability of outperformance strictly increases and the lower figure shows that as the number of grid steps increases, the difference between the outperformance probability from the grid search and that from our explicit solution decreases and converges to 0.

Figure 1. Convergence of Two Optimal Values

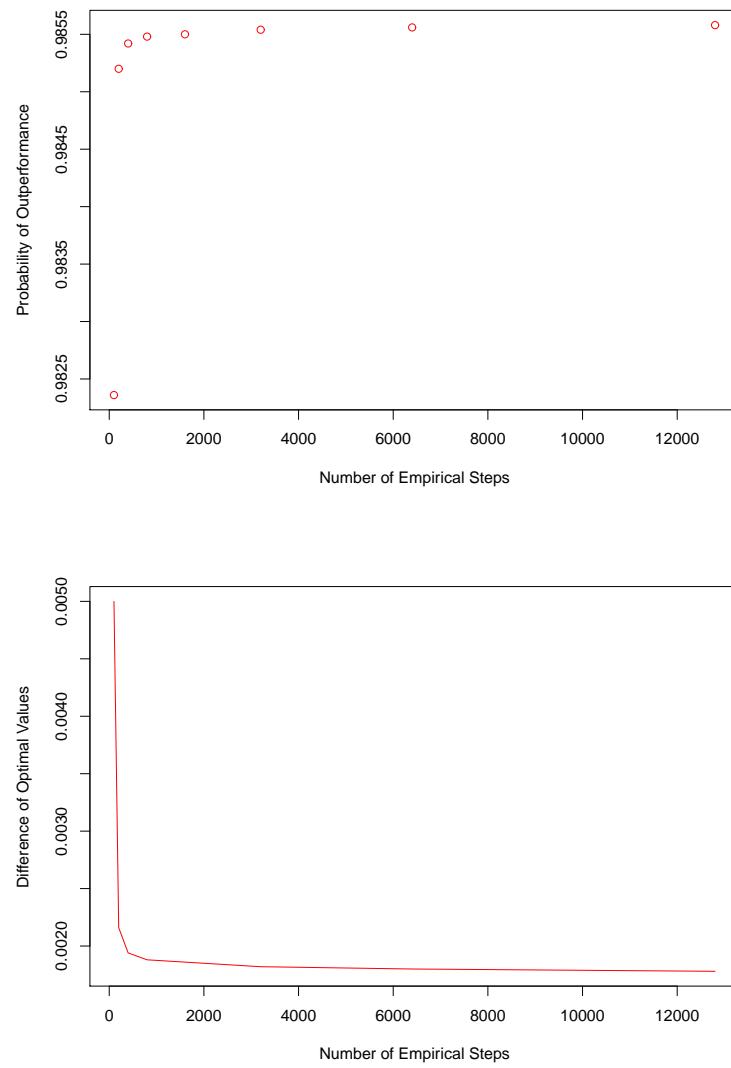
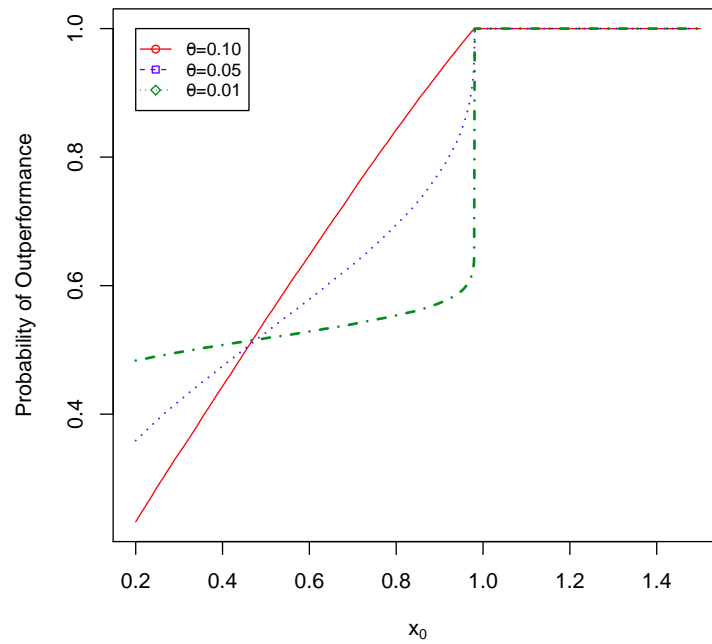


Figure 2. Outperformance Probability versus Initial Capital with Different Sharpe Ratios

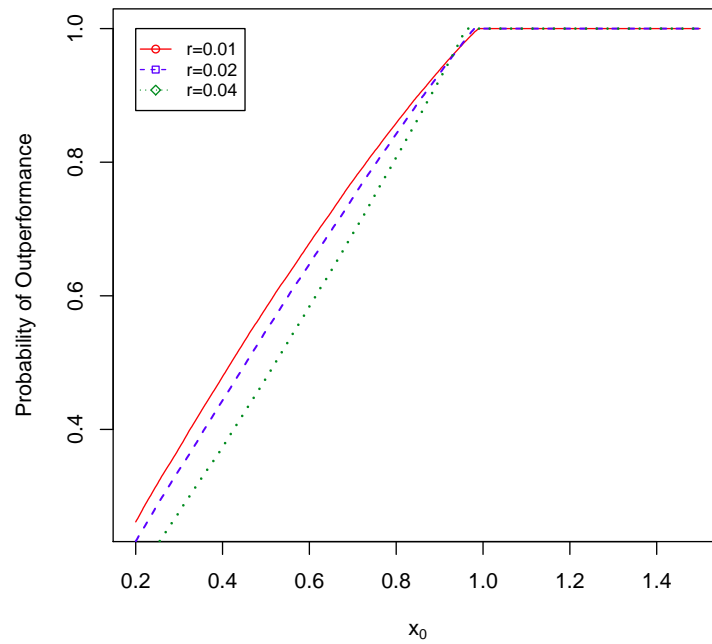


The Outperformance probability versus initial capital x_0 with different Sharpe ratios: $\theta(t) = 0.1$, $\theta(t) = 0.05$, and $\theta(t) = 0.01$ are shown in Figure 2.

It shows that as the initial capital increases, the probability of outperformance increases to 1. When the initial capital is large enough, value of portfolios could exceed the benchmark for sure. For portfolio with higher Sharpe Ratio, it is more volatile, and has higher chance to

exceed the benchmark.

Figure 3. Outperformance Probability versus Initial Capital with Different Interest Rates



The Outperformance probability versus initial capital x_0 with different interest rates: $r(t) = 0.01$, $r(t) = 0.02$, and $r(t) = 0.04$ are shown in Figure 3.

For higher interest rate, it leads to a lower Sharpe Ratio, and it has lower chance to exceed the benchmark.

CHAPTER 3

YAARI'S DUAL MODEL

In this chapter, we discuss the optimal solution to Yaari's Model. The pricing kernel here could contain atoms and thus is more general than non-atomic ones. Under the assumption that probability distortion/weighting is differentiable, we derive a modified optimization problem that contains left-continuous quantile function of the pricing kernel and terminal case flow. We propose an algorithm, called Search-and-Cut Algorithm to find the optimal solution, which is good for cases where the weighting/pricing-kernel ratio consists of a finite number of monotone pieces. We also prove the existence and uniqueness of the optimal solution as well. Finally, we derive an optimal solution of Yaari's dual model for more general pricing kernels and probability distortions.

There are two main differences between Yaari's dual model and goal reaching model in last chapter. One is Yaari's dual model includes a probability distortion w and the other one is it does not have a fixed benchmark b .

3.1 Introduction and Model Setup

Following He and Zhou (2011), we consider Yaari's Dual problem as follows

$$\begin{aligned} \max_X \quad & \psi(X) = \int_0^\infty w(P(X > x))dx \\ \text{subject to} \quad & E[\rho X] = x_0, X \geq 0, X \in \mathcal{F}_T \end{aligned} \tag{3.1}$$

where $w : [0, 1] \rightarrow [0, 1]$ is a probability distortion or weighting function, which may represent a subjective inflation or deflation of the true probability. Typically, we assume $w(0) = 0$, $w(1) = 1$ and w is nondecreasing on $[0, 1]$ and differentiable on $(0, 1)$. In this problem, $\rho \in \mathcal{F}_T$ is a nonnegative random variable, such that $0 < \rho < \infty$ and $0 < E(\rho) < \infty$.

Consider a special case $(\Omega, \mathcal{F}_T, \mathbf{P}) = ([0, 1], \mathcal{B}[0, 1], \lambda)$, where $\mathcal{B}[0, 1]$ is the Borel σ -field on $[0, 1]$, λ is Lebesgue measure. In this probability space, a random variable is a measurable function on $[0, 1]$. Note that if the original probability space is not $([0, 1], \mathcal{B}[0, 1], \lambda)$, one can always replace ρ with

$$F_\rho^{\leftarrow}(x) := \inf\{y : F_\rho(y) \geq x\}, \quad x \in (0, 1)$$

where $F_\rho(\cdot)$ is the cumulative distribution function of ρ . Then ρ and F_ρ^{\leftarrow} share the same distribution function. Therefore, we may assume ρ is nondecreasing as a function on $[0, 1]$.

Therefore, we first consider the modified optimization problem as follows:

$$\begin{aligned} \max_X \quad & \psi(X) = \int_0^\infty w(\lambda(X > x))dx \\ \text{subject to} \quad & E[\rho X] = \int_0^1 \rho(t)X(t)dt = x_0 \end{aligned} \quad (3.2)$$

where both ρ and X are nonnegative measurable functions on $[0, 1]$, and ρ is nondecreasing on $[0, 1]$, and positive on $(0, 1)$.

Note that the domain of F_ρ^{\leftarrow} could be extended to $[0, 1]$ with $F_\rho^{\leftarrow}(0) = -\infty$ and $\inf \emptyset = +\infty$.

3.2 Properties of an Optimal Solution

Theorem 3 *Assume that ρ is a nonnegative nondecreasing function on $[0, 1]$, and the optimization (Equation 3.2) yields a solution X , then X is nonincreasing.*

3.2.1 Properties of Optimal Solution on Non-atomic ρ

If ρ contains no atom on $[0, 1]$, one can derive the conclusion of Theorem 3 based on Lemma 2.5 in He and Zhou (2011) or Theorem 1 in Jin and Zhou (2010)(or Theorem B.1 in Jin and Zhou (2008)). If X is a solution of (Equation 3.2), one can find a probability distribution function G , such that $P(X \leq x) = G(x)$.

Define

$$X_2^* = G^{-1}(1 - F_\rho(\rho)),$$

where $G^{-1}(x) := \inf\{y : G(y) \geq x\}$, $x \in [0, 1]$. Let

$$Z := 1 - F(\rho),$$

which follows uniform distribution on $[0, 1]$, since ρ contains no atom.

Note that

$$P(X_2^* \leq x) = P(G^{-1}(1 - F_\rho(\rho)) \leq x) = P(Z \leq G(x)) = G(x),$$

that is X and X_2^* have the same distribution, thus $\psi(X) = \psi(X_2^*)$. Since $G^{-1}(1 - F_\rho(\rho))$ is nonincreasing, following Lemma 1 in Jin and Zhou (2010),

$$E[X_2^*\rho] \leq E[X\rho] = x_0,$$

for any $P(X \leq x) = G(x)$, where the inequality becomes equality if and only if $X = X_2^*$ a.s. on $[0, 1]$. If $E[X_2^*\rho] < E[X\rho] = x_0$, then solution of (Equation 3.2) can be improved to

$$X^* = X \cdot \frac{E[X\rho]}{E[X_2^*\rho]},$$

such that $E[X^*\rho] = x_0$, and $\psi(X^*) > \psi(X)$. In this case, X is not a solution, which leads to a contradiction. Therefore, for ρ containing no atom on $[0, 1]$, an optimal solution of (Equation 3.2) must be nonincreasing a.s. on $[0, 1]$.

3.2.2 Properties of Optimal Solution on Atomic ρ

Now we assume ρ contains atom(s), that is, ρ is not strictly increasing. Since ρ is non-decreasing on $[0, 1]$, the number of its atoms are at most countably infinite, and those atoms occupy non-overlapped intervals on $[0, 1]$. Thus, we may denote the countable atomistic intervals by A_1, A_2, A_3, \dots , such that $\rho(A_i) = a_i > 0$, *i.e.*, $\rho(x) = a_i, \forall x \in A_i$. Note that the number of A_i could be finite or countably infinite. Even that A_i 's are non-overlapped, they may not be able to be arranged in increasing order according to their left-end (also the number of A_i may

be countable infinite, see Example 2). Define $C = (\cup A_i)^c$, on which ρ contains no atom. Then X can be rewritten as $X = X\mathbf{1}_C + X\mathbf{1}_{A_1} + X\mathbf{1}_{A_2} + \dots$.

To prove Theorem 3, we need five lemmas as follows. As a direct corollary of Theorem 1 in Jin and Zhou (2010), we have:

Lemma 1 *Any optimal solution X to (Equation 3.2) is nonincreasing a.s. on C .*

Lemma 2 *Suppose X is a nonnegative, integrable function on $[a, b]$. If for any t_1, t_2, h , such that $a \leq t_1 < t_1 + h \leq t_2 < t_2 + h \leq b$, we always have $\int_{t_1}^{t_1+h} X(s)ds \geq \int_{t_2}^{t_2+h} X(s)ds$, then X is nonincreasing a.s. on $[a, b]$.*

Proof of Lemma 1:

Let $F(t) = \int_0^t X(s)ds$, $0 \leq t \leq 1$. According to Theorem 31.3 Billingsley (1995), there exist an $A \subseteq [a, b]$, such that $\lambda(A^c) = 0$ and $F'(t) = X(t)$, $\forall t \in A$.

For any $t_1, t_2 \in A$, such that $t_1 < t_2$,

$$\begin{aligned} X(t_1) &= F'(t_1) = \lim_{h \rightarrow 0} \frac{F(t_1 + h) - F(t_1)}{h} = \lim_{h \rightarrow 0} \frac{1}{h} \int_{t_1}^{t_1+h} X(s)ds \\ &\geq \lim_{h \rightarrow 0} \frac{1}{h} \int_{t_2}^{t_2+h} X(s)ds = \lim_{h \rightarrow 0} \frac{F(t_2 + h) - F(t_2)}{h} = F'(t_2) \\ &= X(t_2) \end{aligned}$$

Therefore, X is nonincreasing a.s. on $[a, b]$. \square

Lemma 3 *If X is a solution of (Equation 3.2), one can always find an X^* , which shares the same distribution with X and is nonincreasing a.s. on each A_i .*

Proof of Lemma 3:

Fix any A_i , let a, b be the ending points of A_i . Then $X\mathbf{1}_{[a,b]}$ can be regarded as a random variable on the probability space $([a, b], \mathcal{B}[a, b], \frac{\lambda}{b-a})$, which has the distribution function:

$$F_i(x) = P(X\mathbf{1}_{[a,b]} \leq x) = \frac{1}{b-a} \lambda((X \leq x) \cap [a, b]), \quad -\infty < x < \infty$$

Define

$$X^*(t) = \begin{cases} F_i^{\leftarrow}\left(\frac{b-t}{b-a}\right), & t \in (a, b) \\ X(t), & \text{otherwise} \end{cases}$$

where $F_i^{\leftarrow}(t) := \inf\{y : F_i(y) \geq t\}$, $0 < t < 1$, is nondecreasing. Let

$$U(t) = \frac{b-t}{b-a}, \quad a \leq t \leq b.$$

Then $U \sim Unif(0, 1)$ as a random variable on $([a, b], \mathcal{B}[a, b], \frac{\lambda}{b-a})$. Actually,

$$\begin{aligned} P(U \leq x) &= \frac{1}{b-a} \lambda(\{a \leq t \leq b : \frac{b-t}{b-a} \leq x\}) = \frac{1}{b-a} \lambda(\{a \leq t \leq b : t \geq b - x(b-a)\}) \\ &= \frac{1}{b-a} (b - (b - x(b-a))) = x, \quad 0 < x < 1 \end{aligned}$$

Therefore $F_i^{\leftarrow}(U)$ and $X\mathbf{1}_{[a,b]}$ shares the same distribution on (a, b) . Thus, X and X^* share the same distribution on $[0, 1]$. Then we get

$$E[\rho X^*] - E[\rho X] = a_i E[F_i^{\leftarrow}(U) - X\mathbf{1}_{[a,b]}] = 0.$$

Having $\psi(X^*) = \psi(X)$ and $E[\rho X^*] = x_0$, X^* is also an optimal solution of (Equation 3.2). By the way, one could always find a solution which is nonincreasing on each A_i . \square

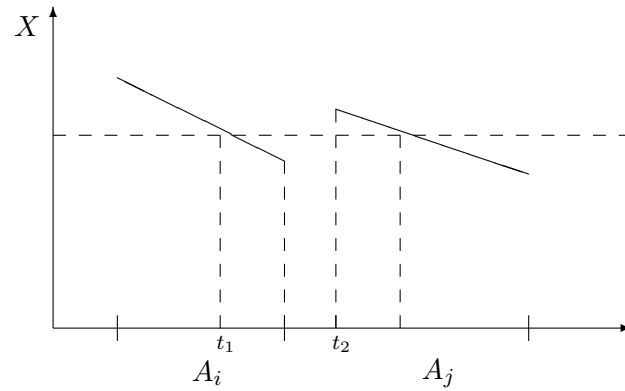
In order to make the later proof simple, on each A_i , we use nonincreasing X^* to replace X .

Lemma 4 *Suppose X is a solution to (Equation 3.2), which is nonincreasing a.s. on each A_i .*

Then X is nonincreasing a.s. on $\cup A_i$.

Proof of Lemma 4:

Given any two atom sets A_i, A_j , we assume $\rho(A_i) = a_i < a_j = \rho(A_j)$ without any loss of generality. Then the interval A_i must locate left to A_j since ρ is nondecreasing. We first show that X is nonincreasing on $A_i \cup A_j$. Otherwise, according to Lemma 2 & 3, \exists non-overlapped, equal length ($h > 0$) intervals I_1, I_2 on $A_i \cup A_j$ s.t. $\int_{I_1} X(s)ds < \int_{I_2} X(s)ds$. From conclusion

Figure 4. Graph of X Non-increasing on $A_i \cup A_j$ 

of Lemma 3, as shown in Figure 4, neither $I_1 \cup I_2 \subset A_i$ nor $I_1 \cup I_2 \subset A_j$ is possible. It can be verified that there always

$$\exists t_1, t_2, h > 0 \text{ s.t. } (t_1, t_1 + h) \subset A_i, (t_2, t_2 + h) \subset A_j,$$

and

$$\int_{t_1}^{t_1+h} X(s)ds < \int_{t_2}^{t_2+h} X(s)ds.$$

Interchanging the values of X on $(t_1, t_1 + h)$ and $(t_2, t_2 + h)$. We define

$$X^*(s) = \begin{cases} X(s + t_2 - t_1), & s \in (t_1, t_1 + h) \\ X(s - t_2 + t_1), & s \in (t_2, t_2 + h) \\ X(s), & \text{otherwise} \end{cases}$$

Then we get

$$\begin{aligned} & E[\rho X^*] - E[\rho X] \\ = & \int_{t_1}^{t_1+h} \rho(s) X^*(s) ds + \int_{t_2}^{t_2+h} \rho(s) X^*(s) ds \\ & - \int_{t_1}^{t_1+h} \rho(s) X(s) ds - \int_{t_2}^{t_2+h} \rho(s) X(s) ds \\ = & a_i \int_{t_1}^{t_1+h} (X^*(s) - X(s)) ds - a_j \int_{t_2}^{t_2+h} (X(s) - X^*(s)) ds \end{aligned}$$

According to the definition of X^* ,

$$\int_{t_1}^{t_1+h} (X^*(s) - X(s)) ds = \int_{t_2}^{t_2+h} (X(s) - X^*(s)) ds > 0$$

Since $a_i < a_j$, then

$$E[\rho X^*] - E[\rho X] = (a_i - a_j) \int_{t_1}^{t_1+h} (X^*(s) - X(s)) ds < 0.$$

Therefore, X can be strictly improved into $X^* \cdot \frac{E[\rho X]}{E[\rho X^]}$. That is, X is not an optimal solution, which leads to a contradiction.

Since X is nonincreasing a.s. on $A_i \cup A_j$, $\forall i \neq j$, and ρ has countable atoms, then X is nonin-

creasing on $\cup A_i$. □

Lemma 5 *Suppose X is an optimal solution to (Equation 3.2), which is nonincreasing a.s. on $\cup_i A_i$. Then X is nonincreasing a.s. on $[0, 1]$.*

Proof of Lemma 5:

Let's fix an arbitrary A_i . There is no harm to assume that $A_i = [t_2, t_2 + h_2]$ with $h_2 > 0$ is a closed interval since the values of ρ at the ending points $\{t_2, t_2 + h_2\}$ won't ruin the conclusion anyway. Recall $C = (\cup A_i)^c$. Let $C_l = [0, t_2) \cap C$ be a subset of C , which lies left to A_i , whereas $C_r = (t_2 + h_2, 1] \cap C$ be another subset of C , which lies right to A_i , as shown in Figure 5.

Clearly, $C_l \cup C_r = C$ and $C_l \cap C_r = \emptyset$. Note that C_l and C_r may contain no interval, even they both have positive Lebesgue measure.

First, we show that X is nonincreasing a.s. on $C_l \cup A_i$.

A special case is $\lambda(C_l) = 0$. Apparently, in this case, X is nonincreasing on $C_l \cup A_i$ a.s. . Now suppose $\lambda(C_l) > 0$, then $t_2 > 0$.

There exists two cases. Case I: $essinf_{t \in C_l} X(t) \geq esssup_{t \in A_i} X(t)$ (as shown in Figure 6), then X is nonincreasing a.s. on $C_l \cup A_i$, where $essinf$ ($esssup$) is the infimum (supremum) for all elements in a set, but rather almost everywhere, that is, except on a set of measure zero.

Let $g(x) = \lambda([t_2 - x, t_2] \cap C)$, $x \in [0, t_2]$, which is continuous, then $g(0) = 0$, $g(t_2) = \lambda(C_l)$.

Figure 5. Graph of C_l , A_i , and C_r

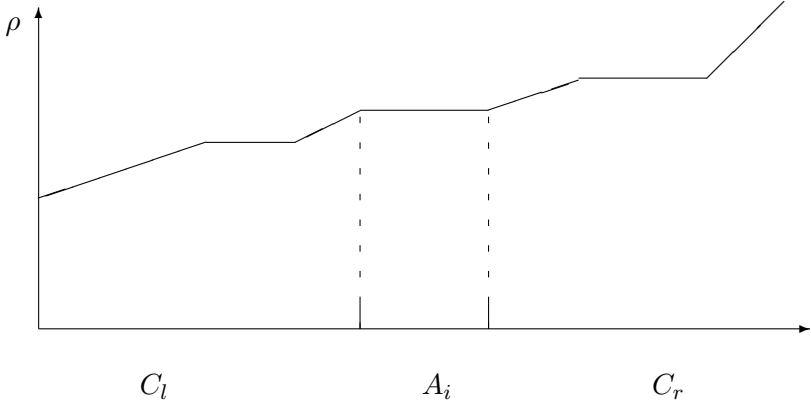


Figure 6. Case I of Lemma 5

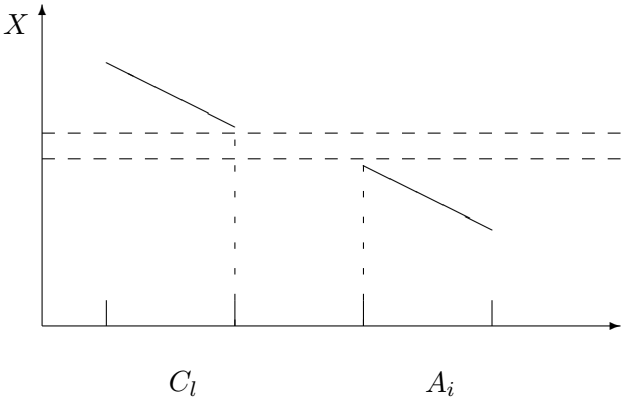
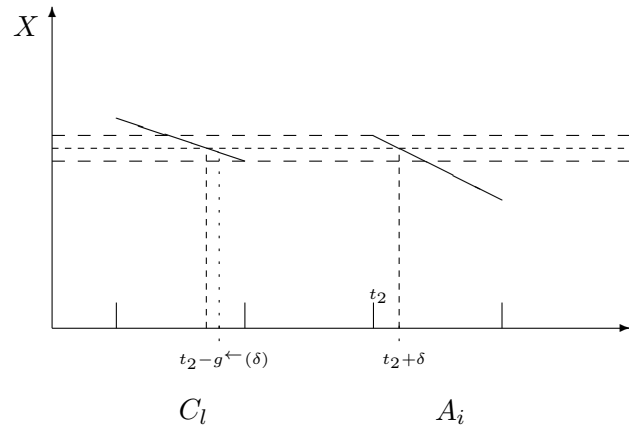


Figure 7. Case II of Lemma 5



Case II: $\text{essinf}X(t)_{t \in C_l} < \text{esssup}X(t)_{t \in A_i}$, (as shown in Figure 7), then $\exists 0 < \delta < h$, such that $X(t_2 - g^{\leftarrow}(\delta)) < X(t_2 + \delta)$, where $g^{\leftarrow}(\delta) := \inf\{x : g(x) = \delta\}$. Let $J_1 = (t_2 - g^{\leftarrow}(\delta), t_2)$ and $J_2 = (t_2, t_2 + \delta)$. Note that $\lambda(J_1 \cap C) = \lambda(J_2) = \delta$, although the lengths of J_1 and J_2 may not be the same.

By interchanging the values of X on $J_1 \cap C$ and J_2 , we get

$$X^*(t) = \begin{cases} X(t_2 + \delta - g(t_2 - t)), & t \in J_1 \cap C \\ X(t_2 - g^{\leftarrow}(t_2 + \delta - t)), & t \in J_2 \\ X(t), & \text{otherwise} \end{cases}$$

Similarly, one can derive

$$\begin{aligned} & E[\rho X^*] - E[\rho X] \\ = & \int_{J_1 \cap C} \rho(s) X^*(s) ds + \int_{J_2} \rho(s) X^*(s) ds \\ & - \int_{J_1 \cap C} \rho(s) X(s) ds - \int_{J_2} \rho(s) X(s) ds \\ = & \int_0^\delta \rho(t_2 - g^{\leftarrow}(\delta - y)) X(t_2 + y) dy \\ & - \int_0^\delta \rho(t_2 - g^{\leftarrow}(\delta - y)) X(t_2 - g^{\leftarrow}(\delta - y)) dy \\ & + \int_0^\delta \rho(t_2 + y) X(t_2 - g^{\leftarrow}(\delta - y)) dy \\ & - \int_0^\delta \rho(t_2 + y) X(t_2 + y) dy \\ = & \int_0^\delta \rho(t_2 - g^{\leftarrow}(\delta - y)) [X(t_2 + y) - X(t_2 - g^{\leftarrow}(\delta - y))] dy \\ & - \int_0^\delta \rho(t_2 + y) [X(t_2 + y) - X(t_2 - g^{\leftarrow}(\delta - y))] dy \\ = & \int_0^\delta [\rho(t_2 - g^{\leftarrow}(\delta - y)) - \rho(t_2 + y)] [X(t_2 + y) - X(t_2 - g^{\leftarrow}(\delta - y))] dy \\ < & 0 \end{aligned}$$

Note that $g^{\leftarrow}(\delta) = \delta$ on J_2 . Then X can be strictly improved to $X \cdot \frac{E[\rho X]}{E[\rho X^*]}$, which leads to a contradiction.

Therefore, X is nonincreasing a.s. on $C_l \cup A_i$. Following the same procedure, one can verify that X is nonincreasing a.s. on $C_r \cup A_i$. Therefore X is nonincreasing on a.s. $C \cup A_j$, for any A_j . Since ρ has countable atoms, X is nonincreasing a.s. on $C \cup (\cup_i A_i)$. \square

Combining Lemma 1 \sim Lemma 5, we conclude that if ρ is a nonnegative nondecreasing function, then yields an optimal solution X , which is nonincreasing on $[0, 1]$. It finishes the proof for Theorem 3.

We can also get the following conclusion:

Assume that ρ is a nonnegative nondecreasing function on $[0, 1]$, then for any $X \in \mathfrak{X}$, there exists an $X' \in \mathfrak{X}_d := \{X \in \mathfrak{X} \mid X \text{ is nonincreasing}\}$. In other words, the optimization problem (Equation 3.2) is equivalent to

$$\max_{X \in \mathfrak{X}_d} \psi(X) = \int_0^\infty w(\lambda(X > x)) dx \quad (3.3)$$

3.2.3 Example of Extreme Case: Cantor Function ρ

Example 2 *Cantor function $\rho(\omega) : [0, 1] \rightarrow [0, 1]$. Let $[a_1^{(1)}, b_1^{(1)}]$ be the middle interval of $[0, 1]$ with length $\varepsilon < \frac{1}{3}$. Let $[a_1^{(2)}, b_1^{(2)}], [a_2^{(2)}, b_2^{(2)}]$ be the middle intervals of remaining after deleting $[a_1^{(1)}, b_1^{(1)}]$ from $[0, 1]$ with length ε^2 . Continue in this fashion so that at the n^{th} stage we have*

the intervals $[a_1^{(n)}, b_1^{(n)}], \dots, [a_k^{(n)}, b_k^{(n)}], \dots, [a_{2^n-1}^{(n)}, b_{2^n-1}^{(n)}]$. Then the complement of the union of all these intervals $[a_k^{(n)}, b_k^{(n)}]$ is the Cantor set without the endpoints. Define

$$\rho(\omega) = \frac{2k-1}{2^n}, \omega \in [a_k^{(n)}, b_k^{(n)}]$$

Thus, ρ is atomic on those intervals, and is non-atomic on the Cantor set, see Figure 8. In this case, C is the Cantor set, which has positive Lebesgue measure $\frac{1-3\varepsilon}{1-2\varepsilon}$, but contain no intervals.

3.3 Solution of Modified Optimization Problem

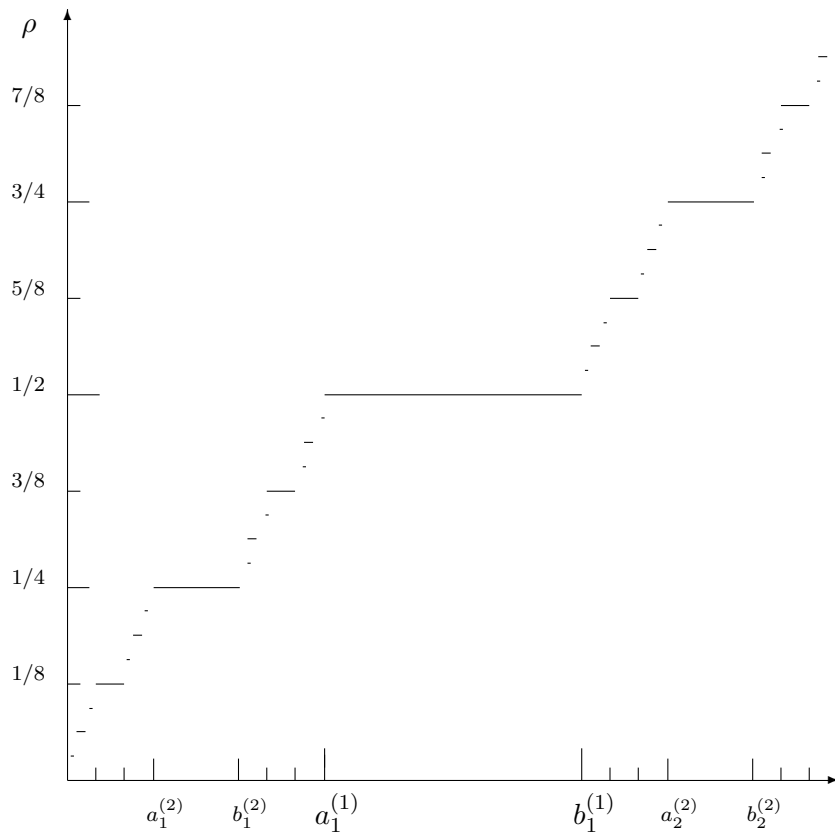
3.3.1 Form of the Optimal Solution to a Modified Optimization Problem

In this section, we will derive the form of the solution for the modified optimization problem, with aid of the following theorem and lemmas.

The preference measure ψ in Yaari's model, via Fubini's theorem, can be rewritten in the following form (He and Zhou, 2011, equation 2.12).

$$\psi(X) = \int_0^\infty w(P(X > x))dx = \int_0^\infty xd[-w(1 - F_X(x))] \quad (3.4)$$

Figure 8. Cantor Function ρ



where $w : [0, 1] \rightarrow [0, 1]$, is a nondecreasing distortion on the c.d.f. of X .

Proof of Equation 3.4:

$$\begin{aligned}
& \int_0^\infty x d[-w(1 - F_X(x))] \\
&= \int_0^\infty \left(\int_0^x 1 ds \right) d[-w(1 - F_X(x))] \text{ since } 1 \geq 0, \text{ by Fubini's theorem} \\
&= \int_0^\infty \left(\int_s^\infty 1 d[-w(1 - F_X(x))] \right) ds \\
&= \int_0^\infty (0 + w(1 - F_X(s))) ds \\
&= \int_0^\infty w(1 - F_X(s)) ds \\
&= \int_0^\infty w(P(X > x)) dx
\end{aligned}$$

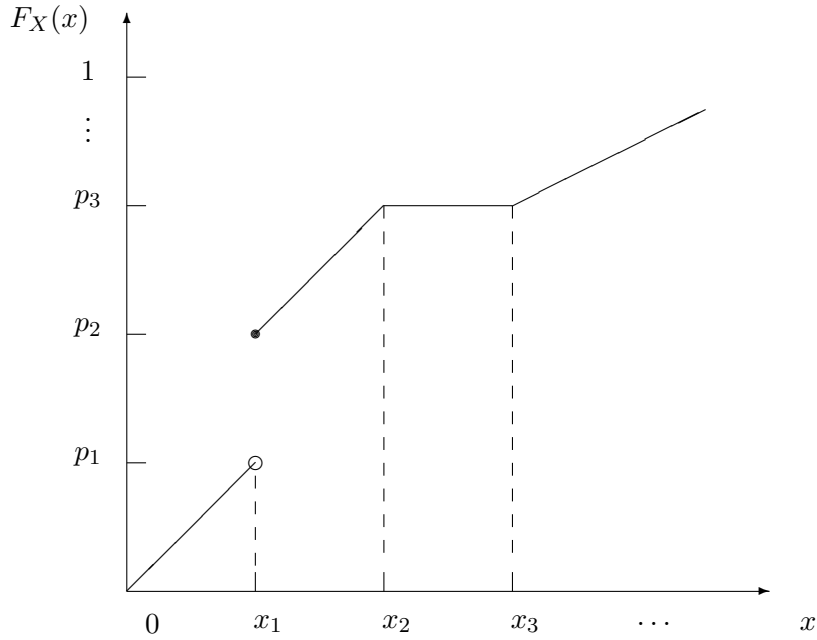
□

Preference measure $\psi(X)$ can be considered as the expected cash flow, with distorted probability distribution.

3.3.2 Discussion on Quantile Functions

If F_X is continuous and non-atomic (i.e. strictly increasing), let $z = F_X(x)$, $z \in [0, 1]$, then the random variable $Z = F_X(x)$ follows the uniform distribution on $[0, 1]$. For more general cases, F_X may have atoms and/or jumps, as illustrated in Figure 9.

We define the (left-continuous) quantile function

Figure 9. Graph of F_X 

$$F_X^{\leftarrow}(z) := \inf\{y : F_X(y) \geq z\}, z \in (0, 1).$$

Graph of $F_X^{\leftarrow}(z)$ is shown in Figure 10.

Meanwhile, define right-continuous quantile function

$$F_X^{\rightarrow}(z) := \sup\{y : F_X(y) \leq z\}, z \in (0, 1).$$

or equivalently

$$F_X^{\rightarrow}(z) := \inf\{y : F_X(y) > z\}, z \in (0, 1).$$

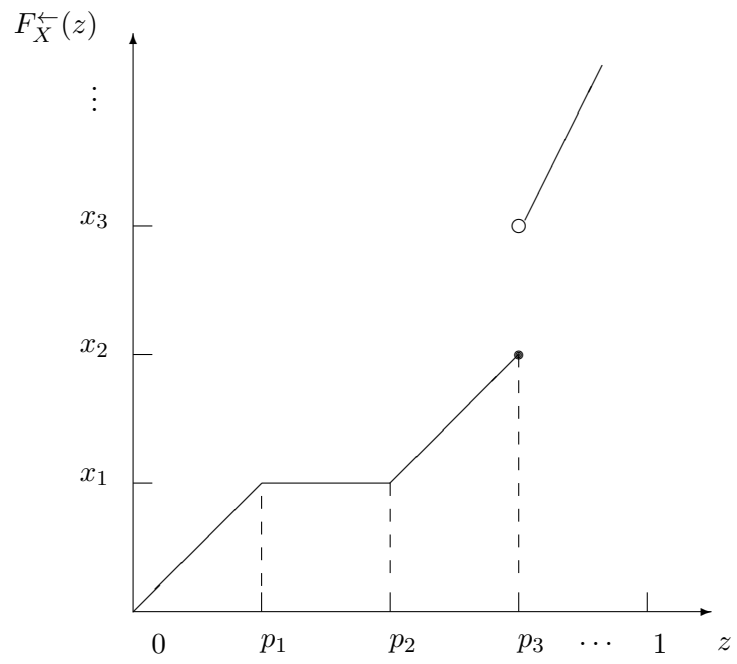
Figure 10. Graph of Left-continuous Quantile Function $F_X^{\leftarrow}(z)$ 

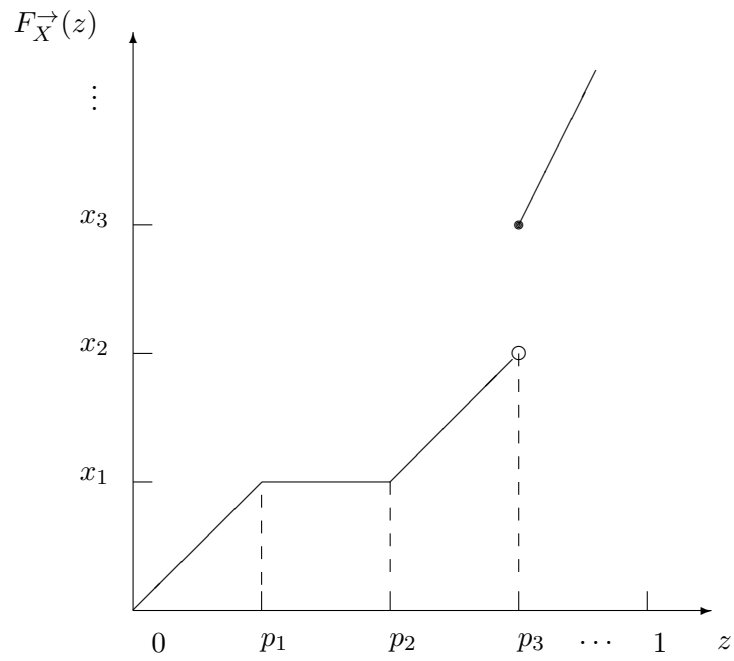
Figure 11. Graph of Right-continuous Quantile Function $F_X^{\rightarrow}(z)$ 

Figure 11 shows the graph of $F_X^{\rightarrow}(z)$.

As we shall see, there are some special points and intervals on the domain of F_X .

We define

$$V_X = \{x : F_X(x-) \neq F_X(x)\}$$

That is, V_X is the set of discontinuous points of F_X . In Figure 9 above, $\{x_1\} \in V_X$.

Let

$$H = \{z \in [0, 1] : F_X^{\leftarrow}(z) \neq F_X^{\rightarrow}(z)\}$$

which is the set of all levels at which the graph of $F_X(x)$ has a horizontal portion. In other words, H is the set of atom points of F_X . In Figure 9 above, $\{p_3\} \in H$. Since H is a set of points projected from a set of disjoint intervals on domain of F_X , i.e. $[0, \infty)$, then H is countable.

We define

$$H_X = \{x : F_X(x) = z, z \in H\} \setminus V_X$$

$$N_X = [0, \infty) \setminus (H_X \cup V_X)$$

By this way, H_X , V_X and N_X form a partition of $[0, \infty)$.

Let

$$V = \cup_{x \in V_X} [F_X(x-), F_X(x)) \setminus H$$

It is the union of all intervals of z , on which the graph of $F_X(x)$ has a jump. In Figure 9, $[p_1, p_2) \subset V$. Note that V_X is countable, and the intervals in V are disjoint.

Let

$$N = [0, 1] \setminus (H \cup V)$$

Then $N \cup H \cup V = [0, 1]$ and $N \cap H = H \cap V = N \cap V = \emptyset$.

Example 3 For the $F_X(x)$ displayed in Figure 9, $V_X = \{x_1\}$, $H = \{p_3\}$, $H_X = \{[x_2, x_3]\}$, $N_X = \{[0, x_1), (x_1, x_2), (x_3, \infty)\}$, $V = \{[p_1, p_2)\}$, and $N = \{[0, p_1), [p_2, p_3), (p_3, 1]\}$.

Lemma 6 Suppose w is left-continuous on $[0, 1]$, then

$$\int_0^\infty x d[-w(1 - F_X(x))] = \int_0^1 F_X^{\leftarrow}(z) d[-w(1 - z)]$$

Proof of Lemma 6:

Since $F_X(x)$ is continuous and strictly increasing on N_X , according to the substitution rule for Lebesgue-Stieltjes integral (see, for example Falkner and Teschl (2012)),

$$\int_{N_X} x d[-w(1 - F_X(x))] = \int_N F_X^{\leftarrow}(z) d[-w(1 - z)].$$

Following Theorem 3 in Falkner and Teschl (2012), we define $W_V(z) = -w_V(1 - z)$ be the increasing function that is obtained from $W(z) = -w(1 - z)$ by removing the jumps, denoted by V_W , that W has at the points of V , whereas $W_H(z) = -w_H(1 - z)$ be the increasing function that is obtained from $W(z) = -w(1 - z)$ by removing the jumps, denoted by H_W , that W has at the points of H . For each point z in V or H , let

$$\Delta W(z, -) = W(z) - W(z-) \text{ and } \Delta W(z, +) = W(z+) - W(z)$$

According to Theorem 3 in Falkner and Teschl (2012),

$$\begin{aligned}
\int_{V_X} x d[-w(1 - F_X(x))] &= \int_V F_X^{\leftarrow}(z) dW_V \\
&+ \sum_{z \in V_W} F_X^{\leftarrow}(z) W(z, -) \\
&+ \sum_{z \in V_W} F_X^{\rightarrow}(z) W(z, +)
\end{aligned}$$

Similarly, one can get

$$\begin{aligned}
\int_{H_X} x d[-w(1 - F_X(x))] &= \int_H F_X^{\leftarrow}(z) dW_H \\
&+ \sum_{z \in H_W} F_X^{\leftarrow}(z) W(z, -) \\
&+ \sum_{z \in H_W} F_X^{\rightarrow}(z) W(z, +)
\end{aligned}$$

Since w is left-continuous, then

$$-w(1 - z) = -w((1 - z)-) = -w(1 - z+)$$

Which implies $W(z, +) = 0$ and

$$\begin{aligned}
\int_{V_X} xd[-w(1 - F_X(x))] &= \int_V F_X^{\leftarrow}(z) dW \\
&= \int_V F_X^{\leftarrow}(z) d[-w(1 - z)] \\
\int_{H_X} xd[-w(1 - F_X(x))] &= \int_H F_X^{\leftarrow}(z) dW \\
&= \int_H F_X^{\leftarrow}(z) d[-w(1 - z)]
\end{aligned}$$

Therefore,

$$\begin{aligned}
\int_0^\infty xd[-w(1 - F_X(x))] &= \int_{N_X} xd[-w(1 - F_X(x))] \\
&+ \int_{V_X} xd[-w(1 - F_X(x))] \\
&+ \int_{H_X} xd[-w(1 - F_X(x))] \\
&= \int_N F_X^{\leftarrow}(z) d[-w(1 - z)] \\
&+ \int_V F_X^{\leftarrow}(z) d[-w(1 - z)] \\
&+ \int_H F_X^{\leftarrow}(z) d[-w(1 - z)] \\
&= \int_0^1 F_X^{\leftarrow}(z) d[-w(1 - z)]
\end{aligned}$$

□

Corollary 2 *Given that w is differentiable,*

$$\int_0^\infty x d[-w(1 - F_X(x))] = E[F_X^{\leftarrow}(Z)w'(1 - Z)] \quad (3.5)$$

for any random variable $Z \sim Unif(0, 1)$.

Corollary 2 confirms equation (2.16) with $u(x) = x$ in He and Zhou (2011).

Instead of finding solutions with respect to random variable X , we can solve the optimal problem regarding quantile function $F_X^{\leftarrow}(z)$.

According to Theorem 3, we can also replace the budget constraint

$$E[\rho X] = x_0$$

by

$$E[F_\rho^{\leftarrow}(1 - Z)F_X^{\leftarrow}(Z)] = x_0$$

Then, the optimal problem (Equation 3.2) is equivalent to:

$$\begin{aligned} \max \quad & \psi(F_X^{\leftarrow}(\cdot)) = E[F_X^{\leftarrow}(Z)w'(1 - Z)] \\ \text{subject to} \quad & E[F_\rho^{\leftarrow}(1 - Z)F_X^{\leftarrow}(Z)] = x_0. \\ & F_X^{\leftarrow}(\cdot) \geq 0 \end{aligned} \quad (3.6)$$

Note that if we consider $F_X^{\leftarrow}(\cdot)$ as a random variable on $[0, 1]$, then $\psi(F_X^{\leftarrow}(\cdot))$ in (Equation 3.6) is equivalent to $\psi(X)$ in (Equation 3.4), because they share the same distribution function. Using β as the Lagrange multiplier, given that $E[F_\rho^{\leftarrow}(1 - Z)F_X^{\leftarrow}(Z)] = x_0$, the preference measure can be rewritten into

$$\begin{aligned}\psi(X) &= E[F_X^{\leftarrow}(Z)w'(1 - Z)] \\ &= E[F_X^{\leftarrow}(Z)(w'(1 - Z) - \beta F_\rho^{\leftarrow}(1 - Z))] + \beta E[F_\rho^{\leftarrow}(1 - Z)F_X^{\leftarrow}(Z)] \\ &= E[F_X^{\leftarrow}(Z)(w'(1 - Z) - \beta F_\rho^{\leftarrow}(1 - Z))] + \beta x_0\end{aligned}$$

Let $v(x_0)$ be the optimal value of (Equation 3.6). Optimization problem (Equation 3.6) can be solved via the following family of optimization problems with parameter β :

$$\begin{aligned}\max_{F_X^{\leftarrow}(\cdot)} \quad & \psi_\beta(F_X^{\leftarrow}(\cdot)) = E[F_X^{\leftarrow}(Z)(w'(1 - Z) - \beta F_\rho^{\leftarrow}(1 - Z))] + \beta x_0 \\ \text{subject to} \quad & F_X^{\leftarrow}(\cdot) \in \mathbb{M}\end{aligned}\tag{3.7}$$

where $\mathbb{M} = \{F_X^{\leftarrow}(\cdot) \geq 0 \mid F_X^{\leftarrow}(\cdot) \text{ is nondecreasing left-continuous on } [0, 1]\}$. Let $v_\beta(x_0)$ be the optimal value of (Equation 3.7). Solution to the original problem can be obtained via $v(x_0) = \inf_\beta v_\beta(x_0)$. Since both $\psi(F_X^{\leftarrow}(Z))$ and $E[F_\rho^{\leftarrow}(1 - Z)F_X^{\leftarrow}(Z)]$ increase along with $F_X^{\leftarrow}(Z)$, $v_\beta(x_0) = \infty$ for $\beta \leq 0$, we only need to consider $\beta > 0$.

Assumption 2 $w(0) = 0$, $w(1) = 1$, w is nondecreasing and differentiable on $[0, 1]$,

$\limsup_{z \uparrow 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} < \infty$, moreover $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ has finite number of monotonic pieces on $z \in [0, 1]$, that is, $\exists 0 = z_0 < z_1 < \dots < z_n = 1$, s.t. $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ is monotone on (z_i, z_{i+1}) for $i = 0, 1, \dots, n-1$.

In Assumption 2, $\limsup_{z \downarrow 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} < \infty$ guarantees existence of $v(x_0)$, the optimal value of (Equation 3.6).

Assumption 2 generalizes Assumption 3.5 in He and Zhou (2011) where $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ was assumed to be continuous, strictly increasing, and then strictly decreasing on $(0, 1)$.

Lemma 7 Under Assumption 2, for any given $\beta \in (\inf_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}, \sup_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)})$, there always exists a $z^*(\beta) \in [0, 1]$ such that

$$v_\beta(x_0) = \sup_{F_X^{\leftarrow}(\cdot) \in \mathbb{F}_\beta} \psi_\beta(F_X^{\leftarrow}(\cdot)),$$

that is, the optimal solution of (Equation 3.7) can be attained in the subclass \mathbb{F}_β , which is defined by

$$\mathbb{F}_\beta := \{F_X^{\leftarrow}(\cdot) \in \mathbb{M} : F_X^{\leftarrow}(z) = b \mathbf{1}_{\{z^*(\beta) < z \leq 1\}}, b > 0\}$$

Proof of Lemma 7 is in the following two sections.

3.3.3 Algorithm: Search-and-Cut

We describe the procedure of finding $z^*(\beta)$ in the following three cases:

1. $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ is continuous, and have a finite number of strictly monotonic pieces on $z \in [0, 1]$.
2. $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ is continuous, and have a finite number of monotonic pieces on $z \in [0, 1]$, but not necessarily strictly monotonic.
3. $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ may not be continuous, but still have a finite number of monotonic pieces on $z \in [0, 1]$.

Under Case 1, we propose an algorithm called “Search-and-Cut” to obtain $z^*(\beta)$:

Algorithm: Search-and-Cut

1° Initialization: $z_0(\beta) = 0$ and $\tilde{z}(\beta) = 0$.

2° $\underline{z}(\beta) := \inf\{z \in (\tilde{z}(\beta), 1] : \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} - \beta = 0\}$,

$$\lim_{z_- \nearrow z} \frac{w'(1-z_-)}{F_\rho^{\leftarrow}(1-z_-)} - \beta \nearrow 0, \text{ and } \lim_{z_+ \searrow z} \frac{w'(1-z_+)}{F_\rho^{\leftarrow}(1-z_+)} - \beta \searrow 0\}$$

If $\underline{z}(\beta) := \inf \emptyset$, then skip Step 3 and go to Step 4. Otherwise, go to Step 3.

3° (a) If $\int_{z_0(\beta)}^{\underline{z}(\beta)} (w'(1-z) - \beta F_\rho^{\leftarrow}(1-z)) dz < 0$, then let

$$z_0(\beta) = \underline{z}(\beta),$$

$$\tilde{z}(\beta) = \underline{z}(\beta),$$

and go back to Step 2.

(b) If $\int_{z_0(\beta)}^{\underline{z}(\beta)} (w'(1-z) - \beta F_\rho^{\leftarrow}(1-z)) dz \geq 0$, then let

$$\tilde{z}(\beta) = \underline{z}(\beta),$$

and go back to Step 2.

4° $z^*(\beta) := z_0(\beta)$

□

In the next section, we provide the proof of Lemma 7 under Case 1, 2 and 3 based on Algorithm: Search-and-Cut for Case 1, 2 and 3. Note that Case 2 is a generalization of Case 1, and Case 3 is a generalization of Case 2.

3.3.4 Proof of Lemma 7

We first prove Lemma 7 under Case 1. After getting $z^*(\beta)$ from the above algorithm, we consider the optimal value, $v_\beta(x_0)$.

(a) Since $\beta, v_\beta(x_0)$ is the optimal value, given β , then

$$\sup_{F_X^{\leftarrow}(\cdot) \in \mathbb{F}_\beta} \psi_\beta(F_X^{\leftarrow}(\cdot)) \leq v_\beta(x_0)$$

(b) On the other hand, it can be shown that the $z^*(\beta)$ obtained from the ‘‘Search-and-Cut’’ algorithm has the following property:

$$\int_0^{z^*(\beta)} (w'(1-z) - \beta F_\rho^{\leftarrow}(1-z)) dz < 0, \text{ if } z^*(\beta) > 0$$

Define $\bar{z}(\beta) := \sup\{z \in (\tilde{z}(\beta), 1] : \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} - \beta = 0\}$,

$$\lim_{z_- \nearrow z} \left(\frac{w'(1-z_-)}{F_\rho^{\leftarrow}(1-z_-)} - \beta \right) \searrow 0, \text{ and } \lim_{z_+ \searrow z} \left(\frac{w'(1-z_+)}{F_\rho^{\leftarrow}(1-z_+)} - \beta \right) \nearrow 0\}$$

If $\bar{z}(\beta) := \sup \emptyset$, then let $\bar{z}(\beta) = 1$. We define $\alpha := \lim_{z \nearrow \bar{z}(\beta)} F_X^{\leftarrow}(z) = F_X^{\leftarrow}(\bar{z}(\beta))$.

Note that for any triple $0 \leq z_p < z_{+,-} < z_n \leq 1$, such that

$$w'(1-z) - \beta F_\rho^{\leftarrow}(1-z) \begin{cases} > 0, & \text{if } z \in (z_p, z_{+,-}) \\ = 0, & \text{if } z = z_{+,-} \\ < 0, & \text{if } z \in (z_{+,-}, z_n) \end{cases}$$

the following inequality holds:

$$\begin{aligned} & \int_{z_p}^{z_n} F_X^{\leftarrow}(z)(w'(1-z) - \beta F_\rho^{\leftarrow}(1-z))dz \\ = & \int_{z_p}^{z_{+,-}} F_X^{\leftarrow}(z)(w'(1-z) - \beta F_\rho^{\leftarrow}(1-z))dz \\ + & \int_{z_{+,-}}^{z_n} F_X^{\leftarrow}(z)(w'(1-z) - \beta F_\rho^{\leftarrow}(1-z))dz \\ \leq & F_X^{\leftarrow}(z_{+,-}) \int_{z_p}^{z_{+,-}} (w'(1-z) - \beta F_\rho^{\leftarrow}(1-z))dz \\ + & F_X^{\leftarrow}(z_{+,-}) \int_{z_{+,-}}^{z_n} (w'(1-z) - \beta F_\rho^{\leftarrow}(1-z))dz \\ = & F_X^{\leftarrow}(z_{+,-}) \int_{z_p}^{z_n} (w'(1-z) - \beta F_\rho^{\leftarrow}(1-z))dz \end{aligned} \tag{3.8}$$

For any feasible $F_X^{\leftarrow}(\cdot) \in \mathbb{M}$, one can get,

$$\begin{aligned}
\psi_\beta(F_X^{\leftarrow}(\cdot)) &= E[F_X^{\leftarrow}(Z)(w'(1-Z) - \beta F_\rho^{\leftarrow}(1-Z))] + \beta x_0 \\
&= \int_0^{z^*(\beta)} F_X^{\leftarrow}(z)(w'(1-z) - \beta F_\rho^{\leftarrow}(1-z))dz \\
&\quad + \int_{z^*(\beta)}^{\tilde{z}(\beta)} F_X^{\leftarrow}(z)(w'(1-z) - \beta F_\rho^{\leftarrow}(1-z))dz \\
&\quad + \int_{\tilde{z}(\beta)}^1 F_X^{\leftarrow}(z)(w'(1-z) - \beta F_\rho^{\leftarrow}(1-z))dz + \beta x_0 \\
&\leq 0 + F_X^{\leftarrow}(\tilde{z}(\beta)) \int_{z^*(\beta)}^{\tilde{z}(\beta)} (w'(1-z) - \beta F_\rho^{\leftarrow}(1-z))dz \quad (3.9) \\
&\quad + F_X^{\leftarrow}(\tilde{z}(\beta)) \int_{\tilde{z}(\beta)}^1 (w'(1-z) - \beta F_\rho^{\leftarrow}(1-z))dz + \beta x_0 \\
&= \alpha \int_{z^*(\beta)}^1 (w'(1-z) - \beta F_\rho^{\leftarrow}(1-z))dz + \beta x_0 \\
&\leq \sup_{H(\cdot) \in \mathbb{F}_\beta} \psi_\beta(H(\cdot))
\end{aligned}$$

Note that the function $w'(1-z) - \beta F_\rho^{\leftarrow}(1-z)$ on $(z^*(\beta), \tilde{z}(\beta))$ can be split into a few subintervals each equipped with a triple $(z_p, z_{+,-}, z_n)$. Within each subinterval (z_p, z_n) inequality (Equation 3.8) holds. The inequality in (Equation 3.9) can be verified by replacing all $F_X^{\leftarrow}(z_{+,-})$ with $F_X^{\leftarrow}(\tilde{z}(\beta))$. Similarly, one can verify that the inequality holds for z on $(\tilde{z}(\beta), 1)$ when replacing $F_X^{\leftarrow}(z)$ by $F_X^{\leftarrow}(\tilde{z}(\beta))$.

The inequality in (Equation 3.9) attains equality, if and only if

$$F_X^{\leftarrow}(z) = \alpha \mathbf{1}_{\{z^*(\beta) < z \leq 1\}}$$

After combining (a) and (b), the optimal problem (Equation 3.7) has the optimal value:

$$v_\beta(x_0) = \sup_{F_X^\leftarrow \in \mathbb{F}_\beta} \psi_\beta(F_X^\leftarrow(\cdot)),$$

with an optimal solution taking the form of

$$\hat{F}_X^\leftarrow(\cdot) = b \mathbf{1}_{\{z^*(\beta) < z \leq 1\}} \text{ for some } b > 0.$$

Now we prove Lemma 7 under Case 2 and Case 3. The proof follows the similar procedure under Case 1 up to some adjustments, specifically on finding $\underline{z}(\beta)$ in Step 2 of the algorithm.

Under Case 2, let's define

$$\mathbb{Z}_{const} = \{[z_l, z_r] \subset (\tilde{z}(\beta), 1] : z_l \leq z_r, \forall z \in [z_l, z_r], \frac{w'(1-z)}{F_\rho^\leftarrow(1-z)} - \beta = 0, \text{ and}$$

$$\lim_{z_- \nearrow z_l} \frac{w'(1-z_-)}{F_\rho^\leftarrow(1-z_-)} - \beta \nearrow 0, \text{ and } \lim_{z_+ \searrow z_r} \frac{w'(1-z_+)}{F_\rho^\leftarrow(1-z_+)} - \beta \searrow 0\}$$

Then \mathbb{Z}_{const} is a collection of non-overlapped closed intervals or single points. For any $[z_l, z_r] \in \mathbb{Z}_{const}$, $\frac{w'(1-z)}{F_\rho^\leftarrow(1-z)} = \beta$, $\forall z \in [z_l, z_r]$. Let $[z_{l(1)}, z_{r(1)}]$ be the interval in \mathbb{Z}_{const} with the smallest values, which means

$$[z_{l(1)}, z_{r(1)}] \in \mathbb{Z}_{const} \text{ and } \forall [z_l, z_r] \neq [z_{l(1)}, z_{r(1)}] \in \mathbb{Z}_{const}, z_l > z_{l(1)}$$

The set \mathbb{Z}_{const} and interval $[z_{l(1)}, z_{r(1)}]$ will be updated along with $\tilde{z}(\beta)$. Those $z \in [z_{l(1)}, z_{r(1)}]$ will also be taken into consideration when finding $\underline{z}(\beta)$ through iteration.

Under Case 2, we only need to change the definition of $\underline{z}(\beta)$ in Step 2 into:

$$\underline{z}(\beta) = z_{l(1)}$$

and the iteration of $\tilde{z}(\beta)$ in Step 3 into:

$$\tilde{z}(\beta) = z_{r(1)}$$

The rest of the algorithm remains the same, which results in $z^*(\beta)$ for each β , such that Lemma 7 holds.

Note for Case 2, the optimal value is unique, but optimal solutions may not be unique. In the expression $\hat{F}_X^{\leftarrow}(\cdot) = b\mathbf{1}_{\{z^*(\beta) < z \leq 1\}}$, $z^*(\beta)$ can be any value in $[z_{l(1)}, z_{r(1)}]$, which comes from the last iteration in the algorithm.

As for Case 3, Step 2 in the ‘‘Search-and Cut’’ algorithm can be generalized by defining

$$\mathbb{Z}_{upcross} = \{z \in (\tilde{z}(\beta), 1] :$$

$$\lim_{z_- \nearrow z} \frac{w'(1 - z_-)}{F_\rho^{\leftarrow}(1 - z_-)} - \beta \nearrow 0 \text{ or } \lim_{z_- \nearrow z} \frac{w'(1 - z_-)}{F_\rho^{\leftarrow}(1 - z_-)} - \beta < 0,$$

$$\begin{aligned}
& \lim_{z_+ \searrow z} \frac{w'(1-z_+)}{F_\rho^{\leftarrow}(1-z_+)} - \beta \searrow 0 \text{ or } \lim_{z_+ \searrow z} \frac{w'(1-z_+)}{F_\rho^{\leftarrow}(1-z_+)} - \beta > 0 \} \\
& \cup \{ [z_l, z_r] \subset (\tilde{z}(\beta), 1] : z_l < z_r, \forall z \in (z_l, z_r), \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} - \beta = 0, \\
& \lim_{z_- \nearrow z_l} \frac{w'(1-z_-)}{F_\rho^{\leftarrow}(1-z_-)} - \beta \nearrow 0 \text{ or } \lim_{z_- \nearrow z_l} \frac{w'(1-z_-)}{F_\rho^{\leftarrow}(1-z_-)} - \beta < 0, \\
& \lim_{z_+ \searrow z_r} \frac{w'(1-z_+)}{F_\rho^{\leftarrow}(1-z_+)} - \beta \searrow 0 \text{ or } \lim_{z_+ \searrow z_r} \frac{w'(1-z_+)}{F_\rho^{\leftarrow}(1-z_+)} - \beta > 0 \}
\end{aligned}$$

Note that $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ may not be continuous, but it has finite number of monotone pieces, which means

$$\limsup_{z \nearrow z_0} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} = \liminf_{z \nearrow z_0} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$$

That is, $\lim_{z \nearrow z_0} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ must exist. Similarly, $\lim_{z \searrow z_0} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ must exist.

Then $\mathbb{Z}_{upcross}$ is a collection of non-overlapped closed intervals or single points. For any component (closed interval or single point) of $\mathbb{Z}_{upcross}$, as a function of z , $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} - \beta$ passes upwards through 0. Let $[z_{l(1)}, z_{r(1)}]$ be the interval or point in $\mathbb{Z}_{upcross}$ with the smallest value(s), which means

$$[z_{l(1)}, z_{r(1)}] \in \mathbb{Z}_{upcross}, \forall z \in \mathbb{Z}_{upcross}, z \geq z_{l(1)},$$

$$\text{and } \forall [z_l, z_r] \neq [z_{l(1)}, z_{r(1)}] \in \mathbb{Z}_{upcross}, z_l > z_{l(1)}$$

The set $\mathbb{Z}_{upcross}$ and interval $[z_{l(1)}, z_{r(1)}]$ will be updated along with $\tilde{z}(\beta)$.

Under Case 3, we only need to change the definition of $\underline{z}(\beta)$ in Step 2 into:

$$\underline{z}(\beta) = z_{l(1)}$$

and the iteration of $\tilde{z}(\beta)$ in Step 3 into:

$$\tilde{z}(\beta) = z_{r(1)}$$

The rest of the algorithm remains the same, which results in $z^*(\beta)$ for each β , such that Lemma 7 holds. □

Note that Case 3 is a generalization of Case 2, whereas Case 2 is a generalization of Case 1. There are cases that are not covered by those three cases, since function $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ may have infinite number of monotonic pieces (see Example 4).

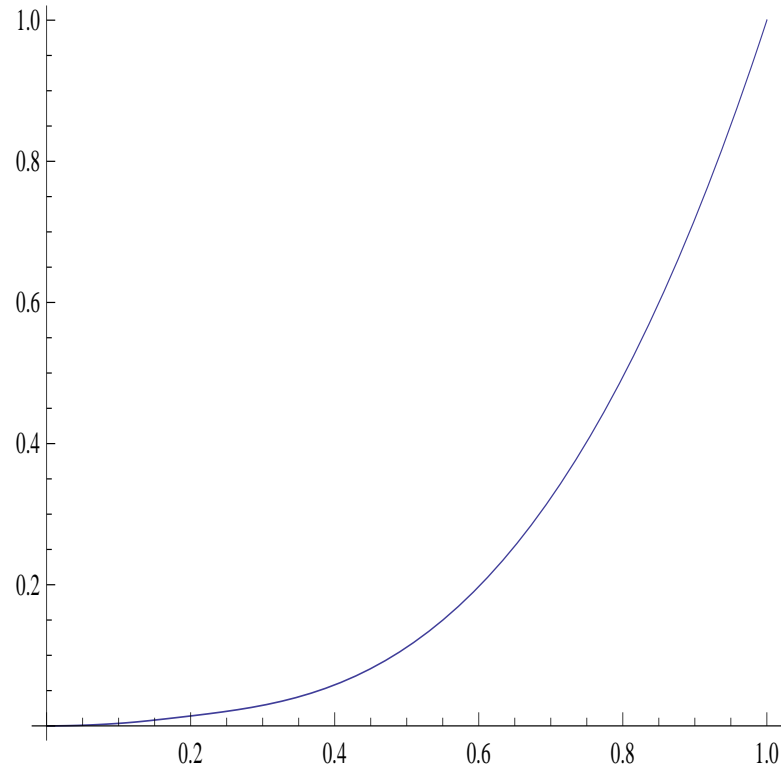
3.3.5 Example of Extreme Case: Infinite Number of Monotone Pieces

Example 4

$$\rho(x) \equiv x, x \in [0, 1], \text{ thus } F_\rho^{\leftarrow}(1-z) = 1-z$$

$$w(z) = \frac{\pi + 6z^2 + 2z((2z^2 - 1) \cos \frac{1}{z} + z(6z^2 - 1) \sin \frac{1}{z}) - 2 \int_0^{\frac{1}{z}} \frac{\sin t}{t} dt}{6 + \pi + 2 \cos(1) + 10 \sin(1) - 2 \int_0^1 \frac{\sin t}{t} dt}$$

Figure 12. Graph of an Artificial Distortion Function w

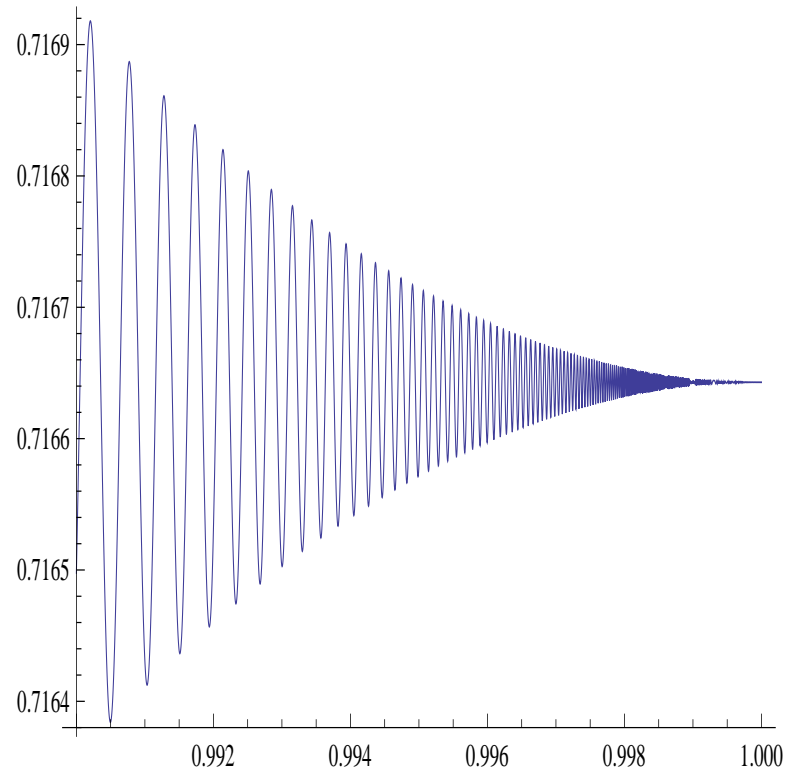


which satisfy $w(0) = 0$, $w(1) = 1$, and w is differentiable. Note that $2 \int_0^\infty \frac{\sin t}{t} dt = \frac{\pi}{2}$. Graph of w is shown in Figure 12,

One can verify that:

$$\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} = d((1-z)^2 \sin \frac{1}{1-z} + \frac{1}{4}), \quad z \in (0, 1)$$

Figure 13. Graph of $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ with Infinite Number of Monotonic Pieces in Example 4



where d

$$d = \frac{48}{6 + \pi + 2 \cos(1) + 10 \sin(1) - 2 \int_0^1 \frac{\sin t}{t} dt}$$

The graph of $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ is shown in Figure 13.

When $\beta = 0.71665$, $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} - \beta$ would pass upward through 0 infinite times. The algorithm we conduct won't work for this case.

Although cases like Example 4 do exist, in practice one can always replace the problematic part (say w on $(1 - 10^{-6}, 1)$) with a polynomial curve. Then the modified w falls into Case 2 and the optimal solution for the modified w has ignorable difference from the one based on the original w . Therefore Lemma 7 is good enough in practice.

Detailed derivation of Example 4:

Starting from the function of $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ and derive backward:

$$\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} = d((1-z)^2 \sin \frac{1}{1-z} + \frac{1}{4}), z \in (0, 1)$$

Since $F_\rho^{\leftarrow}(1-z) = 1-z$, one can get:

$$w'(1-z) = d \cdot (1-z)((1-z)^2 \sin \frac{1}{1-z} + \frac{1}{4})$$

Replace $1-z$ by x , $x \in (0, 1)$, which is

$$w'(x) = d \cdot x(x^2 \sin \frac{1}{x} + \frac{1}{4}), x \in (0, 1)$$

The distortion probability w can be expressed as:

$$w(x) = d \int_0^x t(t^2 \sin \frac{1}{t} + \frac{1}{4})dt + c$$

where c is a constant, knowing $w(0) = 0$ and $w(1) = 1$, one can get

$$c = 0$$

and d satisfies

$$d \int_0^1 (x^2 \sin \frac{1}{x} + \frac{1}{4})dx = 1$$

The value of d is:

$$d = \frac{48}{6 + \pi + 2 \cos(1) + 10 \sin(1) - 2 \int_0^1 \frac{\sin(t)}{t} dt}$$

Thus, given β , the function $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} - \beta$ has infinite number of points up-crossing 0.

3.3.6 Derivation of Optimal Solution

Theorem 4 *Under Assumption 2, the optimization problem (Equation 3.6) can be solved and*

(a) *The optimal value is*

$$v(x_0) = \beta^* x_0$$

where $\beta^* \in (\inf_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}, \sup_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)})$ s.t.

$$\int_{z^*(\beta^*)}^1 (w'(1-z) - \beta^* F_\rho^{\leftarrow}(1-z)) dz = 0$$

and $z^*(\beta)$ is defined as in Lemma 7 for a given β .

(b) The optimal solution is $F_{X^*}^{\leftarrow}(z) = b^* \mathbf{1}_{\{z^*(\beta^*) < z \leq 1\}}$ with $b^* = \frac{x_0}{\int_{z^*(\beta^*)}^1 F_\rho^{\leftarrow}(1-z) dz}$

Proof of (a) of Theorem 4 Based on Lemma 7,

$$\begin{aligned} v_\beta(x_0) &= \sup_{F_X^{\leftarrow}(\cdot) \in \mathbb{F}_\beta} \psi_\beta(F_X^{\leftarrow}(\cdot)) \\ &= \sup_{b > 0} [b \int_{z^*(\beta)}^1 (w'(1-z) - \beta F_\rho^{\leftarrow}(1-z)) dz] + \beta x_0 \end{aligned}$$

Note that for given β , $\int_{z^*(\beta)}^1 (w'(1-z) - \beta F_\rho^{\leftarrow}(1-z)) dz$ and βx_0 are fixed.

We denote $g(\beta) = \int_{z^*(\beta)}^1 (w'(1-z) - \beta F_\rho^{\leftarrow}(1-z)) dz$.

Then the optimization problem (Equation 3.7) can be solved as follows

$$\left\{ \begin{array}{ll} v_\beta(x_0) = +\infty. & \text{if } g(\beta) > 0; \\ v_\beta(x_0) = \beta x_0, \text{ and } F_X^{\leftarrow}(z) \equiv 0, z \in (0, 1] & \text{if } g(\beta) < 0; \\ v_\beta(x_0) = \beta x_0, \text{ and } F_X^{\leftarrow}(z) = b \mathbf{1}_{\{z^*(\beta) < z \leq 1\}} & \text{if } g(\beta) = 0. \end{array} \right.$$

Therefore, $\inf_{\beta > 0} v_\beta(x_0) = \beta^* x_0$.

(1) Note that $v(x_0) \leq \inf_{\beta > 0} v_\beta(x_0) = \beta^* x_0$

(2) On the other hand, if we take $\beta = \beta^*$, and $\hat{F}_X^{\leftarrow}(z) = b\mathbf{1}_{\{z^*(\beta^*) < z \leq 1\}}$, $0 \leq z \leq 1$, such that

$$E[F_\rho^{\leftarrow}(1-Z)\hat{F}_X^{\leftarrow}(Z)] = x_0.$$

The objective function of (Equation 3.7) becomes

$$\begin{aligned} \psi(\hat{F}_X^{\leftarrow}(\cdot)) &= E[\hat{F}_X^{\leftarrow}(Z)(w'(1-Z) - \beta^*F_\rho^{\leftarrow}(1-Z))] + \beta^*E[\hat{F}_\rho^{\leftarrow}(1-Z)\hat{F}_X^{\leftarrow}(Z)] \\ &= E[\hat{F}_X^{\leftarrow}(Z)(w'(1-Z) - \beta F_\rho^{\leftarrow}(1-Z))] + \beta^*x_0 \\ &= \int_{z^*(\beta^*)}^1 (w'(1-z) - \beta F_\rho^{\leftarrow}(1-z))dz + \beta^*x_0 \\ &= \beta^*x_0 \end{aligned}$$

Since $v(x_0) \geq \psi(\hat{F}_X^{\leftarrow}(\cdot))$, then $\beta^*x_0 = \psi_\beta(\hat{F}_X^{\leftarrow}(\cdot)) \leq v(x_0)$.

Combine (1) and (2), one can get $v(x_0) = \psi_\beta(\hat{F}_X^{\leftarrow}(\cdot)) = \beta^*x_0$ which is the optimal value for (Equation 3.7). \square

3.3.7 Existence and Uniqueness of Optimal Solution

Lemma 8 *Under Assumption 2, $z^*(\beta)$ is increasing and left continuous on $\beta \in (\inf_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}, \sup_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)})$.*

Proof of Lemma 8:

Suppose $\beta_0, \beta_1 \in (\inf_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}, \sup_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)})$ and $\beta_0 < \beta_1$. According to the Search-and-Cut algorithm

$$\int_0^{z^*(\beta_0)} (w'(1-z) - \beta_0 F_\rho^{\leftarrow}(1-z)) dz < 0.$$

Since $F_\rho^{\leftarrow}(1-z)$ is always positive on $z \in [0, 1)$,

$$\int_0^{z^*(\beta_0)} (w'(1-z) - \beta_1 F_\rho^{\leftarrow}(1-z)) dz < 0.$$

Based on the definition of $\mathbb{Z}_{upcross}$ in Lemma 7, one could find $z_{l(1),\beta_1}$ by plugging in $\tilde{z}(\beta) = z^*(\beta_0)$ and $\beta = \beta_1$. Here $z_{l(1),\beta_1}$ is the smallest value of z , such that $z > z^*(\beta_0)$ and $w'(1-z) - \beta_1 F_\rho^{\leftarrow}(1-z)$ goes up crossing 0. Then

$$\int_{z^*(\beta_0)}^{z_{l(1),\beta_1}} (w'(1-z) - \beta_1 F_\rho^{\leftarrow}(1-z)) dz < 0.$$

Therefore,

$$\int_0^{z_{l(1),\beta_1}} (w'(1-z) - \beta_1 F_\rho^{\leftarrow}(1-z)) dz < 0.$$

Based on the Search-and-Cut algorithm, either $z^*(\beta_1) = z_{l(1),\beta_1}$ or there is a jump over one or a few points or intervals in $\mathbb{Z}_{upcross}$, which leads to $z^*(\beta_1) > z_{l(1),\beta_1}$.

Therefore, $z^*(\beta)$ could be either continuously increasing on β or, as discussed above, jump over points or intervals in $\mathbb{Z}_{upcross}$, but still be increasing on β .

In either case, $z^*(\beta)$ is left continuous. □

Lemma 9 *Under Assumption 2, $g(\beta)$ is strictly decreasing on $\beta \in (\inf_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}, \sup_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)})$.*

Proof of Lemma 9:

Following the proof of Lemma 8, if $z^*(\beta_1) = z_{l(1),\beta_1}$, it is clear that

$$\int_{z^*(\beta_0)}^{z^*(\beta_1)} (w'(1-z) - \beta_0 F_\rho^{\leftarrow}(1-z)) dz > 0.$$

If there is a jump, that is, $z^*(\beta_1) > z_{l(1),\beta_1}$, one could find $z_{r(1),\beta_0}$ which is the biggest value of z , such that $z < z^*(\beta_1)$ and $w'(1-z) - \beta_0 F_\rho^{\leftarrow}(1-z)$ upcrosses 0 at z . Based on the algorithm, one can get

$$\int_{z^*(\beta_0)}^{z_{r(1),\beta_0}} (w'(1-z) - \beta_0 F_\rho^{\leftarrow}(1-z)) dz > 0, \text{ and}$$

$$\int_{z_{r(1),\beta_0}}^{z^*(\beta_1)} (w'(1-z) - \beta_0 F_\rho^{\leftarrow}(1-z)) dz > 0$$

Thus, for both two cases

$$\int_{z^*(\beta_0)}^{z^*(\beta_1)} (w'(1-z) - \beta_0 F_\rho^{\leftarrow}(1-z)) dz > 0.$$

Then

$$\begin{aligned}
g(\beta_0) &= \int_{z^*(\beta_0)}^{z^*(\beta_1)} (w'(1-z) - \beta_0 F_\rho^{\leftarrow}(1-z)) dz \\
&+ \int_{z^*(\beta_1)}^1 (w'(1-z) - \beta_0 F_\rho^{\leftarrow}(1-z)) dz \\
&> \int_{z^*(\beta_1)}^1 (w'(1-z) - \beta_0 F_\rho^{\leftarrow}(1-z)) dz \\
&> \int_{z^*(\beta_1)}^1 (w'(1-z) - \beta_1 F_\rho^{\leftarrow}(1-z)) dz \\
&= g(\beta_1)
\end{aligned}$$

That is, $g(\beta)$ is strictly decreasing on β .

Lemma 10 *Under Assumption 2, $g(\beta)$ is continuous on $(\inf_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}, \sup_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)})$.*

Proof of Lemma 10:

If $z^*(\beta)$ continuous at β , it is straight forward that $g(\beta)$ is continuous at β too. For the other case, as discussed in the proof of Lemma 8, for example, $z^*(\beta)$ has a jump, at point β_J , then the continuity of $g(\beta)$ at β_J can be verified as follows.

First, for $\beta_- \nearrow \beta_J$, since $z^*(\beta_-) \nearrow z^*(\beta_J)$ due to Lemma 8,

$$\lim_{\beta_- \nearrow \beta_J} \int_{z^*(\beta_-)}^{z^*(\beta_J)} (w'(1-z) - \beta_- F_\rho^{\leftarrow}(1-z)) dz = 0 \text{ and}$$

$$\lim_{\beta_- \nearrow \beta_J} \int_{z^*(\beta_J)}^1 (w'(1-z) - \beta_- F_\rho^{\leftarrow}(1-z)) dz - \int_{z^*(\beta_J)}^1 (w'(1-z) - \beta_J F_\rho^{\leftarrow}(1-z)) dz = 0,$$

then

$$\begin{aligned} \lim_{\beta_- \nearrow \beta_J} g(\beta_-) - g(\beta_J) &= \lim_{\beta_- \nearrow \beta_J} \int_{z^*(\beta_-)}^{z^*(\beta_J)} (w'(1-z) - \beta_- F_\rho^{\leftarrow}(1-z)) dz \\ &+ \lim_{\beta_- \nearrow \beta_J} \int_{z^*(\beta_J)}^1 (w'(1-z) - \beta_- F_\rho^{\leftarrow}(1-z)) dz \\ &- \int_{z^*(\beta_J)}^1 (w'(1-z) - \beta_J F_\rho^{\leftarrow}(1-z)) dz \\ &= 0 + 0 = 0 \end{aligned}$$

That is,

$$\lim_{\beta_- \nearrow \beta_J} g(\beta_-) = g(\beta_J).$$

Secondly, for $\beta_+ \searrow \beta_J$, if $z^*(\beta)$ has a jump at β_J , one could always find $z_{r(1),\beta_J}$, which is the biggest value of z , such that $z < z^*(\beta_+)$ and $w'(1-z) - \beta_J F_\rho^{\leftarrow}(1-z)$ upcross 0. According to the Search-and-Cut algorithm,

$$\int_{z^*(\beta_J)}^{z_{r(1),\beta_J}} (w'(1-z) - \beta_J F_\rho^{\leftarrow}(1-z)) dz = 0,$$

$$\lim_{\beta_+ \searrow \beta_J} \int_{z_{r(1),\beta_J}}^{z^*(\beta_+)} (w'(1-z) - \beta_J F_\rho^{\leftarrow}(1-z)) dz = 0, \text{ and}$$

$$\lim_{\beta_+ \searrow \beta_J} \int_{z^*(\beta_+)}^1 (w'(1-z) - \beta_J F_\rho^{\leftarrow}(1-z)) dz - \int_{z^*(\beta_+)}^1 (w'(1-z) - \beta_+ F_\rho^{\leftarrow}(1-z)) dz = 0.$$

Then

$$\begin{aligned} \lim_{\beta_+ \searrow \beta_J} [g(\beta_J) - g(\beta_+)] &= \int_{z^*(\beta_J)}^{z_r(1), \beta_J} (w'(1-z) - \beta_J F_\rho^{\leftarrow}(1-z)) dz \\ &+ \lim_{\beta_+ \searrow \beta_J} \int_{z_r(1), \beta_J}^{z^*(\beta_+)} (w'(1-z) - \beta_J F_\rho^{\leftarrow}(1-z)) dz \\ &+ \lim_{\beta_+ \searrow \beta_J} \int_{z^*(\beta_+)}^1 (w'(1-z) - \beta_J F_\rho^{\leftarrow}(1-z)) dz \\ &- \lim_{\beta_+ \searrow \beta_J} \int_{z^*(\beta_+)}^1 (w'(1-z) - \beta_+ F_\rho^{\leftarrow}(1-z)) dz \\ &= 0 + 0 + 0 = 0 \end{aligned}$$

Therefore,

$$\lim_{\beta_+ \searrow \beta_J} g(\beta_+) \nearrow g(\beta_J).$$

That is, $g(\beta)$ is always continuous on $(\inf_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}, \sup_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)})$, regardless of the continuity of $z^*(\beta)$. \square

Combining Lemmas 8, 9, 10, we conclude that $g(\beta)$ is continuous and strictly decreasing on

$$(\inf_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}, \sup_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}).$$

To discuss the existence and uniqueness of the solution to Equation 3.6, we define $\beta_{sup} =$

$esssup_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$, and $\beta_{inf} = essinf_{0 \leq z \leq 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$.

A trivial case, (see Example 5) is that $\beta_{sup} = \beta_{inf}$, then $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} = c$ a.s. and $\beta^* = \beta_{sup} = \beta_{inf}$, then the optimal value is $\beta^* x_0$.

Example 5 $\rho \sim Unif(0, 1)$, $F_\rho^{\leftarrow}(x) \equiv x$, $0 < x \leq 1$. $w(x) = x^2$, $0 \leq x \leq 1$, which satisfies $w(0) = 0$, $w(1) = 1$ and w is differentiable. One can get

$$\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} \equiv 2, 0 < z \leq 1$$

In this case, $\beta^* = 2$, the optimal value is $2x_0$, and optimal solutions are not unique (refer to Case 2 in Lemma 7). Since $z^*(\beta)$ can be any value in $(0, 1)$, optimal solution is $\hat{F}_X^{\leftarrow}(\cdot) = b \mathbf{1}_{\{z^*(\beta) < z \leq 1\}}$, where b can be determined by the budget constraint and value of $z^*(\beta)$.

For non-trivial case, we suppose $\beta_{sup} > \beta_{inf}$, since we assume $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)}$ has finite number of monotonic pieces on $[0, 1]$, then $\{0 \leq z \leq 1 \mid \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} \geq \beta_{sup}\}$ and $\{0 \leq z \leq 1 \mid \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} \leq \beta_{inf}\}$ are at most finite unions of single points and intervals.

First we assume $\limsup_{z \rightarrow 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} < \beta_{sup}$. We claim that there always exists $\epsilon > 0$, such that $g(\beta_{sup} - \epsilon) < 0$ and $g(\beta_{inf} + \epsilon) > 0$.

Actually, for any given $\epsilon_0 > 0$, $\frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} < \beta_{sup} - \epsilon_0$ for $1 - \delta_0 < z < 1$.

For any $0 < \epsilon < \epsilon_0$, if $z^*(\beta_{sup} - \epsilon) \geq 1 - \delta_0$, then

$$\begin{aligned} g(\beta_{sup} - \epsilon) &= \int_{z^*(\beta_{sup} - \epsilon)}^1 (w'(1-z) - (\beta_{sup} - \epsilon)F_\rho^{\leftarrow}(1-z))dz \\ &< \int_{z^*(\beta_{sup} - \epsilon)}^1 (w'(1-z) - (\beta_{sup} - \epsilon_0)F_\rho^{\leftarrow}(1-z))dz \\ &< 0 \end{aligned}$$

Otherwise $z^*(\beta_{sup} - \epsilon) < 1 - \delta_0$. We still have

$$\begin{aligned} \lim_{\epsilon \searrow 0} g(\beta_{sup} - \epsilon) &= \lim_{\epsilon \searrow 0} \int_{z^*(\beta_{sup} - \epsilon)}^1 (w'(1-z) - (\beta_{sup} - \epsilon)F_\rho^{\leftarrow}(1-z))dz \\ &= \lim_{\epsilon \searrow 0} \int_{z^*(\beta_{sup} - \epsilon)}^{1-\delta_0} (w'(1-z) - (\beta_{sup} - \epsilon)F_\rho^{\leftarrow}(1-z))dz \\ &\quad + \lim_{\epsilon \searrow 0} \int_{1-\delta_0}^1 (w'(1-z) - (\beta_{sup} - \epsilon)F_\rho^{\leftarrow}(1-z))dz \\ &< \lim_{\epsilon \searrow 0} \epsilon(1 - \delta_0 - z^*(\beta_{sup} - \epsilon))F_\rho^{\leftarrow}(1 - z^*(\beta_{sup} - \epsilon)) + \lim_{\epsilon \searrow 0} (\epsilon - \epsilon_0)\delta_0 \\ &= 0 - \epsilon_0\delta_0 \\ &< 0 \end{aligned}$$

$F_\rho^{\leftarrow}(1 - z^*(\beta_{sup} - \epsilon))$ is decreasing when $\epsilon \searrow 0$. Therefore, $\lim_{\epsilon \searrow 0} g(\beta_{sup} - \epsilon) < 0$.

Now we consider $g(\beta_{inf} - \epsilon)$. Let

$$z_0 = \inf\{z : \int_z^1 (w'(1-z) - \beta_{inf}F_\rho^{\leftarrow}(1-z))dz = 0\},$$

based on the algorithm Search-and-Cut, we always have $z^*(\beta_{inf} + \epsilon) < z_0$.

$$\begin{aligned}
& \lim_{\epsilon \searrow 0} \int_{z^*(\beta_{inf} + \epsilon)}^1 (w'(1-z) - \beta_{inf} F_\rho^{\leftarrow}(1-z)) dz \\
&= \lim_{\epsilon \searrow 0} \int_{z^*(\beta_{inf} + \epsilon)}^{z_0} (w'(1-z) - \beta_{inf} F_\rho^{\leftarrow}(1-z)) dz \\
&+ \lim_{\epsilon \searrow 0} \int_{z_0}^1 (w'(1-z) - \beta_{inf} F_\rho^{\leftarrow}(1-z)) dz \\
&> 0
\end{aligned}$$

$$\begin{aligned}
\lim_{\epsilon \searrow 0} g(\beta_{inf} + \epsilon) &= \lim_{\epsilon \searrow 0} \int_{z^*(\beta_{inf} + \epsilon)}^1 (w'(1-z) - (\beta_{inf} + \epsilon) F_\rho^{\leftarrow}(1-z)) dz \\
&= \lim_{\epsilon \searrow 0} \int_{z^*(\beta_{inf} + \epsilon)}^1 (w'(1-z) - \beta_{inf} F_\rho^{\leftarrow}(1-z)) dz \\
&+ \lim_{\epsilon \searrow 0} \int_{z^*(\beta_{inf} + \epsilon)}^1 -\epsilon F_\rho^{\leftarrow}(1-z) dz \\
&> 0
\end{aligned}$$

Then

$$\lim_{\epsilon \searrow 0} g(\beta_{inf} + \epsilon) > 0$$

Therefore, $g(\beta) = 0$ is continuous, strictly decreasing, having positive and negative values on $\beta \in (\beta_{inf}, \beta_{sup})$. Then β^* in Theorem 4, which is the root of $g(\beta) = 0$, always exists and is unique.

A special case is $\limsup_{z \rightarrow 1} \frac{w'(1-z)}{F_\rho^{\leftarrow}(1-z)} = \beta_{sup}$. Under this case, for any $\beta < \beta_{sup}$, $g(\beta) > 0$,

and when $\beta = \beta_{sup}$, $g(\beta) = 0$. Hence, β_{sup} would be the optimal value for β , the root of $\beta^{(*)}$ exists and is unique.

To verify the optimal solution, consider the constraint

$$E[F_{\rho}^{\leftarrow}(1 - Z)\hat{F}_X^{\leftarrow}(Z)] = x_0$$

When $X \sim F_X^{\leftarrow}(\cdot) = b\mathbf{1}_{\{z^*(\beta^*) < z \leq 1\}}$, the constraint can be written into:

$$b \int_{z^*(\beta^*)}^1 F_{\rho}^{\leftarrow}(1 - z) dz = x_0$$

Then,

$$b^* = \frac{x_0}{\int_{z^*(\beta^*)}^1 F_{\rho}^{\leftarrow}(1 - z) dz}$$

The corresponding optimal solution is

$$X^* = b^*\mathbf{1}_{\{z^*(\beta^*) < z \leq 1\}}$$

□

3.4 Optimal Solution to Yaari's Dual Model

Denote $z^*(\beta^*)$ and b^* same as in Theorem 4. Let $c = F_{\rho}^{\leftarrow}(1 - z^*(\beta^*))$.

Theorem 5 *Under Assumption 2, there are two cases as follows.*

Case I: $P[\rho = c] = 0$ i.e. $F_\rho(F_\rho^{\leftarrow}(1 - z^*(\beta^*))) = 1 - z^*(\beta^*)$. Then an optimal solution to Equation 3.1 is

$$X^* = b^* \mathbf{1}_{\{\rho \leq c\}}.$$

Case II: $P[\rho = c] > 0$ i.e. $F_\rho(F_\rho^{\leftarrow}(1 - z^*(\beta^*))) > 1 - z^*(\beta^*)$. Under Assumption 1, an optimal solution to (Equation 3.1) is

$$X^* = b^* \mathbf{1}_{\{[\rho < c] \cup ([\rho = c] \cap [Y \leq \mu^*])\}}$$

where μ^* is determined by

$$P([\rho = c] \cap [Y \leq \mu^*]) = \frac{1}{c}(x_0 - E[\rho b^* \mathbf{1}_{\{\rho < c\}}]),$$

and $Y \in \mathcal{F}_T$ has a continuous c.d.f. according to Assumption 1.

Proof of Theorem 5:

Based on Theorem 4,

$$X^* = b^* \mathbf{1}_{\{z^*(\beta^*) < z \leq 1\}}.$$

Under Case I, $F_\rho(F_\rho^{\leftarrow}(1 - z^*(\beta^*))) = 1 - z^*(\beta^*)$,

$$X^* = b^* \mathbf{1}_{\{F_\rho^{\leftarrow}(1-z) < F_\rho^{\leftarrow}(1-z^*(\beta^*))\}}.$$

According to the equivalence of optimization problems (Equation 3.2) and (Equation 3.6),

$$X^* = b^* \mathbf{1}_{\{\rho < c\}} \in \sigma(\rho) \subset \mathcal{F}_T,$$

where $\sigma(\rho)$ is the smallest σ -field generated by ρ . Therefore, it is also an solution of optimization problem (Equation 3.1).

Under Case II, $F_\rho(F_\rho^{\leftarrow}(1 - z^*(\beta^*))) > 1 - z^*(\beta^*)$, let

$$\begin{aligned} X^* &= b^* (\mathbf{1}_{\{F_\rho^{\leftarrow}(1-z) < F_\rho^{\leftarrow}(1-z^*(\beta^*))\}} \\ &+ b^* \mathbf{1}_{\{[F_\rho^{\leftarrow}(1-z) = F_\rho^{\leftarrow}(1-z^*(\beta^*))] \cap [Y \leq \mu^*]\}} \end{aligned}$$

Under Assumption 1, $Y \in \mathcal{F}_T$ has a continuous *c.d.f.* and μ^* is determined by

$$\begin{aligned} &P([F_\rho^{\leftarrow}(1-z) = F_\rho^{\leftarrow}(1-z^*(\beta^*))] \cap [Y \leq \mu^*]) \\ &= \frac{1}{c} (x_0 - E[F_\rho^{\leftarrow}(1-z) b^* \mathbf{1}_{\{F_\rho^{\leftarrow}(1-z) < F_\rho^{\leftarrow}(1-z^*(\beta^*))\}}]) \end{aligned}$$

which satisfies the constraint

$$E[F_\rho^{\leftarrow}(1-Z) F_{X^*}^{\leftarrow}(Z)] = x_0.$$

According to the equivalence of (Equation 3.2) and (Equation 3.6),

$$X^* = b^* \mathbf{1}_{\{\rho < c \cup ([\rho = c] \cap [Y \leq \mu^*])\}} \in \sigma(\{\rho, Y\}) \subset \mathcal{F}_T$$

and it is also an solution of optimization problem (Equation 3.1). □

CHAPTER 4

CONCLUSION

The quantile approach in the literature can be used to solve portfolio choice models, but highly depends on the prerequisite that the pricing kernel is atomless. In this thesis, for the goal reaching model, we propose an approach based on a generalized Neyman-Pearson Lemma, which can solve problem with a general pricing kernel. For Yaari's dual model, we show that an optimal terminal case flow is nonincreasing with respect to the pricing kernel, which may contain atoms and thus is more general than the atomless case. We develop an algorithm, Search-and-Cut, to find the optimal value of an internal parameter, and the optimal solution as well.

Back tests are done, and results are quite satisfying: the explicit solution based on our results always performs better than the numerical one based on grid search, and as the number of grids increases, the numerical value of the problem converges to our explicit one.

There are some potential applications of the approaches we propose in this thesis. The first one is, they may be used for other portfolio choice models as well, e.g., Lopes' SP/A Theory, Kahneman and Tversky's Prospect Theory.

Secondly, they may be applied on problems which require the atomless assumption when solving by the quantile approaches.

In future, we could consider some extensions for this research, e.g., estimating the pricing

kernels from real financial trading data, and solving the portfolio optimization problem under incomplete market, where pricing kernels are not unique.

CITED LITERATURE

- Browne, S. (1999) Reaching goals by a deadline: digital options and continuous time active portfolio management, *Adv. Appl. Probab.* 31, 551–577
- Browne, S. (2000) Risk-constrained Dynamic Active Portfolio Management, *Manage. Sci.* 46(9), 1188–1199.
- Bellman, R., Dreyfus, S (1962) Applied Dynamic Programming, Princeton University Press, Princeton, NJ
- Bayraktar, E., Huang, Y.J., Song, Q. (2012) Outperforming the market portfolio with a given probability, *Ann. Appl. Probab.* 22, 1465–1494
- Billingsley, P. (1995) Probability and Measure, 3rd edition, John Wiley and Sons, Inc.
- Bellman, R. (1957) Dynamic Programming, Princeton University Press, Princeton, NJ
- Cvitanic, J. (2000) Minimizing expected loss of hedging in incomplete and constrained markets, *SIAM J. Control Optim.* 38, 1050–1066
- Cvitanic, J., Karatzas, I. (1992), Convex Duality in Constrained Portfolio Optimization, *Ann. Appl. Probab.*, 2(4), 767–818

- Cvitanic, J., Karatzas, I. (2001), Generalized Neyman-Pearson Lemma via convex duality, *Bernoulli*, Vol. 7, 79–97
- Dovgosheya, O., Martiob, O., Ryazanova, V., Vuorinenc M. (2006) The Cantor function, *Expositiones Mathematicae* 24, 1–37
- Delbaen, F., Schachermayer, W. (1994) A general version of the fundamental theorem of asset pricing, *Math. Ann.* 300, 463–520
- Delbaen, F., Schachermayer, W. (2006) *The Mathematics of Arbitrage*, Springer, Berlin
- El Karoui, N.E., Quenez, M.C. (1995) Dynamic programming and pricing of contingent claims in an incomplete market, *SIAM J. Control Optim.* 33, 29–66
- Falkner, N., Teschl, G. (2012) On the substitution rule for Lebesgue-Stieltjes integrals, *Expositiones Mathematicae* 30, 412–418
- Fllmer, H., Leukert, P. (1999) Quantile hedging, *Finance Stoch.* 3, 251–273
- Fllmer, H., Schied, A. (2002) Convex measures of risk and trading constraints, *Finance Stoch.* 6, 429–447
- Goll, T., Ruschendorf, L (2001) Minimax and Minimal Distance Martingale Measures and Their Relationship to Portfolio Optimization, *Finance Stochas.* 5(4),

557–581.

Hogg, R., McKean, J., Craig, A. (2005) Introduction to mathematical statistics, *Pearson Education*

Hamada, M., Sherris, M., Van Der Hoek J. (2006): Dynamic Portfolio Allocation, the Dual Theory of Choice and Probability Distortion Functions, *ASTIN Bull.* 36(1), 187–217.

He, X., Zhou, X. (2011) Portfolio choice via quantiles, *Mathematical Finance*, 21(2):203–231

Jarrow, R. (2010) Understanding the risk of leveraged ETFs, *Finance Res. Lett.* 7, 135–139

Jeanblanc, M., Yor, M., Chesney, M. (2009) Mathematical Methods for Financial Markets, *Springer*, London

Jin, H., Zhou, X. (2008) Behavioral portfolio selection in continuous time, *Mathematical Finance*, 18, 385–426

Jin, H., Zhou, X. (2010) Erratum to “Behavioral portfolio selection in continuous time”, *Mathematical Finance*, 20, 521–525

- Kulldorff, M. (1993) Optimal Control of Favorable Games with a Time Limit, *SIAM J. Contrl. Optimizat.* 31(1), 5269
- Krutchenko, R.N., Melnikov, A.V. (2001) Quantile hedging for a jump-diffusion financial market model. In: Kohlmann, M., Tang, S. (eds.) Trends in Mathematics: Workshop of the Mathematical Finance Research Project, Konstanz, Germany, 5–7 October 2000, pp. 215–229. Birkhuser, Basel
- Kramkov, D., Schachermayer, W. (1999) The Asymptotic Elasticity of Utility Functions and Optimal Investment in Incomplete Markets, *Ann. Appl. Probab.* 9(3), 904–950.
- Lehmann, E.L. (1986) Testing Statistical Hypotheses (2nd edition). *John Wiley & Sons, Inc.*, New York.
- Lehmann, E.L., Romano, J.P. (2005) Testing Statistical Hypotheses, 3rd edn, *Springer*, New York
- Leung, T., Song, Q., Yang, J. (2013) Outperformance Portfolio Optimization via the Equivalence of Pure and Randomized Hypothesis Testing, *Finance and Stochastics*, vol 17, 839–870
- Merton, R.C. (1969) Lifetime Portfolio Selection under Uncertainty: The Continuous-time Case, *Rev. Econ. Stat.* , 51(3), 247–257.

- Merton, R.C. (1971) Optimum Consumption and Portfolio Rules in a Continuous-time Model, *J. Econ. Theory* 3, 373–413.
- Ma, J., Yong, J. (1999) Forward-Backward Stochastic Differential Equations and Their Applications. Lecture Notes in Mathematics, vol. 1702. *Springer*, Berlin
- Ma, J., Zhang, J. (2002) Representation theorems for backward stochastic differential equations, *Ann. Appl. Probab.* 12, 1390–1418
- Romano, M., Touzi, N. (1997) Contingent claims and market completeness in a stochastic volatility model, *Math. Finance* 7, 399–410
- Rudloff, B. (2007) Convex hedging in incomplete markets, *Appl. Math. Finance* 14, 437–452
- Rudloff, B., Karatzas, I. (2010) Testing composite hypotheses via convex duality, *Bernoulli* 16, 1224–1239
- Samuelson, P.A. (1969) Lifetime Portfolio Selection by Dynamic Stochastic Programming, *Rev. Econ. Stat.* 51(3), 239–246
- Schied, A. (2004) On the Neyman–Pearson problem for law-invariant risk measures and robust utility functionals, *Ann. Appl. Probab.* 14, 1398–1423

- Schied, A.(2005) Optimal investments for robust utility functionals in complete market models, *Math. Oper. Res.* 30, 750–764
- Sekine, J. (1999) On a robustness of quantile-hedging: complete markets case, *Asia-Pac. Financ. Mark.* 6, 195–201
- Spivak, G., Cvitanic, J. (1999) Maximizing the probability of a perfect hedge, *Ann. Appl. Probab.* 9, 1303–1328
- Yaari, M.E. (1987) The Dual Theory of Choice under Risk, *Econometrica* 55(1), 95–115

VITA

EDUCATION:

Ph.D. Candidate: Mathematics

University of Illinois at Chicago

M.S.: Statistics

University of Chicago, Dec, 2008

B.S.: Financial Engineering

Central University of Finance and Economics, Jun, 2007

WORK EXPERIENCE:

Teaching Assistant, University of Illinois at Chicago, Chicago, IL, 08/2010–12/2013

Senior Analytics Associate, TMX Credit, Evanston, IL, 05/2012–08/2012

Market Research Intern, Blue Cross Blue Shield Association, Chicago, IL, 05/2011–08/2011

Instructor, Bugil Academy, South Korea, 01/2009–07/2010

Research Assistant, University of Chicago, Booth School of Business, Chicago, IL, 06/2008–12/2008

Auditor, PricewaterhouseCoopers, Beijing, China, 05/2007–08/2007

AWARDS:

International Interdisciplinary Contest in Modeling (ICM)- Honorable Mention (2007)

International Mathematical Contest in Modeling (MCM)- Honorable Mention (2006)

SEMINAR TALKS:

Portfolio Choice with General Pricing Kernel: An Application of a Generalized Neyman-Pearson

Lemma, Statistics Graduate Seminar, University of Illinois at Chicago. 10/2013

Cluster Analysis in Consumer Segmentation, Statistics Research Seminar, University of Illinois at Chicago. 09/2011

PUBLICATIONS:

Z.Zhang, Q.Song, J. Yang, Portfolio Choice with General Pricing Kernel (2014), in preparation

K. Tam, S. Li, Z. Zhang, C. Yu, Foreign investment in China and Qualified Foreign Institutional Investor (QFII)(2010), *Asian Business and Management*, Vol. 9, 425–448

C.Li, Z. Zhang, C. Zhang, The Independent Innovation in Chinese Economy (2007), *Journal of Contemporary Economics* (ISSN 1007-9378), 2007–06S

PROFESSIONAL AFFILIATION:

American Statistical Association (ASA)

Institute of Mathematical Statistics (IMS)