OPTIMUM ASSET ALLOCATION WITH BEHAVIORAL UTILITIES : A PLAN FOR ACQUIRING AND CONSUMING RETIREMENT FUNDS

A DISSERTATION
SUBMITTED TO THE PROGRAM IN SCIENTIFIC COMPUTING & COMPUTATIONAL MATHEMATICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Aparna Gupta

May 2000

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Walter Murray
(Principal Co-Advisor)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William F. Sharpe
(Principal Co-Advisor)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Gene H. Golub

Approved for the University Committee on Graduate Studies:

Abstract

The question of optimal strategic asset allocation for investors with behavioral utilities saving for retirement is addressed. To date this problem has been addressed assuming that an investor is rational in the sense when making investment decisions the preference relation of the investor satisfies all the axioms of choice. Research in behavioral science indicates that not all the investment related decisions of an average person satisfy the axioms of choice. In this thesis a broader class of smooth utilities is considered that includes those satisfying the axioms of choice, but allows some to be violated.

We also discuss the optimal consumption problem of how "best" a retiree should invest and consume her retirement funds. We show that mathematically this problem and that of how best to save for retirement can be formulated as large-scale nonlinear constrained optimization problems. Although both can be solved using the same algorithm, there are important differences in how the two problems are modeled.

Acknowledgements

I wish to sincerely thank my advisors, Professor Walter Murray and Professor William F. Sharpe, for their constant inspiration and guidance. I owe most to Professor Murray for his insights into the problem, and his immense patience. Despite being a very busy person, he most willingly made room for discussions with me. Most pleasurable aspect of which was that my discussions with him did not stay restricted to the research problem, I got much valuable professional advice from him, discussed scores of topics such as, Indian food, the amazing variety of English accents, etc., and for the last one year, his most adorable daughter Sofia.

Professor Sharpe gave me not only his valuable time, but will be a constant source of inspiration throughout my life. I am truly grateful to him for making time for discussions with me in his extremely busy schedule, helping me structure the problem addressed in this thesis and giving key insights into the financial concepts of the problem. An additional attraction of visiting him was to admire the beautiful paintings by his wife in his office. I am also very thankful to him for his guidance in chalking out my future career plans.

Many thanks to Professor Gene Golub for not only taking the time to serve on my reading committee, but also to provide me the opportunity of benefiting from his ocean of experience. As for many of his students, he will always be a guiding force for me. I also wish to thank Professor Bernard Roth, Professor Steve Grenadier and Professor Manju Puri for serving on my oral committee, and especially Professor Puri

Contents

Abstract **iv**

Acknowledgements **v**

1 Introduction **1**

1.1 The Investment Problem 2

1.1.1 Utility Theory 2

1.1.2 Normative vs. Descriptive(Behavioral) Choice 3

1.1.3 The Portfolio Problem 5

1.1.4 Overview of Optimization Methods 6

1.2 The Consumption Problem 8

1.2.1 Basic Issues 8

1.2.2 Overview of Optimization Methods 9

1.2.3 Contrast with the Investment Problem 10

2 Modeling the Problem **12**

2.1 Asset-price Dynamics - Binomial Tree Set-up 13

2.2 Optimal Asset Allocation Problem 15

3 Solving the Problem **17**

3.1 Introduction 17

3.2 Computational Challenges 18
3.3 Optimization Techniques 21
3.3.1 Univariate Optimization 21
3.3.2 Multivariate Optimization 22
3.3.3 Quasi-Newton Algorithm 23
3.3.4 Sequential Quadratic Programming Algorithm 24
3.3.5 Application to the Investment Problem 26
3.4 Modeling Variations 27
3.4.1 Wealth as Variables 28
3.4.2 Non-Uniform Time Periods 30
3.4.3 Intermediate Contributions 31
3.5 Comparison of Computational Effort 31
3.5.1 Dynamic Programming 31
3.5.2 Stochastic Optimal Control 33
3.5.3 Stochastic Optimization 34

4 Application to Investment Problem **36**
4.1 Introduction - Utility Theory 36
4.1.1 Descriptive, Normative, and Prescriptive Decision Making . . 39
4.2 Behavioral Utilities - Challenge 42
4.3 Optimal Results for Behavioral Utilities 43
4.3.1 Loss Aversion Utility 44
4.3.2 Piece-wise Linear Utility 46
4.3.3 α - t Semivariance Utility 51
4.3.4 Probability of Reaching a Goal 53
4.3.5 Variant of Probability of Reaching a Goal 57
4.4 Verification of Results from Standard Utilities 58

4.5 Comparisons and Conclusions 62
4.6 Contribution Schemes 64
4.6.1 Problem Formulation for Contributions 65
4.6.2 Results 66

5 Consumption Problem 68
5.1 Introduction 68
5.2 Problem Set-up 71
5.3 Computational Issues 74
5.4 Results 75
5.4.1 Negative Exponential utility 77
5.5 Conclusions 83

6 An Alternate Approach 84
6.1 Introduction 84
6.2 The Asset Allocation Problem using a Behavioral Utility 85
6.2.1 Discovering the Structure in Solution 88
6.2.2 The Structure 89
6.2.3 The Central Idea : Use the Structure 90
6.3 Applying Parametric Optimization 91
6.4 Sensitivity to Asset-Dynamics Parameters 94
6.5 Conclusion 95

7 Future Work 99
7.1 Asset-return Dynamics Models 100
7.2 Solution Technique 101
7.3 Investment Problem 102
7.4 Consumption Problem 103

7.5 Improving Usability of Tool 103

A Sequential Quadratic Programming Algorithms **105**

A.1 Problem 105

A.2 Introduction 106

A.3 Merit Function and Line Search 107

A.4 Multiplier Estimates 108

A.5 Quasi-Newton Update of the Hessian 109

A.6 Active-set Approach 109

A.7 A Simple SQP Algorithm 110

B ADIFOR **111**

C NPSOL Package **112**

D SNOPT Package **114**

Bibliography **116**

List of Tables

List of Figures

2.1 One period transition rule for binomial-tree model 14
2.2 Ordering of nodes in the binomial tree 16

4.1 Plot of loss-aversion utility against wealth, W 45
4.2 Optimal Stocks investment weights for a 9 periods problem using Loss aversion utility 47
4.3 Optimal Stocks investment weights for a 9 periods problem using Loss aversion utility, plotted by wealth levels at nodes for each time period 47
4.4 Optimal Stocks investment weights for a 9 periods problem using Loss aversion utility, with different reference points (continuous for 8% yearly return, dashed line for 5% yearly return) 48
4.5 Optimal Stocks investment weights for a 9 periods problem using Loss aversion utility, with different reference points (continuous for 8% yearly return, dashed line for starting wealth) 48
4.6 Optimal Stocks investment weights for a 9 periods problem using Loss aversion utility, with non-uniform time periods (4 quarters, 2 half year, 1 one year, 1 two year and 1 four year periods) 49
4.7 Plot of piece-wise linear utility against wealth 49
4.8 Optimal Stocks investment weights for a 9 periods problem using Piece-wise Linear utility, plotted by wealth levels at nodes for each time period 51

4.9 Plot of α-t semivariance utility against wealth 52
4.10 Optimal Stocks weights for a 9 periods problem using α-t semivariance utility, plotted by wealth levels at each time period 52
4.11 Optimal Stocks investment weights for a 9 periods problem using α-t semivariance utility, where slope is zero above the reference 53
4.12 Optimal Stocks weights for a 7 periods problem using Probability of Goal utility, plotted by wealth levels for each time period 55
4.13 Histogram for wealth at the planning horizon, $T = 7$ 56
4.14 Number of nodes at which target wealth is achieved for a range of target wealth values 56
4.15 Optimal Stocks investment weights for a 9 periods problem using Negative Exponential utility, plotted by wealth levels at nodes for each time period 60
4.16 Plot of quadratic utility against wealth 61
4.17 Optimal Stocks investment weights for a 9 periods problem using quadratic utility, plotted by wealth levels at nodes for each time period 62
4.18 Plot of Kritzman's combination utility 63
4.19 Optimal Stocks investment weights for a 9 periods problem using Kritzman's combination utility, plotted by wealth levels at nodes for each time period 63
4.20 Optimal Stocks investment weights for a 6 periods problem using loss aversion utility, plotted by wealth levels at nodes for each time period 67

5.1 Negative exponential utility plot for a 8 periods problem 77
5.2 Optimal consumption decisions for a 8 periods problem for negative exponential utility 78

5.3 Optimal investment decisions for a 8 periods problem for negative exponential utility 79
5.4 Consumption for each time, t above the problem bound $\rho\overline{C}_t$ and above the reference lower bound $\overline{C}_t$ 79
5.5 Minimum, Maximum and Mean Consumption level for each time, t. . 80
5.6 Number of nodes above the reference lower bound $\overline{C}_t$ for a range of problem bounds $\rho\overline{C}_t$, ρ ranging from 0.1 to 1.0 81
5.7 Plot of utility for consumption for 4 levels for the reference lower bound relaxation parameter, relative to the utility for consumption at the reference lower bound level 82

6.1 Prospect theory loss-aversion utility of wealth 87
6.2 9 Periods plot of optimal allocation in the risky asset against nodes in the binomial tree 88
6.3 9 Periods plot of optimal allocation in the risky asset against wealth levels for each time period 89
6.4 Optimal stock allocation for a one period and a four period problem as a function of wealth and time horizon 92
6.5 Comparison of the parametric solution with the true solution for a four periods problem 93
6.6 Sensitivity analysis: Variation in true solution due to 1% misspecification in the stock mean parameter, μ. The approximate parametric solution falls well within the error due to 1% misspecification of μ. . . 96
6.7 Variation in true solution due to 1% misspecification in the stock volatility parameter, σ. The approximate parametric solution falls well within the error due to 1% misspecification of σ. 97

Chapter 1

Introduction

The investment problem of how an investor should invest for meeting long-term financial targets and goals is viewed as an important problem and has been the subject of considerable research. In such studies there are assumptions made about the planner's preferences, assets available for investment and rules for the asset-price evolution. Here a behavioral aspect of the problem is also addressed. We obtain the optimal investment strategy for investors with a class of behavioral utilities. The traditional utilities can also be solved for with this approach. The main emphasis of this work is on an individual's retirement planning problem.

We also address the problem of optimal consumption and investment of one's retirement funds during retirement. This problem has received less attention, but in our view it is of equal importance; with increasing life-expectancy the number of retirement years are approaching the number of pre-retirement investment years. Moreover, it can be expected that a retiree will take on much more active role than they could while saving for retirement. This is not only because they are likely to have much more time on their hands, but also because consumption decisions after retirement have a much more immediate impact on their quality of lives. The consumptions problem is complementary to the long-term investment problem, and the mathematical formulation of the problems are also similar.

1.1 The Investment Problem

We begin with exploring the modeling considerations for the retirement planning problem, formalizing the decision making criterion and framing the optimal choice problem as an optimization problem.

1.1.1 Utility Theory

An individual faced with a decision chooses from a set of alternatives. Let X denote this set of alternatives. Among all elements of X the decision-maker weighs the merit and de-merits of each option against others and attempts to select the option that suits her best. This requires a precise definition of "best."

In more rigorous terms, we need to define an ordering for the set of alternatives that makes it possible to compare every element of the set with every other element. We denote this ordering by '$\succeq$', then if $x, y \in X$, $x \succeq y$ will imply that the decision-maker prefers x to y, and $x \succ y$ would imply x is strictly preferred to y.

The preference relation defined above is said to be rational if it satisfies the *axioms of choice*. These axioms are reflexivity, completeness and transitivity and are defined as follows (see Varian [56], MasColell [38]] :

Reflexivity For all $x \in X$ we have $x \succeq x$.

Completeness For all $x, y \in X$ we have either $x \succeq y$, $y \succeq x$ or both.

Transitivity For all $x, y, z \in X$, if $x \succeq y$, $y \succeq z$ then $x \succeq z$.

Although at first sight these axioms may look quite reasonable, it is worth taking a closer look. Imposition of these axioms has enormous implications in practical terms. Reflexivity is the weakest and most acceptable axiom. The completeness axiom implies that the decision-maker has done the introspection of all the alternatives, however far they may be from one's realm of common experience. It is easy to imagine that this may need serious work and reflection on one's preferences and may in some cases be almost impossible. The transitivity axiom, on the other hand, implies that in a sequence of pairwise choices there is no possibility of cycles, which is to say that however the options are framed or presented, the decision-maker is mature enough to be able to rank all of them in an order that contains no cycles. Besides the above three axioms of choice, preference relations are assumed to satisfy additional axioms like monotonicity, continuity and convexity of preferences.

1.1.2 Normative vs. Descriptive(Behavioral) Choice

The axioms of choice listed in the previous section form the normative theory of choice. A decision-maker who satisfies these axioms is said to be rational. It is possible to represent the preferences in terms of a utility function, which is a map from the set of alternatives to the real line. Further assumptions of continuity, monotonicity and convexity of the preference relation imply that the utility function obtained is a continuous, increasing and concave function. Most research in economics and finance has developed elegant and elaborate theory on the basis of the assumption that the decision-makers' preferences satisfy these axioms.

As was observed in the previous section, it may be a difficult proposition for an average decision-maker to satisfy all the axioms of choice. Following this thought, many behavioral scientists and economists have attempted to understand the decision-making processes among ordinary people through experiments and studies of various kinds (see Wright [60], Bell [2]). These attempts also try to quantify the choice making process, but try to be parsimonious in their assumptions. They constitute what is called the descriptive theory of choice. The five major phenomena of choice that violate the standard model of normative theory are listed as : Framing effects, Nonlinear preferences, Source dependence, Risk seeking and Loss aversion. These have been confirmed in a number of experiments, with both real and hypothetical payoff and are defined as follows in Tversky & Kahnemann [55, 54], Tversky [53].

Framing effects Lack of description invariance implying that variations in the framing of options yield systematically different preferences.

Nonlinear preferences The expectation principle of utility theory states that utility of risky prospects is linear in outcome probabilities. However, experimental evidence indicates that people tend to transform probabilities nonlinearly, overweighing small probabilities and underweighing moderate and high ones.

Source dependence Willingness to bet on an uncertain event depends not only on the degree of uncertainty but also on its source.

Risk Seeking As opposed to the generally assumed risk aversion in economic analysis, in certain situations people prefer more risk to less. For instance, people prefer a small probability of winning a large prize over the expected value of that prospect.

Loss Aversion Carriers of value are gains and losses defined relative to a reference point. The losses loom larger than the gains, that is, an amount of loss elicits more "unhappiness" than the same amount of gain elicits "happiness."

Behavioral finance is a fast growing field of research. In this framework researchers have made an attempt to understand and explain the investment style and preferences of individual investors as well as other market characteristics from a descriptive stand point. It is hoped that these descriptive studies will help bridge the gap between what is rational and perfect and what is not. Bridging this gap will also help 'educate' people to make more rational decisions.

1.1.3 The Portfolio Problem

The portfolio problem is a problem involving choice. The choice set typically contains elements such as how much to invest, what to invest in, what point in one's life-cycle should one invest in what assets, how should one invest in order to have maximum tax benefits over the life-cycle, etc. An investor saves from current earnings for future consumption and invests the savings in investment vehicles, broadly classified into asset classes like stocks, bonds, gold, real estate, etc. The investor evaluates the performance of the investments using some criterion; we call this the investor's objective. Let us denote the investor's objective by a function, $U(.)$; this can be a function of returns or wealth or any other aspect of the investment the investor values. Moreover, since the future returns of the asset classes is uncertain, the wealth one will have at future times is uncertain; in this case the objective is taken to be the expected value of the function, $U(.)$. The portfolio problem may be described in broad terms as determining the set of decisions to maximize this objective function $E[U(.)]$. Mathematically this may be written as

$$\max_{\mathbf{X_t}, C_t, t=0,T-1} \sum_{t=0}^{T} E[U_t(W_t, C_t, \mathbf{R_t}, ...)],$$

where T is the planning horizon, W_t is the wealth at time t, C_t is the consumption at time t, $\mathbf{R_t}$ is a vector of returns on the asset classes. $U_t(.)$ is the utility function at time t for Wealth, Return on assets, Consumption or any other properties at each decision point t. The vector $\mathbf{X_t}$ is the wealth invested in each asset class at time t; i.e., if there are N asset classes, $\mathbf{X_t}$ will be a vector of size N. If instead of a discrete-time problem this were a continuous-time problem, the $\sum U(.)$ would be replaced by an $\int_0^T U(.)$.

The problem as stated above is in its most general form, other than assuming time-additivity for the objective function. In order to solve the problem, it is simplified by making assumptions about the investor's preferences; for instance, the utility may be assumed to be a function of only wealth and consumption, or the investor may not consume prior to the planning horizon, T. There are also bounds and constraints imposed on these variables to make them physically feasible; for instance, you may not be allowed to consume more than the wealth you have at any time.

1.1.4 Overview of Optimization Methods

Most commonly the optimal investment problem is stated as

$$\max_{X_t, C_t, t=0,T-1} \sum_{t=0}^{T} E[U_t(W_t, C_t)].$$

The problem has been solved by several researchers under different assumptions about planning horizon, utility function, asset classes and their price dynamics. In this section we give a brief overview of this work.

Among the early works in the field, Markowitz [36] solved a one period portfolio selection problem for an investor who jointly maximizes the mean of the portfolio return and minimizes its variance. Such preference is obtained when either the investor has a quadratic utility or the assets-returns are normally distributed. The problem becomes one of solving a quadratic program in the decision variables, which are the investment weights for all the assets or asset-classes. Samuelson [48] poses the problem in a multi-period framework and obtains the optimal portfolio selection strategy using stochastic dynamic programming. The problem is solved for the logarithmic and power utilities (constant relative risk aversion utilities) using the dynamic programming backward recursion. Merton [39, 40] supplemented and extended Samuelson [48] by posing the problem as an optimal stochastic control problem in continuous time. Closed form solutions were obtained for constant relative risk aversion utilities and utilities with positive and hyperbolic absolute risk aversion by solving for the value function in the Hamilton-Jacobi-Bellman equation.

Among the more recent works, Brennan et al [9] model the rate of return on the portfolio to be governed by a set of state variables. The optimal portfolio choice problem is stated as an optimal stochastic control problem. The nonlinear partial differential equation from the Hamilton-Jacobi-Bellman (HJB) equation for the value function is solved using implicit finite difference approximation. Constant relative risk aversion utilities are assumed. Karatzas et al [27] solve the optimal portfolio choice problem in continuous-time as an optimal control problem and obtain a closed form solution of the HJB equation for general strictly increasing and concave utilities. Duffie and Zariphopoulou [15] generalize the above by introducing a stochastic income process in the problem that is undiversifiable. They obtained a weak solution to the HJB equation, referred to as a viscosity solution.

In Maranas et al [35] the Markowitz's mean-variance model is extended to a multi-period framework. Discrete time-steps and scenarios are considered and the optimal "dynamically balanced" strategy is obtained by maximizing a linear combination of mean and variance of return, with the coefficient of variance being negative. The resulting problem is a non-convex nonlinear problem, which is solved using a global optimization algorithm. Their algorithm is based on a convex lower bounding of the objective and successive refinement of converging lower and upper bounds by means of a standard branch-and-bound procedure. Birge [6], Dantzig and Infanger [13], Infanger [25] approach the problem in the stochastic optimization framework. The problem is posed as a large linear program and solved using decomposition techniques, such as Bender's decomposition. Shoven [50] considers an allocation-location problem of wealth in pension and savings accounts. The major issues addressed are the optimal allocation of wealth between risky and riskfree assets in the pension funds and the optimal location of these funds in order to have maximum tax-benefits. Shoven performed his study using Monte Carlo simulation using data from the past seventy years for parameters estimation.

1.2 The Consumption Problem

The consumption problem is also a problem involving choice. Most researchers have studied this problem in an integrated fashion with the investment problem. We choose to study it separately since in the retirement planning framework the two stages, namely the time in life when one plans one's retirement and the time when one is retired have very different characteristics, concerns and preferences. A key difference between pre-retirement and post-retirement decisions is that post-retirement decisions have a much more immediate impact.

1.2.1 Basic Issues

Achieving retirement goals requires a well-defined plan. The most important considerations for this planning would be

- retirement age,
- lifestyle and related expenses,
- life expectancy,
- sources of income at retirement,
- marital status,
- and bequest.

Clearly, a lot depends on how early and how well an individual plans for one' retirement and accumulates the retirement capital. We assume that the planner has done the best in her capacity to accumulate the retirement funds, and address how the individual should consume the funds year by year after retirement and how the remaining funds should be invested in order to keep the best flow of income during the retirement years.

Modeling life expectancy is a critical aspect of the problem. At this stage of life the uncertainty related to conditions of health and life is at the largest and modeling of the problem should take this into account. Most other considerations listed above for post-retirement planning can be taken as parameters of the model to an extent. For instance, marital status can be accommodated by considering larger consumption volumes; similar adjustment could be made for life-style and expenses. The problem is formulated, similar to the investment problem, as

$$\max_{X_t, C_t, t=0,\tilde{T}-1} \sum_{t=0}^{\tilde{T}} E[U_t(W_t, C_t)],$$

where C_t is the consumption at time, t and X_t is the investment decision. One critical difference in this problem is that the planning horizon, $\tilde{T}$ is not known with certainty. One option for dealing with this uncertainty is to use the expected life-expectancy, $\overline{T}$ for an individual in place of the random value, $\tilde{T}$ and use other modeling techniques to accommodate some of the uncertainty related with $\tilde{T}$.

1.2.2 Overview of Optimization Methods

Most research work on this problem has been in an integrated framework along with the investment problem. In this section we give a brief overview of these approaches.

Bodie, Merton and Samuelson [8] consider a life-cycle model with an optimal investment choice, choice of labor and consumption over the life-time of the planner. The problem is posed as an optimal control problem in continuous time and solved by assuming the wages are perfectly hedged. Relations are drawn between the optimal investment choice and the human capital of the planner over the planning period. Campbell and Viceira [10] consider an infinitely lived investor with the Epstein-Zin-Weil utility (a recursive generalization of the power utility) and solve the investor's optimal consumption/investment problem. They replace the first-order conditions (Euler equation) and budget constraints of the problem with approximate equations and explore an analytical solution for the approximate problem. They analyze trends like optimal growth rate of consumption, optimal consumption-wealth ratio, etc. Viceira [57] extends Campbell-Viceira by including an undiversifiable labor income process to the problem. Investors are infinitely lived but have a certain probability, π_r of transition to a retired state. After reaching a retired state investors are assumed to stay retired thereafter. Power utility for consumption is assumed and the solution technique is similar to Campbell-Viceira. Analysis is done of the effect of retirement on the portfolio choice.

1.2.3 Contrast with the Investment Problem

With some basic understanding of both the investment and the consumption problem, it is now possible to compare and contrast the characteristics of the two problems. Contrast in the property of the planning horizon for the two problems stands out. While the planning horizon, or the time when one plans to retire, is known with a substantial degree of certainty in the investment problem, in the consumption problem a challenging aspect of the problem is to model the uncertainty related with the planning horizon. It is not only the horizon that is uncertain, but also the state and cost of health related expenses needed by the planner along the planning period that is not predictable. The state of a person's health impacts the person's utility.

In the investment problem the utility is essentially independent of time, since utility is only for the final wealth, i.e., the wealth one is able to accumulate in one's retirement fund by the time one retires, while in the consumption problem practically all the utility comes from the consumption of wealth along the planning period and very little, if any, utility comes from what is left at the end of the planning horizon. In the investment problem the place of the consumption variables is taken by the intermediate contribution variables, which is the yearly contribution one makes to the retirement fund. Typically these contribution levels are not variables, since most often a person invests the maximum amount allowed free of taxation. Further, the consumption levels in the consumption problem are much higher than the periodic contributions one makes in the investment problem. In the consumption problem the retiree can greatly benefit from monitoring the consumptions, and may wish to exercise some discretion on immediate consumption. The asset-classes available for investment and their price-dynamics remain the same; therefore the investment variables are the same in both problems.

Chapter 2

Modeling the Problem

An investment portfolio's growth is driven by how much we save and invest, what we invest in, how long we invest, what return we earn and how much of that return we surrender to investment costs (transaction costs) and to taxes. As an individual investor we usually have a great amount of control over every factor listed above, except one, the asset-returns.

The returns we get on our investments are uncertain. The past performances may be indicative of the future performance, but in no way are fully predictive. This fact makes modeling the asset-return dynamics particularly challenging. In this chapter we lay out the model we use for asset-return dynamics and formulate the optimal investment/consumption problem in terms of it.

The models one comes across in literature for asset-return or price dynamics can be classified in two groups, continuous-time models and discrete-time models. In continuous time, asset-price dynamics are modeled as being governed by a stochastic differential equation, which in turn is driven by an underlying brownian motion or a set of factors that are driven by brownian motion. Various additional features such as mean-reversion and stochastic volatility are introduced to the process to make it more realistic. With all these features these models have had varying degrees of success in their applications; however, they become quite complicated and analytical solutions may not be possible, especially for more general classes of utilities or assumptions. (See Duffie [15], Lamberton [30] for more details on continuous-time models.)

In discrete-time models the asset-return dynamics are primarily governed by a transition rule, i.e., a rule that dictates how the price or return changes from one period to the other. A discrete-time model could in some cases be considered to be a time-discretization of a continuous-time stochastic differential equation model obtained by discretizing the brownian motion or other sources of randomness therein. There may not always be a one-to-one map for all discrete-time models to a continuous-time model, since this depends on whether or not there is a limit of the discrete-time transition rule as the time-grid size shrinks to zero.

Among the discrete-time models it is usual to consider discrete-time, discrete-space asset-return models with a finite number of states. These models are often represented by a "tree." If each state at each time is depicted by a node, an arc can be drawn from this node to every node at the next time period to which transition is possible. The entire set of nodes and arcs so drawn for a finite number of time-periods has an appearance of a tree with its branches.

2.1 Asset-price Dynamics - Binomial Tree Set-up

We have chosen to work with a discrete-time, discrete-space asset-return dynamics model for our problem. At each point of time in this model, transition is possible to two states for asset-returns – we call one of the states an "up" and the other "down." The "up" refers to a return greater than 1, or in other words, the price of the asset goes up, and the "down" refers to the return being less than 1. At first sight this may look like a simplistic approach, but this model has enjoyed enormous success due to its 1) ease of use and 2) its capacity to capture important characteristics of the problem.

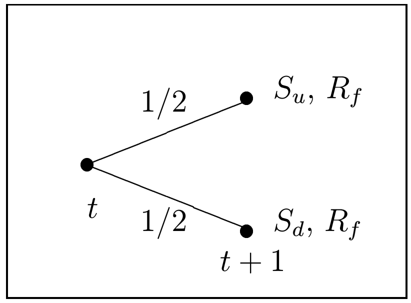

Figure 2.1: One period transition rule for binomial-tree model

We consider it to be equally probable for the returns to go up or down. Due to the two-way branching at each node, this model is commonly known as the binomial tree model (Sharpe [49]). It can be considered to be a version of the simple symmetric random walk, the limit of which, as the size of time-interval shrinks to zero, is the brownian motion. This further entails that the asset-price process dictated by the binomial tree model is a geometric brownian motion in the limit (see Luenberger [32], Ross [46] for details).

Throughout this work we consider two asset-classes, a risky one and a riskfree one. At each point of time the risky asset's return goes up or down with a probability 1/2, whereas the return of the riskfree asset stays a constant, hence the name "riskfree." Let us denote the return of the risky asset on the "up" by S_u and S_d on the "down." Let the return of the riskfree asset be R_f. If one invests W_s dollars in the risky asset and W_b in the riskfree asset in a period, the wealth one would have in the next period will be $W_sS_u + W_bR_f$ if an "up" occurs and $W_sS_d + W_bR_f$ if a "down" occurs. See figure (2.1) for a pictorial description of the binomial tree model.

2.2 Optimal Asset Allocation Problem

Based on the binomial-tree model described in the previous section, the optimal asset allocation problem can be formulated as follows. Given a time-horizon, T for the planning problem, at each time point t lying in $\{0, 1, ..., T-1\}$, a decision regarding the amount of wealth to invest in the two asset-classes needs to be made. Given the investor knows the state of the world at $t = 0$, there are $2^T - 1$ number of nodes between $t = 0$ and $t = T - 1$. Each node represents a decision variable, namely what fraction of wealth to invest in the risky asset and the riskfree asset. The nodes are ordered as indicated in figure (2.2). Let each decision variable be denoted by $X_{t,n}$, where t is the time period and n is the number of the node in the ordering. The investor invests an initial amount of wealth, W_0 at $t = 0$. There is also an amount of wealth associated with each node in the tree, denoted by $W_{t,n}$ by the same convention as for the investment decision variable. This is the wealth the investor has at that node given that she starts with a wealth W_0 at $t = 0$ and invests according to a particular set of values for the decision variables, $X_{t,n}$. The relation between the decision variables $X_{t,n}$ and state variables $W_{t,n}$, S_u, S_d and R_f can be written as follows:

$$X_{t,n_1} W_{t,n_1} S_u + (1 - X_{t,n_1}) W_{t,n_1} R_f = W_{t+1,n_2} \quad (2.1)$$

$$X_{t,n_1} W_{t,n_1} S_d + (1 - X_{t,n_1}) W_{t,n_1} R_f = W_{t+1,n_3}, \quad (2.2)$$

where equation (2.1) is for an "up" and equation (2.2) for a "down." $X_{t,n}$ is taken to be the fraction of one's wealth at time t and node n one invests in the risky asset. This makes $(1 - X_{t,n})$ the fraction of wealth invested in the riskfree asset at time t and node n, assuming that the investor invests all her wealth at each node. The relation between the nodes n_1, n_2 and n_3 is indicated in figure (2.2).

Utility is from the wealth at the planning horizon, since the investor does not

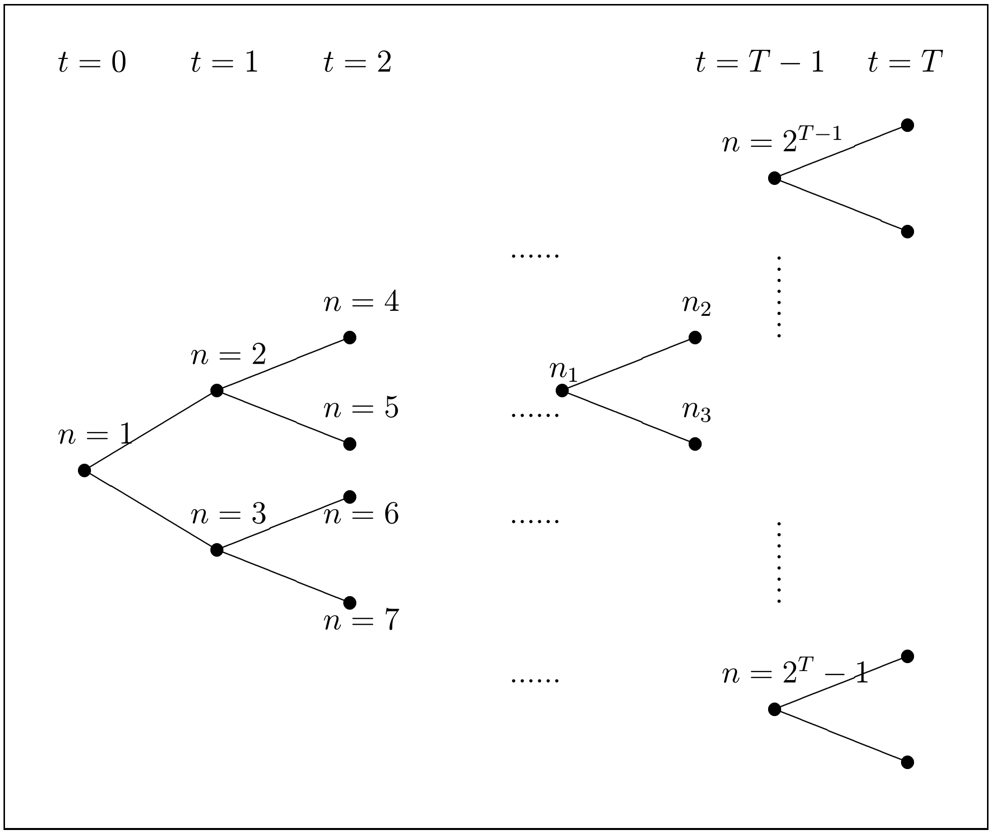

Figure 2.2: Ordering of nodes in the binomial tree

consume prior to the planning horizon. The optimal asset allocation problem is written as follows:

$$\max_{X_{t,n},t=0,T-1,n=1,N} \sum_{m=q}^{N} [pU(X_{T-1,m}W_{T-1,m}S_u + (1 - X_{T-1,m})W_{T-1,m}R_f) + pU(X_{T-1,m}W_{T-1,m}S_d + (1 - X_{T-1,m})W_{T-1,m}R_f)],$$

where p is the probability of each node at the planning horizon. Index m ranges over nodes one period prior to planning horizon, which are nodes 2^{T-1} to $2^T - 1$ in figure (2.2).

Chapter 3

Solving the Problem

3.1 Introduction

In the long-term financial planning problem, also known as the strategic asset allocation problem, the term “long” is a relative one. In some set-ups it may refer to a year long planning period, whereas in others a 30 year period would be considered appropriate. From a retirement planning perspective it seems reasonable to consider a planning period ranging from 10 to 30 years. During this planning period, if the planner makes yearly investment decisions, there are 10 to 30 decisions to make. This may look like a relatively easy task, but unfortunately, there is more to the problem. Future investment decisions depend on future economic scenarios, which are uncertain. Therefore, what one really needs to know is one’s decision for every possible economic scenario in future. It is easily imaginable that attempting to select a decision for every economic scenario for a period of 10 to 30 years is by no means an easy task.

In the previous chapter a model for possible future economic scenarios was presented. The scenarios were defined in terms of returns on the two asset-classes available for investment. In this chapter we first sketch the computational challenges associated with a scenario based decision approach and how they relate specifically to our model for the problem. Before describing our approach for addressing these challenges, a brief overview of general optimization techniques is given. This is followed by a description of methods used for finding the optimal solution for our problem. Finally, we present some modeling variations that facilitate computation of the optimal solution, and give a comparison of our computational effort with other approaches.

3.2 Computational Challenges

In the previous chapter we described the binomial-tree model for evolution of asset returns over time. In this model, in each time period asset returns have a transition to two possible states; we called one of the states an "up" and the other a "down." The decision variables representing the fraction of wealth invested in the risky asset at each node of the tree is denoted by $X_{t,n}$. Assuming that the investor invests all the wealth at each node, $(1-X_{t,n})$ is the fraction invested in the riskfree asset. Evolution of wealth, starting with wealth W_0 at $t = 0$, was given by equations

$$X_{t,n_1}W_{t,n_1}S_u + (1 - X_{t,n_1})W_{t,n_1}R_f = W_{t+1,n_2} \tag{3.1}$$

$$X_{t,n_1}W_{t,n_1}S_d + (1 - X_{t,n_1})W_{t,n_1}R_f = W_{t+1,n_3}, \tag{3.2}$$

where equation(3.1) is for an "up" and equation(3.2) for a "down." Nodes n_2 and n_3 at time $t + 1$ are the nodes connected to node n_1 at time t. Therefore, at a typical node at the planning horizon, T, wealth is given by

$$X_{T-1,m}W_{T-1,m}R_r + (1 - X_{T-1,m})W_{T-1,m}R_f, \quad (3.3)$$

where R_r takes the value S_u or S_d depending on whether the transition from $T-1$ to T is an "up" or a "down." Assuming that the planner derives utility from the wealth at the planning horizon, the optimum asset allocation problem can be written as:

$$\begin{aligned} \max_{X_{t,n}, t=0,T-1,n=1,N} \quad & \sum_{m=q}^{N}[pU(W^u_{T,m}) + pU(W^d_{T,m})], \\ \ni \quad & 0 \leq X_{t,n} \leq 1, \forall t, n, \\ \text{where} \quad W^u_{T,m} \quad = \quad & X_{T-1,m}W_{T-1,m}S_u + (1 - X_{T-1,m})W_{T-1,m}R_f, \\ W^d_{T,m} \quad = \quad & X_{T-1,m}W_{T-1,m}S_d + (1 - X_{T-1,m})W_{T-1,m}R_f, \end{aligned}$$

where p is the probability of each node at the planning horizon and $U(.)$ is the investor's utility for wealth at the planning horizon. Index m ranges over all the outcomes at one period prior to the planning horizon. In our problem formulation we permit no short sale, i.e., borrowing of assets is not allowed. This restriction translates to the above bounds on the decision variables. The restriction reflects regulatory constraints on positions the planner may take on the asset classes. In principle when short-selling is allowed, the principal issue of concern is that the investor may default on the borrowed assets. In other words, it may be necessary to impose constraints on the wealth at each node of the tree to stay positive. In practice, individual investors invest under restrictions on the positions they can take in terms of short-sale or borrowing of assets.

The investment problem formulated as an optimization problem is a nonlinear optimization problem with bounds on variables. Major sources of complexity in the above optimization problem are 1) size of the problem, 2) nonlinearity of the objective and 3) the possibility of a less well-behaved objective. People typically need to plan for retirement 20 to 30 years prior to retirement. This long planning period results in a very large problem size, for instance, in a 20 period binomial tree there are $1,048,575$ ($= 2^{20} - 1$) nodes, hence as many decision variables. An optimization problem with this many variables is indeed a large problem.

Nonlinearity of the objective is observed when the wealth term, $W_{T-1,m}$ in the objective is replaced using equation (3.1) or (3.2) in terms of W_{T-2,m_1}, where m_1 at $T-2$ is a node connected to node m at $T-1$ in the binomial tree. This process can be recursively continued by repeatedly using equation (3.1) or (3.2) and expressing W_{T-2,m_1} in terms of W_{T-3,m_2}, where m_2 at $T-3$ is a node connected to node m_1 at $T-2$ in the binomial tree, and so on. At the end of this substitution process each term at which the utility function $U(.)$ is evaluated is a T-degree polynomial of appropriate $X_{t,n}$'s. Nonlinearity in the utility function further enhances nonlinearity of the objective.

In Chapter 1 a brief description of different choice criteria was given, differentiating the normative from the descriptive or behavioral class. Normative choice is defined in terms of some regularizing principles that are called the axioms of choice. A consequence of satisfying these axioms is that the utility function representing the preference relation is a very well-behaved function. A preference relation in the descriptive or behavioral category, however, may violate some of the axioms that define normative preferences. As a result, the corresponding utility function may not be a very well-behaved function. This will effectively mean that the objective of our optimization problem may not be a smooth, concave function.

3.3 Optimization Techniques

We now give a brief overview of the general optimization techniques for finding a solution for a nonlinear optimization problem. Presentation is tailored for our problem and is not an attempt to be comprehensive. Some excellent sources for detailed discussion on optimization theory, numerical optimization and related practical issues are Gill, Murray and Wright [23], Bersekas [5], Luenberger [31]. A general nonlinearly constrained optimization problem can be written as:

$$\begin{aligned} \max_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \ni \quad & \mathbf{l} \leq \mathbf{x} \leq \mathbf{u} \\ & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{c}(\mathbf{x}) \leq \mathbf{0}, \end{aligned}$$

where $\mathbf{x}$, $\mathbf{l}$ and $\mathbf{u}$ are vectors of dimension m, $\mathbf{A}$ is a matrix of size mxn, $\mathbf{c}(.)$ is a vector of nonlinear constraints and $f(.)$ is the nonlinear objective function. Therefore, the above optimization problem has a nonlinear objective along with bounds on the variables and linear & nonlinear inequality constraints. An optima for this problem is characterized by a certain set of optimality conditions. Numerical algorithms to solve the problem are developed by focusing on satisfying these conditions.

3.3.1 Univariate Optimization

If the optimization problem is one-dimensional, i.e., $x \in \Re$, and has no constraints or bounds, a solution x^* is characterized by the following conditions,

$$f'(x^*) = 0 \text{ and } f''(x^*) < 0,$$

assuming sufficient differentiability of function f.

If derivatives of the objective are available, an approach for finding an optima is to find a zero of the function $f'(.)$, which is a necessary condition for an optima. Conditions for sufficiency can then be checked at these points. Some standard and well-known methods for finding the zeros of a function are the bisection method, the secant method and Newton's method. These are iterative methods based on reducing the interval of uncertainty at each iteration. The bisection method reduces the interval of uncertainty by half in each iteration by comparing the function values at the end points of the two subintervals. Newton's method and the secant method approximate the function by a linear function and the iterates are taken to be the zero of the linear approximations.

Newton's method can also be interpreted as a polynomial fitting method. In these methods a polynomial is fitted to the objective function, $f(x)$ using the function and derivative information at the iterates and the optima of the approximating polynomial is taken as the next iterate. In this class, Newton's method corresponds to locally fitting a quadratic polynomial to the objective. Bounds on the variable do not result in much complication in one-dimensional problems, since they can be separately checked for being the optima.

3.3.2 Multivariate Optimization

The problem of finding an optima becomes much harder when going from a one dimensional problem to a higher dimensional one. In a single dimension there are a total of two directions to choose from as a search direction; in higher dimensions however, this choice set becomes infinite.

Line search methods form an important class of methods for multivariate optimization problems. Major considerations for iteratively finding the optima by line search methods are (in each iteration) : 1) finding a search direction and 2) finding an appropriate length to move in the search direction to obtain the next iterate. Convergence of the iterates (to a local optima) is guaranteed if the direction of search is one of ascent (locally the objective function increases in this direction) and the step-length results in a "sufficient" increase in the function value. Newton's method for the univariate case generalizes for the multivariate problem as follows. An approximating quadratic function is obtained by a second-order Taylor expansion of the objective function at each iterate. The search direction is computed by solving the exact optimal for the approximating quadratic function. A step of an appropriate length is then taken in the search direction so that there is a "sufficient" increase in the objective value.

The approximating quadratic function has a strict maxima only if the second derivative of the objective at the iterate, $\nabla^2 f(x_k)$ (known as the Hessian), is negative definite. If this is not the case, the Hessian of the approximating quadratic function is taken to be a modified negative definite version of the true Hessian of the objective. This method is called the modified Newton's method.

3.3.3 Quasi-Newton Algorithm

If the Hessian of the objective is either difficult or impossible to compute, the curvature information can be built up as the iterations progress. The objective and gradient information gained at each iteration is used. This class of methods is called the quasi-Newton methods. The initial Hessian approximation B_0 is usually taken to be the identity matrix or whatever information is initially available on the Hessian. In every subsequent iteration k, the Hessian is updated by adding an updating matrix U_k to the Hessian. The updating matrix is such that the symmetric, negative-definiteness of the approximate Hessian is not destroyed.

So far we have considered methods for unconstrained optimization problems. For problems with bounds and constraints additional considerations will be required in the development of an algorithm. For problems with simple bounds the search direction is selected so that feasibility is maintained. Variables are classified as free or fixed; free variable implies that the variable is not on its bound, and fixed if the variable is on its bound. At each iterate of Newton's method or the quasi-Newton method the descent direction is computed using only gradients and Hessian corresponding to the free variables, and Hessian updates are required only for the free variables. Depending on the step length, the variables either stay fixed/free or change from fixed to free or vice versa, hence changing the working set. We next describe the sequential quadratic programming methods that are considered the most robust and efficient methods for solving nonlinearly constrained optimization problems.

3.3.4 Sequential Quadratic Programming Algorithm

A brief description of sequential quadratic programming (SQP) algorithms follows (a more detailed description is given in Appendix A). An optimization problem with nonlinear equality constraints can be stated as:

$$\begin{aligned} \max_{x} \quad & f(\mathbf{x}) \\ \ni \quad & \mathbf{c}(\mathbf{x}) = 0 \end{aligned}$$

where $\mathbf{x}$ is a vector of dimension m, $\mathbf{c}(.)$ is a vector of nonlinear constraints and $f(.)$ is the nonlinear objective function. An associated function to the above optimization problem, which gets used in finding the optima, is the Lagrangian function for the problem and is defined as

$$L(x, \lambda) = f(x) - \lambda^T c(x).$$

Optimality conditions for this problem are as follows: x^* is a local maximizer of the problem if there exist multipliers λ^* such that

$$\begin{aligned} c(x^*) &= 0 && (3.4)\\ \hat{J}^T\lambda^* &= g(x^*), \ \text{[or } Z^T g(x^*) = 0] && (3.5)\\ \lambda^* &> 0 && (3.6)\\ Z^T(x^*)W(x^*,\lambda^*)Z(x^*) & \quad \text{is negative definite,} && (3.7) \end{aligned}$$

where $g = \nabla f(x)$, $\hat{J}$ is the constraint Jacobian, $Z(x)$ is a basis for the nullspace of $\hat{J}(x)$, and $W(x, \lambda)$ is the Hessian of the Lagrangian $L(x, \lambda)$. These conditions are called the second-order KKT (Karush-Kuhn-Tucker) conditions, the first three being first-order conditions. These conditions are used for developing a method for finding a solution for the constrained optimization problem.

Some early techniques for finding an optima for a nonlinearly constrained problem involved converting the constrained problem into an unconstrained one. This is done by including the nonlinear constraints within the objective in an appropriate way, such as by the use of a penalty function or a barrier function. A penalty function approach allows the iterates to become infeasible, but incurs a penalty that increases with infeasibility, whereas a barrier function is a modification of the objective so that the iterates are always feasible.

SQP methods are considered the most robust and efficient algorithms for solving nonlinearly constrained optimization problems. They are iterative procedures and can be viewed as a generalization of Newton's method to the (inequality) constrained problem. The basic idea is to construct a constrained quadratic model around the current iterate, x_k. Maximization of the quadratic model gives an update p_k that approximates the error $(x^* - x_k)$, where x^* is a local optima of the objective function, f. The new iterate is taken as $x_{k+1} = x_k + \alpha_k p_k$, where α_k is usually determined by a line search method applied to a merit function. Trust-region methods may also be used.

Any algorithm for constrained optimization needs to ensure that the algorithm converges to a point that is both feasible and a constrained minimizer. Therefore, to measure how good the next iterate is, simply looking at the objective is not enough because that does not say anything about how close the iterate is to the feasible set. A standard approach today is to define a merit function, which combines both the objective and the constraint violations into one function. For this reason merit functions are closely related to penalty functions. The two most popular merit functions for SQP methods are the l_1 merit function

$$M(x) = F(x) + \rho \|c(x)\|_1,$$

and the augmented Lagrangian

$$M(x, \lambda) = F(x) - \lambda^T c(x) + \frac{\rho}{2} c(x)^T c(x),$$

where $\rho > 0$ is a penalty parameter and λ is a set of Lagrange multiplier estimates.

Both the objective of the approximating quadratic problem and the merit function for computing step-length require the Lagrange multipliers. These are approximated and updated from the solution of the quadratic subproblem.

3.3.5 Application to the Investment Problem

Our observations of the previous section indicated that the optimal investment choice problem formulated using a binomial-tree model for asset returns is a nonlinear problem with bounds on the variables.

The smoothness properties of our objective are largely dependent on the smoothness properties of the utility function $U(.)$, since the expression at which utility is evaluated is a smooth function of the decision variables. Even if the utility is given to be a smooth, differentiable function at all but a finite number of points, it is very cumbersome to explicitly compute the gradients of the objective. Our options were to either use an automatic differentiation software package such as ADIFOR (see Appendix B) to compute the derivatives, or use optimization methods that use only objective function values or perform a finite-difference approximation for the gradient of the objective to compute a descent direction. Alternatively, we could consider reformulating the problem such that computing the exact derivatives is simplified.

We used the NPSOL [22] software package, a Fortran package for nonlinear programming developed at the Systems Optimization Laboratory at Stanford University. (See Appendix C for more details on NPSOL.) NPSOL implements an SQP algorithm [21] and is a general-purpose system for solving optimization problems involving constraints. It maximizes a linear or nonlinear function subject to bounds on the variables and linear or nonlinear constraints. NPSOL finds solutions that are locally optimal. Ideally any nonlinear functions should be smooth and users should provide gradients, although this is not essential. Unknown gradients are estimated by finite differences. Discontinuities in the function gradients can often be tolerated if they are not too close to an optimum. Each quadratic program is solved using a quadratic programming package with several features that improve the efficiency of an SQP algorithm.

3.4 Modeling Variations

In this section we first consider some variations in the modeling of the optimization problem that facilitate computation of an optima. This is followed by some variations in the model for the investment problem.

3.4.1 Wealth as Variables

In the earlier sections of the chapter we formulated the optimal investments problem as a nonlinear programming problem in terms of the decision variables, $X_{t,n}$. This formulation makes the computation of the objective gradients expensive and cumbersome. For simplicity we can use an optimizer that either requires only the function values for computing search directions or uses finite-difference approximation of the objective gradients. Both approaches would be inefficient, since they do not use the exact first derivative information.

It is possible to rectify this by considering the wealth at each node to be a variable in the optimization problem. Since the wealth values at each node are fully determined once the investment decisions $X_{t,n}$ are known, we need to include constraints that impose this characteristic of the investment problem. There will be $2(2^T - 1)$ variables in the new optimization problem. In the terminology of optimal control, the wealth variables are the state-variables of the problem, whereas $X_{t,n}$'s are the control variables. In a typical stochastic control problem the number of state variables is much larger than the number of control variables, but in our case they are almost equal. This makes it feasible to treat them both as variables of the optimization problem. However, we now need to solve a nonlinearly constrained optimization problem. The nonlinear constraints of the problem are given by (3.1) & (3.2), where X_{t,n_1}, W_{t,n_1}, W_{t+1,n_2} and W_{t+1,n_3} are now variables $\forall t, n_1, n_2, n_3$ relevant indices. This implies that the problem now has as many nonlinear constraints as there are decision variables, i.e., $2^T - 1$ of them, where T is the planning horizon. These constraints are very structured; each constraint involves only three variables and is bilinear, i.e., each term in the left-hand-side of the constraint is a product of two variables. These observations imply that the constraint Jacobian is very easy to compute and is sparse and structured. All these properties can be put to good use in an optimization algorithm. The most attractive advantage of this modeling variation is that now it is possible to compute all the objective gradients with great ease. It is, in fact, also possible to compute the exact Hessian of the objective and constraint functions, which enables computing the Hessian of the Lagrangian. The Hessian of the Lagrangian is also sparse and highly structured. In the original formulation, although the Hessian of the objective is half the size, it is dense and hard to compute. The gains in efficiency from this modeling variation outweigh the losses from an increase in problem size.

A question that arises at this point is should the wealth at the planning horizon, T also be treated as variable? If we do so, the number of nonlinear constraints increases to $2^T - 1 + 2^T$, which is an increase of 2^T, and the number of variables increases to $2(2^T - 1) + 2^T$, which is also an increase of 2^T. The objective function of the problem becomes:

$$\max_{X_{t,n},t=0,T-1,n=1,N,W_{t,n},t=0,T,n=1,N_1} \sum_{m=q}^{N_1} [pU(W_{T,m})],$$

where p is the probability of each final node and m ranges over all the final nodes of the tree. Clearly, this further reduces the nonlinearity of the objective. However, we observe that the benefits from this change do not outweigh the loss of efficiency from an increased number of variables and constraints. This modeling approach may become desirable if we need to impose additional bounds on the wealth at the planning horizon.

As was observed above, the alternate formulation made it possible to explicitly compute the objective gradients and constraint Jacobian, and these were both structured and sparse. We used SNOPT [20] software package, a Fortran based package for nonlinear programming from the Systems Optimization Laboratory at Stanford University. SNOPT implements an SQP algorithm [21] and is a general-purpose system for solving optimization problems involving many variables and constraints. It maximizes a linear or nonlinear function subject to bounds on the variables and sparse linear or nonlinear constraints. SNOPT finds solutions that are locally optimal, and ideally any nonlinear functions should be smooth and users should provide gradients, although this is not essential. Unknown gradients are estimated by finite differences. Discontinuities in the function gradients can often be tolerated if they are not too close to an optimum. A more detailed description of the features of SNOPT are discussed in Appendix D.

3.4.2 Non-Uniform Time Periods

The investment problem is solved on a rolling basis, hence each time an investment decision needs to be made the optimization problem is solved taking the currently owned wealth as the starting wealth of the problem. The optimal decision for $t = 0$ is implemented. This implies that the optimal decisions suggested for subsequent periods are not of direct relevance to the planner. Such usage of the model suggests that there should be a finer time-resolution in the beginning of the planning horizon compared to later. In other words, the model for the problem should have shorter time-periods in the beginning of the problem, and these time periods should progressively become larger towards the end. If a planner intends to make half-yearly investment decisions, the model may first have 5 years spanned by half-year periods (a total of 10 optimization problem periods), then have another 3 years spanned by one year periods (additional 3 optimization problem periods) and finally have 12 years spanned by 2 two year periods and 2 four year periods (which are another 4 optimization problem periods). Consequently, there will be 17 periods in this problem spanning a total of 20 years. When the problem is rolled over, in subsequent periods a different selection of time-period lengths may be selected, perhaps with a greater number of shorter-length periods. Therefore, there is considerable flexibility in this framework for planning period size and planning horizon; this flexibility can be used to the planner's advantage.

This change in the model of the investment problem does not structurally change the character of the optimization problem. The parameters of the problem change, specifically the asset returns for a period vary depending on the length of the period.

3.4.3 Intermediate Contributions

Typically people do not make a single large contribution to their retirement funds. Instead, it is more common to make periodic contributions to the fund. This feature of the problem is included in the model by introducing a contribution variable denoting the dollar amount invested in the retirement fund in each period. The contribution variables can be taken to be state and time-dependent, or be only time-dependent, alternately for further simplicity, may be the same for all states and time.

While these changes introduce additional variables, the number of nonlinear constraints remains the same. The nonlinear constraints will now have an extra linear term. We may also introduce additional bounds and constraints on the contribution variables.

3.5 Comparison of Computational Effort

In this section we compare the computational complexity & effort of our approach with some other approaches, such as dynamic programming, stochastic optimal control solved by solving the nonlinear partial differential equation and stochastic optimization.

3.5.1 Dynamic Programming

Dynamic programming is a general technique for solving dynamic optimization problems, such as the investment problem, using a recursion structure. The basic idea of this technique is to optimize the decision for the first period, assuming that subsequently optimal decisions are made. Since the subsequent optimal decisions are not known, a backward recursion approach is adopted in an implementation of the technique. The problem is formulated by identifying the set of actions the planner may take, denoted by A, and a set that describes all possible states of the world for each time period, denoted by S. If the problem is a finite horizon problem, we denote the planning horizon by T. If $V_t(s)$ is defined as the maximum expected "reward" that can be earned starting from time t and state s up to T, then the recursion relation may be written as

$$V_t(s) = \max_{a \in A}(r(s,a) + E[V_{t+1}(s)]), \qquad \text{where } \ s \in S. \tag{3.8}$$

The term $r(s,a)$ denotes the reward obtained in a single period in state s, when action a is taken. The function V is commonly known as the *value function* for the problem and the above equation is the *Hamilton-Jacobi-Bellman* (HJB) equation. In order to solve the problem one also needs to define an end condition, such as $V_T(s) = f(s)$. Using the end condition, equation (3.8) needs to be solved for each time $t = T - 1 : 0$ and for all values s, the state of the world, may take.

In the retirement planning problem, the planner's utility is for the wealth she is able to accumulate in the retirement fund at the time of retirement. The states of the world are defined by the wealth level in the retirement fund at each time period, which in turn depends on the investment decisions and the returns obtained on the investments. The actions clearly are the investment decisions, which in our problem are to decide at each time period what fraction of wealth to invest in the risky asset. There is no immediate reward, hence the term $r(s,a)$ is zero. The backward recursion equation becomes

$$V_T(W_T) = U(W_T) \tag{3.9}$$

$$V_t(W_t) = \max_{x_t \in [0,1]} (E[V_{t+1}(x_t W_t S_t + (1-x_t)W_t R_f)]), \quad \text{for} \quad t = T-1:0, \tag{3.10}$$

where S_t is the per dollar return on the risky asset and R_f is the per dollar return on the riskfree one. To solve the problem the recursion equation (3.10) needs to be solved for a range of levels of wealth, W_t to determine the functional form for V_t, since this will be required for the recursion step of $t-1$. If V_t is not obtainable in closed form, the computations become considerably intensive. If the utility function $U(.)$ is not well-behaved, the optima in each step of recursion may not be easy to obtain. Issues related to uniqueness of solution are also of concern here.

3.5.2 Stochastic Optimal Control

When instead of a discrete-time formulation, the problem is solved in continuous-time, the above approach becomes a stochastic control problem. In continuous-time the HJB equation given above becomes the following nonlinear partial differential equation.

$$\sup_{a \in A} D^a J(y,t) = 0, \qquad (y,t) \in (Y, [0,T]), \tag{3.11}$$

where,

$$D^a J(y,t) = g(a,y) J_y(y,t) + J_t(y,t) + \frac{1}{2} h(a,y)^2 J_{yy}(y,t), \tag{3.12}$$

along with the boundary condition

$$J(y,T) = U(y), \qquad y \in Y. \tag{3.13}$$

The functions $g(a, y)$ and $h(a, y)$ above arise from the equations that govern the evolution of the state of the world given by y. Analytical solutions for this problem are hard to obtain, and may be obtainable for only certain simple utility functions. For more complex problems numerical approaches are required. The nonlinear partial differential equations are solved using finite-difference methods. These are usually rather large, complex numerical problems, and sophisticated methods need to be employed.

3.5.3 Stochastic Optimization

The stochastic optimization approach is a scenario based approach for a finite horizon problem where the uncertain components of the problem are taken to have a finite number of outcomes in each time period. Decision variables are associated with each outcome at each time period. The problem is formulated by introducing a constraint for each scenario at each time period in order to keep decisions of a time period consistent with those taken in the prior period(s). This results in a large constrained optimization problem. The constraint matrix consists of structured blocks. In order to solve the problem when the objective is linear, the structure is exploited using decomposition techniques, such as the L-shaped method, which is essentially a Dantzig-Wolfe decomposition [14] of the dual or a Benders' decomposition [3] of the primal. A variation in the decomposition based method is one in which instead of exactly computing the recourse function it is approximated using sampling techniques. (See [12, 14].)

As the planning horizon or the number of sources of uncertainty or the number of possible outcomes for each source of uncertainty increases, the problem size grows enormously. This makes these problems computationally very challenging. Besides, these solution techniques were developed for problems with linear objectives. Using these approaches for a general nonlinear problem has definite challenges that are being currently addressed (see [6, 7]).

Chapter 4

Application to Investment Problem

A planner laying out a financial plan must have a criterion for comparing various investment strategies that may be adopted, and possibly choose a strategy that is best with respect to the criterion. In this chapter an outline for the theory of choice and the utility theory is presented. We make distinctions between normative, descriptive and prescriptive choice and present the optimal investment strategy with respect to a class of descriptive or behavioral utilities. Finally, some results from the standard normative utilities are presented and a comparison of results from the two categories is discussed.

4.1 Introduction - Utility Theory

The process by which people make their wealth related decisions has been the subject of study for a very long time, and can be at least dated back to Bernoulli [4]. In the twentieth century, among others, work by Friedman and Savage [18], von Neumann and Morgenstern [58], Markowitz [37] and Samuelson [51] forms the basis of research in this field.

An individual faced with a decision chooses from a set of alternatives. Let X

denote this set of alternatives. Among all elements of X the decision-maker weighs the merit and de-merits of each option against others and attempts to select the option that suits her "best." This requires a precise definition of "best."

In more rigorous terms, we need to define an ordering on the set of alternatives that makes it possible to compare every element of the set with every other element. We denote this ordering by '$\succeq$', then if $x, y \in X$, $x \succeq y$ will imply that the decision-maker prefers x to y, and $x \succ y$ would imply x is strictly preferred to y.

The preference relation defined above is said to be rational if it satisfies the *axioms of choice*. These axioms are reflexivity, completeness and transitivity and are defined as follows (see Varian [56], MasColell [38]) :

Reflexivity For all $x \in X$ we have $x \succeq x$.

Completeness For all $x, y \in X$ we have either $x \succeq y$, $y \succeq x$ or both.

Transitivity For all $x, y, z \in X$, if $x \succeq y$, $y \succeq z$ then $x \succeq z$.

For the existence of an ordinal utility function that maps the elements of X to the real line, $\Re$ ($U : X \mapsto \Re$), another axiom called the continuity axiom is required.

Continuity For all $x \in X$, the sets $\{y\colon\ x \succeq y\}$ and $\{y\colon\ y \succeq x\}$ are closed sets.

Given the above four axioms, existence of an ordinal utility function can be proved. This utility function is a continuous function such that

$$\begin{aligned} U(x) > U(y) \quad &\Leftrightarrow \quad x \succ y \\ U(x) = U(y) \quad &\Leftrightarrow \quad x \sim y. \end{aligned}$$

It is important to note that this representation for the preference relation "$\succ$" is not unique, since $\forall\ f$, monotonic and mapping $\Re \mapsto \Re$, $f(U(x))$ is also a representation for "$\succ$."

Although at first sight these axioms may look reasonable, it is worth taking a closer look. Imposition of these axioms has enormous implications in practical terms. Reflexivity is the weakest and most acceptable axiom. The completeness axiom implies that the decision-maker has done the introspection of all the alternatives, however far removed they may be from one's realm of common experience. It is easy to imagine that satisfying the completeness axiom may need serious work and reflection on one's preferences, and may in some cases be almost impossible. The transitivity axiom, on the other hand, implies that in a sequence of pairwise choices there is no possibility of cycles, which is to say that however the options are framed or presented, the decision-maker is mature enough to be able to rank all of them in an order that contains no cycles. Besides the above axioms of choice, preference relations are assumed to satisfy more axioms, such as monotonicity, convexity.

So far the alternatives from which the decision-maker chooses had no uncertainty associated with them. For a study of choice under uncertainty we will need to frame the axioms in terms of "*lotteries.*" Lotteries, L, are defined by outcomes $\mathbf{x} = \{x_1, x_2, ..., x_m\}$ and the corresponding probabilities $\pi = \{\pi_1, \pi_2, ..., \pi_m\}$. The probability of each outcome is taken to be the objective probability of the event. The axioms of choice in this case become:

Reflexivity For every lottery L, $L \succeq L$.

Completeness For every pair of lotteries L_1, L_2, either $L_1 \succeq L_2$ or $L_2 \succeq L_1$.

Transitivity For all L_1, L_2, L_3, if $L_1 \succeq L_2$ and $L_2 \succeq L_3$, then $L_1 \succeq L_3$.

Continuity If $x_1 \succeq x_2 \succeq x_3$, then there exists a probability $\pi, 0 \leq \pi \leq 1$, such that $x_2 \sim \{(x_2, x_3), (\pi, 1 - \pi)\}$. The probability π is unique unless $x_1 \sim x_3$.

Independence Let $L_1 = \{(x_1, .., x_\nu, .., x_m), \pi\}$ and $L_2 = \{(x_1, .., z, .., x_m), \pi\}$. If $x_\nu \sim z$, then $L_1 \sim L_2$. z may be either a simple outcome or another lottery. If

z is a lottery, $z = \{(x_1^\nu, x_2^\nu, ..., x_n^\nu), \pi^\nu\}$, then

$$L_1 \sim L_2 \sim \{(x_1, ..., x_{\nu-1}, x_1^\nu, ..., x_n^\nu, x_{\nu+1}, ..., x_m),$$
$$(\pi_1, ..., \pi_{\nu-1}, \pi_\nu \pi_1^\nu, ..., \pi_\nu \pi_n^\nu, \pi_{\nu+1}, ..., \pi_m)\}$$

Dominance Let $L_1 = \{(x_1, x_2), (\pi_1, 1 - \pi_1)\}$ and $L_2 = \{(x_1, x_2), (\pi_2, 1 - \pi_2)\}$. If $x_1 \succ x_2$, then $L_1 \succ L_2$ if and only if $\pi_1 > \pi_2$.

Assuming that the above axioms are satisfied, it is possible to prove the existence of an ordinal utility, such that the decision-maker chooses lotteries by comparing expected utility of the lotteries with respect to this ordinal utility and the objective probabilities, π. This is called the von Neumann-Morgenstern [58] utility after its originators. Therefore, if U is the ordinal utility for the outcomes, the von Neumann-Morgenstern utility becomes $\mathbf{U}(x) = E[U(x)]$. Utility functions so defined are determined uniquely up to a positive linear transformation, i.e., $\mathbf{U}(x)$ and $a\mathbf{U}(x) + b$ are equivalent for $a > 0$.

These axioms imply that only the final pay-offs matter and there is no thrill or aversion for suspense or gambling *per se*. Among all the above axioms the independence axiom is most difficult to satisfy in practical terms (see Machina [34] for more details).

4.1.1 Descriptive, Normative, and Prescriptive Decision Making

The study of the process of decision making can be classified into three major categories (see Bell, Raiffa and Tversky [2] for more details):

Descriptive Descriptive models, also called behavioral decision making models, are concerned with how and why real people think and act the way they do. How do they perceive uncertainties? What are their internal conflicts, biases or hang-ups? Do they decompose complex problems, think separately about component parts of problems, and then integrate these separate analyses? Or do they think more holistically and intuitively? The study of these questions may at times involve intricate mathematical modeling and require clinical, experimental or sophisticated statistical analysis. In short, it is an abstract system that purports to describe or predict behavior.

Normative A normative model can be defined as an abstract system that attempts to capture how ideal people behave. The hallmarks of normative analysis are rationality, coherence and consistency, as captured by certain prescribed axioms, basic principles and fundamental desiderata motivated by what is logical, rational and intelligent behavior. There is not much space for the cognitive concerns, internal turmoil, shifting values, anxieties, disappointments and regrets of real people, or space for the fact that real people may not always be capable of performing intricate computations etc. in this analysis.

Prescriptive Prescriptive analysis addresses the question, what could be done to help real people make better choices? It involves devising aids for helping people lessen their cognitive concerns and internal turmoil, and as a result helps them make more consistent, coherent decisions.

The axioms of choice listed in the previous section form the basis for the normative theory of choice. Decision-makers whose preferences satisfy these axioms are said to be economically rational and it is possible to represent the preference relation by a utility function, which is a map from the set of alternatives to the real line. If the preference relation also satisfies convexity and monotonicity properties, the utility function obtained is an increasing, smooth and concave function. Most research in economics and finance is developed on the basis of the assumption that decision-makers' preferences satisfy these axioms of choice.

In practical terms, it may be a difficult proposition for an average decision-maker to satisfy all the above axioms of choice. Decision-makers are not economic automatons, but instead make mistakes, have remorse, suffer anxiety and sometimes cannot make-up their minds. By such non-rational behavior we don't mean people with diminished capacity, but ordinary people who get caught in cognitive traps from time to time.

Several behavioral scientists have attempted to understand the decision-making process among ordinary people through experiments and studies of various kinds (see Bell, Raiffa and Tversky [2], Edwards and Tversky [16], Wright [60]). The five major phenomena observed and confirmed in a number of experiments with real or hypothetical payoffs that violate the standard models of normative choice are : Framing effects, Nonlinear preferences, Source dependence, Risk seeking and Loss aversion. These phenomena are defined in Tversky & Kahnemann [55, 54], Tversky [53] as follows.

Framing effects Lack of description invariance implying that variations in the framing of options yield systematically different preferences.

Nonlinear preferences The expectation principle of utility theory states that the utility of risky prospects is linear in outcome probabilities. However, experimental evidence indicates that people tend to transform probabilities nonlinearly, overweighing small probabilities and underweighing moderate and high ones.

Source dependence Willingness to bet on an uncertain event depends not only on the degree of uncertainty but also on its source.

Risk Seeking As opposed to the generally assumed risk aversion in economic analysis, in certain situations people prefer more risk to less. For instance, people prefer a small probability of winning a large prize over the expected value of that prospect.

Loss Aversion Carriers of value are gains and losses defined relative to a reference point. The losses loom larger than the gains, that is, an amount of loss elicits more "unhappiness" than the same amount of gain elicits "happiness."

In the rest of this chapter we explore the optimal investment decision for an investor with a behavioral utility. Clearly, this does not fall within the realm of a descriptive study, since no attempt is being made to see how people with such utilities actually make their investment decisions. It also does not fall under the normative framework, since these utilities do not satisfy the axioms of choice. This work may best be described as a prescriptive model or a second-order approximation of a descriptive model (assuming normative models are the first-order approximations). On the other hand, it is hoped that computation of optimal strategy corresponding to these behavioral utilities will help bridge the gap between what is rational and perfect and what is not. Bridging this gap will help "educate" people to make more rational decisions.

4.2 Behavioral Utilities - Challenge

In the normative decision models a monotonic, twice-differentiable, concave utility function is obtained. Such a well-behaved utility may not be obtainable when one or more of the assumptions of the normative theory are violated. Descriptive studies indicate that real people do not always act rationally as claimed by the normative models. They are hindered in making coherent, consistent and rational decisions due to short-falls in their cognitive capacities. As a result, their utility or value functions may lack the good properties associated with normative utility functions. Therefore, attempting to find an optima for behavioral utilities creates challenges.

People think in terms of summary statistics for their future wealth; for example, they may be interested in maximizing the probability of their future wealth exceeding a "Goal" value. They may determine their "Goal" value as an estimate of the amount of wealth that will allow them to lead a comfortable life after retirement. In a discrete model this will imply that the value function is discontinuous, making computation of objective derivatives almost impossible. Alternatively, besides desiring to reach the Goal value, they may also want to mitigate losses as much as possible. In other words, there may be multiple-objectives involving both maximizing the probability of reaching a Goal and maximizing say, the fifth percentile of future wealth distribution. People may evaluate their wealth outcomes with respect to a reference point, judging doing better than the reference differently than doing worse. This will result in a different value function on either side of the reference point, and hence a possible loss of differentiability at the reference point.

Special considerations will be required for finding the optimal investment decisions for all these different types of utilities or value functions. In the following sections we explore and present our results for the problem.

4.3 Optimal Results for Behavioral Utilities

Average decision-makers are not economic automatons, instead from time to time they are affected by emotions and cognitive hindrances in making rational decisions. For these reasons their utility functions may differ from the ones implied by the normative theory of choice. In what follows it is assumed that although this may be the case, the decision-makers should nevertheless make decisions that are optimal with respect to their utilities. Wherever relevant throughout this section, we will also assume that the decision-maker is an expected utility maximizer.

The optimization problem formulated in the previous chapter, with binomial tree model for asset-returns, is to find the optimal value of the investment decision variables, $X_{t,n}$, so that the expected utility of wealth at the planning horizon is maximized. Therefore, the following problem is solved for specific utility functions, $U(.)$:

$$\begin{aligned} \max_{X_{t,n}, t=0,T-1, n=1,N} \sum_{m=q}^{N} [pU(X_{T-1,m}W_{T-1,m}S_u &+ (1-X_{T-1,m})W_{T-1,m}R_f) + \\ pU(X_{T-1,m}W_{T-1,m}S_d &+ (1-X_{T-1,m})W_{T-1,m}R_f)] \\ \ni 0 \leq X_{t,n} \leq 1, &\qquad \forall t, n, \end{aligned}$$

where p is the probability of each node at the planning horizon, T is the planning horizon and $U(.)$ is the investor's utility for wealth at the planning horizon. The index m ranges over all the outcomes at one period prior to the planning horizon. In our problem formulation we permit no short sale, i.e., borrowing of assets is not allowed. This restriction translates to the above bounds on the decision variables. These bounds reflect regulatory constraints on positions the planner can take.

In the above formulation it is possible to solve for the optimal asset allocation decisions for a large class of utilities; only some of these are presented in what follows. Also, the size of problems for which results are presented is chosen to be moderate for ease of presentation. The results from the non-uniform time-periods framework is presented only for the loss-aversion utility, although other utilities may be used.

4.3.1 Loss Aversion Utility

Tversky and Kahnemann developed an alternative theory to the expected utility theory of normative choice, called the prospect theory. The salient features of this alternative theory are (see Tversky and Kahneman [54] for more details) :

Reference dependence The carriers of utility are gains and losses defined relative

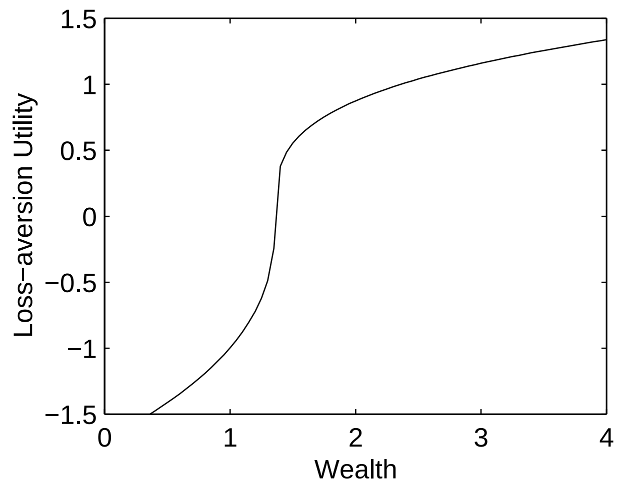

Figure 4.1: Plot of loss-aversion utility against wealth, W

to a reference point.

Loss aversion The function is steeper in the negative than in the positive domain, implying that losses loom larger than corresponding gains.

Diminishing sensitivity The marginal utility of both gains and losses decreases with their size.

This utility function is continuous and increasing, but the above properties give it an asymmetric S-shape (see figure 4.1). The utility is convex and much steeper below the reference point and concave and less steep above it, hence the derivatives don't match at the reference point.

Figures (4.2) and (4.3) give the optimal asset allocation for loss-aversion utility for $T = 9$ and the reference point is taken as the 8% annual return on the starting wealth. In figure (4.2) the optimal allocation for nodes at each time-period are plotted from "top" to "bottom." "Top" is the scenario where return is high in all prior time periods and "bottom" is where return has been low in all prior periods. In figure

(4.3) the allocation is plotted by wealth level at the nodes in each time period. Due to overlapping fewer points are visible in these plots than there are nodes in the tree at each time-period.

It is interesting to observe the effect of the reference point on the optimal asset allocation. In figure (4.4) and (4.5) the optimal asset allocation for three different reference points for the loss aversion utility is given. The first reference point is as before, the second corresponds to the riskfree annual rate of return of 5% and the third is the starting wealth.

Note that in the above formulation of the problem all the characteristics of prospect theory modeling of utility are not followed; in particular, we do not weight the probabilities associated with the final nodes in the tree according to Kahneman-Tversky's prospect theory.

Results for non-uniform time periods are given in figure (4.6). A 9 periods problem is solved with 4 quarters, 2 half-year periods, 1 one year period, 1 two years period and 1 four years period. Therefore, a total of 9 years is spanned with a higher resolution for periods close to $t = 0$.

4.3.2 Piece-wise Linear Utility

The piece-wise linear utility is also a utility with a reference point; there is a kink at the reference point. The slope is higher below the kink than above it. Figure (4.7) plots the utility.

Formulation of the problem as done for loss-aversion utility does not work for a piece-wise linear objective. SNOPT tests for the correctness of the objective gradients

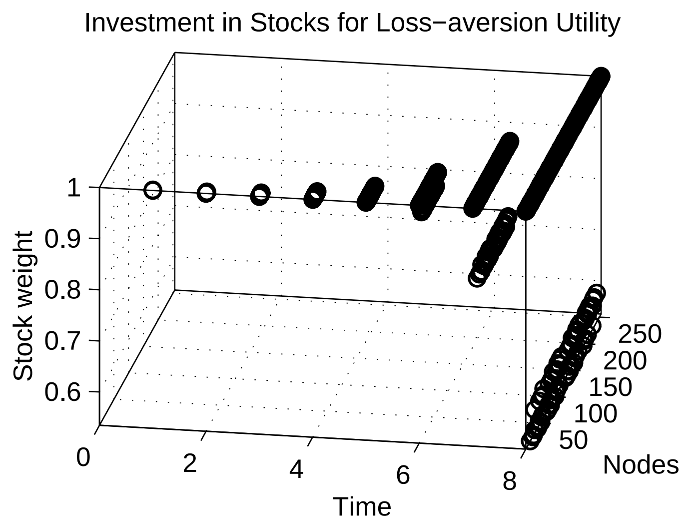

Figure 4.2: Optimal Stocks investment weights for a 9 periods problem using Loss aversion utility

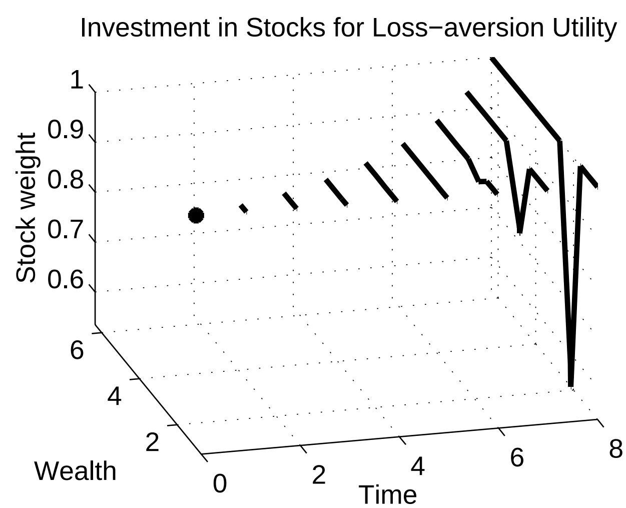

Figure 4.3: Optimal Stocks investment weights for a 9 periods problem using Loss aversion utility, plotted by wealth levels at nodes for each time period

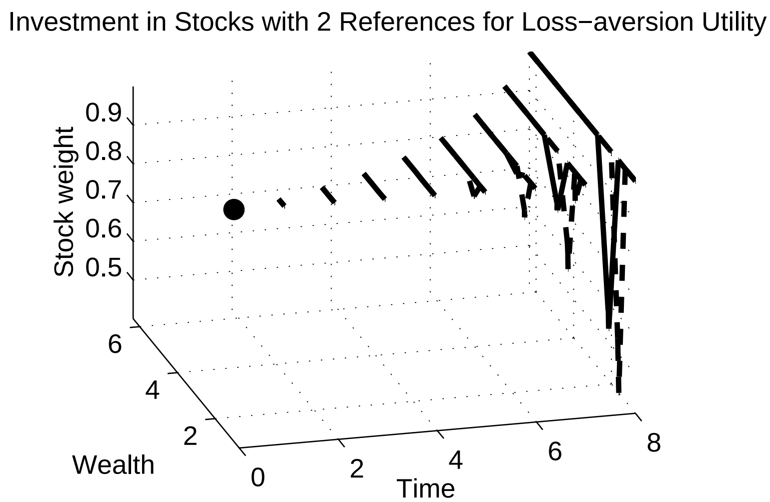

Figure 4.4: Optimal Stocks investment weights for a 9 periods problem using Loss aversion utility, with different reference points (continuous for 8% yearly return, dashed line for 5% yearly return)

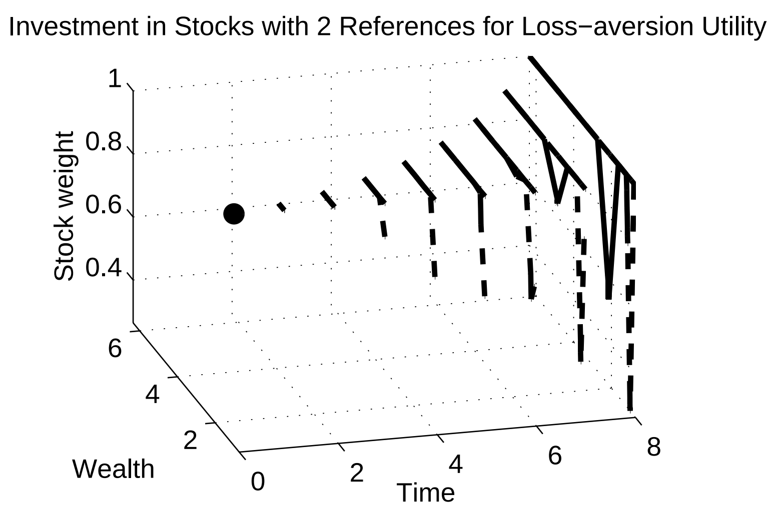

Figure 4.5: Optimal Stocks investment weights for a 9 periods problem using Loss aversion utility, with different reference points (continuous for 8% yearly return, dashed line for starting wealth)

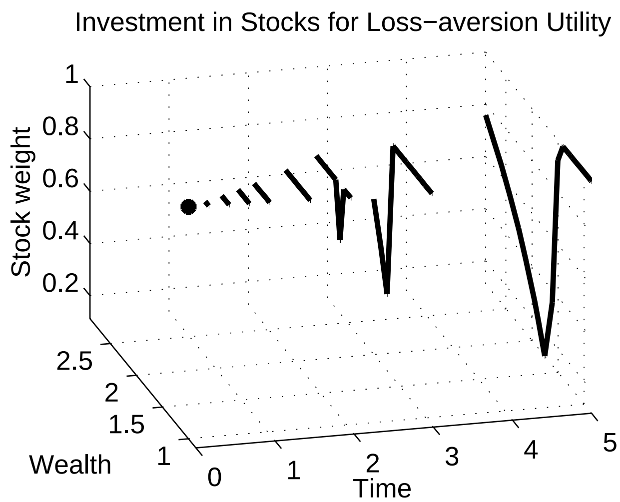

Figure 4.6: Optimal Stocks investment weights for a 9 periods problem using Loss aversion utility, with non-uniform time periods (4 quarters, 2 half year, 1 one year, 1 two year and 1 four year periods)

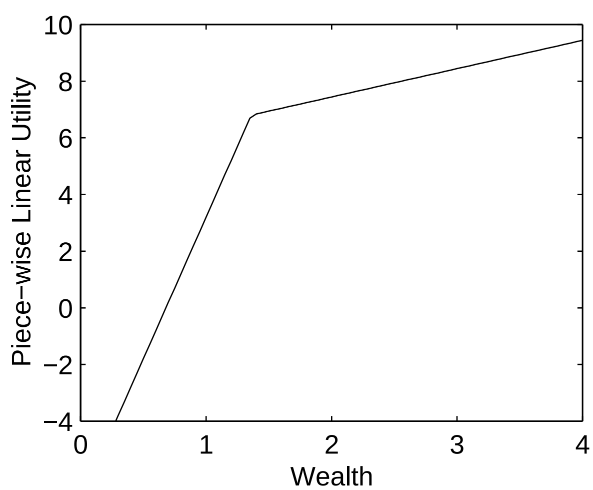

Figure 4.7: Plot of piece-wise linear utility against wealth

provided using finite-difference approximation (since it assumes that the objective is a smooth function). The finite difference approximation of derivative does not match those given to the solver and hence it detects an error. In order to be able to solve the problem for a piece-wise linear utility, we first alter the formulation of the problem as follows.

Wealth levels at all nodes are taken as variables, including the wealth at the planning horizon T. Two extra variables are defined for each node at the planning horizon of the tree, i.e., there are now $(2^T - 1)$ $X_{t,n}$ variables, $(2^{T+1} - 1)$ $W_{t,n}$ variables and $2(2^T)$ additional variables. Let these additional variables be $Z_1(1 : 2^T)$ and $Z_2(1 : 2^T)$. If the slope of each linear in the utility is c_1 and c_2 and the reference point is $\hat{w}$, then the objective for the new, equivalent problem formulation becomes

$$\max_{X_{t,n}, W_{t,n}, Z_1, Z_2} \sum_{i=1}^{2^T} \quad p(c_1 Z_1(i) + c_2 Z_2(i)) \tag{4.1}$$

$$\ni W_{T,2^T+i-1} = Z_1(i) + Z_2(i) \text{ for } i = 1 : 2^T \tag{4.2}$$

$$0 \leq Z_1(i) \leq \hat{w} \text{ for } i = 1 : 2^T \tag{4.3}$$

$$0 \leq Z_2(i) \leq \infty \text{ for } i = 1 : 2^T, \tag{4.4}$$

along with the constraints (3.1) & (3.2) involving $X_{t,n}$ and $W_{t,n}$ variables. The bounds (4.2), (4.3) & (4.4) make sure that the variables $Z_1(1 : 2^T)$ and $Z_2(1 : 2^T)$ are consistent with the variables $W_{T,n}$ for $n = 2^T : 2^{T+1} - 1$. In the new formulation the objective is linear.

It may be thought that increasing the number of variables in this manner would have a serious impact on the efficiency of the solver. However, at every iteration it can be shown that at least one of the two new variables, $Z_1(i)$ and $Z_2(i)$, must be on a bound. Consequently, the degrees of freedom does not increase at each iteration and this factor dominates the efficiency of SNOPT. In figure (4.8) the optimal asset

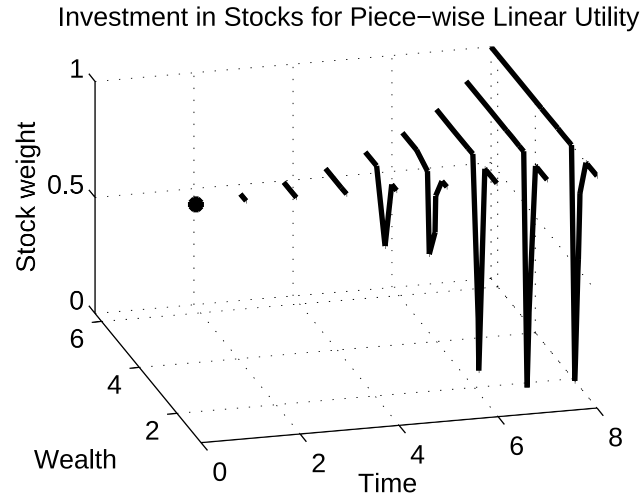

Figure 4.8: Optimal Stocks investment weights for a 9 periods problem using Piece-wise Linear utility, plotted by wealth levels at nodes for each time period

allocation is given for $T = 9$. We compare the solutions for shifted reference points, taking them to be the starting wealth and riskfree return on the starting wealth. The results are very similar to the loss aversion utility.

4.3.3 α - t Semivariance Utility

The α - t semivariance utility also has a reference point and is quadratic below the reference and linear above it. Figure (4.9) plots the utility. Two cases for this utility are considered - one flat above the reference, the other increasing above the reference. In figures (4.10) & (4.11) the optimal asset allocation is given for the two cases of α-t semivariance utility for a 9 periods problem.

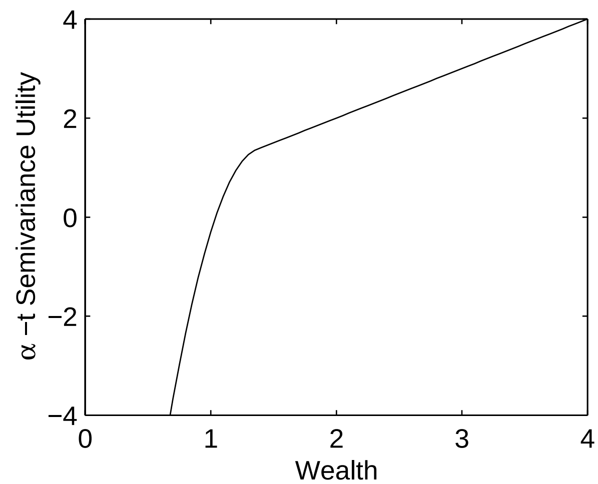

Figure 4.9: Plot of α-t semivariance utility against wealth

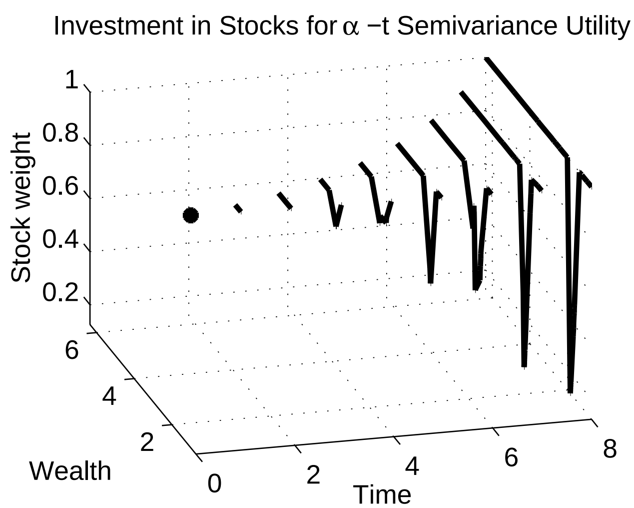

Figure 4.10: Optimal Stocks weights for a 9 periods problem using α-t semivariance utility, plotted by wealth levels at each time period

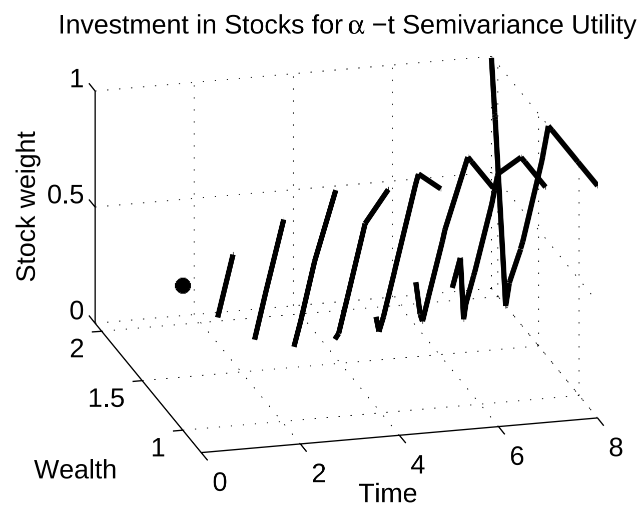

Figure 4.11: Optimal Stocks investment weights for a 9 periods problem using α-t semivariance utility, where slope is zero above the reference

4.3.4 Probability of Reaching a Goal

This utility evaluates the performance of wealth accumulated at the planning horizon by the probability of it exceeding a prespecified target wealth level, referred to as *Goal*. The decision-makers may come up with a target wealth level by estimating the total expenses they expect to have after retirement. Because of its intuitive appeal, financial advisors like to talk in terms of this utility with their clients. Mathematically the utility may be written as

$$Prob(W_T > G), \tag{4.5}$$

where W_T is the random variable denoting the wealth at the planning horizon and G is the target wealth level.

In a discrete model for asset-returns this utility is a discontinuous one, hence the problem needs to be formulated so that the objective becomes a continuous function. Our approach is as follows. Adopt an investment strategy such that the target value (Goal) is achieved at a certain number of nodes at the planning horizon. Then select nodes at the planning horizon in a certain order in order to maximize the wealth attained at them. This formulation of the problem results in the objective being continuous.

A good starting point is crucial in this problem. We present two strategies to get a good starting point.

Strategy 1 The investor initially invests only in the risky asset unless a transition node is reached. At the transition node part of wealth is invested in the riskfree asset and at all subsequent nodes all wealth is invested in the riskfree asset. The determination of the transition node and the proportion of wealth to invest in the risky asset is determined by assuring that the target wealth is reached at the node at the planning horizon.

Strategy 2 The investor initially invests both in the risky and the riskfree asset unless a transition node is reached. At the transition node investment in the two assets is made such that at all subsequent nodes wealth needs to be invested only in the riskfree asset to achieve the target wealth. Prior to the transition node investment in the two asset classes is made so that at each node the *expected* returns for the remaining time periods just attains the Goal level.

Note that both these starting points are feasible and that the total number of active constraints is close to the number of variables. After adopting one of the above two strategies a set of nodes at the planning horizon is at the Goal level. We next pick the highest non-Goal node and maximize the wealth level at this node along with constraints that the nodes that have already achieved the Goal stay at Goal. Optimization could be done on a sub-tree or on the entire tree. In figure (4.12) investments for a planning horizon, $T = 7$ are given, and in (4.13) a histogram for the corresponding wealth at the planning horizon is given. Note that a considerable number (almost 50%) of nodes are at Goal, while on the other hand there are some

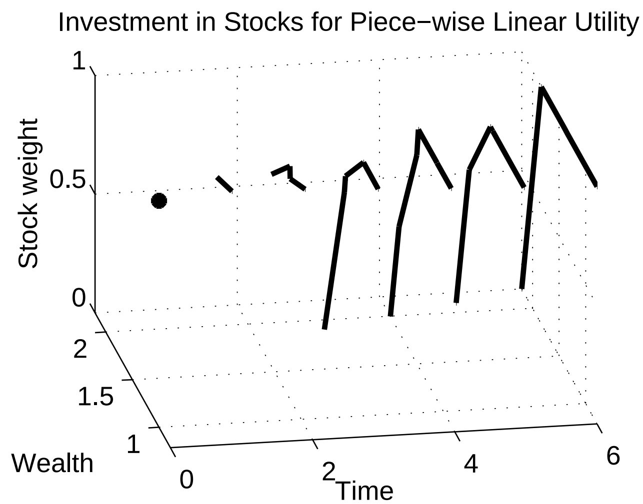

Figure 4.12: Optimal Stocks weights for a 7 periods problem using Probability of Goal utility, plotted by wealth levels for each time period

nodes that are even below the starting wealth level of 1 unit. It is interesting to observe the number of nodes at the planning horizon at which Goal is attained for different values of Goal levels. Figure (4.14) gives a plot of number of the nodes at which Goal is attained against different Goal levels. A conspicuous jump occurs in the number of Goal-nodes for a Goal value around 2.23.

A variant of the above problem is that the investor has a rough idea of what her Goal value is, but will like to maximize the Goal value once a certain number of nodes at which it is attained has been determined. For this problem we again adopt one of the above two strategies and perform the above mentioned optimization on number of nodes that attain the Goal. We then maximize the Goal value itself, under constraints that the number of nodes that attained the Goal does not decrease.

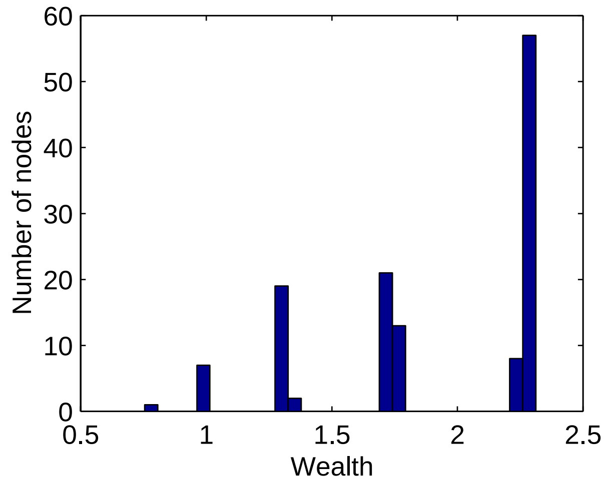

Figure 4.13: Histogram for wealth at the planning horizon, $T = 7$

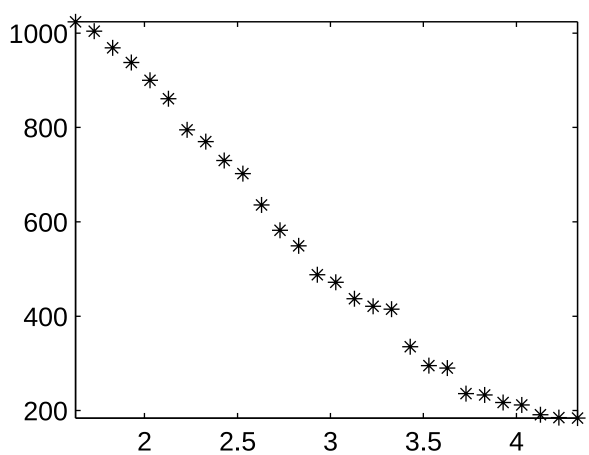

Figure 4.14: Number of nodes at which target wealth is achieved for a range of target wealth values

4.3.5 Variant of Probability of Reaching a Goal

A variant of the previous utility may be when the investor also cares for the downside risk besides caring for reaching a target wealth level. As before, in a discrete model for asset-returns this utility is a discontinuous one; hence the problem needs to be formulated so that the objective becomes a continuous function. We adopt a similar strategy to that described earlier: we adopt an investment strategy such that the target value, *Goal*, is achieved at a certain number of nodes at the planning horizon, along with no node being below a lower bound, say L.

A good starting point is again crucial. We modify the two strategies from the previous section for this problem as

Strategy 3 Following strategy 1 the investor initially invests only in the risky asset unless a transition node is reached. At the transition node part of wealth is invested in the riskfree asset and at all subsequent nodes all wealth is invested in the riskfree asset. A transition node on the downside is the node where the wealth level becomes low enough that investing in the riskfree asset will barely make it stay above L at the planning horizon. Therefore, when a transition node is hit on the downside, the investor invests only in the riskfree asset in all subsequent nodes.

Strategy 4 Following strategy 2 the investor initially invests both in the risky and the riskfree asset unless a transition node is reached. The strategy is then identical to that of strategy 3.

Therefore, there are now two types of transition nodes: one is identical to that described in strategy 1 and the other is one at which continuing investing in the risky asset can no longer ensure the final wealth to satisfy its lower bound. As with the first type of transition node, the investment is split between the two asset-classes and at all subsequent nodes all investment is in the riskfree asset.

After adopting one of the above two strategies a set of nodes at the planning horizon is at the Goal level and none of them is below L. We next pick the highest non-Goal node and maximize the wealth level at the selected node, along with constraints that the nodes that already achieved the Goal stay at Goal and none goes below L. Optimization could be done on a sub-tree or on the entire tree.

As before, another variant of the above problem is that the investor would like to maximize the Goal value once a certain number of nodes at which it is attained is figured out. For this problem we again adopt one of the above two strategies and perform the above mentioned optimization on the number of nodes that attain the Goal. We then maximize the Goal value itself, under constraints that the nodes that attained the Goal stay so and no node at the planning horizon goes below the level L.

4.4 Verification of Results from Standard Utilities

Research in finance and economics has primarily focused on utilities in the normative class. In order to validate our modeling and solution method, we applied our method to utilities in the normative class and checked if our results qualitatively match those of others. We begin with the constant relative risk aversion utility.

Power Utility

Absolute risk aversion, $A(W)$ as defined by Arrow and Pratt, is given as

$$\frac{-U''(W)}{U'(W)},$$

for twice continuously differentiable utilities. It is a measure of aversion to risk for different values of wealth. On the other hand, relative risk aversion, $R(W)$ is defined as

$$\frac{-WU''(W)}{U'(W)},$$

and is a measure of aversion to risk relative to the wealth level (see Huang and Litzenberger [19]). As the name suggests, relative risk aversion is a constant for constant relative risk aversion(CRRA) utilities, i.e., aversion to risk decreases with wealth.

Power utility is a constant relative risk aversion utility. In continuous time, given two asset classes with prices governed by a constant coefficient stochastic differential equation, the optimal allocation is given as :

$$X_t^* = \frac{\alpha - r}{\sigma^2(1-\gamma)},$$

where r is the riskfree rate, α, σ are coefficients of the stochastic differential equation governing asset-price dynamics and $1-\gamma$ is the relative risk aversion (see Merton [41]). Therefore, the optimal allocation for the power utility is a constant over time. Optimal allocation from our computations is a constant over all time-periods, $X_{t,n} = 1$ for all t and n. This is consistent with the continuous-time result, since in our model $r = 1.05$, $\alpha = 1.11$, $\sigma = 0.15$ and $1-\gamma$ is taken as 1.5. Therefore, the optimal constant allocation to the risky asset should be $\frac{0.06}{0.15^2(1.5)} = 1.78$. Due to the bounds imposed on the decision variables, our solution is $X_{t,n} = 1$.

Logarithmic Utility

Logarithmic utility is a special case of the power utility with the characteristic that zero wealth level elicits infinite displeasure (it is not defined for negative wealth levels). The results are similar to that of a power utility.

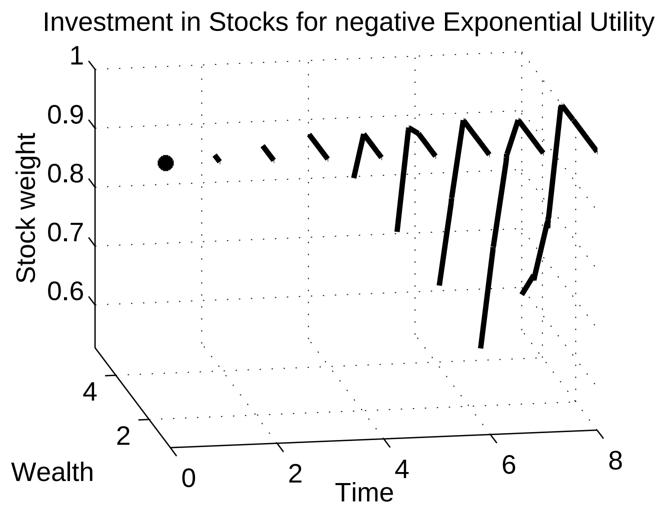

Figure 4.15: Optimal Stocks investment weights for a 9 periods problem using Negative Exponential utility, plotted by wealth levels at nodes for each time period

Negative Exponential Utility

The (negative) exponential utility belongs to the constant absolute risk aversion class (CARA), which implies that with increasing levels of wealth the aversion to risk remains the same. The utility asymptotes to a horizontal line as wealth goes to infinity. Continuous-time result for the negative exponential utility is

$$X_t^* = \frac{\alpha - 1}{\eta r \sigma^2 W_t},$$

where η is the absolute risk aversion. Therefore, the optimal dollar amount invested in stocks is a constant. Our results for a 9 periods problem are in figure (4.15). The optimization is done with constraints that allow no short selling, hence the saturation at low wealth levels. There are not as many points visible in the plot as there are number of nodes in each time-period of the tree, because several of the points overlap. This is because wealth levels are the same at these nodes.

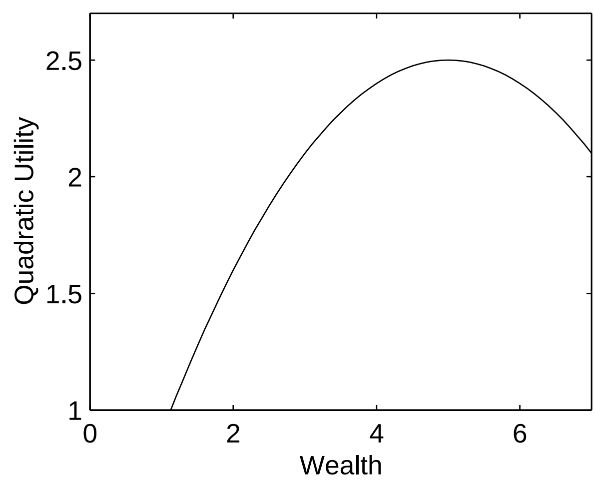

Figure 4.16: Plot of quadratic utility against wealth

Quadratic Utility

Quadratic utility, which is a concave, quadratic function of wealth, has the undesirable property of saturating at a wealth level that corresponds to the maxima of the function, after which there is dis-utility for more wealth (see figure (4.16)). This property of the quadratic utility seems somewhat counter-intuitive.

One way to remedy this property is to replace the decreasing part of the quadratic by an increasing linear function. This modification makes the utility resemble the α-t semivariance type utility studied before. Results from our computations using the quadratic utility can be seen in figure (4.17).

Kritzman's Combination Utility

This utility is a combination of the power and logarithmic utility; more specifically, the utility function is $U(W) = \log W - \frac{1}{W}$. A plot of the utility function is given in figure (4.18). Results for the Kritzman's Combination Utility can be seen in figure (4.19).

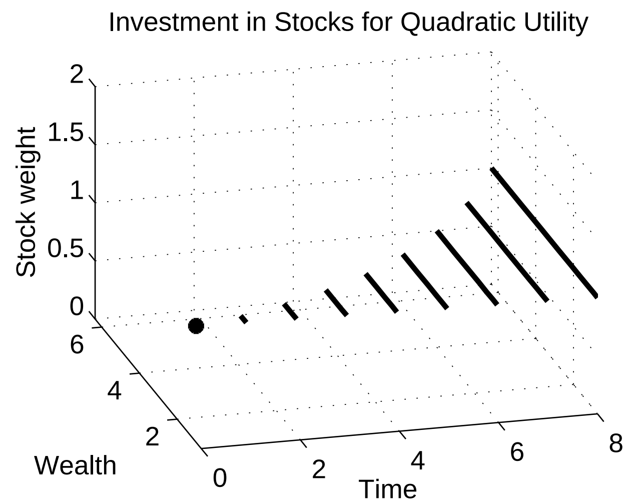

Figure 4.17: Optimal Stocks investment weights for a 9 periods problem using quadratic utility, plotted by wealth levels at nodes for each time period

4.5 Comparisons and Conclusions

Using our formulation for the investment problem it is possible to obtain the optimal investment decisions for several different types of utility functions, results for some of which were presented in the previous sections. We solved the problem and presented results for some behavioral utilities and some of their variants. This was followed by results for some more traditional utilities from the normative class of utilities.

We now make some general remarks comparing the characteristics of the optimal investment decisions for utilities in the two classes, normative and descriptive. The two classes are large and contain rather diverse utilities, so it is not possible to make very rigorous comparative statements. Among the behavioral utilities it was observed that there were considerable shifts from investing all the wealth in the risky asset, i.e., at several wealth levels investments in the riskfree asset was favored. However, among the normative utilities, except for the negative exponential utility, investments

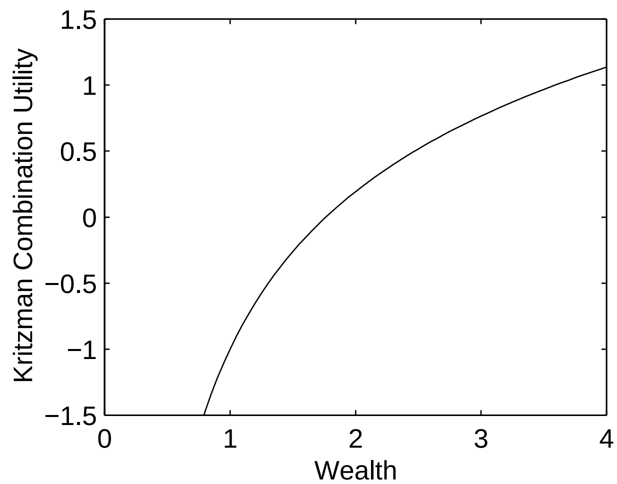

Figure 4.18: Plot of Kritzman's combination utility

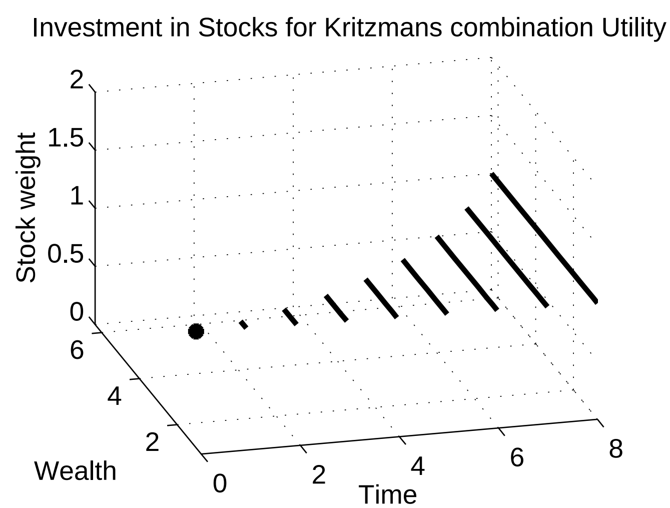

Figure 4.19: Optimal Stocks investment weights for a 9 periods problem using Kritzman's combination utility, plotted by wealth levels at nodes for each time period

only in the risky asset for all times and all levels of wealth was obtained as optimal.

In the problem formulation and results presented earlier we imposed bounds on the investment decision variables allowing no short-selling or borrowing of the asset classes. It is interesting to solve the problem for both normative and behavioral utilities and observe the impact of relaxing these bounds so that some amount of short-selling is permitted. The results in this case for the referential behavioral utilities, such as the loss-aversion utility, piece-wise linear utility etc., have similar characteristics as before. The V-shaped "dip" in the allocation to the risky asset that appears at a range of wealth level persists. Outside the dip, the "saturation" level is now at the new upper bound of the investment decision variables. For the normative utilities such as power or logarithmic utility, consistent with the analytic results, the optimal investment decision is a constant for all times and wealth levels.

It was observed earlier that the exercise of obtaining optimal investment decisions for the behavioral utilities may be viewed as a prescriptive step. This is, however, a very subjective analysis. To someone whose preferences are represented by the Prospect theory loss-aversion utility the corresponding optimal investment decisions make look reasonable and satisfactory, while some others may feel uncomfortable about, for instance, investing all the wealth in the risky asset at low levels of wealth. As a consequence, the subsequent prescriptive steps will continue to be subjective.

4.6 Contribution Schemes

In the results seen so far the optimization problem for the investment problem was formulated with the assumption that the investor deposits all the wealth she will ever invest in the retirement fund at $t = 0$, i.e., at the beginning of the planning period. This is clearly far from reality, since in reality people make periodic contributions to the retirement fund. We now look at this version of the problem.

Greatest degree of flexibility in modeling the contributions will be to make them both state and time dependent variables. This will double the total number of decision variables in the problem. It is reasonable to impose an upper bound on the contribution variables, since there are restrictions in practice on how much one can contribute yearly to a retirement fund. It is also reasonable to expect most utility functions to be increasing in wealth, since it is highly unlikely that one should prefer to have less wealth rather than more. As a result, the contribution variables will always be on their upper bounds in the solution of the optimization problem. This suggests that there is not much advantage in treating them as variables, and one only needs to select sensible levels of upper bounds for contribution levels in formulating the problem. In other words, it suffices to model the problem with a fixed level of contribution amount for each time period of the problem.

The cases where the contribution variables may not be on their upper bound are either when the utility function is not an increasing function of wealth or when the problem is formulated within a holistic financial planning framework, i.e., one in which the planner considers trade-offs between contributing to retirement fund and spending the wealth now. In the latter case the objective of the problem will also have terms that represent utility from spending the wealth now along with utility derived from saving and investing the wealth towards retirement.

4.6.1 Problem Formulation for Contributions

The problem is now formulated as follows. The contribution levels for each time period is taken as $K_{t,n}$ and the upper bound on this variable is denoted by $\bar{K}_t$. The transfer of wealth equations become

$$X_{t,n_1}(W_{t,n_1} + K_{t,n_1})S_u + (1 - X_{t,n_1})(W_{t,n_1} + K_{t,n_1})R_f = W_{t+1,n_2} \qquad (4.6)$$

$$X_{t,n_1}(W_{t,n_1} + K_{t,n_1})S_d + (1 - X_{t,n_1})(W_{t,n_1} + K_{t,n_1})R_f = W_{t+1,n_3}, \qquad (4.7)$$

where W_{t,n_1} denotes the wealth in the retirement fund, and K_{t,n_1} denotes the dollar contribution made at time t and node n_1 to the retirement fund. The optimization problem becomes

$$\begin{aligned} \max_{X_{t,n},K_{t,n}t=0,T-1,n=1,N} \quad & E_0[U(\tilde{W}_T] \\ \ni 0 \leq X_{t,n} \leq 1, \quad & \forall t,n \\ 0 \leq K_{t,n} \leq \bar{K}_t \quad & \forall t,n. \end{aligned}$$

This problem formulation is very similar to the consumption problem that will be studied in greater detail in the next chapter. The only major difference between the two problems is that here the objective contains no utility term for the contributions whereas in the consumption problem the objective primarily consists of utility from consumptions.

4.6.2 Results

Following the argument presented earlier regarding the contribution levels ending up on their upper bounds when the problem is solved for a utility function that is increasing in wealth, we modify the problem as follows. We fix the contribution levels at their problem upper bounds and solve the following optimization problem.

$$\begin{aligned} \max_{X_{t,n},t=0,T-1,n=1,N} \quad & E_0[U(\tilde{W}_T] \\ \ni 0 \leq X_{t,n} \leq 1, \quad & \forall t,n \\ X_{t,n_1}(W_{t,n_1} + \bar{K}_t)S_u \quad + \quad & (1 - X_{t,n_1})(W_{t,n_1} + \bar{K}_t)R_f = W_{t+1,n_2} \\ X_{t,n_1}(W_{t,n_1} + \bar{K}_t)S_d \quad + \quad & (1 - X_{t,n_1})(W_{t,n_1} + \bar{K}_t)R_f = W_{t+1,n_3}. \end{aligned}$$

In figure (4.20) the results are plotted for the prospect theory loss-aversion utility in a three dimensional plot by the wealth level at each node; wealth level is taken as net

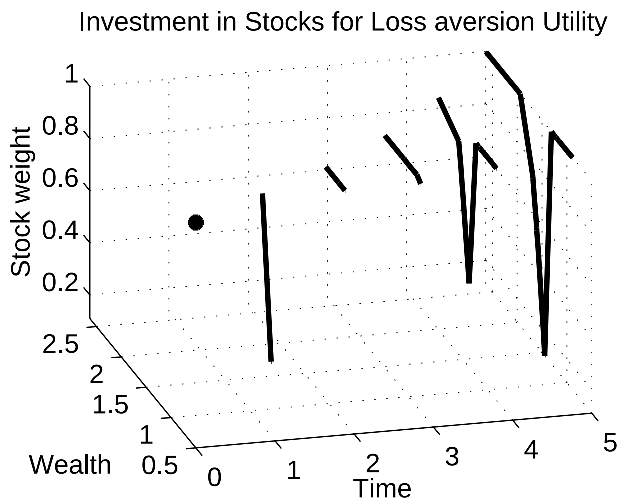

Figure 4.20: Optimal Stocks investment weights for a 6 periods problem using loss aversion utility, plotted by wealth levels at nodes for each time period

of the contribution made at the node.

Here the starting wealth is taken to be 0.5 units of wealth, along with 0.1 units of wealth contributed in each subsequent period. The pattern for optimal investing in the later periods is very similar to what was seen in the non-contributions case, but in earlier periods, specifically for $t = 0$ and $t = 1$, there is considerable difference in investment levels. The investment shift from investing in stocks because the wealth levels in the retirement fund are comparatively low towards the beginning of the planning period.

Chapter 5

Consumption Problem

5.1 Introduction

The consumption problem refers to the problem of finding the optimal consumption strategy for wealth after retirement. An accompanying, supporting optimal investment strategy also needs to be computed. Most researchers have studied this problem in an integrated fashion with the investment problem. We choose to study it separately since in the retirement planning framework, the two stages, namely the time in life when one plans one's retirement and the time when one is retired have very different characteristics, concerns and preferences. A key difference between pre-retirement and post-retirement decisions is that post-retirement decisions have a much more immediate impact.

Achieving retirement goals requires a well-defined plan. The most important considerations for this planning would be:

- when one retires or the retirement age,
- expected lifestyle and related expenses after retirement,
- what is one's life expectancy,

- what different sources of income one has after retirement,
- marital status,
- and whether or not there is a bequest motive.

Clearly, a lot depends on how early and how well an individual plans for one's retirement and accumulates the retirement capital. The consumption problem assumes that the planner has done the best in her capacity to accumulate the retirement funds, and addresses how the individual should consume the funds year by year after retirement and how the remaining funds should be invested in order to keep a continuous, assured flow of income during the retirement years.

Modeling life expectancy is a critical aspect of the problem. A related issue is a retiree's health. This may impact both consumption due to health care costs and a retiree's ability to enjoy life. The uncertainty related with health and life is at the largest in one's life at this stage and modeling of the problem should take this into account. Most other considerations listed above for retirement planning can be taken as parameters of the model to an extent– for instance, marital status can be accommodated by considering larger consumption volumes; a similar adjustment could be made for lifestyle and related expenses. The problem is formulated very similar to the investment problem as

$$\max_{X_t, C_t, t=0,T-1} \sum_{t=0}^{\tilde{T}} E[U_t(W_t, C_t)]$$

with the difference that the planning horizon, $\tilde{T}$, is not known with certainty and there is an extra set of variables, C_t, representing consumption over time.

As was mentioned earlier, most work on this problem has been in an integrated framework along with the investment problem. We give a brief overview of these approaches. Bodie, Merton and Samuelson [8] consider a life-cycle model with an optimal investment choice, choice of labor and consumption over the life-time of a planner. The problem is posed as an optimal control problem in continuous time and solved by assuming the wages to be perfectly hedged. Relations are drawn between the optimal investment choice and human capital of the planner over the planning period. Campbell and Viceira [10] consider an infinitely-lived investor with the Epstein-Zin-Weil utility (a recursive generalization of the power utility) and solve the investor's optimal consumption/investment problem. They replace the first-order conditions (Euler equation) and budget constraints of the problem with approximate equations and explore an analytical solution for the approximate problem. They analyze trends such as optimal growth rate of consumption, optimal consumption-wealth ratio, etc. Viceira [57] extends Campbell-Viceira by including an undiversifiable labor income process to the problem. Investors are infinitely lived, but have a certain probability, π_r of transition to a retired state. After reaching a retired state investors are assumed to stay retired thereafter. The power utility for consumption is assumed and the solution technique is similar to Campbell-Viceira. Analysis is done on the effect of retirement on the portfolio choice.

There are several other researchers who have performed empirical analyses to understand retirement behavior among individuals. We briefly outline this work. Hurd and McGarry [24] study the evolution of subjective survival probabilities and their ability to predict actual mortality. They find that subjective probabilities reasonably predict actual survival. In Rust [47] a dynamic programming based model for retirement behavior is developed and solved using a nested fixed-point algorithm. At each time step the worker chooses how much to consume and whether to work full-time, part-time, or exit the labor force. The objective is to maximize expected discounted utility over the worker's remaining life-time. Lumsdaine, Stock and Wise [33] study the predictive performance of three models of retirement behavior - "option value," dynamic programming, and probit. They advocate the use of a simpler model with good predictability, and find that in this sense the "option value" model does better than the dynamic programming one.

5.2 Problem Set-up

The problem set-up is similar to the investment problem. Investment decisions of how much to invest in the two asset-classes, risky and riskfree, need to be made. Returns from the risky asset are modeled using the binomial-tree model. There are new variables for the consumption decisions over time, denoted by $C_{t,n}$. $C_{t,n}$ represents the dollar amount withdrawn from the retirement fund at time t and node n for consumption during the subsequent period. The planner derives utility from consumption, hence the objective of the problem is to maximize the expected utility of consumption. Since there is uncertainty in the planning horizon of the problem, the planner is likely to prefer consuming earlier in the planning period rather than delay it for later use. The planner may have plans for activities, such as traveling after retirement, which they are likely to be more fit to do early in retirement rather than later for health related reasons. This may be another motivation for preferring early consumption. To model the preference for early consumption, a time-preference parameter, δ is introduced. Utility for consumption at time, t, is taken to be $\delta^t U(C_t)$, where δ lies in the interval $(0, 1)$. As a result, utility from consumption in periods farther out in the planning period get incrementally discounted. There may also be a bequest term in the objective, which is taken as the utility for wealth remaining at the planning horizon. The objective has several additional terms and is more complicated than before. At each node the wealth remaining in the retirement fund, after the consumption money is set aside, is invested in the two asset-classes. The transfer of wealth equations for the consumption problem become:

$$X_{t,n_1}(W_{t,n_1} - C_{t,n_1})S_u + (1 - X_{t,n_1})(W_{t,n_1} - C_{t,n_1})R_f = W_{t+1,n_2} \quad (5.1)$$

$$X_{t,n_1}(W_{t,n_1} - C_{t,n_1})S_d + (1 - X_{t,n_1})(W_{t,n_1} - C_{t,n_1})R_f = W_{t+1,n_3}, \quad (5.2)$$

where equation (5.1) is for an "up" and equation (5.2) for a "down." X_{t,n_1} is taken to be the fraction of wealth at time t and node n_1 invested in the risky asset and $(1-X_{t,n_1})$ the fraction invested in the riskfree asset. This is assuming that the investor invests all the wealth at each node after setting aside the consumption money. Nodes n_2 and n_3 at time $t+1$ are connected to node n_1 at time t. C_{t,n_1} is the dollar amount of wealth withdrawn from the investment fund for consumption between time t and $t+1$ at node n_1. The optimal asset allocation problem is written as:

$$\max_{X_{t,n}, C_{t,n}, t=0,\tilde{T}-1, n=1,N} \sum_{t=0}^{\tilde{T}-1} E_0[\delta^t U(C_{t,.})] + E_0[\delta^{\tilde{T}} U(W_{\tilde{T},.})]$$

$$W_{\tilde{T},m_1} = X_{\tilde{T}-1,m}(W_{\tilde{T}-1,m} - C_{\tilde{T}-1,m})S + (1 - X_{\tilde{T}-1,m})(W_{\tilde{T}-1,m} - C_{\tilde{T}-1,m})R_f,$$

where S takes the value S_u if the transition from m to m_1 is an "up" and the value S_d if the transition is a "down." $U(.)$ is the utility for consumption and for wealth at the planning horizon, δ is the time-preference parameter for consumption, and $\tilde{T}$ is the random planning horizon. We have taken the utility for consumption and wealth to be the same, although these may in some cases be different. For instance, if there is no bequest motive, there may be a dis-utility for a positive wealth remaining at the planning horizon. Corresponding changes in the objective are easy to make.

In the above optimization problem formulation we also need to impose bounds on the decision variables. As in the investment problem, the problem is formulated so that short selling is not permitted, i.e., borrowing of assets is not allowed. This restriction translates to bounds on the investment decision variables, namely $0 \leq X_{t,n} \leq 1$. The restriction reflects regulatory constraints on positions the planner can take on the asset classes. If some short selling is allowed, it would not seriously alter either the nature of the problem or the method for solving it. One set of sensible bounds on the consumption decisions would be $C_{t,n} \geq 0$. Although a lower bound of 0 is meaningful for the consumption variables, we believe that a more strict bound may be desirable. The lower bound should reflect the minimum amount one requires to maintain a normal subsistence level. With this in view, the lower bound will depend on the life-style, standard of living, or marital status parameters of the planner. The bound may be taken to be increasing with time to adjust for inflation, and also perhaps, for reasons such as increased health related expenses as one gets older.

We mentioned in the previous section that a critical modeling aspect of this problem has to do with the planning horizon of the problem. Although the planning horizon in the investment problem was not an absolutely certain event, the planner had a considerable amount of control over it. In the consumption problem, however, the planner has little or no control over the planning horizon. This makes it important to capture this characteristic of the problem in the modeling of the problem. One option for dealing with this uncertainty is to use the expected life-expectancy, $\overline{T}$, for an individual in place of the random value, $\tilde{T}$, and use other modeling techniques to accommodate some of the uncertainty related with $\tilde{T}$.

We propose to have non-uniform time periods for this problem, beginning with shorter periods that progressively become long as the planning horizon is reached. For instance, we may start with quarter or half-year long periods, and increase them to periods of two years or four years length. The discount factor gets accordingly adjusted by being raised to the appropriate power. This implies that for the later, longer periods, utility of consumption is much lower than that in the early periods. The problem is solved on a rolling basis; therefore it is possible to adjust for all the parameters of the problem according to how the outcomes turn out to be. One may wonder that imposing this decreased utility of consumption as time progresses will entail not having much money left for later consumption, in case one happens to live much longer than the expected life-expectancy. This is where the lower bound on the consumption variables for $t = 0 : T$ plays an important role; by making sure that one will be able to maintain the minimum desired living standards. Life expectancy is adjusted between rolling solves of the problem, and hence will be modeled accurately.

5.3 Computational Issues

The consumption problem stated above as an optimization problem is a nonlinear programming problem, similar to the investment problem. In the investment problem there were three major sources of complexity, size of the problem, nonlinearity of the objective and possibility of a less well-behaved utility function. In the original formulation the objective gradients were hard to compute. An alternate formulation of the problem, in which we treated the wealth level at each node of the tree to be a variable, made the objective gradient simple and easy to compute. An upshot of it, however, was we introduced as many nonlinear constraints in the problem as there are decision nodes in the tree. The consumption problem is formulated similarly. The complexities of the investment problem are enhanced in the consumption problem. There are an additional $2^T - 1$ consumption variables, $C_{t,n}$, which also appear nonlinearly in the objective.

If the expected life-expectancy of a planner is 25 years, we model a 25 years problem using non-uniform time-intervals, and span the 25 years using, say, 20 optimization problem periods. In a 20 period binomial-tree there are $(2^{20} - 1 =)\ 1,048,575$ nodes, and hence twice as many decision variables. All the variables have bounds, and there are an additional $1,048,575$ nonlinear constraints in the problem. The alternate formulation made the objective gradients easy to compute, but since the consumption variables appear in the objective, they are more complicated compared to the investment problem. The constraints also include the consumption variables, hence the constraint Jacobian matrix is more complicated and relatively dense than the investment problem. However, the matrix is still sparse and structured.

We used the SNOPT (Sparse Nonlinear OPTimizer) software [20] developed at the Systems Optimization Laboratory at Stanford University to solve the investment problem. SNOPT is very suitable for our problem, and yet not too specific for the current structure of the problem. In seeking a method to solve our problem, we did not want the method to be too specific to the models currently under consideration. What we wanted to establish is that a powerful general purpose algorithm could solve our current models. By using a general purpose algorithm we will not need to switch algorithms when additional complexity is added to the model. The consumption problem is slightly more complex, but the overall structure and properties of the problem are very similar to the investment problem, hence we use SNOPT to solve the problem. Exact function derivatives are provided to the solver.

5.4 Results

We now solve the problem using a specific utility function and present the results in this section. Consider the case when the planner invests only in the riskfree asset after retirement; this will guarantee a standard of living. Assuming that there is no bequest motive, this implies that she can consume at the highest possible rate such that no wealth is left at the planning horizon. To find the rate at which consumption is supported with this scheme, solve the following equations for C_f.

$$\begin{aligned} W_1 &= (W_0 - C_f)R_f \\ W_2 &= (W_1 - C_f R_c)R_f \\ W_3 &= (W_2 - C_f R_c^2)R_f \\ \ldots & \\ W_T &= (W_{T-1} - C_f R_c^{T-1})R_f \\ W_T &= 0 \end{aligned}$$

In the above equations it is assumed that consumption in every subsequent period increases at a rate R_c and the riskfree rate is denoted by R_f. Increase in consumption with time is justified to not only adjust for inflation (since consumption is taken in nominal terms), but also to take into account increased expected expenses, such as health related expenses, as one gets older. If we solve the consumption problem using a utility function that is increasing in consumption levels, along with the above rate of consumption as the lower bound on the consumption variables, it is no surprise that all the consumption levels will be at their bounds. We denote these bounds by $\overline{C}_t$ and call them the *reference lower bound*. For simplicity take $R_c = R_f$. Note that this lower bound is the highest feasible lower bound on the consumption variables.

Now consider the case when the planner invests both in the risky and the riskfree asset. The lower bound on the consumption variables is reduced from the reference level by a fraction ρ. The new lower bounds are $\rho\overline{C}_t$, where $\rho \in (0, 1)$. The consumption problem is solved with this reduced lower bound. We now apply this approach to a specific utility function.

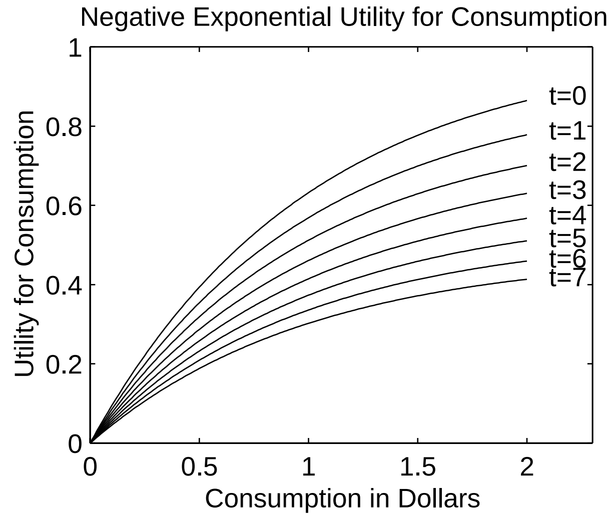

Figure 5.1: Negative exponential utility plot for a 8 periods problem

5.4.1 Negative Exponential utility

Negative exponential utility is a constant absolute risk aversion (CARA) utility. Absolute risk aversion, $A(W)$ as defined by Arrow and Pratt is given as

$$\frac{-U''(W)}{U'(W)}, \tag{5.3}$$

for twice continuously differentiable utilities. Absolute risk aversion is a measure of aversion for risk, and for the negative exponential utility is a constant for all levels of wealth. This implies that a planner's appetite for risk is a constant for whatever their level of consumption or wealth. Time preference parameter, δ, is taken to be 0.9. A plot of the utility for consumption for an 8 periods problem is given in figure (5.1).

We solve an 8 periods problem with the lower bound relaxed by taking the bound relaxation parameter, ρ equal to 0.8. In figure (5.2) a three dimensional plot of the optimal consumption variables is presented. In figure (5.3) the corresponding optimal

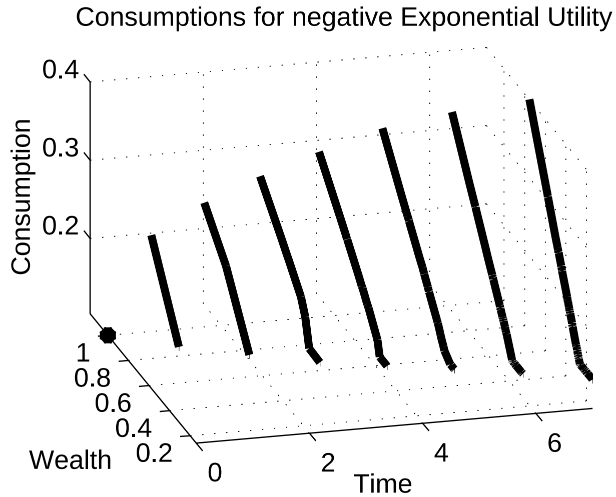

Figure 5.2: Optimal consumption decisions for a 8 periods problem for negative exponential utility

investment variables are plotted. Note that at several nodes the optimal investment is to invest 100% in the risky asset. A question one may ask is what is the benefit in term of consumption levels from taking these risky investment positions. One way to measure the "benefit" would be to see the number of nodes in the tree that are above the corresponding reference lower bound level. In figure (5.4) percentage of consumption variables that are above their problem bounds, $\rho\overline{C}_t$, and the reference lower bound $\overline{C}_t$ are plotted for each time period. The point of this comparison is to see the improvement in consumption from investing in the risky asset. Figure (5.5) is a plot of minimum, maximum and mean consumption levels for each time period for an 8 periods problem.

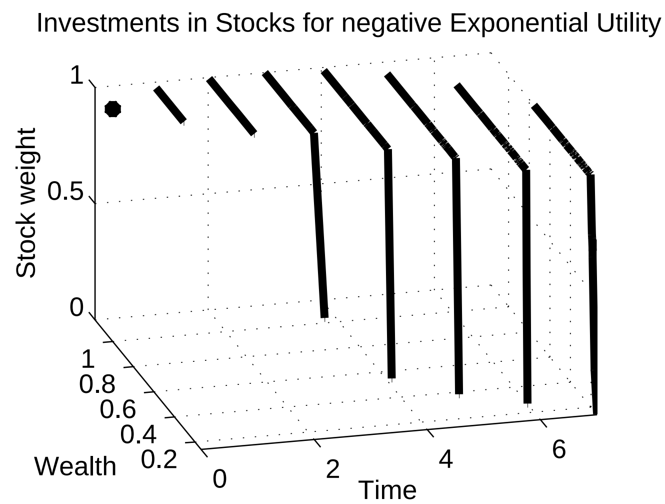

Figure 5.3: Optimal investment decisions for a 8 periods problem for negative exponential utility

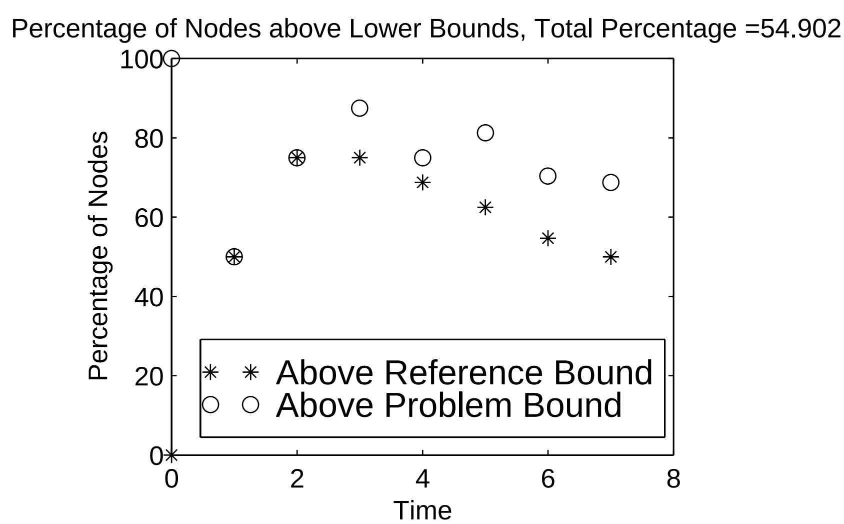

Figure 5.4: Consumption for each time, t above the problem bound $\rho\overline{C}_t$ and above the reference lower bound $\overline{C}_t$

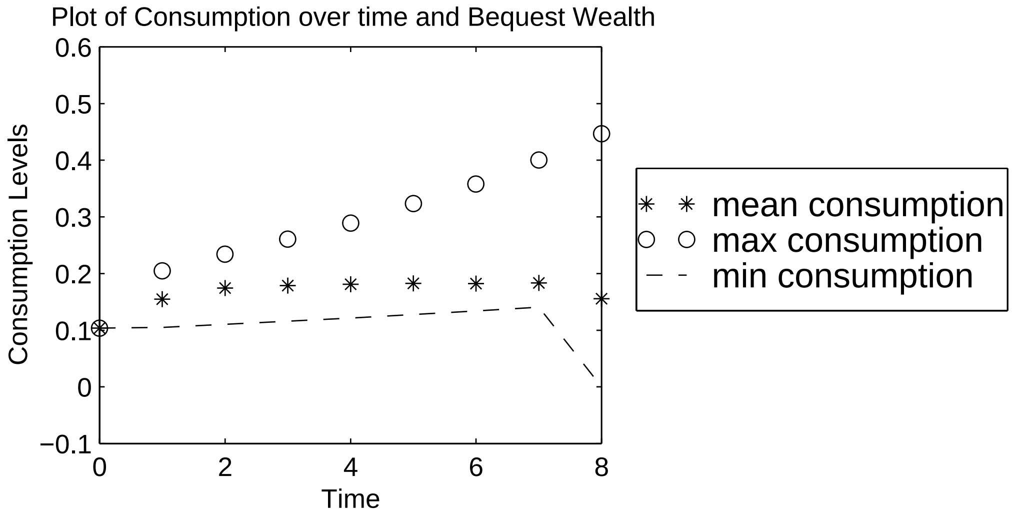

Figure 5.5: Minimum, Maximum and Mean Consumption level for each time, t.

The results for the consumption problem with the bound relaxation parameter taken as 0.8 indicated that approximately 55% of the nodes exceed the reference lower bound. We now address the question of how this percentage changes as the relaxation parameter varies in the range 0 to 1. Figure (5.6) is a plot of a range of values for the bound relaxation parameter and the corresponding percentage of nodes that are above the reference lower bound. The plot has an interesting characteristic, for a ρ value around 0.8 the percentage of nodes that exceed the reference lower bound peaks and then asymptotes to 35%, approximately.

We take a closer look at four levels for the bound relaxation parameter, ρ, in order to understand the characteristic of this plot. Figure (5.7) plots the consumption levels on the tree by their utility. The effect on the consumption levels is visible when the relaxation bound is lowered from 0.8 to 0.7. During this change the effect from the time-preference criterion for consumption overtakes the effect from the lower bound on consumption; as a result, consumption in the early periods is higher leaving less money for later consumption. As a consequence, the percentage of nodes that are

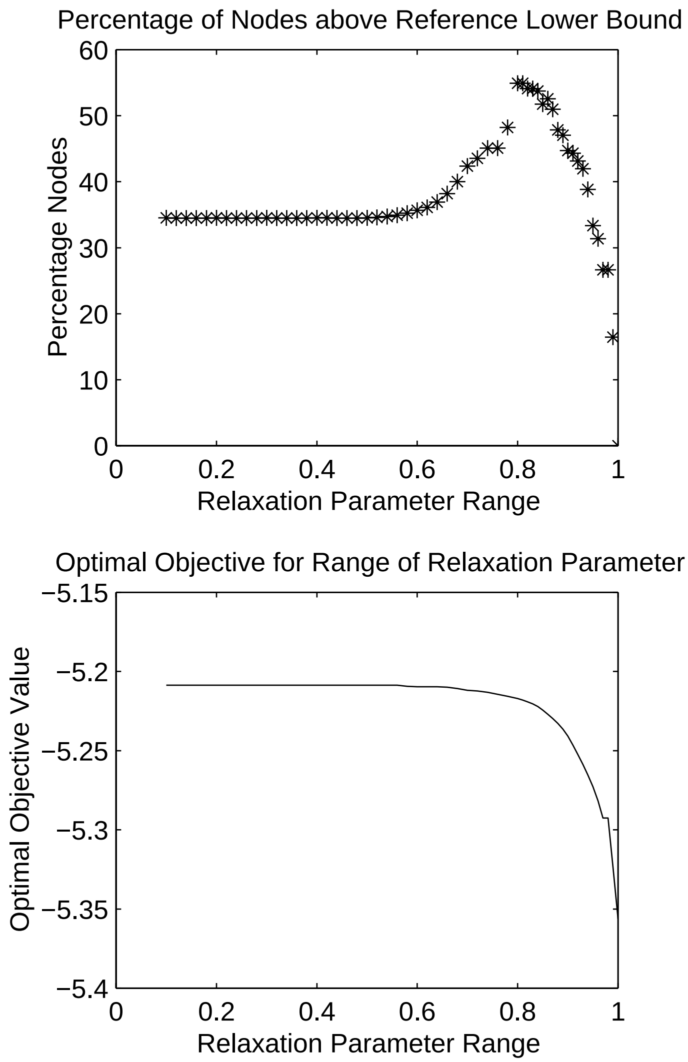

Figure 5.6: Number of nodes above the reference lower bound $\overline{C}_t$ for a range of problem bounds $\rho\overline{C}_t$, ρ ranging from 0.1 to 1.0

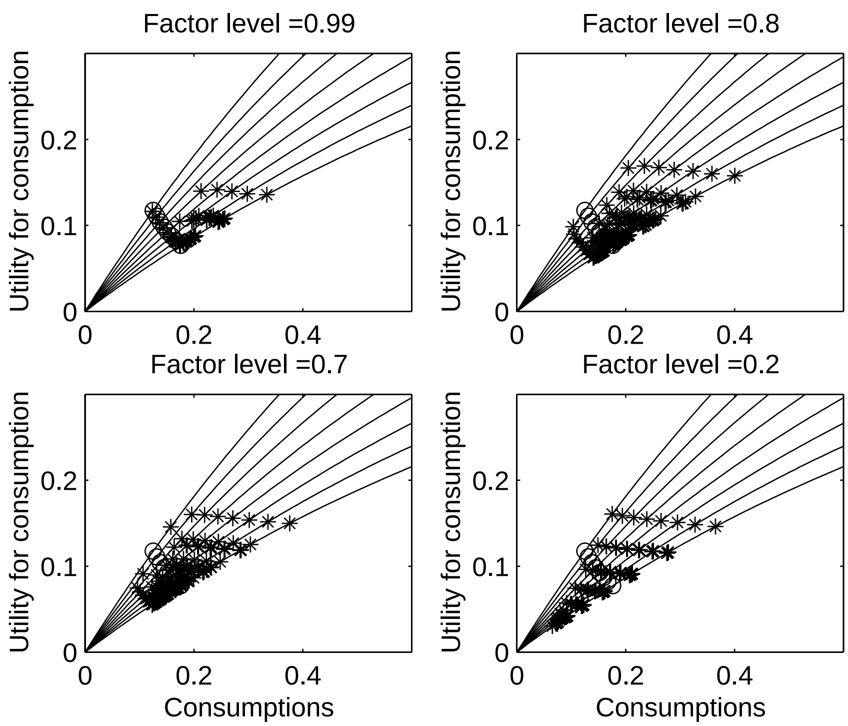

Figure 5.7: Plot of utility for consumption for 4 levels for the reference lower bound relaxation parameter, relative to the utility for consumption at the reference lower bound level

able to exceed the reference lower bound begins to decrease as the bound relaxation parameter is further decreased.

5.5 Conclusions

It is a better choice to invest in the risky asset with higher expected returns after retirement along with the riskfree investments. Although riskfree investing guarantees a level of consumption, exposing oneself – up to an extent – to a risky asset brings the opportunity of improving consumption substantially. Interestingly, the results show that reducing one's guaranteed income below approximately 80% of the assured income gives little additional benefit.

Even though the models used for the consumption problem discussed in the chapter were simple, some interesting characteristics of the problem are observed. We could further enhance the models by including a model for randomness in the planning horizon that better reflects the nature of the problem. Other extensions to this problem are discussed in greater detail in the final chapter.

Chapter 6

An Alternate Approach

6.1 Introduction

The problem of finding optimal investment decisions for long-term financial planning becomes computationally intractable in its most general form. This could either be due to the asset-dynamics models chosen or the properties of the utility function of the investors. Closed form analytical solutions are obtainable only in special cases. To solve this important problem for other cases, various simplifying assumptions are made. Some well-behaved and well- investigated utility functions are log utility (Mossin [42]), mean-variance (Markowitz [36]), Mean Absolute Deviation (Stone [52], Worzel et al [59]), and concave piece-wise linear utility (Dantzig and Infanger [13], Konno [29]). Asset dynamics have been simplified by discretizing the asset returns using multinomial trees (Mulvey and Ruszczynski [43]), network flows (Mulvey and Valdimirou [44]), and scenario based modeling (Rockafellar and Wets [45], Carino et al. [11]). Monte Carlo importance sampling is used to reduce drastically the effort needed to approximate a large multinomial tree (Dantzig and Infanger [13]).

In the previous chapters a binomial-tree based model was developed and optimal investment and consumption decisions were obtained for a more general class of utilities. In this chapter we present an alternate approach to finding the optimal investment decisions. For scenario based methods, including binomial trees, the number of decision variables grows exponentially with the number of time periods. Inspection of the optimal solution using these methods can, however, reveal a structure that can be exploited to simplify the problem. We develop an approach that exploits the observed structure in the solution by approximating the optimal decision in each period with a low dimensional parameterization. The problem is reformulated as a nonlinear optimization in the parameter space, its dimension increasing linearly with the number of time periods. This approach allows us to find optimal investments for a more general class of utilities than the standard concave ones.

6.2 The Asset Allocation Problem using a Behavioral Utility

In Chapter 1 the asset allocation problem was viewed as one involving choice, where the planner chooses the "best" mix of assets over time in their portfolio to meet a certain objective at the planning horizon. The preference relations were taken to be represented by a utility function, $U(.)$. In the past years there has been considerable interest in finance literature in utility functions studied in the psychology literature. These utility functions, referred to as *behavioral*, are a result of experiments conducted with people involving real or imaginary pay-offs. The subjects are presented with different choice options under different settings and inferences are drawn from their choice behavior. In the behavioral framework, Prospect theory, developed by Kahneman and Tversky [26, 55, 54], has earned particular attention. Essential characteristics of the prospect theory utility are :

Reference dependence The carriers of utility are gains and losses defined relative to a known reference point. People perceive and evaluate different choice options as a gain or a loss from a reference point.

Loss aversion The function is steeper in the negative rather than in the positive domain, implying that losses loom larger than corresponding gains. An amount of loss elicits more "unhappiness" than the same amount of gain elicits "happiness."

Diminishing sensitivity The marginal utility of both gains and losses decreases with their size. Gain or loss of an extra dollar results in a lesser impact as the amount of overall gain or loss gets larger.

The implication of these characteristics on the utility function are that it has an asymmetric S-shape that is concave above a reference and convex below it (see figure 6.1). The utility is referential in nature, which is the other contrasting property with the standard utility functions. Lack of concavity and differentiability of the prospect theory utility make computation of the corresponding optimal investment decision more challenging.

From a retirement planning point of view the problem addressed is to find an optimal asset-class mix for an investor's portfolio over a planning period so that the expected utility at the planning horizon is maximized. As before, the utility is taken to be for wealth, in particular for the wealth at planning horizon. This is a reasonable assumption, since in the retirement planning framework there is a tax penalty associated with premature withdrawals from the retirement fund. We will continue to consider a two asset-classes problem, namely, risky and riskfree. A more general problem may be considered, for instance, one in which utility is also for intermediate wealth and consumptions or when there are more than 2 asset-classes.

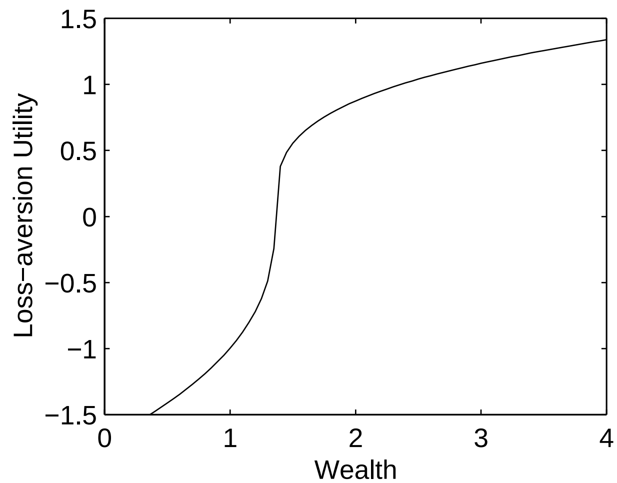

Figure 6.1: Prospect theory loss-aversion utility of wealth

The planning horizon for the problem is when the planner intends to retire, and it is assumed that the planner has a fair idea of when she plans to retire. We denote the planning horizon by T; therefore the planning period is from 0 to T, where 0 represents the beginning of the planning period. x_t is taken to be the fraction of wealth at time t, denoted by W_t, invested in the risky asset. (The notation is purposefully changed to make a distinction with the variables in the binomial tree approach.) Hence assuming that all the wealth at each time is either invested in the risky or the riskfree asset, $(1 - x_t)$ is the fraction of wealth invested in the riskfree asset. The problem is formulated as:

$$\max_{x_t, t=0,T-1} E_0[U(W_T)]. \tag{6.1}$$

Therefore, the problem is to find the optimal investment weight for the risky (and the riskfree) asset over the planning period so that the expected utility for wealth at the planning horizon is maximized.

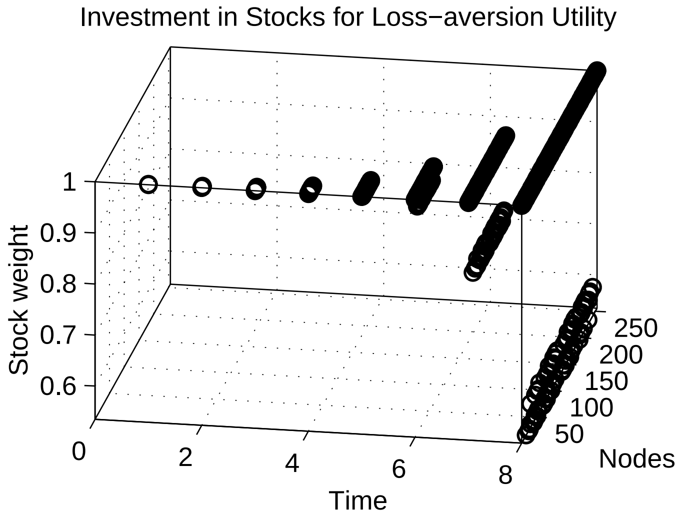

Figure 6.2: 9 Periods plot of optimal allocation in the risky asset against nodes in the binomial tree

6.2.1 Discovering the Structure in Solution

In chapter 4 the above problem was solved using a binomial-tree model for asset-return dynamics. In this set-up the risky asset return follows a simple random-walk; in each period there is an equal probability of either a high or a low return. The high-low return values are chosen to match the appropriate mean and standard deviation. An investment decision is made at each node of the tree, and is chosen so that the expected utility of wealth at T is maximized. In figures (6.2) and (6.3) some results from chapter 4 for the loss aversion utility are presented.

Optimal investment weights for the risky asset are denoted by $X_{t,n}$, where t denotes the time and n the ordering of nodes of the binomial tree as given in figure (2.2). The problem is solved taking the planning horizon, $T = 9$ and the reference point of the utility function is taken to be the wealth one would obtain if an 8% yearly return

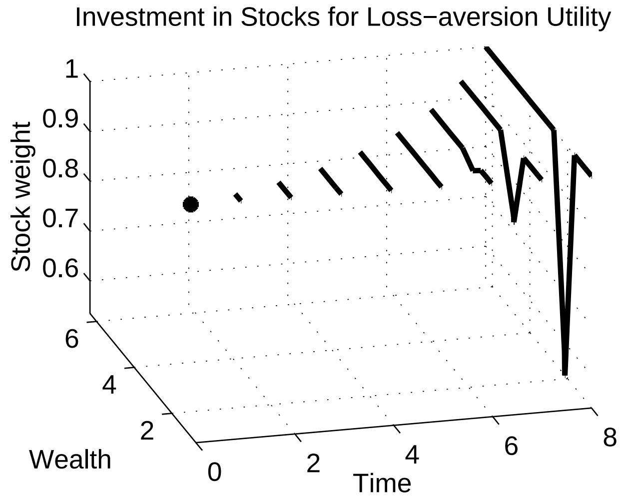

Figure 6.3: 9 Periods plot of optimal allocation in the risky asset against wealth levels for each time period

is obtained each year for the entire planning period. In figure (6.2) the investment weights are plotted by the nodes of the tree for each time period. The plot is not much more instructive than indicating that there are a small number of scenarios at periods $t = 4$ to $t = 8$ in which the optimal allocation in the risky asset is less than 100%. When the asset dynamics are "memoryless" the optimal allocation is expected to be state dependent, rather than path dependent. Figure (6.3) confirms this. It shows that not only is the optimal asset allocation a function of wealth, it is a very simple function of wealth.

6.2.2 The Structure

The structure observed above may be described as follows. At extreme low levels of wealth the optimal allocation to the risky asset is 100%. As wealth increases, the optimal allocation to the risky asset falls, then rises back to 100%. In a multi-period setting the "dip" in the allocation to the risky asset gets shallower and moves towards lower wealth levels as the number of time periods to the planning horizon increases (see figure 6.3). The dip eventually disappears completely, indicating that at sufficiently many periods prior to the planning horizon it is optimal to be fully invested in the risky asset at all levels of wealth. The dips also have a dependence on the reference point of the utility, as was observed in chapter 4, shifting together with the reference wealth level.

The "saturation" for extreme levels of wealth is due to the constraint imposed on the asset allocation variables to allow no short selling. Relaxing these bounds to allow some short-selling (borrowing of assets) retains the structure with a different saturation level. For instance, if a short-sell position could be taken up to the current wealth level, the relaxed bounds for allocation weights will be -1 and 2. The saturation level will now be at 200% level instead of the earlier 100% level.

This time varying dip is also observed for other utilities like the piece-wise linear, α-t target semi-variance (Fishburn [17], Adachi [1]), and other S-shaped smooth utilities. This implies that for a large class of utility functions, the optimal investment strategies may be both time and wealth level dependent. Hence, for utility functions in this class strategies that are time-invariant (buy & hold) or wealth-invariant will be sub-optimal.

6.2.3 The Central Idea : Use the Structure

In our approach for solving for the optimal investment strategy we make use of the above observed characteristic of the optimal asset allocation by parametrizing $x(t, W_t)$ with respect to W_t with a few parameters for each t.

The method is as follows. First solve a one or two period problem using any brute force technique, such as dynamic programming. This step may be required only once for a class of utilities that have similar functional form for optimal asset allocation. We then choose a set of approximating functions for the optimal solution as a function of wealth. The multi-period problem is solved by maximizing the expected utility of wealth at the planning horizon with respect to the parameters of the approximating function. This is a nonlinear optimization problem. Scenario generation for objective function evaluation is done using Monte Carlo simulation. In the next section an example problem is solved.

6.3 Applying Parametric Optimization

For illustration we apply the above outlined method to the prospect theory loss-aversion utility, along with bounds on the investment weights allowing no short-selling. The true solution for a range of starting wealths for the loss-aversion utility is given in figure (6.4). For each time period this can be approximated with a linear for the left edge of the dip and a quadratic function for the right edge (along with bounds that the allocation is bound above by 1). This entails 5 parameters (2 for linear and 3 for the quadratic) for each time-period $t = 0$ to $T - 1$. Therefore, the functional forms of the approximating functions are:

$$l(t, W_t) = a_1(t) + b_1(t)W_t \quad \text{(Linear)} \tag{6.2}$$

$$q(t, W_t) = a_2(t) + b_2(t)W_t + c_2(t)W_t^2 \quad \text{(Quadratic).} \tag{6.3}$$

The linear function $l(t, W_t)$ approximates the left edge of the dip and the quadratic function $q(t, W_t)$ approximates the right. It may be desirable to introduce some more bounds and constraints – for instance, the slope of the linear should be negative, while the quadratic should be a concave function. Optimize $E[U(W_T)]$ w.r.t. the $5T$ parameters. Note that the number of parameters, and hence variables in the optimization problem, grow only linearly with T.

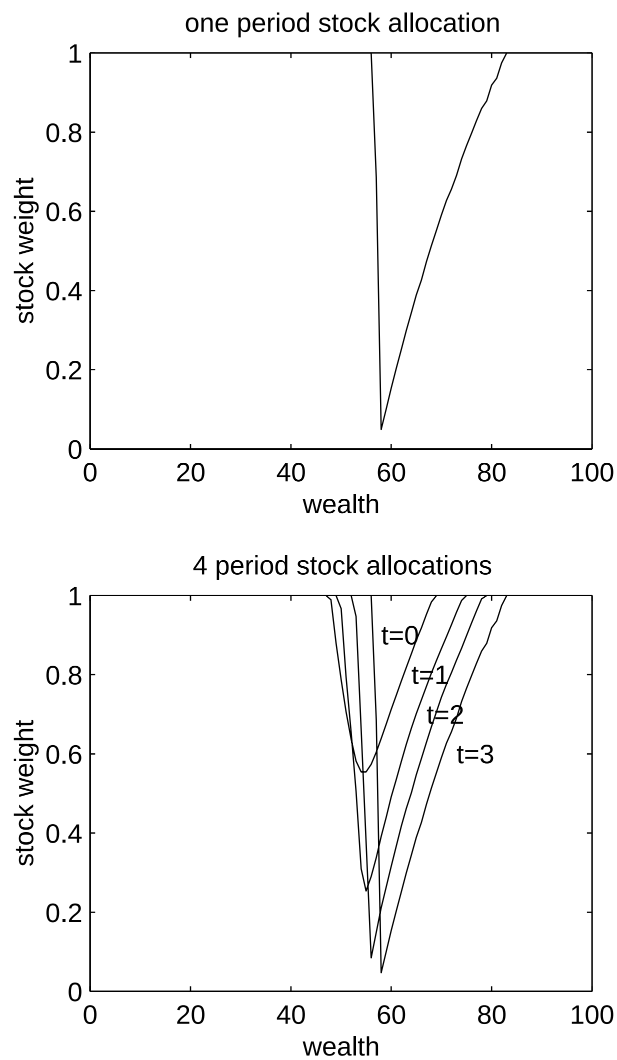

Figure 6.4: Optimal stock allocation for a one period and a four period problem as a function of wealth and time horizon

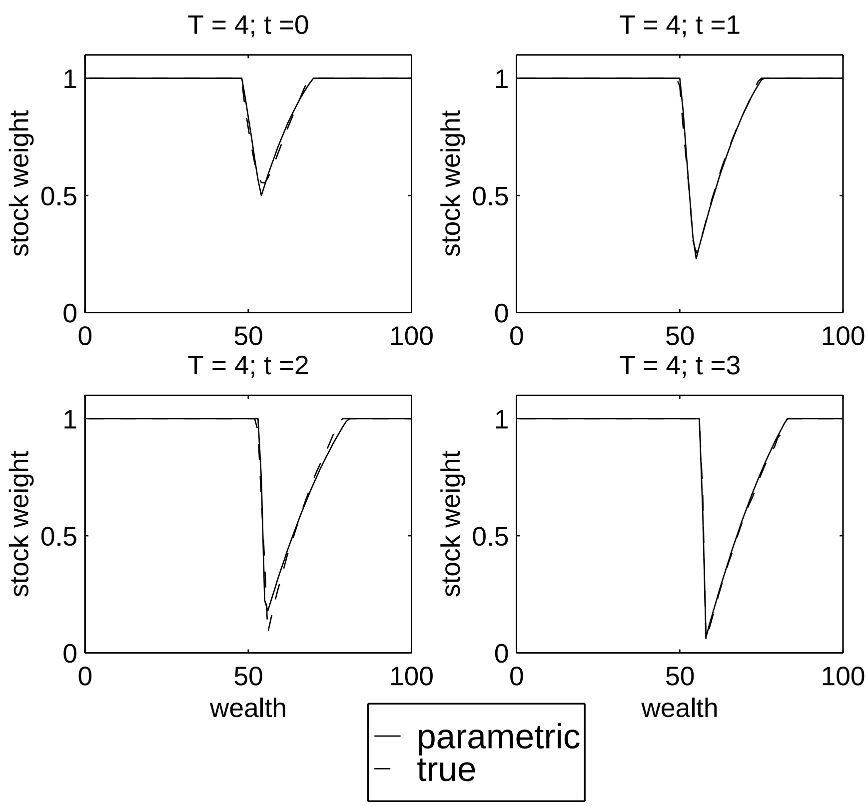

Figure 6.5: Comparison of the parametric solution with the true solution for a four periods problem

We set up the nonlinear programming problem with $(a_1(t), b_1(t), a_2(t), b_2(t), c_2(t); t = 0, T-1)$ being the variables of optimization as follows.

$$\max_{a_1(t),b_1(t),a_2(t),b_2(t),c_2(t);t=0,T-1} \quad E_0[U(W_T)], \tag{6.4}$$

$$x(t, W_t) = min(1, \quad max(l(t, W_t), q(t, W_t))), \tag{6.5}$$

$$W_{t+1} = x(t, W_t)W_t \quad \tilde{S} + (1 - x(t, W_t))W_t R_f, \tag{6.6}$$

where $\tilde{S}$ is the stochastic return from the risky asset and R_f is the return on the riskfree asset. Here the bounds on the variables are suppressed for simplicity. The problem is solved for the optimal investment decisions using MATLAB's nonlinear optimizer routine. As in most iterative procedures, getting a good starting point is critical for this problem. We adopted certain steps involving optimizing a subproblem, such as fixing some of the variables, to get a good starting point. See figure (6.5) for plots of parametric versus true solution for a four periods problem using the loss aversion utility. For function evaluation Monte Carlo simulation was used, taking the price process of the risky asset to be lognormally distributed, with mean, $\mu = 1.11$ and standard deviation, $\sigma = 0.15$. The price evolution is given as

$$S_t = S_{t-1} \exp^{(\mu + \sigma \tilde{N})}, \tag{6.7}$$

where $\tilde{N}$ is a standard normal. The riskfree rate, $R_f = 1.05$. A sample size of 10000 is used in the optimization.

6.4 Sensitivity to Asset-Dynamics Parameters

There is an inherent uncertainty associated with mean, standard deviation, etc., of future asset-returns. As a result, the apparent precision of a solution is somewhat deceptive since the problem may itself be viewed as not so well defined. In this section, we compare our approximation error (parametric solution relative to the true solution) to the problem specification error (due to uncertainty in the asset dynamics).

We obtain the true and the parametric optimal asset allocation for a certain set of values for the asset-dynamics parameters, R_f, μ, and σ. Let us call this true solution, the *reference*. If, however, these values are not known precisely, but say within 1% accuracy, we obtain a different set of true optimal allocations. These are plotted in figures (6.6) and (6.7) in the top plots for a 4 period problem and allocations at $t = 3$, by $\pm 1\%$ shift in the μ and σ values, respectively. The bottom figures in (6.6) and (6.7) are plots of parametric optimal along with the true curves corresponding to the $\pm 1\%$ shifts in value of μ and σ, respectively. Similar plots can also be obtained for $t = 0, 1, 2$ and for shifts in riskfree return, R_f.

Our approximate parametric optimal is well within the $\pm 1\%$ absolute shift of optimal asset allocation for μ, σ and R_f. And the combined effect of change in these asset-dynamics parameters will only result in greater shifts in the true solutions. Therefore, precision of the parametric solution is a problem only if the parameters for asset-dynamics are known with greater precision. If this is the case, one can select a different parametrization for a better approximation.

6.5 Conclusion

In this chapter we developed a method for obtaining near optimal multi-period investment decisions in a concise parametric form for a range of state variable levels. In Keane and Wolpin [28] the method suggested for making problems in discrete-time

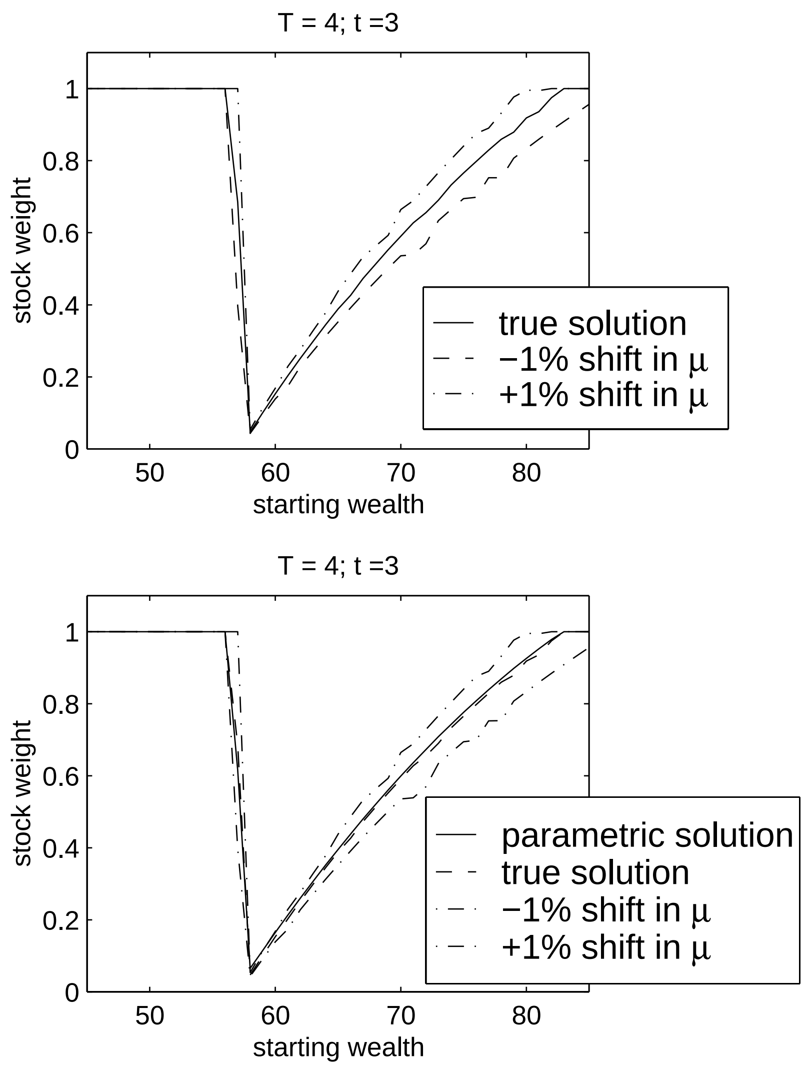

Figure 6.6: Sensitivity analysis: Variation in true solution due to 1% misspecification in the stock mean parameter, μ. The approximate parametric solution falls well within the error due to 1% misspecification of μ.

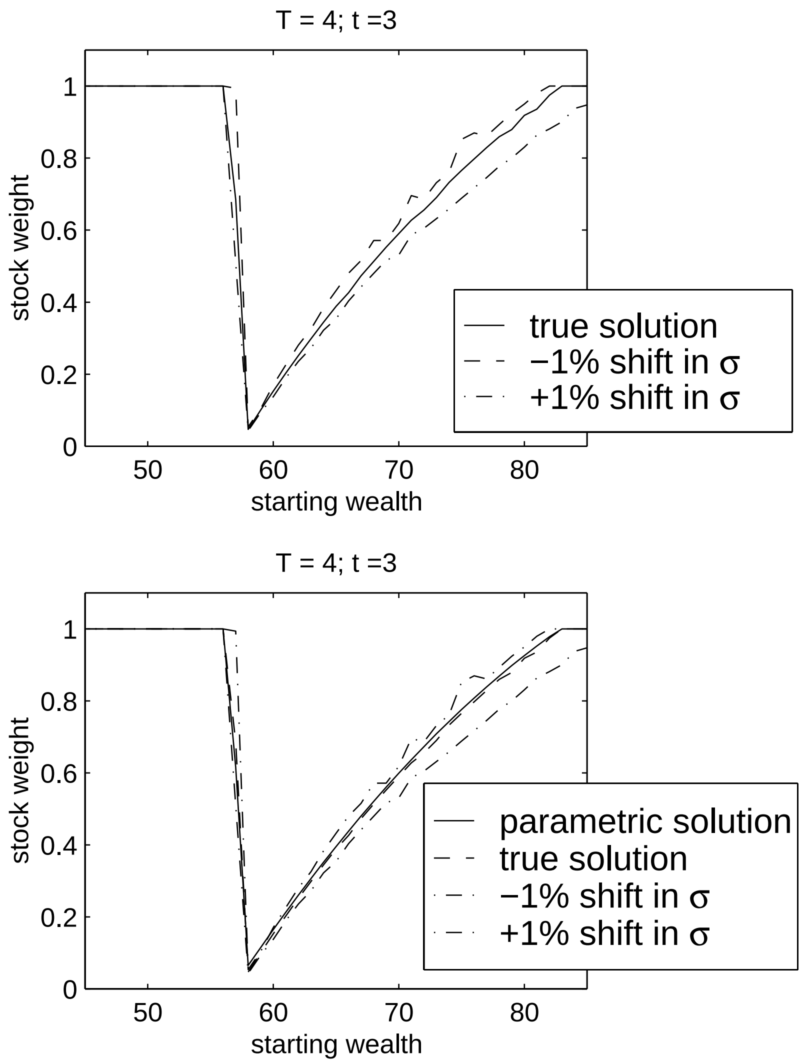

Figure 6.7: Variation in true solution due to 1% misspecification in the stock volatility parameter, σ. The approximate parametric solution falls well within the error due to 1% misspecification of σ.

stochastic control tractable is similar in flavor, although they simplify the computational requirement using a simulation and interpolation scheme for the value function of the problem.

In the binomial tree approach each $x(W_t)$ is a variable, resulting in an exponential increase in the number of variables as the planning horizon increases. In the parametric formulation, however, the number of variables increases only linearly with the planning horizon. An assumption in this approach is that the user has an *a priori* knowledge of the type of solution the problem may have in order to come up with a plausible set of parametric forms for it. But, this is not too demanding since a one or two periods problem is usually tractable using dynamic programming or other brute force methods.

Another attractive feature of this approach is greater freedom to model the asset-dynamics process. Correlations of the risky asset with the (relatively) riskfree one, or two risky assets (for example, small-cap/large-cap, stock/bond index etc.) can be easily accommodated. Either of which will square the complexity for a binomial lattice or a tree approach, but here the change will be only in the sample generation process; the number of variables in optimization problem will not be greatly affected.

In the parametric framework it is also possible to have multiple sources of risk which will get incorporated into the simulation process, not disturbing any other modeling aspects. Serial correlation in any of the price processes may introduce path dependence in the optimal asset allocations, thus adding some complexity to the parametrization. The approach can be extended to a problem with more than 2 assets (mutual funds, stock sectors etc.). This would make the problem greater than one-dimensional. Contributions and consumption can be introduced with some minor changes in the modeling. This method of parametrizing the action in terms of states of the world can also be used in other problems where the parametrization is possible.

Chapter 7

Future Work

A framework for financial planning is developed in the previous chapters, with a specific focus on retirement planning. The retirement related planning problem is viewed as comprised of two subproblems, the problem of acquiring wealth for retiring by saving and optimal investing, and the problem of optimal withdrawals of wealth from the retirement fund after retirement. There are several extensions possible to the problems and the framework developed so far. The extensions can be classified into the following categories.

Asset-return dynamics

Solution technique

Investment problem

Consumption problem

Usability of the tool

We now explore these extensions in greater detail.

7.1 Asset-return Dynamics Models

Assets available for investing were classified into two major categories, risky and riskfree. Returns on the risky asset were modeled using a binomial tree model, where in every period the returns were either high or low with equal probability. Returns on the riskfree asset are known with certainty for all time. In reality this may not be totally satisfactory, since every type of asset has some degree of uncertainty or risk associated with its outcomes. In order to incorporate this in the model, multiple branching at each node of the tree will be required, especially if the two asset classes are not perfectly correlated. The asset prices are known to mean-revert; this can be modeled by making the transition probabilities time dependent. For instance, if there is an “up” in a period, the probability of a subsequent “up” may be lower. Besides, there is also a need to allow asset returns to have more than two possible outcomes in each period; this is particularly important for the asset with a higher variance.

The planner may wish to classify assets into more than two asset classes, differing from each other in the levels of riskiness and returns. The categories may be of stock-bonds and cash, or of different mutual funds that the planner is required to select from. Modeling asset-returns for a greater number of asset classes will result in a greater degree of branching at each node of the tree.

Prospect theory, as discussed in earlier chapters, is developed by two psychologist, Daniel Kahneman and Amos Tversky. It is an alternate theory to the expected utility theory. One of the results of the theory is that people tend to transform probabilities in a nonlinear fashion when evaluating uncertain prospects. The low probabilities get overweighed while the higher ones get underweighed. In order to provide a comprehensive tool for use in a descriptive (behavioral) framework, we will need to extend our models so that this characteristic of people’s decision making is incorporated.

A direct implication of the extensions in the asset-return dynamics models suggested in this section is an exponential increase in the optimization problem size. In order to be able to address larger problems we will need to explore improved solution techniques, and this is the topic of discussion in the next section.

7.2 Solution Technique

In Chapter 3 solution method for the optimization problem arising in resource allocation for the investment and consumption problem was outlined. The optimization problem was formulated as a nonlinearly constrained optimization problem. Sequential quadratic programming algorithms are a robust and efficient set of algorithms to solve problems in this category. There was substantial sparsity and structure in our problem, and we used SNOPT software to solve the optimization problem. SNOPT (Sparse Nonlinear OPTimizer), developed at the Systems Optimization Laboratory at Stanford, is a sparse implementation of sequential quadratic programming algorithms. The current version of SNOPT does not use second derivative information for the functions of a problem. In our problem formulation, where the wealth level at each node of the tree is treated as a variable (along with a nonlinear constraint for each node to satisfy the budgetary constraints), the second derivative of the objective and constraints are easy to compute. By being able to use this additional information, we hope to achieve greater efficiency for solving the problem, and potentially allowing the capability of solving larger sized problems.

Another attractive feature of the optimization problem developed here is that it yields very well to decomposition techniques. At each node there are only three linking variables between the root of the tree and the subtree starting from the node, which makes it easy to decompose the problem into subproblems. Decomposition techniques adopted in collaborative or multidisciplinary optimization, used in applications such as aircraft design can be very beneficially implemented here.

There are 2^t subtrees starting at each node at time t. An optimization problem may be separately solved on each of the subtrees. Since there are only three variables in the base of the tree that are linked with any single subtree, matching these common (between the base and the subtree) variables with those at the base of the tree is easy. For several utility functions considered in previous chapters, it was optimal to invest 100% in the risky asset in the earlier periods. For these utility functions matching variables gets even more simplified.

7.3 Investment Problem

In the investment problem optimal decisions are computed for saving and investing to acquire retirement funds. The objective for the problem was taken to be the expected utility of wealth at the planning horizon, where the planning horizon was the time one expects to retire. Therefore, the planner evaluated her overall performance in terms of the wealth at the planning horizon. This may not usually be the case. According to prospect theory people like to evaluate outcome of their single decision, rather than outcome of a sequence of decisions. To implement this will imply that the objective of the optimization problem will not only involve expected utility of wealth at the planning horizon, but also that of wealth at all intermediate time periods. In the case of utility functions with a reference point, the reference point for intermediate time periods may vary with time.

In implementing any investment recommendations there are transaction costs involved that should be taken into consideration in the overall asset allocation problem. This is necessary for the recommendations to stay realistic. There is also a need for tax related considerations in locating and allocating wealth in different asset classes.

7.4 Consumption Problem

The consumption problem addressed the post-retirement resources allocation problem of how best to consume the wealth saved in the retirement fund, while optimally investing the remaining wealth. Uncertainty associated with planning horizon of the problem was addressed in modeling for the problem. Results were obtained for the negative exponential utility, a constant absolute risk aversion utility. Exploring the optimal consumption decisions when the utility for consumption is descriptive in nature is an interesting direction. Some additional questions that will need to be addressed with the behavioral utilities are how does the time preference criterion change and what is an appropriate reference point in the utility function for each time period. Transaction costs and tax related issues need to be addressed in this problem also.

7.5 Improving Usability of Tool

To provide a seamless planning tool, the pre-retirement planning and the post-retirement planning will need to be merged so that there is a smooth transition from one planning phase to the other. The problem will be viewed with three regions of planning, the first when the time of retirement is far out, and hence the main focus is on saving and investing. This will be followed by a region where retirement is near and the planner will continue to plan the saving and investing for acquiring adequate retirement funds, but also begin to plan the consumption process after retirement. And the third region will be the purely post-retirement planning region when the planner would have retired and will be concerned with the optimal consumption decisions and investment decisions for the remaining wealth in the retirement fund.

In order to make this tool usable, there is a need for a good user interface. The user interface will need to fulfill several functions. First, it will be necessary to help the planner figure out their utility functions or objectives for retirement planning. Second, good methods will need to be devised for communicating the results of the optimization problem to the user, so that its implications and consequences are understood. Third, some capability will need to be built for the planner to do analysis of the investment recommendations obtained from the optimizer. And lastly, if the tool is to be used as a prescriptive tool, some features will need to be developed that aid the re-assessment process of the planner's preference criteria.

Appendix A

Sequential Quadratic Programming Algorithms

A.1 Problem

A typical nonlinear optimization problem with nonlinear constraints may be written as

$$\begin{aligned} \mathbf{NLP} : \min_{\mathbf{x}} \quad & F(\mathbf{x}) \\ \ni \mathbf{c}(\mathbf{x}) \quad & \geq \quad \mathbf{0}, \end{aligned}$$

where $F(\mathbf{x})$, the objective function, is a nonlinear function, and $\mathbf{c}(\mathbf{x})$ is a m_N-vector of nonlinear constraint functions.

Most optimization algorithms are designed to try and satisfy the optimality conditions for optima. Before stating the necessary and sufficient conditions for optimality, define a local minimizer to the NLP to be a point x^* such that $\mathbf{c}(x^*) \geq 0$ and there exists a $\delta > 0$ such that $F(x) \geq F(x^*)$ for all x satisfying

$$\|x - x^*\| \leq \delta \text{ and } \mathbf{c}(x) \geq 0. \tag{A.1}$$

The necessary conditions for x^* to be a local minimizer are that there exist multipliers λ^* such that

$$\begin{aligned}
c(x) &\geq 0 && \text{(A.2)}\\
\hat{J}^T(x^*)\lambda^* &= g(x^*), \;\; [\text{or } \; Z^T g(x^*) = 0] && \text{(A.3)}\\
\lambda^* &\geq 0 && \text{(A.4)}\\
Z^T(x^*)W(x^*,\lambda^*)Z(x^*) & \quad \text{is positive semidefinite,} && \text{(A.5)}
\end{aligned}$$

where $g = \nabla F(x)$, $\hat{J}$ is the constraint Jacobian, $Z(x)$ is a basis for the nullspace of $\hat{J}(x)$, and $W(x,\lambda)$ is the Hessian of the Lagrangian $L(x,\lambda)$. These conditions are called the second-order KKT (Karush-Kuhn-Tucker) conditions. The first three are called the first-order conditions. If the last two conditions are replaced by

$$\begin{aligned}
\lambda^* &> 0 && \text{(A.6)}\\
Z^T(x^*)W(x^*,\lambda^*)Z(x^*) & \quad \text{is positive definite,} && \text{(A.7)}
\end{aligned}$$

the conditions become sufficient for optimality.

A.2 Introduction

Sequential quadratic programming (SQP) algorithms can be viewed as a generalization of Newton's method to the (inequality-) constrained problem. The basic idea is to construct a constrained quadratic model around the current iterate x_k. Minimization of this quadratic model gives an update p_k that approximates the error $x^* - x_k$. The new iterate is $x_{k+1} = x_k + \alpha_k p_k$, where α_k is usually determined be a line search. Trust-region methods may also be used.

In other words, SQP methods solve a sequence of quadratic programs (QPs). Each QP has linear constraints and a quadratic objective. The QP constraints are a linearization of the nonlinear constraints at x_k. It is important to note that the quadratic objective is an approximation to the Lagrangian $L(x, \lambda)$, and not to the objective function $F(x)$. The kth QP subproblem for NLP can be stated as

$$\begin{aligned} \mathbf{QP}: \min_{p} \quad & g_k^T p + p^T H_k p \\ \ni \quad & J_k p \geq -c_k, \end{aligned}$$

where $g_k = \nabla F(x_k)$, J_k is the Jacobian of c at x_k, and H_k is an approximation to the Hessian of the Lagrangian.

A.3 Merit Function and Line Search

Any algorithm for constrained optimization needs to ensure that the algorithm converges to a point that is both feasible and a constrained minimizer. Therefore, to measure how good a given point is, simply looking at the objective is not enough because that does not say anything about how close the iterate is to the feasible set. The standard approach today is to define a merit function which combines both the objective and the constraint violations into one function. For this reason, merit functions are closely related to penalty functions. For simplicity let us assume all constraints in **NLP** are equalities of the form $c(x) = 0$, the two most popular merit functions for SQP methods are the l_1 merit function

$$M(x) = F(x) + \rho \|c(x)\|_1, \tag{A.8}$$

and the augmented Lagrangian

$$M(x, \lambda) = F(x) - \lambda^T c(x) + \frac{\rho}{2} c(x)^T c(x), \tag{A.9}$$

where $\rho > 0$ is a penalty parameter and λ is a set of Lagrange multiplier estimates. In practice, a more complicated version of the augmented Lagrangian merit function is used.

We now focus on algorithms that use a line search as opposed to the trust region approach. That is, after a search direction p has been determined, we search along the line $x + \alpha p$ in order to find a good step length α. The merit function is used in the line search to determine the step α. Most convergence proofs for SQP methods rely on proving a sufficient decrease in the merit function at each iteration.

A.4 Multiplier Estimates

The augmented Lagrangian merit function based SQP methods rely on Lagrange multiplier estimates. Instead of only searching in the primal space for x^*, these methods generate iterates (x_k, λ_k) that need to converge to (x^*, λ^*). Hence the search needs to be done in a higher-dimensional space than that of x. This basic idea is also the foundation of primal-dual methods. It is important that we compute accurate multiplier estimates, since x cannot converge to x^* any faster than $\lambda \to \lambda^*$.

In SQP methods we need to distinguish between the multipliers of the nonlinear problem and those of the QP subproblems. In each QP minor iteration, the QP multiplier estimates are updated. After solving a QP, the QP multipliers can then be used to define the search direction in the Lagrange multiplier (dual) space.

A.5 Quasi-Newton Update of the Hessian

Most SQP methods do not require exact second derivatives because these may be unavailable or expensive to compute. Another potential problem is that the exact Hessian may be indefinite. Standard SQP methods assume H is positive semidefinite and thus the QP subproblems are convex. We focus on SQP methods where at each iteration, a positive definite quasi-Newton approximation of the Hessian is used.

A.6 Active-set Approach

There are two main strategies for handing constraints in optimization problems. In interior-point (barrier) methods, the iterates are strictly feasible with respect to inequality constraints. The problem of finding an initial feasible point remains. If no such point is known, the problem is typically augmented with an artificial variable that creates an obvious feasible point for the augmented problem. When both equalities and inequalities are present, not all interior-point methods are feasible-point methods; infeasible-interior-point methods also exist.

A different approach is taken in active-set methods. The idea is to predict which constraints are active (that is, equality holds) at the solution x^*. The algorithm maintains a set of constraints that are forced to be active at each iteration. However, active-set methods are often very efficient in practice because the working set changes little when the iterates are close to minimizer.

A.7 A Simple SQP Algorithm

Based on the previous sections, we outline a simplified SQP algorithm; real implementations of SQP algorithms tend to be much more complicated.

Algorithm A.1: Simple SQP algorithm

Choose x_0, λ_0 and H_0. Set $k = 0$.

While the KKT conditions are not satisfied

> Set up a QP subproblem at x_k. Compute a search direction p_k and a Lagrange multiplier μ_k.
>
> Compute a step length α that reduces the merit function.
>
> Update $x_{k+1} = x_k + \alpha p_k$.
>
> Use μ_k to update the multiplier estimates λ_k.
>
> Set $k = k + 1$.
>
> Evaluate $c(x_k)$, $g(x_k)$, $J(x_k)$.
>
> Update H_k, the Hessian approximation.

end

Appendix B

ADIFOR

ADIFOR is a tool for the automatic differentiation of Fortran 77 programs. Given a Fortran 77 source code and a user's specification of dependent and independent variables, ADIFOR will generate an augmented derivative code that computes the partial derivatives of all of the specified dependent variables with respect to all of the specified independent variables in addition to the original result. ADIFOR is a collaborative project between the Mathematics and Computer Science Division at Argonne National Laboratory and the Center for Research on Parallel Computation at Rice University.

Appendix C

NPSOL Package

NPSOL is a set of Fortran subroutines designed to minimize a smooth function subject to constraints, which may include bounds on the variables, linear constraints and smooth nonlinear constraints. The problem is assumed to be stated in the following form:

$$\begin{aligned} &\min_{\mathbf{x}} \quad F(\mathbf{x}) \\ \ni \mathbf{l_1} \leq &\quad \mathbf{x} \quad \leq \mathbf{u_1} \\ \mathbf{l_2} \leq &\quad \mathbf{Ax} \quad \leq \mathbf{u_2} \\ \mathbf{l_3} \leq &\quad \mathbf{c}(\mathbf{x}) \quad \leq \mathbf{u_3}, \end{aligned}$$

where $F(\mathbf{x})$, the objective function, is a nonlinear function, $\mathbf{A}$ is a constant matrix of general constraints, and $\mathbf{c}(\mathbf{x})$ is a m_N-vector of nonlinear constraint functions. NPSOL may also be used for unconstrained, bound-constrained and linearly constrained optimization problems, in which case $\mathbf{A}$ or $\mathbf{c}$ may be empty. The user needs to provide subroutines that define the objective and constraint functions and (optionally) their gradients; the unspecified objective derivatives are approximated by finite-differences. The objective function F and the constraints functions are assumed to be smooth, i.e., at least twice-continuously differentiable. All matrices are treated as dense, and hence NPSOL is not intended for large sparse problems.

NPSOL uses a sequential quadratic programming (SQP) algorithm, in which the search direction is the solution of a quadratic programming subproblem. The algorithm treats bounds, linear constraints and nonlinear constraints separately. The Hessian of each QP subproblem is a positive-definite quasi- Newton approximation to the Hessian of the Lagrangian function. The steplength at each iteration is required to produce a sufficient decrease in an augmented Lagrangian merit function. Each QP subproblem is solved using a quadratic programming package with several features that improve the efficiency of an SQP algorithm.

Appendix D

SNOPT Package

SNOPT is a general-purpose system for solving optimization problems involving many variables and constraints. It minimizes a linear or nonlinear function subject to bounds on the variables and sparse linear or nonlinear constraints. The problem is assumed to be stated in the form:

$$\begin{aligned} &\min_{\mathbf{x}} \quad f(\mathbf{x}) \\ \ni\ \mathbf{l_1} \leq &\quad \mathbf{x} \quad \leq \mathbf{u_1} \\ \mathbf{l_2} \leq &\quad \mathbf{Gx} \quad \leq \mathbf{u_2} \\ \mathbf{l_3} \leq &\quad \mathbf{C}(\mathbf{x}) \quad \leq \mathbf{u_3}, \end{aligned}$$

where $\mathbf{l_1}, \mathbf{l_2}, \mathbf{l_3}$ are lower bounds and $\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}$ are upper bounds, f is a smooth scalar objective function, $\mathbf{G}$ is a sparse matrix, $\mathbf{C}(\mathbf{x})$ is a vector of smooth nonlinear constraint functions. It is suitable for large-scale linear and quadratic programming and linearly constrained optimization as well as for general nonlinear programs.

SNOPT finds solutions that are locally optimal, and ideally any nonlinear functions should be smooth and users should provide gradients. It is often more widely useful. For example, local optima are often global solutions, and discontinuities in the function gradients can often be tolerated if they are not too close to an optimum. Unknown gradients are estimated by finite differences.

If nonlinear functions are absent SNOPT applies the primal simplex method. If the objective is nonlinear, SNOPT uses a sparse sequential quadratic programming (SQP) algorithm that obtains search direction from a sequence of quadratic programming subproblems (using limited-memory quasi-Newton approximations to the Hessian of the Lagrangian). Each QP subproblem minimizes a quadratic model of a certain Lagrangian function subject to a linearization of the constraints. An augmented Lagrangian merit function is reduced along each search direction to ensure convergence from any starting point.

SNOPT is most efficient if only some of the variables enter nonlinearly, or if the number of active constraints (including simple bounds) is nearly as large as the number of variables. SNOPT requires relatively few evaluations of the problem functions. Hence, it is especially effective if the objective or constraint functions (and their gradients) are expensive to evaluate. In cases where objective and constraint functions are structured in the sense that they are formed from sums of linear and nonlinear functions, the structure can be exploited by SNOPT.

The source code for SNOPT is suitable for any machine with a Fortran compiler. SNOPT may be called from a driver program (typically in Fortran, C or MatLab). SNOPT may also be used as a stand-alone package, reading data in the MPS format used by commercial mathematical programming systems.

Bibliography

[1] Jeffery M. Adachi. *Combining Stochastic Linear Programming and Dynamic Programming to Solve Multiperiod Portfolio Problems.* PhD thesis, Department of Operations Research, Stanford University, 1996.

[2] David E. Bell, Howard Raiffa, and Amos Tversky. *Decision Making - Descriptive, Normative, and Prescriptive Interactions.* Cambridge University Press, Cambridge, U.K., 1988.

[3] J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4:238–252, 1962.

[4] Daniel Bernoulli. Specimen theoriae novae de mensura sortis. *Comentarii Academiae Scientiarum Imperiales Petropolitanae*, 5:175–192, 1738.

[5] Dimitri P. Bertsekas. *Nonlinear Programming.* Athena Scientific, Belmont, Massachusetts, 1995.

[6] John R. Birge. Stochastic programming computation and applications. *INFORMS Journal on Computing*, 9(2):111–133, 1997.

[7] John R. Birge and Charles H. Rosa. Parallel decomposition of large-scale stochastic nonlinear programs. *Annals of Operations Research*, 64:39–65, 1996.

[8] Zvi Bodie, Robert C. Merton, and William F. Samuelson. Labor supply flexibility and portfolio choice in life-cycle mode. *Journal of Economic Dynamics and Control*, 16:427–449, 1992.

[9] Michael J. Brennan, Eduardo S. Schwartz, and Ronald Lagnado. Strategic asset allocation. *Journal of Economic Dynamics and Control*, 21:1377–1403, 1997.

[10] John Y. Campbell and Luis M. Viceira. Consumption and portfolio decisions when expected returns are time varying. *NBER Working paper No. 5857*, 1996.

[11] David R. Carino, Terry Kent, David H. Meyers, Celine Stacy, Mike Sylvanus, Andrew L. Turner, Kouji Watanabe, and William T. Ziemba. The russell-yasuda kasai model: An asset/liability model for a japanese insurance company using multistage stochastic programming. *Interfaces*, 24(1):29–49, 1994.

[12] George B. Dantzig and Peter Glynn. Parallel processors for planning under uncertainty. *Annals of Operations Research*, 22:1–21, 1990.

[13] George B. Dantzig and Gerd Infanger. Multi-stage stochastic linear programs for portfolio optimization. *Annals of Operations Research*, 45:59–76, 1993.

[14] George B. Dantzig and P. Wolfe. The decomposition principle for linear programs. *Operations Research*, 8:101–111, 1960.

[15] Darrell Duffie and Thaleia Zariphopoulou. Optimal investment with undiversifiable income risk. *Mathematical Finance*, 3(2):135–148, 1993.

[16] Ward Edwards and Amos Tversky. *Decision Making - Selected Readings*. Penguin Books, Baltimore, Maryland, 1967.

[17] P.C. Fishburn. Mean-risk analysis with risk associated with below-target returns. *American Economic Review*, 67(2):116–126, 1977.

[18] Milton Friedman and L. J. Savage. The utility analysis of choices involving risk. *The Journal of Political Economy*, 56(4):279–304, 1948.

[19] Chi fu Huang and Robert H. Litzenberger. *Foundations for Financial Economics.* Prentice Hall, Inc., Englewood Cliffs, New Jersy, 1988.

[20] Philip E. Gill, Walter Murray, and Michael A. Saunders. *User's Guide for SNOPT(Version 4.0): A Fortran Package for Sparse Nonlinear Programming.* Technical Report SOL 86-2, 1998.

[21] Philip E. Gill, Walter Murray, and Michael A. Saunders. Snopt: An algorithm for large-scale constrained optimization. *SIAM Journal on Optimization*, To appear.

[22] Philip E. Gill, Walter Murray, Michael A. Saunders, and Margaret H. Wright. *User's Guide for NPSOL(Version 4.0): A Fortran Package for Nonlinear Programming.* Technical Report SOL 86-2, 1986.

[23] Philip E. Gill, Walter Murray, and Margaret H. Wright. *Practical Optimization.* Academic Press, Harcourt Brace and Company, Publishers, London, 1981.

[24] Michael D. Hurd and Katheleen McGarry. The predictive validity of subjective probabilities of survival. *NBER Working Paper No 6193*, 1997.

[25] Gerd Infanger. *Planning Under Uncertainty - Sovling Large-Scale Stochastic Linear Programs.* Boyd & Fraser Publishing Company, Danvers, Massachusetts, 1994.

[26] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decisions under risk. *Econometrica*, 47:313–327, 1979.

[27] Ioannis Karatzas, John P. Lehoczky, Suresh Sethi, and Steven E. Shreve. Explicit solution of a general consumption/investment problem. *Mathematics of Operations Research*, 11(2):261–294, 1986.

[28] Michael P. Keane and Kenneth I. Wolpin. The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: monte carlo evidence. *The Review of Economics and Statistics*, 76(4):648–672, 1994.

[29] Hiroshi Konno and Katsunari Kobayashi. An integrated stock-bond portfolio optimization model. *J. Econ. Dyn. Control*, 21(8-9):1405–1425, 1997.

[30] Lamberton and Lapeyre. *Introduction to Stochastic Calculus Applied to Finance.* Chapman and Hall, London, UK, 1996.

[31] David G. Luenberger. *Introduction to Linear and Nonlinear Programming.* Addison-Wesley, Reading, MA, 1984.

[32] David G. Luenberger. *Investment Science.* Oxford University Press, New York, 1998.

[33] Robin L. Lumsdaine, James H. Stock, and Davis A. Wise. Three models of retirement: Computational complexitiy versus predictive validity. *NBER Working Paper No 3558*, 1990.

[34] Mark Machina. Expected utility analysis without the independence axiom. *Econometrica*, 50:277–323, 1982.

[35] C.D. Maranas, I.P. Androulakis, C.A. Floudas, A.J. Berger, and J.M. Mulvey. Solving long-term financial planning problems via global optimization. *Journal of Economic Dynamics and Control*, 21(8-9):1405–1425, 1997.

[36] Harry Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.

[37] Harry Markowitz. The utility of wealth. *Journal of Political Economy*, 60(2):151–158, 1952.

[38] Andreu Mas-Colell, Michael D. Winston, and Jerry R. Green. *Microeconomic Theory*. Oxford University Press, New York, 1995.

[39] Robert C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics*, 51(3):247–257, 1969.

[40] Robert C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373–413, 1971.

[41] Robert C. Merton. *Continuous-Time Finance*. Basil Blackwell Inc., Cambridge, Massachusetts 02142, 1990.

[42] Jan Mossin. Optimal multiperiod portfolio policies. *Journal of Business*, 41(2):215–229, 1968.

[43] J.M. Mulvey and Andrzej Ruszczynski. A new scenario decomposition method for large-scale stochastic optimization. *Operations Research*, 43(3):447–490, 1995.

[44] J.M. Mulvey and H. Vladimirou. Stochastic network optimization models for investment planning. *Annals of Operations Research*, 20:187–217, 1989.

[45] R.T. Rockafellar and Roger J-B. Wets. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16:119–147, 1991.

[46] Sheldon M. Ross. *Stochastic Processes*. John Wiley & Sons, Inc., New York, 1996.

[47] John Rust. A dynamic programming model of retirement behavior. *NBER Working paper No. 2470*, 1987.

[48] Paul A. Samuelson. Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*, 51:239–246, 1969.

[49] William F. Sharpe, Gordon J. Alexander, and Jeffery V. Bailey. *Investments*. Prentice Hall, Inc. Englewood Cliffs, New Jersy, 1995.

[50] John B. Shoven. The location and allocation of assets in pension and conventional savings accounts. *NBER Working Paper No. 7007*, 1999.

[51] Joseph E. Stiglitz, editor. *Collected Scientific Papers of Paul A. Samuelson*. Number Chapter 13. The MIT Press, Cambridge, MA, 1966.

[52] Bernell K. Stone. A general class of three-parameter risk measures. *Journal of Finance*, 28:675–685, 1973.

[53] Amos Tversky. Additivity, utility an subjective probability. *Journal of Mathematical Psychology*, 4:175–202, 1967.

[54] Amos Tversky and Daniel Kahneman. Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics*, 106(4):1039–1061, 1991.

[55] Amos Tversky and Daniel Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5:297–323, 1992.

[56] Hal R. Varian. *Microeconomic Theory*. W.W. Norton and Company, Inc., New York, 1992.

[57] Luis M. Viceira. *Optimal Consumption and Portfolio Choice for Long-Horizon Investors*. PhD thesis, Department of Economics, Harvard University, 1998.

[58] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton N.J., 1944.

[59] Kenneth J. Worzel, Christiana Vassiadou-Zeniou, and Stavros A. Zenios. Integrated simulation and optimization models for tracking indices of fixed-income securities. *Operations Research*, 42:223–233, 1994.

[60] George Wright. *Behavioral Decision Making*. Plenum Press, New York, N.Y., 1985.