

# Essays in Financial Engineering

by

Martin B. Haugh

B.Sc. Mathematics and Statistics, University College Cork, 1992

M.Sc. Mathematics, University College Cork, 1993

M.Sc Applied Statistics, University of Oxford, 1994

Submitted to the Sloan School of Management  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2001

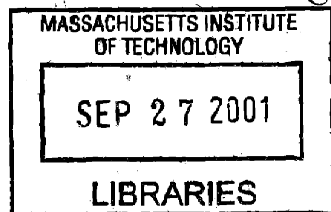
© Massachusetts Institute of Technology 2001. All rights reserved.

Author .....  
Sloan School of Management  
August 10, 2001

Certified by .....  
Andrew W. Lo  
Harris & Harris Group Professor of Finance  
Thesis Supervisor

Accepted by .....  
James B. Orlin  
Edward Pennell Brooks Professor of Operations Research  
Codirector, Operations Research Center

Dewey



# Essays in Financial Engineering

by

Martin B. Haugh

B.Sc. Mathematics and Statistics, University College Cork, 1992

M.Sc. Mathematics, University College Cork, 1993

M.Sc Applied Statistics, University of Oxford, 1994

Submitted to the Sloan School of Management  
on August 10, 2001, in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy in Operations Research

## Abstract

This thesis consists of three essays that apply techniques of operations research to problems in financial engineering. In particular, we study problems in portfolio optimization and options pricing.

The first essay is motivated by the fact that derivative securities are equivalent to specific dynamic trading strategies in complete markets. This suggests the possibility of constructing buy-and-hold portfolios of options that mimic certain dynamic investment policies, e.g., asset-allocation rules. We explore this possibility by solving the following problem: given an optimal dynamic investment policy, find a set of options at the start of the investment horizon which will come closest to the optimal dynamic investment policy. We solve this problem for several combinations of preferences, return dynamics, and optimality criteria, and show that under certain conditions, a portfolio consisting of just a few European options is an excellent substitute for considerably more complex dynamic investment policies.

In the second essay, we develop a method for pricing and exercising high-dimensional American options. The approach is based on approximate dynamic programming using nonlinear regression to approximate the value function. Using the approximate dynamic programming solutions, we construct upper and lower bounds on the option prices. These bounds can be evaluated by Monte Carlo simulation, and they are general enough to be used in conjunction with other approximate methods for pricing American options. We characterize the theoretical worst-case performance of the pricing bounds and examine how they may be used for hedging and exercising the option. We also discuss the implications for the design of the approximate pricing algorithm and illustrate its performance on a set of sample problems where we price call options on the maximum and the geometric mean of a collection of stocks.

The third essay explores the possibility of solving high-dimensional portfolio optimization problems using approximate dynamic programming. In particular, we employ approximate

value iteration where the portfolio strategy at each time period is obtained using quadratic approximations to the approximate value function. We then compare the resulting solution to the best heuristic strategies available. Though the approximate dynamic programming solutions are often competitive, they are sometimes dominated by the best heuristic strategy. On such occasions we conclude that inaccuracies in the quadratic approximations are responsible for the poor performance. Finally, we compare our results to other recent work in this area and suggest possible methods for improving these algorithms.

Thesis Supervisor: Andrew W. Lo  
Title: Harris & Harris Group Professor of Finance

## Acknowledgments

I have many people to thank for helping me arrive at this moment. My thesis committee members, Dimitris Bertsimas and Jiang Wang, deserve my sincere thanks for the many ways in which they helped me. I am also very grateful to Leonid Kogan with whom I have had many useful and interesting conversations. But of course, I owe the greatest debt to my advisor, Andrew Lo. Andrew always encouraged me to work on problems that appealed to me and during my time at MIT I learnt a great deal about the entire research process from him. He always made himself available and working with him was a pleasure.

I would also like to thank all the students, staff and codirectors who were at the OR Center during my time there. It is because of them that the OR Center is such a great to be.

Since moving to Boston in 1996, I have had the privilege of meeting many people who have since become very good friends. I hope and am sure that we shall remain friends in the years ahead. Many of these people were in the Operations Research Center, and in MIT in general. Others, to their benefit I'm sure, had no affiliation with MIT. To you all, I say thank you. You know who you are. To my friends and family in Ireland, I also say thank you.

In true academic style, I should dedicate this thesis to somebody or something. My favourite watering hole in Boston, Grafton Street, surely deserves a mention. I certainly spent enough time and money there.

But of course, I can only dedicate this thesis to my parents who have done so much for me.

# Contents

<b>1</b>	<b>Asset Allocation and Derivatives</b>	<b>8</b>
1.1	Introduction . . . . .	9
1.2	Literature Review . . . . .	11
1.3	The Optimal Buy-and-Hold Portfolio . . . . .	12
1.3.1	Maximizing Expected Utility . . . . .	13
1.3.2	Minimizing Mean-Squared Error . . . . .	15
1.3.3	Minimizing Weighted Mean-Squared Error . . . . .	17
1.4	Three Leading Cases . . . . .	18
1.4.1	Geometric Brownian Motion . . . . .	19
1.4.2	The Ornstein-Uhlenbeck Process . . . . .	20
1.4.3	A Bivariate Linear Diffusion Process . . . . .	23
1.5	Numerical Results . . . . .	24
1.5.1	Geometric Brownian Motion . . . . .	28
1.5.2	The Ornstein-Uhlenbeck Process . . . . .	33
1.5.3	A Bivariate Linear Diffusion Process . . . . .	36
1.6	Discussion . . . . .	38
1.6.1	Practical Considerations . . . . .	38
1.6.2	Other Derivative Securities . . . . .	39
1.6.3	Other Preferences and Price Processes . . . . .	40
1.6.4	The Predictability Gap . . . . .	40
1.7	Conclusion . . . . .	41

<b>2 Pricing American Options: A Duality Approach</b>	<b>55</b>
2.1 Introduction . . . . .	56
2.2 Problem Formulation . . . . .	57
2.2.1 The Model . . . . .	57
2.2.2 The Dual Problem . . . . .	58
2.3 Option Price Approximation . . . . .	61
2.3.1 The Q-Value Function . . . . .	62
2.3.2 Approximate Q-Value Iteration . . . . .	63
2.3.3 The Lower Bound on the Option Price . . . . .	65
2.3.4 The Upper Bound on the Option Price . . . . .	67
2.4 Numerical Results . . . . .	68
2.4.1 Call on the Maximum of 5 Assets . . . . .	69
2.4.2 Call on the Geometric Mean of 5 Assets . . . . .	69
2.4.3 Approximate Q-Value Iteration Revisited . . . . .	69
2.5 Conclusions and Further Research . . . . .	70
<b>3 Portfolio Optimization and Approximate Dynamic Programming</b>	<b>74</b>
3.1 Introduction . . . . .	74
3.2 Problem Formulation . . . . .	75
3.3 Approximate Value Iteration . . . . .	77
3.3.1 Quadratic Approximations to the Value Function . . . . .	78
3.4 Numerical Results . . . . .	81
3.5 Conclusions and Further Research . . . . .	84
<b>A Solution of HJB Equations for Ornstein-Uhlenbeck Processes</b>	<b>92</b>
A.1 Trending Ornstein-Uhlenbeck Value Function . . . . .	92
A.2 Non-Trending Ornstein-Uhlenbeck Value Function . . . . .	94
<b>B Approximate Q-Value Iteration</b>	<b>98</b>
B.1 Low Discrepancy Sequences . . . . .	98
B.2 Training Point Selection and Evaluation . . . . .	99

B.3	Training the Neural Network . . . . .	101
B.4	Policy Fixing and Feature Extraction . . . . .	102
<b>C</b>	<b>Linear Programming and Dynamic Programming Bounds</b>	<b>104</b>
C.1	Optimal Stopping . . . . .	105
C.2	Finite Horizon Dynamic Programming . . . . .	106
	<b>Bibliography</b>	<b>109</b>

# Chapter 1

## Asset Allocation and Derivatives

### Abstract

The fact that derivative securities are equivalent to specific dynamic trading strategies in complete markets suggests the possibility of constructing buy-and-hold portfolios of options that mimic certain dynamic investment policies, e.g., asset-allocation rules. We explore this possibility by solving the following problem: given an optimal dynamic investment policy, find a set of options at the start of the investment horizon which will come closest to the optimal dynamic investment policy. We solve this problem for several combinations of preferences, return dynamics, and optimality criteria, and show that under certain conditions, a portfolio consisting of just a few options is an excellent substitute for considerably more complex dynamic investment policies.

*Co-author:* **Andrew Lo**



## 1.1 Introduction

It is now well-known that under certain conditions, complex financial instruments such as options and other derivative securities can be replicated by sophisticated dynamic trading strategies involving simpler securities such as stocks and bonds. This “delta-hedging” strategy—for which Robert Merton and Myron Scholes shared the Nobel Memorial Prize in Economics in 1998—is largely responsible for the multi-trillion-dollar derivatives industry and is now part of the standard toolkit of every derivatives dealer in the world.

The essence of delta-hedging is the ability to actively manage a portfolio continuously through time, and to do so in a “self-financing” manner, i.e., no cash inflows or outflows after the initial investment, so that the portfolio’s value tracks the value of the derivative security without error at each point in time, until the maturity date of the derivative. If such a portfolio strategy were possible, then the cost of implementing it must equal the price of the derivative, otherwise an arbitrage opportunity would exist. Black and Scholes (1973) and Merton (1973) used this argument to deduce the celebrated Black-Scholes option-pricing formula, but an even more significant outcome of their research was the insight that there exists a correspondence between dynamic trading strategies over a period of time and complex securities at a single point in time.

In this paper, we consider the reverse implications of this correspondence by constructing an optimal portfolio of complex securities at a single point in time to mimic the properties of a dynamic trading strategy over a period of time. Specifically, we focus on dynamic investment policies, i.e., asset-allocation rules, that arise from standard dynamic optimization problems in which an investor maximizes the expected utility of his end-of-period wealth, and we pose the following problem: given an investor’s optimal dynamic investment policy for two assets, stocks and bonds, construct a “buy-and-hold” portfolio—a portfolio that involves no trading once it is established—of stocks, bonds, and options at the start of the investment horizon that will come closest to the optimal dynamic policy. By defining “closest” in three distinct ways—expected utility, mean-squared error of terminal wealth, and utility-weighted mean-squared error of terminal wealth—we propose three sets of numerical algorithms for solving this problem in general, and characterize specific solutions for several sets of preferences (constant relative risk-aversion, constant absolute risk-aversion) and return dynamics (geometric Brownian motion, mean-reverting processes).

The optimal buy-and-hold problem is an interesting one for several reasons. First, it is widely acknowledged that the continuous-time framework in which most of modern finance has been developed is an approximation to reality—it is currently impossible to trade continuously, and even if it were possible, market frictions would render continuous trading infinitely costly. Consequently, any practical implementation of continuous-time asset-allocation policies invariably requires some discretization in which the investor’s portfolio is rebalanced only a finite number of times, typically at equally spaced time intervals, with the number of intervals chosen so that the discrete asset-allocation policy “approximates” the optimal continuous-time policy in some metric. However, Merton’s (1973) insight suggests that it may be possible to approximate a continuous-time trading strategy in a different manner, i.e., by including a few well-chosen options in the portfolio at the outset and trading considerably less frequently. In particular, Merton (1995) observes that derivatives can be an effective substitute for dynamic open-market operations of central banks seeking to engage in interest-rate stabilization policies. Therefore, in the presence of transactions costs, deriva-

tive securities may be an efficient way to implement optimal dynamic investment policies.<sup>1</sup> Indeed, we find that under certain conditions, a buy-and-hold portfolio consisting of just a few options is an excellent substitute for considerably more complex dynamic investment policies.

Second, the approximation errors between the optimal dynamic policy and the buy-and-hold policy will reveal the importance of dynamic trading, the “completeness” of financial markets, and the ability of investors to achieve certain financial goals in a cost-effective manner.<sup>2</sup> In particular, the conditions that guarantee dynamic completeness are nontrivial restrictions on market structure and price dynamics (see, for example, Duffie and Huang, 1985), hence there are situations in which exact replication is impossible. These instances of *market incompleteness* are often attributable to institutional rigidities and market frictions—transactions costs, periodic market closures, and discreteness in trading opportunities and prices—and while the pricing of derivative securities can still be accomplished in some cases via equilibrium arguments,<sup>3</sup> this still leaves open the question of how expensive it is to achieve certain financial objectives, or how close one can come to those objectives for a given budget?

Finally, the optimal buy-and-hold portfolio can be used to develop a measure of the risks associated with the corresponding dynamic investment policy that the buy-and-hold portfolio is designed to replicate. While there is general agreement in the financial community regarding the proper measurement of risk in a static context—the market beta from the Capital Asset Pricing Model—there is no consensus regarding the proper measurement of risk for dynamic investment strategies. Market betas are notoriously unreliable in a multi-period setting,<sup>4</sup> and other measures such as the Sharpe ratio, the Sortino ratio, and maximum draw-down have been used to capture different risk exposures of dynamic investment strategies. By developing a correspondence between a dynamic investment strategy and a buy-and-hold portfolio, it may be possible to construct a more comprehensive set of risk measures for the dynamic strategy through the characteristics of the buy-and-hold portfolio and the approximation error.

In Section 1.2 we provide a brief review of the strands of the asset-allocation and derivatives pricing literature that are most relevant to our problem. We describe the buy-and-hold alternative to the standard asset-allocation problem in Section 1.3 and propose three methods for solving it: maximization of expected utility, minimization of mean-squared error, and

---

<sup>1</sup>Taxes can be viewed as another type of transactions cost, and the optimal buy-and-hold portfolio offers several additional advantages over the optimal dynamic investment policy for taxable investors.

<sup>2</sup>Financial markets are said to be “complete” (in the Arrow-Debreu sense) if it is possible to construct a portfolio of securities at a point in time which guarantees a specific payoff in a specific state of nature at some future date. The notion of “dynamic completeness” is the natural extension of this idea to dynamic trading strategies. See Harrison and Kreps (1979) and Duffie and Huang (1985) for a more detailed discussion.

<sup>3</sup>Examples of continuous-time incomplete-markets models include Breeden (1979), Duffie and Shafer (1985, 1986), Föllmer and Sonderman (1986), Duffie (1987), and He and Pearson (1991). Examples of discrete-time incomplete-markets models include Scheinkman and Weiss (1986), Aiyagari and Gertler (1991), Heaton and Lucas (1992, 1996), Weil (1992), Telmer (1993), Aiyagari (1994), Lucas (1994), and He and Modest (1995). Other aspects of pricing and hedging in incomplete markets have been considered by Magill and Quinzii (1996), Kallsen (1999), Kramkov and Schachermeyer (1999), Bertsimas, Kogan, and Lo (2000), Goll and Rueschendorf (2000), and Schael (2000).

<sup>4</sup>See, for example, the short-put strategy described in Lo (2000).

a hybrid of the two (minimization of utility-weighted mean-squared error). While the first approach is the most direct, it is also the most computationally intensive. The latter two approaches are simpler to implement, however, they do not maximize expected utility and as a result, the portfolios that they generate may be suboptimal. These issues are addressed in more detail in Sections 1.4 and 1.5 where we implement the three methods for geometric Brownian motion, the Ornstein-Uhlenbeck process, and a bivariate linear diffusion process with a stochastic mean-reverting drift. Extensions, qualifications, and other aspects of the optimal buy-and-hold portfolio are discussed in Section 1.6, and we conclude in Section 1.7.

## 1.2 Literature Review

The literature on asset allocation is vast and addresses a broad set of issues, many that are beyond the scope of this paper's main focus.<sup>5</sup> Most studies that consider derivatives in the context of asset allocation use option-pricing methods to gauge the economic value of market-timing skills, e.g., Merton (1981), Henriksson and Merton (1981), and Evnine and Henriksson (1987). Carr, Jin and Madan (2000) solve the asset-allocation problem in an economy where derivatives are required to complete the market. Carr and Madan (2000) consider a single-period model where agents are permitted to trade the stock, bond and European options with a continuum of strikes. Because of the inability to trade dynamically, options constitute a new asset class and the impact of beliefs and preferences on the agent's positions in the three asset classes is studied. In a general equilibrium framework, they derive conditions for mutual-fund separation where some of the separating funds are composed of derivative securities. None of these papers explores the possibility of substituting a simple buy-and-hold portfolio for a dynamic investment policy.

Three other strands of the literature are relevant to our paper: Merton's (1995) functional approach to understanding the dynamics of financial innovation,<sup>6</sup> the literature on dynamic portfolio choice with transactions costs, and the literature on option replication.

Among the many examples contained in Merton (1995) illustrating the importance of function in determining institutional structure is the example of the German government's issuance in 1990 of ten-year *Schuldschein* bonds with put-option provisions. Merton (1995) observes that the put provisions have the same effect as an interest-rate stabilization policy in which the government repurchases bonds when bond prices fall and sells bonds when bond prices rise. More importantly, Merton (1995) writes that "... the put bonds function as the equivalent of a dynamic, 'open market', trading operation without any need for actual transactions". This automatic stabilization policy is a "proof of concept" for the possibility of substituting a buy-and-hold portfolio for a particular dynamic investment strategy, and the optimal buy-and-hold portfolio of Section 1.3 may be viewed as a generalization of Merton's automatic stabilization policy to the asset-allocation problem.

Magill and Constantinides (1976) were among the first to point out that in the presence of transactions costs, trading occurs only at discrete points in time. More recent studies by

---

<sup>5</sup>See Sharpe (1987), Arnott and Fabozzi (1992), and Bodie, Kane and Marcus (1999) for more detailed expositions of asset allocation.

<sup>6</sup>See, also, Bodie and Merton (1995) and Merton (1997).

Davis and Norman (1990), Aiyagari and Gertler (1991), Heaton and Lucas (1992, 1996), and He and Modest (1995) have contributed to the growing consensus that trading costs have a significant impact on investment performance and, therefore, investor behavior. Despite the recent popularity of internet-based day-trading, it is now widely accepted that buy-and-hold strategies such as indexation are difficult to beat—transactions costs and management fees can quickly dissipate the value-added of many dynamic asset-allocation strategies.

The option-replication literature is relevant to our paper primarily because of the correspondence between a complex security and a dynamic trading strategy in simpler securities, an insight which gave rise to this literature. The classic references are Black and Scholes (1973), Merton (1973), Cox and Ross (1976), Harrison and Kreps (1979), Duffie and Huang (1985), and Huang (1985a,b). More recently, several studies have considered the option-replication problem directly, in some cases using mean-squared error as the objective function to be minimized,<sup>7</sup> and in other cases with transactions costs.<sup>8</sup> In the latter set of studies, the existence of transactions costs induces discrete trading intervals, and the optimal replication problem is solved for some special cases, e.g., call and put options on stocks with geometric Brownian motion or constant-elasticity-of-variance price dynamics, or for more general derivative securities under vector-Markov price processes.

We take these somewhat disparate literatures as our starting point. Merton's (1995) automatic stabilization policy illustrates the possibility of substituting a static buy-and-hold portfolio for a specific dynamic trading strategy, i.e., an interest-rate stabilization policy. The fact that trading is costly implies that continuous asset-allocation is not feasible, and that alternatives to frequent trading are important to investors. The technology for replicating options is clearly well established, and a natural generalization of that technology is to construct portfolios of options that replicate more general dynamic trading strategies. We begin developing this generalization in the next section.

### 1.3 The Optimal Buy-and-Hold Portfolio

The asset-allocation problem has become one of the classic problems of modern finance, thanks to Samuelson's (1969) and Merton's (1969) pioneering studies over three decades ago. The simplest formulation—one without intermediate consumption—consists of an investor's objective to maximize the expected utility  $E[U(W_T)]$  of end-of-period wealth  $W_T$  by allocating his wealth  $W_t$  between two assets, a risky security (the "stock") and a riskless security (the "bond"), over some investment horizon  $[0, T]$ . The bond is assumed to yield a riskless instantaneous return of  $r dt$  and with an initial market price of \$1, the bond price at any date  $t$  is simply  $\exp(rt)$ . The stock price is denoted by  $P_t$  and is typically assumed to

---

<sup>7</sup>See, for example, Duffie and Jackson (1990), Schweizer (1992, 1995, 1996), Schäl (1994), Delbaen and Schachermeyer (1996), and Bertsimas, Kogan, and Lo (2000a).

<sup>8</sup>See Leland (1985), Hodges and Neuberger (1989), Bensaid, et al. (1992), Boyle and Vorst (1992), Davis, Panas, and Zariphopoulou (1993), Edirisinghe, Naik, and Uppal (1993), Henrotte (1993), Avellaneda and Paras (1994), Neuberger (1994), Whalley and Wilmott (1994), Grannan and Swindle (1996), and Toft (1996).

satisfy an Itô stochastic differential equation:

$$dP_t = \mu(P_t, t) P_t dt + \sigma(P_t, t) P_t dB_t \quad (1.3.1)$$

where  $B_t$  is standard Brownian motion and  $\mu(P_t, t)$  and  $\sigma(P_t, t)$  satisfy certain regularity conditions that ensure the existence of a solution to (1.3.1). The standard asset-allocation problem is then:

$$\text{Max}_{\{\omega_t\}} E[U(W_T)] \quad (1.3.2)$$

subject to

$$dW_t = [r + \omega_t(\mu - r)]W_t dt + \omega_t W_t \sigma dB_t \quad (1.3.3)$$

where  $\omega_t$  is the fraction of the investor's portfolio invested in the stock at time  $t$  and (1.3.3) is the budget constraint that wealth  $W_t$  must satisfy at all times  $t \in [0, T]$ .<sup>9</sup>

Denote by  $\{\omega_t^*\}$  the optimal dynamic investment policy, i.e., the solution to (1.3.2)–(1.3.3), and let  $W_T^*$  denote the end-of-period wealth generated by the optimal policy. The question we wish to answer in this paper is: how close can we come to this optimal policy with a buy-and-hold portfolio of stocks, bonds, and options? We measure closeness in three ways: a direct approach in which we maximize the expected utility of the buy-and-hold portfolio, and two indirect approaches in which we minimize the mean-squared error and weighted mean-squared error between  $W_T^*$  and the end-of-period wealth of the buy-and-hold portfolio. These three approaches are described in Sections 1.3.1–1.3.3, respectively.

### 1.3.1 Maximizing Expected Utility

Our reformulation of the standard asset-allocation problem (1.3.2)–(1.3.3) contains only two modifications: (1) we allow the investor to include up to  $n$  European call options in his portfolio at date 0 which expire at date  $T$ ;<sup>10</sup> and (2) we do not allow the investor to trade after setting up his initial portfolio of stocks, bonds and options. Specifically, denote by  $D_i$  the date- $T$  payoff of a European call option with strike price equal to  $k_i$ , hence:

$$D_i = (P_T - k_i)^+ . \quad (1.3.4)$$

<sup>9</sup>See Merton (1992, Chapter 5) for details.

<sup>10</sup>Without loss of generality, we focus exclusively on call options for expositional simplicity. Parallel results for put options can be easily derived via put-call parity (see, for example, Cox and Rubinstein (1985)).

Then the “buy-and-hold” asset-allocation problem for the investor is given by:

$$\text{Max}_{\{a,b,c_i,k_i\}} E[U(V_T)] \quad \text{subject to} \quad (1.3.5)$$

$$V_T \equiv a \exp(rT) + b P_T + c_1 D_1 + c_2 D_2 + \cdots + c_n D_n \quad (1.3.6)$$

$$W_0 = \exp(-rT) E^Q[V_T] \quad (1.3.7)$$

where  $a$  and  $b$  denote the investor’s position in bonds and stock, and  $c_1, \dots, c_n$  the number of options with strike prices  $k_1, \dots, k_n$ , respectively. Note that we use  $V_T$  instead of  $W_T$  to denote the investor’s end-of-period wealth to emphasize the distinction between this case and the standard asset-allocation problem in which stocks and bonds are the only assets considered and intermediate trading is allowed.

The budget constraint is given by (1.3.7), where  $E^Q[\cdot]$  is the conditional expectation operator under the equivalent martingale measure  $Q$ .<sup>11</sup> This constraint is highly nonlinear in the option strikes  $\{k_i\}$ , creating significant computational challenges for any optimizer. Moreover, for certain utility functions, it is necessary to impose solvency constraints to avoid bankruptcy, and such constraints add to the computational complexity of the problem.

For these reasons, our approach for solving (1.3.5)–(1.3.7) consists of two steps. In the first step, we assume that the strike prices  $\{k_i\}$  are fixed, in which case (1.3.5)–(1.3.7) reduces to maximizing a concave objective function subject to linear constraints. Such a problem has a unique global optimum that is generally quite easy to find. This is done by discretizing the distribution of  $P_T$  and solving the Karush-Kuhn-Tucker conditions which, in this case, are sufficient for an optimal solution.<sup>12</sup> We will refer to this problem—where the strikes  $\{k_i\}$  are fixed—as the “sub-problem”.

The second step involves determining the best set of strikes. We propose to solve this problem by specifying in advance a large number,  $N \gg n$ , of possible strikes where the  $N$  strikes are chosen to be representative of the distribution of  $P_T$ . We then solve the sub-problem for each of the  $\binom{N}{n}$  possible combinations of options and select the best combination.

In selecting the set of  $N$  strikes, we must ensure that their range spans a significant portion of the support of  $P_T$ . Therefore, the distribution of  $P_T$  must be taken into account in specifying the strikes. Given a distribution for  $P_T$ , we select an interval of its support and choose  $N$  points—spaced either evenly (for simplicity) or according to the probability mass of the distribution of  $P_T$  (for efficiency)—so that approximately 4 to 6 standard deviations of  $P_T$  are contained within the interval.

In solving each sub-problem, we discretize the distribution of  $P_T$ . This yields a straightforward nonlinear optimization problem with a concave objective function and linear constraints, which can be solved relatively quickly.

For concreteness, we provide a detailed analysis of this approach in Section 1.4 for two

<sup>11</sup>Note that specifying  $Q$  yields pricing formulas for all the options contained in our optimal buy-and-hold portfolio since  $\exp(-rT)E^Q[D_i]$  is the date-0 price of option  $i$ . Therefore, option-pricing formulas are implicit in (1.3.7). For example, it is easy to verify that under geometric Brownian motion,  $\exp(-rT)E^Q[D_i]$  reduces to the celebrated Black-Scholes formula.

<sup>12</sup>See, for example, Bertsekas (1999).

specific utility functions, constant-relative-risk-aversion (CRRA) and constant-absolute-risk-aversion (CARA) utility:

$$U(V_T) = \frac{V_T^\gamma}{\gamma} \quad (\text{CRRA}) \quad (1.3.8)$$

$$U(V_T) = -\frac{\exp(-\gamma V_T)}{\gamma} \quad (\text{CARA}). \quad (1.3.9)$$

One subtlety arises for CRRA utility: the function is not defined for negative wealth. In such cases, the following  $n+2$  solvency constraints must be imposed along with the budget constraint to ensure non-negative wealth:

$$\begin{aligned} 0 &\leq a \exp(rT) \\ 0 &\leq a \exp(rT) + bk_1 \\ 0 &\leq a \exp(rT) + (b + c_1)k_2 - c_1k_1 \\ &\vdots \\ 0 &\leq a \exp(rT) + (b + c_1 + \dots + c_{n-1})k_n - (c_1k_1 + \dots + c_{n-1}k_{n-1}) \\ 0 &\leq b + c_1 + \dots + c_n \\ 0 &\leq k_1 \leq k_2 \leq \dots \leq k_n. \end{aligned} \quad (1.3.10)$$

### 1.3.2 Minimizing Mean-Squared Error

In situations where the computational demands of the buy-and-hold asset-allocation problem of Section 1.3.1 is too great, a less demanding alternative is to use mean-squared error as the metric for measuring the closeness of the end-of-period wealth  $V_T$  of the buy-and-hold portfolio of stocks, bonds, and options with the end-of-period wealth  $W_T^*$  of the optimal portfolio. In addition, for dynamic investment policies that are not derived from maximization of expected utility, e.g., dollar-cost averaging, a mean-squared-error objective function may be appropriate. In this case, the buy-and-hold portfolio problem becomes:

$$\text{Min}_{\{a,b,c_i,k_i\}} \text{E}[(W_T^* - V_T)^2] \quad \text{subject to} \quad (1.3.11)$$

$$V_T \equiv a \exp(rT) + bP_T + c_1D_1 + c_2D_2 + \dots + c_nD_n \quad (1.3.12)$$

$$W_0 = \exp(-rT) \text{E}^Q[V_T] \quad (1.3.13)$$

If  $W_T^*$  depends only on the terminal stock price  $P_T$  and not on any of its path  $\{P_t\}$ —as is the case when  $\{P_t\}$  follows a geometric Brownian motion and  $W_T^*$  is the end-of-period wealth from an optimization of an investor's expected utility—it can be shown that  $V_T$  can be made arbitrarily close to  $W_T^*$  in mean-square as the number of options  $n$  in the buy-and-hold portfolio increases without bound. If we do not impose any additional constraints beyond the budget constraint (such as the solvency constraints (1.3.10) of Section 1.3.1), the corresponding sub-problems for (1.3.11)–(1.3.13) can be solved very quickly, and the first order conditions, which are necessary and sufficient, merely amount to solving a series of linear equations.

Specifically, the sub-problem associated with (1.3.11)–(1.3.13) consists of selecting portfolio weights for stocks, bonds, and options to minimize the mean-squared error between  $W_T^*$  and  $V_T$ , holding fixed the strike prices  $\{k_i\}$  of the  $n$  options available to the investor. It is clear from (1.3.11)–(1.3.13) that for fixed strike prices, the objective function is convex so the first-order conditions are sufficient to characterize an optimal solution. These conditions may be written as

$$\begin{bmatrix} \exp(rT) & E[S_T] & E[D_1] & \cdots & E[D_n] \\ \exp(rT)E[S_T] & E[S_T^2] & E[D_1S_T] & \cdots & E[D_nS_T] \\ \exp(rT)E[D_1] & E[S_TD_1] & E[D_1^2] & \cdots & E[D_nD_1] \\ \exp(rT)E[D_2] & E[S_TD_2] & E[D_1D_2] & \cdots & E[D_nD_2] \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \exp(rT)E[D_n] & E[S_TD_n] & E[D_1D_n] & \cdots & E[D_n^2] \end{bmatrix} \begin{bmatrix} a \\ b \\ c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} E[W_T^*] \\ E[W_T^*P_T] \\ E[W_T^*D_1] \\ E[W_T^*D_2] \\ \vdots \\ E[W_T^*D_n] \end{bmatrix} \quad (1.3.14)$$

or, in matrix notation:

$$\Sigma \eta = \epsilon. \quad (1.3.15)$$

Inverting (1.3.15) to compute

$$\hat{\eta} = \Sigma^{-1}\epsilon \quad (1.3.16)$$

and then substituting  $\hat{\eta} \equiv [\hat{a} \ \hat{b} \ \hat{c}_1 \ \cdots \ \hat{c}_n]'$  into the objective function (1.3.11) yields the optimal value for a given sub-problem. Repeating this procedure for all  $\binom{N}{n}$  sub-problems and selecting the best of these solutions gives an approximate solution to (1.3.11)–(1.3.13).

However, for some utility functions, it is necessary to impose the solvency constraints (1.3.10), in which case the solution to the sub-problem cannot be simplified according to (1.3.16).



### 1.3.3 Minimizing Weighted Mean-Squared Error

A third alternative to the two approaches outlined in Sections 1.3.1 and 1.3.2 is to maximize expected utility but where we substitute an approximation for the utility function. This yields a *weighted* mean-squared-error objective function where the weighting function is the second derivative of the utility function evaluated at the optimal end-of-period wealth  $W_T^*$ . This is a hybrid of the two approaches proposed above that provides important economic motivation for mean-squared error, and approximates the direct approach of maximizing expected utility described in Section 1.3.1.

Specifically, consider the sub-problem of Section 1.3.1 in which we maximize expected utility holding fixed the strike prices  $\{k_i\}$ :

$$\text{Max}_{\{a,b,c_i\}} \text{E}[U(V_T)]$$

subject to the budget (1.3.7) and solvency constraints (1.3.10). Take a Taylor expansion of  $U(W_T^* \pm \lambda(W_T^* - V_T))$  about the global optimal  $W_T^*$ :

$$\begin{aligned} \text{E}[U(W_T^* \pm \lambda(W_T^* - V_T))] &\approx \text{E}[U(W_T^*)] \pm \lambda \text{E}[(W_T^* - V_T)U'(W_T^*)] + \\ &\quad \frac{\lambda^2}{2} \text{E}[(W_T^* - V_T)^2 U''(W_T^*)]. \end{aligned} \quad (1.3.17)$$

If  $V_T$  were “budget feasible”, by which we mean that  $\exp(-rT)\text{E}^Q[V_T] = W_0$ , and  $V_T$  were sufficiently close to  $W_T^*$ , then this implies that  $\pm\lambda(W_T^* - V_T)$  is a feasible direction of travel from  $W_T^*$ . For sufficiently small  $\lambda$ , (1.3.17) implies that

$$\text{E}[(W_T^* - V_T)U'(W_T^*)] = 0$$

under certain regularity conditions. Therefore, maximizing  $\text{E}[U(V_T)]$  should be equivalent to maximizing

$$\frac{1}{2} \text{E}[(W_T^* - V_T)^2 U''(W_T^*)] \quad (1.3.18)$$

for  $V_T$  sufficiently close to  $W_T^*$ . This gives rise to a third approach to the buy-and-hold asset-allocation problem, one that involves approximating  $W_T^*$  in mean-square rather than explicitly maximizing expected utility:

$$\text{Min}_{\{a,b,c_i,k_i\}} \text{E}[-U''(W_T^*)(W_T^* - V_T)^2] \quad (1.3.19)$$

subject to

$$V_T \equiv a \exp(rT) + b P_T + c_1 D_1 + c_2 D_2 + \dots + c_n D_n \quad (1.3.20)$$

$$W_0 = \exp(-rT) E^Q[V_T] . \quad (1.3.21)$$

For CRRA utility, we still need to impose solvency constraints, but even with such constraints we can solve the sub-problem much more quickly in the weighted mean-squared error case than in the maximization of expected utility proposed in Section 1.3.1. Indeed, the computational challenges for the weighted mean-squared error approach are comparable to the mean-squared error approach of Section 1.3.2.

A potential difficulty with the utility-weighted mean-squared-error approach is that some of the expectations in (1.3.19) may not be defined. Even when the expectations are defined, it is possible that some of them are very difficult to compute when they are ill-conditioned, i.e., “close” to being undefined. In such cases the approach either will not work or will be very difficult to implement. This typically occurs for low values of relative risk-aversion. Fortunately, it is precisely for low values of risk aversion that a direct maximization of expected utility works best. The reason is that the discretization of the support of  $P_T$  leads to approximation errors that can be extreme for high values of risk aversion. In particular, the discretized distribution has finite support, hence the optimal buy-and-hold strategy obtained with this distribution may perform poorly outside this finite support. The power-law specification of CRRA preferences will magnify small approximation errors of this type when the risk-aversion parameter is large.

Therefore, the maximization of expected utility and the minimization of utility-weighted mean-squared-error complement each other. As we will see in Section 1.5, when both approaches work well, they result in almost identical portfolios and certainty equivalents. Therefore, in the numerical examples of Section 1.5, we will maximize expected utility for low values of relative risk-aversion and minimize utility-weighted mean-squared-error for higher values when computing the utility-optimal buy-and-hold portfolios.

## 1.4 Three Leading Cases

To derive the optimal buy-and-hold portfolios according to the three criteria of Section 1.3, we require a few auxiliary results that depend on the specific utility function of the investor and the stochastic process for stock prices. In this section, we derive these results for CRRA and CARA utility under three leading cases for the stock-price process: geometric Brownian motion (Section 1.4.1), the trending Ornstein-Uhlenbeck process (Section 1.4.2), a bivariate linear diffusion process with a stochastic mean-reverting drift (Section 1.4.3).

In the case of geometric Brownian motion, the required results are straightforward—we are able to characterize  $W_T^*$  explicitly for both CRRA and CARA preferences, and all three approaches to the optimal buy-and-hold portfolio can be readily implemented. However, for the other two stochastic processes, the optimal dynamic asset-allocation strategies are path dependent, which implies that no buy-and-hold portfolio of stocks, bonds, and European call options can ever achieve the same certainty equivalents as the optimal dynamic strategies.

In such situations, we propose an alternative to  $W_T^*$  as a target for the optimal buy-and-hold portfolio, and derive this alternative explicitly in Sections 1.4.2 and 1.4.3.

### 1.4.1 Geometric Brownian Motion

In the case of geometric Brownian motion, the stock price  $P_t$  satisfies the following stochastic differential equation (SDE):

$$dP_t = \mu P_t dt + \sigma P_t dB_t \quad (1.4.1)$$

where  $B_t$  is a standard Brownian Motion. Recall that the standard asset-allocation problem in the absence of derivatives is given by (1.3.2)–(1.3.3):

$$\text{Max}_{\{\omega_t\}} E[U(W_T)]$$

subject to the budget equation

$$dW_t = [r + \omega_t(\mu - r)]W_t dt + \omega_t W_t \sigma dB_t$$

where  $\omega_t$  is the fraction of the investor's portfolio invested in the stock at time  $t$  (see Merton, 1969, 1971 for a more detailed exposition). For concreteness, we consider two specific utility functions: constant absolute risk-aversion (CARA) and constant relative risk-aversion (CRRA) utility. These are well-known utility functions for which we have closed-form solutions to the standard asset-allocation problem. In particular, for CRRA utility, we have:

$$U(W_T) = \frac{W_T^\gamma}{\gamma} \quad (1.4.2)$$

$$W_T^* = W_0 \exp\left(rT - \frac{\xi^2 T (2\gamma - 1)}{2(1 - \gamma)^2} + \frac{\xi B_T}{(1 - \gamma)}\right) \quad (1.4.3)$$

$$\omega_t^* = \frac{\mu - r}{(1 - \gamma)\sigma^2} \quad (1.4.4)$$

and for CARA utility,

$$U(W_T) = -\frac{\exp(-\gamma W_T)}{\gamma} \quad (1.4.5)$$

$$W_T^* = \frac{\gamma W_0 \exp(rT) + \xi^2 T + \xi B_T}{\gamma}, \quad \xi \equiv \frac{\mu - r}{\sigma} \quad (1.4.6)$$

$$\omega_t^* = \frac{\exp(-r(T-t))\xi}{\gamma\sigma W_t} \quad (1.4.7)$$

Given these closed-form solutions, we can make explicit comparisons of the optimal buy-and-hold portfolio of stocks, bonds and options with the standard optimal asset-allocation strategies for the two utility functions.

## 1.4.2 The Ornstein-Uhlenbeck Process

If stock prices are predictable to some degree, the asset-allocation problem becomes considerably more challenging since the optimal investment strategy is path-dependent. This implies that  $W_T^*$  is also path-dependent and extremely difficult to compute explicitly, hence the mean-squared-error approaches of Sections 1.3.2 and 1.3.3 are not feasible. However, in certain cases, it is possible to derive an upper bound on the certainty equivalent of the optimal buy-and-hold portfolio of stocks, bonds and options, which provides some indication of the benefits of options in replicating dynamic investment strategies. We present such an upper bound in this section for the case where log-prices  $X_t \equiv \log P_t$  follow a trending Ornstein-Uhlenbeck process:<sup>13</sup>

$$dX_t = [-\delta(X_t - \mu t - X_0) + \mu]dt + \sigma dB_t, \quad \delta > 0. \quad (1.4.8)$$

which has the solution:

$$X_t = X_0 + \mu t + \sigma \exp(-\delta t) \int_0^t \exp(\delta s) dB_s. \quad (1.4.9)$$

The solution to the standard asset-allocation problem (1.3.2)–(1.3.3) in this case is characterized by the following Hamilton-Jacobi-Bellman equation:

$$0 = \text{Max}_{\omega_t} \left\{ J_t + W_t J_W \left( r + \omega_t [-\delta(X_t - \mu t - X_0) + \mu + \frac{\sigma^2}{2} - r] \right) + J_X (-\delta(X_t - \mu t - X_0) + \mu) + \frac{1}{2} \omega_t^2 \sigma^2 W_t^2 J_{WW} + \frac{1}{2} \sigma^2 J_{XX} + \sigma^2 \omega_t W_t J_{XW} \right\} \quad (1.4.10)$$

<sup>13</sup>See Lo and Wang (1995) for a more detailed exposition of its properties. We also derive results for the standard Ornstein-Uhlenbeck process (without trend), which are included in the Appendix.

where

$$J(W_t, X_t, t) \equiv \text{Max}_{\omega_t} E_t[U(W_T)] . \quad (1.4.11)$$

The solutions to (1.4.10) for CRRA and CARA utility are given in the Appendix.

Because  $W_T^*$  is path-dependent in this case, even if we allow the number of options  $n$  in the buy-and-hold portfolio to increase without bound, the certainty equivalent of the buy-and-hold portfolio will never approach the certainty equivalent  $W_T^*$ . However, an upper bound on the certainty equivalent of any buy-and-hold portfolio can be derived by allowing the investor to purchase an unlimited number of options at all possible strike prices. The certainty equivalent of the end-of-period wealth in this case, which we denote by  $V_T^\infty$ , is clearly an upper bound for any buy-and-hold portfolio containing a finite number  $n$  of options.

To derive  $V_T^\infty$ , we require the conditional state-price density of the terminal stock price  $P_T$ , defined as:

$$\pi_T^b \equiv E[\pi_T | P_T = b] \quad (1.4.12)$$

where  $\pi_T$  is the unconditional state-price density of the terminal stock price.<sup>14</sup> The economic interpretation of  $\pi_T^b$  is the price per unit probability of 1 unit of wealth at time  $T$  in the event that  $P_T = b$ . By definition,  $\pi_T^b$  is given by:

$$\pi_T^b = E[\pi_T | P_T = b] = \frac{E[\pi_T 1_{\{P_T=b\}}]}{E[1_{\{P_T=b\}}]} . \quad (1.4.13)$$

The numerator of (1.4.13) is computed by applying Girsanov's Theorem and noting that the Radon-Nikodym derivative  $dQ/dF$  of the equivalent martingale measure  $Q$  with respect to the true probability measure  $F$  is equal to  $\exp(rT)\pi_T$ . Under  $Q$ , the stock price at time  $T$  is given by

$$P_T^Q = \exp(Z_T) \equiv P_0 \exp\left(\left(r - \frac{\sigma^2}{2}\right)T + \sigma \widehat{B}_T\right) . \quad (1.4.14)$$

where  $\widehat{B}_T$  is a standard Brownian motion under  $Q$ . Under the true probability measure,  $F$ ,

---

<sup>14</sup>See Duffie (1996) for a more detailed exposition of state-price densities.

recall that the stock price at time  $T$  is given by

$$P_T = \exp(X_T) \equiv \exp\left(X_0 + \mu T + \sigma e^{-\delta T} \int_0^T e^{\delta s} dB_s\right). \quad (1.4.15)$$

With this in mind, we can write (1.4.13) as

$$\pi_T^b = \frac{\exp(-rT) f_{P_T}^Q(b)}{f_{P_T}(b)} \quad (1.4.16)$$

where  $f_{P_T}$  and  $f_{P_T}^Q$  denote the log-normal density functions of  $P_T$  under  $F$  and  $Q$  respectively. Simplifying (1.4.16) yields:

$$\pi_T^b = \left(\frac{\sigma_x}{\sigma_z}\right) \exp\left(-rT - \frac{1}{2} \left[ \left(\frac{\log b - \mu_x}{\sigma_z}\right)^2 - \left(\frac{\log b - \mu_x}{\sigma_x}\right)^2 \right]\right) \quad (1.4.17)$$

where

$$\begin{aligned} \mu_x &= X_0 + \mu T & , & \quad \sigma_x^2 = \frac{\sigma^2}{2\delta} (1 - \exp(-2\delta T)) \\ \mu_z &= X_0 + \left(r - \frac{\sigma^2}{2}\right) T & , & \quad \sigma_z^2 = \sigma^2 T. \end{aligned} \quad (1.4.18)$$

Using  $\pi_T^b$  as the state-price density process, we can derive the optimal buy-and-hold portfolio in which options of all possible strikes may be included. Using the approach proposed in Cox and Huang (1989) for the case of CRRA utility, the problem reduces to:

$$\max E \left[ \frac{(V_T)^\gamma}{\gamma} \right] \quad \text{subject to} \quad E [\pi_T^b V_T] = W_0 \quad (1.4.19)$$

which has the solution:

$$V_T^\infty = \frac{W_0 (\pi_T^b)^{\frac{1}{\gamma-1}}}{E \left[ (\pi_T^b)^{\frac{\gamma}{\gamma-1}} \right]} \quad (1.4.20)$$

where

$$\mathbb{E} \left[ \left( \pi_T^b \right)^{\frac{\gamma}{\gamma-1}} \right] = \frac{\sigma_o}{\sigma_x} \left( \frac{\sigma_x}{\sigma_z} \right)^{\frac{\gamma}{\gamma-1}} \exp \left( \frac{-rT\gamma}{\gamma-1} + \frac{\gamma [\mu_x - \mu_z]^2}{2[\gamma-1][\gamma\sigma_x^2 - \sigma_z^2]} \right)$$

and

$$\sigma_o^2 = \frac{\sigma_x^2 \sigma_z^2 (\gamma - 1)}{(\gamma \sigma_x^2 - \sigma_z^2)}. \quad (1.4.21)$$

This, in turn, implies:

$$\begin{aligned} U^\infty &\equiv \mathbb{E} \left[ \frac{(V_T^\infty)^\gamma}{\gamma} \right] = \frac{W_0^\gamma}{\gamma} \mathbb{E} \left[ \left( \pi_T^b \right)^{\frac{\gamma}{\gamma-1}} \right]^{1-\gamma} \\ \text{CE}(V_T^\infty) &= (\gamma U^\infty)^{\frac{1}{\gamma}} \end{aligned}$$

where  $\text{CE}(\cdot)$  denotes the certainty equivalent operator.

The case of CARA utility can also be handled in a similar manner.

Having solved for the optimal buy-and-hold portfolio and its certainty equivalent in the infinite options case, we can now compare this upper bound to the optimal buy-and-hold portfolios with a finite number of options. We use the same method as in the geometric Brownian motion case (see Section 1.4.1), hence we omit the details.

### 1.4.3 A Bivariate Linear Diffusion Process

We now turn to a third set of price dynamics for  $P_t$ , one in which there are two sources of uncertainty, implying that markets are incomplete. Nevertheless, we are still able to compute optimal buy-and-hold portfolios of stocks, bonds and options, and can also derive the upper bound to the buy-and-hold certainty equivalents as in Section 1.4.2. Specifically, let  $X_t \equiv \log P_t$  satisfy the following bivariate linear diffusion process:

$$dX_t = \left( \mu_t - \frac{\sigma_1^2}{2} \right) dt + \sigma_1 dB_{1t} \quad (1.4.22)$$

$$d\mu_t = \kappa (\theta - \mu_t) dt + \sigma_2 dB_{2t} \quad (1.4.23)$$

where  $B_{1t}$  and  $B_{2t}$  are two standard Brownian motions with instantaneous correlation coefficient  $\rho$ . Kim and Omberg (1996, 1998) derive the optimal value function for the standard asset-allocation problem with these price dynamics for an investor with HARA utility. Despite the fact that markets are incomplete, it is clear that options can be replicated using

trading strategies in only the stock and the bond,<sup>15</sup> hence options can be priced by arbitrage in this case. Therefore, we can perform the same analysis for these dynamics as we did for geometric Brownian motion in Section 1.4.1 and the Ornstein-Uhlenbeck process in Section 1.4.2.

To derive  $V_T^\infty$  for the bivariate diffusion (1.4.22)–(1.4.23), we perform a similar set of calculations as in Section 1.4.2. We begin by solving (1.4.22) and observing that  $P_T$  is lognormally distributed with parameters:

$$\mu_X = X_0 + \left(\theta - \frac{\sigma_1^2}{2}\right)T + \frac{\theta - \mu_o}{\kappa}(\exp(-\kappa T) - 1) \quad (1.4.24)$$

$$\sigma_x^2 = \sigma_1^2 T + \frac{2\sigma_1\sigma_2\rho}{\kappa} \left[ T + \frac{\exp(-\kappa T)}{\kappa} - \frac{1}{\kappa} \right] + \frac{\sigma_2^2}{\kappa^3} \left[ T\kappa - \frac{3}{2} + 2\exp(-\kappa T) - \frac{\exp(-2\kappa T)}{2} \right]. \quad (1.4.25)$$

The conditional state-price density then follows in the same manner as (1.4.17):

$$\pi_T^b = \left(\frac{\sigma_x}{\sigma_z}\right) \exp\left(-rT - \frac{1}{2} \left[ \left(\frac{\log b - \mu_z}{\sigma_z}\right)^2 - \left(\frac{\log b - \mu_x}{\sigma_x}\right)^2 \right]\right) \quad (1.4.26)$$

where

$$\mu_z = X_0 + \left(r - \frac{\sigma^2}{2}\right)T, \quad \sigma_z^2 = \sigma^2 T \quad (1.4.27)$$

With the conditional state-price density in hand,  $V_T^\infty$  and its certainty equivalent are readily derived.

## 1.5 Numerical Results

To illustrate the practical relevance of our optimal buy-and-hold portfolio, we provide numerical results in this section for CRRA preferences under each of the three stochastic processes of Section 1.4 using the nonlinear programming solver LOQO and the algebraic mathematical programming language AMPL.<sup>16</sup> Before turning to those results, we begin with a simple

<sup>15</sup>For further discussion, see Lo and Wang (1995).

<sup>16</sup>AMPL is described in Fourer, Gay, and Kernighan (1999). Information on LOQO can be obtained from <http://www.princeton.edu/loqo/>.



example to motivate our analysis. Let

$$U(W_T) = \frac{W_T^\gamma}{\gamma} , \quad \gamma = -4 , \quad W_0 = \$100,000 , \quad T = 20 \text{ years}$$

$$P_0 = \$50 , \quad r = .05 , \quad \mu = .15 , \quad \sigma = .20$$

which implies a relative risk-aversion coefficient of 5, a portfolio weight  $\omega_t^*$  of 50% for the stock in the optimal dynamic asset-allocation policy (1.4.4), and a certainty equivalent of \$448,169 for  $W_T^*$ . Now consider the problem of constructing an optimal buy-and-hold portfolio containing stocks, bonds, and a maximum of 2 options, assuming that there are only 4 possible options to choose from, with the following strikes:

$$k_1 = \$176 , \quad k_2 = \$976 , \quad k_3 = \$1,775 , \quad k_4 = \$2,575 .$$

For the approach outlined in Section 1.3.1, we maximize the expected utility:

$$\text{Max}_{\{a,b,c_i,k_i\}} E[U(V_T)] \quad \text{subject to}$$

$$V_T \equiv a \exp(rT) + b P_T + c_1 D_1 + c_2 D_2$$

$$W_0 = \exp(-rT) E^Q[V_T]$$

and the corresponding solvency constraints. We discretize the support of  $P_T$  using a grid of 4,000 points, chosen in such a way that the weight associated with each point in the objective function is equal to 1/4000. A direct optimization then yields the following certainty equivalents for sub-problems of the optimal buy-and-hold problem for the various combinations of strikes:

Options Used:	1 and 2	1 and 3	1 and 4	2 and 3	2 and 4	3 and 4	(1.5.1)
CE( $V_T^*$ ):	\$447,307	\$447,137	\$447,067	\$437,971	\$437,850	\$436,506	

From (1.5.1), it is apparent that the optimal buy-and-hold strategy is to use options with strikes  $k_1 = 176$  and  $k_2 = 976$ , and the optimal portfolio positions are:

$$a^* = \$36,097 , \quad b^* = 1,521 , \quad c_1^* = -907 , \quad c_2^* = -353 . \quad (1.5.2)$$

With only two options, the optimal buy-and-hold portfolio yields an estimated certainty equivalent of \$447,307,<sup>17</sup> which is 99.8% of the certainty equivalent of the optimal dynamic asset-allocation strategy, a strategy that requires continuous trading over a 20-year period!

Note that the portfolio weights implied by the positions (1.5.2) are 36.1% in bonds, 76.1% in stocks, and -12.2% in options. The optimal buy-and-hold portfolio consists of shorting options 1 and 2, and investing the proceeds—approximately \$12,100—in stocks and bonds along with the initial wealth of \$100,000.

Alternatively, we can minimize the mean-squared error between  $V_T$  and  $W_T^*$  according to Section 1.3.2:

$$\text{Min}_{\{a,b,c_i,k_i\}} E[(W_T^* - V_T)^2] \quad \text{subject to}$$

$$\begin{aligned} V_T &\equiv a \exp(rT) + b P_T + c_1 D_1 + c_2 D_2 \\ W_0 &= \exp(-rT) E^Q[V_T] \end{aligned}$$

and also subject to the solvency constraints (1.3.10). The root-mean-squared-error (RMSE) (as a percentage of  $E[W_T^*]$ ) of each of the sub-problems is given by:

Options Used:	1 and 2	1 and 3	1 and 4	2 and 3	2 and 4	3 and 4
RMSE (%):	6.27	4.73	5.69	6.47	5.75	9.95

Under the mean-squared-error criterion, the optimal buy-and-hold portfolio consists of a different set of options than under the expected-utility criterion—in this case, options 1 and 3—and the optimal positions are:

$$a^* = \$20,928 \quad , \quad b^* = 1,980 \quad , \quad c_1^* = -1,508 \quad , \quad c_2^* = -291 \quad . \quad (1.5.3)$$

With such a buy-and-hold portfolio, the root-mean-squared-error is 4.73% of the expected value of  $W_T^*$ , and the certainty equivalent of this portfolio is \$436,034, which is 97.3% of the certainty equivalent of the optimal dynamic asset-allocation strategy. Despite the fact that (1.5.3) is only an indirect method of approximating  $W_T^*$ , the certainty equivalent is almost identical to that of the optimal dynamic strategy. The portfolio weights corresponding to (1.5.3) are 20.9% in bonds, 99.0% in stocks, and -19.9% in options.

Finally, if we minimize the weighted mean-squared-error according to Section 1.3.3,

$$\text{Min}_{\{a,b,c_i,k_i\}} E[-U''(W_T^*)(W_T^* - V_T)^2]$$

<sup>17</sup>The estimation error is due to the discretization of the distribution of  $P_T$ . Once we obtain the strategy (1.5.2), we can compute the certainty equivalent exactly, and in this case, it is \$446,034, which is 99.5% of the certainty equivalent of the optimal dynamic asset-allocation strategy.

subject

$$\begin{aligned} V_T &\equiv a \exp(rT) + b P_T + c_1 D_1 + c_2 D_2 \\ W_0 &= \exp(-rT) E^Q[V_T] \end{aligned}$$

and the solvency constraints (1.3.10), we obtain the following weighted RMSE's for the various sub-problems:

Options Used:	1 and 2	1 and 3	1 and 4	2 and 3	2 and 4	3 and 4
Weighted RMSE	0.738	0.764	0.777	1.830	1.839	2.013

which yields an optimal buy-and-hold portfolio containing options 1 and 2 and positions:

$$a^* = \$35,321 \quad , \quad b^* = 1,523 \quad , \quad c_1^* = -930 \quad , \quad c_2^* = -349 \quad . \quad (1.5.4)$$

Although the weighted RMSE of the optimal buy-and-hold portfolio, 0.738, is somewhat difficult to interpret, the certainty of equivalent of the portfolio is \$445,967 which is 99.5% of the certainty equivalent of the optimal dynamic asset-allocation strategy. With portfolio weights of 35.3% in bonds, 76.2% in stocks, and -11.5% in options, the minimum utility-weighted mean-squared-error approach yields an almost-identical solution to the maximum expected-utility approach (recall that the portfolio weights of the latter are 36.1% in bonds, 76.1% in stocks, and -12.2% in options). Therefore, the hybrid approach provides an excellent approximation to the maximization of expected utility.

In Sections 1.5.1–1.5.3, we perform more computationally intensive optimizations for the three stochastic processes of Section 1.4 under CRRA preferences using the three approaches described in Section 1.3: maximizing expected utility, and minimizing mean-squared error and weighted mean-squared error. In particular, for each stochastic process, we compute two optimal buy-and-hold portfolios for each of six different values of the relative risk aversion coefficient (RRA = 1, 2, 5, 10, 15, 20): a utility-optimal buy-and-hold portfolio obtained by either direct maximization of expected utility or minimization of utility-weighted mean-squared error (as in Sections 1.3.1 and 1.3.3, respectively), and a mean-square-optimal buy-and-hold portfolio (as in Section 1.3.2). For each stochastic process and each value of the relative risk-aversion coefficient, we consider  $N = 45$  possible strike prices and up to  $n = 3$  options for the utility-optimal buy-and-hold portfolios and up to  $n = 5$  options for the mean-square-optimal buy-and-hold portfolios. This yields up to  $\binom{45}{3} = 14,190$  and  $\binom{45}{5} = 1,221,759$  sub-problems for each of the two optimizations, respectively.

The strikes are selected in the following way. Letting  $\mu_x$  and  $\sigma_x$  denote the mean and variance of  $X_T \equiv \log P_T$ , we partition the interval  $[\exp(\mu_x - 3\sigma_x), \exp(\mu_x + 3\sigma_x)]$  into 45 evenly spaced points which we denote by  $s_1 \equiv \exp(\mu_x - 3\sigma_x), \dots, s_{45} \equiv \exp(\mu_x + 3\sigma_x)$ . We then use these points as our strikes,  $k_i = s_i$ ,  $i = 1, \dots, 45$ . Such a procedure for choosing the set of strikes  $\{k_i\}$  is simple to implement, however, more sophisticated methods can be

employed to improve the performance of the overall optimization process.

To facilitate comparisons across different optimal buy-and-hold portfolios we use one set of 45 strikes for each of the three stochastic processes considered in Sections 1.5.1–1.5.3, i.e., for each stochastic process, we construct one set of 45 strikes and keep these fixed as we vary the values of relative risk-aversion and the number of options  $n$  in the buy-and-hold portfolio. This is clearly suboptimal—for example, when  $n = 1$ , we can optimize the buy-and-hold portfolio over several thousand possible strike prices very quickly—but holding the strikes fixed allows us to gauge the impact of other parameters such as the risk-aversion coefficient and the number of options on the objective function being optimized. In practical applications, the set of possible strikes should be optimized for each specification of the buy-and-hold problem; in our limited experience, simple heuristics for optimizing the set of strikes can lead to substantial improvements in overall performance.

For each of the three cases considered in Sections 1.5.1–1.5.3, we maintain the following set of assumptions:

$$\begin{aligned}
 U(W_T) &= \frac{W_T^\gamma}{\gamma} \\
 \gamma &= 0, -1, -4, -9, -14, -19 \\
 W_0 &= \$100,000, \quad T = 20 \text{ years} \\
 P_0 &= \$50, \quad r = 0.05
 \end{aligned}
 \tag{1.5.5}$$

$$E[\log(P_t/P_{t-1})] = 0.15$$

$$\text{Var}[\log(P_t/P_{t-1})] = 0.20^2$$

where the values of  $\gamma$  correspond to relative risk-aversion coefficients of 1, 2, 5, 10, 15, and 20, respectively.

### 1.5.1 Geometric Brownian Motion

For geometric Brownian motion (1.4.1), we set the parameters  $(\mu, \sigma)$  to match the mean and variance of continuously compounded returns specified in (1.5.5). Based on our algorithm for constructing the set of strike prices from the distribution of  $\log P_T$ , we select the strikes

for our  $n$  options from among the following 45 possibilities (in dollars):

69	401	733	1,066	1,398	1,731	2,063	2,396	2,728
3,061	3,393	3,725	4,058	4,390	4,723	5,055	5,388	5,720
6,052	6,385	6,717	7,050	7,382	7,715	8,047	8,379	8,712
9,044	9,377	9,709	10,042	10,374	10,706	11,039	11,371	11,704
12,036	12,369	12,701	13,033	13,366	13,698	14,031	14,363	14,696

### Utility-Optimal Buy-and-Hold Portfolios

Table 1a reports the utility-optimal buy-and-hold portfolios for various levels of risk aversion and, for each risk-aversion parameter, for the number of options  $n$  varying from 0 to 3. For example, the first panel of Table 1a contains results for the log-utility case ( $\gamma = 0$ , or  $\text{RRA} = 1$ ). This is a very low level of risk aversion—by most empirical and experimental accounts, an unrealistically low level—and implies that the investor’s objective is to maximize the expected geometric average rate of return of his portfolio. Examples of investors with such preferences are proprietary traders and hedge-fund managers. The results for the  $\text{RRA} = 1$  panel were obtained by maximizing expected utility directly using a discretized distribution for  $P_T$  (see Section 1.3.1). The results for the remaining five panels of Table 1a were obtained by minimizing the utility-weighted mean-squared error (see Section 1.3.3).

The first row of Table 1a’s first panel corresponds to the optimal buy-and-hold portfolio with no options ( $n = 0$ )—for log-utility, the optimal portfolio is to put 100% of the investor’s wealth into the stock.<sup>18</sup> Not surprisingly, the certainty equivalent of such a strategy is only 20.2% of the certainty equivalent of the optimal dynamic strategy  $\text{CE}(W_T^*)$ . By not allowing the investor to trade at all over the 20-year period, and without access to any options, the investor’s welfare is reduced by approximately 80%. As the number of options is increased, his welfare increases so that for  $n = 3$  options the certainty equivalent of the optimal buy-and-hold is 92.2% of  $\text{CE}(W_T^*)$ .

For log utility, it is interesting to note that the RMSE is approximately 3,650% even for  $n = 3$  and despite the fact that the certainty equivalent of the optimal buy-and-hold portfolio is close to that of the optimal dynamic investment policy. This, and the very slow rate at which the RMSE decreases as we increase  $n$  from 0 to 3, suggests that it may be possible to obtain an excellent approximation to the optimal dynamic strategy—in terms of expected utility—without being able to approximate  $W_T^*$  very well in mean-square.

Note that within each relative risk-aversion panel of Table 1a, the RMSEs decrease monotonically as the number of options  $n$  increases from 0 to 3. This, of course, need not be the case since we are maximizing expected utility, not minimizing RMSE. In fact, it is quite possible for the RMSE to increase as we increase  $n$ . However, the fact that they do decrease monotonically suggests that there is some correlation between smaller RMSE and a more

<sup>18</sup>In fact, in the absence of solvency constraints, the optimal portfolio weight for the stock would be much greater than 100%, i.e., for CRRA preferences, the solvency constraints are binding.

preferred buy-and-hold portfolio. Of course, as  $n$  becomes arbitrarily large, the RMSE must converge to 0.

Perhaps the most interesting feature of Table 1a is how the results fall naturally into two distinct groups. The first group consists of the first two panels, corresponding to investors who are not very risk averse (relative risk-aversion coefficients of 1 and 2, respectively) and who, in the standard dynamic asset-allocation framework, would optimally hold more than 100% of their wealth in the risky asset. In a buy-and-hold portfolio without options ( $n=0$ ), these investors are bound by the solvency constraints (1.3.10), making it difficult for them to approximate  $CE(W_T^*)$  very well (the certainty equivalents  $CE(V_T^*)$  of the optimal buy-and-hold portfolios are only 20.2% of  $CE(W_T^*)$  for the log-utility investor and 81.9% for the investor with  $RRA=2$ ). Options are of particular benefit to these investors, who purchase call options so that they can increase their risk exposure.<sup>19</sup> They do not invest in bonds at all, but divide their wealth between stocks and options. As the number of options allowed increases, the fraction of wealth devoted to options in the optimal buy-and-hold portfolio for the log-utility investor also increases, from 60.4% for  $n=1$  to 99.3% for  $n=3$ . For a relative risk-aversion coefficient of 2, the proportion of the optimal buy-and-hold portfolio devoted to options declines slightly as  $n$  increases, apparently stabilizing at approximately 59% for  $n=3$ .

The second group consists of the remaining four panels, which correspond to investors who, in the standard dynamic asset-allocation framework, would optimally hold less than 100% of their wealth in the risky asset. For these investors a buy-and-hold portfolio with no options has a certainty equivalent that is approximately 97% of  $CE(W_T^*)$ . It is remarkable that a well chosen buy-and-hold portfolio in the stock and the bond can do so well over a 20-year horizon.

When just 1 or 2 options are added to the buy-and-hold portfolio in these cases, the certainty equivalents  $CE(V_T^*)$  of the optimal portfolios increase to approximately 99.7% of  $CE(W_T^*)$ . In contrast to the first two panels, investors with higher risk-aversion parameters are net sellers of call options, forgoing some of the upside gain in order to limit losses on the downside. The value of these option positions ranges from 24% to 37% of their initial wealth. The optimal buy-and-hold portfolios invest the option premia, together with the initial wealth of \$100,000, in stocks and bonds.

The combination of a short position in a call option and a long position in the underlying stock is often called a "hedged position" since the gains (losses) of one security offset to some degree the losses (gains) of the other. Figure 1-1 provides an example of such a hedged position: a long position in one share of stock and a short position in a call option on that stock with strike price  $k$ . The combination yields a payoff that has limited upside—beyond  $k$ , the payoff is constant at  $k$ —which a sufficiently risk-averse investor might find attractive, since he receives cash now in exchange for an uncertain upside.

For risk-aversion coefficients greater than or equal to 5, Table 1a shows that the optimal buy-and-hold portfolios all include hedged positions in which part of the upside potential in the stock is relinquished in exchange for option premia that are invested in stocks and bonds. For a relative risk-aversion coefficient of 10, the optimal buy-and-hold portfolio with 3 options consists of a -37.6% investment in options, 90.1% in the stock, and 47.5% in bonds.

---

<sup>19</sup>Call options are generally more risky than the underlying stock on which they are based. See, for example, Cox and Rubinstein (1985).

Since this portfolio yields an excellent approximation to the optimal dynamic investment strategy (it has a certainty equivalent  $CE(V_T^*)$  of 99.7%), we can be fairly confident that these rather unorthodox positions do, in fact, accurately represent the investor's preferences. Indeed, by graphing the payoff diagram of this optimal buy-and-hold portfolio along the lines of Figure 1-1, we can obtain a visual representation of the investor's dynamic risk exposures at a single point in time.

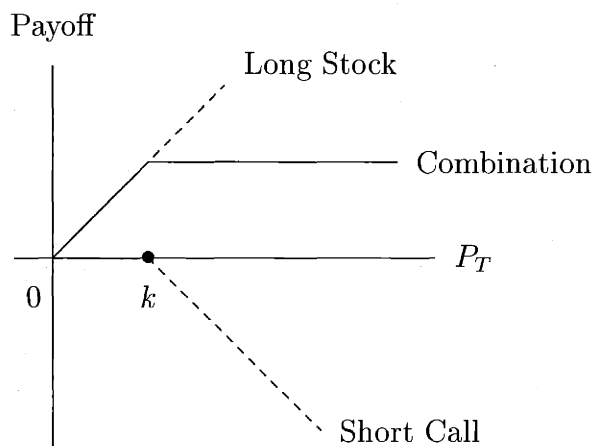


Figure 1-1: Payoff diagram of hedged position (long stock and short call).

A common characteristic in all of the panels of Table 1a is the optimal strike prices of the options in the buy-and-hold portfolio. Despite the fact that the possible strikes range from \$69 to \$14,696, the highest strike selected by the optimization algorithm is \$1,731. Under geometric Brownian motion, the expected stock price 20 years into the future is:

$$E_0[P_T] = P_0 \exp(\mu T) = \$50 \times \exp(0.17 \times 20) = \$1,498 .$$

Therefore, almost all of the options selected by the optimal buy-and-hold portfolio are in-the-money relative to the expected terminal price  $E_0[P_T]$ , which characterizes another aspect of the investor's risk profile.

Also, the fact that among the 45 possible strikes, only 5 are employed in the optimal buy-and-hold portfolios over the range of relative risk-aversion coefficients from 1 to 20 suggests the possibility of standardizing a small number of "canonical" long-dated options that will appeal to a broad set of investors.

### Mean-Square-Optimal Buy-and-Hold Portfolios

Table 1b reports the mean-square-optimal buy-and-hold portfolios for various levels of risk aversion and, for each risk-aversion parameter, for the number options  $n$  varying from 0 to 5. We use a larger number of options in this case to illustrate the fact that even with a

larger number of options, a mean-square-optimal portfolio need not come close in certainty equivalence to the optimal dynamic investment policy.

The first row of Table 1b's first panel corresponds to the optimal buy-and-hold portfolio with no options ( $n=0$ ), which is identical to the first row of Table 1a's first panel. As the number of options  $n$  is increased, the investor's welfare increases, so that for  $n=5$ , the certainty equivalent of the optimal buy-and-hold strategy is 34.9% of  $CE(W_T^*)$ . Although this is a considerable improvement over the  $n=0$  case, it is still quite far below the optimal dynamic strategy's certainty equivalent. This is not unexpected in light of the fact that we are minimizing mean-squared-error, not maximizing expected utility. As  $n$  increases beyond 5, this approximation will improve eventually, but the optimization process becomes considerably more challenging for larger  $n$ . For example, the  $n=15$  case involves  $\binom{45}{15} = 344,867,425,584$  sub-problems, and if each sub-problem requires 0.01 seconds to solve, the overall optimization would take approximately 109.4 years to complete.

Unlike Table 1a, in Table 1b the certainty equivalents of the optimal buy-and-hold portfolio,  $CE(V_T^*)$ , do not increase monotonically with the number of options  $n$ . For example, in the case of log utility ( $RRA=1$ ),  $CE(V_T^*)$  is 20.4% of  $CE(W_T^*)$  for  $n=1$  option, but declines to 10.2% for  $n=2$  options. This underscores the fact that we are minimizing mean-squared-error in the optimal buy-and-hold portfolios of Table 1b, not maximizing expected utility. In fact, it is possible for a buy-and-hold portfolio to exhibit a small RMSE and a small certainty equivalent at the same time.<sup>20</sup> Therefore, while RMSE must decline monotonically with  $n$ , the certainty equivalents need not. Of course, as the number of options  $n$  increases without bound,  $CE(V_T^*)$  will approach  $CE(W_T^*)$  eventually, even if not monotonically.

The option positions in the optimal buy-and-hold portfolios provide additional insight into the differences between maximizing expected utility and minimizing mean-squared-error in constructing the optimal buy-and-hold portfolio. As  $n$  increases from 0 to 1 in the first panel of Table 1b, the optimal buy-and-hold portfolio changes from 100% stocks to 99.8% stocks and 0.2% options, with a huge position (29.0 million) in the option with strike price \$14,696. Given a current stock price of \$50, this option is obviously deeply out-of-the-money, hence its price is extremely close to zero, so close that 29.0 million options amount to only 0.2% of the investor's initial portfolio. Moreover, recall that these are 20-year options, hence a strike price of \$14,696 should be compared not only with the current stock price but with the expected stock price at maturity,  $P_T$ . Recall that under geometric Brownian motion, the expected stock price 20 years into the future is \$1,498. Therefore, even taking into account the expected appreciation in the stock over the next 20 years, the strikes are still extraordinarily high.

The  $n=2$  case differs dramatically from the  $n=1$  case. When given the opportunity to include 2 options in the buy-and-hold portfolio, the optimal weights become 62.9% in options, -4.0% in the stock, and the remaining 41.1% in bonds. The optimal buy-and-hold portfolio involves shorting \$4,000 of the stock and putting the proceeds, as well as the original \$100,000, into bonds and options. The options component consists of two positions: 331,561 options with a strike of \$1,398, and 28.3 million options with a strike of \$14,696. The latter position is similar to that of the  $n=1$  case, and accounts for a relatively small part of the portfolio. The majority of the 69.2% allocated to options is due to the former position in

<sup>20</sup>This typically occurs when the buy-and-hold strategy results in a final wealth  $V_T^*$  that is close to zero over some interval of  $P_T$ .



options of a much lower strike price. The lower strike price implies a higher option price, hence the cost of 331,561 of these options dwarfs the cost of 28.3 million of the higher-strike options. While this buy-and-hold portfolio is indeed optimal from a mean-squared-error criterion, the certainty equivalent reported in Table 1b shows that the investor's welfare has actually declined by half, as compared to the  $n=1$  case. Moreover, the RMSE declines only slightly, suggesting that we treat this case cautiously and with a certain degree of skepticism.

As the investor's risk-aversion parameter increases, Table 1b shows that the optimal buy-and-hold portfolio performs considerably better in terms of certainty equivalence, in most cases attaining 90% or more of the certainty equivalent of the optimal dynamic strategy. For risk-aversion coefficients greater than 2, the RMSE of the buy-and-hold portfolio is less than 5% with only one or two options. The intuition for this pattern follows from the fact that investors with higher risk aversion invest a smaller proportion of their wealth in the stock market, hence their final wealth  $W_T^*$  has lower variance which makes it easier to approximate  $W_T^*$  with a buy-and-hold strategy.

The option positions in optimal buy-and-hold portfolios are also different for higher levels of risk aversion, consisting of fewer options and at lower strike prices. To see why, observe that for risk-aversion coefficients of 5 and greater, the optimal buy-and-hold portfolios with no options ( $n=0$ ) consist largely of bonds (75.6% in bonds for RRA=5, 95.3% for RRA=10, 96.3% for RRA=15, and 97.5% for RRA=20). When options are allowed in the buy-and-hold portfolios, additional risk-reduction possibilities become feasible and the optimization algorithm takes advantage of such opportunities. In particular, for risk-aversion levels of 5 and greater, the option positions are generally negative—the optimal buy-and-hold portfolios consist of selling options and investing the proceeds as well as the original \$100,000 initial wealth in stocks and bonds. For example, the third panel of Table 1b shows that with a risk-aversion coefficient of 5, the optimal buy-and-hold portfolio with 5 options is 59.0% in stocks, 43.1% in bonds, and -2.1% in options, with short positions in all 5 options, and where the optimal strikes range from \$401 to \$13,698. These results correspond well with those of Table 1a, in which the optimal buy-and-hold portfolios of investors with higher risk-aversion coefficients contained hedged positions (long positions in the stock and short positions in options).

## 1.5.2 The Ornstein-Uhlenbeck Process

To calibrate the parameters of the trending Ornstein-Uhlenbeck process (1.4.8), we observe that the moments of the stationary distribution of  $\{P_t\}$  are given by:

$$\begin{aligned} E[\log(P_t/P_{t-1})] &= \mu \\ \text{Var}[\log(P_t/P_{t-1})] &= \frac{\sigma^2}{\delta}(1 - \exp(-\delta)) \\ \text{Corr}[\log(P_t/P_{t-1}), \log(P_{t-1}/P_{t-2})] &= -\frac{1}{2}(1 - \exp(-\delta)) . \end{aligned}$$

Therefore, using the parameters in (1.5.5) and setting the first-order autocorrelation coefficient equal to  $-0.05$  uniquely calibrates the parameter vector  $(\mu, \sigma, \delta)$ . The distribution of

$\log P_T$  implied by these parameters yields the following 45 possible strikes (in dollars) from which we select our  $n$  options in the optimal buy-and-hold portfolio:

265	346	426	506	587	667	748	828	908
989	1,069	1,150	1,230	1,310	1,391	1,471	1,552	1,632
1,712	1,793	1,873	1,954	2,034	2,114	2,195	2,275	2,356
2,436	2,517	2,597	2,677	2,758	2,838	2,919	2,999	3,079
3,160	3,240	3,321	3,401	3,481	3,562	3,642	3,723	3,803

Note that the distribution of possible strikes lies in a much narrower range in this case than in the geometric Brownian motion case of Section 1.5.1: 265 to 3,723 for the trending Ornstein-Uhlenbeck process versus 69 to 14,363 for geometric Brownian motion. This is an implication of the mean-reverting nature of the trending Ornstein-Uhlenbeck process, a stochastic process in which log-prices are stationary about a deterministic trend, in contrast to geometric Brownian motion in which log-prices are difference-stationary. In the former case, the variance of the log-price process is bounded as the horizon increases without bound, whereas in the latter case, the variance is proportional to the horizon, implying a wider range of strikes.

Recall from Section 1.4.2 that because the optimal dynamic asset-allocation strategy is path-dependent under (1.4.8), the certainty equivalent of  $V_T^*$  will not approach the certainty equivalent of  $W_T^*$  as the number of options  $n$  in the buy-and-hold portfolio increases without bound. Indeed, there is an upper bound for  $\text{CE}(V_T^*)$ , which is the certainty equivalent of the optimal buy-and-hold portfolio with an infinite number of options,  $\text{CE}(V_T^\infty)$ , and for path-dependent dynamic portfolio strategies,  $\text{CE}(V_T^\infty)$  is strictly less than  $\text{CE}(W_T^*)$ . In the case of the trending Ornstein-Uhlenbeck process (1.4.8) and CRRA preferences, we have an explicit expression for  $V_T^\infty$  (see Section 1.4.2), hence we can construct a mean-square optimal buy-and-hold portfolio where the benchmark is  $V_T^\infty$ , not  $W_T^*$ .

### Utility-Optimal Buy-and-Hold Portfolios

Table 2a summarizes the utility-optimal buy-and-hold portfolios for the same combination of risk-aversion parameters and number of options  $n$  as in the geometric Brownian motion case of Table 1a. The results for the panels with  $\text{RRA} = 1, 2, 5$  were obtained by maximizing expected utility directly using a discretized distribution for  $P_T$  (see Section 1.3.1), and the results for the remaining three panels of Table 2a were obtained by minimizing the utility-weighted mean-squared error (see Section 1.3.3).

Note that for each level of risk aversion, the certainty equivalent  $\text{CE}(W_T^*)$  of the optimal dynamic strategy is considerably larger than that of the geometric Brownian motion case. The presence of predictability can be exploited by the investor and in doing so, his expected utility is increased dramatically, e.g., from a certainty equivalent of \$9,948,433 in the geometric Brownian motion case to \$13,162,500 in the Ornstein-Uhlenbeck case for log-utility. A more direct measure of the economic value of predictability can be obtained by considering

the difference between the certainty equivalents of the optimal dynamic strategy and those of the optimal buy-and-hold portfolio with an infinite number of options. For a log-utility investor, this difference is \$745,150 or 5.6% of  $CE(W_T^*)$ , a significant amount. As the level of risk aversion increases, this difference declines in absolute terms—less wealth is allocated to the risky asset, hence predictability has less of an impact—but is relatively stable as a percentage of  $CE(W_T^*)$ , fluctuating between 4% and 6%.

The most interesting feature of Table 2a is that the certainty equivalents of the buy-and-hold portfolios do not approach  $CE(V_T^\infty)$  as quickly as the certainty equivalents of Table 1a. This is most easily seen in the third panel ( $RRA = 5$ ) in which the certainty equivalent of the optimal buy-and-hold portfolio with 3 options is only 83.2% of  $CE(V_T^\infty)$ . However, as we remarked earlier, the data for this panel was computed by maximizing expected utility through a discretization of the distribution of  $P_T$  using a grid of 4,000 points. Because of the relatively high value of  $RRA$ , any interval in the support of  $P_T$  where  $W_T(P_T)$  is close to 0 will result in a large negative contribution to the certainty equivalent. We can address this issue by using a finer grid, but only at the expense of computational complexity.<sup>21</sup>

Another interesting feature of Table 2a is that there is no investment in the bond in any of the buy-and-hold portfolios in the first four panels ( $RRA = 1, 2, 5, 10$ ). While this is not unexpected for low levels of risk aversion—such investors seek higher expected returns by nature of their risk preferences—it is quite surprising for investors with  $RRA = 10$ . The intuition for this result comes from the fact that stock returns are predictable in this case, hence there is greater value to be gained from investing in stocks for each level of risk aversion. Alternatively, the predictability in stock returns make stocks less risky, *ceteris paribus*, hence even a risk-averse investor will hold a larger fraction of his wealth in stocks in this case.

As in Tables 1a and 1b, the optimal buy-and-hold portfolios for less risk-averse investors ( $RRA = 1, 2, 5$ ) are net positive in options, ranging from 98.8% when  $RRA = 1$  to 18.9% when  $RRA = 5$ , for  $n = 3$ . However, unlike the geometric Brownian motion case, the optimal buy-and-hold portfolios do contain short positions in some options, even for these lower levels of risk aversion. For example, when  $RRA = 2$  and  $n = 3$ , the optimal buy-and-hold portfolio consists of long positions in the \$265-strike and \$506-strike options, but a short position of 15,246 options in the \$1,391-strike option. For higher levels of risk aversion, the situation is reversed: the optimal buy-and-hold portfolios are net negative in options, but they do contain long positions in certain options. For example, when  $RRA = 20$  and  $n = 3$ , the optimal buy-and-hold portfolio consists of short positions in the \$265-strike and \$506-strike options, but a long position of 1,753 options in the \$346-strike option.

These long and short positions underscore the complexity of an investor's ideal risk exposures, and may provide a useful benchmark for comparing different dynamic investment policies at a single point in time. In particular, it may be possible to re-interpret these option positions as classic spread trades, e.g., bull/bear and butterfly spreads, or combinations, e.g., strips, straps, straddles, and strangles.<sup>22</sup> By doing so, we may be able to gain insight into the implicit bets that a particular dynamic asset-allocation strategy contains, and develop a

<sup>21</sup>Since these numerical results are mainly for illustrative purposes, we have not endeavored to optimize them within each panel. Instead, to ensure comparability across risk-aversion parameters and other specifications, we have attempted to hold fixed as many aspects of the optimization process as possible.

<sup>22</sup>For this purpose, it may be useful to convert some of the call-option positions into their put-option equivalents using the put-call parity relation (see, for example, Cox and Rubinstein, 1985).

standard lexicon for comparing those bets across investment policies.

### Mean-Square-Optimal Buy-and-Hold Portfolios

Table 2b summarizes the mean-square-optimal buy-and-hold portfolios for the same combination of strikes, risk-aversion parameters, and number of options  $n$  as in Table 2a. Table 2b shows that the RMSE of the optimal buy-and-hold portfolio declines rapidly. With only one or two options, the optimal buy-and-hold portfolio is typically within 5% of the upper bound  $CE(V_T^\infty)$ . For example, in the case where relative risk-aversion is 2, the RMSE of the optimal buy-and-hold portfolio with no options is 87.5%; with 1 option, the RMSE declines to 28.9%; and with 2 options, the RMSE is 5.0%. With 5 options, the RMSE is less than 2.0% for all but the lowest level of risk aversion ( $RRA = 1$ , for which the RMSE is 2.7%). But as in Table 1b, the certainty equivalents of the optimal buy-and-hold portfolio do not increase monotonically as the number of options increases, since we are optimizing mean-squared-error, not expected utility. For example, in the second panel ( $RRA = 2$ ) the certainty equivalent drops precipitously from 64.4% to 37.4% of the upper bound  $CE(V_T^\infty)$  as the number of options increases from 4 to 5. However, for higher levels of risk aversion, the certainty equivalents do tend to increase with the number of options in the portfolio (and are guaranteed to converge to the upper bound  $CE(V_T^\infty)$  as  $n$  increases without bound).

As risk aversion increases, the optimal buy-and-hold portfolios behave in a similar manner to those in Table 1b: options are used to hedge long positions in the stock. For risk-aversion levels of 10 or greater, all options positions are negative.

### 1.5.3 A Bivariate Linear Diffusion Process

We calibrate the parameters  $(\kappa, \theta, \sigma_1, \sigma_2, \rho)$  of the bivariate linear diffusion (1.4.22)–(1.4.23) using the following values:

$$\begin{aligned}
 E[\log(P_t/P_{t-1})] &= 0.15 \\
 \text{Var}[\log(P_t/P_{t-1})] &= 0.04 \\
 \text{Var}[\mu_t] &= .025^2 \\
 \text{Corr}[\mu_t, \mu_{t-1}] &= 0.05 \\
 \rho &= 0 .
 \end{aligned}
 \tag{1.5.6}$$

The first two moments are calibrated with the same values as those in the geometric Brownian motion and trending Ornstein-Uhlenbeck cases. The value for the variance of  $\mu_t$  implies a standard deviation of 250 basis points for the conditional mean  $\mu_t$ , and we assume that  $\mu_t$  is only slightly autocorrelated over time, and not correlated at all with the Brownian motion driving prices. This calibration implies the following 45 possible strikes (in dollars) from

which we select our  $n$  options in the optimal buy-and-hold portfolio:

68	403	737	1,071	1,405	1,739	2,073	2,407	2,742
3,076	3,410	3,744	4,078	4,412	4,746	5,081	5,415	5,749
6,083	6,417	6,751	7,085	7,420	7,754	8,088	8,422	8,756
9,090	9,425	9,759	10,093	10,427	10,761	11,095	11,429	11,764
12,098	12,432	12,766	13,100	13,434	13,768	14,103	14,437	14,771

Note the similarity between the range of these strikes and that of geometric Brownian motion in Section 1.5.1. This suggests that the economic properties of the bivariate linear diffusion process are close to those of geometric Brownian motion, which will be borne out by the optimal buy-and-hold portfolios described below.

As in the case of the trending Ornstein-Uhlenbeck process, under (1.4.22)–(1.4.23) the optimal dynamic asset-allocation strategy is path-dependent. Therefore, we shall again use the upper bound  $V_T^\infty$  as the benchmark in our mean-square-optimal buy-and-hold portfolio, and compare its certainty equivalent  $\text{CE}(V_T^*)$  to  $\text{CE}(V_T^\infty)$ .

### Utility-Optimal Buy-and-Hold Portfolios

Table 3a reports the optimal buy-and-hold portfolios under (1.4.22)–(1.4.23) for CRRA preferences with the same risk aversion levels as in Tables 1 and 2. The results of the first two panels of Table 3a were computed by maximizing expected utility according to Section 1.3.1 and the results of the remaining panels were computed by minimizing utility-weighted mean-squared-error according to Section 1.3.3.

Table 3a contains certain features in common with Tables 1a and 2a, but also exhibits some important differences. As in Table 2a, the certainty equivalents of  $V_T^\infty$  are lower than their counterparts for  $W_T^*$ , but in Table 3a the gap declines monotonically as risk aversion increases. For log-utility,  $\text{CE}(V_T^\infty)$  is 15.5% less than  $\text{CE}(W_T^*)$ , but this difference is only 7.5% when  $\text{RRA} = 2$ , 3.1% when  $\text{RRA} = 5$ , and 0.8% when  $\text{RRA} = 20$ . In contrast, the gap between  $\text{CE}(W_T^*)$  and  $\text{CE}(V_T^\infty)$  in Table 2a is still 4.2% when  $\text{RRA} = 20$ . This underscores the fact that the predictability of the bivariate linear diffusion is of a different form than that of the trending Ornstein-Uhlenbeck process.

Indeed, there are striking similarities between Tables 3a and 1a, another indication that the terminal stock price  $P_T$  and option prices corresponding to the two stochastic processes—as we have calibrated them—have much in common. However, note that the certainty equivalents in Table 1a are relative to  $\text{CE}(W_T^*)$ , not  $\text{CE}(V_T^\infty)$ . Nevertheless, even the values of  $\text{CE}(V_T^\infty)$  in Table 3a are extremely close to the values of  $\text{CE}(W_T^*)$  in Table 1a. This close correspondence suggests that for all practical purposes, the bivariate process (1.4.22)–(1.4.23) offers the same buy-and-hold investment opportunities to the investor as geometric Brownian motion.

## Mean-Square-Optimal Buy-and-Hold Portfolios

Table 3b reports the mean-square-optimal buy-and-hold portfolios under (1.4.22)–(1.4.23) for CRRA preferences with the same risk aversion levels as in Table 3a. These results match those in Table 1b quite closely. Specifically, as in Table 1b, the optimal buy-and-hold portfolio is a particularly poor approximation to both  $W_T^*$  and  $V_T^\infty$  in the log-utility case, with RMSE's greater than 3,500%, certainty equivalents  $CE(V_T^*)$  no greater than 35% of  $CE(V_T^\infty)$ , and large swings in portfolio weights as  $n$  is changed from 1 to 2 and from 2 to 3. For higher levels of risk aversion, the optimal buy-and-hold portfolios in Table 3b are remarkably close to those in Table 1b in terms of portfolio weights, option positions, and certainty equivalents, providing further confirmation that the bivariate linear diffusion, calibrated according to (1.5.6), share many of the same economic properties as geometric Brownian motion.

## 1.6 Discussion

For expositional purposes, we have made a number of simplifying assumptions, many of which can be relaxed at the expense of notational and computational complexity. In Section 1.6.1, we consider some practical issues regarding the implementation of the optimal buy-and-hold portfolio. We discuss the advantages of using more complex derivative securities in Section 1.6.2, and in Section 1.6.3 we consider extending our analysis to other preferences and price processes. Finally, in Section 1.6.4 we argue that the gap between  $CE(W_T^*)$  and  $CE(V_T^\infty)$  is a useful measure of the economic value of predictability, and discuss the role of taxes and transactions costs in interpreting the gap.

### 1.6.1 Practical Considerations

An obvious prerequisite to any practical implementation of the optimal buy-and-hold portfolio proposed in Section 1.3 is the existence of options with the appropriate maturity  $T$  and strike prices  $\{k_i^*\}$ . These two issues—time-to-maturity and the set of available strikes—are related, since a longer time-to-maturity generally implies a greater dispersion for the optimal strikes (to accommodate the greater dispersion in the terminal stock-price distribution). For horizons less than one year, there are relatively liquid options on the S&P 500 and other indexes, usually with a reasonable number of strikes above and below the spot price, hence the possibility of replacing certain dynamic investment strategies with an optimal buy-and-hold portfolio is plausible. However, for longer maturities such as the 20-year horizons proposed in the numerical examples of Section 1.5, exchange-traded options do not exist.

This might seem to be a serious impediment to implementing the optimal buy-and-hold strategy for realistic investment horizons. However, we think there is hope for several reasons. First, longer-maturity index options are always available through custom OTC derivatives contracts, although this is admittedly a very expensive alternative. Second, the scarcity of longer-maturity contracts is a reflection of existing demand—if optimal buy-and-hold portfolios become popular, this will create new demand for such contracts, leading to increased supply. Recent legislative debate regarding the privatization of the US social

security system suggests the possibility of a huge increase in demand for such products and services. Third, insurance companies now provide various policies that have similar features to long-dated options, e.g., annuities with call and put features, contingent life-insurance policies, etc., hence they may be a natural supplier of optimal buy-and-hold portfolios. And finally, an imperfect alternative to long-dated options is a carefully managed sequence of shorter-term options, and it may be possible to derive a dynamic trading strategy consisting of a sequence of overlapping options contracts that will yield the same investment profile as the optimal buy-and-hold strategy.<sup>23</sup> A dynamic trading strategy seems contrary to our motivation for constructing a buy-and-hold alternative to the standard optimal dynamic asset-allocation policy. However, the inclusion of a few well-chosen short-maturity options from time to time in an otherwise passive buy-and-hold portfolio might be a very cost-effective and efficient alternative to the optimal dynamic policy, and we are investigating this possibility in our current research program.

Another issue that arises in the practical implementation of the optimal buy-and-hold strategy is computational challenges associated with the optimization procedure. As discussed in Section 1.5, there are limits to the number of sub-problems that can be handled in a reasonable amount of time, which imposes limits on the number of possible strikes that can be considered, as well as the number of options  $n$  in the buy-and-hold portfolio. In our numerical examples, we have made no attempt to optimize our algorithm for numerical and computational efficiency, preferring instead to maintain consistency across examples to facilitate comparisons. For example, when solving for the optimal buy-and-hold portfolio with  $n = 1$  option, there was no need to limit ourselves to just 45 possible strikes. In fact, this problem can be solved very efficiently even if we were to consider several thousand possible strikes. In addition, by selecting the range of strikes as a function of the relative risk-aversion parameter, it is possible to obtain considerably better results than those of Tables 1–3.<sup>24</sup> Therefore, the numerical results of Section 1.5 should be taken as illustrative only, and not necessarily indicative of the best possible performance of the optimal buy-and-hold portfolios.

## 1.6.2 Other Derivative Securities

For simplicity, we have used only European call options in our buy-and-hold strategies. A natural extension is to include more complex derivatives, perhaps with path dependencies such as knock-out or average-rate options. This extension may be especially relevant in the presence of predictability, since in such cases we cannot attain  $CE(W_T^*)$  with a buy-and-hold strategy even if we include an infinite number of European options. In fact, the specific form of predictability may suggest a class of derivatives that are particularly suitable. For

---

<sup>23</sup>See Bertsimas, Kogan, and Lo (2000b) for an example of how such a strategy might be derived.

<sup>24</sup>Specifically, having selected the  $N$  possible strikes, we solve the  $\binom{N}{n}$  sub-problems as described in Section 1.3.1. Once this is completed, we use the strikes  $\{\hat{k}_i\}$  from the sub-problem with the largest optimum to select another set of  $N$  possible strikes. This new set of  $N$  strikes spans a smaller interval than the original set, but still contains  $\{\hat{k}_i\}$ . We then solve another  $\binom{N}{n}$  sub-problems and select the largest optimum as our solution, and denote the corresponding strikes as  $\{k_i^*\}$ . This two-stage procedure for determining the set of possible strikes often yields significant improvements in the objective function.

example, in the case of the trending Ornstein-Uhlenbeck process (1.4.8), it seems reasonable to conjecture that derivatives whose payoffs depend on

$$\int_0^T h(|X_t - X_0 - \mu t|) dt \quad (1.6.1)$$

for some function  $h(\cdot)$  would be most useful for approximating  $W_T^*$  in a buy-and-hold portfolio. This should be true more generally for other mean-reverting stock-price processes. On the other hand, if the stock-price process displays some type of persistence or “momentum”, a different class of derivatives might be more appropriate.

### 1.6.3 Other Preferences and Price Processes

Although we have confined much of our analysis in Sections 1.4 and 1.5 to the special cases of CRRA and CARA preferences under three specific price processes, we wish to emphasize that the optimal buy-and-hold portfolio can be derived for many other preferences and price processes. For example, the class of hyperbolic absolute risk-aversion (HARA) preferences can be accommodated, as well as any price process for which the conditional state-price density can be computed. Even more general preferences and price processes are allowable at the expense of computational complexity. For example, for price processes that do not admit closed-form expressions for the conditional state-price densities, these can be estimated nonparametrically as in Ait-Sahalia and Lo (1998).

### 1.6.4 The Predictability Gap

As we have seen in Sections 1.4.2 and 1.4.3, in the presence of predictability in the stock-price process, buy-and-hold portfolios of stocks, bonds, and European call options cannot approximate  $W_T^*$  arbitrarily well, even as the number of options increases without bound. We use the term “predictability gap” to denote the difference between  $CE(W_T^*)$  and  $CE(V_T^\infty)$ , which depends on the investor’s preferences as well as the parameters of the stock-price process.

The natural question to ask is how significant is this predictability gap? Given that the end-of-period wealth  $W_T^*$  of the optimal dynamic asset-allocation policy is generally unattainable due to transactions costs and other market frictions,  $CE(W_T^*)$  can be viewed as a theoretical upper bound on how well an investor can do. On the other hand, if  $V_T^\infty$  is well approximated by an optimal buy-and-hold portfolio  $V_T^*$  with just a few options, it is more likely to be attainable in practice given that only a few trades are required to establish the portfolio and there are few costs to bear thereafter. Therefore, if the predictability gap is small, the buy-and-hold portfolio may well be optimal even in the presence of predictable stock returns. To investigate this possibility, we must consider the impact of transactions costs on  $CE(W_T^*)$ .

Most of the studies in the transactions costs literature ignore predictability, assuming



independently and identically distributed (IID) returns instead.<sup>25</sup> Such studies may underestimate the impact of transaction costs because the presence of predictability provides another motive for trade, i.e., time-varying investment opportunity sets. Therefore, we might expect transactions costs—as a percentage of initial wealth  $W_0$ —to be higher if stock returns are predictable.

Balduzzi and Lynch (1999) do consider transactions costs in the case of predictability, computing the impact on investor's expected utility when transactions costs exist but are ignored by the investor. They do not report the difference between the certainty equivalent of the optimal asset-allocation policy in an economy without transactions costs and the certainty equivalent of the optimal policy in an economy with transactions costs (though their framework should allow them to do so easily). They do mention, however, that "... the presence of transaction costs ... decreases the utility cost of ignoring predictability". This suggests that  $CE(W_T^*)$  might be reduced significantly, reducing the predictability gap and providing more compelling motivation for the optimal buy-and-hold portfolio.

An even more compelling motivation for the optimal buy-and-hold portfolio is the presence of taxes. For taxable investors,  $CE(W_T^*)$  is reduced by the present value of the sequence of capital gains taxes that are generated by an optimal dynamic asset-allocation strategy. In contrast, all of the capital gains taxes are deferred until date  $T$  in a buy-and-hold portfolio. Therefore, the economic value of predictability is likely to be even lower for taxable investors, and the optimal buy-and-hold portfolio that much more attractive.

## 1.7 Conclusion

In this paper, we compare optimal buy-and-hold portfolios of stocks, bonds, and options to the standard optimal dynamic asset-allocation policies involving only stocks and bonds. Under certain conditions, buy-and-hold portfolios are excellent approximations—in terms of certainty equivalence and mean-squared-error of end-of-period wealth—to their dynamic counterparts, suggesting that in those cases, dynamic trading strategies may be "automated" by simple buy-and-hold portfolios with just a few options. Cases where the approximation breaks down are also of interest, since such situations highlight the importance of dynamic trading opportunities.

There are a number of extensions of this research that may be worth pursuing. The most obvious is to perform similar analyses for other stochastic processes and preferences, those that are more consistent with the empirical evidence. The main challenge in this case is, of course, tractability and computational complexity.

A more important extension is to consider approximating other dynamic investment strategies with buy-and-hold portfolios of derivatives. Although we focus on optimal dynamic asset-allocation strategies in this paper, there is no reason to confine our attention to such a narrow class of strategies. For example, deriving optimal buy-and-hold strategies to approximate dollar-cost averaging strategies or other popular dynamic investment strategies—strategies that need not be based on expected utility maximization—might be

---

<sup>25</sup>For example, Magill and Constantinides (1976), Constantinides (1986), Davis and Norman (1990), Dumas and Luciano (1991), and Gennotte and Jung (1992) all assume IID return-generating processes.

of considerably broader interest.

Finally, the composition of the optimal buy-and-hold portfolio provides an interesting summary of the risk exposures of the optimal dynamic asset-allocation policy that the buy-and-hold portfolio approximates. By examining the payoff structure of the optimal buy-and-hold portfolio, and its sensitivities to various market factors and economic shocks, we can develop insights into the risks of dynamic investment policies using measures computed at a single point in time. We hope to explore these and other extensions in the near future.

Table 1a

Utility-optimal buy-and-hold portfolios of stocks, bonds, and  $n$  European call options for CRRA utility under geometric Brownian motion stock-price dynamics with drift  $\mu = 15\%$  and volatility  $\sigma = 20\%$ . Other calibrated parameters include: riskless rate  $r = 5\%$ , initial stock price  $P_0 = \$50$ , initial wealth  $W_0 = \$100,000$ , and time period  $T = 20$  years. 'RRA' denotes the coefficient of relative risk aversion, 'CE( $W_T^*$ )' denotes the certainty equivalent of the optimal dynamic stock/bond policy, and 'CE( $V_T^*$ )' denotes the certainty equivalent of the optimal buy-and-hold portfolio, reported as a percentage of CE( $W_T^*$ ).

$n$	Options (%)	Stock (%)	CE( $V_T^*$ ) (%)	RMSE (%)	Option Positions in Optimal Portfolio with $n$ Options		
					Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)
CE( $W_T^*$ ) = \$9,948,433 , RRA = 1 (Log Utility)							
0	0.0	100.0	20.2	3,659.6			
1	60.4	39.6	68.7	3,653.4	54,653		
					733		
2	80.0	20.0	87.7	3,642.4	13,371	143,901	
					401	1,398	
3	99.3	0.7	92.2	3,642.4	786	12,518	144,890
					69	401	1,398
CE( $W_T^*$ ) = \$1,644,465 , RRA = 2							
0	0.0	100.0	81.9	206.2			
1	63.3	36.7	94.5	188.6	2,214		
					69		
2	59.1	40.9	99.2	146.2	1,647	3,040	
					69	401	
3	59.3	40.7	99.4	97.4	1,661	2,795	4,120
					69	401	1,398
CE( $W_T^*$ ) = \$558,453 , RRA = 5							
0	0.0	62.0	97.3	143.5			
1	-46.3	131.9	99.1	103.8	-1,620		
					69		
2	-36.9	120.4	99.8	35.9	-1,207	-602	
					69	401	
3	-37.0	120.5	99.8	14.8	-1,215	-573	-197
					69	401	1,398

Table 1a (continued)

$n$	Options (%)	Stock (%)	CE( $V_T^*$ ) (%)	RMSE (%)	Option Positions in Optimal Portfolio with $n$ Options		
					Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)
$CE(W_T^*) = \$389,619$ , RRA = 10							
0	0.0	26.5	96.6	154.9			
1	-47.1	102.0	98.9	104.5	-1,647 69		
2	-37.5	90.0	99.7	25.4	-1,258 69	-387 401	
3	-37.6	90.1	99.7	5.9	-1,262 69	-377 401	-91 1,731
$CE(W_T^*) = \$345,561$ , RRA = 15							
0	0.0	16.2	97.2	124.0			
1	-37.1	76.6	99.1	82.3	-1,297 69		
2	-29.6	67.1	99.8	17.6	-1,000 69	-262 401	
3	-29.7	67.2	99.8	4.2	-1,003 69	-256 401	-50 1,731
$CE(W_T^*) = \$325,437$ , RRA = 20							
0	0.0	11.6	97.7	101.0			
1	-30.0	60.7	99.3	66.5	-1,048 69		
2	-24.0	53.0	99.8	13.3	-812 69	-196 401	
3	-24.0	53.1	99.8	3.2	-814 69	-192 401	-34 1,731

Table 1b

Mean-square-optimal buy-and-hold portfolios of stocks, bonds, and  $n$  European call options for CRRA utility under geometric Brownian motion stock-price dynamics with parameters  $(\mu, \sigma)$  calibrated to match the following moments:  $E[\log(P_t/P_{t-1})] = 0.15$ ,  $\text{Var}[\log(P_t/P_{t-1})] = 0.04$ . Other calibrated parameters include: riskless rate  $r = 5\%$ , initial stock price  $P_0 = \$50$ , initial wealth  $W_0 = \$100,000$ , and time period  $T = 20$  years. 'RRA' denotes the coefficient of relative risk aversion, 'CE( $W_T^*$ )' denotes the certainty equivalent of the optimal dynamic stock/bond policy, and 'CE( $V_T^*$ )' denotes the certainty equivalent of the optimal buy-and-hold portfolio, reported as a percentage of CE( $W_T^*$ ).

$n$	Options (%)	Stock (%)	CE( $V_T^*$ ) (%)	RMSE (%)	Option Positions in Optimal Portfolio with $n$ Options				
					Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)
CE( $W_T^*$ ) = \$9,948,433 , RRA = 1 (Log Utility)									
0	0.0	100.0	20.2	3,659.6					
1	0.2	99.8	20.4	2,889.6	29.0 × 10 <sup>6</sup> 14,696				
2	62.9	-4.0	10.2	2,886.7	331,561 1,398	28.3 × 10 <sup>6</sup> 14,696			
3	2.8	97.2	23.1	2,870.6	2,431,277 5,388	-68.2 × 10 <sup>6</sup> 14,363	94.7 × 10 <sup>6</sup> 14,696		
4	8.0	92.0	28.1	2,869.5	943,747 3,393	2,987,657 8,712	-85.9 × 10 <sup>6</sup> 14,363	111.0 × 10 <sup>6</sup> 14,696	
5	15.5	84.5	34.9	2,869.3	465,463 2,396	1,415,411 6,052	2,917,259 10,374	-92.0 × 10 <sup>6</sup> 14,363	116.2 × 10 <sup>6</sup> 14,696
CE( $W_T^*$ ) = \$1,644,465 , RRA = 2									
0	0.0	100.0	81.9	206.2					
1	0.8	99.2	83.4	57.2	14,846 2,063				
2	3.9	96.1	86.6	26.6	9,233 1,066	15,764 9,044			
3	7.6	92.4	89.3	15.8	6,873 733	8,120 4,390	15,063 14,696		
4	19.6	68.4	89.4	13.8	4,908 401	4,815 2,063	6,273 5,388	14,315 14,696	
5	17.5	82.5	94.0	13.3	4,342 401	3,797 1,731	3,753 3,725	4,430 6,717	13,501 14,696
CE( $W_T^*$ ) = \$558,453 , RRA = 5									
0	0.0	24.4	84.7	28.3					
1	-0.1	42.5	94.2	7.8	-543 1,398				
2	-0.6	50.8	96.6	3.6	-554 733	-225 4,058			
3	-2.5	61.8	98.4	2.2	-624 401	-246 1,731	-159 6,385		
4	-2.3	60.5	98.2	1.4	-569 401	-218 1,398	-134 3,725	-110 10,374	
5	-2.1	59.0	98.1	1.0	-500 401	-191 1,066	-134 2,396	-102 5,388	-91 13,698

Table 1b (continued)

$n$	Options (%)	Stock (%)	$CE(V_T^*)$ (%)	RMSE (%)	Option Positions in Optimal Portfolio with $n$ Options				
					Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)
$CE(W_T^*) = \$389,619$ , $RRA = 10$									
0	0.0	6.7	86.4	27.0					
1	-0.1	17.7	94.8	6.4	-297 1,066				
2	-1.8	30.0	98.0	3.0	-450 401	-107 2,396			
3	-1.5	27.9	97.7	1.7	-375 401	-101 1,398	-52 4,723		
4	-1.4	27.3	97.6	1.3	-343 401	-93 1,066	-54 2,728	-32 7,715	
5	-68.0	139.6	96.5	0.9	-2,345 69	-236 401	-100 1,066	-54 2,728	-32 7,715
$CE(W_T^*) = \$345,561$ , $RRA = 15$									
0	0.0	3.7	89.6	21.5					
1	-0.3	14.0	97.0	5.0	-248 733				
2	-1.2	19.2	98.3	2.2	-306 401	-60 2,396			
3	-1.0	18.0	98.1	1.3	-260 401	-60 1,398	-27 4,723		
4	-47.6	96.4	98.5	1.0	-1,637 69	-189 401	-62 1,398	-27 4,723	
5	-51.2	102.2	98.0	0.7	-1,767 69	-163 401	-56 1,066	-29 2,396	-19 6,385
$CE(W_T^*) = \$325,437$ , $RRA = 20$									
0	0.0	2.5	91.7	17.5					
1	-0.2	10.0	97.5	3.9	-180 733				
2	-0.9	14.1	98.5	1.7	-230 401	-41 2,396			
3	-0.8	13.2	98.4	1.1	-198 401	-40 1,398	-18 4,390		
4	-37.8	75.7	98.9	0.8	-1,304 69	-142 401	-42 1,398	-18 4,390	
5	-40.7	80.2	98.5	0.5	-1,405 69	-122 401	-40 1,066	-20 2,396	-12 6,385

Table 2a

Utility-optimal buy-and-hold portfolios of stocks, bonds, and  $n$  European call options for CRRA utility under a trending Ornstein-Uhlenbeck stock-price process where the parameters  $(\sigma, \mu, \delta)$  have been calibrated to match the following moments:  $E[\log(P_t/P_{t-1})] = 0.15$ ,  $\text{Var}[\log(P_t/P_{t-1})] = 0.04$ ,  $\text{Corr}[\log(P_t/P_{t-1}), \log(P_{t-1}/P_{t-2})] = -0.05$ . Other calibrated parameters include: riskless rate  $r = 5\%$ , initial stock price  $P_0 = \$1$ , initial wealth  $W_0 = \$100,000$ , and time period  $T = 20$  years. 'RRA' denotes the coefficient of relative risk aversion, 'CE( $W_T^*$ )' denotes the certainty equivalent of the optimal dynamic stock/bond policy, 'CE( $V_T^\infty$ )' denotes the certainty equivalent of the optimal buy-and-hold portfolio with a continuum of options, and 'CE( $V_T^*$ )' denotes the certainty equivalent of the optimal buy-and-hold portfolio with a finite number  $n$  of options, reported as a percentage of CE( $V_T^\infty$ ).

$n$	Options (%)	Stock (%)	CE( $V_T^*$ ) (%)	RMSE (%)	Option Positions in Optimal Portfolio with $n$ Options		
					Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)
CE( $W_T^*$ ) = \$13,162,500 , CE( $V_T^\infty$ ) = \$12,417,350 , RRA = 1 (Log Utility)							
0	0.0	100.0	16.2	115.7			
1	96.6	3.4	89.1	42.3	24,955 426		
2	98.8	1.2	96.2	40.3	6,251 346	30,824 587	
3	98.8	1.2	97.8	12.6	6,047 346	33,494 587	-39,564 1,793
CE( $W_T^*$ ) = \$6,166,222 , CE( $V_T^\infty$ ) = \$5,814,196 , RRA = 2							
0	0.0	100.0	31.3	87.5			
1	83.9	16.1	89.0	29.1	10,204 265		
2	83.0	17.0	90.5	34.2	7,822 265	4,810 426	
3	83.0	17.0	91.6	6.7	8,283 265	6,623 506	-15,246 1,391
CE( $W_T^*$ ) = \$2,011,701 , CE( $V_T^\infty$ ) = \$1,874,790 , RRA = 5							
0	0.0	100.0	72.2	30.8			
1	16.7	83.3	79.8	56.6	2,036 265		
2	19.5	80.5	82.7	26.9	5,777 265	-5,062 346	
3	18.9	81.1	83.2	13.9	5,304 265	-4,202 346	-2,385 989

Table 2a (continued)

$n$	Options (%)	Stock (%)	$CE(V_T^*)$ (%)	RMSE (%)	Option Positions in Optimal Portfolio with $n$ Options		
					Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)
$CE(W_T^*) = \$957,797$ , $CE(V_T^\infty) = \$900,296$ , RRA = 10							
0	0.0	100.0	91.9	117.3			
1	-6.6	106.6	96.8	11.2	-1,712 426		
2	-7.0	107.0	97.1	3.8	-1,047 346	-951 748	
3	-6.8	106.8	97.1	2.5	-958 346	-641 587	-525 1,150
$CE(W_T^*) = \$681,834$ , $CE(V_T^\infty) = \$647,654$ , RRA = 15							
0	0.0	75.8	87.9	151.9			
1	-15.9	108.7	95.3	10.3	-1,937 265		
2	-15.5	108.1	95.2	2.0	-1,869 265	-285 1,150	
3	-16.9	110.2	95.7	3.9	-2,844 265	1,435 346	-689 587
$CE(W_T^*) = \$560,880$ , $CE(V_T^\infty) = \$537,074$ , RRA = 20							
0	0.0	54.1	87.9	138.6			
1	-13.9	89.0	93.2	5.1	-1,686 265		
2	-15.5	91.7	93.7	11.6	-2,385 265	740 346	
3	-15.7	92.0	93.8	4.0	-2,868 265	1,753 346	-643 506



Table 2b

Mean-square-optimal buy-and-hold portfolios of stocks, bonds, and  $n$  European call options for CRRA utility under a trending Ornstein-Uhlenbeck stock-price process with parameters  $(\sigma, \mu, \delta)$  calibrated to match the following moments:  $E[\log(P_t/P_{t-1})] = 0.15$ ,  $\text{Var}[\log(P_t/P_{t-1})] = 0.04$ ,  $\text{Corr}[\log(P_t/P_{t-1}), \log(P_{t-1}/P_{t-2})] = -0.05$ . Other calibrated parameters include: riskless rate  $r = 5\%$ , initial stock price  $P_0 = \$50$ , initial wealth  $W_0 = \$100,000$ , and time period  $T = 20$  years. 'RRA' denotes the coefficient of relative risk aversion, 'CE( $W_T^*$ )' denotes the certainty equivalent of the optimal dynamic stock/bond policy, 'CE( $V_T^\infty$ )' denotes the certainty equivalent of the optimal buy-and-hold portfolio with a continuum of options, and 'CE( $V_T^n$ )' denotes the certainty equivalent of the optimal buy-and-hold portfolio with a finite number  $n$  of options, reported as a percentage of CE( $V_T^\infty$ ).

$n$	Options (%)	Stock (%)	CE( $V_T^*$ ) (%)	RMSE (%)	Option Positions in Optimal Portfolio with $n$ Options				
					Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)
CE( $W_T^*$ ) = \$13,162,500 , CE( $V_T^\infty$ ) = \$12,417,350 , RRA = 1									
0	0.0	100.0	16.2	115.7					
1	91.7	8.3	87.4	37.5	32,697 506				
2	86.8	13.2	81.3	9.3	44,428 587	-44,691 1,712			
3	87.0	13.0	81.5	6.3	44,694 587	-55,819 1,793	10,865 3,803		
4	88.7	-3.4	56.7	4.5	45,736 587	-33,938 1,632	-29,561 2,195	17,831 3,803	
5	90.2	9.8	56.7	2.7	24,169 506	24,603 748	-32,245 1,552	-33,388 2,114	16,665 3,803
CE( $W_T^*$ ) = \$6,166,222 , CE( $V_T^\infty$ ) = \$5,814,196 , RRA = 2									
0	0.0	100.0	31.3	87.5					
1	87.2	12.8	88.5	28.9	10,599 265				
2	75.4	24.6	79.8	5.0	14,198 346	-14,689 1,471			
3	82.7	-9.0	24.3	3.6	15,686 346	-6,392 1,150	-9,114 1,632		
4	79.8	7.4	64.4	2.3	15,136 346	-8,530 1,230	-9,152 1,873	2,399 3,803	
5	82.1	-5.9	37.4	1.6	15,582 346	-5,987 1,150	-6,895 1,552	-6,095 2,195	3,515 3,803
CE( $W_T^*$ ) = \$2,011,701 , CE( $V_T^\infty$ ) = \$1,874,790 , RRA = 5									
0	0.0	100.0	72.2	30.8					
1	-0.2	100.2	72.4	26.8	-2,005 1,793				
2	17.1	82.9	81.4	6.3	2,421 265	-3,647 908			
3	20.6	79.4	81.5	2.8	3,139 265	-3,022 667	-1,705 1,391		
4	26.4	73.6	80.5	1.8	5,151 265	-3,730 426	-1,731 908	-1,162 1,552	
5	29.0	71.0	79.9	1.4	7,665 265	-5,731 346	-1,481 748	-1,063 1,150	-810 1,632

Table 2b (continued)

$n$	Options (%)	Stock (%)	$CE(V_T^*)$ (%)	RMSE (%)	Option Positions in Optimal Portfolio with $n$ Options				
					Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)
$CE(W_T^*) = \$957,797$ , $CE(V_T^\infty) = \$900,296$ , $RRA = 10$									
0	0.0	39.0	67.9	28.4					
1	-1.7	74.7	86.5	4.5	-1,377 748				
2	-4.0	92.4	93.0	1.8	-1,367 506	-464 1,230			
3	-5.2	98.1	94.8	1.0	-1,200 426	-485 828	-277 1,471		
4	-7.4	107.4	97.2	0.7	-1,166 346	-479 667	-288 989	-215 1,552	
5	-7.1	106.2	96.9	0.6	-1,061 346	-418 587	-278 828	-209 1,150	-158 1,632
$CE(W_T^*) = \$681,834$ , $CE(V_T^\infty) = \$647,654$ , $RRA = 15$									
0	0.0	22.0	70.3	25.4					
1	-2.1	53.9	87.2	3.7	-1,003 587				
2	-4.0	64.4	90.7	1.4	-1,008 426	-259 1,069			
3	-6.1	72.9	93.0	0.8	-1,028 346	-283 748	-146 1,391		
4	-5.9	71.8	92.8	0.5	-974 346	-233 667	-136 989	-93 1,552	
5	-9.6	84.4	95.4	0.4	-1,037 265	-290 506	-168 748	-106 1,069	-86 1,552
$CE(W_T^*) = \$560,880$ , $CE(V_T^\infty) = \$537,074$ , $RRA = 20$									
0	0.0	15.0	73.7	22.6					
1	-1.5	37.6	86.8	3.0	-708 587				
2	-3.0	46.4	89.9	1.2	-760 426	-155 1,069			
3	-4.7	53.6	91.8	0.6	-808 346	-177 748	-86 1,391		
4	-8.5	66.4	94.4	0.4	-971 265	-191 587	-98 908	-67 1,471	
5	-7.8	63.9	94.1	0.3	-866 265	-193 506	-105 748	-64 1,069	-50 1,552

Table 3a

Utility-optimal buy-and-hold portfolios of stocks, bonds, and  $n$  European call options for CRRA utility under a bivariate linear diffusion stock-price process with parameters  $(\sigma_1, \sigma_2, \rho, \kappa, \theta)$  of the steady-state distribution calibrated to match the following moments:  $E[\log(P_t/P_{t-1})] = 0.15$ ,  $\text{Var}[\log(P_t/P_{t-1})] = 0.04$ ,  $\text{Var}[\mu_t] = 0.025^2$ ,  $\text{Corr}[\mu_t, \mu_{t-1}] = 0.05$ , and  $\rho = 0$ . Other calibrated parameters include: riskless rate  $r = 5\%$ , initial stock price  $P_0 = \$50$ , initial wealth  $W_0 = \$100,000$ , and time period  $T = 20$  years. ‘RRA’ denotes the coefficient of relative risk aversion, ‘ $\text{CE}(W_T^*)$ ’ denotes the certainty equivalent of the optimal dynamic stock/bond policy, ‘ $\text{CE}(V_T^\infty)$ ’ denotes the certainty equivalent of the optimal buy-and-hold portfolio with a continuum of options, and ‘ $\text{CE}(V_T^*)$ ’ denotes the certainty equivalent of the optimal buy-and-hold portfolio with a finite number  $n$  of options, reported as a percentage of  $\text{CE}(V_T^\infty)$ .

$n$	Options (%)	Stock (%)	$\text{CE}(V_T^*)$ (%)	RMSE (%)	Option Positions in Optimal Portfolio with $n$ Options		
					Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)
$\text{CE}(W_T^*) = \$11,861,394$ , $\text{CE}(V_T^\infty) = \$10,142,498$ , RRA = 1 (Log Utility)							
0	0.0	100.0	19.8	4,346.6			
1	60.3	39.7	68.2	4,341.5	56,231 737		
2	79.9	20.1	87.5	4,332.2	13,577 403	151,097 1,405	
3	99.3	0.7	91.9	4,332.2	786 68	12,725 403	152,096 1,405
$\text{CE}(W_T^*) = \$1,778,906$ , $\text{CE}(V_T^\infty) = \$1,645,135$ , RRA = 2							
0	0.0	100.0	81.8	214.7			
1	64.6	35.4	94.4	197.4	2,258 68		
2	58.0	42.0	99.2	155.4	1,600 68	3,159 403	
3	58.2	41.8	99.4	106.3	1,624 68	2,811 403	4,379 1,405
$\text{CE}(W_T^*) = \$575,004$ , $\text{CE}(V_T^\infty) = \$557,315$ , RRA = 5							
0	0.0	61.8	97.4	141.8			
1	-45.9	131.2	99.1	103.0	-1,604 68		
2	-36.7	119.8	99.8	35.4	-1,201 68	-595 403	
3	-36.8	120.0	99.8	14.8	-1,207 68	-568 403	-192 1,405

Table 3a (continued)

$n$	Options (%)	Stock (%)	$CE(V_T^*)$ (%)	RMSE (%)	Option Positions in Optimal Portfolio with $n$ Options		
					Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)
$CE(W_T^*) = \$395,205$ , $CE(V_T^\infty) = \$389,080$ , RRA = 10							
0	0.0	26.5	96.7	154.7			
1	-46.8	101.5	98.9	105.1	-1,635 68		
2	-37.3	89.5	99.7	25.3	-1,250 68	-386 403	
3	-37.4	89.6	99.7	6.1	-1,254 68	-376 403	-90 1,739
$CE(W_T^*) = \$348,828$ , $CE(V_T^\infty) = \$345,214$ , RRA = 15							
0	0.0	16.2	97.3	124.2			
1	-36.9	76.2	99.1	82.9	-1,288 68		
2	-29.5	66.8	99.8	17.6	-994 68	-261 403	
3	-29.5	66.8	99.8	4.3	-996 68	-255 403	-49 1,739
$CE(W_T^*) = \$327,732$ , $CE(V_T^\infty) = \$325,182$ , RRA = 20							
0	0.0	11.6	97.7	101.2			
1	-29.8	60.3	99.3	67.1	-1,041 68		
2	-23.8	52.8	99.8	13.3	-807 68	-196 403	
3	-23.9	52.8	99.8	3.3	-809 68	-191 403	-33 1,739

Table 3b

Mean-square-optimal buy-and-hold portfolios of stocks, bonds, and  $n$  European call options for CRRA utility under a bivariate linear diffusion stock-price process with parameters  $(\sigma_1, \sigma_2, \rho, \kappa, \theta)$  of the steady-state distribution calibrated to match the following moments:  $E[\log(P_t/P_{t-1})] = 0.15$ ,  $\text{Var}[\log(P_t/P_{t-1})] = 0.04$ ,  $\text{Var}[\mu_t] = 0.025^2$ ,  $\text{Corr}[\mu_t, \mu_{t-1}] = 0.05$ , and  $\rho = 0$ . Other calibrated parameters include: riskless rate  $r = 5\%$ , initial stock price  $P_0 = \$50$ , initial wealth  $W_0 = \$100,000$ , and time period  $T = 20$  years. 'RRA' denotes the coefficient of relative risk aversion, 'CE( $W_T^*$ )' denotes the certainty equivalent of the optimal dynamic stock/bond policy, 'CE( $V_T^\infty$ )' denotes the certainty equivalent of the optimal buy-and-hold portfolio with a continuum of options, and 'CE( $V_T^*$ )' denotes the certainty equivalent of the optimal buy-and-hold portfolio with a finite number  $n$  of options, reported as a percentage of CE( $V_T^\infty$ ).

$n$	Options (%)	Stock (%)	CE( $V_T^*$ ) (%)	RMSE (%)	Option Positions in Optimal Portfolio with $n$ Options				
					Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)
CE( $W_T^*$ ) = \$11,861,394 , CE( $V_T^\infty$ ) = \$10,142,498 , RRA = 1									
0	0.0	100.0	19.8	4,346.6					
1	0.2	99.8	20.0	3,579.8	35.36 × 10 <sup>6</sup> 14,771				
2	100.0	0.0	0.0	3,578.5	247,608 1,071	34.8 × 10 <sup>6</sup> 14,771			
3	2.8	97.2	22.7	3,564.7	2,695,502 5,415	-75.6 × 10 <sup>6</sup> 14,437	108.2 × 10 <sup>6</sup> 14,771		
4	8.1	91.9	27.7	3,563.8	1,023,256 3,410	3,412,389 8,756	-96.3 × 10 <sup>6</sup> 14,437	127.2 × 10 <sup>6</sup> 14,771	
5	15.6	84.4	34.5	3,563.7	497,357 2,407	1,586,801 6,083	3,353,435 10,427	-103.3 × 10 <sup>6</sup> 14,437	133.3 × 10 <sup>6</sup> 14,771
CE( $W_T^*$ ) = \$1,778,906 , CE( $V_T^\infty$ ) = \$1,645,135 , RRA = 2									
0	0.0	100.0	81.8	214.7					
1	0.5	99.5	82.9	62.2	16,875 2,407				
2	2.0	98.0	85.0	29.3	11,060 1,405	18,392 10,761			
3	7.6	92.4	89.3	17.9	7,084 737	8,847 4,412	17,361 14,771		
4	19.8	66.3	88.7	16.0	5,044 403	5,127 2,073	6,858 5,415	16,539 14,771	
5	17.5	82.5	94.0	15.5	4,417 403	4,021 1,739	4,093 3,744	4,828 6,751	15,662 14,771
CE( $W_T^*$ ) = \$575,004 , CE( $V_T^\infty$ ) = \$557,315 , RRA = 5									
0	0.0	24.6	85.0	27.9					
1	-0.1	42.5	94.3	7.7	-536 1,405				
2	-0.6	50.6	96.6	3.5	-546 737	-223 4,078			
3	-2.4	61.5	98.4	2.1	-615 403	-243 1,739	-158 6,417		
4	-2.2	60.2	98.3	1.3	-561 403	-215 1,405	-133 3,744	-109 10,427	
5	-2.0	58.8	98.1	1.0	-493 403	-189 1,071	-132 2,407	-101 5,415	-91 13,768

Table 3b (continued)

$n$	Options (%)	Stock (%)	$CE(V_T^*)$ (%)	RMSE (%)	Option Positions in Optimal Portfolio with $n$ Options				
					Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)	Quantity Strike (\$)
$CE(W_T^*) = \$395,204$ , $CE(V_T^\infty) = \$389,080$ , $RRA = 10$									
0	0.0	6.7	86.5	26.9					
1	-0.1	17.7	94.9	6.4	-295 1,071				
2	-1.7	29.8	98.0	3.0	-446 403	-107 2,407			
3	-1.5	27.8	97.7	1.7	-371 403	-101 1,405	-53 4,746		
4	-1.4	27.2	97.6	1.3	-340 403	-93 1,071	-54 2,742	-32 7,754	
5	-68.4	140.1	96.4	0.9	-2,357 68	-235 403	-100 1,071	-53 2,742	-33 7,754
$CE(W_T^*) = \$348,828$ , $CE(V_T^\infty) = \$345,214$ , $RRA = 15$									
0	0.0	3.7	89.7	21.4					
1	-0.3	13.9	97.1	5.0	-246 737				
2	-1.2	19.1	98.3	2.2	-304 403	-60 2,407			
3	-1.0	17.9	98.1	1.3	-257 403	-59 1,405	-27 4,746		
4	-47.8	96.7	98.5	1.0	-1,644 68	-188 403	-62 1,405	-27 4,746	
5	-51.5	102.6	97.9	0.7	-1,776 68	-162 403	-56 1,071	-29 2,407	-19 6,417
$CE(W_T^*) = \$327,732$ , $CE(V_T^\infty) = \$325,182$ , $RRA = 20$									
0	0.0	2.5	91.8	17.5					
1	-0.2	9.9	97.6	3.9	-179 737				
2	-0.9	14.0	98.6	1.7	-228 403	-41 2,407			
3	-0.8	13.2	98.4	1.1	-197 403	-40 1,405	-18 4,412		
4	-38.2	76.2	98.9	0.8	-1,316 68	-140 403	-44 1,405	-18 4,746	
5	-40.9	80.5	98.5	0.5	-1,411 68	-121 403	-40 1,071	-20 2,407	-12 6,417

## Chapter 2

# Pricing American Options: A Duality Approach

### Abstract

We develop a method for pricing and exercising high-dimensional American options. Our approach is based on approximate dynamic programming using nonlinear regression to approximate the value function. Our main theoretical result is a new representation of the American option price as a solution of a dual minimization problem. This dual problem involves pricing a look-back European-style option with a properly chosen process for the strike price. Based on this dual characterization of the price function, we construct upper and lower bounds on the option price, which can be evaluated by Monte Carlo simulation and are general enough to be used in conjunction with other approximate methods for pricing American options. We characterize the theoretical worst-case performance of the pricing bounds and discuss the implications for the design of the approximate pricing procedure. We illustrate the performance of our procedure on a set of sample problems where we price call options on the maximum and the geometric mean of a collection of stocks. Our numerical results are encouraging and suggest directions for future research.

*Co-author:* Leonid Kogan

## 2.1 Introduction

Valuation and optimal exercise of American options remains one of the most challenging practical problems in option pricing theory. The computational cost of traditional valuation methods, such as lattice and tree-based techniques, increases rapidly with the number of underlying securities and other payoff-related variables. Due to this well-known curse of dimensionality, practical applications of such methods are limited to problems of low dimension.

In recent years, several methods have been proposed to address this curse of dimensionality. Instead of using traditional deterministic approaches, these methods use Monte Carlo simulation to estimate option prices. Tilley (1993) was the first to demonstrate that American options could be priced using simulation techniques. Other important work in this literature includes Barraquand and Martineau (1995), Carriere (1996), Raymar and Zwecher (1997), Ibanez and Zapatero (1999) and Garcia (1999). Longstaff and Schwartz (2001) have proposed an approximate dynamic programming approach that can compute good price estimates and is very fast in practice. Tsitsiklis and Van Roy (1999, 2000) provide theoretical results that help explain the success of approximate dynamic programming methods. The price estimates these techniques construct, however, typically only give rise to lower bounds on the true option price. As a result, there is usually no formal method for evaluating the accuracy of the price estimates.

In an important contribution to the literature, Broadie and Glasserman (1997a,b) develop two *stochastic mesh* methods that allow them to generate lower and upper bounds on the option price that converge asymptotically to the true option price. The complexity of the first method, however, is exponential in the number of exercise periods. The second approach does not suffer from this drawback but nonetheless appears to be computationally demanding. In an effort to address this drawback, Boyle, Kolkiewicz and Tan (2001) generalize the stochastic mesh method of Broadie and Glasserman (1997b) using *low discrepancy sequences* to improve the efficiency of the approach.

In recent independent work, Rogers (2001) establishes a dual representation of American option prices similar to ours and applies the new representation to compute upper bounds on several types of American options using Monte Carlo simulation. However, he does not provide a systematic procedure for generating tight upper bounds and does not derive estimates of the accuracy of the bounds on the option price.

The contribution of this paper is two-fold. First, we develop a new representation of the American option price as a solution of a properly defined minimization problem. Using initial estimates of the option price that are generated using approximate dynamic programming techniques, we then apply the dual representation to construct upper and lower bounds on the option price. If the initial price estimates coincide with the true price, then the two bounds coincide with the price as well. Numerical values of the bounds can be estimated using simulation. We show that in the worst case the difference between the upper bound and the true price is given by a linear function of the approximation errors of the initial estimate of the option price. A similar result is also true of the lower bound. Our method for estimating the bounds on the option price is general in nature and can be used to enhance other valuation procedures proposed in the literature.

Second, we implement a fast and accurate valuation method based on approximate dynamic programming (see Bertsekas and Tsitsiklis 1996) where we use non-linear regression



techniques to approximate the value function. Unlike most procedures that use Monte Carlo simulation to estimate the continuation value of the option, our method is deterministic and relies on low discrepancy sequences as an alternative to Monte Carlo simulation. For the examples considered in this paper, we find that low discrepancy sequences provide a significant computational improvement over simulation. A further advantage of the dual representation of the American option price can be found by examining the functional form of the upper bound to which it gives rise. This functional form suggests how the approximate dynamic programming algorithm can be adapted to reduce the computational load that is required to obtain a given level of pricing accuracy.

The rest of the paper is organized as follows. In Section 2.2, we derive a new duality result for American options and use it to derive an upper bound on the option price. In Section 2.3, we describe our procedure for approximating the option price and in Section 2.4, we report results of numerical experiments, illustrating the performance of our procedure. We conclude in Section 2.5.

## 2.2 Problem Formulation

In this section we derive a new duality result for American options, which can be used to estimate the upper bound on the American option price. This section begins with the definition of the valuation problem in Section 2.2.1, followed by the duality result in Section 2.2.2.

### 2.2.1 The Model

**Information Set.** We consider an economy with a set of dynamically complete financial markets, described by the underlying probability space,  $\Omega$ , the sigma algebra  $\mathcal{F}$ , and the risk-neutral valuation measure  $\mathcal{Q}$ . It is well known (see Harrison and Kreps 1979) that if financial markets are dynamically complete, then under mild regularity assumptions there exists a unique risk-neutral measure, allowing one to compute prices of all state-contingent claims as the expected value of their discounted cash flows. The information structure in this economy is represented by the augmented filtration  $\{\mathcal{F}_t : t \in [0, T]\}$ . More specifically, we assume that  $\mathcal{F}_t$  is generated by  $Z_t$ , a  $d$ -dimensional standard Brownian motion, and the state of the economy is represented by an  $\mathcal{F}_t$ -adapted Markovian process  $\{X_t \in \mathbb{R}^d : 0 \leq t \leq T\}$ .

**Option Payoff.** Let  $h_t = h(X_t)$  be a nonnegative adapted process representing the payoff of the option, so that the holder of the option receives  $h_t$  if the option is exercised at time  $t$ . We also define a riskless account process  $B_t = \exp\left(\int_0^t r_s ds\right)$ , where  $r_t$  denotes the instantaneously risk-free rate of return. We assume that the discounted payoff processes satisfies the following integrability condition

$$E_0 \left[ \max_{t=0,1,\dots,T} \left| \frac{h_t}{B_t} \right| \right] < \infty \quad (2.2.1)$$

**Exercise Dates.** The American feature of the option allows its holder to exercise it at any of the pre-specified exercise dates  $\mathcal{T} = \{0, 1, \dots, T\}$ , equally spaced between 0 and  $T$ . This assumption is made to simplify the notation and is not restrictive.

**Option Price.** The value process of the American option,  $V_t$ , is the price process of the option conditional on it not having been exercised before  $t$ . It satisfies

$$V_t = \sup_{\tau \geq t} E_t^Q \left[ \frac{B_t h_\tau}{B_\tau} \right]. \quad (2.2.2)$$

where  $\tau$  is any stopping time with values in the set  $\mathcal{T} \cap [t, T]$  and  $E_t[\cdot]$  denotes the expected value under the risk-neutral measure  $Q$ , conditional on the time  $t$  information,  $\mathcal{F}_t$ . This is a well-known characterization of American options (see, for example, Bensoussan 1984, Karatzas 1988 and Pliska 1997).

## 2.2.2 The Dual Problem

The problem of pricing an American option, the *primal* problem, is that of computing

$$V_0 = \sup_{\tau \in \mathcal{T}} E_0 \left[ \frac{h_\tau}{B_\tau} \right]. \quad (2.2.3)$$

For an arbitrary adapted supermartingale  $\pi_t$ , we define the *dual* function to be

$$F(0, \pi) = E_0 \left[ \sup_{t \in \mathcal{T}} \left( \frac{h_t}{B_t} - \pi_t \right) \right] + \pi_0. \quad (2.2.4)$$

Then the dual problem is to minimize the dual function over all supermartingales  $\pi_t$ . Let  $U_0$  denote the optimal value of the dual problem, so that

$$U_0 = \inf_{\pi} F(0, \pi) = \inf_{\pi} E_0 \left[ \sup_{t \in \mathcal{T}} \left( \frac{h_t}{B_t} - \pi_t \right) \right] + \pi_0 \quad (2.2.5)$$

The following theorem shows that the optimal values of the dual and primal problems coincide.

**Theorem 1 (Duality Relation)** *The optimal values of the primal problem (2.2.3) and the dual problem (2.2.5) are equal, i.e.,  $V_0 = U_0$ . Moreover, an optimal solution of the dual problem is given by  $\pi_t^* = V_t/B_t$ , where  $V_t$  is the value process for the American option,*

$$V_t = \sup_{\tau \in \{t, T\} \cap \mathcal{T}} E_t \left[ \frac{h_\tau B_t}{B_\tau} \right].$$

**Proof.** For any supermartingale  $\pi_t$ ,

$$\begin{aligned} V_0 &= \sup_{\tau \in \mathcal{T}} \mathbb{E}_0 \left[ \frac{h_\tau}{B_\tau} \right] = \sup_{\tau \in \mathcal{T}} \mathbb{E}_0 \left[ \frac{h_\tau}{B_\tau} - \pi_\tau + \pi_\tau \right] \leq \sup_{\tau \in \mathcal{T}} \mathbb{E}_0 \left[ \frac{h_\tau}{B_\tau} - \pi_\tau \right] + \pi_0 \\ &\leq \mathbb{E}_0 \left[ \max_{t \in \mathcal{T}} \left( \frac{h_t}{B_t} - \pi_t \right) \right] + \pi_0 \end{aligned} \quad (2.2.6)$$

where the first inequality follows from the optional sampling theorem for supermartingales (see Billingsley 1995) and condition (2.2.1). Taking the infimum over all supermartingales,  $\pi_t$ , on the right hand side of (2.2.6) implies  $V_0 \leq U_0$ . On the other hand, the process  $V_t/B_t$  is a supermartingale, which implies

$$U_0 \leq \mathbb{E}_0 \left[ \max_{t \in \mathcal{T}} (h_t/B_t - V_t/B_t) \right] + V_0.$$

Since  $V_t \geq h_t$  for all  $t$ , we conclude that  $U_0 \leq V_0$ . Therefore,  $V_0 = U_0$ , and equality is attained when  $\pi_t^* = V_t/B_t$ . ■

Theorem 1 shows that an upper bound on the price of the American option can be constructed simply by evaluating the dual function over an arbitrary supermartingale  $\pi_t$ . In particular, if such a supermartingale satisfies  $\pi_t \geq h_t/B_t$ , the option price  $V_0$  is bounded above by  $\pi_0$ . Theorem 1 therefore implies the following well-known characterization of the American option price (see, for example, Pliska 1997).

**Corollary 1 (Option Price Characterization)** *The discounted option price process  $V_t/B_t$  is the smallest supermartingale,  $\pi_t$  that satisfies  $\pi_t \geq h_t/B_t$ .*

## Dynamic Replication

The pricing problem is closely related to the problem of dynamic replication of the American option, which is equally important in practice. While various methods for approximating American option prices have been suggested in the literature, computing reliable replication strategies has remained a challenging problem. The result of Theorem 1 suggests a method for super-replicating the American option. Super-replicating trading strategies are important for hedging the option since they almost surely dominate the payoff function of the option.

Corresponding to each supermartingale,  $\pi_t$ , there is a super-replicating strategy with initial cost equal to the dual function  $F(0, \pi)$ . This is seen by extending the definition of the dual function to all times between 0 and  $T$  so that

$$\frac{F(t, \pi)}{B_t} := \mathbb{E}_t \left[ \sup_{s \in \{t, T\} \cap \mathcal{T}} \left( \frac{h_s}{B_s} - \pi_s \right) \right] + \pi_t. \quad (2.2.7)$$

Both terms on the right of (2.2.7) are supermartingales, implying that  $F(t, \pi)/B_t$  is also a supermartingale. Since financial markets in our model are dynamically complete, this implies

(see Karatzas and Shreve 1998) that there exists a self-financing trading strategy with an initial cost  $F(0, \pi)$  which almost surely dominates  $F(t, \pi)$ . Since  $F(t, \pi)$  dominates the price of the option and hence its payoff at exercise, such a trading strategy super-replicates the payoff of the American option. Using an approximation to the option price, we can define  $\pi_t$  explicitly so that super-replicating the option can be a relatively straightforward task. In particular, Boyle et al. (1997) and Garcia et al. (2000) describe Monte Carlo methods for estimating the portfolio strategies replicating the present value process of a state contingent claim. This claim could correspond to a derivative security or some consumption process. Their results are therefore directly applicable to (2.2.7).

### Tightness of the Upper Bound

When the supermartingale  $\pi_t$  happens to coincide with the discounted option value process,  $V_t/B_t$ , the upper bound  $F(0, \pi)$  equals the true price of the American option. This suggests that a tight upper bound can be obtained by using an accurate approximation to the true option value process,  $\tilde{V}_t$ , to define  $\pi_t$ . In particular, we suggest two alternative definitions of  $\pi_t$ :

$$\pi_0 = \tilde{V}_0 \tag{2.2.8}$$

$$\pi_{t+1} = \pi_t + \frac{\tilde{V}_{t+1}}{B_{t+1}} - \frac{\tilde{V}_t}{B_t} - E_t \left[ \frac{\tilde{V}_{t+1}}{B_{t+1}} - \frac{\tilde{V}_t}{B_t} \right]^+ \tag{2.2.9}$$

or

$$\pi_{t+1} = \pi_t \left( \frac{B_t \tilde{V}_{t+1}}{B_{t+1} \tilde{V}_t} - E_t \left[ \frac{B_t \tilde{V}_{t+1}}{B_{t+1} \tilde{V}_t} - 1 \right]^+ \right) \tag{2.2.10}$$

where  $(x)^+ := \max(x, 0)$ . By construction,  $E_t[\pi_{t+1} - \pi_t] \leq 0$  for either definition of  $\pi_t$ , implying  $\pi_t$  is an adapted supermartingale. For the remainder of this paper, we will take  $\pi_t$  to be defined by (2.2.8) and (2.2.9). While we cannot say a priori that the upper bound corresponding to this supermartingale is tighter than the one determined by (2.2.8) and (2.2.10), it is considerably easier to analyze its performance. In particular, the following theorem relates the worst-case performance of the upper bound determined by (2.2.8) and (2.2.9) to the performance of the original approximation  $\tilde{V}_t$ .

**Theorem 2 (Tightness of the Upper Bound)** *Consider any approximation to the American option price,  $\tilde{V}_t$ , satisfying  $\tilde{V}_t \geq h_t$ , and let  $\bar{V}_0$  denote the upper bound corresponding to (2.2.8) and (2.2.9). Then,*

$$\bar{V}_0 \leq V_0 + 2 \sum_{t=0}^T E_0 \left[ \left| \frac{\tilde{V}_t}{B_t} - \frac{V_t}{B_t} \right| \right]. \tag{2.2.11}$$

**Proof.** Simplifying (2.2.8) and (2.2.9), and using Theorem 1, we obtain

$$\bar{V}_0 = \tilde{V}_0 + E_0 \left[ \sum_{j=1}^T \left( E_{j-1} \left[ \frac{\tilde{V}_j}{B_j} - \frac{\tilde{V}_{j-1}}{B_{j-1}} \right] \right)^+ \right] \quad (2.2.12)$$

as an upper bound on the price of the American option.<sup>1</sup> We then have

$$\begin{aligned} \bar{V}_0 &\leq \tilde{V}_0 + E_0 \left[ \sum_{j=1}^T \left( E_{j-1} \left[ \frac{\tilde{V}_j}{B_j} - \frac{V_j}{B_j} + \frac{V_j}{B_j} - \frac{V_{j-1}}{B_{j-1}} + \frac{V_{j-1}}{B_{j-1}} - \frac{\tilde{V}_{j-1}}{B_{j-1}} \right] \right)^+ \right] \\ &\leq \tilde{V}_0 + E_0 \left[ \sum_{j=1}^T \left( E_{j-1} \left[ \frac{\tilde{V}_j}{B_j} - \frac{V_j}{B_j} + \frac{V_{j-1}}{B_{j-1}} - \frac{\tilde{V}_{j-1}}{B_{j-1}} \right] \right)^+ \right] \\ &\leq V_0 + |\tilde{V}_0 - V_0| + E_0 \sum_{j=1}^T \left( E_{j-1} \left[ \left| \frac{\tilde{V}_j}{B_j} - \frac{V_j}{B_j} \right| \right] + \left| \frac{\tilde{V}_{j-1}}{B_{j-1}} - \frac{V_{j-1}}{B_{j-1}} \right| \right), \end{aligned}$$

where the second inequality is due to the supermartingale property of the discounted option price process,  $V_t$ , and the last step follows from the triangle inequality. The result now follows. ■

Theorem 2 shows that the upper bound remains tight as long as the approximation  $\tilde{V}_t$  is close to the true option price  $V_t$ . Accuracy of  $\tilde{V}_t$  depends on the details of the approximation procedure. If for example,  $\tilde{V}_t$  is constructed using the approximate dynamic programming algorithm of Longstaff and Schwartz (2000), then Tsitsiklis and Van Roy (2000) show that, in theory, the approximation error can be made arbitrarily small. This then suggests that the upper bound can be made arbitrarily tight, though the computational effort required might be prohibitive. According to (2.2.11), the worst-case estimate of the upper bound on the option price increases linearly with the number of exercise periods, holding the maturity date of the option fixed. This suggests that in practice the upper bound can be estimated more accurately for problems with a small number of exercise periods. Our numerical experiments confirm this intuition, as we shall see in Section 2.4.

## 2.3 Option Price Approximation

In this section we describe the approximate dynamic programming algorithm that we use for estimating the value function,  $V_t$ . Motivated by the upper bound, we will show in section 2.4.3 we will show how this algorithm might be improved.

---

<sup>1</sup>This bound may also be derived using the linear programming formulation of dynamic programming problems. See Appendix 3 where we also discuss the possibility of finding upper bounds for dynamic programming problems that are more general than optimal stopping problems.

The problem is to compute

$$V_0 = \sup_{\tau \in \mathcal{T}} E_0 \left[ \frac{h_\tau}{B_\tau} \right]. \quad (2.3.13)$$

In theory this problem is easily solved using value iteration. We solve for the value functions,  $V_t$ , recursively so that

$$V_T = h(X_T), \quad (2.3.14)$$

$$V_t = \max \left( h(X_t), E_t \left[ \frac{B_t}{B_{t+1}} V_{t+1}(X_{t+1}) \right] \right). \quad (2.3.15)$$

The price of the option is then given by  $V_0(X_0)$  where  $X_0$  is the initial state of the economy. In practice, however, if  $d$  is large so that  $X_t$  is high dimensional, then the 'curse of dimensionality' implies that value iteration is not practical.

Because finding the exact solution is intractable, we instead attempt to find an approximate solution. Before describing the approximation algorithm in detail, however, we first describe the Q-value function.

### 2.3.1 The Q-Value Function

A concept that is closely related to the value function is the Q-value function, which has a number of desirable properties that the value function does not possess. In the context of optimal stopping, the Q-value function is defined as

$$Q_t(X_t) = E_t \left[ \frac{B_t}{B_{t+1}} V_{t+1}(X_{t+1}) \right] \quad (2.3.16)$$

so that the Q-value at time  $t$  is simply the expected value of the option, conditional on it not being exercised at time  $t$ . The value of the option is then given by

$$V_t(X_t) = \max(h(X_t), Q_t(X_t)). \quad (2.3.17)$$

We can also write

$$Q_t(X_t) = E_t \left[ \frac{B_t}{B_{t+1}} \max(h(X_{t+1}), Q_{t+1}(X_{t+1})) \right] \quad (2.3.18)$$

and equation (2.3.18) clearly gives a natural extension of value iteration to Q-value iteration. The algorithm we use in this paper consists of performing an approximate Q-value iteration, an approach similar to the one used by Longstaff and Schwartz (2000) and Tsitsiklis and Van Roy (2000).

There are a number of reasons for why it is preferable to concentrate on approximating  $Q_t$  rather than approximating  $V_t$  directly. Letting  $\tilde{Q}_t$  and  $\tilde{V}_t$  denote our estimates of  $Q_t$  and  $V_t$  respectively, we can write the defining equations for approximate Q-value and value iteration as follows:

$$\tilde{Q}_t = E_t \left[ \frac{B_t}{B_{t+1}} \max \left( h(X_{t+1}), \tilde{Q}_{t+1}(X_{t+1}) \right) \right] \quad (2.3.19)$$

$$\tilde{V}_t = \max \left( h(X_t), E_t \left[ \frac{B_t}{B_{t+1}} \tilde{V}_{t+1}(X_{t+1}) \right] \right). \quad (2.3.20)$$

First, the functional forms of (2.3.19) and (2.3.20) suggest that  $Q_t$  is smoother than  $V_t$ , and therefore more easily approximated. But the most important reason for concentrating on  $Q_t$  is the following. In order to obtain a lower bound, we need to simulate sample paths originating from  $X_0$  and apply the stopping strategy, as determined by  $\{\tilde{V}_t\}$  or  $\{\tilde{Q}_t\}$ , to these paths. The average discounted payoffs from these paths will then be an unbiased lower bound on the true price of the option. In order to see why it is important that we estimate  $\tilde{Q}_t$  and not  $\tilde{V}_t$ , consider a sample path,  $P$  say. Suppose at time  $t$  we have not yet exercised the option and the current state is  $X_t$ . We now need to determine whether or not the option should be exercised. If we only have  $\tilde{V}_t$  available to us, then we need to compare  $\tilde{V}_t(X_t)$  with  $h(X_t)$ . If  $\tilde{V}_t(X_t) > h(X_t)$  then we do not exercise. However if  $\tilde{V}_t(X_t)$  is only marginally greater than  $h(X_t)$ , then it may be the case that  $\tilde{V}_t(X_t)$  is actually attempting to approximate  $h(X_t)$ . However we misinterpret  $\tilde{V}_t(X_t)$  and assume that it is optimal to continue when in fact it is optimal to exercise. This problem can be quite severe when there are relatively few exercise periods because in this case, there is often a significant difference between the value of exercising now and the continuation value. When we have an estimate of  $Q_t(X_t)$  this problem does not arise because we have a direct estimate of the continuation value, an estimate which  $\tilde{V}_t(X_t)$  cannot provide.

### 2.3.2 Approximate Q-Value Iteration

In this section we describe the procedure we use for approximating the Q-value functions,  $\tilde{Q}_t$ . This algorithm is standard in the approximate dynamic programming literature. The first step is to select an *approximation architecture*,  $\left\{ \tilde{Q}_t(\cdot; \beta_t) : \beta_t \in \mathfrak{R}^N \right\}$ , which is a class of functions from which we select  $\tilde{Q}_t$ . This class is parametrized by the vector  $\beta_t \in \mathfrak{R}^N$  so that the problem of determining  $\tilde{Q}_t$  is reduced to the problem of selecting  $\beta_t$  where  $\beta_t$  is chosen to minimize some approximation error.

Possible architectures are the linearly parametrized architectures of Longstaff and Schwartz (2000) and Tsitsiklis and Van Roy (2000), or non-linearly parametrized architectures such as neural networks or support vector machines (see Vapnik 1999). In this paper we use a

multi-layered perceptron with a single hidden layer, a particular class of neural networks. Multi-layered perceptrons with a single hidden layer are known to possess the universal approximation property so that they are able to approximate any continuous function over a compact set arbitrarily well, provided that a sufficient number of neurons are used (see Hornick, Stinchcombe and White 1989). Further details may be found in Appendix B which is available upon request from the authors.

The second step in the procedure is to select for each  $t = 1, \dots, K - 1$ , a set

$$S_t := \{P_1^t, \dots, P_{N_t}^t\} \quad (2.3.21)$$

of *training points* where  $P_n^t \in \mathfrak{R}^d$  for  $n = 1, \dots, N_t$ . The sets  $S_t$  are typically chosen in such a way that they are representative of the probability distribution of  $X_t$ .

Finally, we carry out an approximate Q-value iteration. Defining  $\tilde{Q}_T \equiv 0$ , we begin with  $t = K - 1$  and for  $n = 1, \dots, N_t$ , we use  $\hat{Q}_t(P_n^t)$  to estimate  $Q_t(P_n^t)$  where

$$\hat{Q}_t(P_n^t) := \hat{E}_t \left[ \frac{B_t}{B_{t+1}} \max \left( h(X_{t+1}), \tilde{Q}_{t+1}(X_{t+1}) \right) \right]. \quad (2.3.22)$$

The operator  $\hat{E}[\cdot]$  in (2.3.22) is intended to approximate the expectation operator,  $E[\cdot]$ . This is necessary as it is usually not possible to compute the expectation exactly on account of the high dimensionality of the state space. For example,  $\hat{E}[\cdot]$ , could correspond to Monte Carlo simulation where we simulate from the distribution of  $X_{t+1}$  given that  $X_t = P_n^t$ . In this paper, however, we use low discrepancy sequences (see Appendix B) rather than simulation to evaluate (2.3.22) as their use significantly improved the performance of the approximate dynamic programming algorithm.

We then estimate  $Q_t$  with  $\tilde{Q}_t(\cdot; \hat{\beta}_t)$  where

$$\hat{\beta}_t = \arg \min_{\beta_t} \sum_{n=1}^{N_t} \left( \hat{Q}_t(P_n^t) - \tilde{Q}_t(P_n^t; \beta_t) \right)^2. \quad (2.3.23)$$

In practice, we usually minimize a variant of the quantity in (2.3.23) in order to avoid the difficulties associated with overfitting. Once  $\tilde{Q}_t$  has been found, we then iterate in the manner of value iteration until we have found  $\tilde{Q}_0$ . Further details are summarized in Appendix B.

Once we have approximated  $Q_t$  for  $t = 0, \dots, T - 1$ , we need to compute an estimate of the price of the American option. We could use

$$\tilde{V}_0(X_0) = \max \left( h(X_0), \tilde{Q}_0(X_0) \right) \quad (2.3.24)$$

but such an estimate, however, is of limited value because we can say very little about the



estimation error. Even if we knew that this was a good estimate of the option price, it tells us very little about how to hedge the option. We address these issues next by constructing two price estimators whose expected values,  $\underline{V}_0$  and  $\bar{V}_0$ , are lower and upper bounds respectively on the true price of the option.

### 2.3.3 The Lower Bound on the Option Price

We can use the  $\tilde{Q}_t(\cdot)$ 's for  $t = 1, \dots, T - 1$  to construct a lower bound on the true option price. For each sample path originating from  $X_0$ , we find the first exercise period  $t$ , if it exists, in which  $h(X_t) \geq \tilde{Q}_t(X_t)$ . The option is then exercised at this time and the discounted payoff of the path is given by  $h(X_t)/B_t$ . Since this is a feasible  $\mathcal{F}_t$ -adapted exercise policy, it is clear that the expected discounted payoff from following this policy defines a lower bound  $\underline{V}_0$  on the true option price,  $V_0$ . Formally,  $\tilde{\tau} = \min\{t \in \mathcal{T} : \tilde{Q}_t \leq h_t\}$  and

$$\underline{V}_0 = E_0 \left[ \frac{h_{\tilde{\tau}}}{B_{\tilde{\tau}}} \right].$$

By replacing the expected value of discounted payoffs with an average over a large number of paths, we obtain an unbiased estimate of the lower bound.

The estimate of the lower bound performs very well in our numerical experiments, as reported in Section 2.4. The main reason for this is that it is only necessary that  $\tilde{Q}_t$  be a good approximation to  $Q_t$  when  $X_t$  is close to the optimal exercise frontier. To see the intuition behind this observation, consider how an error can be made on some arbitrary sample path. The only way an error can be made is if there is a point on the path,  $X_t$ , such that the option has not already been exercised before time  $t$ , and

$$Q_t(X_t) < h(X_t) < \tilde{Q}_t(X_t). \quad (2.3.25)$$

or

$$Q_t(X_t) > h(X_t) > \tilde{Q}_t(X_t) \quad (2.3.26)$$

If an error takes place and  $\tilde{Q}_t(X_t)$  is reasonably close to  $Q_t(X_t)$ , then  $h(X_t) \approx Q_t(X_t)$  and therefore the magnitude of the error will be small. Furthermore, in problems where there are relatively few exercise periods, the probability that  $h(X_t) \approx Q_t(X_t)$  at any point  $X_t$  on the sample path is often very low. Therefore, when there are relatively few exercise periods, the low probability of an error and the likely small magnitude of any such error that does occur, suggest that the lower bound should be a very good approximation to the true price of the option. This intuition is confirmed by the following theorem, characterizing the worst-case performance of the lower bound.

**Theorem 3 (Tightness of the Lower Bound)** *The lower bound on the option price satisfies*

$$\underline{V}_0 \geq V_0 - E_0 \left[ \sum_{t=0}^T \frac{|\tilde{Q}_t - Q_t|}{B_t} \right]. \quad (2.3.27)$$

**Proof.** At time  $t$ , the following six mutually exclusive events are possible: (i)  $\tilde{Q}_t \leq Q_t \leq h_t$ , (ii)  $Q_t \leq \tilde{Q}_t \leq h_t$ , (iii)  $\tilde{Q}_t \leq h_t \leq Q_t$ , (iv)  $Q_t \leq h_t \leq \tilde{Q}_t$ , (v)  $h_t \leq Q_t \leq \tilde{Q}_t$ , (vi)  $h_t \leq \tilde{Q}_t \leq Q_t$ . We define  $\tilde{\tau}_t = \min\{s \in [t, T] \cap \mathcal{T} : \tilde{Q}_s \leq h_s\}$  and

$$\underline{V}_t = B_t E_t \left[ \frac{h_{\tilde{\tau}_t}}{B_{\tilde{\tau}_t}} \right].$$

For each of the six scenarios above, we establish a relation between the lower bound and the true option price.

(i),(ii) The algorithm for estimating the lower bound correctly prescribes immediate exercise of the option so that  $V_t - \underline{V}_t = 0$ .

(iii) In this case the option is exercised incorrectly.  $\underline{V}_t = h_t$  and  $V_t = Q_t$  implying  $V_t - \underline{V}_t \leq |\tilde{Q}_t - Q_t|$ .

(iv) In this case the option is not exercised though it is optimal to do so. Therefore

$$\underline{V}_t = \frac{B_t}{B_{t+1}} E_t [V_{t+1}]$$

while

$$V_t = h_t \leq Q_t + (\tilde{Q}_t - Q_t) = \frac{B_t}{B_{t+1}} E_t [V_{t+1}] + (\tilde{Q}_t - Q_t).$$

This implies

$$V_t - \underline{V}_t \leq |\tilde{Q}_t - Q_t| + \frac{B_t}{B_{t+1}} E_t [V_{t+1} - \underline{V}_{t+1}].$$

(v),(vi) In this case the option is correctly left unexercised so that

$$V_t - \underline{V}_t = \frac{B_t}{B_{t+1}} E_t \left[ V_{t+1} - \underline{V}_{t+1} \right].$$

Therefore by considering the six possible scenarios, we find that

$$V_t - \underline{V}_t \leq \left| \tilde{Q}_t - Q_t \right| + \frac{B_t}{B_{t+1}} E_t \left[ V_{t+1} - \underline{V}_{t+1} \right].$$

Iterating and using the fact that  $\underline{V}_T = V_T$  implies the result. ■

While this theorem might suggest that the performance of the lower bound deteriorates linearly in the number of exercise periods, in our numerical experiments we find that this is not the case. This is not inconsistent with Theorem 3 since this theorem describes the worst case performance of the bound. In particular, in order for the exercise strategy that defines the lower bound to achieve the worst case performance, it is necessary that at each exercise period the condition (2.3.25) is satisfied. For this to happen, it must be the case that at *each* exercise period, the underlying state variables are close to the optimal exercise boundary. In addition,  $\tilde{Q}_t$  must systematically overestimate the true value  $Q_t$  so that the option is not exercised while it is optimal to do so. In practice, the variability of the underlying state variables,  $X_t$ , might suggest that  $X_t$  spends little time near the optimal exercise boundary. This suggests that as long as  $\tilde{Q}_t$  is a good approximation to  $Q_t$  near the optimal exercise frontier, the lower bound should be a good estimate of the true price, regardless of the number of exercise periods.

### 2.3.4 The Upper Bound on the Option Price

In Section 2.2.2 we derived an expression for an upper bound on the price of the American option. This was done by defining an appropriate supermartingale as in (2.2.8) and (2.2.9). However, in (2.2.8) we can alternatively set  $\tilde{V}_0 = \underline{V}_0$ . In this case, the upper bound is given by

$$\bar{V}_0 = \underline{V}_0 + E_0 \left[ \sum_{j=1}^T \left( E_{j-1} \left[ \frac{\tilde{V}_j}{B_j} - \frac{\tilde{V}_{j-1}}{B_{j-1}} \right] \right)^+ \right]. \quad (2.3.28)$$

Then (2.3.28) provides a natural decomposition of the upper bound  $\bar{V}_0$  into the sum of two components. The first component is the lower bound,  $\underline{V}_0$ , while the second component measures the extent to which  $\tilde{V}_t/B_t$  does not behave as a supermartingale, a property that is possessed by the true discounted option price process.

We estimate  $\bar{V}_0$  by simulating sample paths of the state variables, evaluating

$$\sum_{j=1}^T \left( E_{j-1} \left[ \frac{\tilde{V}_j}{B_j} - \frac{\tilde{V}_{j-1}}{B_{j-1}} \right] \right)^+ \quad (2.3.29)$$

along each path and taking the average over all paths. Evaluating (2.3.29) numerically is time consuming since at each point,  $(t, X_t)$ , on the path, we need to compute

$$\left( E_t \left[ \frac{\tilde{V}_{t+1}}{B_{t+1}} - \frac{\tilde{V}_t}{B_t} \right] \right)^+ . \quad (2.3.30)$$

According to Jensen's inequality, any unbiased noisy estimate of the expectation in (2.3.30), however, will result in an upwards biased estimate of the upper bound. It is therefore important that accurate estimates of the expectation in (2.3.30) can be computed. To do this we again use low discrepancy sequences.

## 2.4 Numerical Results

In this section we illustrate our method by pricing call options on the maximum and geometric mean of a collection of underlying stocks. The numerical results of this section were obtained using the methodology described earlier. In addition, however, problem specific information was also used to improve the initial estimates of the option price. For example, it is well known that the American option price is greater than or equal to the price of the corresponding European option. It is straightforward to incorporate such information into the approximate dynamic programming algorithm and as a result, the initial approximation,  $\tilde{V}_t$ , is often significantly improved. Further details of this technique, known as policy fixing (see Broadie and Glasserman 1997b), are given in Appendix B.

We assume that the market has  $N$  traded securities with price processes given by

$$dS_t^i = S_t^i [(r - \delta_i)dt + \sigma_i dz_t^i] \quad (2.4.31)$$

where  $z_t^i$  is a standard Brownian motion and the instantaneous correlation of  $z_t^i$  and  $z_t^j$  is  $\rho_{ij}$ . Each security pays dividends at a continuous rate of  $\delta_i$ . We assume that the option expires at time  $T$  and that there are  $n + 1$  equally spaced exercise dates in the interval  $[0, T]$ . The first date occurs at  $t = 0$  which we call the  $0^{th}$  exercise period and the  $n^{th}$  exercise date occurs at  $t = T$ . We use  $k$  to denote the strike price of the option and let  $r$  be the annual continuously compounded interest rate.

### 2.4.1 Call on the Maximum of 5 Assets

We assume that there are 5 assets,  $r = 0.05$ ,  $T = 3$  years and  $k = 100$ . We let  $\delta_i = 0.1$ ,  $\sigma_i = 0.2$  and  $\rho_{ij} = 0$  for  $i, j = 1, \dots, 5$ . All stocks are assumed to have the same initial price  $S_0$ . The “true” prices of the American options in Table 2.1 are in fact estimates of the true prices that are taken from Broadie and Glasserman (1997b). We let the number of exercise periods be 4, 7, and 10.

The lower bound provides a very accurate estimate of the option price and the difference between the upper and the lower bound is small relative to the early exercise premium of the American option under consideration. The upper bound therefore provides useful information about the precision of our estimate of the option price. Consistent with our discussion in Section 2.3.4, the upper bound is tighter for problems with fewer exercise periods.

### 2.4.2 Call on the Geometric Mean of 5 Assets

For the American call option on the geometric mean of a collection of stocks the true price of the option can be computed using a standard binomial tree, since the stochastic process that describes the evolution of the geometric mean is itself a geometric Brownian motion.

Table 2.2 shows that the lower bound is very close to the true price of the option, even when the total number of exercise periods is 101. It is a more precise estimate of the true price than the upper bound, which is consistent with our discussion in Sections 2.3.3 and 2.4.3. When there are 101 exercise periods the upper bound is relatively tight, although it is clearly not as good as the upper bound for the problem where there are only 11 exercise periods.

### 2.4.3 Approximate Q-Value Iteration Revisited

Perhaps the obvious way to compute the lower and upper bounds is in a sequential fashion so that after estimating the Q-value functions, we simulate a number of sample paths to compute the lower bound, and simulate another set to compute the upper bound. One difficulty with this strategy, however, is that the difference between  $\underline{V}_0$  and  $\bar{V}_0$  might be significant so that there is a large duality gap. When this occurs, we are forced to re-estimate the  $Q_t$ 's, possibly using more training points or a more flexible approximation architecture. This process may need to be repeated a number of times before we obtain a sufficiently small duality gap. Such an ad hoc approach might be very inefficient and we now propose a solution to address this problem.

The lower bound is typically a considerably better estimate than the upper bound of the true option price. This is borne out by our numerical experiments so that even when the upper bound is a very good estimate, the lower bound is still better. The reason for this is quite straightforward. We argued previously that while the worst-case performance of the lower bound is linear in the number of exercise periods, this worst-case performance is very rarely attained. Indeed, good performance of the lower bound depends mainly on how well the Q-value function is approximated when the state variables are close to the optimal exercise boundary. In contrast, the functional form of the upper bound suggests

that its performance deteriorates linearly in the number of time periods. Furthermore, the functional form of the upper bound also suggests that it is important for the Q-value function to be well approximated throughout the state space, rather than just the subset of the state space where it is close to the optimal exercise boundary.

This then suggests that when there is a large duality gap it is usually because the upper bound is not sufficiently close to  $V_0$ . The expression for the upper bound, as given in (2.3.28), suggests that if  $E_{t-1} [\tilde{V}_t/B_t - \tilde{V}_{t-1}/B_{t-1}]^+$  tends to be large, then the upper bound will not be very tight. Based on this observation, we propose the following modification to the approximate Q-value iteration.

After  $\tilde{Q}_t$  has been computed, we do not proceed directly to computing  $\tilde{Q}_{t-1}$ . Instead, we simulate a number of points from the distribution of  $X_t$  and for each point, we compute  $E_t [\tilde{V}_{t+1}/B_{t+1} - \tilde{V}_t/B_t]^+$ . If the average value of these quantities is below some threshold,  $\epsilon_t$ , then believing that we have a good estimate of  $\tilde{V}_t$ , we proceed to estimate  $\tilde{Q}_{t-1}$ . Otherwise we re-estimate  $\tilde{Q}_t$ , either by increasing the number of time  $t$  training points or by refining the approximation architecture, depending on the remedy that seems more appropriate. We then repeat the process until the average is less than  $\epsilon_t$ . The resulting estimates of the Q-value functions should lead to tight lower and upper bounds.

A further advantage of this proposal is that it allows us to determine how much computational effort is required to obtain a good solution. In particular, we can now determine online how many training points are needed or how complex the approximation architecture needs to be in order to obtain good estimates of the option price.

## 2.5 Conclusions and Further Research

In this paper we have developed a method for pricing and exercising high-dimensional American options. Our approach is based on approximate dynamic programming using nonlinear regression to approximate the value function. Our main theoretical result is a new representation of the American option price as a solution of a dual minimization problem. Based on this dual characterization of the price function, we construct upper and lower bounds on the option prices that can be evaluated by Monte Carlo simulation and are general enough to be used in conjunction with other approximate methods for pricing American options. Our procedure performs well on a set of sample problems where we price call options on the maximum and the geometric mean of a collection of stocks.

Estimating tight upper bounds on the option price for problems with numerous exercise periods remains a challenging and important aspect of the pricing problem. We are currently pursuing several alternative approaches. First, we could employ an extrapolation procedure such as Richardson extrapolation, to extend estimates of the upper bound for problems with relatively few exercise periods to those with many exercise periods. This technique has been used successfully in other financial applications to price American options. (See, for example, Geske and Johnson 1984 and Broadie, Glasserman and Jain 1997).

Such an approach could be quite effective in practice, but it suffers from an important limitation since estimates of the upper bound obtained by applying the extrapolation procedure are no longer upper bounds on the true option price.

Another approach is to estimate the option price and a corresponding supermartingale,  $\pi_t$ , for an auxiliary problem with relatively few exercise periods. Then we extend the process  $\pi$  to all exercise dates of the original problem. For instance, if  $t_1$  and  $t_2$  are two consecutive exercise dates for the auxiliary problem and  $t \in (t_1, t_2)$  is the exercise date of the original problem, one would define  $\pi_t = E_t[\pi_{t_2}]$ . Such an approach might prove useful, particularly for estimating prices of out-of-the-money options.

Another obvious direction for future research is to use the method to price more complicated American options. Our procedure is applicable to any American option as long as the underlying state process is Markovian. In addition, identifying the types of problems on which our procedure works well would be of considerable value.

Another subject for future research is to compare the performance of the two possible upper bounds as determined by the formulations in (2.2.9) and (2.2.10). If neither upper bound dominates the other, then it would be useful to determine the circumstances under which one bound is superior to the other.

Table 2.1: Call on the maximum

Table 2.1 lists the estimates of the price of an American call option on the maximum of 5 assets. We use the following set of parameter values:  $r = 0.05$ ,  $T = 3$ ,  $k = 100$ ,  $\delta_i = 0.1$ ,  $\sigma_i = 0.2$  and  $\rho_{ij} = 0$  for  $i, j = 1, \dots, 5$ . All stocks are assumed to have the same initial price  $S_0$ . "CI Lower Bd" and "CI Upper Bd" denote confidence intervals for the lower and upper bounds, respectively. The column "True Price\*" contains the estimates of the true prices of the American options taken from Broadie and Glasserman (1997b). We let the number of exercise periods be 4, 7, and 10.

$S_0$	95% CI Lower Bd	95% CI Upper Bd	True Price*	European Price
4 exercise periods				
90	[15.990, 16.020]	[16.041, 16.049]	16.006	14.586
100	[25.260, 25.296]	[25.325, 25.338]	25.284	23.052
110	[35.666, 35.708]	[35.778, 35.791]	35.695	32.685
7 exercise periods				
90	[16.449, 16.478]	[16.617, 16.642]	16.474	14.586
100	[25.902, 25.937]	[26.036, 26.054]	25.920	23.052
110	[36.467, 36.507]	[36.606, 36.627]	36.797	32.685
10 exercise periods				
90	[16.627, 16.655]	[16.819, 16.845]	16.659	14.586
100	[26.138, 26.172]	[26.261, 26.280]	26.158	23.052
110	[36.762, 36.801]	[37.041, 37.083]	36.782	32.685



Table 2.2: Call on the geometric average

Table 2.2 lists the estimates of the price of an American call option on the geometric average of 5 assets. We use the following set of parameter values:  $r = 0.03$ ,  $T = 1$ ,  $k = 100$ ,  $\delta_i = 0.05$ ,  $\sigma_i = 0.4$  and  $\rho_{ij} = 0$  for  $i, j = 1, \dots, 5$ . All stocks are assumed to have the same initial price  $S_0$ . “CI Lower Bd” and “CI Upper Bd” denote confidence intervals for the lower and upper bounds, respectively. We let the number of exercise periods be 11 and 101.

$S_0$	95% CI Lower Bd	95% CI Upper Bd	True Price	European Price
11 exercise periods				
90	[1.358, 1.364]	[1.388, 1.391]	1.362	1.172
100	[4.284, 4.294]	[4.353, 4.358]	4.291	3.445
110	[10.204, 10.216]	[10.265, 10.270]	10.211	7.521
101 exercise periods				
90	[1.381, 1.401]	[1.486, 1.488]	1.392	1.172
100	[4.352, 4.385]	[4.5222, 4.5224]	4.371	3.445
110	[10.402, 10.437]	[10.552, 10.553]	10.432	7.521

# Chapter 3

## Portfolio Optimization and Approximate Dynamic Programming

### 3.1 Introduction

In this chapter we explore some of the issues that arise when approximate dynamic programming (ADP) techniques are used to solve high-dimensional portfolio optimization problems. In particular, we study a five-dimensional finite horizon portfolio optimization problem where the agent wishes to maximize the expected utility of terminal wealth. There are two traded assets in the economy, an instantaneously risk-free bond and a risky stock. The short term interest rate,  $r_t$ , and the Sharpe ratio,  $s_t$ , are both stochastic. The agent holds a non-traded asset which is held until expiration and in each time period he receives a labor income. The value of the non-traded asset and the labor income are also stochastic.

We use an ADP algorithm to solve this problem and this leads naturally to a lower bound on the true value function. As is typically the case with ADP solutions, however, it is difficult to determine how good our approximate solution is. We attempt to overcome this problem by comparing the ADP solution to some good heuristic solutions which we describe in Section 3.4. This stands in contrast to the problem of pricing American options where we were able to compute upper bounds on the true value function.

Finally, we compare the work here with the recent work by Brandt et al (2001) who also solve portfolio optimization problems using ADP techniques.<sup>1</sup> This leads to possible suggestions for improving these algorithms and it also highlights the importance of accurately assessing the quality of an ADP solution.

In Section 3.2 we describe the portfolio optimization problem in more detail before describing the solution approach in Section 3.3. We present our numerical results in Section

---

<sup>1</sup>Other work in portfolio optimization that employs reinforcement learning and approximate dynamic programming techniques includes Neuneier (1996), Moody et al (1998) and Van Roy (1999). However, the paper by Brandt et al (2001) is easily the most closely related to the work in this chapter.

3.4 and conclude with suggestions for further research in Section 3.5.

## 3.2 Problem Formulation

We assume that the agent's utility function over terminal wealth is given by a power utility function,  $u(\cdot)$ , so that

$$u(W_T, H_T) = \frac{(W_T + H_T)^{1-\gamma}}{1-\gamma}$$

where  $\gamma > 0$  is the coefficient of relative risk aversion (CRRA).  $W_t$  and  $H_t$  are the values of the agent's liquid wealth and non-traded asset respectively, at time  $t$ , so that  $W_t + H_t$  is the total wealth at time  $t$ . We assume that trading takes place at discrete points in time and we use  $h$  to denote the length of the interval between trades.

The value function at time  $t$ ,  $J(\cdot)$ , is a function of  $X_t$  where  $X_t$  is the five-dimensional state vector given by

$$X_t \equiv [W_t \ H_t \ Y_t \ r_t \ s_t]. \quad (3.2.1)$$

The term  $Y_t$  in (3.2.1) is the flow of labour income that the agent receives at time  $t$  so that the agent receives  $Y_t h$  immediately prior to trading at time  $t + h$ . As mentioned above,  $r_t$  and  $s_t$  respectively denote the short-term interest rate and Sharpe ratio of the risky asset.

Letting  $R_t \equiv \log(r_t)$  and  $S_t \equiv \log(s_t)$ , the evolution of the state variables is described by the following stochastic differential equations <sup>2</sup>

$$\begin{aligned} dH_t &= \mu_H H_t dt + \sigma_H H_t dB_t^H \\ dY_t &= \mu_Y Y_t dt + \sigma_Y Y_t dB_t^Y \\ dR_t &= -\lambda_R (R_t - \bar{R}) dt + \sigma_R dB_t^R \\ dS_t &= -\lambda_S (S_t - \bar{S}) dt + \sigma_S dB_t^S \end{aligned}$$

where  $dB_t^i dB_t^j = \rho_{ij} dt$ . The price of the risky asset,  $P_t$ , is described by

$$dP_t = (r_t + \sigma_P s_t) P_t dt + \sigma_P P_t dB_t^P$$

---

<sup>2</sup>Since trading occurs at discrete intervals we could equally well use discrete versions of all the stochastic processes.

and the budget constraint implies that the evolution of liquid wealth is given by

$$dW_t = \left( [r_t + \theta_t \sigma_P s_t] W_t + Y_t \right) dt + \theta_t W_t \sigma_P dB_t^P \quad (3.2.2)$$

where  $\theta_t$  is the proportion of liquid wealth that is invested in the risky asset at time  $t$ . Since power utility is not defined over negative wealth and since trading occurs only at discrete intervals, it is necessary to impose the constraint that  $\theta_t \in [0, 1]$  for all  $t$ . This constraint is typical of the situation that many individuals face in practice.

Using Itô's Lemma, it is easily seen that

$$H_{t+h} = H_t \exp\left(\left(\mu_H - \frac{\sigma_H^2}{2}\right)h + \sigma_H(B_{t+h}^H - B_t^H)\right) \quad (3.2.3)$$

$$Y_{t+h} = Y_t \exp\left(\left(\mu_Y - \frac{\sigma_Y^2}{2}\right)h + \sigma_Y(B_{t+h}^Y - B_t^Y)\right) \quad (3.2.4)$$

$$r_{t+h} = \exp\left(\bar{R}(1 - e^{-\lambda_R h}) + e^{-\lambda_R h} R_t + \sigma_{R^e} e^{-\lambda_R(t+h)} \int_t^{t+h} e^{-\lambda_R u} dB_u^R\right) \quad (3.2.5)$$

$$s_{t+h} = \exp\left(\bar{S}(1 - e^{-\lambda_S h}) + e^{-\lambda_S h} S_t + \sigma_{S^e} e^{-\lambda_S(t+h)} \int_t^{t+h} e^{-\lambda_S u} dB_u^S\right) \quad (3.2.6)$$

$$P_{t+h} = P_t \exp\left(\left(r_t + \sigma_P s_t - \frac{\sigma_P^2}{2}\right)h + \sigma_P(B_{t+h}^P - B_t^P)\right). \quad (3.2.7)$$

We note that conditional on information at time  $t$ , the state variables at time  $t + h$  are all log-normally distributed and that their means and covariances are easily calculated.

The portfolio optimization problem is now given by

$$\max_{\theta_t \in [0,1]} E_0 \left[ \frac{(W_T + H_T)^{1-\gamma}}{1-\gamma} \right] \quad (3.2.8)$$

subject to the evolution of the state variables given by (3.2.3) to (3.2.7), and the discrete analogue of the budget constraint in (3.2.2).

Before describing the ADP algorithm, we note a standard property of the true value function that is commonly used in portfolio optimization problems and one that will assist us in choosing a suitable approximation architecture for the value function. The property states that we can scale some of the variables in the value function so as to reduce the dimensionality of the problem by one. It relies on the fact that we are assuming power utility and implies that we may write

$$J_t(W_t, H_t, Y_t, r_t, s_t) = \frac{W_t^{1-\gamma}}{1-\gamma} g_t\left(\frac{H_t}{W_t}, \frac{Y_t}{W_t}, r_t, s_t\right) \quad (3.2.9)$$

for some function  $g_t(\cdot)$ .<sup>3</sup> This is easily confirmed either by considering the Bellman equation or, in continuous time, by examining the HJB partial differential equation.

### 3.3 Approximate Value Iteration

The ADP algorithm we use is almost identical to that used in Chapter 2. We begin by choosing the approximation architecture that we use to approximate the true value function. Equation (3.2.9) implies that approximating  $J_t(\cdot)$  amounts to approximating  $g_t(\cdot)$  so that we need to estimate the four-dimensional function  $g_t(\cdot)$  as opposed to the five-dimensional value function.

We use the set of polynomial functions in  $V_t$  that are of degree less than or equal to  $n$ , where

$$V_t \equiv \left[ \frac{H_t}{W_t}, \frac{Y_t}{W_t}, r_t, s_t \right]. \quad (3.3.10)$$

We consider values of  $n$  equal to 1, 2 and 3. These give rise to linear, quadratic and cubic approximations to  $g_t(\cdot)$  respectively. For ease of exposition, we only describe the linear case but the quadratic and cubic cases are analogous though somewhat more expensive from a computational viewpoint.

If we assume a set of linear basis functions for the approximation architecture, then we approximate  $J_t(\cdot)$  with

$$\tilde{J}_t(W_t, H_t, Y_t, r_t, s_t) = \frac{W_t^{1-\gamma}}{1-\gamma} \left( b_0 + b_1 \frac{H_t}{W_t} + b_2 \frac{Y_t}{W_t} + b_3 r_t + b_4 s_t \right) \quad (3.3.11)$$

so that selecting  $\tilde{J}_t$  is equivalent to selecting the  $b_i$ 's.<sup>4</sup>

The next step is to determine a set of training points  $\{X_t^i\}$  for each  $t$  and for  $i = 1, \dots, N$ , where we recall that  $X_t$  denotes a vector of state variables  $[W_t H_t Y_t r_t s_t]$ . The training points are obtained by simulating  $N$  paths beginning from the initial state vector,  $X_0$ . A potential difficulty that arises is the endogeneity of the liquid wealth,  $W_t$ . In order to simulate sample values of  $W_t$ , we need to use some trading strategy. If the trading strategy is not close to optimal then our training points may not be sufficiently representative of the sample space and result in a poor ADP solution. We overcome this by using a good heuristic trading strategy which we describe in Section 3.4 to generate the training points. In order to ensure that these training points are in fact 'good' we could also use each training point to generate an additional training point. For example, if  $[W_t H_t Y_t r_t s_t]$  is a training point that we obtain from following a particular heuristic trading strategy, then the additional training

---

<sup>3</sup>We could also have chosen to scale by  $H_t$  or  $Y_t$  in (3.2.9), but for the purpose of selecting a good strategy, it is more convenient to scale by  $W_t$ .

<sup>4</sup>The  $b_i$ 's depend of course on  $t$ , but we suppress this dependence for notational simplicity.

point might be  $[W_t(1 + \epsilon_t) H_t Y_t r_t s_t]$  where  $\epsilon_t > 0$  is some constant. We found that this procedure for selecting training points did not produce an improvement to the results of Section 3.4, thereby suggesting that the heuristic strategy we used was sufficiently close to optimal for the purpose of selecting training points.

The third step in the algorithm is to perform an approximate value iteration. Beginning with  $t = T - h$ , we find the optimal strategy at each training point,  $X_t^i$ . This is equivalent to finding  $\hat{\theta}_t(X_t^i)$ , the proportion of liquid wealth that is invested in the risky asset. This is done by approximating the term  $W_t^{1-\gamma}/(1-\gamma)$  in the approximate value function with a second order Taylor expansion and is described in further detail in Section 3.3.1.

Once  $\hat{\theta}_t(X_t^i)$  has been found, we need to estimate  $J_t(X_t^i)$ . This is done by simulating successor points to  $X_t^i$ , assuming that  $\hat{\theta}_t(X_t^i)$  of  $W_t^i$  has been invested at time  $t$  in the risky asset. Using the time  $t + h$  approximation to the value function<sup>5</sup>,  $\tilde{J}_{t+h}(\cdot)$ , we average  $\tilde{J}_{t+h}(\cdot)$  over all successor points and take this average as our estimate for  $J_t(X_t^i)$ . This estimate is denoted by  $\hat{J}_t(X_t^i)$ .

Of course, we are not restricted to using Monte Carlo simulation for generating the successor points. As was the case with the problem of pricing American options, we used low discrepancy sequences to generate the successor points in the numerical results of Section 3.4.

When the estimates  $\hat{J}_t(X_t^i)$  have been determined, we determine the  $b_i$ 's in (3.3.11) by regressing the  $\hat{J}_t(X_t^i)$ 's (scaled by  $W_t^{1-\gamma}/(1-\gamma)$ ) on the basis functions in (3.3.10).

With  $\tilde{J}_t(\cdot)$  now determined, we move to time  $t - h$  in the usual manner of value iteration, and continue until we have found  $\tilde{J}_h(\cdot)$ . At this point, we could proceed to find  $\tilde{J}_0(\cdot)$ , and use this as an estimate of the value function. However, we can say very little about the quality of this estimate. Instead, we simulate the strategy as determined by  $\{\tilde{J}_t(\cdot)\}$  and use the simulations to construct an estimate of the value function. Since this is a feasible adapted trading strategy, the expected value from following the strategy constitutes a lower bound on the optimal value function. The estimate that we obtain from simulating the strategy is therefore an unbiased lower bound and this is the value that we report in the numerical results of Section 3.4.

We now describe the procedure for computing  $\hat{\theta}_t(X_t^i)$  using the estimated value function,  $\tilde{J}_{t+h}(\cdot)$ .

### 3.3.1 Quadratic Approximations to the Value Function

The approximate dynamic programming algorithm of the previous section requires the computation of  $\hat{\theta}_t(X_t^i)$  at each training point,  $X_t^i$ . This is also true when we wish to simulate the ADP solution. It is therefore desirable that we be able to compute  $\hat{\theta}_t(X_t^i)$  (or an approximation to it) very quickly. Unfortunately, given the functional form of  $\tilde{J}_t(\cdot)$  it is not possible to find  $\hat{\theta}_t(X_t^i)$  explicitly. One possibility is to use a numerical procedure though this would typically be quite slow. For the problem we are considering, we might attempt to find

<sup>5</sup>If  $t = T - h$ , then we use the true value function at time  $t + h = T$ .

$\widehat{\theta}_t(X_t^i)$  using an exhaustive search, for example. This would certainly be very slow, however, and would become impractical for problems where there were several risky assets.

Instead, we use an approximation to  $\widetilde{J}_t(\cdot)$  that enables us to quickly obtain an explicit estimate of  $\widehat{\theta}_t(X_t^i)$ . The approximation we use is obtained by replacing terms of the form  $W_t^\delta$  in  $\widetilde{J}_t(\cdot)$  with quadratic functions in  $W_t$ .

This type of approximation has also been examined in Brandt et al (2001)<sup>6</sup>. They conclude that while the quadratic approximations work well in some contexts, in general they do not provide a sufficiently good approximation. Instead they suggest using a quartic approximation to  $W_t^\delta$ . This gives rise to an implicit equation for  $\widehat{\theta}_t(X_t^i)$  which they solve iteratively, using the quadratic solution as a starting point for the iteration.

There are a number of reasons for why we use the quadratic approximation in this chapter. First, the solution to the quadratic approximation can be found much more quickly than the solution to the quartic approximation. (Later we will describe some situations where being able to solve the problem quickly might be important.)

Second, when Brandt et al take the quadratic approximation at time  $t + h$ , say, they expand about the point  $W_t \exp(r_t h)$ . This is a somewhat naive expansion since  $W_{t+h}$  will typically be larger than  $W_t \exp(r_t h)$ . In fact,  $W_{t+h}$  may actually be considerably larger than  $W_t \exp(r_t h)$  when  $h$ , the interval between trading, is large. Instead, better results might be obtained by expanding about  $\overline{W}_{t+h}$ , where  $\overline{W}_{t+h}$  is the expected wealth at time  $t + h$  when a good heuristic strategy is chosen at time  $t$ . This is the expansion that we use in this chapter and we shall see in Section 3.4 that it often results in a significantly superior solution.

Of course a third reason for exploring the quadratic expansion is that we would like to test it on a class of problems for which it hasn't been used before. In fact we shall see that some of the qualitative properties of the quadratic solution that were found by Brandt et al (2001) are reversed in Section 3.4. We now describe the quadratic expansion in more detail.

Assume that at time  $t$  the state vector is given by  $X_t$  and that we would like to find  $\widehat{\theta}_t(X_t)$ . We do this using  $\widetilde{J}_{t+h}$  which, as before, is given by<sup>7</sup>

$$\begin{aligned} \widetilde{J}_{t+h}(W_{t+h}, H_{t+h}, Y_{t+h}, r_{t+h}, s_{t+h}) &= \frac{W_{t+h}^{1-\gamma}}{1-\gamma} \left( b_0 + b_1 \frac{H_{t+h}}{W_{t+h}} + b_2 \frac{Y_{t+h}}{W_{t+h}} + b_3 r_{t+h} + b_4 s_{t+h} \right) \\ &= \frac{W_{t+h}^{1-\gamma}}{1-\gamma} \left( b_0 + b_3 r_{t+h} + b_4 s_{t+h} \right) + \frac{W_{t+h}^{-\gamma}}{1-\gamma} \left( b_1 H_{t+h} + b_2 Y_{t+h} \right) \end{aligned}$$

---

<sup>6</sup>See Brandt et al (2001) for other references to work that studies quadratic approximations to utility functions in portfolio optimization problems.

<sup>7</sup>When  $t + h = T$  then we know  $J_T(W_T, H_T) = \frac{W_T^{1-\gamma}}{1-\gamma} \left( 1 + \frac{H_T}{W_T} \right)^{1-\gamma}$ . However, in order to compute  $\widehat{\theta}_{T-h}(X_{T-h})$  explicitly, we instead use an approximation to  $J_T$ . This is done by substituting a cubic polynomial for  $\left( 1 + \frac{H_T}{W_T} \right)^{1-\gamma}$  in the expression for  $J_T$ . This cubic approximation is used regardless of whether or not the basis function we use in estimating  $J_t$  for  $t \leq T - h$  are linear, quadratic or cubic polynomials. Of course, there is no particular reason to choose a cubic. Indeed, a higher order polynomial could have been chosen and we could still compute  $\widehat{\theta}_{T-h}(X_{T-h})$  explicitly since the state variables are log-normally distributed.

We now approximate  $W_{t+h}^{1-\gamma}$  and  $W_{t+h}^{-\gamma}$  with second order Taylor expansions where we expand about

$$\bar{W}_{t+h} = W_t \left( \mu(X_t) e^{(r_t + \sigma_P s_t)h} + (1 - \mu(X_t)) e^{r_t h} \right) + Y_t h. \quad (3.3.12)$$

The term  $\mu(X_t)$  in (3.3.12) is the fraction of wealth invested in the risky asset by a 'good' heuristic trading strategy which we describe in Section 3.4. In particular, we have

$$\frac{W_{t+h}^{1-\gamma}}{1-\gamma} \approx \frac{\gamma(1+\gamma)}{2(1-\gamma)} \bar{W}_{t+h}^{1-\gamma} + (1-\gamma) \bar{W}_{t+h}^{-\gamma} W_{t+h} - \frac{\gamma}{2} \bar{W}_{t+h}^{-\gamma-1} W_{t+h}^2$$

and

$$\frac{W_{t+h}^{-\gamma}}{1-\gamma} \approx \frac{(\gamma+2)(\gamma+1)}{2(1-\gamma)} \bar{W}_{t+h}^{-\gamma} - \frac{\gamma(2+\gamma)}{1-\gamma} \bar{W}_{t+h}^{-\gamma-1} W_{t+h} + \frac{\gamma(1+\gamma)}{2(1-\gamma)} \bar{W}_{t+h}^{-\gamma-2} W_{t+h}^2$$

Simplifying our notation, we write

$$\begin{aligned} \frac{W_{t+h}^{1-\gamma}}{1-\gamma} &\approx q_1(X_t) + q_2(X_t) W_{t+h} + q_3(X_t) W_{t+h}^2 \\ \frac{W_{t+h}^{-\gamma}}{1-\gamma} &\approx q_4(X_t) + q_5(X_t) W_{t+h} + q_6(X_t) W_{t+h}^2 \end{aligned}$$

so that

$$\begin{aligned} \tilde{J}_{t+h}(W_{t+h}, H_{t+h}, Y_{t+h}, r_{t+h}, s_{t+h}) &\approx \bar{J}_{t+h}(W_{t+h}, H_{t+h}, Y_{t+h}, r_{t+h}, s_{t+h}) \\ &\equiv \left( q_1(X_t) + q_2(X_t) W_{t+h} + q_3(X_t) W_{t+h}^2 \right) (b_0 + b_3 r_{t+h} + b_4 s_{t+h}) \\ &\quad + \left( q_4(X_t) + q_5(X_t) W_{t+h} + q_6(X_t) W_{t+h}^2 \right) (b_1 H_{t+h} + b_2 Y_{t+h}) \end{aligned}$$

Estimating  $\hat{\theta}_t(X_t)$  now amounts to solving the first order conditions for

$$\max_{\theta_t \in [0,1]} \text{E}_t \left[ \bar{J}_{t+h}(W_{t+h}, H_{t+h}, Y_{t+h}, r_{t+h}, s_{t+h}) \right].$$



These conditions are easily solved explicitly<sup>8</sup>.

### 3.4 Numerical Results

In this section we describe the numerical results that were obtained using the ADP algorithm of Section 3.3. In order to assess the quality of the ADP solutions it is necessary to have some benchmark against which we may judge them. Ideally, we would be able to use the  $\{\tilde{J}_t\}$ 's to construct lower and upper bounds on the true value function, as was the case in Chapter 2. The duality gap would then constitute a useful measure of performance. At this moment, however, we do not know how to compute upper bounds so instead we attempt to find good heuristic strategies. The benchmark for the ADP solutions will then be whether or not they outperform the heuristic strategies.

We consider two classes of heuristic strategies, both of which are motivated by the solution to the standard portfolio optimization problem where the agent has power utility and where the returns on the risky asset in each time period are independent and identically distributed. In this case the agent invests a constant fraction of his wealth in the risky asset in each time period. Since the Sharpe ratio for this problem is constant, we could also say the agent invests a constant fraction of the Sharpe ratio in the risky asset in each time period. This observation gives rise to two heuristic strategies. In the first strategy the agent invests a constant fraction,  $\theta_c$ , of his liquid wealth in the risky asset in each time period. In the second strategy, the agent invests a constant fraction,  $C_s$ , of the *current* Sharpe ratio in the risky asset in each time period. We refer to these two heuristic strategies as the constant and Sharpe strategies, respectively. Of course these heuristic strategies are also subject to the constraint that  $\theta \in [0, 1]$ . We mention that we expect the constant strategy to be outperformed by the Sharpe strategy, which in turn should often be close to optimal<sup>9</sup>. In the numerical results of this section, the values of  $\theta_c$  and  $C_s$  that we report are the best values that we could find. However, the performances of the heuristic strategies are not very sensitive to  $\theta_c$  and  $C_s$  so that the values we report are only estimates of the optimal values.

In the following tables, we solve the ADP problem for three different sets of parameters. For each set, we consider  $T = 2$  and  $T = 5$  years, and assume that trading takes place on a quarterly basis. We also consider three different values of the coefficient of relative risk aversion (CRRA = 2, 5 and 10). Each of the three parameter sets has the following parameters in common<sup>10</sup>.

---

<sup>8</sup>The second order conditions should automatically be satisfied if  $\tilde{J}_{t+h}$  is a good approximation to  $J_{t+h}$  which we assume to be the case.

<sup>9</sup>We cannot prove this statement of course, so for now it is merely a conjecture.

<sup>10</sup>No effort was made to estimate these parameters from real data. They were chosen only so that they appeared 'reasonable'.

### Parameter Sets 1, 2, 3

$$\begin{aligned}
 W_0 &= \$100,000 & H_0 &= \$50,000 & Y_0 &= \$10,000 & r_0 &= .05 & s_0 &= .4 \\
 \sigma_H &= .15 & \sigma_Y &= .2 & \sigma_R &= .2 & \sigma_S &= .5 & \sigma_P &= .2 \\
 \mu_H &= .1 & \mu_Y &= .08 & \lambda_R &= .2 & \lambda_S &= .3 \\
 \bar{R} &= \log(.05) & \bar{S} &= \log(.4) \\
 h &= .25 \text{ Years}
 \end{aligned}$$

The parameter sets only differ in the instantaneous correlations that drive each of the Brownian motions.

#### Parameter Set 1

$$\begin{aligned}
 \rho_{HY} &= .5 & \rho_{HR} &= -.5 & \rho_{HS} &= .5 & \rho_{HP} &= .5 & \rho_{YR} &= 0 \\
 \rho_{YS} &= 0 & \rho_{YP} &= .5 & \rho_{RS} &= -.5 & \rho_{RP} &= -.5 & \rho_{SP} &= 0
 \end{aligned}$$

#### Parameter Set 2

$$\begin{aligned}
 \rho_{HY} &= .5 & \rho_{HR} &= 0 & \rho_{HS} &= .5 & \rho_{HP} &= -.25 & \rho_{YR} &= .5 \\
 \rho_{YS} &= 0 & \rho_{YP} &= -.25 & \rho_{RS} &= -.25 & \rho_{RP} &= .5 & \rho_{SP} &= 0
 \end{aligned}$$

#### Parameter Set 3

$$\begin{aligned}
 \rho_{HY} &= .75 & \rho_{HR} &= .75 & \rho_{HS} &= .75 & \rho_{HP} &= .75 & \rho_{YR} &= .75 \\
 \rho_{YS} &= .75 & \rho_{YP} &= .75 & \rho_{RS} &= .75 & \rho_{RP} &= .75 & \rho_{SP} &= .75
 \end{aligned}$$

The results for parameter sets 1, 2 and 3 are presented in Tables 3.1 to 3.4, Tables 3.5 to 3.8, and Tables 3.9 to 3.12, respectively. A number of features are worth noting that are common to the three sets of parameters.

First, it is clear that the solutions based upon using quadratic or cubic basis functions to approximate the value function significantly outperform the solution based upon linear basis functions. This is in contrast to Brandt et al (2001) where linear basis function marginally outperform quadratic basis functions<sup>11</sup>. It is also true that the cubic basis functions generally appear to be superior to the quadratic basis functions though the difference, as measured by difference in certainty equivalents, is marginal.

A second observation is how the quadratic expansion based upon the 'good' heuristic

<sup>11</sup>This distinction is somewhat ambiguous since Brandt et al (2001) use linear basis functions to estimate both the value function *and* derivatives of the value function.

strategy is superior to the ‘naive’ expansion where  $W_{t+h}^\delta$  is expanded about  $W_t \exp(r_t h)$ .<sup>12</sup> In some cases the difference between the two appears to be significant. This suggests that for some of the problems considered by Brandt et al (2001) the ‘good’ quadratic expansion might actually lead to a solution that is effectively optimal.

As expected for each parameter set, the Sharpe heuristic strategy is always superior to the constant strategy. The relationship between the Sharpe and ADP strategies is more interesting, however. For parameter set 1, the best ADP strategy is usually superior to the Sharpe strategy. However, this fails to be true for parameter set 2 where in some cases the Sharpe strategy significantly outperforms the best ADP strategy. In addition, the difference between the two increases as  $\gamma$ , the coefficient of relative risk aversion, increases. The reason for this may be explained as follows.

As Brandt et al (2001) point out, there are two opposing effects that occur as  $\gamma$  increases when we use the quadratic expansion to approximate  $W_{t+h}^\delta$ . The first effect is that the quadratic approximation becomes less accurate as  $\delta$  increases in magnitude and becomes more negative. On the other hand, as  $\delta$  increases in magnitude (and becomes more negative), the agent typically wishes to invest less in the risky asset. This has the effect of reducing the variance of  $W_{t+h}$  and so this tends to improve the performance of the approximation. The net effect for the examples of Brandt et al (2001) is that the ADP solution *improves* as  $\gamma$  increases.

For parameter set 2, however, the fraction of liquid wealth invested in the risky asset *increases* with  $\gamma$ . (We can confirm this by computing the average fraction of liquid wealth held in the risky asset for each value of  $\gamma$ . However, it may also be confirmed by observing how the values of  $C_s$  increase with  $\gamma$  in Tables 3.5 and 3.7.) This is due to a hedging effect that is explained by the fact that in parameter set 2,  $\rho_{HP}$  and  $\rho_{YP}$  are both negative. As a result, a more risk averse investor will seek to hedge more and for parameter set 2, this has the net effect of causing the more risk averse investor to invest more in the risky asset. As a result, the ADP solutions for parameter set 2 deteriorate with  $\gamma$ . Furthermore, they do not do nearly as well as the ADP solutions of parameter set 1 when compared to the heuristic strategies. Again, this may be explained by the negativity of  $\rho_{HP}$  and  $\rho_{YP}$  in parameter set 2 which causes agents to hold more of the risky asset. We can also confirm these observation by considering Tables 3.9 to 3.12 where parameter set 3 was used. For these tables,  $\rho_{HP} = \rho_{YP} = .75$  which is larger than their common value of .5 in parameter set 1. In this case we would expect the large positive correlation of  $P_t$  with  $H_t$  and  $Y_t$  to have the effect of reducing the holdings in the risky asset thereby improving the performance of the ADP solutions. This may be confirmed by observing how the ADP solutions of parameter set 3 do considerably better than parameter sets 1 and 2 when compared to their respective heuristic strategies.

So far we have suggested that the quadratic approximation to  $W_{t+h}^\delta$  is responsible when the ADP solution does not do as well as the Sharpe strategy. In addition to the explanations provided above we also confirmed this by examining other possible sources of error. For example when we increased the number of training points we found that there was little or no improvement in the ADP solution. This was also the case when we increased the number

---

<sup>12</sup>The ‘good’ heuristic strategy that we use in this section is the best of the two heuristic strategies that we consider. In every case, this turns out to be the Sharpe strategy.

of points in the low discrepancy sequence that were used for training point evaluation<sup>13</sup>.

Another possible explanation for why the Sharpe strategy sometimes outperforms the ADP strategy is error propagation. This refers to the phenomenon where errors in estimating the value function are propagated back in time during the approximate value iteration.<sup>14</sup> This propagation would have the cumulative effect of magnifying errors so that the estimate of  $V_t(\cdot)$  deteriorates in quality as  $t$  decreases. However, as may be seen from the mean  $R^2$  values that are reported, it is unlikely that error propagation could be entirely responsible. For example, even in Table 3.8 when CRRA = 2 the mean  $R^2$  value was 99.998%. Since there are 20 trading intervals, this might suggest<sup>15</sup> that the total propagated error in any one period is bounded by  $1 - .99998^{20} = .04\%$ . It seems unlikely that this could account for the fact that the Sharpe strategy outperforms the best ADP strategy by approximately \$900.

In general, the  $R^2$  values appear to be very good. It is also worth mentioning that when the  $R^2$  values are not very good then it may be the case that the quadratic approximation to  $W_{t+h}^\delta$  is again to blame. The reason for this is that poorer estimates of  $\theta_t(X_t)$  will result in poorer estimates of  $\hat{J}_t(X_t)$ . This in turn might result in an approximation  $\tilde{J}_t(\cdot)$  that is not very smooth and therefore difficult to approximate with our set of basis functions.

Evidence for this conjecture is provided by the fact that for all three sets of parameters, the mean  $R^2$  values decrease as the CRRA increases. Admittedly, this relationship between  $R^2$  and  $\gamma$  could be explained if  $J(\cdot)$  became less smooth, and therefore 'harder' to approximate as the CRRA increases. In that event we might then expect the cubic approximations to outperform the quadratic approximations when  $\gamma = 10$ . There is no evidence in the data, however, to support this conclusion. It therefore appears to be the case that the quadratic approximation is the dominant source of error in the ADP algorithm.

### 3.5 Conclusions and Further Research

In this chapter we have discussed some of the issues that arise when we use approximate dynamic programming algorithms to solve portfolio optimization problems. We have seen how identifying good heuristic strategies can significantly improve the performance of the quadratic expansion method for finding  $\hat{\theta}_t(X_t)$ . The same heuristic strategy should also prove useful when a quartic expansion is used.

From a qualitative viewpoint, we have shown that some of the properties observed by Brandt et al (2001) do not necessarily generalize to other classes of problems. For example, we saw that using quadratic or cubic basis functions instead of linear basis functions for

---

<sup>13</sup>All of the results in this section were obtained using 2500 training points at each time period and using 1500 low discrepancy points for each training point evaluation.

<sup>14</sup>We use the term 'error propagation' here to refer mainly to errors that are made due to a poor choice of approximation architecture. This appears to be the type of error propagation that concerns Brandt et al (2001) and is the reason for why their algorithm is quadratic in the number of trading periods. Of course, errors due to poor choices of  $\theta_t(X_t)$  will also cause errors to propagate. Such errors, however, can be minimized if good estimates of  $\theta_t(X_t)$  can be found.

<sup>15</sup>Of course this is merely a back of the envelope type calculation and may not be possible to justify rigorously.

value function approximation can often result in a significantly superior solution. We also saw that the performance of the quadratic approximation need not necessarily improve as the coefficient of relative risk aversion increases.

An important conclusion we can therefore draw from this work is that a particular property of an ADP solution may not hold when applied to a different portfolio optimization problem. This merely emphasizes the importance of being able to construct lower and upper bounds on the true value function since we cannot conclude that an algorithm will work well on all portfolio optimization problems simply because it works well on a subset of them. Using approximate dynamic programming solutions to construct upper bounds on the true value function is a subject of ongoing research.

At this point we can also conjecture that an improved ADP algorithm could be found by using the best features of this work and Brandt et al (2001). This algorithm would use the quartic expansion and regression methods of Brandt et al (2001)<sup>16</sup>. However, instead of expanding  $W_{t+h}^\delta$  about  $W_t \exp(r_t h)$ , a good heuristic strategy should be used to determine the particular quartic expansion that is used.

It would also seem wise to use quadratic or cubic polynomials in the state variables as basis functions for approximating the value function. While Brandt et al (2001) found that linear basis functions performed best, they were only marginally superior to quadratic basis functions. On the other hand, in this chapter we found that quadratic and cubic basis functions significantly outperform linear basis functions<sup>17</sup>.

Finally, the algorithm could either use the value iteration procedure of this chapter, or instead use the methodology of Brandt et al (2001) which aims to avoid propagation of errors when estimating the value function. Their methodology, however, is quadratic in the number of trading intervals. As we have pointed out, it is not at all clear that a significant propagation of errors takes place when we carry out approximate value iteration. If this is the case, and computational efficiency is a consideration, then approximate value iteration may well be superior<sup>18</sup>. This is also a possible direction for future research.

---

<sup>16</sup>The regression procedure of Brandt et al (2001) effectively combines the training point evaluation and linear regression into a single step and should be quicker than the method we describe in this chapter. We did not use this method here as our focus at this point is not on computational efficiency.

<sup>17</sup>It should be pointed out that when Brandt et al (2001) compare linear and quadratic basis function, they do so using the quartic expansion to compute  $\hat{\theta}_t(X_t)$ . It remains to be seen if their conclusion would still hold had they instead used a quadratic expansion.

<sup>18</sup>There are a number of possible situations where computational efficiency could be important. For example, a hedge fund using this methodology might wish to be able to react very quickly to sudden changes in the economic environment. As a second example, a financial institution may want to solve its clients' portfolio optimization problems and trade according to the solutions they obtain. In this scenario, the institution would need to solve many portfolio optimization problems and so computational efficiency would again be a factor.

Table 3.1: Heuristic Strategies: Parameter Set 1,  $T = 2$  Years

Table 3.1 contains approximate 95% confidence intervals for the certainty equivalents of the two heuristic trading strategies for each of three different power utility utility functions (CRRA = 2,5,10). The column ‘Constant’ refers to the best constant proportion ( $\theta_c$ ) trading strategy while the column ‘Sharpe’ refers to the best strategy that invests a constant fraction ( $C_s$ ) of the current Sharpe ratio in the risky asset. The second row in each panel contains the constants  $\theta_c$  and  $C_s$ , respectively. The certainty equivalents are computed by simulating each strategy along 4 million sample paths.

The other parameters are as given in parameter set 1 with  $T = 2$  years.

CRRA	Constant	Sharpe
2	[205490, 205584] 1.00	[205690, 205779] 1.23
5	[194896, 194950] 0.44	[195842, 195899] 1.12
10	[189971, 190005] 0.13	[190303, 190339] 0.74

Table 3.2: ADP Strategies: Parameter Set 1,  $T = 2$  Years

Table 3.2 contains approximate 95% confidence intervals for the certainty equivalents of the ADP strategies for each of three different power utility utility functions (CRRA = 2,5,10). There are two classes of ADP strategies. In the first class the optimal fraction of wealth that is invested in the risky asset at time  $t$  is computed using a quadratic expansion of the estimated value function about  $W_t \exp(r_t h)$ . The second class uses a quadratic expansion about  $W_t \exp(R_t^s h)$  where  $\exp(R_t^s h)$  is the expected return from the heuristic Sharpe strategy. Within each class, there are three strategies corresponding to the linear, quadratic and cubic approximations to the value function. The second row of each panel contains the average  $R^2$  value where the average is taken over each of the regressions that are used to estimate the value function from  $t = h$  to  $t = T - h$ . The certainty equivalents are computed by simulating each strategy along 4 million sample paths.

The other parameters are as given in parameter set 1 with  $T = 2$  years.

CRRA	Expansion About $W_t e^{r_t h}$			Expansion About $W_t e^{R_t^s h}$		
	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic
2	[ 204978, 205066] 99.669	[ 205688, 205777] 99.996	[ 205688, 205776] 99.999	[ 205188, 205277] 99.723	[ 205692, 205781] 99.996	[ 205692, 205781] 99.999
5	[ 194711, 194767] 99.267	[ 195692, 195742] 99.971	[ 195809, 195861] 99.998	[ 194981, 195040] 99.293	[ 195838, 195891] 99.970	[ 195872, 195926] 99.998
10	[ 189808, 189845] 97.828	[ 190319, 190351] 99.767	[ 190307, 190341] 99.972	[ 189883, 189920] 97.822	[ 190367, 190400] 99.766	[ 190356, 190390] 99.976

Table 3.3: Heuristic Strategies: Parameter Set 1,  $T = 5$  Years

Table 3.3 contains approximate 95% confidence intervals for the certainty equivalents of the two heuristic trading strategies for each of three different power utility utility functions (CRRA = 2,5,10). The column ‘Constant’ refers to the best constant proportion ( $\theta_c$ ) trading strategy while the column ‘Sharpe’ refers to the best strategy that invests a constant fraction ( $C_s$ ) of the current Sharpe ratio in the risky asset. The second row in each panel contains the constants  $\theta_c$  and  $C_s$ , respectively. The certainty equivalents are computed by simulating each strategy along 4 million sample paths.

The other parameters are as given in parameter set 1 with  $T = 5$  years.

CRRA	Constant	Sharpe
2	[ 323584, 323825] 1.00	[ 325138, 325362] 1.22
5	[ 281322, 281462] 0.45	[ 286614, 286762] 1.14
10	[ 262822, 262914] 0.14	[ 264885, 264988] 0.86

Table 3.4: ADP Strategies: Parameter Set 1,  $T = 5$  Years

Table 3.4 contains approximate 95% confidence intervals for the certainty equivalents of the ADP strategies for each of three different power utility utility functions (CRRA = 2,5,10). There are two classes of ADP strategies. In the first class the optimal fraction of wealth that is invested in the risky asset at time  $t$  is computed using a quadratic expansion of the estimated value function about  $W_t \exp(r_t h)$ . The second class uses a quadratic expansion about  $W_t \exp(R_t^s h)$  where  $\exp(R_t^s h)$  is the expected return from the heuristic Sharpe strategy. Within each class, there are three strategies corresponding to the linear, quadratic and cubic approximations to the value function. The second row of each panel contains the average  $R^2$  value where the average is taken over each of the regressions that are used to estimate the value function from  $t = h$  to  $t = T - h$ . The certainty equivalents are computed by simulating each strategy along 4 million sample paths.

The other parameters are as given in parameter set 1 with  $T = 5$  years.

CRRA	Expansion About $W_t e^{r_t h}$			Expansion About $W_t e^{R_t^s h}$		
	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic
2	[ 315405, 315610] 99.709	[ 325115, 325335] 99.994	[ 325016, 325235] 99.999	[ 316886, 317095] 99.746	[ 325149, 325371] 99.994	[ 325086, 325307] 99.999
5	[ 279611, 279747] 99.399	[ 285536, 285664] 99.952	[ 286179, 286310] 99.991	[ 281319, 281462] 99.402	[ 286527, 286665] 99.950	[ 286586, 286725] 99.991
10	[ 263255, 263358] 98.566	[ 264840, 264933] 99.761	[ 264653, 264748] 99.968	[ 263783, 263889] 98.484	[ 265108, 265206] 99.744	[ 265057, 265155] 99.969

Table 3.5: Heuristic Strategies: Parameter Set 2,  $T = 2$  Years

Table 3.5 contains approximate 95% confidence intervals for the certainty equivalents of the two heuristic trading strategies for each of three different power utility utility functions (CRRA = 2,5,10). The column ‘Constant’ refers to the best constant proportion ( $\theta_c$ ) trading strategy while the column ‘Sharpe’ refers to the best strategy that invests a constant fraction ( $C_s$ ) of the current Sharpe ratio in the risky asset. The second row in each panel contains the constants  $\theta_c$  and  $C_s$ , respectively. The certainty equivalents are computed by simulating each strategy along 4 million sample paths.

The other parameters are as given in parameter set 2 with  $T = 2$  years.

CRRA	Constant	Sharpe
2	[ 209466, 209545] 1.00	[ 209514, 209591] 1.64
5	[ 199960, 200020] 0.73	[ 200878, 200936] 1.83
10	[ 193737, 193780] 0.38	[ 194255, 194297] 1.97

Table 3.6: ADP Strategies: Parameter Set 2,  $T = 2$  Years

Table 3.6 contains approximate 95% confidence intervals for the certainty equivalents of the ADP strategies for each of three different power utility utility functions (CRRA = 2,5,10). There are two classes of ADP strategies. In the first class the optimal fraction of wealth that is invested in the risky asset at time  $t$  is computed using a quadratic expansion of the estimated value function about  $W_t \exp(r_t h)$ . The second class uses a quadratic expansion about  $W_t \exp(R_t^s h)$  where  $\exp(R_t^s h)$  is the expected return from the heuristic Sharpe strategy. Within each class, there are three strategies corresponding to the linear, quadratic and cubic approximations to the value function. The second row of each panel contains the average  $R^2$  value where the average is taken over each of the regressions that are used to estimate the value function from  $t = h$  to  $t = T - h$ . The certainty equivalents are computed by simulating each strategy along 4 million sample paths.

The other parameters are as given in parameter set 2 with  $T = 2$  years.

CRRA	Expansion About $W_t e^{r_t h}$			Expansion About $W_t e^{R_t^s h}$		
	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic
2	[ 208217, 208292] 99.543	[ 209498, 209573] 99.992	[ 209493, 209568] 99.999	[ 208446, 208521] 99.566	[ 209507, 209583] 99.992	[ 209502, 209578] 99.999
5	[ 198305, 198361] 98.259	[ 200432, 200484] 99.887	[ 200497, 200548] 99.990	[ 198630, 198688] 98.274	[ 200649, 200704] 99.886	[ 200721, 200775] 99.990
10	[ 191029, 191080] 95.840	[ 193619, 193658] 99.448	[ 193754, 193791] 99.873	[ 191177, 191229] 95.671	[ 193778, 193819] 99.438	[ 193953, 193992] 99.842



Table 3.7: **Heuristic Strategies: Parameter Set 2,  $T = 5$  Years**

Table 3.7 contains approximate 95% confidence intervals for the certainty equivalents of the two heuristic trading strategies for each of three different power utility utility functions (CRRA = 2,5,10). The column ‘Constant’ refers to the best constant proportion ( $\theta_c$ ) trading strategy while the column ‘Sharpe’ refers to the best strategy that invests a constant fraction ( $C_s$ ) of the current Sharpe ratio in the risky asset. The second row in each panel contains the constants  $\theta_c$  and  $C_s$ , respectively. The certainty equivalents are computed by simulating each strategy along 4 million sample paths.

The other parameters are as given in parameter set 2 with  $T = 2$  years.

CRRA	Constant	Sharpe
2	[ 339161, 339383] 1.00	[ 339634, 339848] 1.68
5	[ 297940, 298111] 0.72	[ 302845, 303008] 1.91
10	[ 273170, 273310] 0.36	[ 275877, 276017] 2.01

Table 3.8: **ADP Strategies: Parameter Set 2,  $T = 5$  Years**

Table 3.8 contains approximate 95% confidence intervals for the certainty equivalents of the ADP strategies for each of three different power utility utility functions (CRRA = 2,5,10). There are two classes of ADP strategies. In the first class the optimal fraction of wealth that is invested in the risky asset at time  $t$  is computed using a quadratic expansion of the estimated value function about  $W_t \exp(r_t h)$ . The second class uses a quadratic expansion about  $W_t \exp(R_t^s h)$  where  $\exp(R_t^s h)$  is the expected return from the heuristic Sharpe strategy. Within each class, there are three strategies corresponding to the linear, quadratic and cubic approximations to the value function. The second row of each panel contains the average  $R^2$  value where the average is taken over each of the regressions that are used to estimate the value function from  $t = h$  to  $t = T - h$ . The certainty equivalents are computed by simulating each strategy along 4 million sample paths.

The other parameters are as given in parameter set 2 with  $T = 5$  years.

CRRA	Expansion About $W_t e^{r_t h}$			Expansion About $W_t e^{R_t^s h}$		
	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic
2	[ 326812, 327013] 99.580	[ 338573, 338787] 99.978	[ 338464, 338677] 99.998	[ 328227, 328430] 99.598	[ 338718, 338933] 99.978	[ 338577, 338790] 99.998
5	[ 288046, 288181] 98.796	[ 295304, 295476] 99.808	[ 298259, 298400] 99.973	[ 289920, 290063] 98.757	[ 296454, 296634] 99.801	[ 299519, 299669] 99.971
10	[ 264233, 264372] 97.392	[ 262743, 262912] 99.235	[ 270234, 270362] 99.807	[ 265400, 265551] 97.174	[ 263536, 263715] 99.159	[ 270807, 270951] 99.782

Table 3.9: Heuristic Strategies: Parameter Set 3,  $T = 2$  Years

Table 3.9 contains approximate 95% confidence intervals for the certainty equivalents of the two heuristic trading strategies for each of three different power utility utility functions (CRRA = 2,5,10). The column ‘Constant’ refers to the best constant proportion ( $\theta_c$ ) trading strategy while the column ‘Sharpe’ refers to the best strategy that invests a constant fraction ( $C_s$ ) of the current Sharpe ratio in the risky asset. The second row in each panel contains the constants  $\theta_c$  and  $C_s$ , respectively. The certainty equivalents are computed by simulating each strategy along 4 million sample paths.

The other parameters are as given in parameter set 3 with  $T = 2$  years.

CRRA	Constant	Sharpe
2	[ 202827, 202937] 1.00	[ 203181, 203283] 1.09
5	[ 191937, 191986] 0.24	[ 192494, 192545] 0.72
10	[ 187842, 187874] 0.00	[ 187842, 187874] 0.00

Table 3.10: ADP Strategies: Parameter Set 3,  $T = 2$  Years

Table 3.10 contains approximate 95% confidence intervals for the certainty equivalents of the ADP strategies for each of three different power utility utility functions (CRRA = 2,5,10). There are two classes of ADP strategies. In the first class the optimal fraction of wealth that is invested in the risky asset at time  $t$  is computed using a quadratic expansion of the estimated value function about  $W_t \exp(r_t h)$ . The second class uses a quadratic expansion about  $W_t \exp(R_t^s h)$  where  $\exp(R_t^s h)$  is the expected return from the heuristic Sharpe strategy. Within each class, there are three strategies corresponding to the linear, quadratic and cubic approximations to the value function. The second row of each panel contains the average  $R^2$  value where the average is taken over each of the regressions that are used to estimate the value function from  $t = h$  to  $t = T - h$ . The certainty equivalents are computed by simulating each strategy along 4 million sample paths.

The other parameters are as given in parameter set 3 with  $T = 2$  years.

CRRA	Expansion About $W_t e^{r_t h}$			Expansion About $W_t e^{R_t^s h}$		
	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic
2	[ 202976, 203078] 99.849	[ 203174, 203273] 99.995	[ 203187, 203288] 99.998	[ 203047, 203148] 99.877	[ 203184, 203285] 99.995	[ 203190, 203291] 99.998
5	[ 191785, 191832] 99.428	[ 192579, 192624] 99.982	[ 192654, 192702] 99.999	[ 191871, 191919] 99.442	[ 192634, 192680] 99.982	[ 192672, 192720] 99.999
10	[ 186719, 186766] 96.551	[ 187944, 187977] 99.629	[ 187907, 187941] 99.944	[ 186719, 186766] 96.551	[ 187944, 187977] 99.629	[ 187907, 187941] 99.944

Table 3.11: Heuristic Strategies: Parameter Set 3,  $T = 5$  Years

Table 3.11 contains approximate 95% confidence intervals for the certainty equivalents of the two heuristic trading strategies for each of three different power utility utility functions (CRRA = 2,5,10). The column ‘Constant’ refers to the best constant proportion ( $\theta_c$ ) trading strategy while the column ‘Sharpe’ refers to the best strategy that invests a constant fraction ( $C_s$ ) of the current Sharpe ratio in the risky asset. The second row in each panel contains the constants  $\theta_c$  and  $C_s$ , respectively. The certainty equivalents are computed by simulating each strategy along 4 million sample paths.

The other parameters are as given in parameter set 3 with  $T = 5$  years.

CRRA	Constant	Sharpe
2	[ 306357, 306656] 1.00	[ 309271, 309534] 1.02
5	[ 264476, 264589] 0.16	[ 266514, 266632] 0.56
10	[ 249238, 249330] 0.00	[ 249238, 249330] 0.00

Table 3.12: ADP Strategies: Parameter Set 3,  $T = 5$  Years

Table 3.12 contains approximate 95% confidence intervals for the certainty equivalents of the ADP strategies for each of three different power utility utility functions (CRRA = 2,5,10). There are two classes of ADP strategies. In the first class the optimal fraction of wealth that is invested in the risky asset at time  $t$  is computed using a quadratic expansion of the estimated value function about  $W_t \exp(r_t h)$ . The second class uses a quadratic expansion about  $W_t \exp(R_t^s h)$  where  $\exp(R_t^s h)$  is the expected return from the heuristic Sharpe strategy. Within each class, there are three strategies corresponding to the linear, quadratic and cubic approximations to the value function. The second row of each panel contains the average  $R^2$  value where the average is taken over each of the regressions that are used to estimate the value function from  $t = h$  to  $t = T - h$ . The certainty equivalents are computed by simulating each strategy along 4 million sample paths.

The other parameters are as given in parameter set 3 with  $T = 5$  years.

CRRA	Expansion About $W_t e^{r_t h}$			Expansion About $W_t e^{R_t^s h}$		
	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic
2	[ 303555, 303788] 99.792	[ 309216, 309468] 99.995	[ 309315, 309572] 99.998	[ 304564, 304801] 99.814	[ 309321, 309576] 99.995	[ 309352, 309611] 99.998
5	[ 263475, 263616] 99.396	[ 267233, 267338] 99.965	[ 267206, 267317] 99.995	[ 263740, 263883] 99.387	[ 267430, 267536] 99.964	[ 267320, 267432] 99.996
10	[ 246640, 246831] 97.271	[ 248434, 248544] 99.556	[ 249093, 249188] 99.914	[ 246640, 246831] 97.271	[ 248434, 248544] 99.556	[ 249093, 249188] 99.914

# Appendix A

## Solution of HJB Equations for Ornstein-Uhlenbeck Processes

In this Appendix, we derive the optimal value function  $J(\cdot)$  from the Hamilton-Jacobi-Bellman equation (1.4.10) for the trending and standard Ornstein-Uhlenbeck processes.

### A.1 Trending Ornstein-Uhlenbeck Value Function

We present here the solution to the Hamilton-Jacobi-Bellman equation (1.4.10) of Section 1.4.2. Recall that this equation is given by:

$$0 = \text{Max}_{\omega_t} \left\{ J_t + W_t J_W \left( r + \omega_t [-\delta(X_t - \mu t - X_0) + \mu + \frac{\sigma^2}{2} - r] \right) + J_X (-\delta(X_t - \mu t - X_0) + \mu) + \frac{1}{2} \omega_t^2 \sigma^2 W_t^2 J_{WW} + \frac{1}{2} \sigma^2 J_{XX} + \sigma^2 \omega_t W_t J_{XW} \right\} .(A.1)$$

Solving for  $\omega_t$  and substituting back into A.1 yields the following partial differential equation (PDE):

$$0 = J_t J_{WW} + [-\delta(X_t - \mu t - X_0) + \mu] J_X J_{WW} + r W J_W J_{WW} + \frac{\sigma^2}{2} J_{XX} J_{WW} - \frac{\sigma^2}{2} J_{XW}^2 - \left( -\delta(X_t - \mu t - X_0) + \mu + \frac{\sigma^2}{2} - r \right) J_W J_{XW} -$$

$$\frac{\left(-\delta(X_t - \mu t - X_0) + \mu + \frac{\sigma^2}{2} - r\right)^2}{2\sigma^2} J_W^2 \quad (\text{A.2})$$

subject to  $J(W, X, T) = U(W)$ . We solve this PDE by conjecturing that

$$J(W, X, t) = U(W \exp[r(T-t)]) \exp(\alpha(t) + \beta(t)X + \zeta(t)X^2)$$

where  $\alpha(T) = \beta(T) = \zeta(T) = 0$ . Therefore solving the PDE reduces to solving three ordinary differential equations. We then solve these differential equations for  $\alpha(t)$ ,  $\beta(t)$  and  $\zeta(t)$ .

## CRRA Utility

For an investor with the CRRA utility function,  $U(W) = W^\gamma/\gamma$ , it is only possible to solve explicitly for  $\beta(t)$  and  $\zeta(t)$ . Solving for  $\alpha(t)$  required evaluating a number of definite integrals for which there did not seem to be analytic solutions. These integrals are easy to solve numerically, however, and it is therefore possible to find a very good numerical solution to the value function,  $J(W, X, t)$ . We present here the solutions for  $\beta(t)$  and  $\zeta(t)$ . Let

$$\begin{aligned} a &= 1 + \sqrt{1-\gamma} \quad , \quad b = 1 - \sqrt{1-\gamma} \quad , \quad q = \frac{-2\delta}{\sqrt{1-\gamma}} \quad , \quad H = \frac{-\gamma\delta^2\mu}{\sigma^2\sqrt{1-\gamma}} \\ I &= \frac{(2r-\sigma^2)\gamma\delta}{2\sigma^2} \quad , \quad J = \frac{\gamma\delta(2r-2\mu-\sigma^2)}{2\sigma^2\sqrt{1-\gamma}} \quad , \quad K = \frac{-\gamma\delta\mu}{\sigma^2} \end{aligned} \quad (\text{A.3})$$

Then  $\beta(t)$  and  $\zeta(t)$  are given by:

$$\beta(t) = \frac{\sqrt{1-\gamma}}{\delta(a - b \exp[q(T-t)])} \left[ (Ht + K + I + J) - (Ht - K + J - I) \exp(q(T-t)) - 2(K + I) \exp\left(\frac{q(T-t)}{2}\right) \right] \quad (\text{A.4})$$

$$\zeta(t) = \frac{\gamma\delta}{2\sigma^2} \left[ \frac{1 - \exp(q(T-t))}{a - b \exp(q(T-t))} \right] \quad (\text{A.5})$$

## CARA Utility

For the CARA utility function,  $U(W) = -\exp(-\gamma W)/\gamma$ ,  $\alpha(t)$ ,  $\beta(t)$ ,  $\zeta(t)$  are:

$$\alpha(t) = \frac{\Gamma_1(t^3 - T^3)}{3} + \frac{\Gamma_2(t^2 - T^2)}{2} + \Gamma_3(t - T) \quad (\text{A.6})$$

$$\beta(t) = \Delta_1(t^2 - T^2) + \Delta_2(T - t) \quad (\text{A.7})$$

$$\zeta(t) = \frac{\delta^2 (t - T)}{2\sigma^2} \quad (\text{A.8})$$

where

$$\Delta_1 = \frac{\delta^2}{4} - \frac{r\delta^2}{2\sigma^2} - \frac{\delta^2\mu}{2\sigma^2} \quad (\text{A.9})$$

$$\Delta_2 = \frac{\delta}{2\sigma^2} (\delta\sigma^2 T - 2r\delta T + 2\mu + \sigma^2 - 2r) \quad (\text{A.10})$$

$$\Gamma_1 = \frac{\delta^2\mu^2}{2\sigma^2} + \left(\frac{\sigma^2}{2} - r\right) \Delta_1 \quad (\text{A.11})$$

$$\Gamma_2 = \frac{\delta\mu}{\sigma^2} \left(\mu - r + \frac{\sigma^2}{2}\right) - \left(\frac{\sigma^2}{2} - r\right) \Delta_2 - \frac{\delta^2}{2} \quad (\text{A.12})$$

$$\Gamma_3 = \frac{\left(\mu - r + \frac{\sigma^2}{2}\right)^2}{2\sigma^2} + \Delta_1 T^2 \left(r - \frac{\sigma^2}{2}\right) - \Delta_2 T \left(r - \frac{\sigma^2}{2}\right) + \frac{\delta^2 T}{2} \quad (\text{A.13})$$

## A.2 Non-Trending Ornstein-Uhlenbeck Value Function

Recall that  $X_t \equiv \log P_t$  and let  $X_t$  satisfy the following stochastic differential equation:

$$dX_t = -\delta (X_t - \alpha) dt + \sigma dB_t \quad (\text{A.14})$$

where  $\alpha$  and  $\delta$  are both positive. The solution to (A.14) is given by:

$$X_t = \alpha + \exp(-\delta t) [X_0 - \alpha] + \sigma \exp(-\delta t) \int_0^t \exp(\delta s) dB_s \quad (\text{A.15})$$

and the corresponding Hamilton-Jacobi-Bellman equation is given by

$$\begin{aligned} 0 = \text{Max}_\omega & \left[ J_t + W J_W \left( (1 - \omega)r - \delta\omega (\log P - \alpha) + \frac{\omega\sigma^2}{2} \right) + \right. \\ & P J_P \left( \frac{\sigma^2}{2} - \gamma (\log P - \alpha) \right) + \frac{1}{2} W^2 \sigma^2 J_{WW} \omega^2 + \\ & \left. \frac{1}{2} P^2 \sigma^2 J_{PP} + W P \sigma^2 \omega J_{PW} \right]. \end{aligned} \quad (\text{A.16})$$

We solve this PDE by conjecturing that the value function is of the form:

$$J(W, X, t) = U(W \exp[r(T-t)]) \exp(\alpha(t) + \beta(t)X + \zeta(t)X^2) \quad (\text{A.17})$$

where  $\alpha(T) = \beta(T) = \zeta(T) = 0$ .

### CRRA Utility

For  $U(W) = W^\gamma/\gamma$ , we have the following system of ODE's:

$$\frac{d\alpha}{dt} = \frac{\gamma}{\gamma-1} \left[ \frac{\sigma^2}{2\delta} \beta^2 + \frac{\Delta_3}{2} \beta + \frac{\Delta_2}{2\sigma^2} - \sigma^2 \zeta \right] \quad (\text{A.18})$$

$$\frac{d\beta}{dt} = (2\sigma^2 \zeta - \delta) \frac{\beta}{\gamma-1} + \frac{\gamma}{\gamma-1} \left[ \Delta_3 \zeta + \frac{\Delta_1}{2\sigma^2} \right] \quad (\text{A.19})$$

$$\frac{d\zeta}{dt} = \frac{2\sigma^2}{\gamma-1} \zeta^2 - \frac{2\delta}{\gamma-1} \zeta + \frac{\delta^2 \gamma}{2\sigma^2 (\gamma-1)} \quad (\text{A.20})$$

where

$$\Delta_1 = 2r\delta - \sigma^2\delta - 2\alpha\delta^2 \quad (\text{A.21})$$

$$\Delta_2 = \sigma^4/4 + \alpha^2\delta^2 + r^2 - r\sigma^2 + \sigma^2\delta\alpha - 2r\delta\alpha \quad (\text{A.22})$$

$$\Delta_3 = \sigma^2 - 2r + \frac{2\alpha\delta}{\gamma} \quad (\text{A.23})$$

Then

$$\beta(t) = \frac{2(d_1 + d_2) - d_1(e^{s(T-t)} + e^{-s(T-t)}) - d_2(ae^{s(T-t)} + be^{-s(T-t)})}{s(ae^{s(T-t)} - be^{-s(T-t)})} \quad (\text{A.24})$$

$$\zeta(t) = \frac{\delta\gamma}{2\sigma^2} \left[ \frac{1 - e^{q(T-t)}}{a - be^{q(T-t)}} \right] \quad (\text{A.25})$$

where

$$\begin{aligned} d_1 &= \frac{\gamma^2 \Delta_3 \delta}{2(\gamma-1)\sigma^2} & , & & d_2 &= \frac{\gamma \Delta_1}{2(\gamma-1)\sigma^2} & , & & a &= 1 + \sqrt{1-\gamma} \\ b &= 1 - \sqrt{1-\gamma} & , & & s &= \frac{\delta(a-\gamma)}{(1-\gamma)a} & , & & q &= \frac{-2\delta}{\sqrt{1-\gamma}} \end{aligned} \quad (\text{A.26})$$

To define  $\alpha(t)$ , let:

$$\begin{aligned} f_1 &= \frac{\sigma^2}{2(\gamma-1)} \quad , \quad f_2 = \frac{\gamma\Delta_3}{2(\gamma-1)} \quad , \quad f_3 = \frac{\gamma\Delta_2}{2(\gamma-1)\sigma^2} \\ \rho_1 &= \frac{2(d_1+d_2)}{s} \quad , \quad \rho_2 = \frac{-(d_1+ad_2)}{s} \quad , \quad \rho_3 = \frac{-(d_1+bd_2)}{s} \end{aligned} \quad (\text{A.27})$$

and

$$I_1 = [2as (ae^{2s(T-t)} - b)]^{-1} \quad (\text{A.28})$$

$$I_2 = \frac{-1}{2s} \left[ \frac{1}{a^2} \log \left( \frac{a - be^{-2s(T-t)}}{be^{-2s(T-t)}} \right) - \frac{1}{a(a - be^{-2s(T-t)})} \right] \quad (\text{A.29})$$

$$I_3 = \frac{-1}{2s} \left[ \frac{1}{b^2} \log \left( \frac{ae^{2s(T-t)}}{ae^{2s(T-t)} - b} \right) - \frac{1}{b(ae^{2s(T-t)} - b)} \right] \quad (\text{A.30})$$

$$I_4 = \frac{1}{2s} \left[ \frac{e^{s(T-t)}}{b(ae^{2s(T-t)} - b)} + \frac{1}{b\sqrt{-ab}} \tan^{-1} \left( \sqrt{\frac{-a}{b}} e^{s(T-t)} \right) \right] \quad (\text{A.31})$$

$$I_5 = \frac{1}{2s} \left[ \frac{e^{s(T-t)}}{a(ae^{2s(T-t)} - b)} - \frac{1}{a\sqrt{-ab}} \tan^{-1} \left( \sqrt{\frac{-a}{b}} e^{s(T-t)} \right) \right] \quad (\text{A.32})$$

$$I_7 = \frac{-1}{s\sqrt{-ab}} \tan^{-1} \left( \sqrt{\frac{-a}{b}} e^{s(T-t)} \right) \quad (\text{A.33})$$

$$I_6 = I_1 \quad (\text{A.34})$$

$$I_8 = \frac{-1}{2as} \log (ae^{2s(T-t)} - b) \quad (\text{A.35})$$

$$I_9 = \frac{-1}{2bs} \log (a - be^{-2s(T-t)}) \quad (\text{A.36})$$

$$I_{10} = \frac{\delta\gamma}{2\sigma^2} \left[ \frac{1}{aq} \log (ae^{-q(T-t)} - b) - \frac{1}{bq} \log (a - be^{q(T-t)}) \right] \quad (\text{A.37})$$

Then

$$\begin{aligned} \alpha(t) &= f_1 [\rho_1^2 I_1 + \rho_2^2 I_2 + \rho_3^2 I_3 + 2\rho_1\rho_2 I_4 + 2\rho_1\rho_3 I_5 + 2\rho_2\rho_3 I_6] + \\ &\quad f_2 [\rho_1 I_7 + \rho_2 I_8 + \rho_3 I_9] - \sigma^2 I_{10} + f_3 t + G \end{aligned} \quad (\text{A.38})$$

where  $G$  is the constant defined by the condition  $\alpha(T) = 0$ .



## CARA Utility

The solution of (A.16) for CARA utility,  $U(W) = -\exp(-\gamma W)/\gamma$  is given by:

$$\alpha(t) = \frac{\delta^2}{6\sigma^2} \left( \frac{\sigma^2}{2} - r \right)^2 (t-T)^3 + \left[ \frac{\Delta_1}{4\sigma^2} \left( \frac{\sigma^2}{2} - r \right) - \frac{\delta^2}{4} \right] (T-t)^2 - \frac{\Delta_2}{2\sigma^2} (T-t) \quad (\text{A.39})$$

$$\beta(t) = \frac{\delta^2}{2} \left( \frac{1}{2} - \frac{r}{\sigma^2} \right) (t-T)^2 - \frac{\Delta_1}{2\sigma^2} (T-t) \quad (\text{A.40})$$

$$\zeta(t) = \frac{\delta^2}{2\sigma^2} (t-T) \quad (\text{A.41})$$

$$(\text{A.42})$$

where the solution is of the form:

$$J(W, X, t) = U(W \exp[r(T-t)]) \exp(\alpha(t) + \beta(t)X + \zeta(t)X^2). \quad (\text{A.43})$$

# Appendix B

## Approximate Q-Value Iteration

In Section 2.3.2 we described the basic algorithm for carrying out the approximate Q-value iteration. In this section we discuss some aspects of the algorithm in more detail as well as describing some important improvements to the algorithm that we make. We begin by discussing a tool from numerical analysis, low discrepancy sequences, and then proceed to describe how such sequences can be used for both training point selection and evaluation. Then, after describing the method by which we train the neural network, we proceed to describe policy fixing and feature extraction, both of which can significantly improve the algorithm's performance.

### B.1 Low Discrepancy Sequences

A low discrepancy sequence is a deterministic sequence of points that is evenly dispersed in some fixed domain. Often, and without loss of generality, we take this domain to be the unit cube  $[0, 1]^d$ . Because the points in a low discrepancy sequence are evenly dispersed, they are often used to numerically integrate some function  $f(\cdot)$  over  $[0, 1]^d$  so that

$$\int_{[0,1]^d} f(x) dx \approx \frac{\sum_{i=1}^N f(y_i)}{N} \quad (\text{B1})$$

where  $\{y_i : i = 1, \dots, N\}$ , is a set of  $N$  consecutive terms from the low discrepancy sequence. An important property that low discrepancy sequences possess is that as new terms are added, the sequence remains evenly dispersed. This property implies that, in contrast to other numerical integration schemes, the term  $N$  in (B1) need not be determined in advance and can therefore be chosen according to some termination criterion. Because these sequences are evenly dispersed, their use in numerical integration often results in a rate of convergence that is much faster than Monte Carlo simulation where the convergence rate is  $O(\frac{1}{\sqrt{N}})$ . For the technical definition of a low discrepancy sequence and a more detailed introduction to

their properties and financial applications, see Boyle, Broadie and Glasserman (1997). See Birge (1994), Joy, Boyle and Tan (1996) and Paskov and Traub (1995) for some of these applications.

In this paper we use low discrepancy sequences for training point selection and training point evaluation. The low discrepancy sequences are of particular value for training point evaluation since a good estimate of

$$\mathbb{E}_t \left[ \frac{B_t}{B_{t+1}} \max \left( h(X_{t+1}), \tilde{Q}_{t+1}(X_{t+1}) \right) \right] \quad (\text{B2})$$

can usually be computed much faster by using a low discrepancy sequence in place of Monte Carlo simulation.

Even though a low discrepancy point  $y \in [0, 1]^d$  is deterministic it can be useful to interpret it as being sampled from a uniform distribution in  $[0, 1]^d$ . With this in mind, it is then straightforward to convert  $y$  into a point,  $x$ , that is representative of any random variable  $X$  with cumulative distribution function  $F(\cdot)$ . For example, suppose  $X$  is a  $d$ -dimensional standard normal random variable with correlation matrix equal to the identity. We can then construct a point  $x \in \mathbb{R}^d$  that is representative of  $X$  by setting

$$x = F^{-1}(y) \quad (\text{B3})$$

where the operation  $F^{-1}$  is taken componentwise in (B3). In finance applications, random variables are often lognormally distributed, but since transforming a normal random variable into a lognormal random variable is easy, we can do likewise with  $x$ . Therefore, we can easily convert a  $d$ -dimensional low discrepancy sequence into a sequence of points that is representative of some fixed  $d$ -dimensional probability distribution. Hereafter, we will assume whenever a low discrepancy sequence is used, that the points in the sequence have already been transformed so that they represent some fixed probability distribution. This probability distribution will always be a lognormal distribution whose parameters should be clear from the context.

## B.2 Training Point Selection and Evaluation

In the approximate Q-value iteration algorithm of Section 2.3.2, we need to define a set of training points

$$S_t := \{P_1^t, \dots, P_{N_t}^t\} \text{ for } t = 1, \dots, K - 1 \quad (\text{B4})$$

that we use to train the neural networks. In this section we describe how these sets may be selected. As we use  $\{\hat{Q}_t(P_1^t), \dots, \hat{Q}_t(P_{N_t}^t)\}$  to train the time  $t$  neural network, it makes

sense that we should choose  $S_t$  so that it is in some way representative of the distribution of  $X_t$ . With this in mind, the obvious solution is to simulate points from the distribution of  $X_t$ . This is in the spirit of Longstaff and Schwartz (2000) and Tsitsiklis and Van Roy (2000) who use simulated trajectories to estimate Q-value functions. We could also use simulated trajectories to determine  $S_t$ , but in so doing, we would be giving up the flexibility of allowing the number of training points to vary with the exercise period. For this reason, it might be preferable instead to select  $S_t$  by simply simulating from the distribution of  $X_t$  so that  $S_t$  is independent of  $S_r$  for  $r \neq t$ .

Another method for choosing the training points is to use a low discrepancy sequence. Having chosen the number of training points,  $N$  say, we simply take  $N$  terms from a low discrepancy sequence and use these as our training points. Our limited experience shows that both simulation and low discrepancy sequences work very well in practice. The performance of the low discrepancy sequences, however, appeared to be marginally superior when applied to the problems we consider in this paper.

For each training point,  $P_i^t$ , we need to compute

$$\hat{Q}_t(P_i^t) := \hat{E}_t \left[ \frac{B_t}{B_{t+1}} \max \left( h(X_{t+1}), \tilde{Q}_{t+1}(X_{t+1}) \right) \right] \quad (\text{B5})$$

where, as mentioned earlier, the  $\hat{E}[\cdot]$  operator is intended to approximate the expectation operator,  $E[\cdot]$ . An obvious way of defining  $\hat{E}[\cdot]$  is to set

$$\hat{E}_t \left[ \frac{B_t}{B_{t+1}} \max \left( h(X_{t+1}), \tilde{Q}_{t+1}(X_{t+1}) \right) \right] = \frac{B_t}{NB_{t+1}} \sum_{i=1}^N \max \left( h(x_i), \tilde{Q}_{t+1}(x_i) \right) \quad (\text{B6})$$

where the  $x_i$ 's are drawn randomly from the conditional distribution of  $X_{t+1}$ . The problem with this method is that the rate of convergence to the true expectation is  $O(\frac{1}{\sqrt{N}})$  which can be too slow for our purposes. Instead, we use a low discrepancy sequence to generate the  $x_i$ 's. We do not choose  $N$  in advance but instead use the following termination criterion adapted from Paskov and Traub (1995).

Let  $E(N)$  denote the estimate in (B6) when  $N$  low discrepancy points are used. We then examine  $E(1000i)$  for  $i = 1, \dots, M$  and terminate either when

$$|E(1000(i+1)) - E(1000i)| < \epsilon \quad (\text{B7})$$

or when  $i = M$ , if the condition in (B7) is not satisfied for any  $i < M$ . In the problems we consider,  $\epsilon$  ranges from 5 to 10 cents and  $M$  is set equal to 100. We find that using low discrepancy sequences instead of Monte Carlo simulation for training point evaluation significantly improves the performance of the algorithm. In the results of Section 2.4 we use a particular class of low discrepancy sequences, namely *Sobol* sequences, for both training point selection and evaluation. This means that our algorithm for estimating the Q-value

functions is in fact deterministic. Monte Carlo simulation is only employed when we use the estimated Q-value functions to obtain upper and lower bounds on the true price of the option.

### B.3 Training the Neural Network

In the approximate Q-value iteration algorithm of Section 2.3.2, we approximate the Q-value function,  $Q_t$ , with a multilayer perceptron,  $\tilde{Q}_t$ , whose functional form is given by

$$\tilde{Q}_t(x; \beta_t) = \sum_{j=1}^N r(j) \sigma \left( b(j) + \sum_{l=1}^d (r(j, l) x(l)) \right) \quad (\text{B8})$$

where the  $b(j)$ 's,  $r(j)$ 's and  $r(j, l)$ 's constitute the parameter vector  $\beta_t$ , and  $\sigma(\cdot)$  is the logistic function so that

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \quad (\text{B9})$$

In attempting to find the  $\hat{\beta}_t$  that minimizes the sum of squares in (2.3.23) it is important to avoid overfitting the data. Two common methods for addressing this issue are *regularization* and *cross validation*. When regularization is used, the objective function to be minimized is given by

$$\gamma \sum_{i=1}^{N_t} \left( \hat{Q}_t(P_i^t) - \tilde{Q}_t(P_i^t; \beta) \right)^2 + (1 - \gamma) f(\beta_t) \quad (\text{B10})$$

where  $0 < \gamma < 1$ , and  $f(\beta_t)$  is a function that increases with the magnitude of  $\beta_t$ . The intuition behind (B10) is that by penalizing large values of  $\beta_t$ , the network is penalized for overfitting. This approach can be justified theoretically but in practice it is difficult to determine the weight  $\gamma$  that should be applied in (B10).

The approach that we use in this paper is cross validation. This approach requires the training points to be divided into three sets, namely training, validation and test sets. Initially, only the training and validation sets are used in the minimization so that the quantity

$$\sum \left( \hat{Q}_t(P_i^t) - \tilde{Q}_t(P_i^t; \beta_t) \right)^2 \quad (\text{B11})$$

is minimized where the sum in (B11) is taken over points in the training set. The minimization is performed using the *Levenberg Marquardt* method for least squares optimization (see Bertsekas and Tsitsiklis 1996). At each iteration of the minimization, the error in the validation set is also computed and as long as overfitting is not taking place, then the validation error should decrease along with the training set error. However, if the validation error starts increasing at any point then it is likely that overfitting is taking place. The algorithm then terminates if the validation error increases for a prespecified number of iterations, and  $\hat{\beta}_t$  is then set equal to the value of  $\beta_t$  in the last iteration of the minimization before the validation error began to increase.

There is one further difficulty with the neural network architecture that needs to be addressed. The neural network will typically have many local minima and it is often the case that the algorithm will terminate at a local minimum that is far from the global minimum. In this case, it is necessary to repeat the minimization again, this time using a different starting value for  $\beta_t$ . This may be repeated until a satisfactory local minimum has been found. Sometimes a number of good local minima are available and when this occurs, the test set can be used to determine the optimal  $\hat{\beta}_t$ . The test set may also be used for other diagnostic and model checking purposes.

We use this training algorithm for finding  $\tilde{Q}_{K-1}$ . For the remaining Q-value functions, however, the problem is now somewhat simplified since it is usually the case that  $\tilde{Q}_t \approx \tilde{Q}_{t-1}$ . We can therefore use  $\hat{\beta}_t$  as the initial solution for training the time  $t - 1$  neural network. In practice, this means that the other neural networks can be trained very quickly and that we only need to perform the minimization once. It also means that we can dispense with the need for having a test set for all but the terminal neural network.

## B.4 Policy Fixing and Feature Extraction

We can improve the Q-value approximation algorithm considerably by exploiting any useful problem specific information that we may have. In this section we describe *policy fixing*, a term defined by Broadie and Glasserman (1997b), and *feature extraction*, a widely used technique in the field of artificial intelligence.

The idea underlying policy fixing is very simple: if we know that the price of the American option is bounded by some quantity, then this information should be used throughout the algorithm. For example, the value of the American option is always greater than or equal to the value of the corresponding European option. It is often the case that the value of this European option can be computed very quickly, and when this is the case, it makes sense to use this information. It is also possible that the value of the American option may be bounded by the value of a different, but related, European or American option. For example, consider the price of an American option that, upon exercise, pays  $\max(0, \hat{S} - k)$  where  $\hat{S}$  is the maximum of  $N$  underlying securities and  $k$  is the strike price. Then the value of this option at time  $t$  is bounded below by the value of an American option whose payoff function is  $\max(0, S_{Max} - k)$ , where  $S_{Max}$  is the stock that had the highest value at time  $t$ . In order to exploit this kind of information it is necessary to trade off the prospect of better bounds with the computational time that is required to compute these bounds. Choosing a good tradeoff is more of an art than a science but a good understanding of the problem should

usually lead to a good choice.

Once a bound has been chosen, we need to incorporate it into the algorithm. This is easily achieved. Suppose for example that  $Q_t \geq b_t \forall t$ . Then we simply redefine  $\hat{Q}_t$  so that now

$$\hat{Q}_t(X_t) = \max \left( b_t(X_t), E_t \left[ \frac{B_t}{B_{t+1}} \max \left( h(X_{t+1}), \tilde{Q}_{t+1}(X_{t+1}) \right) \right] \right). \quad (\text{B12})$$

Similarly,  $\tilde{Q}_{t+1}(X_{t+1})$  is no longer defined to be equal to the value of the time  $t + 1$  neural network. Instead, it is set equal to the maximum of the value of the time  $t + 1$  neural network and  $b_{t+1}(X_{t+1})$ .

Another technique that can significantly improve performance is feature extraction. Feature extraction is the process of taking a function,  $f(\cdot)$  say, of  $X_t$  and using  $f(X_t)$  as an additional input to the time  $t$  neural network. Of course, we can use multiple functions if necessary. Because of the universal approximation property of the neural network architecture, the idea that feature extraction can be useful seems counterintuitive. After all, from an informational point of view, we are providing no new information to the neural network. However, the universal approximation property depends on the neural network,  $\tilde{Q}_t$ , containing an arbitrarily large number of neurons,  $N$  say, in order for it to be a sufficiently good approximation to  $Q_t$ . If we use good features for the neural network, however, the number of neurons that are required to attain the same degree of approximation will typically be considerably less than  $N$ . In practice, we find that adding good features to the neural network can often result in a considerably more accurate estimate of the option price. For example, when  $X_t$  represents the prices of  $d$  underlying stocks, it is often a good idea to order the stock prices before using them as inputs to the neural network. Other features that often prove useful are the same functions that are used for policy fixing. These include European option prices, the intrinsic value of the option, and sometimes the values of related, but low dimensional, American options.

# Appendix C

## Linear Programming and Dynamic Programming Bounds

In this appendix we derive the upper bound of Chapter 2 using the linear programming formulation of dynamic programming problems. We also explore whether or not this bound for optimal stopping problems might be generalized to finite horizon dynamic programming problems. First, we introduce some notation.

Suppose that there are  $T + 1$  dates,  $\{0, 1, \dots, T\}$ , and without loss of generality, that there are  $m$  possible states at each date. We also assume that there is a finite set of controls available,  $U = \{u_1, \dots, u_l\}$ .

Let  $g^t(i, u, j)$  be the immediate payoff if action  $u$  is taken in state  $i$ , date  $t$ , and the successor state is state  $j$ . Similarly, let  $p^t(i, u, j)$  be the probability that the successor state is state  $j$  given action  $u$  is taken at date  $t$ , in state  $i$ .

Denote by  $J_i^t$  the value function in state  $i$ , time  $t$ . We also introduce a terminal state at date  $T + 1$  so that  $J^{T+1} := 0$ .

Suppose the dynamic program (DP) we are trying to solve is a maximization problem so that the Bellman equation is

$$J_i^t = \max_u \sum_{k=1}^m p^t(i, u, k)(g^t(i, u, k) + J_k^{t+1}), \quad t = 0, \dots, T.$$

It is well known that the DP can be expressed as the following linear program (LP):

$$\min \sum_{t=0}^T \sum_{i=1}^m w_i^t J_i^t \tag{C1}$$



subject to

$$\begin{aligned} J_i^t &\geq \sum_{k=1}^m p^t(i, u, k)(g^t(i, u, k) + J_k^{t+1}) \quad \forall u, i, t = 0, \dots, T \\ J^{T+1} &= 0, \end{aligned}$$

where  $\{w_i^t\}$  is any set of positive weights.<sup>1</sup> Suppose now that an estimate of the value function  $\{\tilde{J}_i^t\}$  is available. If we can construct  $\{V_i^t\}$  from  $\{\tilde{J}_i^t\}$  so that  $\{V_i^t\}$  is feasible for the LP then  $V^0$  constitutes an upper bound for  $J^0$ .<sup>2</sup>

## C.1 Optimal Stopping

In the special case of optimal stopping problems, there are only two possible controls, *stop* and *continue*, so the LP is

$$\min \sum_{t=0}^T \sum_{i=1}^m J_i^t$$

subject to

$$J_i^t \geq g^t(i) \quad \forall i, t = 1, \dots, T \tag{C2}$$

$$J_i^t \geq \sum_{k=1}^m p^t(i, k) J_k^{t+1} \quad \forall i, t = 1, \dots, T. \tag{C3}$$

Note that these constraints simply say that the value function is the smallest supermartingale that dominates the payoff function,  $g$ . Now suppose we have  $\{\tilde{J}_i^t\}$  available that by construction satisfy (C2).<sup>3</sup> We then define

$$\begin{aligned} V^{T+1} &= 0 \\ V^T &= \tilde{J}^T \\ &\vdots \\ V^{t-1} &= \mathbf{E}_{t-1}[V^t] - \mathbf{E}_{t-1}[\tilde{J}^t - \tilde{J}^{t-1}] + (\mathbf{E}_{t-1}[\tilde{J}^t - \tilde{J}^{t-1}])^+ \end{aligned}$$

<sup>1</sup>See, for example, Bertsekas and Tsitsiklis (1996).

<sup>2</sup>This is easily seen by examining the LP formulation and considering different sets of  $\{w_i^t\}$ .

<sup>3</sup>Given any approximation,  $\{\tilde{J}_i^t\}$ , this condition is easily enforced.

By construction,  $V$  satisfies (C3). It also satisfies (C2) since  $V^t \geq \tilde{J}^t$  for all  $t^4$  and since we assumed that  $\tilde{J}^i$  satisfied (C2). Therefore  $V^0$  is an upper bound and it is easy to obtain

**Theorem 4** *An upper bound for the optimal stopping problem is given by*

$$V^0 = \tilde{J}^0 + E_0\left[\sum_{t=1}^T (E_{t-1}[\tilde{J}^t - \tilde{J}^{t-1}])^+\right].$$

This is precisely the same bound that we obtained in Chapter 2 using the Optional Sampling Theorem for Supermartingales. The general linear programming formulation also suggests the possibility that this bound could be generalized to finite horizon dynamic programming problems. From a financial engineering point of view, this might be of value since many portfolio optimization problems, for example, can be formulated as finite horizon dynamic programming problems.

## C.2 Finite Horizon Dynamic Programming

Consider the portfolio optimization problem where there is a finite horizon,  $T$ , and where lifetime utility is derived from intertemporal consumption and terminal wealth. We let  $g^t$  denote the intertemporal consumption at time  $t$  and let  $E_t^{s_t}$  denote the conditional expected value given time  $t$  information and assuming that strategy  $s_t$  is used at date  $t$ .

The LP then takes the form

$$\min \sum_{t=0}^T \sum_{i=1}^m J_i^t$$

subject to

$$\begin{aligned} J_i^t &\geq \sum_{k=1}^m p^t(i, u, k) [g^t(i, u) + J_k^{t+1}] \quad \forall u, i, t = 0, \dots, T \\ J^{T+1} &= 0. \end{aligned} \tag{C4}$$

Assuming that we have an approximate solution,  $\tilde{J}$ , available that satisfies  $\tilde{J}_i^T = g^T(i, u) =$

---

<sup>4</sup>This may be seen using induction starting at  $t = T$ .

$g^T(i) = J_i^T$  for all  $u$ , we proceed to construct a feasible solution,  $V$ , to (C4). We define

$$\begin{aligned} V^{T+1} &= 0 \\ V^T &= \tilde{J}^T \\ &\vdots \\ V^{t-1} &= E_{t-1}^{bs_{t-1}}[g^{t-1} + V^t] - E_{t-1}^{bs_{t-1}}[g^{t-1} + \tilde{J}^t - \tilde{J}^{t-1}] + \max_u(E_{t-1}^u[g^{t-1} + \tilde{J}^t - \tilde{J}^{t-1}])^+. \end{aligned} \quad (C5)$$

where the action  $bs_{t-1}$  at date  $t-1$  is the best action with respect to  $\{p^{t-1}(i, u, k)\}$  and  $\{V^t\}$ . By construction,  $V$  satisfies (C4) so that  $V^0$  is an upper bound for the true value function.<sup>5</sup> Note that if  $\tilde{J}^t = J^t$ , the true value function, then  $V^t = J^t$  and the upper bound is in fact tight. The expression for  $V^0$  can easily be computed to give

**Theorem 5** *An upper bound for the dynamic program is given by*

$$V^0 = \tilde{J}^0 + E_0^{bs} \left[ \sum_{t=1}^T \max_u (E_{t-1}[g^{t-1} + \tilde{J}^t - \tilde{J}^{t-1}])^+ \right].$$

where the strategy  $bs$  denotes the strategy that takes action  $bs_t$  at each date  $t$ .

In the special case where there is no intertemporal consumption so that utility is only derived from terminal wealth we obtain the following LP:

$$\min \sum_{t=0}^T \sum_{i=1}^m J_i^t$$

subject to

$$J_i^t \geq \sum_{k=1}^m p^t(i, u, k) J_k^{t+1} \quad \forall u, i, t = 0, \dots, T-1 \quad (C6)$$

$$\begin{aligned} J_i^T &\geq u^T(i) \quad \forall i \\ J^{T+1} &= 0 \end{aligned} \quad (C7)$$

where  $u_i^T$  denotes the utility of terminal wealth in state  $i$ .

---

<sup>5</sup>Note that if we omit the second and third expressions in the right hand side of (C5) we would still have a feasible solution. In this case, however, finding  $V$  amounts to solving the dynamic program exactly which is assumed to be very difficult. Defining  $V$  as in (C5) may make it easy to compute an upper bound,  $V_0$ .

Again, assuming we have available  $\{\tilde{J}^t\}$  such that  $\tilde{J}^T$  satisfies (C7), we can construct  $\{V^t\}$  so that they satisfy (C6) and (C7). We define

$$\begin{aligned} V^{T+1} &= 0 \\ V^T &= \tilde{J}^T \\ &\vdots \\ V^{t-1} &= E_{t-1}^{bs_{t-1}}[V^t] - E_{t-1}^{bs_{t-1}}[\tilde{J}^t - \tilde{J}^{t-1}] + \max_u(E_{t-1}^u[\tilde{J}^t - \tilde{J}^{t-1}])^+. \end{aligned}$$

By construction (C6) and (C7) are satisfied for all  $u, i$  and  $t$ , so that  $V^0$  is an upper bound for the true value function. As before we can write the upper bound explicitly to get

**Corollary 2** *An upper bound for the portfolio optimization problem without intertemporal consumption is given by*

$$V^0 = \tilde{J}^0 + E_0^{bs} \left[ \sum_{t=1}^T \max_u (E_{t-1}[\tilde{J}^t - \tilde{J}^{t-1}])^+ \right].$$

## Discussion

Theorem 5 appears to suggest a method for constructing upper bounds for general finite horizon dynamic programming problems. Having found an approximate solution,  $\tilde{J}^t$ , we simply estimate the upper bound using simulation. Unfortunately, we need to simulate with respect to the measure as determined by  $bs$ , and this appears to be difficult since we do not know  $\{V^t\}$ . On the other hand, since we do know  $\{\tilde{J}^t\}$  it is very easy to simulate with respect to the probability measure that is determined by optimizing with respect to  $\tilde{J}^{t+1}$  at each date,  $t$ . We expect this measure to be close to the measure determined by  $bs$  if  $\{\tilde{J}^t\}$  is a good approximation. Nevertheless, it is not the appropriate measure and therefore does not allow us to estimate the upper bound.

One problem where this difficulty can be overcome is when we can replace the measure determined by  $bs$  with another measure,  $m$ , so that the expectation under  $m$  is at least as large as the expectation under the measure determined by  $bs$ . We can do this for example in optimal stopping problems where the measure  $m$  corresponds to the strategy of never exercising early. We then retrieve the bound of Theorem 4. Interestingly, this implies that a superior upper bound for the American option problem than that given in Chapter 2 is given by Corollary 2. This bound also uses  $\{\tilde{J}^t\}$  but we do not seem to be able to estimate it as we cannot simulate with respect to the correct probability measure.

## Bibliography

- Ait-Sahalia, Y. and A. Lo, 1998, "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices", *Journal of Finance* 52, 499–548.
- Aiyagari, R., 1994, "Uninsured Idiosyncratic Risk and Aggregate Saving", *Quarterly Journal of Economics* 109, 659–684.
- Aiyagari, R. and M. Gertler, 1991, "Asset Returns with Transactions Costs and Uninsured Individual Risk", *Journal of Monetary Economics* 27, 311–331.
- Arnott, R. and F. Fabozzi, 1992, *Active Asset Allocation*. Chicago, IL: Probus Publishing Company.
- Avellaneda, M. and A. Paras, 1994, "Dynamic Hedging Portfolios for Derivative Securities in the Presence of Large Transactions Costs", *Applied Mathematical Finance* 1, 165–193.
- Balduzzi, P. and A. Lynch, 1999, "Transaction Costs and Predictability: Some Utility Cost Calculations", *Journal of Financial Economics* 52, 47–78.
- Barraquand, J. and D. Martineau, 1995, "Numerical Valuation of High-Dimensional Multivariate American Securities", *Journal of Financial and Quantitative Analysis* 30, 383–405.
- Bensaid, B., Lesne, J., Pages, H. and J. Scheinkman, 1992, "Derivative Asset Pricing with Transaction Costs", *Mathematical Finance* 2, 63–86.
- Bensoussan, A., 1984, "On the Theory of Option Pricing", *Acta Applicandae Mathematicae*, 2, 139–158.
- Bertsekas, D., 1999, *Nonlinear Programming*. Belmont, MA: Athena Scientific.
- Bertsekas, D. and J. Tsitsiklis, 1996, *Neuro-Dynamic Programming*, Athena Scientific, Belmont, MA.
- Bertsimas, D., Kogan, L. and A. Lo, 2000a, "When Is Time Continuous?", *Journal of Financial Economics* 55.
- Bertsimas, D., Kogan, L. and A. Lo, 2000b, "Hedging Derivative Securities and Incomplete Markets: An  $\epsilon$ -Arbitrage Approach", to appear in *Operations Research*.
- Billingsley, P., 1995, *Probability and Measure*, Third Edition, John Wiley & Sons.
- Birge, J. R., 1994, "Quasi-Monte Carlo Approaches to Option Pricing", Technical Report 94–19. Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109.
- Black, F. and M. Scholes, 1973, "Pricing of Options and Corporate Liabilities", *Journal of Political Economy* 81, 637–654.
- Bodie, Z., Kane, A. and A. Marcus, 1999, *Investments*, Fourth Edition. Boston, MA: Irwin/McGraw-Hill.

- Bodie, Z. and R. Merton, 1995, "Financial Infrastructure and Public Policy: A Functional Perspective", in D. Crane, K. Froot, S. Mason, A. Perold, R. Merton, Z. Bodie, E. Sirri, and P. Tufano, eds., *The Global Financial System: A Functional Perspective*. Boston, MA: Harvard Business School Press.
- Boyle, P., M. Broadie, and P. Glasserman, 1997, "Monte Carlo Methods for Security Pricing", *Journal of Economic Dynamics and Control*, 21, 1267–1321.
- Boyle, P., A.W. Kolkiewicz and K.S. Tan, 2001, "Pricing American Style Options Using Low Discrepancy Mesh Methods", Working Paper, University of Waterloo, Waterloo.
- Boyle, P. and T. Vorst, 1992, "Option Replication in Discrete Time with Transaction Costs", *Journal of Finance* 47, 271–294.
- Brandt, M.W., A. Goyal and P. Santa-Clara, 2001, "Dynamic Portfolio Choice: A Simulation Approach", Working Paper, University of Pennsylvania.
- Breeden, D. 1979, "An Intertemporal Capital Pricing Model with Stochastic Investment Opportunities", *Journal of Financial Economics* 7, 265–296.
- Broadie, M., P. Glasserman, 1997a, "Pricing American-style Securities Using Simulation", *Journal of Economic Dynamics and Control*, 21, 1323–1352.
- Broadie, M., P. Glasserman, 1997b, "A Stochastic Mesh Method for Pricing High-Dimensional American Options", Working Paper, Columbia University.
- Broadie, M., P. Glasserman, and G. Jain, 1997, "Enhanced Monte Carlo Estimates for American Option Prices", *Journal of Derivatives* 5, 25–44.
- Carr, P. and D. Madan, 2000, "Optimal Positioning in Derivative Securities", to appear in *Review of Derivatives Research*.
- Carr, P., Jin, X. and D. Madan, 2000, "Optimal Consumption and Derivative Investment Rules in Continuous Time", to appear in *Finance and Stochastics*.
- Carriere, J., 1996, "Valuation of Early-Exercise Price of Options using Simulations and Nonparametric Regression", *Insurance: Mathematics and Economics* 19, 19–30.
- Constantinides, G., 1986, "Capital Market Equilibrium with Transaction Costs", *Journal of Political Economy* 94, 842–862.
- Cox, J. and C. Huang, 1989, "Optimum Consumption and Portfolio Policies When Asset Prices Follow a Diffusion Process", *Journal of Economic Theory* 49, 33–83.
- Cox, J. and S. Ross, 1976, "The Valuation of Options for Alternative Stochastic Processes", *Journal of Financial Economics* 3, 145–166.
- Cox, J. and M. Rubinstein, 1985, *Options Markets*. Englewood Cliffs, NJ: Prentice-Hall.
- Delbaen, F. and W. Schachermeyer, 1996, "The Variance Optimal Martingale Measure for Continuous Processes", *Bernoulli* 9, 81–105.
- Davis, M. and A. Norman, 1990, "Portfolio Selection with Transactions Costs", *Mathematics of Operations Research* 15, 676–713.
- Davis, M., Panas, V. and T. Zariphopoulou, 1993, "European Option Pricing with Transaction Costs", *SIAM Journal of Control and Optimization* 31, 470–493.

- Duffie, D., 1996, *Dynamic Asset Pricing Theory*, Second Edition. Princeton, NJ: Princeton University Press.
- Duffie, D., 1987, "Stochastic Equilibria with Incomplete Financial Markets", *Journal of Economic Theory* 41, 405–416.
- Duffie, D. and C. Huang, 1985, "Implementing Arrow-Debreu Equilibria by Continuous Trading of Few Long-Lived Securities", *Econometrica* 53, 1337–1356.
- Duffie, D. and M. Jackson, 1990, "Optimal Hedging and Equilibrium in a Dynamic Futures Market", *Journal of Economic Dynamics and Control* 14, 21–33.
- Duffie, D. and W. Shafer, 1985, "Equilibrium in Incomplete Markets I: A Basic Model of Generic Existence", *Journal of Mathematical Economics* 14, 285–300.
- Duffie, D. and W. Shafer, 1986, "Equilibrium in Incomplete Markets II: Generic Existence in Stochastic Economies", *Journal of Mathematical Economics* 15, 199–216.
- Dumas, B. and E. Luciano, 1991, "An Exact Solution to the Portfolio Choice Problem Under Transactions Costs", *Journal of Finance* 46, 577–595.
- Edirisinghe, C., Naik, V. and R. Uppal, 1993, "Optimal Replication of Options with Transaction Costs and Trading Restrictions", *Journal of Financial and Quantitative Analysis* 28, 117–138.
- Evnine, J. and R. Henriksson, 1987, "Asset Allocation and Options", *Journal of Portfolio Management* 14, 56–61.
- Föllmer, H. and D. Sonderman, 1986, "Hedging of Non-Redundant Contingent-Claims", in W. Hildebrand and A. Mas-Colell, eds., *Contributions to Mathematical Economics, in Honor of Gérard Debreu*. Amsterdam: North-Holland.
- Fourer, R., Gay, D. and B. Kernighan, 1999, *AMPL: A Modeling Language for Mathematical Programming*. San Francisco, CA: Scientific Press.
- Garcia, D., 1999, "A Monte Carlo Method for Pricing American Options", Working Paper, University of California at Berkeley.
- Garcia, R., J. Detemple and M. Rindisbacher, 2000, "A Monte-Carlo Method for Optimal Portfolios", Working Paper.
- Gennotte, G. and A. Jung, 1992, "Commissions and Asset Allocation", *Journal of Portfolio Management* 19, 12–23.
- Geske, R. and H.E. Johnson, 1984, "The American Put Options Valued Analytically", *Journal of Finance* 39, 1511–1524.
- Goll, T. and L. Rueschendorf, 2000, "Minimax and Minimal Distance Martingale Measures and Their Relationship to Portfolio Optimization", unpublished working paper, Institut für Mathematische Stochastik, Universität Freiburg.
- Goonatilake, S. and P. Treleaven, Eds., 1995, *Intelligent Systems for Finance and Business*, John Wiley & Sons.
- Grannan, E. and G. Swindle, 1996, "Minimizing Transaction Costs of Option Hedging Strategies", *Mathematical Finance* 6, 341–364.

- Harrison, J. M. and D. Kreps, 1979, "Martingales and Arbitrage in Multiperiod Securities Markets", *Journal of Economic Theory* 20, 381-408.
- Harrison, J. M. and S. R. Pliska 1981, "Martingales and Stochastic Integrals in the Theory of Continuous Trading", *Stochastic Processes and their Applications* 11, 215-260.
- He, H. and D. Modest, 1995, "Market Frictions and Consumption-Based Asset Pricing", *Journal of Political Economy* 103, 94-117.
- He, H. and N. Pearson, 1991, "Consumption and Portfolio Policies with Incomplete Markets and Short-Sale Constraints", *Journal of Economic Theory* 54, 259-304.
- Heaton, J. and D. Lucas, 1992, "The Effects of Incomplete Insurance Markets and Trading Costs in a Consumption-Based Asset Pricing Model", *Journal of Economic Dynamics and Control* 16, 601-620.
- Heaton, J. and D. Lucas, 1996, "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing", *Journal of Political Economy* 104, 443-487.
- Henriksson, R. and R. Merton, 1981, "On Market Timing and Investment Performance II: Statistical Procedures for Evaluating Forecasting Skills", *Journal of Business* 54, 513-533.
- Henrotte, P., 1993, "Transactions Costs and Duplication Strategies", unpublished working paper, Stanford University.
- Hodges, S. and A. Neuberger, 1989, "Optimal Replication of Contingent Claims Under Transaction Costs", *Review of Futures Markets* 8, 222-239.
- Hornick, K., M. Stinchcombe and H. White, 1989, "Multilayer feedforward Networks are Universal Approximators", *Neural Networks* Vol 2, 359-366.
- Huang, C., 1985a, "Information Structure and Equilibrium Asset Prices", *Journal of Economic Theory* 53, 33-71.
- Huang, C., 1985b, "Information Structures and Viable Price Systems", *Journal of Mathematical Economics* 14, 215-240.
- Hutchinson, A., A. Lo and T. Poggio, 1994, "A Nonparametric Approach to the Pricing and Hedging of Derivative Securities Via Learning Networks", *Journal of Finance* 49, 851-889.
- Ibanez, A. and F. Zapatero, 1998, "Monte Carlo Valuation of American Options through Computation of the Optimal Exercise Frontier", Instituto Tecnológico Autónomo de México.
- Joy, C., P.P. Boyle, and K.S. Tan, 1996, "Quasi-Monte Carlo Methods in Numerical Finance", *Management Science* 42, 926-938.
- Kim, T. and E. Omberg, 1996, "Dynamic Nonmyopic Portfolio Behavior", *Review of Financial Studies* 9, 141-161.
- Kim, T. and E. Omberg, 1996, "Dynamic Nonmyopic Portfolio Behavior", Working Paper, San Diego State University.
- Karatzas, I, 1998, "On the Pricing of American Options", *Applied Mathematics and Optimization* 17, 37-60.



- Karatzas, I. and S. Shreve, 1998, *Methods of Mathematical Finance* Springer-Verlag New York.
- Leland, H., 1985, "Option Pricing and Replication with Transaction Costs", *Journal of Finance*, 40, 1283–1301.
- Liu, J., 1999, "Portfolio Selection in Stochastic Environments", PhD Thesis, Graduate School of Business, Stanford University.
- Lo, A., 2000, "Risk Management For Hedge Funds: An Introduction and Overview", unpublished manuscript.
- Longstaff, F. and E. Schwartz, "Valuing American Options by Simulation: A Simple Least-Squares Approach", *Review of Financial Studies*, forthcoming.
- Lucas, D., 1994, "Asset Pricing With Undiversifiable Income Risk and Short Sales Constraints: Deepening the Equity Premium Puzzle", *Journal of Monetary Economics* 34, 325–341.
- Magill, M. and G. Constantinides, 1976, "Portfolio Selection with Transactions Costs", *Journal of Economic Theory* 13, 245–263.
- Magill, M. and M. Quinzii, 1996, *Theory of Incomplete Markets*. Cambridge, MA: MIT Press.
- Merton, R., 1969, "Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case", *Review of Economics and Statistics* 51, 247–257.
- Merton, R., 1971, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model", *Journal of Economic Theory* 3, 373–413.
- Merton, R., 1973, "Theory of Rational Option Pricing", *Bell Journal of Economics and Management Science* 4, 141–183.
- Merton, R., 1981, "On Market Timing and Investment Performance I: An Equilibrium Theory of Value for Market Forecasts", *Journal of Business* 54, 363–406.
- Merton, R., 1992, *Continuous-Time Finance*. Cambridge, MA: Blackwell Publishers.
- Merton, R., 1995, "Financial Innovation and the Management and Regulation of Financial Institutions", *Journal of Banking and Finance* 19, 461–481.
- Merton, R., 1997, "On the Role of the Wiener Process in Finance Theory and Practice: The Case of Replicating Portfolios", *Proceedings of Symposia in Pure Mathematics* 60, 218–219.
- Moody, J. et al., 1998, "Performance Functions and Reinforcement Learning for Trading Systems and Portfolios", *Journal of Forecasting* 17, 441–470.
- Neuberger, A., 1994, "Option Replication with Transaction Costs: An Exact Solution for the Pure Jump Process", in D. Chance and R. Trippi, eds., *Advances in Futures and Options Research*, Volume 7. Greenwich, CT: JAI Press.
- Neuneier, R., 1996, "Optimal Asset Allocation Using Adaptive Dynamic Programming", *Advances in Neural Information Processing Systems*, D. Touretzky, M. Mozer, and M. Hasselmo, eds., MIT Press, Cambridge, Mass.

- Niederreiter, H., 1992, "Random Number Generation and Quasi-Monte Carlo Methods", CBMS-NSF 63, SIAM, Philadelphia, PA.
- Paskov, S. and J. Traub, 1995, "Faster Valuation of Financial Derivatives", *Journal of Portfolio Management* 22, Fall, 113–120.
- Pliska, S., 1997, *Introduction to Mathematical Finance*, Blackwell Publishers Inc., Malden, MA.
- Press, W.H., S.A. Teukolsky, S.A. Vetterling and W.T. Flannery, 1992, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd Ed, Cambridge University Press, Cambridge.
- Raymar, S. and M. Zwecher, 1997, "A Monte Carlo Valuation of American Call Options on the Maximum of Several Stocks", *Journal of Derivatives* 5, 7–23.
- Rogers, L.C.G., 2001, Monte Carlo Valuation of American Options, Working Paper, University of Bath.
- Samuelson, P., 1969, "Lifetime Portfolio Selection by Dynamic Stochastic Programming", *Review of Economics and Statistics* 51, 239–246.
- Schäl, M., 1994, "On Quadratic Cost Criteria for Option Hedging" *Mathematics of Operations Research* 19, 121–131.
- Scheinkman, J. and L. Weiss, 1986, "Borrowing Constraints and Aggregate Economic Activity", *Econometrica* 54, 23–45.
- Schweizer, M., 1992, "Mean-Variance Hedging for General Claims", *Annals of Applied Probability* 2, 171–179.
- Schweizer, M., 1995, "Variance-Optimal Hedging in Discrete Time", *Mathematics of Operations Research* 20, 1–31.
- Schweizer, M., 1996, "Approximation Pricing and the Variance-Optimal Martingale Measure", *Annals of Probability* 24, 206–236.
- Sharpe, W., 1987, "Integrated Asset Allocation", *Financial Analysts Journal* 43, 25–32.
- Telmer, C., 1993, "Asset-Pricing Puzzles and Incomplete Markets", *Journal of Finance* 48, 1803–1832.
- Tezuka, S., 1995, *Uniform Random Numbers: Theory and Practice*, Kluwer Academic Publishers, Boston.
- Tilley, J.A., 1993, "Valuing American Options in a Path Simulation Model", *Transactions of the Society of Actuaries* 45, 83–104.
- Toft, K., 1996, "On the Mean-Variance Tradeoff in Option Replication with Transactions Costs", *Journal of Financial and Quantitative Analysis* 31, 233–263.
- Tsitsiklis, J., and B. Van Roy, 1999, "Optimal Stopping of Markov Processes: Hilbert Space Theory, Approximation Algorithms, and an Application to Pricing Financial Derivatives", *IEEE Transactions on Automatic Control* 44, 1840–1851.
- Tsitsiklis, J., and B. Van Roy, 2000, "Regression Methods for Pricing Complex American-Style Options", Working Paper, MIT.

- Van Roy, B., 1999, "Temporal Difference Learning and Applications in Finance", *Computational Finance*, Y.S. Abu-Mostafa et al., eds, MIT Press, Cambridge, Mass.
- Vapnik, V.N., 1999, *The Nature of Statistical Learning Theory (Statistics for Engineering and Information Science)*, 2nd Ed, Springer Verlag.
- Weil, P., 1992, "Equilibrium Asset Prices with Undiversifiable Labor Income Risk", *Journal of Economic Dynamics and Control* 16, 769-790.
- Whalley, E. and P. Wilmott, 1994, "Hedge with an Edge", *RISK* 7.

