Portfolio Optimization with Position Constraints: an Approximate Dynamic Programming Approach

Martin B. Haugh*

Leonid Kogan

Sloan School of Management, MIT, Massachusetts, MA 02142, lkogan@mit.edu.

Zhen Wu

Department of IE and OR, Columbia University, New York, NY 10027, zw2002@columbia.edu.

*This research for this paper was completed while the first author was at the Department of Industrial Engineering and Operations Research at Columbia University. He can be contacted at <mhaugh@alum.mit.edu>.

Abstract

We analyze dynamic portfolio choice problems using an approximate dynamic programming (ADP) algorithm. We extend the algorithm to the case of constraints on borrowing and implement a duality-based simulation procedure for estimating bounds on the true value function. We demonstrate that the ADP solution exhibits a high degree of accuracy in the considered examples, indicating that this is a promising approach to tackling challenging practical problems in the area of asset allocation and portfolio choice. We present additional evidence on the performance of the duality-based method for estimating performance of approximate portfolio rules, showing that it provides a valuable tool in conjunction with ADP-style algorithms.

Subject Classifications: Finance: portfolio optimization. Dynamic Programming: optimal control, duality theory.

1 Introduction

Recent years have seen increased interest in the problem of optimal portfolio choice. This interest has been fueled partly by methodological advances, and partly by the growing practical importance of such problems, particularly since the increased emphasis on defined contribution pension plans has recently shifted the burden of life-time asset allocation onto individuals.

A significant effort in academic literature has been directed towards obtaining explicit solutions to selected problems (e.g., Merton 1971, 1990, Karatzas, Lehoczky, Sethi and Shreve, 1986, Kim and Omberg, 1986, Cox and Huang, 1989, Liu, 1998, Wachter, 2002) and analyzing calibrated versions of simple models (e.g., Brennan, Schwartz and Lagnado, 1997, Munk, 2000, Campbell and Viceira, 1999, 2002, Brennan and Xia, 2002, Barberis, 2002, Campbell, Chan and Viceira, 2003, Chacko and Viceira, 2005, Wachter and Sangvinatsos, 2005), avoiding the challenge posed by lack of analytical tractability and high dimensionality of many problems of practical interest. Recently, Brandt, Goyal, Santa-Clara and Stroud (2005) suggested a computational algorithm, based on approximate dynamic programming (ADP) ideas, aimed at high-dimensional and computationally intensive problems. Their algorithm relies on functional approximations to the value function and applies to problems with incomplete financial markets. An important unexplored aspect of their algorithm is the quality of approximation. While the procedure can be shown to perform well on a few sample problems with known solutions, it is clearly not a guarantee that it will fair as well on more challenging problems for which solutions cannot be obtained using standard methods.

In this paper, we show how ADP approach can be strengthened by a rigorous duality-based simulation method for evaluating the quality of approximate solutions, developed by Haugh, Kogan, and Wang (2006) (HKW). Their simulation method can handle problems with position constraints, including incomplete markets. For the examples we consider, the ADP algorithm delivers accurate approximations to the optimal policy, as verified by our simulation technique. Second, our analysis also demonstrates the practical potential of the duality-based method of HKW. The latter demonstrated that their simulation method shows promise in several examples based on a simple myopic portfolio policy. In the problems considered, the myopic solution was close to the optimum, leaving open the question of how well the algorithm would perform on problems with a significant hedging component in the optimal policy. In this paper, we show that the simulation method of HKW also performs well on problems for which myopic strategies are far from optimal.

2 The Model

We consider a portfolio choice problem under incomplete markets and portfolio constraints. We formulate our model in discrete time: the economy exists between 0 and T, and trading takes place at equally spaced periods $t_i \in [0, T]$, $i = (0, 1, ..., I)$, $t_0 = 0$, $t_I = T$. In later sections, when we compute duality-based performance bounds, we define the discrete-time model as an Euler approximation to a continuous-time diffusion model.

The investment opportunity set

There are N risky assets and a riskfree asset. Without loss of generality, we assume that the assets pay no dividends. The instantaneous moments of asset returns depend on the M-dimensional vector of state variables X_t. We assume that the vector of state variables X_t follows a general Markov process. The vector of risky returns is assumed to be conditionally log-normally distributed, i.e., return on the risky asset n over the time interval $[t_i, t_{i+1})$ is given by $R_i^{(n)} = \exp(r_i^{(n)})$, where the vector $r_i = (r_i^{(1)}, r_i^{(2)}, ..., r_i^{(N)})^\top$ has a conditional multi-variate normal distribution, given the state X_i at time t_i. Thus, in order to specify the return process, we only need to define the vector of conditional means and the variance-covariance matrix of r_i. We denote the return on the riskfree asset over the same time interval by $R_i^{(f)} = \exp(r_i^{(f)})$.

The objective function

We assume that the portfolio policy is chosen to maximize the expected utility of wealth at the terminal date T, $\mathrm{E}_0[U(W_T)]$. We further assume that the utility function is of constant relative risk aversion (CRRA) type, so that $U(W) = W^{1-\gamma}/(1-\gamma)$. In addition to being a very popular specification of preferences, the CRRA utility function has an advantage that optimal portfolio policies are independent of wealth, which makes the problem more tractable computationally. We let θ denote the vector of portfolio shares invested in the risky assets, so that the return on the portfolio is given by

$$R_i^{(f)} + \theta^\top (R_i - R_i^{(f)} \imath),$$

where $\imath = (1, ..., 1)^\top$.

In summary, the portfolio choice problem is to solve for

$$\sup_{\{\theta_i\}} \mathrm{E}_0 \left[\frac{1}{1-\gamma} W_T^{1-\gamma} \right] \tag{1}$$

subject to the budget constraint

$$W_{i+1} = W_i \left(R_i^{(f)} + \theta^\top (R_i - R_i^{(f)} \imath) \right).$$

3 Approximate Dynamic Programming for Portfolio Optimization

We now describe our method for constructing approximate solutions to the portfolio choice problems. We use the ADP algorithm for incomplete markets, developed by Brandt et al. (2005). We also extend their method to handle no-borrowing constraints.

Brandt et al (2005) propose solving the portfolio choice problem by approximating the value function $V_i(X_i, W_i)$. The value function satisfies the Bellman equation

$$V_i(X_i, W_i) = \max_{\theta_i} E_i \left[V_{i+1} \left(X_{i+1},\ W_i \left(R_i^{(f)} + \theta_i^\top \left(R_i - R_i^{(f)} \imath \right) \right) \right) \right]. \tag{2}$$

with the terminal condition $V_T(X_T, W_T) = W_T^{1-\gamma}/(1-\gamma)$.

Since in most cases the dynamic program (2) cannot be solved exactly we instead look for an approximate solution. Following Brandt et al (2005), we use a 4^{th}-order Taylor expansion of V_{i+1} around $W_i R_i^{(f)}$ in the right-hand-side of (2) and obtain

$$\begin{aligned} V_i(X_i, W_i) \ \approx\ \max_{\theta_i}\ E_i \Big[& V_{i+1}(X_{i+1}, W_i R_i^{(f)}) \ +\ \partial_2 V_{i+1}(X_{i+1}, W_i R_i^{(f)}) W_i \widetilde{R}_i^\top \theta_i \\ & + \frac{1}{2} \partial_2^2 V_{i+1}(X_{i+1}, W_i R_i^{(f)}) (W_i \widetilde{R}_i^\top \theta_i)^2 \ +\ \frac{1}{6} \partial_2^3 V_{i+1}(X_{i+1}, W_i R_i^{(f)}) (W_i \widetilde{R}_i^\top \theta_i)^3 \\ & + \frac{1}{24} \partial_2^4 V_{i+1}(X_{i+1}, W_i R_i^{(f)}) (W_i \widetilde{R}_i^\top \theta_i)^4 \Big] \end{aligned} \tag{3}$$

where $\widetilde{R}_i = \left(R_i - R_i^{(f)} 1 \right)$ is the vector of excess returns on the risky assets and $\partial_2^n V_i$ denotes the n-th order partial derivative of the value function with respect to its second argument, wealth.

If we use $\hat{\theta}_i$ to denote the optimal weights in (3), then we can characterize the partial derivatives of the value function as

$$\begin{aligned} V_{i+1}(X_{i+1}, W_i R_i^{(f)}) &= \frac{(W_i R_i^{(f)})^{1-\gamma}}{1-\gamma} E_{i+1}[H_{i+1}] \\ \partial_2 V_{i+1}(X_{i+1}, W_i R_i^{(f)}) &= (W_i R_i^{(f)})^{-\gamma} E_{i+1}[H_{i+1}] \\ \partial_2^2 V_{i+1}(X_{i+1}, W_i R_i^{(f)}) &= -\gamma (W_i R_i^{(f)})^{-\gamma-1} E_{i+1}[H_{i+1}] \\ \partial_2^3 V_{i+1}(X_{i+1}, W_i R_i^{(f)}) &= \gamma(\gamma+1) (W_i R_i^{(f)})^{-\gamma-2} E_{i+1}[H_{i+1}] \\ \partial_2^4 V_{i+1}(X_{i+1}, W_i R_i^{(f)}) &= -\gamma(\gamma+1)(\gamma+2) (W_i R_i^{(f)})^{-\gamma-3} E_{i+1}[H_{i+1}] \end{aligned}$$

where

$$H_{i+1} = \prod_{k=i+1}^{I-1} \left(R_k^{(f)} + \widetilde{R}_k^\top \hat{\theta}_k \right)^{1-\gamma}.$$

We can then rewrite (3) to obtain

$$
\begin{aligned}
V_i(X_i, W_i) \;=\; W_i^{1-\gamma} \; \max_{\theta_i} E_i \Bigg[& \frac{1}{1-\gamma}\left(R_i^{(f)}\right)^{1-\gamma} E_{i+1}[H_{i+1}] \;+\; \left(R_i^{(f)}\right)^{-\gamma} E_{i+1}[H_{i+1}]\widetilde{R}_i^\top \theta_i \\
& - \frac{1}{2}\gamma\left(R_i^{(f)}\right)^{-\gamma-1} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \theta_i)^2 \;+\; \frac{1}{6}\gamma(\gamma+1)\left(R_i^{(f)}\right)^{-\gamma-2} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \theta_i)^3 \\
& - \frac{1}{24}\gamma(\gamma+1)(\gamma+2)\left(R_i^{(f)}\right)^{-\gamma-3} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \theta_i)^4 \Bigg]. \quad (4)
\end{aligned}
$$

The optimal solution to (4) is clearly independent of W_i. The first-order optimality conditions corresponding to (4) are given by

$$
\begin{aligned}
E_i \Bigg[& \left(R_i^{(f)}\right)^{-\gamma} E_{i+1}[H_{i+1}]\widetilde{R}_i - \gamma\left(R_i^{(f)}\right)^{-\gamma-1} E_{i+1}[H_{i+1}]\widetilde{R}_i^\top \hat{\theta}_i \widetilde{R}_i \\
& + \tfrac{1}{2}\gamma(\gamma+1)\left(R_i^{(f)}\right)^{-\gamma-2} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \hat{\theta}_i)^2 \widetilde{R}_i \\
& - \tfrac{1}{6}\gamma(\gamma+1)(\gamma+2)\left(R_i^{(f)}\right)^{-\gamma-3} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \hat{\theta}_i)^3 \widetilde{R}_i \Bigg] \;=0
\end{aligned}
$$

implying in particular that

$$
\hat{\theta}_i = \left(E_i[B_{i+1}]\right)^{-1}\left\{\frac{1}{\gamma}E_i[A_{i+1}] + \frac{1}{2}(1+\gamma)E_i[C_{i+1}(\hat{\theta}_i)] - \frac{1}{6}(1+\gamma)(2+\gamma)E_i[D_{i+1}(\hat{\theta}_i)]\right\} \quad (5)
$$

where

$$
\begin{aligned}
A_{i+1} &= \left(R_i^{(f)}\right)^{-\gamma} E_{i+1}[H_{i+1}]\widetilde{R}_i & (6)\\
B_{i+1} &= \left(R_i^{(f)}\right)^{-\gamma-1} E_{i+1}[H_{i+1}]\widetilde{R}_i\widetilde{R}_i^\top & (7)\\
C_{i+1}(\theta_i) &= \left(R_i^{(f)}\right)^{-\gamma-2} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \theta_i)^2 \widetilde{R}_i & (8)\\
D_{i+1}(\theta_i) &= \left(R_i^{(f)}\right)^{-\gamma-3} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \theta_i)^3 \widetilde{R}_i. & (9)
\end{aligned}
$$

We discuss how functions A_{i+1}, B_{i+1}, C_{i+1} and D_{i+1} can be estimated in Section 3.1 below.

For now, assume that these quantities can be estimated with sufficient accuracy. Then, we can solve (5) by a simple iterative procedure. In particular, if we start with a good initial approximation to $\hat{\theta}_i$, θ_i^0, then we can substitute θ_i^0 into the right-hand-side of (5) and obtain a new estimate, θ_i^1. We iterate in this manner until the sequence of estimates, θ_i^k, converges.

There are many possible ways to find accurate initial approximations, θ_i^0, for the iteration. Brandt et al. (2005) suggest using the solution, to the ADP algorithm that is based on using a 2^{nd} order Taylor expansion of the value function in (3). It is easily seen that θ_i^0 is then given explicitly as

$$\theta_i^0 = (E_i[B_{i+1}])^{-1} \{\frac{1}{\gamma} E_i[A_{i+1}]\}. \tag{10}$$

Alternatively, one can use simple suboptimal portfolio strategies, such as the static and myopic strategies of Section 5.2, as starting points for iteration.

There is no theoretical guarantee that the iterative procedure above converges. sometimes, one may need to use a higher order approximation to achieve convergence. For example, the corresponding iteration for the 5^{th}-order approximation is given by

$$\begin{aligned}\theta_i^{k+1} = (E_i[B_{i+1}])^{-1} \{\frac{1}{\gamma} E_i[A_{i+1}] + \frac{1}{2}(1+\gamma)E_i[C_{i+1}(\theta_i^k)] - \frac{1}{6}(1+\gamma)(2+\gamma)E_i[D_{i+1}(\theta_i^k)] \\ + \frac{1}{24}(1+\gamma)(2+\gamma)(3+\gamma)E_i[F_{i+1}(\theta_i^k)]\}\end{aligned} \tag{11}$$

where $F_{i+1}(\theta_i^k) = \left(R_i^{(f)}\right)^{-\gamma-4} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \theta_i^k)^4 \widetilde{R}_i$.

The procedure described above is suggested in Brandt et al. (2005). In the following sections, we generalize their method to handle portfolio constraints. We introduce our technique using several examples with common portfolio constraints. However, in order to motivate our procedure for handling constraints, we first discuss the method for estimating conditional expectations in (6–9).

3.1 Numerical Implementation: Estimating Conditional Expectations

In order to find the optimal $\hat{\theta}_i$ via the iterative procedures of (5) or (15) it is necessary to estimate the quantities $E_i[A_{i+1}]$, $E_i[B_{i+1}]$, $E_i[C_{i+1}]$ and $E_i[D_{i+1}]$ where A_{i+1}, B_{i+1}, C_{i+1} and D_{i+1} are given by equations (6), (7), (8) and (9), respectively. Indeed since C_{i+1} and D_{i+1} are actually functions of θ_i it is necessary to re-estimate them in each iteration of (5) and (15). The key to a useful computational algorithm is to be able to estimate required conditional expectations efficiently. A naive approach of using a full-scale "within simulation" Monte Carlo simulation loop to estimate each conditional expectation is too costly computationally and would make multi-period problems extremely difficult to solve. Besides, in order to estimate conditional expectations, one would have to have information about the optimal policy along each of the paths in the internal simulation loop, which is a non-trivial obstacle. To get around the need to conduct simulations within simulations, Brandt et al. (2005) suggest a clever across-path regression procedure, based on the approaches of Longstaff and Schwartz (2001) and Tsitsiklis and Van Roy (2001) for pricing American options. We now summarize this approach and then discuss how it can be extended to constrained problems.

The Across-Path Regression Approach

The algorithm is based upon simulating S sample paths of the underlying state variables from $t = 0$ to $t = T$. We now describe how to estimate $h(X_i) = E_i[G]$, where G is a random variable, depending on the future values of state variables.

The S sample paths provide us with S realizations of the pair, (X_i, G). Using (X_i^s, G^s) for $s = 1, \ldots, S$ to denote these realizations, we may now consider a linear regression of the form

$$G \;=\; \beta_i^\top f(X_i) + \epsilon_i, \tag{12}$$

where $f(X_i)$ is a matrix containing S rows, each of which is a vector of values of basis functions and ϵ_i is an orthogonal error term. Each basis function should be a function of time i information, X_i. We choose a complete set of basis functions, which guarantees an arbitrarily accurate approximation with a sufficient number of basis functions and a large enough number of simulated paths. For example, using polynomial basis functions, and truncating our expansion at the second order, we can take a constant, linear functions, and quadratic functions to form our basis, in which case an n-th row of $f(X_i)$ is given by

$$f_n(X_i) = \left[1 \quad (X_i^n)^\top \quad \left(X_i^n \bigotimes X_i^n \right)^\top \right]$$

Higher-order terms of the form $(X_i^n \bigotimes X_i^n ... \bigotimes X_i^n)^\top$ can be added to improve approximation accuracy. Of course, to avoid over-fitting, the number of basis functions should remain sufficiently small compared to the number of simulated paths, S.

After estimating the regression (12), we use the fitted values $\hat{\beta}_i$ to approximate the conditional expectation, $h(X_i)$:

$$h(X_i) \approx \hat{h}(X_i) \;=\; \hat{\beta}_i^\top f(X_i)$$

3.2 No-borrowing Constraint

The across-path regression approach of the previous Section is an effective tool for handling incomplete-markets problems, for which explicit expressions (5) are available. However, a straightforward application of such an approach to problems with portfolio constraints runs into significant computational obstacles: as soon as constraints on the portfolio composition, θ, are imposed, one has to solve a constrained optimization problem for every simulation path at every trading period. This precludes one from estimating the optimal portfolio policy using an efficient regression approach.

In this paper we do not attempt to provide a general treatment of portfolio constraints. Instead we consider the special case of no-borrowing constraints and solve the problem using an intuitive application of Lagrangian relaxation. In particular, we relax constraints using Lagrange multipliers and obtain explicit expressions for the optimal policy in terms of such multipliers. We then perform iterations, similar to the case of incomplete markets, updating the multipliers at each step. We do it in a manner that guarantees feasibility of our approximate solution. The resulting iterative procedure is well suited for implementation using the across-path regressions. We can then apply the duality-based performance evaluation algorithm to the constrained portfolio choice problem.

When the no-borrowing constraint is imposed, the optimization problem corresponding to the 4^{th}-order approximation to the value function takes form

$$\begin{aligned} V_i(X_i, W_i) \; = \; \max_{\imath^\top \theta_i \le 1} E_i \Bigg[\frac{1}{1-\gamma} \left(R_i^{(f)}\right)^{1-\gamma} E_{i+1}[H_{i+1}] \; + \; \left(R_i^{(f)}\right)^{-\gamma} E_{i+1}[H_{i+1}] \widetilde{R}_i^\top \theta_i \\ - \frac{1}{2}\gamma \left(R_i^{(f)}\right)^{-\gamma-1} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \theta_i)^2 \; + \; \frac{1}{6}\gamma(\gamma+1) \left(R_i^{(f)}\right)^{-\gamma-2} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \theta_i)^3 \\ - \frac{1}{24}\gamma(\gamma+1)(\gamma+2) \left(R_i^{(f)}\right)^{-\gamma-3} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \theta_i)^4 \Bigg] \end{aligned} \tag{13}$$

Let $\alpha \le 0$ denote a Lagrange multiplier on the portfolio constraint in (13). We relax the no-borrowing constraint to obtain a problem of the form

$$\begin{aligned} V_i(X_i, W_i) \; = \; \max_{\theta_i} E_i \left[\frac{\left(R_i^{(f)}\right)^{1-\gamma}}{1-\gamma} E_{i+1}[H_{i+1}] \; + \; \left(R_i^{(f)}\right)^{-\gamma} E_{i+1}[H_{i+1}] \widetilde{R}_i^\top \theta_i \right. \\ - \frac{1}{2}\gamma \left(R_i^{(f)}\right)^{-\gamma-1} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \theta_i)^2 \; + \; \frac{1}{6}\gamma(\gamma+1) \left(R_i^{(f)}\right)^{-\gamma-2} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \theta_i)^3 \\ \left. - \frac{1}{24}\gamma(\gamma+1)(\gamma+2) \left(R_i^{(f)}\right)^{-\gamma-3} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \theta_i)^4 + \alpha(\imath^\top \theta_i - 1) \right] \end{aligned} \tag{14}$$

Note that the effect of the term $\alpha(\imath^\top \theta_i - 1)$ in (14) is to penalize the objective function when the no-borrowing constraint is violated.

The first-order conditions for the optimization problem in (14) are given by

$$\begin{aligned} E_i \Bigg[\left(R_i^{(f)}\right)^{-\gamma} E_{i+1}[H_{i+1}] \widetilde{R}_i - \gamma \left(R_i^{(f)}\right)^{-\gamma-1} E_{i+1}[H_{i+1}] \widetilde{R}_i^\top \hat{\theta}_i \widetilde{R}_i \\ + \frac{1}{2}\gamma(\gamma+1) \left(R_i^{(f)}\right)^{-\gamma-2} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \hat{\theta}_i)^2 \widetilde{R}_i \\ - \frac{1}{6}\gamma(\gamma+1)(\gamma+2) \left(R_i^{(f)}\right)^{-\gamma-3} E_{i+1}[H_{i+1}](\widetilde{R}_i^\top \hat{\theta}_i)^3 \widetilde{R}_i + \alpha \imath \Bigg] = 0. \end{aligned}$$

As was the case for the incomplete markets problem without no-borrowing constraints, we can rewrite these conditions in the form

$$\hat{\theta}_i^{nb} = \hat{\theta}_i^{inc} + \frac{\alpha}{\gamma} \left(E_i[B_{i+1}]\right)^{-1} \imath \tag{15}$$

where $\hat{\theta}_i^{inc}$ refers to the expression for $\hat{\theta}_i$ in the incomplete market without any constraints. Starting with a feasible initial approximation to the portfolio policy, θ_i^0, we then employ the same iterative procedure as in Section 3 to compute the optimal $\hat{\theta}_i$. We maintain feasibility of the sequence $\theta_i^0, \theta_i^1, \ldots, \theta_i^k$ by varying α with each iteration. In particular, it is easy to see that we can maintain feasibility of θ_i^k if we let α_k be the multiplier for the k-th iteration and choose it to satisfy

$$\alpha_k = \begin{cases} \frac{\gamma(1-\imath^\top \theta_i^{inc,k})}{\imath^\top (E_i[B_{i+1}])^{-1}\imath}, & \text{if } \imath^\top \theta_i^{inc,k} \geq 1, \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

The multiplier α defined by (16) remains non-positive after each iteration, since the matrix B_{i+1} is positive-definite.

It is easy to check that our choice of the multiplier in (16) is such that, if the k-th iteration $\theta_i^{nb,k-1}$ were already optimal, then we would also recover the optimal solution as $\theta_i^{nb,k}$, and our iterative procedure would indicate convergence. This is because, by construction, the solution $\theta_i^{nb,k}$ is feasible for the original problem, satisfies the first-order optimality conditions for the relaxed problem, satisfies complimentary slackness conditions together with α_k, and the multiplier α_k is non-positive. Thus, to guarantee that $\theta_i^{nb,k}$ is optimal, it suffices that $E_i[B_{i+1}]$ is the same as under the optimal portfolio policy, which is the case if $\theta_i^{nb,k-1}$ is optimal.

4 The Algorithm

Estimating Derivatives of Conditional Expectations

In order to compute the upper bound on the true value function using our approximation, $\tilde{V}_i$, it is necessary to compute partial derivatives of $\tilde{V}_i$. One of the advantages of using a power utility function is that we only need to compute $\partial\tilde{V}_i/\partial x$ and do not need higher order derivatives. It is easy to see that finding $\partial\tilde{V}_i/\partial x$ amounts to finding $\partial E_i[H_i]/\partial x$. While differentiation is an inherently unstable procedure, we obtained satisfactory results by using the derivative of our approximation to $E_i[H_i]$ as our approximation to $\partial E_i[H_i]/\partial x$. This observation may be explained in part by the fact that in our numerical examples we found the magnitude of the θ component contributing to the market-price-of-risk to be significantly larger than the partial derivative component.

4.1 The ADP Algorithm

1. Simulate m paths of the state variables originating from X_0.

2. Starting from $t_n = T - \Delta$, work backwards computing (and storing: see the Section

below on constructing the policy) approximately optimal $\hat{\theta}_i$'s. We do this at time t_i for each of the m paths by:

(a) We choose a feasible starting point of $\theta_i^{\mathbf{0}}$ for each sample path. Set $k = 0$.

(b) We estimate $E_i[A_{i+1}]$, $E_i[B_{i+1}]$, $E_i[C_{i+1}(\theta_i^{\mathbf{k}})]$ and $E_i[D_{i+1}(\theta_i^{\mathbf{k}})]$ for each sample path using the across-path regression approach. Then using (5) (or (15) in the no-borrowing case), we obtain $\theta_i^{\mathbf{k+1}}$ for each sample path.

(c) If $\max_{1\leq j\leq m} ||\theta_i^{j,\mathbf{k+1}} - \theta_i^{j,\mathbf{k}}||_\infty$ is sufficiently small, we terminate at this time step. Otherwise we continue iterating using equation (5), or equation (15) in the no-borrowing case.

Constructing an Approximate Portfolio Policy

The ADP algorithm generates an approximate optimal policy along each sample path. In order to evaluate the overall quality of the approximate policy, as well as for practical applications of the algorithm, one must be able to extrapolate the approximate trading policy to arbitrary points in the state space. Clearly, there are many ways to perform such an extrapolation. We use a simple method, based on an idea of local averaging. In particular, we generate n_j equal intervals along the j^{th} dimension of the state space. Together with I equal intervals, this creates a total of $In_1n_2\cdots n_M$ $(M+1)$-dimensional rectangles, which cover the parts of the state space most likely to be visited by the state vector X_t. Then, if a point (t, X_t) falls inside one of the rectangles, we define the corresponding extrapolated value of the portfolio policy as an arithmetic average over the values at all the points used by the ADP algorithm that happen to fall inside the same rectangle. If a point (t, X_t) happens to be outside of the pre-defined rectangles, or if we cannot find a single path used in the ADP algorithm with a point inside the rectangle corresponding to (t, X_t), then we simply assign the value of the approximate portfolio policy based on a simple analytical approximation (e.g., a myopic approximation, see below). In our numerical experiments, we found that we end up using such an analytical approximation infrequently, typically less than 2% of the time.

5 Numerical Experiments

We perform several numerical experiments, in which we illustrate the performance of the ADP algorithm. Based on the results in Haugh, Kogan and Wang (2006), we know that the myopic portfolio policy often provides a good approximation to the optimal policy. The main promise of the ADP algorithm is that it can be used to obtain high-quality approximate solutions in situations where the optimal policy is far from myopic, i.e., the hedging component of the optimal policy is significant. With that in mind, we design some of our experiments so that, a priori, the myopic component of the optimal policy is small relative to the hedging component. As we shall see below, the ADP algorithm captures the hedging demand quite well. In order to further evaluate the performance of the ADP algorithm in our experiments, we compare the resulting lower and upper bounds on the value function with those corresponding to simple analytical approximation, the myopic policy defined below. In the numerical experiments in Haugh, Kogan and Wang (2006), the bounds based on the myopic policy are relatively tight. This indicates that for the portfolio choice problems considered there the hedging demand can be safely ignored, resulting in little loss of utility. Below we consider settings in which capturing the hedging demand is crucial, and the ADP algorithm is able to accomplish that with significant accuracy.

5.1 The Data-Generating Process

Recall that there are N risky securities and one short-term risk-free bond (a bank account) with M state variable processes driving their short term returns. We vary the numbers of assets and state variables in our experiments, but the general structure of the data-generating process is always the same.

We define all the processes in continuous time, and then generate the corresponding discrete-time dynamics using Euler approximations. Euler approximations add a certain amount of numerical error to the estimates of the upper bound on the value function of the discrete-time problem, since the theoretical results in Haugh, Kogan and Wang (2006) apply to continuous-time models. However, by shrinking the time between trading periods, we ensure that numerical approximation errors are quite small. The reported results where obtained by applying the ADP algorithm with a time-step equal to 1/15, and using a time-step of 1/100 in out-of-sample simulations to estimate the bounds on the true value function.

In continuous time, the vector of state variables X_t is characterized by a linear stochastic differential equation, driven by a J-dimensional vector of independent standard Brownian motions z_t:

$$dX_t = -K\, X_t\, dt + \sigma_X dz_t, \tag{17}$$

where X_t is an M by 1 vector, K is an M by J matrix, and σ_X is an M by J matrix.

We assume that the instantaneously risk-free rate is stochastic and given by $r_t = \delta_0 + \delta_1 X_t$, where δ_0 and δ_1 are constant. In order to define return processes for the risky assets, we first introduce their market price of risk:

$$\Lambda_t = \lambda_1 + \lambda_2 X_t, \tag{18}$$

where λ_1 is a constant N by 1 vector and λ_2 is a constant N by M matrix. Then, returns on the risky assets satisfy

$$\frac{dP_{it}}{P_{it}} = (r_t + \sigma_i \Lambda_t)\, dt + \sigma_i\, dz_t, \tag{19}$$

where σ_i is the 1 by N diffusion vector for asset i. We allow σ_i to be a deterministic function of time.

Our definition of risky assets is quite general and can be used to describe returns on both stocks and bonds. In fact, in some of our numerical experiments, we assume that the portfolio consists of both stocks and bonds and use this to calibrate our model.

Calibration

We now consider two special cases of our general data-generating process. Our first model corresponds to the market with two assets: a risk-free short-term bond (a bank account) and a risky long-term bond. We define parameter values so that the myopic component of the optimal policy is identically equal to zero. We then evaluate the ability of the ADP algorithm to recover the hedging component, which completely characterizes the optimal portfolio policy.

Model I: There are three state variables and one risky asset in our first model, therefore $M = 3$ and $N = 1$. The risky asset is a long-term bond maturing at time T. Given that the risk-free rate is an affine function of the state vector X_t, it follows that the market price of risk in (18) is also an affine function of the state vector X_t. The risk-neutral process for the risk-free rate is therefore Gaussian, and hence the term structure of interest rates is affine (e.g., Duffie 2001, Chapter 7). Specifically, the diffusion matrix of the long-term bond, σ_1, is an explicit function of time, given by

$$\begin{aligned} \sigma_1 &= b\,(1 - e^{-\tau\, M}), \\ b &= -\delta_1 Q^{-1}, \\ Q &= K + \sigma_X\, \lambda_2, \end{aligned} \tag{20}$$

where $e^{-\tau\, Q}$ denotes matrix exponentiation and $\tau = T - t$.

To calibrate the model, we adopt parameter values for the state-variable process from Wachter and Sangvinatsos (2005). We deviate in one respect: we set the market price of risk to zero, to guarantee that the myopic component of demand is identically equal to zero. Our parameter values are summarized in Table 1.

Model II: Our second model is a more realistic example, in which we construct an optimal portfolio of the risk-free short-term bond, two bonds of three- and ten-year maturity, and a stock index. The three- and ten-year bonds are in fact dynamic roll-over strategies, according to which one must constantly reinvest the funds to maintain the duration of the bonds at three and ten years respectively. Specifically, an investment at time t into a three-year bond maturing at $t + 3$ must be liquidated at time $t + \Delta t$ and reinvested in another bond, maturing at $t + \Delta t + 3$. Both stock and bond returns are predictable by a vector of state variables. The vector of state variables is now four-dimensional. Its first three components are the same as in our first model and are taken from Wachter and Sangvinatsos (2005). The fourth component represents the logarithm of the dividend yield on the NYSE value-weighted portfolio, which we use to predict returns on the stock market, the first risky asset in our model. We assume a Gaussian model for stock returns and, furthermore, we assume that the shocks driving changes in the dividend yield and stock returns are independent of the shocks to the first three components of the state vector and are not priced. This assumption is justified mostly by its convenience: since we are not striving to have an empirically accurate model of the joint behavior of stocks and bonds, the simplifying assumption allows us to use parameter estimates of Wachter and Sangvinatsos (2005) to describe the behavior of bonds in our model. The diffusion coefficients of bond returns can be computed according to (20). The diffusion coefficients of stock returns, denoted by σ_S, are give in Table 2, together with other model parameters.

5.2 Numerical Results

We first define formally the myopic portfolio policy, which we use to initiate the iterative solution procedure described in Section 4.1. We also compare the performance of the portfolio strategy generated by the ADP algorithm, captured by the resulting bounds on the maximal expected utility, to the bounds corresponding to the myopic strategy.

Let Σ_t denote the diffusion matrix of returns on the risky assets. Let μ_t denote the vector of instantaneous expected returns. Given our linear specification of the market prices of risk, $\mu_t = \mu_1 + \mu_2 X_t$. When no-borrowing constraints are imposed, we limit feasible portfolios to the set $\mathbf{K} = \{\imath^\top \theta_t \leq 1\}$. Then, the myopic portfolio policy is defined as

$$\theta_t^{myopic} = \arg\max_{\theta_t \in \mathbf{K}} \theta_t^\top (\mu_t - r_t) - \frac{1}{2} \theta_t^\top \Sigma_t \Sigma_t^\top \theta_t. \tag{21}$$

The results for Model I are summarized in Table 3. We choose a rather high value of risk aversion, $\gamma = 15$, to make sure that the myopic policy is far from the optimum. As we know from prior literature (e.g., Wachter, 2003), long-horizon investors with high risk aversion tend to gravitate towards holding a bond with maturity matching their investment horizon. Thus, for comparison, we include the results for a strategy of investing 100% of the portfolio into the long-term bond. As we see, the ADP algorithm produces a relatively tight set of bounds on the optimal expected utility. In contrast, the myopic policy performs rather poorly, as expected: due to the lack of risk premium, the myopic policy prescribes holding 100% of the portfolio in the short-term risk-free instrument, which is pretty much the opposite of what a long-horizon highly risk-averse investor would like to do. Another aspect of the results in Table 3 is that the duality-based method for estimating upper bounds on the value function performs well when the optimal solution is far from the myopic approximation. This provides further evidence of the method's potential.

The results for Model II are summarized in Table 4. Our calibration implies a significant degree of predictability in stock returns, therefore the certainty-equivalent returns achievable by the constructed portfolios are quite high. We see a clear pattern across the different risk aversion values: the ADP policy outperforms the myopic policy in terms of the achieved expected utility, which is expressed as a lower bound on the optimal value. The difference is not large, due to the fact that expected returns on the stock index in our model are quite volatile and capturing this predictability through a myopic policy seems to have a first-order affect on expected utility. Moreover, the variable predicting stock returns is not well hedged by any of the risky assets, diminishing the importance of the hedging component of demand. It is interesting that the upper bound obtained from the ADP policy is not always superior to the one based on the myopic policy. This must be due to the fact that to evaluate the upper bound one must approximate the partial derivatives of the value function with respect to the state variables. If such approximation turns out to be inaccurate, one can obtain tighter bounds by using the myopic strategy and simply setting the above partial derivatives to zero.

6 Conclusion

As our results indicate, the ADP algorithm is a viable tool for solving practically interesting portfolio choice problems. The method is accurate in the considered examples, whether the myopic policy is close to optimality or not. As our numerical experiments demonstrate, the accuracy of the ADP-based approximate solutions can be reliably gauged by the duality-based simulation method of HKW. Together, the two algorithms form a powerful set of tools which can be used to tackle difficult, high-dimensional problems.

Barberis, N. 2002. Investing for the Long Run When Returns are Predictable. *Journal of Finance.* **55** 225-264.

Brandt, M., A. Goyal, P. Santa-Clara, and R. Stroud. 2005. A Simulation Approach to Dynamic Portfolio Choice with an Application to Learning About Return Predictability. *Review of Financial Studies* **18**, 831-873.

Brennan, M. and Y. Xia. 2002. Dynamic Asset Allocation Under Inflation. *Journal of Finance.* **57** 1201-1238.

Brennan, M., E. Schwartz, and R. Lagnado. 1997. Strategic Asset Allocation. *Journal of Economic Dynamics and Control.* **21** 1377-1403.

Campbell, J. and L. Viceira. 1999. Consumption and Portfolio Decisions When Expected Returns are Time Varying. *The Quarterly Journal of Economics.* **114** 433-495.

Campbell, J. and L. Viceira. 2002. *Strategic Asset Allocation – Portfolio Choice for Long-Term Investors.* Oxford University Press, USA.

Campbell, J., Y. Chan and L. Viceira. 2003. A Multivariate Model of Strategic Asset Allocation. *Journal of Financial Economics.* **67** 41-80.

Chacko, G. and L. Viceira. 2005. Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets. *Review of Financial Studies.* **18** 1369-1402.

Cox, J. and C.F. Huang. 1989. Optimal Consumption and Portfolio Policies When Asset Prices Follow a Diffusion Process. *Journal of Economic Theory.* **49** 33-83.

Duffie, D. 2001. *Dynamic Asset Pricing Theory* . 3rd ed. Princeton University Press, USA.

Haugh, M., L. Kogan and J. Wang. 2006. Portfolio Evaluation: A Duality Approach. Forthcoming, *Operations Research.*

Karatzas, I., J. Lehoczky, S. Sethi and S. Shreve. 1986. Explicit Solution of a General Consumption/Investment Problem. *Mathematics of Operationas Research.* **11** 261-294.

Kim, T. and E. Omberg. 1986. Dynamic Nonmyopic Portfolio Behavior. *Review of Financial Studies.* **9** 141-61.

Merton, R. 1971. Optimum Consumption and Portfolio Rules in a Continuous-Time Model, *Journal of Economic Theory* **3**, 373-413.

Merton, R. 1990. *Continuous-Time Finance.* New York: Basil Blackwell.

Munk, C. 2000. Optimal Consumption/Investment Policies with Undiversifiable Income Risk and Liquidity Constraints. *Journal of Economic Dynamics and Control.* **24** 1315-1343.

Liu, J. 1998. Dynamic Portfolio Choice and Risk Aversion. Working paper. Stanford University, Palo Alto.

Longstaff, F. and E. Schwartz. 2001. Valuing American Options by Simulation: A Simple Least-Squares Approach. *Review of Financial Studies* **14** 113-147.

Tsitsiklis, J., and B. Van Roy. 2001. Regression Methods for Pricing Complex American–Style Options. *IEEE Transactions on Neural Networks* **12** 694-703.

Wachter, J. 2002. Portfolio and Consumption Decisions under Mean-Reverting Returns: An Exact Solution for Complete Markets. *Journal of Financial and Quantitative Analysis* **37** 63-91.

Wachter, J. 2003. Risk Aversion and Allocation to Long-Term Bonds. *Journal of Economic Theory* **112** 325-333.

Wachter, J. and A. Sangvinatsos. 2005. Does the Failure of the Expectations Hypothesis Matter for Long-Term Investors?, *Journal of Finance* **60** 179-230.

Parameter	Values		
K	0.5760	0	0
	0	3.3430	0
	-0.4210	0	0.0830
σ_X	1.0000	0	0
	0	1.0000	0
	0	0	1.0000
δ_0			0.0560
δ_1	0.0180	0.0070	0.010
$\lambda_1^\top$	0	0	0
λ_2	0	0	0
	0	0	0
	0	0	0

Table 1: Model I, calibrated parameters. The market consists of the risk-free asset and a long-term bond. The data-generating process is Gaussian. See Section 5.1 for details.

Parameter	Values				
K		0.5760	0	0	0
		0	3.3430	0	0
		-0.4210	0	0.0830	0
		0	0	0	0.0800
σ_X	1.0000	0	0	0	0
	0	1.0000	0	0	0
	0	0	1.0000	0	0
	0	0	0	-0.1985	0.3473
σ_S	-0.0126	0.0057	-0.0295	0.0143	0.000
δ_0					0.0560
δ_1		0.0180	0.0070	0.0100	0
$\lambda_1^\top$	-0.5630	-0.2450	-0.2190	0.4400	0
λ_2		0	1.7540	0	0
		0	-1.8150	0	0
		0.5370	0.3760	-0.0820	0
		0.1110	0.3050	-0.0170	0.200
		0	0	0	0

Table 2: Model II, calibrated parameters. The market consists of the risk-free asset, two long-term bonds, and a stock index. Stock returns are predictable and the process for the predictive variable is independent of the processes driving the term structure of interest rates. The data-generating process is Gaussian. See Section 5.1 for details.

	Incomplete Markets
LB^{ADP}	5.22
	[5.22,5.23]
UB^{ADP}	5.53
	[5.53,5.54]
LB^{m}	4.42
	[4.41,4.42]
UB^{m}	5.59
	[5.58,5.60]
LB^{LT}	5.51
	[5.51,5.51]
UB^{LT}	5.53
	[5.52,5.53]
	No-borrowing
LB^{ADP}	5.29
	[5.28,5.29]
UB^{ADP}	5.67
	[5.65,5.69]
LB^{m}	4.42
	[4.41,4.42]
UB^{m}	6.87
	[6.80,6.94]
LB^{LT}	5.51
	[5.51,5.51]
UB^{LT}	5.58
	[5.57,5.59]

Table 3: Model I, portfolio performance. The parameter set is defined in Table 1 and the problem horizon is $T = 5$ years. The risk aversion coefficient is $\gamma = 15$. All results are computed for the initial value of the state variable $X_0 = 0$. The rows marked LB^{ADP} and UB^{ADP} report the estimates of the bounds on the expected utility achieved by using the ADP portfolio strategy. The ADP algorithm uses fourth-order approximations and relies on a myopic policy for initial approximation and extrapolation. The rows marked LB^{m} and UB^{m} report the corresponding results based on the myopic portfolio strategy. Expected utility is reported as a continuously compounded certainty equivalent return. Approximate 95% confidence intervals are reported in brackets. The rows marked LB^{LT} and UB^{LT} report the bounds based on the policy of holding all of the portfolio in a long-term bond maturing at time T. We report the results both for incomplete markets and the case of no-borrowing constraints.

	$\gamma = 1.5$	$\gamma = 3$	$\gamma = 5$
		Incomplete Markets	
LB^{ADP}	56.85	34.73	24.26
	[56.78,56.91]	[34.66,34.81]	[24.18,24.34]
UB^{ADP}	60.43	37.86	28.77
	[60.31,60.56]	[37.45,38.28]	[27.85,29.73]
LB^m	56.95	34.39	23.80
	[56.88,57.03]	[34.33,34.45]	[23.75,23.85]
UB^m	60.43	37.55	27.84
	[60.31,60.56]	[37.14,37.98]	[26.82,28.90]
		No-borrowing	
LB^{ADP}	31.72	19.72	15.23
	[31.67,31.77]	[19.69,19.76]	[15.14,15.32]
UB^{ADP}	33.47	21.59	16.58
	[33.39,33.55]	[21.41,21.76]	[16.34,16.83]
LB^m	30.57	19.46	14.69
	[30.52,30.61]	[19.44,19.49]	[14.67,14.71]
UB^m	33.09	21.19	15.91
	[33.01,33.18]	[21.01,21.38]	[15.57,16.26]

Table 4: Model II, portfolio performance. The parameter set is defined in Table 2 and the problem horizon is $T = 5$ years. The risk aversion coefficient takes values $\gamma = 1.5, 3, 5$. All results are computed for the initial value of the state variable $X_0 = 0$. The rows marked LB^{ADP} and UB^{ADP} report the estimates of the bounds on the expected utility achieved by using the ADP portfolio strategy. The ADP algorithm uses fourth-order approximations and relies on a myopic policy for initial approximation and extrapolation. The rows marked LB^m and UB^m report the corresponding results based on the myopic portfolio strategy. Expected utility is reported as a continuously compounded certainty equivalent return. Approximate 95% confidence intervals are reported in brackets. We report the results both for incomplete markets and the case of no-borrowing constraints.

A Upper bound on expected utility: a quadratic subproblem

In this section we derive explicit solutions to constrained quadratic optimization problems involved in estimating an upper bound on the optimal expected utility. See HKW for a general formulation and definitions related to the dual formulation. We consider here the case of incomplete markets and no-borrowing constraints.

We start with a simpler case of Model I. The quadratic programming problem that needs to be solved repeatedly during Monte Carlo simulations is

$$\begin{aligned} \min \quad & \frac{1}{2}||\tilde{\Lambda} - \hat{\Lambda}^{\nu}||_2 \\ s.t \quad & \sigma_1 \hat{\Lambda}^{\nu} = \sigma_1 \Lambda + \nu \\ & \nu \text{ is feasible} \end{aligned}$$

Above, $\tilde{\Lambda}$ is the candidate for a market price of risk in a fictitious market, obtained from an approximate value function produced by the ADP algorithm (see HKW for definitions). $\hat{\Lambda}^{\nu}$ is the market price of risk in a fictitious complete market, which we use to compute an upper bound on the expected utility. ν is a scalar in this case, which parameterizes feasible fictitious markets.

When markets are incomplete, the only feasible value of ν is zero. We are thus faced with a simple projection problem,

$$\begin{aligned} \min \quad & \frac{1}{2}||\tilde{\Lambda} - \hat{\Lambda}^{\nu}||_2 \\ s.t \quad & \sigma_1 (\hat{\Lambda}^{\nu} - \Lambda) = 0 \end{aligned}$$

which has an explicit solution:

$$\hat{\Lambda}^{\nu} = \tilde{\Lambda} - \left[(\sigma_1 \sigma_1^{\top})^{-1} \sigma_1 (\tilde{\Lambda} - \Lambda) \right] \sigma_1^{\top}. \tag{22}$$

In the case of no-borrowing constraints, the feasible set is $\nu \leq 0$. Therefore, we are solving

$$\begin{aligned} \min \quad & \frac{1}{2}||\tilde{\Lambda} - \hat{\Lambda}^{\nu}||_2 \\ s.t \quad & \sigma_1 (\hat{\Lambda}^{\nu} - \Lambda) \leq 0. \end{aligned}$$

We are thus faced with two possibilities. If $\sigma_1(\tilde{\Lambda} - \Lambda) \leq 0$, then $\hat{\Lambda}^{\nu} = \tilde{\Lambda}$. Otherwise, $\hat{\Lambda}^{\nu}$ is given by (22).

We now consider Model II. The quadratic subproblem has a similar form. Let Σ denote the diffusion matrix of returns on risky assets and define variables $y = \hat{\Lambda}^{\nu} - \tilde{\Lambda}$ and $b = \Sigma(\Lambda - \tilde{\Lambda})$. Then, the quadratic subproblem takes form

$$\begin{array}{ll} \min & \frac{1}{2}||y||_2 \\ s.t. & \Sigma y - \nu \imath = b \\ & \nu \text{ is feasible} \end{array}$$

$\imath$ is a vector of ones of the same dimension as the number of risky assets and, again, ν is a constant. When markets are incomplete, $\nu = 0$ and the optimization problem reduces to

$$\begin{array}{ll} \min & \frac{1}{2}||y||_2 \\ s.t. & \Sigma y = b \end{array}$$

which has the solution

$$y = \Sigma^{\top}(\Sigma\Sigma^{\top})^{-1}b.$$

In the case of no-borrowing constraints, the feasible set is given by $\nu \leq 0$. We are now facing two distinct cases. Relaxing the equality constraints with Lagrange multipliers π, we see that if the system of linear equations

$$\begin{array}{rcl} y - \Sigma^{\top}\pi & = & 0 \\ \imath^{\top}\pi & = & 0 \\ \Sigma y - \nu \imath & = & b \end{array}$$

has a solution with $\nu \leq 0$, then we have an optimal y. Otherwise, we must set $\nu = 0$, and the optimal y is given by the solution for the case of incomplete markets.