A Simulation Approach to Dynamic Portfolio Choice with an Application to Learning About Return Predictability

Michael W. Brandt
Fuqua School of Business, Duke University, and NBER

Amit Goyal
Goizueta Business School, Emory University

Pedro Santa-Clara
The Anderson School, UCLA, and NBER

Jonathan R. Stroud
The Wharton School, University of Pennsylvania

We present a simulation-based method for solving discrete-time portfolio choice problems involving non-standard preferences, a large number of assets with arbitrary return distribution, and, most importantly, a large number of state variables with potentially path-dependent or non-stationary dynamics. The method is flexible enough to accommodate intermediate consumption, portfolio constraints, parameter and model uncertainty, and learning. We first establish the properties of the method for the portfolio choice between a stock index and cash when the stock returns are either iid or predictable by the dividend yield. We then explore the problem of an investor who takes into account the predictability of returns but is uncertain about the parameters of the data generating process. The investor chooses the portfolio anticipating that future data realizations will contain useful information to learn about the true parameter values.

After years of relative neglect in the academic literature, dynamic portfolio choice is again attracting attention. The renewed interest in the topic follows the recent empirical evidence of return predictability and is fueled by practical issues, including parameter and model uncertainty, learning, transaction costs, taxes, and background risks.[1] Unfortunately, the realism of portfolio choice problems considered until now has been

We thank Doron Avramov, John Birge, Michael Brennan, John Cochrane, René Garcia, Jun Liu, Francis Longstaff, Ross Valkanov, and Yihong Xia for helpful discussions. We are also grateful to Campbell Harvey (the editor) and two anonymous referees for their detailed comments and suggestions. Financial support from the Rodney L. White Center for Financial Research at the Wharton School is gratefully acknowledged. Address correspondence to Amit Goyal, Goizueta Business School, Emory University, 1300 Clifton Rd., Atlanta, GA 30322-2722 or e-mail: Amit_Goyal@bus.emory.edu.

[1] See Campbell and Viceria (2002) and Brandt (2003) for surveys of the dynamic portfolip choice literature.

© The Author 2005. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oupjournals.org
doi:10.1093/rfs/hhi019 Advance Access publication May 25, 2005

constrained by the scarcity of analytic results and by the limited power of existing numerical methods.

In this article, we present a simulation-based method for solving more realistic discrete-time portfolio choice problems involving non-standard preferences and a large number of assets with arbitrary return distribution. The main advantage of our approach relative to other numerical solution methods is that it can handle problems in which the conditional return distribution depends on a large number of state variables with potentially path-dependent or non-stationary dynamics. The method can also accommodate intermediate consumption, portfolio constraints, parameter and model uncertainty, and learning.

The original literature on dynamic portfolio choice, pioneered by Merton (1969, 1971) and Samuelson (1969) in continuous time and by Fama (1970) in discrete time, produced many important insights into the properties of optimal portfolio policies. Unfortunately, closed-form solutions are available only for a few special parameterizations of the investor's preferences and return dynamics, as exemplified by Kim and Omberg (1996), Liu (1999a), and Wachter (2002). The recent literature therefore uses a variety of numerical and approximate solution methods to incorporate realistic features into the dynamic portfolio problem. For example, Brennan, Schwartz, and Lagnado (1997) solve numerically the PDE characterizing the solution to the dynamic optimization. Campbell and Viceira (1999) log-linearize the first-order conditions (FOCs) and budget constraint to obtain approximate closed-form solutions. Das and Sundaram (2000) and Kogan and Uppal (2000) perform different expansions of the value function for which the problem can be solved analytically. By far the most popular approach involves discretizing the state space, which is done by Balduzzi and Lynch (1999), Brandt (1999), Barberis (2000), and Dammon, Spatt, and Zhang (2001), among many others. Once the state space is discretized, the value function can be evaluated by a choice of quadrature integration (Balduzzi and Lynch), simulations (Barberis), binomial discretizations (Dammon, Spatt, and Zhang), or nonparametric regressions (Brandt), and then the dynamic optimization can be solved by backward recursion.

These numerical and approximate solution methods share some important limitations. Except for the nonparametric approach of Brandt (1999), they assume unrealistically simple return distributions. All of the methods rely on constant relative risk aversion (CRRA) preferences, or its extension by Epstein and Zin (1989), to eliminate the dependence of the portfolio policies on wealth and thereby make the problem path-independent. Most importantly, the methods cannot handle the large number of state variables with complicated dynamics which arise in many realistic portfolio choice problems. A partial exception is Campbell, Chan, and Viceira (2003), who use log-linearization to solve a problem with many state variables but linear dynamics.

Our simulation method overcomes these limitations. The first step of the method entails simulating a large number of hypothetical sample paths of asset returns and state variables. We simulate these paths from the known, estimated, or bootstrapped joint dynamics of the returns and state variables. Alternatively, we simulate the sample paths from the investor's posterior belief about the joint distribution of the returns and state variables to incorporate parameter and model uncertainty in a Bayesian context. The key feature of the simulation stage is that the joint dynamics of the asset returns and state variables can be high-dimensional, arbitrarily complicated, path-dependent, and even non-stationary.

Given the set of simulated paths of returns and state variables, we solve for the optimal portfolio (and consumption) policies recursively in a standard dynamic programming fashion. Starting one period before the end of the investor's horizon, at time $T-1$, we compute for each simulated path the portfolio weights that maximize a Taylor series expansion of the investor's expected utility. This approximated problem has a straightforward (semi-) closed-form solution involving conditional (on the state variables) moments of the utility function, its derivatives, and the asset returns. We compute these conditional moments with least squares regressions of the realized utility, its derivatives, and the asset returns at time T on basis functions of the realizations of the state variables at time $T-1$ *across* the simulated sample paths. The algorithm then proceeds recursively until time zero. Each period and for each simulated path, we find the portfolio weights that maximize the conditional expectation of the investor's utility, given the optimal portfolios for all future periods until the end of the horizon. To summarize, our method consists of simulating the asset returns and state variables, computing a set of across-path regressions for each period, and then evaluating the closed-form solution of the approximate portfolio problem.

Our approach is inspired by the simulation method of Longstaff and Schwartz (2001) for pricing American-style options. Longstaff and Schwartz use regressions across simulated sample paths to evaluate the conditional expectation of the continuation value of the option and compare this expectation to the immediate exercise value at all future dates along each simulated path. We adopt the idea of using across-path regressions to evaluate conditional expectations but take these expectations as inputs to the portfolio optimization along each path. Our approach is also related to the "parameterized expectations" method of den Haan and Marcet (1990) for solving dynamic macroeconomic models.[2] The difference is that we use regressions *across* a large number of sample paths instead of

[2] Despite the name, the innovation of that method is not to parameterize expectations but to evaluate parameterized expectations through regressions on simulated data. The idea of parameterizing expectations goes back to Bellman, Kalaba, and Kotkin (1963), Daniel (1976), and Williams and Wright (1984).

along a single path. The across-path regressions allow us to solve finite-horizon problems, which may be path-dependent and non-stationary, as opposed to only stationary infinite-horizon problems. Finally, our discrete-time approach complements the simulation method of Detemple, Garcia, and Rindisbacher (2003) for solving continuous-time portfolio choice problems in complete markets. Although the methods are complements in their applications, they are conceptually different. Detemple, Garcia, and Rindisbacher rely on the assumption of complete markets to express the dynamic problem as a static one [as in Cox and Huang (1989) and Aït-Sahalia and Brandt (2003)] and then use simulations to evaluate expectations in standard Monte Carlo fashion. We use regressions across simulations to solve recursively a dynamic programming problem.

We first apply our method to the portfolio choice between a stock index and cash when the stock returns are either iid or predictable by the dividend yield of the index. We use this simple setup to analyze the potential sources of error in our approach. In particular, we use the iid returns case to establish that the Taylor series expansion of the expected utility induces only minimal approximation errors. (These results are of independent interest for static portfolio problems with higher-order moments.) We use the predictable returns case, which has been studied by Brennan, Schwartz, and Lagnado (1997), Balduzzi and Lynch (1999), Brandt (1999), Campbell and Viceira (1999), and Barberis (2000), among others, to show that our method delivers the same solution as more traditional approaches. To further assess the numerical accuracy of our method, we also perform a Monte Carlo experiment. We find that the simulation-induced errors of the optimal portfolio weights and the corresponding loss in the certainty equivalent return on wealth are negligible.

As a more challenging and also economically more interesting application of our method, we then explore the problem of an investor who takes into account the predictability of returns but is uncertain about the parameters of the data generating process. The investor chooses the portfolio anticipating the effect of learning about the true parameter values from each new data realization between the initial portfolio choice and the end of the investment horizon. Specifically, we consider the same problem as Barberis (2000) of a Bayesian investor who faces returns which appear marginally predictable by the dividend yield. Unlike Barberis, however, we do not assume that between the initial decision at date 0 and the terminal date T the investor's beliefs about the parameters are unchanged and therefore reflect only the information available at the initial decision. Instead, we allow the investor to update these beliefs and thereby learn about the true parameter values with each new return and dividend yield realization between dates 0 and T.

Learning in discrete time is a challenging application, because it involves a large number of state variables with non-linear and possibly

non-stationary dynamics. For example, the usual VAR model for dividend yield predictability involves seven parameters, and, given uninformative or conjugate priors, the joint posterior distribution of these parameters is characterized by 11 state variables (the dividend yield and, in standard regression notation, the unique elements of $X'X$, $X'Y$, and $(Y-X(X'X)^{-1}X'Y)(Y-X(X'X)^{-1}X'Y)'$). The evolution of the state variables depends on the newly observed return, dividend yield, their squares, and cross-product. It is therefore clearly non-linear. For more elaborate model specifications, such as when the degree of return predictability is allowed to fluctuate, or with economically motivated non-conjugate priors in the spirit of Black and Litterman (1992) and Connor (1997), the posterior distribution of returns is characterized by even more state variables and the dynamics of these state variables may become non-stationary.

Given the proliferation of state variables, fully incorporating learning in a discrete-time portfolio choice problem has until now been deemed computationally infeasible. Brennan (1998) and Barberis (2000) examine the effect of learning about the unconditional risk premium of stocks but ignore the evidence of predictability. They find that for a CRRA investor who is more risk averse than the log utility case, learning about the unconditional risk premium can lead to substantial negative hedging demands (i.e., deviations from the myopic portfolio choice). A long-horizon investor who anticipates learning about the unconditional risk premium in the future allocates substantially less wealth to equities than an investor who is either myopic or does not learn. Xia (2001) studies the effect of learning about predictability with a known unconditional risk premium and a known unconditional mean of the predictive variable.[3] Her results show that learning about predictability can also induce hedging demands but that the sign and magnitude of these hedging demands depend on the horizon and current value of the predictive variable.

Unfortunately, it is difficult, if not impossible, to infer from these existing results the effect of *simultaneously* learning about the unconditional risk premium, predictability, and the moments of the predictor. Intuitively, a new data realization provides information about the unconditional risk premium, the difference between the conditional and unconditional risk premium, or both. Updating on the difference between the conditional and unconditional risk premium, in turn, can be due to learning about predictability or due to learning about the unconditional mean of the predictive variable and hence about the difference between the current value of the predictor and its mean (if the current value of the predictor is further from its updated mean, the conditional risk premium is further from the unconditional risk premium). Furthermore, since the majority of the literature on

[3] Related work on learning about a time-varying risk premium include Detemple (1986), Dothan and Feldman (1986), and Gennotte (1986).

learning is set in continuous time where second moments are typically assumed known, the effect of learning about second moments on a discrete-time portfolio choice is still unexplored. The aim of our application is to start filling these gaps in the literature. This article is, to our knowledge, the first to provide a solution to a reasonably realistic discrete-time portfolio choice problem with learning about all parameters of the return-generating process.

Our empirical results show that both the parameter uncertainty and the effect of learning tend to reduce the investor's allocation to stocks. We find that learning about the parameters of the data generating process induces a negative hedging demand for stocks. The intuition for this hedging demand is that a positive return innovation leads to an upward revision of the unconditional and/or conditional expected return (depending on the value of the dividend yield), which constitutes a subjective improvement in future investment opportunities. By reducing the holdings of stocks, the investor can hedge this effect. In our experiment, the negative hedging demand due to learning tends to dominate the positive hedging demand induced by the dividend yield predictability. This shows that learning can qualitatively change the solution of the dynamic portfolio choice problem. Our results also suggest that partially incorporating learning is better, in terms of certainty equivalent return on wealth, than ignoring this aspect of the portfolio choice problem.

The remainder of this article is structured as follows. We describe the basic algorithm in Section 1. In Section 2, we discuss some implementation issues and extensions. The applications of the method are in Section 3. We conclude in Section 4.

1. The Method

1.1 The investor's problem

Consider the portfolio choice at time t of an investor who maximizes the expected utility of wealth at some terminal date T by trading in N risky assets and a risk-free asset (cash) at times $t, t+1,\ldots T-1$.[4] Formally, the investor's problem is

$$V_t(W_t,Z_t) = \max_{\{x_s\}_{s=t}^{T-1}} \mathrm{E}_t[u(W_T)], \tag{1}$$

subject to the sequence of budget constraints

$$W_{s+1} = W_s\left(x_s' R_{s+1}^e + R^f\right) \ \forall s \geq t \tag{2}$$

[4] We discuss the case of intermediate consumption in Section 2.1.

where x_s is a vector of portfolio weights on the risky assets chosen at time s, R^e_{s+1} is the vector of *excess* returns on the N risky assets from time s to $s+1$, and R^f is the gross return on the risk-free asset (assumed constant for simplicity).[5] The function $u(\cdot)$ measures the investor's utility of terminal wealth W_T and the subscript on the expectation denotes that it is conditional on the information set Z_t available at time t. For concreteness, we can think of the information set Z_t as a vector of state variables.

The intertemporal portfolio choice is a dynamic problem. At time t, the investor chooses the portfolio weights x_t conditional on having wealth W_t and information Z_t. In this decision, the investor takes into account the fact that at every future date s the portfolio weights will be optimally revised conditional on the then available wealth W_s and information Z_s.[6]

The function $V_t(W_t, Z_t)$ denotes the expectation at time t, conditional on the information Z_t, of the utility of terminal wealth W_T generated by the current wealth W_t and the subsequent *optimal* portfolio weights $\{x_s{}^*\}_{s=t}^{T-1}$. $V_t(\cdot,\cdot)$ is termed the value function. It represents the value, in units of expected utils, of the portfolio choice problem to the investor. Another interpretation of the value function is that it measures the investor's investment opportunities. If the current information suggests that investment opportunities are good, meaning that the sequence of optimal portfolio choices is expected to generate an above-average return on wealth with below-average risk, the current value of the portfolio choice problem to the investor is high. If investment opportunities are poor, the value is low.

The dynamic nature of the portfolio choice is best understood by expressing the multiperiod problem (1) as a single-period problem with utility $V_{t+1}(W_{t+1}, Z_{t+1})$ of next period's wealth W_{t+1} and information Z_{t+1}:

$$\begin{aligned} V_t(W_t,Z_t) &= \max_{\{x_s\}_{s=t}^{T-1}} \mathrm{E}_t[u(W_T)] \\ &= \max_{x_t} \mathrm{E}_t\left\{ \max_{\{x_s\}_{s=t+1}^{T-1}} \mathrm{E}_{t+1}[u(w_T)] \right\} \\ &= \max_{x_t} \mathrm{E}_t\left\{ V_{t+1}\left[W_t\left(x_t' R^e_{t+1} + R^f\right), Z_{t+1}\right]\right\}. \end{aligned} \tag{3}$$

The second equality follows from the law of iterated expectations, and the third equality uses the definition of the value function as well as the budget constraint. It is important to recognize that the expectation in the third line is taken over the *joint* distribution of next period's returns R_{t+1} and information Z_{t+1}, conditional on the current information Z_t.

[5] We can relax the constant risk-free rate assumption by including R_t^f in the vector of state variables Zt.

[6] The information Z_s contains the investor's wealth W_s, which means that conditioning on both quantities is technically redundant. However, the investor's wealth is an unusual state variable, because it is endogenous to the portfolio choice. We therefore consider it separately from the exogenous information.

The recursive equation (3) is the so-called Bellman equation and is the basis for any recursive solution of the dynamic portfolio choice problem. The FOCs for an optimum at each time are

$$\mathrm{E}_t\left\{\partial_1 V_{t+1}\left[W_t\left(x_t' R_{t+1}^e + R^f\right), Z_{t+1}\right] R_{t+1}^e\right\} = 0, \tag{4}$$

where $\partial_1 V_{t+1}(\cdot,\cdot)$ denotes the partial derivative with respect to the first argument of the value function. These FOCs make up a system of nonlinear equations involving (possibly highorder) integrals that can in general be solved for x_t only numerically.

For illustrative purposes, consider the case of CRRA utility $u(W_T) = W_T^{1-\gamma}/(1-\gamma)$, where γ denotes the coefficient of relative risk aversion. The Bellman equation then simplifies to

$$\begin{aligned} V_t(W_t, Z_t) &= \max_{x_t} \mathrm{E}_t\left[\max_{\{x_s\}_{s=t+1}^{T-1}} \mathrm{E}_{t+1}\left(\frac{W_T^{1-\gamma}}{1-\gamma}\right)\right] \\ &= \max_{x_t} \mathrm{E}_t\left(\max_{\{x_s\}_{s=t+1}^{T-1}} \mathrm{E}_{t+1}\left\{\frac{\left[W_t \prod_{s=t}^{T-1}\left(x_s' R_{s+1}^e + R^f\right)\right]^{1-\gamma}}{1-\gamma}\right\}\right) \\ &= \max_{x_t} \mathrm{E}_t\left(\underbrace{\frac{\left[W_t\left(x_t' R_{t+1}^e + R^f\right)\right]^{1-\gamma}}{1-\gamma}}_{u(W_{t+1})} \underbrace{\max_{\{x_s\}_{s=t+1}^{T-1}} \mathrm{E}_{t+1}\left\{\left[\prod_{s=t+1}^{T-1}\left(x_s' R_{s+1}^e + R^f\right)\right]^{1-\gamma}\right\}}_{\psi_{t+1}(Z_{t+1})}\right) \end{aligned} \tag{5}$$

With CRRA utility, the value function next period $V_{t+1}(W_{t+1}, Z_{t+1})$ can be expressed as the product of the utility of wealth $u(W_{t+1})$ and a function of the state variables $\psi_{t+1}(Z_{t+1})$. Furthermore, since the utility function is homothetic in wealth we can, without loss of generality, normalize W_t=1. It follows that the value function depends only on the state variables and that the Bellman equation is

$$\frac{1}{1-\gamma}\psi_t(Z_t) = \max_{x_t} \mathrm{E}_t\left[u\left(x_t' R_{t+1}^e + R^f\right)\psi_{t+1}(Z_{t+1})\right] \tag{6}$$

The corresponding FOCs are

$$\mathrm{E}_t\left[\partial u\left(x_t' R_{t+1}^e + R^f\right)\psi_{t+1}(Z_{t+1}) R_{t+1}^e\right] = 0 \tag{7}$$

which, despite being simpler than the general case, can still only be solved numerically.

This CRRA utility example also helps to illustrate the difference between the dynamic and myopic (single-period) portfolio choice. If the excess returns R_{t+1}^e are contemporaneously independent of the innovations to the state variables Z_{t+1}, the $(T-t)$-period portfolio choice is identical to the single-period

portfolio choice at date t because the conditional expectation in equation (6) factors into a product of two conditional expectations. The first expectation is of the utility of next period's wealth $u(W_{t+1})$, and the second expectation is of the function of the state variables ψ_{t+1} (Z_{t+1}). Since the latter does not depend on the portfolio weights, the FOCs of the multiperiod problem are the same as those of the single-period problem. If, in contrast, the excess returns are not independent of the innovations to the state variables, the conditional expectation does not factor, the FOCs are not the same, and, as a result, the dynamic portfolio choice may be substantially different from the single-period portfolio choice. The differences between the dynamic and myopic policies are called *hedging demands* because by deviating from the single-period portfolio choice the investor tries to hedge against changes in the investment opportunities (or, equivalently, in the state variables).

1.2 Step 1: Expanding the value function

We can simplify the problem significantly by expanding the value function in a Taylor series around the current value of wealth growing at the risk-free rate $W_t R^f$. In the case of a second-order expansion, the approximated value function is[7]

$$V_t(W_t, Z_t) \approx \max_{x_t} \mathrm{E}_t\big[V_{t+1}\big(W_t R^f, Z_{t+1}\big) + \partial_1 V_{t+1}\big(W_r R^f, Z_{t+1}\big)\big(W_t x_t' R^e_{t+1}\big) + \tfrac{1}{2}\,\partial_1^2 V_{t+1}\big(W_t R^f, Z_{t+1}\big)\big(W_t x_t' R^e_{t+1}\big)^2\big], \tag{8}$$

where $\partial_1^2 V_{t+1}(\cdot,\cdot)$ denotes the second partial derivative with respect to the first argument of the value function. In this case, the FOCs have a closed-form solution in terms of the joint conditional moments of the value function and the returns in the next period:

$$\begin{aligned} x_t &\approx -\big\{\mathrm{E}_t\big[\partial_1^2 V_{t+1}\big(W_t R^f, Z_{t+1}\big)\big(R^e_{t+1} R^e_{t+1}{}'\big)\big] W_t^2\big\}^{-1} \times \mathrm{E}_t\big[\partial_1 V_{t+1}\big(W_t R^f, Z_{t+1}\big) R^e_{t+1}\big] W_t \\ &\approx -[\mathrm{E}_t(B_{t+1}) W_t]^{-1} \times \mathrm{E}_t(A_{t+1}). \end{aligned} \tag{9}$$

This approximate solution for the optimal weights involves two conditional expectations. The first expectation is the second moment matrix of returns (essentially the covariance matrix) scaled by the second derivative of the value function next period. The second expectation involves the risk premia of the assets scaled by the first derivative of the value function.

Unfortunately, a second-order expansion of the value function is sometimes not sufficiently accurate, especially when the utility function departs significantly from quadratic utility or when the returns are far from

[7] There is an extensive theoretical and empirical literature on approximating utility and value functions with polynomial expansions, including Samuelson (1970), Hakansson (1971), Grauer (1981), Pulley (1981), Kroll, Levy, and Markowitz (1984), Grauer and Hakansson (1993), and Zhao and Ziemba (2000).

Gaussian. We therefore work with a fourth-order expansion that includes adjustments for the skewness and kurtosis of returns and their effects on the utility of the investor. The approximation of the value function is then

$$\begin{aligned} V_t(W_t,Z_t) \approx \max_{x_t} \mathrm{E}_t \Big[& V_{t+1}(W_t R^f, Z_{t+1}) \\ & + \partial_1 V_{t+1}(W_t R^f, Z_{t+1})(W_t x_t' R^e_{t+1}) \\ & + \frac{1}{2}\partial_1^2 V_{t+1}(W_t R^f, Z_{t+1})(W_t x_t' R^e_{t+1})^2 \\ & + \frac{1}{6}\partial_1^3 V_{t+1}(W_t R^f, Z_{t+1})(W_t x_t' R^e_{t+1})^3 \\ & + \frac{1}{24}\partial_1^4 V_{t+1}(W_t R^f, Z_{t+1})(W_t x_t' R^e_{t+1})^4 \Big]. \end{aligned} \tag{10}$$

In this case, the FOCs define only an *implicit* solution for the optimal weights in terms of moments of the value function and the returns in the next period:

$$\begin{aligned} x_t \approx & -\left\{ \mathrm{E}_t\left[\partial_1^2 V_{t+1}(W_t R^f, Z_{t+1})(R^e_{t+1} R^{e\prime}_{t+1})\right] W_t^2 \right\}^{-1} \\ & \times \left\{ \mathrm{E}_t\left[\partial_1 V_{t+1}(W_t R^f, Z_{t+1})(R^e_{t+1})\right] W_t \right. \\ & + \frac{1}{2}\mathrm{E}_t\left[\partial_1^3 V_{t+1}(W_t R^f, Z_{t+1})(x'_t R^e_{t+1})^2 R^e_{t+1}\right] W_t^3 \\ & \left. + \frac{1}{6}\mathrm{E}_t\left[\partial_1^4 V_{t+1}(W_t R^f, Z_{t+1})(x'_t R^e_{t+1})^3 R^e_{t+1}\right] W_t^4 \right\} \\ \approx & -[\mathrm{E}_t[B_{t+1}]W_t]^{-1} \times \left\{ \mathrm{E}_t(A_{t+1}) + \mathrm{E}_t[C_{t+1}(x_t)]W_t^2 \right. \\ & \left. + \mathrm{E}_t[D_{t+1}(x_t)]W_t^3 \right\} \end{aligned} \tag{11}$$

It is easy to solve this implicit expression for the optimal weights in practice. We start by computing the portfolio weights for the second-order expansion of the value function and take this to be an initial "guess" for the optimal weights, denoted $x_t(0)$. We then enter this guess on the right-hand side of equation (11) and obtain a new solution for the optimal weights on the left-hand side, denoted $x_t(1)$. After a few iterations n, the guess $x_t(n)$ is very close to the solution $x_t(n+1)$, and we take this value to be the solution of equation (11).[8]

Consider again the case of CRRA utility with $V_{t+1}(W_{t+1},Z_{t+1}) = u(W_{t+1})\psi_{t+1}(Z_{t+1})$. The solution for the second-order expansion of the value function is

[8] This approximate solution is of independent interest for static portfolio choice problems when taking into account higher order moments is deemed important.

$$
\begin{aligned}
x_t &\approx -\left\{\mathrm{E_t}\left[\partial^2 u(W_t R^f)\psi_{t+1}(Z_{t+1})(R^e_{t+1}R^{e'}_{t+1})\right]W_t\right\}^{-1} \\
&\quad \times \mathrm{E}_t\left[\partial u(W_t R^f)\psi_{t+1}(Z_{t+1})R^e_{t+1}\right] \\
&\approx -\frac{\partial u(W_t R^f)}{\partial^2 u(W_t R^f)W_t}\left\{\mathrm{E_t}\left[\psi_{t+1}(Z_{t+1})(R^e_{t+1}R^{e'}_{t+1})\right]\right\}^{-1} \times \mathrm{E}_t\left[\psi_{t+1}(Z_{t+1})R^e_{t+1}\right] \\
&\approx \frac{1}{\gamma}\left\{\mathrm{E}_t\left[\psi_{t+1}(Z_{t+1})(R^e_{t+1}R^{e'}_{t+1})\right]\right\}^{-1} \times \mathrm{E}_t\left[\psi_{t+1}(Z_{t+1})R^e_{t+1}\right], \qquad (12)
\end{aligned}
$$

where the second line follows from the fact that $u(W_t R^f)$ and its derivatives are deterministic, and the third line uses the definition of relative risk aversion γ.

Equation (12) illustrates more clearly the relation between the dynamic and myopic portfolio policies. If the returns are contemporaneously independent of the innovations to the state variables, both expectations factor into two terms, with one term in common, and, as a result, the portfolio weights reduce to the familiar mean-variance form[9]:

$$
\begin{aligned}
x_t &\approx -\frac{1}{\gamma}\left\{\mathrm{E}_t\left[\psi_{t+1}(Z_{t+1})\right]\mathrm{E}_t\left(R^e_{t+1}R^{e'}_{t+1}\right)\right\}^{-1} \times \mathrm{E}_t\left[\psi_{t+1}(Z_{t+1})\right]\mathrm{E}_t\left(R^e_{t+1}\right) \\
&\approx \frac{1}{\gamma}\left[\mathrm{E}_t\left(R^e_{t+1}R^{e'}_{t+1}\right)\right]^{-1} \times \mathrm{E}_t\left(R^e_{t+1}\right). \qquad (13)
\end{aligned}
$$

Otherwise, the dynamic and myopic portfolio choices differ. The extent to which they differ depends on the contemporaneous correlation of the excess returns, squared returns, and cross products of returns with the innovations to the state variables.

This first step of our method involves two choices: the order of the expansion and the wealth level around which to expand the value function. In theory, the expansion converges as the order increases for any feasible expansion point, suggesting that we want the order of the expansion to be as large as possible and that the expansion point is fairly arbitrary. However, in practice, where we face a trade-off between the accuracy of the approximation and the computational efficiency of our method, these two choices are intimately related. For a low-order expansion to be relatively accurate, the expansion point should be as close as possible to the realized wealth next period. Since little more is known about the accuracy of low-order expansions, we have to rely on experimentation to determine whether a given expansion order and expansion point combination is acceptable for the assumed horizon, rebalancing

[9] The interpretation of the approximate solution in the mean-variance framework is unique to the second-order approximation. Higher-order approximations, such as equation (11), involve higher-order moments.

frequency, return distribution, and utility function. In our experience, a fourth-order expansion around $W_t R^f$ is very accurate for the various problems we have considered (Section 3). Furthermore, the method appears much more sensitive to the order of the expansion than to the expansion point. Indeed, the results from second- and fourth-order expansions can be quite different, whereas expanding the value function around wealth growing at the expected return of the myopic portfolio policy gives virtually identical results to expanding around the wealth growing at the risk-free rate.

1.3 Step 2: Simulating sample paths

The second step of our solution method consists of simulating forward a large number of hypothetical sample paths of the vector $Y_t = \left(R_t^e, Z_t\right)$. Each of these paths is simulated independently from the model:

$$Y_{t+1} = f(Y_t, Y_{t-1}, \ldots; \in_{t+1}), \tag{14}$$

where $\in_{t+1}$ denotes a vector of innovations with conditional distribution $g(\in_{t+1} \mid \in_t, \in_{t-1}, \ldots)$.

If we assume that the investor knows the true form and parameters of the data generating process, the model $f(\cdot;\cdot)$ is either calibrated to or estimated from the data. The innovations $\in_t$ are sampled from a known distribution (such as a multivariate lognormal distribution) or are resampled from the data (bootstrapped). If we relax the unrealistic assumption of a known data generating process, the Y_t are simulated from the joint posterior distribution of the returns and state variables, given the data and the investor's prior distribution over the model and parameter space (Section 3.3 for detail). In either case, we enumerate the simulated paths with $m = 1, 2, \ldots, M$ and denote the realization of the returns and state variables at time s along the mth path $Y_s^m = \left(R_s^{em}, Z_s^m\right)$.

1.4 Step 3: Computing expectations through regressions

We solve recursively for the optimal portfolio weights at each date t for each simulated sample path m. Assume that the optimal portfolio weights from time $t+1$ to $T-1$ have already been computed and are denoted by $\hat{x}_s$, $s = t+1, \ldots, T-1$. We approximate the terminal wealth as the current wealth growing at the risk-free rate for one period [equation (8)] and subsequently growing at the return generated by the optimal portfolio policy to get

$$\hat{W}_T = W_t R^f \prod_{s=t+1}^{T-1} \left(\hat{x}_s' R_{s+1}^e + R^f\right) \tag{15}$$

Equation (9) for the portfolio weights involves conditional expectations at time t of A_{t+1} and B_{t+1}. A_{t+1} and B_{t+1} themselves involve conditional expectations at time $t+1$ as follows:

$$\begin{aligned} A_{t+1} &= \partial_1 V_{t+1}\left(W_t R^f, Z_{t+1}\right) R^e_{t+1} \\ &= \mathrm{E}_{t+1}\left[\partial u\left(\hat{W}_T\right) \prod_{s=t+1}^{T-1} \left(\hat{x}_s' R^e_{s+1} + R^f\right)\right] R^e_{t+1} \end{aligned} \tag{16}$$

$$\begin{aligned} B_{t+1} &= \partial_1^2 V_{t+1}\left(W_t R^f, Z_{t+1}\right) R^e_{t+1} R^{e'}_{t+1} \\ &= \mathrm{E}_{t+1}\left[\partial^2 u\left(\hat{W}_T\right) \prod_{s=t+1}^{T-1} \left(\hat{x}_s' R^e_{s+1} + R^f\right)^2\right] R^e_{t+1} R^{e'}_{t+1} \end{aligned} \tag{17}$$

Note that the expressions inside the conditional expectation operator can be easily calculated for each simulated path.

Consider now the problem of solving for the current portfolio weights, $\hat{x}_t$ given the current wealth W_t^m and the realization of the state variables Z_t^m. The portfolio weight from equation (9) is given by[10]

$$\begin{aligned} x_t &\approx -\left[\mathrm{E}_t(B_{t+1}) W_t\right]^{-1} \times \mathrm{E}_t(A_{t+1}) \\ &= -\left(\mathrm{E}_t\left\{\mathrm{E}_{t+1}\left[\partial^2 u\left(\hat{W}_T\right) \prod_{s=t+1}^{T-1} \left(\hat{x}_s' R^e_{s+1} + R^f\right)^2\right] R^e_{t+1} R^{e'}_{t+1}\right\} W_t\right)^{-1} \\ &\quad \times \mathrm{E}_t\left\{\mathrm{E}_{t+1}\left[\partial u(\hat{W}_T) \prod_{s=t+1}^{T-1} \left(\hat{x}_s' R^e_{s+1} + R^f\right)\right] R^e_{t+1}\right\} \\ &= -\left\{\mathrm{E}_t\left[\partial^2 u\left(\hat{W}_T\right) \prod_{s=t+1}^{T-1} \left(\hat{x}_s' R^e_{s+1} + R^f\right)^2 R^e_{t+1} R^{e'}_{t+1}\right] W_t\right\}^{-1} \\ &\quad \times \mathrm{E}_t\left[\partial u\left(\hat{W}_T\right) \prod_{s=t+1}^{T-1} \left(\hat{x}_s' R^e_{s+1} + R^f\right) R^e_{t+1}\right] \\ &= -\left[\mathrm{E}_t(b_{t+1}) W_t\right]^{-1} \times \mathrm{E}_t(a_{t+1}) \end{aligned} \tag{18}$$

where the second equality follows from the law of iterated expectations. In the last equality, a and b denote the realized values that correspond to A and B. Thus, we can directly compute the portfolio weights from conditional expectations of a_{t+1} or b_{t+1}, which in turn are functions of the investor's utility of terminal wealth, the future optimal portfolio policy, and the return paths. In this way, we avoid writing the optimal portfolio solution as a function of the derivatives of the value function next period.

[10] We use the second-order expansion in the description of the method for expositional ease. The approach extends with obvious modifications to the fourth-order expansion that we actually use in section 3.

The benefit of this is that any errors in evaluating the value function do not propagate in the backward recursions (though there could still be errors in estimating the portfolio weights x_t). The task ahead reduces to evaluating the conditional expectations of the expressions inside the square brackets in equation (18).

We parameterize these expectations as functions of the state variables:

$$\mathrm{E}_t(y_{t+1}) = \varphi(Z_t)'\theta_t \tag{19}$$

where y_{t+1} stands for a generic element of a_{t+1} or b_{t+1}, $\varphi(Z_t)$ denotes a vector of polynomial bases in Z_t, and θ_t are parameters to be estimated. For simplicity, we use as bases a simple power series in Z_t[11]:

$$\varphi(Z_t)' = [1 Z_t \; Z_t^2 \cdots] \tag{20}$$

The key feature of the parameterized expectation (19) is its linearity in the (nonlinear) functions of the state variables. This linearity implies that we can evaluate the parameters θ_t through a simple least squares regression of the realized values y_{t+1}^m at time $t+1$ against the polynomial bases $\varphi(Z_t^m)$ at time t *across* the simulated sample paths.[12] The fitted values of this regression, denoted $\hat{y}_{t+1}^m$, are then used to construct estimates of the conditional expectations of $A_{t+1|t}$ and $B_{t+1|t}$, denoted $\hat{A}_{t+1|t}^m$ and $\hat{B}_{t+1|t}^m$, respectively. These estimates of the conditional expectations, in turn, yield estimates of the optimal portfolio weights at time t for each path m:

$$\hat{x}_t^m = -\left(\hat{B}_{t+1|t}^m W_t^m\right)^{-1} \hat{A}_{t+1|t}^m. \tag{21}$$

From a practical perspective, it is important to recognize that for each date t we only need to compute the fitted values for a *single* set of regressions (one for each unique element of the matrices $A_{t+1|t}$ and $B_{t+1|t}$) which means not only that we can afford a large number of simulations to control the simulation error ($M = 10{,}000$ in our applications) but also that we can vectorize the algorithm to avoid nested loops (in t and m).

Although it is not explicitly required to do so for the calculation of portfolio weight, we can also evaluate the value function at date t and for all simulated paths m. Recall that the value function at time t is the conditional expectation of the final utility of wealth under the sequence of optimal portfolio choices at dates $t, t+1, \ldots, T-1$:

$$V_t(W_t, Z_t) = \mathrm{E}_t\left[u\left(\hat{W}_T\right)\right] = \mathrm{E}_t\left\{u\left[W_t \prod_{s=t}^{T-1}\left(\hat{x}_s' R_{s+1}^e + R^f\right)\right]\right\}, \tag{22}$$

[11] We elaborate on the choice of polynomial bases in Section 2.3.

[12] The idea of using across-path regressions to evaluate conditional expectations in a simulation setting was introduced byLongstaff and Schwartz (2001) in the context of pricing American-style options.

where $\hat{x}_s$ denotes the optimal portfolio weights at date s. We evaluate the conditional expectation in equation (22) using across-path regressions. Specifically, we evaluate the value function at date t by regressing the utility realized at the end of each sample path, denoted $u\left(\hat{W}_T^m\right)$, from the estimated sequence of optimal portfolio weights $\{\hat{x}_s^m\}_{s=t}^{T-1}$ against the polynomial bases of the state variables $\varphi\left(Z_t^m\right)$.

The algorithm proceeds recursively backward. A noteworthy feature of the recursions is that the errors in the approximation of the value function propagate (and cumulate) through the backward recursions only to the extent that the approximation errors in the portfolio weights affect the expected utility of terminal wealth. However, it is well known that even first-order deviations in the optimal portfolio policy have only second-order welfare effects [Cochrane (1989)]. As a result, the propagation and accumulation of the approximation error due to the Taylor series expansion of the value function is kept to a minimum.

2. Extensions and Implementation Issues

2.1 Intermediate consumption

Our simulation method extends to the case of intermediate consumption, where the investor chooses each period the level of consumption and the asset allocation for the wealth that is not consumed. Assuming additive time-separable preferences, the value function is

$$V_t(W_t,Z_t) = \max_{\{x_s,c_s\}_{s=t}^{T}} \mathrm{E}_t\left[\sum_{s=t}^{T} \beta^{s-t} u(c_s W_s)\right]$$

$$= \max_{x_t,c_t} u(c_t W_t) + \beta E_t[V_{t+1}(W_{t+1}, Z_{t+1})], \quad (23)$$

subject to the sequence of budget constraints:

$$W_{s+1} = (1 - c_s) W_s \left(x_s' R_{s+1}^e + R^f\right) \ \forall s \geq t \quad (24)$$

and the terminal condition $c_T = 1$, where c_s denotes the *fraction* of wealth consumed at time s, and β is a subjective discount factor.

Analogous to the case without intermediate consumption, we expand the value function at time $t+1$ around a deterministic wealth of $(1-\overline{c})W_t R^f$. We let $\overline{c}$ be the optimal consumption for a deterministic problem in which the investor's wealth grows at the risk-free rate for the remaining $T-t$ periods (effectively setting $x_s = 0$, for $s = t,\ t+1,\ldots,T-1$).

A second-order expansion of the value function is

$$V_t(W_t,Z_t) \approx \max_{x_t,c_t} u(c_t W_t) + \beta \mathrm{E}_t V_{t+1}\Big\{ \big[(1-\overline{c})W_t R^f, Z_{t+1}\big] + \partial_1 V_{t+1}\big[(1-\overline{c})W_t R^f, Z_{t+1}\big]\big[(1-c_t)W_t x_t' R^e_{t+1} - (c_t-\overline{c})W_t R^f\big] + \frac{1}{2}\partial_1^2 V_{t+1}\big[(1-\overline{c})W_t R^f, Z_{t+1}\big]\big[(1-c_t)W_t x_t' R^e_{t+1} - (c_t-\overline{c})W_t R^f\big]^2 \Big\} \tag{25}$$

Solving the implied FOCs for the optimal portfolio and consumption choices yields

$$\begin{aligned} x_t &\approx \big\{\mathrm{E}_t\big[\partial_1^2 V_{t+1}\big[(1-\overline{c})W_t R^f, Z_{t+1}\big]\big(R^e_{t+1}R^{e\prime}_{t+1}\big)\big](1-c_t)W_t\big\}^{-1} \\ &\quad \times \big(\mathrm{E}_t\big\{\partial_1^2 V_{t+1}\big[(1-\overline{c})W_t R^f, Z_{t+1}\big]R^e_{t+1}\big\}(c_t-\overline{c})W_t R^f \\ &\quad - \mathrm{E}_t\big\{\partial_1 V_{t+1}\big[(1-\overline{c})W_t R^f, Z_{t+1}\big]R^e_{t+1}\big\}\big) \\ &\approx \{\mathrm{E}_t[C_{t+1}](1-c_t)W_t\}^{-1} \times \big\{\mathrm{E}_t[B_{t+1}](c_t-\overline{c})W_t R^f - \mathrm{E}_t[A_{t+1}]\big\} \end{aligned} \tag{26}$$

and

$$\begin{aligned} c_t &\approx \left\{\mathrm{E}_t\left[\partial_1^2 V_{t+1}\big[(1-\overline{c})W_t R^f, Z_{t+1}\big]\left(x_t' R^e_{t+1} + R^f\right)^2\right]W_t\right\}^{-1} \\ &\quad \times \left(\mathrm{E}_t\left\{\partial_1^2 V_{t+1}\big[(1-\overline{c})W_t R^f, Z_{t+1}\big]\left(x_t' R_{t+1} + \overline{c}R^f\right)\left(x_t' R^e_{t+1} + R^f\right)\right\}W_t \right. \\ &\quad \left. - \frac{1}{\beta}\partial u(c_t W)_t + \mathrm{E}_t\left\{\partial_1 V_{t+1}\big[(1-\overline{c})W_t R^f, Z_{t+1}\big]\left(x_t' R^e_{t+1} + R^f\right)\right\}\right) \\ &\approx \{\mathrm{E}_t[F_{t+1}(x_t)]W_t\}^{-1} \times \left\{\mathrm{E}_t[D_{t+1}(x_t)W_t] - \frac{1}{\beta}\partial u(c_t W_t) + \mathrm{E}_t[E_{t+1}(x_t)]\right\} \end{aligned} \tag{27}$$

Similar expressions can be obtained for a fourth-order expansion of the value function.

Equations (26) and (27) must be solved simultaneously.[13] However, in our experience, the following iterative procedure is quite effective. Starting with an initial "guess" for $c_t(0)$, we solve for the portfolio weights $x_t(1)$ from equation (26). We then use these portfolio weights to solve for the consumption choice $c_t(1)$ from equation (27). After a few iterations n, the guesses

[13] A notable exception is the case of CRRA utility, for which the portfolio choice is independent of the consumption choice. The CRRA consumption choice, however, still depends on the portfolio choice.

$\{x_t(n), c_t(n)\}$ are very close to the solutions $\{x_t(n+1), c_t(n+1)\}$, and we take these values to be the solutions to the system of two equations (26) and (27).

2.2 Alternative objective functions and portfolio constraints

Except with CRRA utility, the portfolio choice depends on the investor's current wealth. The current wealth, in turn, depends on the sequence of past returns and past portfolio choices, which are unknown at the current time because we solve the problem recursively backward. To overcome this problem, we recover at each time t and for each path m the full *mapping* from the current wealth W_t^m to the portfolio choice by solving the problem for a *grid* of wealth levels ranging from a lower bound of $\underline{W_t}$ to an upper bound of $\overline{W_t}$. For each current wealth level on this grid, we construct the realized y_t^m for the across-path regressions by interpolating the portfolio choices at all future dates $s = t+1, \ldots, T-1$, which depend on the future wealth realizations, from the sequence of mappings between the wealth and portfolio choice recovered in the previous recursions. Notice that since the wealth resulting from the optimal portfolio choices grows over time, the upper and lower bounds of the wealth grid must also be changed in order to maintain a certain coverage of the wealth distribution at each time. The best way to accomplish this depends on the particular application at hand.[14]

Using this approach, our method can handle virtually any objective function, as long as it is four times differentiable. This includes portfolio choice problems in which the investor derives utility from consumption or wealth in excess of a reference level [Samuelson (1989), Dybvig (1995),Jagannathan and Kocherlakota (1996), and Schroder and Skiadas (1999)] and problems in which the investor exhibits behavioral biases such as loss aversion, ambiguity aversion, or disappointment aversion [Benartzi and Thaler (1995), Liu (1999b), Aït-Sahalia and Brandt (2001), Ang, Bekaert, and Liu (2002), Gomes (2002)]. Our method can also be applied to problems of more practical interest, such as maximizing Sharpe or information ratios, beating or tracking a benchmark, controlling draw-downs, or maintaining a certain value-at-risk Roy (1952), Grossman and Vila (1989), Browne (1999), Basak and Shapiro (2001), Tepla (2001), Alexander and Baptista (2002)].

We can also incorporate portfolio constraints, such as limits on borrowing or short sales. These constraints are important in practice for a variety of investors, including private individuals, mutual funds, and pension plans. To incorporate portfolio constraints, we solve the approximate problem at each future date along each simulated path subject to the constraints. Since the constrained problems typically do not have a closed-form solutions, they must be solved numerically. For this, we rely on an extensive literature

[14] Suppose each period the log returns on the optimal portfolio are approximately normally distributed with mean μp and volatility σ_{p}. An intuitive parmeterization of the upper and lower bounds in this case is $\mathrm{W}_0\exp\{\mu_{\mathrm{p}}t \pm k\sigma_{\mathrm{p}}\sqrt{\mathrm{t}}\}$ for some constant k, where, for example, k = 2.6 yields a 99 percent coverage of the wealth distribution at time t given the initial wealth at time 0.

on effective and fast algorithms for solving high-dimensional constrained optimization problems, including variants of the Newton method [Conn, Gould, and Toint (1988),Moré and Toraldo (1989)], the quasi-Newton or BFGS method [Byrd et al. (1995)], and the sequential quadratic programming approach [Gill, Murray, Saunders (2002)].[15]

2.3 Using regressions for conditional expectations

Longstaff and Schwartz (2001) andTsitsiklis and Van Roy (2001) provide partial convergence results for the use of across-path regressions to value American-style options. These results are extended byClément, Lamberton, and Potter (2001), who show that for a fixed order of the polynomial bases, the Longstaff–Schwartz estimator (of the option price) is normally distributed around the true projection of the price onto the polynomial bases with convergence rate $\sqrt{M}$. Their theorems can be adapted to our setting to show that the portfolio weights from our method converge to the optimal weights of the second- or fourth-order expansions of the investor's utility function.

There are many basis functions we can use for evaluating the conditional expectations, including Chebyshev, Hermite, Laguerre, and Legendre orthogonal polynomials. Longstaff and Schwartz (2001) report numerical evidence that even a simple power series is effective and that the order of the polynomial need not be very high for reliable option pricing. We present similar results in Section 3.1 for the dynamic portfolio choice problem. In any case, increasing the order of the polynomial just increases the number of regressors, which is not a concern because we can simultaneously increase the number of sample paths to maintain the same level of numerical accuracy. For large numbers of state variables, we can follow the suggestions of Judd (1998) and use *complete* polynomials, for which the number of basis functions increases linearly rather than exponentially with the number of state variables.

We need to be somewhat concerned about numerical issues though. It is well understood that OLS regressions are susceptible to outliers, and our regressions are potentially plagued by outliers, because the curvature of the utility function tends to amplify the spread of the terminal wealth. Even sequences of relatively normal returns sometimes result in extreme realizations of the terminal utility and its derivatives. We can address this outliers problem either through robust regression estimators, such as an M-estimator or through more simple α-trimmed OLS regressions, where some fraction α of the extreme observations in both tails are omitted from the OLS regression. It is encouraging, however, that we have encountered serious outliers problems only at extremely long horizons (because the variance of the realized terminal wealth grows approximately linearly with the horizon).

[15] See also the survey of constrained optimization algorithms byConn, Gould, and Toint (1994).

2.4 Computation speed

Despite the use of simulations, our method is relatively fast. In the applications we describe below, it takes less than a couple of minutes to solve a 40-period portfolio choice problem with one risky asset and as many as 10 state variables on a standard PC with a Pentium 1.8-MHz processor and 512.MB of RAM.

3. Applications

3.1 Investing in stocks with IID returns

We first apply our method to the static (or single-period) portfolio choice between a stock index and cash of a CRRA investor with relative risk aversion γ ranging from five to 20 and a holding period ranging from one month to one year. The purpose of this application is to quantify the approximation error due to the Taylor series expansion of the value function. We assume a constant risk-free rate of six percent per year and model the excess stock return as iid lognormal with a mean and volatility that match the sample moments of the historical data from January 1986 through December 1995 on the value-weighted CRSP index. The sample moments are provided for reference in the caption of Table 1. The sample is chosen to match the period studied by Barberis (2000) to make our results comparable to his.

Table 1 summarizes in the columns marked x_*, the *exact* solution of this problem. These columns show the optimal fraction of wealth allo-

Table 1
Error due to the expansion of the value function

γ	x_*	x_2	x_4	x_*	x_2	x_4
		Monthly			Quarterly	
5	0.7592	0.7303 (0.48)	0.7580 (0.00)	0.7484	0.6693 (3.56)	0.7378 (0.06)
10	0.3795	0.3651 (0.23)	0.3791 (0.00)	0.3738	0.3347 (1.74)	0.3701 (0.02)
20	0.1897	0.1826 (0.12)	0.1895 (0.00)	0.1867	0.1673 (0.85)	0.1852 (0.01)
	Mean	0.73%		Mean	2.22%	
	SD	4.42%		SD	7.91%	
		Semi-annual			Annual	
5	0.8059	0.6461 (13.15)	0.7575 (1.21)	1.0000	0.7112 (56.55)	0.8726 (18.14)
10	0.4024	0.3231 (6.41)	0.3843 (0.33)	0.5734	0.3556 (31.30)	0.4556 (9.13)
20	0.2008	0.1615 (3.13)	0.1933 (0.12)	0.2859	0.1778 (15.35)	0.2329 (3.66)
	Mean	4.43%		Mean	8.70%	
	SD	11.03%		SD	13.55%	

This table summarizes the approximation error due to the Taylor series expansion of the utility function for a single-period portfolio choice between a stock index and cash. We consider a constant relative risk aversion investor with relative risk aversion γ and a monthly, quarterly, semi-annual, or annual holding period. The risk-free rate is constant at six percent per year and the excess stock returns iid log-normally distributed with mean and volatility. x_* denotes the exact fraction of wealth allocated to stocks obtained by solving the investor's first-order condition using quadrature integration. x_2 and x_4 denote the approximate solutions from second- and fourth-order expansions of the utility function, respectively. In parentheses is the loss in the certainty equivalent return on wealth due to the approximation, quoted in annualized basis points.

cated to stocks obtained by solving the investor's FOC using quadrature integration.[16] The columns marked x_2 and x_4 show the approximate allocations from a second- and fourth-order Taylor series expansion of the utility function. The sub-panels summarize results for different holding periods, and the rows correspond to different levels of risk aversion. Finally, the numbers in parentheses are the losses in the certainty equivalent return on wealth, quoted in annualized basis points, because of the error in the approximation of the value function.

The results show that the magnitude and economic significance of the approximation error depend on both the return horizon and the curvature of the utility function. Consider first the horizon effect. The second-order approximation leads to little expected utility loss at the monthly horizon (0.12 to 0.48 basis points) but a substantial loss at the annual horizon (as much as 57 basis points). The intuition for this result is that the wealth distribution is more spread out at longer horizons, because the return variance increases approximately linearly with the horizon, causing the higher-order terms of the expansion to be more important. The effect of changing risk aversion is less straightforward. At all horizons, the expected utility loss due to the approximation *decreases* with the level of risk aversion, which is counter-intuitive considering that the curvature of the utility function increases with risk aversion. The reason for this result is that the increase in the curvature of the utility function is only one of two offsetting effects. The other effect is that the investor allocates substantially less wealth to stocks as risk aversion increases (76 percent with $\gamma = \text{five}$ versus 19 percent with $\gamma = 20$ at the monthly horizon). This causes the wealth distribution to be less spread out and, as a result, reduces the importance of the higher-order terms of the expansion. The variance of wealth effect dominates the curvature of the utility function effect, causing the accuracy of the second-order approximation to increase with the level of risk aversion.

Table 1 also summarizes that the fourth-order expansion of the utility function performs substantially better than the second-order expansion. For the monthly, quarterly, and semi-annual holding periods, the difference between the exact and fourth-order approximate allocations is less than 0.05 in magnitude (in most cases even less than 0.01), irrespective of the level of risk aversion, and the associated expected utility loss is negligible. Even for the annual holding period, the difference between the exact and approximate solutions is in all cases less than 0.15 with an expected utility loss of at most 20 basis points.

[16] Although the returns are assumed to be iid, the optimal portfolio weights are not constant across different holding periods. This is because we calibrate the long-horizon return distributions to the corresponding long-horizon sample moments rather than scale the monthly sample moments using the iid assumption. With scaled moments, the horizon-irrelevance result of Merton (1969) and Samuelson (1969) holds exactly.

We conclude from this first application that the second-order expansion of the value function is satisfactory for short holding periods and high levels of risk aversion. The fourth-order expansion performs uniformly better than the second expansion and, more importantly, results in an accurate approximation of the true solution in all cases considered in this article.

3.2 Investing in stocks with predictable returns

We next consider the dynamic (or multiperiod) portfolio choice of a CRRA investor between a stock index and cash when the stock returns are predictable by the dividend yield of the index. This stylized asset allocation problem has received much attention recently [Brennan, Schwartz, and Lagnado (1997), Balduzzi and Lynch (1999), Brandt (1999), Campbell and Viceira (1999), and Barberis (2000)], because it illustrates clearly the differences between the dynamic (or multiperiod) and myopic (or single-period) portfolio choices. The main goal of this application is to establish the properties of our solution method relative to the more traditional approach of discretizing the state space.

We assume the following restricted VAR as the *quarterly* data generating process:

$$\begin{bmatrix} r^e_{t+1} \\ d_{t+1} - p_{t+1} \end{bmatrix} = \begin{bmatrix} 0.227 \\ (0.95) \\ -0.155 \\ (-0.79) \end{bmatrix} + \begin{bmatrix} 0.060 \\ (0.87) \\ 0.958 \\ (17.02) \end{bmatrix} (d_t - p_t) + \begin{bmatrix} \in_{1,t+1} \\ \in_{2,t+1} \end{bmatrix} \tag{28}$$

$$\begin{bmatrix} \in_{1,t} \\ \in_{2,t} \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.0060 & -0.0051 \\ -0.0051 & 0.0049 \end{bmatrix} \right),$$

where $r^e{}_t$ is the log excess return on the value weighted CRSP index and d_t–p_t denotes the log dividend yield of the index, computed from the sum of the past 12 monthly dividends and the current level of the index. The VAR is estimated using quarterly data from January 1986 to December 1995, which again matches the sample of Barberis (2000). In parentheses are Newey and West (1987) adjusted t-statistics.

The restricted VAR in equation (28) is very similar to the model used by Campbell and Viceira (1999) and Barberis (2000). The predictability along with the extreme persistence of the dividend yield induce slow mean-reversion in returns and thereby cause long-horizon returns to be less variable than iid returns [Barberis (2000)]. Furthermore, the strong negative correlation between the return and dividend yield innovations (correlation of $-.95$) induces substantial differences between the myopic and dynamic portfolio policies (hedging demands). The investor can hedge against changes in the dividend yield, which are associated with

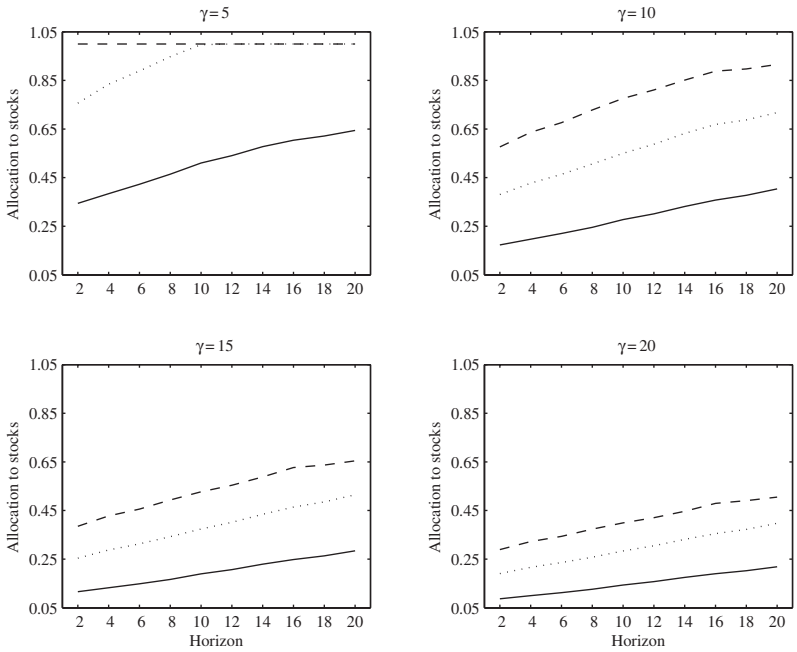

Figure 1
Horizon effect
This figure presents the portfolio choices between a stock index and cash of a constant relative risk aversion investor with relative risk aversion γ ranging from five to 20 as a function of the investment horizon. The portfolios are rebalanced quarterly and the investment horizon, $T-t$ ranges from two quarters to five years. The risk-free rate is constant at six percent per year, and the excess stock returns and log dividend yield evolve as a first-order VAR with parameters given in equation (28). The dashed, dotted, and solid lines represent the optimal allocation of wealth to stocks conditional on the dividend yield being equal to one SD below, at, one SD above the historical average.

changes in future expected returns, by over-allocating wealth to the stock index relative to the myopic portfolio choice [Campbell and Viceira (1999) and Barberis (2000)].

Figure 1 plots the optimal portfolio weight as a function of the investment horizon for γ ranging from five to 20. The dashed, dotted, and solid lines represent the optimal allocation of wealth to stocks conditional on the dividend yield being equal to one SD below, at, or one SD above the historical average, respectively.[17] The horizon ranges from two quarters to five years (20 quarters). The results are based on 10,000 simulated sample paths and a fourth-order expansion of the value function.

Reminiscent of the results of Barberis (2000), the allocation to stocks increases with the horizon. The differences between the quarterly and

[17] The historical average (3.00%) dividend yield is higher than the unconditional average (2.38%) implied by VAR coefficients. We chose to condition on historical averages to be consistent with Barberis (2000).

long-horizon allocations represent the hedging demand or the excess demand for stocks to hedge against unexpected changes in the dividend yield and (through the dividend yield) in future expected returns. The fraction of wealth invested in stocks reaches a steady-state level at a horizon of five to ten years. This steady-state solution corresponds to an infinitely lived investor and is therefore comparable to the results of Campbell and Viceira (1999). The allocation to stocks decreases both with increasing risk aversion and with decreasing values of the dividend yield.

Table 2 examines in more detail the optimal portfolio choice in the case of $\gamma = 10$. The table describes three competing portfolio policies: 'U' denotes the unconditional policy, where the fraction of wealth invested in stocks is chosen according to the unconditional return distribution; 'M' is the myopic policy, where each period the allocation is conditioned on the dividend yield assuming only a single-period horizon; and 'D' denotes the optimal dynamic portfolio policy. For each policy, the table reports the fraction of wealth allocated to stocks and the certainty equivalent return on wealth from following that policy.[18] To assess the numerical accuracy of our solution, we repeat the algorithm 1000 times using different sets of simulations. We report in parentheses the SDs of the portfolio weights and the certainty equivalents of wealth across these repetitions.

Relative to the unconditional portfolio policy, the myopic policy allocates less (more) wealth to stocks when the dividend yield is below (above) its historical mean.[19] Conditioning the investment decision on the dividend yield generates an impressive gain in expected utility. For a five-year horizon, the gain is as much as an annualized 2.6 percent difference in the certainty equivalent return on wealth or an overall 13.69 percent difference in the certainty equivalent of wealth over the five-year horizon. Furthermore, conditioning on the dividend yield in a dynamic instead of myopic fashion produces substantial differences in the portfolio weights. For instance, at the five-year horizon the difference between the dynamic and myopic stock holdings is almost 40 percent. The expected utility gain from investing optimally versus myopically is of the same order of magnitude as the gain from the myopic conditional versus unconditional portfolio choice. For a five-year horizon, the certainty equivalent return on wealth increases by as much as 0.7 percent per

[18] We also calculate the certainty equivalent return on wealth from following the policy in an independent set of simulations or "out-of-sample." To apply the portfolio policy out-of-sample, we first linearize it by regressing the portfolio weights on the state variables across the sample paths. The certainty equivalent expected utility gain computed with the independent set of simulations is within a few basis points of the in-sample expected utility gain. These results are available upon request.

[19] The difference between the two policies when the dividend yield is equal to its historical mean is due to the fact that the unconditional variance of stock returns is greater than the conditional variance.

Table 2
Dynamic portfolio policies with dividend yield predictability

	−1 SD			Mean			+1 SD		
$T-t$	U	M	D	U	M	D	U	M	D
A. Portfolio weights									
2	0.0847 (0.0000)	0.1627 (0.0025)	0.1746 (0.0041)	0.0847 (0.0000)	0.3578 (0.0065)	0.3812 (0.0090)	0.0847 (0.0000)	0.5444 (0.0100)	0.5761 (0.0130)
4	0.0847 (0.0000)	0.1627 (0.0025)	0.2011 (0.0055)	0.0847 (0.0000)	0.3578 (0.0065)	0.4310 (0.0108)	0.0847 (0.0000)	0.5444 (0.0100)	0.6405 (0.0149)
8	0.0847 (0.0000)	0.1627 (0.0025)	0.2549 (0.0075)	0.0847 (0.0000)	0.3578 (0.0065)	0.5161 (0.0133)	0.0847 (0.0000)	0.5444 (0.0100)	0.7385 (0.0167)
20	0.0847 (0.0000)	0.1627 (0.0025)	0.4128 (0.0114)	0.0847 (0.0000)	0.3578 (0.0065)	0.7213 (0.0152)	0.0847 (0.0000)	0.5444 (0.0100)	0.9136 (0.0156)
40	0.0847 (0.0000)	0.1627 (0.0025)	0.5862 (0.0117)	0.0847 (0.0000)	0.3578 (0.0065)	0.8544 (0.0131)	0.0847 (0.0000)	0.5444 (0.0100)	0.9692 (0.0139)
B. Certainly equivalent return on wealth									
2	6.37% (0.00%)	6.47% (0.02%)	6.47% (0.02%)	6.72% (0.00%)	7.49% (0.04%)	7.48% (0.03%)	7.08% (0.00%)	9.26% (0.07%)	9.25% (0.07%)
4	6.37% (0.00%)	6.55% (0.02%)	6.55% (0.02%)	6.71% (0.00%)	7.59% (0.04%)	7.59% (0.04%)	7.s05% (0.00%)	9.39% (0.07%)	9.39% (0.07%)
8	6.37% (0.00%)	6.68% (0.02%)	6.69% (0.03%)	6.69% (0.00%)	7.76% (0.04%)	7.79% (0.05%)	7.00% (0.00%)	9.55% (0.07%)	9.66% (0.09%)
20	6.37% (0.00%)	6.95% (0.02%)	7.12% (0.04%)	6.62% (0.00%)	7.96% (0.03%)	8.36% (0.08%)	6.87% (0.00%)	9.52% (0.04%)	10.25% (0.10%)
40	6.36% (0.00%)	7.08% (0.01%)	7.61% (0.04%)	6.53% (0.00%)	7.84% (0.01%)	8.68% (0.05%)	6.71% (0.00%)	8.94% (0.02%)	10.01% (0.05%)

This table analyses the portfolio choice between a stock index and cash of a constant relative risk aversion investor with relative risk aversion $\gamma = 10$ when the stock returns are predictable by the dividend yield of the index. The portfolios are rebalanced quarterly, and the investment horizon $T - t$ ranges from two quarters to 10 years. The risk-free rate is constant at six percent per year, and the excess stock returns and log dividend yield evolve as a first-order VAR with parameters given in equation (28). We present results for three current realizations of the dividend yield one SD below, at, one SD above the historical average. For each dividend yield and horizon, we evaluate three policies. 'U' is the unconditional policy, 'M' is the myopic conditional policy, and 'D' is the dynamic conditional policy. The first row in each block is the average across 1000 Monte Carlo replications. The second row is the SD of across all the Monte Carlo iterations. Panel A is for the portfolio weight in the stock index. Panel B is for the certainty equivalent return on wealth (annualized) from following the portfolio policy in-sample. The results are based on 10,000 simulated sample paths and a fourth-order expansion of the value function.

year, which translates into an overall 3.55 percent difference in the certainty equivalent of wealth over the five-year horizon.

Table 2 also provides evidence on the precision of our numerical solution method. The standard errors on the portfolio weights computed from the Monte Carlo experiment are always less than 1.67 percent in magnitude and represent no more than 1/15th of the portfolio weight. The standard errors on the certainty equivalent of wealth are negligible with the maximum loss in wealth being only 10 basis points.

Table 3 provides a direct comparison of our solution method to the more traditional dynamic programming approach of discretizing the state space used by Barberis (2000) among others. We discretize the unconditional distribution of the dividend yield into 25 equally spaced values in the interval $\mathrm{E}[d_t - p_t] \pm 3\ \mathrm{Std}[d_t - p_t]$ and solve the dynamic problem recursively backwards on this grid. For each future date and every point in the discretized state space, we evaluate the conditional expectation in equation (6) with 10,000 one-step ahead simulations of the stock return and dividend yield from the VAR, and we interpolate the value function whenever the future realization of the dividend yield does not coincide with one of the grid points. The table reports for both solution methods the optimal allocation to stocks for the case of $\gamma = 10$ when the dividend yield is equal to one SD below, at, or one SD above the historical average at horizons ranging from two quarters to five years (20 quarters).

The table summarizes that the two numerical solution methods yield very similar portfolio weights, with most of the differences between the stock allocations being less than one percent in magnitude.[20] The associated

Table 3
Simulation method versus discretized state space approach

	Simulation			Discretized state space		
$T-t$	−1 SD	Mean	+1 SD	−1 SD	Mean	+1 SD
2	0.1733	0.3806	0.5764	0.1608	0.3447	0.5294
4	0.1969	0.4272	0.6365	0.1765	0.3767	0.5781
8	0.2454	0.5062	0.7284	0.2121	0.4436	0.6727
20	0.4034	0.7176	0.9157	0.3374	0.6349	0.8933

This table compares the portfolio policy of a constant relative risk aversion investor with relative risk aversion $\gamma = 10$ obtained with the simulation method using 10,000 simulated sample paths and a fourth-order expansion of the value function to the solution obtained by discretizing the state space and evaluating the conditional expectations at each point in the state space with 10,000 simulations. The portfolios are rebalanced quarterly, and the investment horizon $T - t$ ranges from two quarters to five years. The risk-free rate is constant at six percent per year, and the excess stock returns and log dividend yield evolve as a first-order VAR with parameters given in equation (28). For both solution methods, the table summarizes the optimal fraction of wealth invested in stocks when the dividend yield being equal to one SD below, at, one SD above the historical average.

[20] The portfolio weights obtained with our simulation method are always somewhat higher than those from the discretized state space approach. It is unclear which method is biased.

differences in expected utility are negligible (and are therefore not reported in the table). An important practical difference between the two methods, which is not apparent in the table, is that our simulation method is much faster (by a factor of almost 10) than the discretized state space approach.

The results in tables 2 and 3 and in Figure 1 are obtained with a *linear* specification of the conditional expectations of $A_{t+1}, B_{t+1}, C_{t+1}(x_t)$, and $D_{t+1}(x_t)$ in equation (11). Although linear expectations do not correspond to a linear value function, since we project the derivatives of the value function on the basis function, it is important to understand how sensitive the solution is to this linearity assumption. In Table 4, we compare the portfolio policies obtained with linear basis functions (which we copy for reference from Table 2) to policies with quadratic basis functions. At short horizons of up to two years, the stock holdings and the annualized in- and out-of-sample certainty equivalent returns on wealth are virtually identical across the two sets of basis functions. At longer horizons of five to ten years, there are small differences in the portfolio weights (as much as one percent). However, the expected utility from implementing the portfolio policy actually deteriorates slightly when we add quadratic regressors. This is probably because the regressions with quadratic basis functions are more sensitive to outliers than with linear basis functions.[21] Both the differences in the computed portfolio weights and the certainty equivalent of wealth are of the order of magnitude of the standard errors on these quantities computed from the Monte Carlo experiment. This suggests that the differences are insignificant so that there is no apparent advantage to increasing the order of the basis functions.

Finally, to assess the magnitude of simulation error in our solution method, we investigate further the role of the number of simulated paths used in the algorithm. Table 5 summarizes standard errors of the portfolio weights and certainty equivalent return on wealth, as we vary the number of simulated paths from 1000 to 10,000. The standard errors are computed from 1000 replications of the algorithm. The results confirm the intuition that the standard errors of both the portfolio weights and the certainty equivalent return on wealth decline with the number of simulations. As few as 10,000 simulation paths are sufficient to reduce the standard errors to negligible levels. For instance, the maximum standard error on the certainty equivalent return on wealth is only 10 basis points for a five-year horizon.

3.3 Investing in stocks while learning about the return model

Having established the properties of our method for a relatively simple portfolio problem, we turn to a more challenging and economically more interesting application. We consider the problem of an investor who takes

[21] Further experimentation reveals that, consistent with this explanation, the expected utility increases when we truncate outliers and deteriorates further when we add cubic terms without truncation.

Table 4
Linear versus quadratic polynomial expectations

	−1 SD		Mean		+1 SD	
$T-t$	Linear	Quadratic	Linear	Quadratic	Linear	Quadratic
A. Portfolio weights						
2	0.1746 (0.0041)	0.1746 (0.0041)	0.3812 (0.0090)	0.3812 (0.0090)	0.5761 (0.0130)	0.5760 (0.0130)
4	0.2011 (0.0055)	0.2008 (0.0055)	0.4310 (0.0108)	0.4301 (0.0109)	0.6405 (0.0149)	0.6384 (0.0150)
8	0.2549 (0.0075)	0.2523 (0.0073)	0.5161 (0.0133)	0.5090 (0.0131)	0.7385 (0.0167)	0.7262 (0.0166)
20	0.4128 (0.0114)	0.3911 (0.0105)	0.7213 (0.0152)	0.6811 (0.0157)	0.9136 (0.0156)	0.8731 (0.0169)
40	0.5862 (0.0117)	0.5498 (0.0129)	0.8544 (0.0131)	0.8140 (0.0159)	0.9692 (0.0139)	0.9521 (0.0160)
B. Certainty equivalent return on wealth						
2	6.47% (0.02%)	6.47% (0.02%)	7.48% (0.03%)	7.48% (0.04%)	9.25% (0.07%)	9.25% (0.07%)
4	6.55% (0.02%)	6.55% (0.02%)	7.59% (0.04%)	7.59% (0.04%)	9.39% (0.07%)	9.39% (0.07%)
8	6.69% (0.03%)	6.69% (0.03%)	7.79% (0.05%)	7.79% (0.05%)	9.66% (0.09%)	9.64% (0.08%)
20	7.12% (0.04%)	7.10% (0.03%)	8.36% (0.08%)	8.30% (0.06%)	10.25% (0.10%)	10.11% (0.08%)
40	7.61% (0.04%)	7.56% (0.04%)	8.68% (0.05%)	8.63% (0.06%)	10.01% (0.05%)	10.09% (0.08%)

This table compares the dynamic portfolio policy of a constant relative risk aversion investor with relative risk aversion $\gamma = 10$ obtained with regressions on linear versus quadratic polynomials of the dividend yield. The portfolios are rebalanced quarterly, and the investment horizon $T - t$ ranges from two quarters to 10 years. The risk-free rate is constant at six percent per year, and the excess stock returns and log dividend yield evolve as a first-order VAR with parameters given in equation (28). The first number in each block is the average across 1000 Monte Carlo replications. The second number in parenthesis is the SD of across all the Monte Carlo replications. Panel A is for the portfolio weight in the stock index. Panel B is for the certainty equivalent return on wealth (annualized) from following the portfolio policy in-sample. The results are based on 10,000 simulated sample paths, a fourth-order expansion of the value function.

Table 5
Number of simulation paths

	−1 SD			Mean			+1 SD		
T–t	1000	5000	10,000	1000	5000	10,000	1000	5000	10,000
A. Portfolio weights									
2	0.1753 (0.0132)	0.1748 (0.0059)	0.1746 (0.0041)	0.3816 (0.0277)	0.3816 (0.0128)	0.3812 (0.0090)	0.5767 (0.0395)	0.5767 (0.0185)	0.5761 (0.0130)
4	0.2021 (0.0181)	0.2015 (0.0079)	0.2011 (0.0055)	0.4303 (0.0351)	0.4312 (0.0156)	0.4310 (0.0108)	0.6389 (0.0474)	0.6406 (0.0213)	0.6405 (0.0149)
8	0.2563 (0.0238)	0.2552 (0.0102)	0.2549 (0.0075)	0.5163 (0.0418)	0.5166 (0.0180)	0.5161 (0.0133)	0.7376 (0.0520)	0.7392 (0.0225)	0.7385 (0.0167)
20	0.4111 (0.0314)	0.4121 (0.0156)	0.4128 (0.0114)	0.7168 (0.0440)	0.7204 (0.0208)	0.7213 (0.0152)	0.9079 (0.0452)	0.9128 (0.0215)	0.9136 (0.0156)
40	0.5804 (0.0353)	0.5851 (0.0162)	0.5862 (0.0117)	0.8502 (0.0412)	0.8541 (0.0180)	0.8544 (0.0131)	0.9570 (0.0361)	0.9671 (0.0184)	0.9692 (0.0139)
B. Certainty equivalent return on wealth									
2	6.49% (0.07%)	6.47% (0.03%)	6.47% (0.02%)	7.51% (0.12%)	7.49% (0.05%)	7.48% (0.03%)	9.28% (0.21%)	9.25% (0.10%)	9.25% (0.07%)
4	6.57% (0.08%)	6.55% (0.03%)	6.55% (0.02%)	7.62% (0.14%)	7.59% (0.06%)	7.59% (0.04%)	9.43% (0.24%)	9.40% (0.11%)	9.39% (0.07%)
8	6.73% (0.09%)	6.70% (0.04%)	6.69% (0.03%)	7.85% (0.16%)	7.80% (0.07%)	7.79% (0.05%)	9.74% (0.27%)	9.67% (0.13%)	9.66% (0.09%)
20	7.19% (0.10%)	7.13% (0.05%)	7.12% (0.04%)	8.46% (0.16%)	8.38% (0.10%)	8.36% (0.08%)	10.34% (0.22%)	10.26% (0.13%)	10.25% (0.10%)
40	7.66% (0.07%)	7.62% (0.05%)	7.61% (0.04%)	8.74% (0.09%)	8.69% (0.06%)	8.68% (0.05%)	10.04% (0.11%)	10.01% (0.06%)	10.01% (0.05%)

This table summarizes the Monte Carlo experiment of the number of simulation paths in the regression used to approximate conditional expectations in the dynamic portfolio policy of a constant relative risk aversion investor with relative risk aversion $\gamma = 10$. The portfolios are rebalanced quarterly, and the investment horizon $T - t$ ranges from two quarters to 10 years. The risk-free rate is constant at six percent per year and the excess stock returns and log dividend yield evolve as a first-order VAR with parameters given in equation (28). The first number in each block is the average across 1000 Monte Carlo replications. The second number in parenthesis is the standard deviation of across all the Monte Carlo replications. The results are based fourth-order expansion of the value function and 1000 Monte Carlo replications.

into account the predictability of returns but is uncertain about the parameters of the return model given by equation (28). The Bayesian portfolio choice literature, pioneered by Zellner and Chetty (1965), Bawa and Klein (1976), and Brown (1978), argues that in the presence of parameter uncertainty, the unknown objective return distribution in the expected utility maximization (1) should be replaced with the investor's subjective posterior return distribution reflecting the information contained in the historical data and the investor's prior beliefs about the parameters. In the case of the restricted VAR in equation (28), this idea has been implemented in a single-period portfolio problem by Kandel and Stambaugh (1996) and partially implemented in a multiperiod problem by Barberis (2000). Both studies demonstrate that parameter uncertainty is an important dimension of risk which can substantially lower the investor's optimal allocation to stocks.

Fully incorporating parameter uncertainty in a multiperiod portfolio problem is more complicated than in an otherwise identical single-period problem. The reason is that the multiperiod solution should take into account the fact that the posterior distribution changes each period as the investor incorporates the information contained in the new data realizations into the posterior beliefs. This learning process increases dramatically the dimensionality of the state space (from one to 11 state variables in the restricted VAR example shown below) and has not been studied before for computational reasons. There are a few exceptions in the literature which deal with learning about individual elements of the parameter vector, including Brennan (1998), an example in Barberis (2000), and Xia (2001). However, due to the limited scope of learning considered, these studies are much more informative about the overall importance of learning in the multiperiod portfolio choice than about the sign and magnitude of the effects in a more realistic setting in which the investor is uncertain about all the parameters of the model. More specifically, there is no study of the effect of *simultaneously* learning about the unconditional risk premium, predictability, mean of the predictor, and second moments of the predictor and return. The proliferation of state variables and the documented economic importance of learning in the mutiperiod portfolio choice make the problem of learning about the entire return-generating process an ideal application of our numerical solution method.

To formalize the setup of this application, we express the restricted VAR in equation (28) generically as

$$Y_{t+1} = BX_t + \epsilon_{t+1} \text{ with } \epsilon_{t+1} \sim N[0,\Sigma] \tag{29}$$

where $Y_{t+1} \equiv [r^e_{t+1}, d_{t+1} - p_{t+1}]'$, $X_t \equiv [1, d_t - p_t]'$, and $\epsilon_{t+1} = [\epsilon_{1,t+1}, \epsilon_{2,t+1}]$. Assuming the investor has standard non-informative prior beliefs about the parameters $[B, \Sigma]$ given by

$$p(B,\Sigma) \propto |\Sigma|^{3/2} \tag{30}$$

the posterior beliefs about the parameters, denoted $p(B, \Sigma|Y, X)$, are characterized by

$$p(\Sigma^{-1}|\mathrm{Y},\mathrm{X}) \sim \mathrm{Wishart}(T - 3, S^{-1})$$

$$p(\mathrm{vec}(B)|\Sigma, Y, X) \sim \mathrm{N}[\mathrm{vec}(\hat{B}), \Sigma \otimes (X'X)^{-1}] \tag{31}$$

with $S \equiv (Y - X\hat{B})'(Y - X\hat{B})$ and $\hat{B} = (X'X)^{-1}(X'Y)$ [Zellner (1971)]. Finally, we obtain the investor's posterior distribution of the data next period by integrating the conditional (on the dividend yield and unknown parameters) distribution implied by the model (29) against the posterior distribution of the parameters in equation .

If we ignore learning, as in Barberis (2000), the Bayesian multiperiod portfolio choice problem is essentially the same as when the investor knows the return-generating process. There is one state variable, the current dividend yield, which fully characterizes the conditional distribution of returns. The only difference is that each period the conditional expectation is evaluated with respect to the investor's posterior distribution of returns (conditional on $d_t - p_t$), which turns out to be a student-t distribution with a variance somewhat larger than the sample variance of returns to reflect the parameter uncertainty. With learning, however, the conditional posterior distribution of returns changes from one period to the next, as new information is incorporated into the investor's beliefs each period. The posterior distribution therefore becomes part of the state space describing the conditional distribution of returns. Specifically, the posterior distributions in equation (31) are fully characterized by the matrices $X'X$, $X'Y$, and S. Sufficient state variables for the investor's multiperiod problem are therefore the unique elements of these matrices along with the current dividend yield, for a total of 11 state variables. This proliferation of state variables explains why fully incorporating parameter uncertainty with learning into the multiperiod portfolio choice has until now been deemed computationally infeasible.

Solving this problem with our simulation-based method is as straightforward as in the previous two applications. Along each simulated path, we update the posterior distributions of the parameters given the previously simulated return and dividend yield and then simulate the next return and dividend yield from the updated posterior predictive distributions. The way the posterior distributions evolve along each path emulates the investor's learning about the return-generating process. We then solve the dynamic optimization problem backward as usual, using as state variables in the across-path regressions the realizations of the dividend yield and the realizations of the unique elements of $X'X$, $X'Y$, and S in each future period.

To get a better sense of the role of learning in the dynamic portfolio choice, we compare three different portfolio problems corresponding to three different subjective data generating processes. The results labeled 'C' are for an investor who takes the parameters as constant at the point estimates given in equation (28), as in the previous application. 'N' denotes the case in which the investor takes into account parameter uncertainty but does not learn from new data realizations. This is the case examined by Barberis (2000). Finally, the results labeled 'L' are for an investor who learns about the return-generating process.

With learning, the conditional moments of the stock return and dividend yield vary along each simulated sample path. The way the first moments change is highly correlated with the new data realizations (discussed further below) but does not exhibit any systematic pattern across horizons. The second moments, in contrast, are highly horizon dependent. Figure 2 illustrates this horizon pattern by plotting the conditional

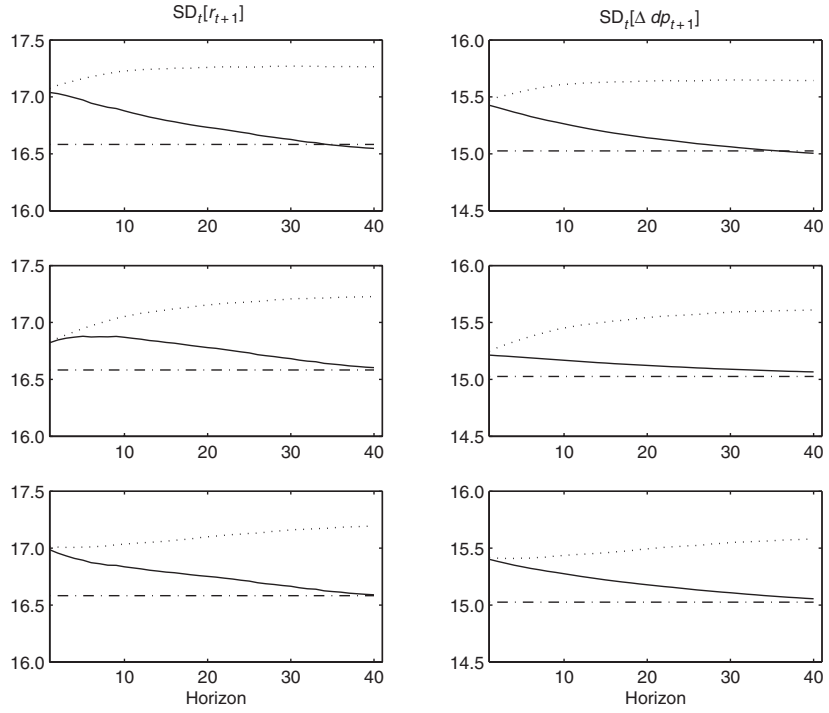

Figure 2
Conditional standard deviation of the return and dividend yield
This figure shows the average conditional annualized standard deviation in percent of the return and dividend yield across the simulated sample paths for three assumptions about parameter uncertainty and learning. The dashed, dotted, and solid lines correspond to constant parameters ('C'), parameter uncertainty without learning ('N'), and parameter uncertainty with learning ('L'), respectively. The three rows are for moments conditional on the dividend yield being equal to one SD below, at, one SD above the historical average, respectively.

second moments of the excess stock return and changes in the dividend yield as functions of the horizon for the three assumptions about parameter uncertainty and learning. More specifically, for each horizon we plot the average conditional SDs (annualized) across the 10,000 simulated sample path. The dashed, dotted, and solid lines correspond to the cases 'C', 'N', and 'L', respectively.

Parameter uncertainty with and without learning ('N' and 'L') causes the conditional second moments to be initially higher than in the constant parameter case ('C'). At the one-quarter horizon, for example, parameter uncertainty adds about 0.5 percent to the annualized volatility of stock returns. As Barberis (2000) explains, without learning the role of parameter uncertainty increases with the horizon because along each sample path the randomness in the parameters is perfectly positively correlated. With learning, in contrast, the role of parameter uncertainty diminishes gradually as the investor learns about the return-generating process. In particular, after 10 years of learning, the effect of parameter uncertainty has vanished, and so the posterior volatility of stock returns (dividend yield) is equal to the volatility of stock returns (dividend yield) under the assumption of constant parameters.

Table 6 examines the dynamic portfolio choice for an investor with $\gamma = 10$. For each assumption about parameter uncertainty and learning, the table summarizes the fraction of wealth allocated to stocks in Panel A. The results are reported for three different starting values of dividend

Table 6
Dynamic policies with learning

	−1 SD			Mean			+1 SD		
$T-t$	C	N	L	C	N	L	C	N	L
A. Portfolio weights									
2	0.1421	0.1324	0.1214	0.3138	0.2978	0.2876	0.4873	0.4516	0.4241
4	0.1612	0.1470	0.1162	0.3493	0.3216	0.2905	0.5353	0.4781	0.4023
8	0.2026	0.1744	0.1121	0.4212	0.3655	0.2918	0.6224	0.5311	0.3775
20	0.3397	0.2380	0.0860	0.6057s	0.4532	0.2454	0.7804	0.6143	0.3127
40	0.5419	0.3288	0.0518	0.7635	0.5450	0.2030	0.8323	0.6925	0.2696
B. Certainty equivalent return on wealth									
2	6.45%	6.46%	6.49%	7.36%	7.37%	7.37%	8.59%	8.67%	8.69%
4	6.47%	6.47%	6.54%	7.29%	7.32%	7.35%	8.02%	8.23%	8.44%
8	6.46%	6.49%	6.61%	6.87%	7.03%	7.26%	6.08%	6.94%	8.11%
20	4.93%	6.05%	6.12%	2.74%	4.37%	6.82%	0.68%	2.23%	7.48%
40	2.13%	2.56%	5.35%	0.99%	1.28%	6.77%	0.15%	0.31%	7.13%

This table summarizes summary statistics for the dynamic conditional policy under three different assumptions about parameter uncertainty and learning. 'C' corresponds to constant parameters, 'N' is for parameter uncertainty without learning, and 'L' corresponds to parameter uncertainty with learning. The portfolios are rebalanced quarterly and the investment horizon $T-t$ ranges from two quarters to five years. The risk-free rate is constant at six percent per year. The investor has constant relative risk aversion preferences with relative risk aversion $\gamma = 10$. Panel A gives the portfolio weights in the stock index. Panel B gives the annualized certainty equivalent return on wealth. Certainty equivalents for 'C' and 'N' policies are evaluated under the 'L' assumption. The results are based on 10,000 simulated sample paths and a fourth-order expansion of the value function.

yield: equal to, one SD below, and one SD above the historical average. The remaining 10 state variables are fixed at values at the end of the sample period in 1995. Panel B summarizes the annualized certainty equivalent return on wealth from following that policy. We evaluate all three portfolio policies under the assumption of learning (i.e., using the simulations for the 'L' case).[22] Therefore, we address the question of how much better off is an investor who ends up learning about the return-generating process by explicitly incorporating parameter uncertainty and learning in the dynamic portfolio optimization. This is the relevant question, because it is always optimal for the investor to learn and committing not to do so is time inconsistent.

Relative to the constant parameter case in the previous application, the portfolio policies which incorporate parameter uncertainty invest substantially less in stocks. For instance, at the 10-year horizon, the 'N' policy invests almost 20 percent less in stocks than the 'C' policy. As Barberis (2000) explains, this result is attributable to the increased posterior variance of returns due to parameter uncertainty. Incorporating the effect of learning further reduces the allocation to stocks. For a 10-year horizon with the initial dividend yield at its historical average, the allocation to stocks with learning is only 20.3 percent, almost 35 percent lower than the allocation in the case of parameter uncertainty without learning and almost one-fourth of the allocation in the constant parameter case.

Table 6 also summarizes that the magnitude of the difference in the stock allocation increases both with the horizon and with the current value of the dividend yield. For instance, the difference between the 'C' and 'L' policies is only six percent at the one-year horizon but a substantial 35 percent at the five-year horizon. As we discuss below, this pattern in the results is mostly due to hedging demands induced by learning.

The benefits of incorporating learning are apparent from the certainty equivalent returns on wealth in Panel B. The greater stock holdings in the constant parameter and parameter uncertainty without learning cases lead to substantial welfare losses relative to the learning case. Furthermore, the expected utility gains from investing optimally versus ignoring parameter uncertainty or ignoring learning are of the same order of magnitude. For a 10-year horizon, the certainty equivalent return on wealth of the 'L' policy is 5.8 percent per year higher than for the 'C' policy and 5.5 percent higher than for the 'N' policy. Over the 10-year period, these gains translate into a more than 70 percent difference in the

[22] We compute the certainty equivalents for the 'C' and 'N' policies by first linearly interpolating the portfolio policy around the dividend yield in own set of simulations. Using cubic splines has no material impact on our results. This portfolio policy "function" is then evaluated at the dividend yield in the 'L' case and then used for generating wealth for the returns under the 'L' scenario.

certainty equivalent of wealth. The certainty equivalent gains increase both with the horizon and with the current value of the dividend yield, which is consistent with the finding above that the differences between the optimal portfolio weights increase in the same dimensions.

Optimal portfolio choice in the case of parameter uncertainty with learning is likely to be different from the case of parameter uncertainty without learning. The revisions in the beliefs about the return-generating process, which determines the future investment opportunities, are strongly correlated with returns, inducing additional hedging demands. Table 7 quantifies these effects. Panel A summarizes the differences between the dynamic and myopic allocations. Panel B summarizes the associated differences in the certainty equivalent return on wealth, evaluated again under the assumption of learning.

Consistent with the prior literature, the hedging demands increase with the horizon for both the 'C' and 'N' policies. This pattern in the optimal allocations is due to the strong negative correlation between the returns and innovations in the dividend yield. As demonstrated originally by Barberis (2000), parameter uncertainty lowers the hedging demand. The reason is that parameter uncertainty lowers the investor's overall exposure to stocks and therefore limits the need to hedge against changes in the conditional return distribution.

Table 7
Hedging demands with learning

	−1 SD			Mean			+1 SD		
$T-t$	C	N	L	C	N	L	C	N	L
A. Hedging demand									
2	0.0105	0.0063	−0.0048	0.0205	0.0129	0.0026	0.0296	0.0184	−0.0091
4	0.0295	0.0209	−0.0099	0.0559	0.0367	0.0055	0.0775	0.0449	−0.0309
8	0.0710	0.0483	−0.0140	0.1279	0.0806	0.0069	0.1646	0.0979	−0.0557
20	0.2080	0.1119	−0.0401	0.3123	0.1682	−0.0395	0.3226	0.1811	−0.1205
40	0.4102	0.2027	−0.0743	0.4701	0.2601	−0.0820	0.3746	0.2593	−0.1635
B. Gain in certainty equivalent return on wealth									
2	−0.00%	−0.00%	−0.00%	−0.01%	−0.00%	−0.00%	−0.03%	−0.01%	0.00%
4	−0.01%	−0.01%	0.00%	−0.03%	−0.01%	0.00%	−0.13%	−0.04%	0.03%
8	−0.04%	−0.02%	0.01%	−0.21%	−0.09%	0.01%	−0.68%	−0.26%	0.05%
20	−1.39%	−0.37%	0.03%	−2.28%	−1.28%	0.30%	−1.97%	−1.36%	0.14%
40	−2.80%	−3.00%	0.86%	−2.00%	−2.28%	1.34%	−1.27%	−1.71%	0.31%

This table summarizes summary statistics for the hedging demands under three different assumptions about parameter uncertainty and learning. 'C' corresponds to constant parameters, 'N' is for parameter uncertainty without learning, and 'L' corresponds to parameter uncertainty with learning. The portfolios are rebalanced quarterly, and the investment horizon $T-t$ ranges from two quarters to five years. The risk-free rate is constant at six percent per year. The investor has constant relative risk aversion preferences with relative risk aversion $\gamma = 10$. Panel A gives the hedging demand in the stock index. Panel B gives the annualized gain in certainty equivalent return on wealth from following the dynamic policy over a myopic policy. Certainty equivalent for 'C' and 'N' policies are evaluated under the 'L' assumption. The results are based on 10,000 simulated sample paths and a fourth-order expansion of the value function.

In the case of learning ('L'), the gain in the certainty equivalent return on wealth from investing optimally as opposed to myopically is about one percent for a 10-year horizon as summarized in Panel B. Ironically, for an investor who ignores learning ('C' or 'N'), investing in a dynamically "optimal" way as opposed to myopically actually results in a certainty equivalent *loss* in utility. The magnitude of this loss increases with the horizon and is of the order of two to three percent for a 10-year horizon. The reason for this utility loss is clear. The dynamic policy without learning allocates more wealth to stocks, especially at long horizons, than the corresponding myopic policy. The myopic policy without learning, in turn, is much more similar to the optimal policy with learning. Therefore, the myopic policy without learning dominates its dynamic counterpart. This result again illustrates the importance of incorporating learning into the dynamic portfolio choice.

The most striking result in Table 7 is that, in the case of parameter uncertainty with learning, the hedging demands are *negative* and even further decreasing with the horizon. This pattern in the optimal allocation is attributable to learning about model parameters. We can break up the set of model parameters into three sets: one related to the intercept and slope of the return predictability equation, the second related to the mean of the dividend yield, and the third related to second moments.

Consider first the impact of learning about the slope and the intercept of the predictability equation. The effects of learning about these two parameters have previously been studied in isolation by Brennan (1998) and Xia (2001). Brennan shows that a higher than expected return realization leads to upward assessment of unconditional mean return, which constitutes a subjective improvement in future investment opportunities. Conversely, a low return has two negative effects on the investor, a decrease in wealth and a perceived deterioration of future investment opportunities. Since stocks are therefore a poor hedge against changes in investment opportunities, the investor chooses to allocate less wealth to stocks than in the myopic case or in the case without learning. Xia, in contrast to Brennan, assumes that the unconditional mean return (intercept) is known but there is uncertainty about predictability (slope). She finds that the hedging demand due to learning depends on where the current dividend yield is relative to its long-term mean, $(d_t - p_t) - \left(\overline{d} - \overline{p}\right)$. She shows that when $(d_t - p_t) > \left(\overline{d} - \overline{p}\right)$, a higher than expected return realization leads to an upward assessment of the slope, leading to an upward assessment of the conditional mean, and thereby a negative hedging demand. On the other hand, there is a positive hedging demand due to learning about predictability when $(d_t - p_t) < \left(\overline{d} - \overline{p}\right)$. The asymmetry in learning about the slope can be understood by plotting a regression line through a scatter plot. When a point is added above the previously fitted line, the slope changes, but the change depends on whether the x-coordinate is higher or lower than

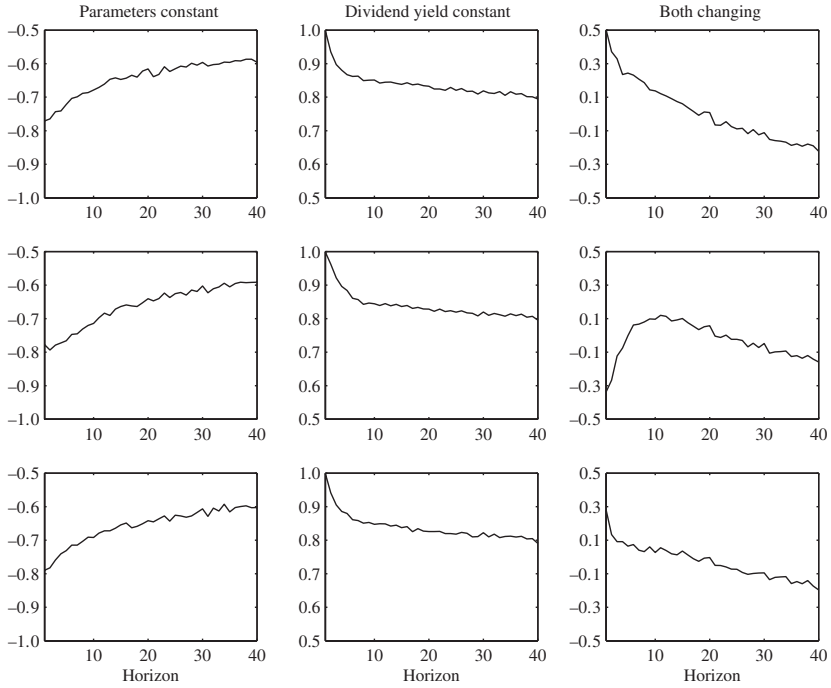

Figure 3
Correlation of conditional mean return with realized return
This figure shows the average correlation across simulated sample paths of three aspects of the conditional mean return with the realized return across simulated sample paths. The columns correspond to the correlations of realized returns from date t to $t+1$ with Parameters Constant: $[B_{11,t}+B_{12,t}(d_{t+1}-p_{t+1})]-[B_{11,t}+B_{12,t}(d_t-p_t)]$. Dividend Yield Constant: $[B_{11,t+1}+B_{12,t+1}(d_t-p_t)]-[B_{11,t}+B_{12,t}(d_t-p_t)]$. Both Changing: $[B_{11,t+1}+B_{12,t+1}(d_{t+1}-p_{t+1})]-[B_{11,t}+B_{12,t}(d_t-p_t)]$. The three rows are for correlations conditional on the dividend yield being equal to one SD below, at, one SD above the historical average, respectively.

the mean, the "pivotal" point of the regression.[23] As we discuss below, our results are consistent with both the Brennan and the Xia effects.

Figure 3 illustrates the relative importance of the positive hedging demand due to the negative correlation of changes in the dividend yield with returns and the negative hedging demand due to the positive correlation of changes in the model parameters with returns. The first column isolates the dividend yield effect. It shows the average correlation at different horizons across the simulated sample paths between changes in the conditional mean return and the realized return, holding constant the parameters of the model between dates t and t+1 but allowing the dividend yield to change. Formally, we compute

$$\mathrm{Corr}\left\{r^e_{t+1}, \left[\hat{B}_{11,t}+\hat{B}_{12,t}(d_{t+1}-p_{t+1})\right]-\left[\hat{B}_{11,t}+\hat{B}_{12,t}(d_t-p_t)\right]\right\}, \quad (32)$$

[23] See also Chapter 5 of Campbell and Viceira (2002) for a detailed discussion of these learning effects.

where $\hat{B}_{ij,t}$ denotes the $\{i, j\}$ element of the posterior mean $\hat{B}$ at date t of the VAR coefficient matrix B. Since this expression reduces to the correlation between excess returns and changes in the dividend yield, it is negative and large in magnitude, consistent with the negative posterior correlation between the VAR residuals in equation (28).[24] This negative correlation implies that, holding the parameters of the return model constant, stocks are a good hedge against changes in the conditional mean return, inducing the usual positive hedging demand.

However, the conditional mean return also changes with the investor's posterior beliefs about the VAR coefficients. The second column of Figure 3 isolates this learning effect. It shows the same correlation between the conditional mean return and realized return, except now holding constant the dividend yield and allowing the model parameters to change:

$$\mathrm{Corr}\left\{r^{e}_{t+1}, \left[\hat{B}_{11,t+1} + \hat{B}_{12,t+1}(d_t - p_t)\right] - \left[\hat{B}_{11,t} + \hat{B}_{12,t}(d_t - p_t)\right]\right\}. \quad (33)$$

Consistent with the intuition that a positive return innovation leads to an upward revision of the unconditional/conditional mean return, this correlation is large and positive (almost perfect at short horizons). Holding constant the dividend yield, returns are then actually a poor hedge against changes in the (subjective) conditional mean return, which means that the learning effect leads to negative hedging demands.

The third column of the figure combines the two effects by showing the average correlation between the conditional mean return and return realization with both the parameters and dividend yield changing. For our sample period, the aggregate correlation is positive but relatively small at short horizons (except when the dividend yield is at its historical average), and it turns negative at horizons of more than five years. This column shows the net impact of the Brennan and Xia effects. However, it is important to note that since Xia's results are state dependent and Brennan's result are not, it is not entirely obvious *a priori* which effect will dominate when the effects are of opposite sign. Furthermore, one learns much slower in discrete time and can hedge much less optimally than in continuous time. Therefore, one cannot directly take Brennan and Xia's continuous time portfolio weights and welfare gains and assume similar results are obtained in discrete time. Our article is the first to quantify the effect first illustrated by Brennan and Xia in discrete time.

[24] The correlation between returns and changes in the dividend yield is less negative than the correlation between the return and dividend yield innovations due to the mean reversion in the dividend yield.

Now consider the impact of learning about the mean of dividend yield, which has not been studied so far.[25] Learning about the mean of the dividend yield induces a positive hedging demand. The intuition is as follows. A positive return innovation is associated with a negative innovation to the dividend yield because of the negative contemporaneous correlation between the two. This leads to a downward assessment of the mean of the dividend yield which in turn reduces the unconditional mean of returns. This worsening of the investment opportunity set when returns are high (and vice versa) creates a positive hedging demand for stocks. Thus, the negative effect of learning studied by Brennan (1998) and Xia (2001) is partially offset when one also takes into account learning about the mean of the predictor variable.

Finally, the effects of learning about second moments have also not been studied in the extant literature.[26] Consider first the impact of learning about the correlation between returns and dividend yield. A positive return innovation is contemporaneously related to a negative dividend yield innovation. However, a high positive return innovation has the effect of reducing this correlation, making it less negative. A less negative correlation implies that the investor is less able to hedge changes in the investment opportunity set, making the investor worse off. This gives rise to positive hedging demand, which again partially offsets the usual negative hedging demands from learning. Consider now the impact of learning about the variance of returns. A low absolute return reduces the assessment of variance, while a high absolute return raises the assessment of variance. Thus, both high positive and high negative returns are associated with bad news on future investment opportunity set, while both low positive and low negative returns are associated with good news on future investment opportunity set. The net effect on hedging demand is unclear. Similarly, learning about the variance of innovations to the dividend yield has an ambiguous impact on portfolio choice.

We analyze the impact of learning about all the different parameters in Table 8. This table compares the portfolio policy with learning from the previous tables, labeled 'L3' here, to two alternative policies for which the investor only conditions on subsets of the 11 state variables. In particular, for policy 'L1' the investor conditions on the coefficients in the first row of $\hat{B} = (X'X)^{-1}(X'Y)$ which determine the conditional mean of the stock return (three state variables including the dividend yield). For policy 'L2', the inves-

[25] Xia (2001) considers the case when both the unconditional mean of stock return and the dividend yield are known. Her predictive regressions are, therefore, expressed in demeaned form, and the question of learning about the mean of dividend yield is moot. Demeaning regressions in our setup is not enough because the "mean" of dividend yield changes over time with learning.

[26] Both Brennan (1998) and Xia (2001) treat the elements of covariance matrix as known constants. Chacko and Viceira (2003) study the optimal portfolio choice of an investor faced with time-varying volatility but there is no learning involved. In the continuous-time framework used in these papers, volatility can be exactly measured due to the continuous-record asymptotic argument of Merton (1980).

Table 8
Dynamic policies with partial conditioning

	−1 SD			Mean			+1 SD		
$T-t$	L1	L2	L3	L1	L2	L3	L1	L2	L3
A. Portfolio weights									
2	0.1216	0.1217	0.1214	0.2882	0.2887	0.2876	0.4272	0.4270	0.4241
4	0.1166	0.1166	0.1162	0.2911	0.2917	0.2905	0.4046	0.4047	0.4023
8	0.1099	0.1108	0.1121	0.2927	0.2941	0.2918	0.3745	0.3776	0.3775
20	0.0808	0.0854	0.0860	0.2462	0.2538	0.2454	0.2995	0.3111	0.3127
40	0.0416	0.0529	0.0518	0.1699	0.1958	0.2030	0.2339	0.2599	0.2696
B. Certainty equivalent return on wealth									
2	6.49%	6.49%	6.49%	7.38%	7.39%	7.37%	8.73%	8.73%	8.69%
4	6.54%	6.55%	6.54%	7.36%	7.37%	7.35%	8.47%	8.47%	8.44%
8	6.62%	6.62%	6.61%	7.24%	7.25%	7.26%	8.05%	8.09%	8.11%
20	6.19%	6.26%	6.12%	6.48%	6.65%	6.82%	7.25%	7.40%	7.48%
40	4.59%	4.79%	5.35%	6.33%	6.52%	6.77%	6.85%	7.01%	7.13%

This table summarizes summary statistics for the dynamic conditional portfolio policies for three different conditioning information sets. Policy 'L1' is conditional on the regression coefficients of the return predictability equation only, policy 'L2' is conditional on the regression coefficients of both the return and dividend yield predictability equations, and policy 'L3' is conditional on all of the state variables (policy 'L3' is the same as policy 'L' in Tables 6 and 7). The portfolios are rebalanced quarterly, and the investment horizon $T-t$ ranges from two quarters to five years. The risk-free rate is constant at six percent per year. The investor has constant relative risk aversion preferences with relative risk aversion $\gamma=10$. Panel A gives the portfolio weights in the stock index. Panel B gives the annualized certainty equivalent return on wealth. The results are based on 10,000 simulated sample paths and a fourth-order expansion of the value function.

tor conditions also on the second row of $\hat{B}$ and therefore also on the conditional means of the dividend yield (five state variables including the dividend yield). 'L3' essentially adds the conditional second moments. In all three cases, we assume the investor learns about the entire data generating process.

The impact of learning about the conditional means of the dividend yield is demonstrated through policies 'L1' and 'L2'. Consistent with our intuition about positive hedging demand due to learning about the mean of dividend yield, we find that policy 'L2' leads to a higher stock allocation than policy 'L1'. The impact of learning about the elements of the covariance matrix is shown through policies 'L2' and 'L3'. We find that policy 'L3' (that learns about second moments) sometimes allocates more and sometimes less than policy 'L2' (that does not learn about second moments) which is consistent with our discussion above.

Panel B of Table 8 summarizes that the utility gains from learning about the mean of dividend yield are of the order of 15 to 20 basis points per year. This effect is substantial especially if we take into account that this calculation understates the true benefit since the data generating process is for the full learning case. This panel also summarizes the economic benefits from learning about second moments of the order of 12 to 56 basis points a year.

These findings also illustrate the economic value of our solution method. Without it, solving a portfolio problem with such high-dimensional state space would still be deemed computationally infeasible.

4. Conclusion

We presented a simulation-based method for solving dynamic portfolio choice problems involving non-standard preferences and a large number of assets with arbitrary return distribution. The main advantage of our approach relative to other numerical solution methods is that our method can handle problems in which the conditional return distribution depends on a large number of state variables with potentially path-dependent or even non-stationary dynamics. The method can also accommodate intermediate consumption, portfolio constraints, parameter and model uncertainty, and learning. Most importantly, our method is easy to implement and works efficiently for large-scale problems.

We first apply the method to the stylized portfolio choice between the stock market and cash when stock returns are either iid or predictable by the dividend yield. We show that our approach replicates the results obtained by previous researchers using alternative numerical methods, giving us confidence in the properties of our method. We then apply our method to the portfolio choice of an investor who realizes that there is some predictability in stock returns but is uncertain about the parameters of the return-generating process. The investor learns about the parameters from the realizations of returns and dividend yield and takes into account the effect of learning from these future data in the optimal portfolio choice. We find that there is a negative hedging demand due to learning that reduces the allocation to stocks beyond the already smaller weight due to the parameter uncertainty. We document substantial welfare gains from conditioning on all the state variables required to learn about the distribution of returns.

References

Aït-Sahalia, Y., and M. W. Brandt, 2001, "Variable Selection for Portfolio Choice," *Journal of Finance*, 56, 1297–1351.

Aït-Sahalia, Y., and M. W. Brandt, 2003, "Portfolio and Consumption Choice with Option-Implied State Prices," Working paper, Princeton University.

Alexander, G. A., and A. M. Baptista, 2002, "Economic Implications of Using a Mean-VaR Model for Portfolio Selection: A Comparison with Mean-Variance Analysis," *Journal of Economic Dynamics and Control*, 26, 1159–1193.

Ang, A., G. Bekaert, and J. Liu, 2002, "Why Stocks May Disappoint," Working paper, Columbia University.

Balduzzi, P., and A. W. Lynch, 1999, "Transaction Costs and Predictability: Some Utility Cost Calculations," *Journal of Financial Economics*, 52, 47–78.

Barberis, N., 2000, "Investing for the Long Run when Returns are Predictable," *Journal of Finance*, 55, 225–264.

Basak, S., and A. Shapiro, 2001, "Value-at-Risk Based Risk Management: Optimal Policies and Asset Prices," *Review of Financial Studies*, 14, 371–405.

Bawa, V. S., and R. W. Klein, 1976, "The Effect of Estimation Risk on Optimal Portfolio Choice," *Journal of Financial Economics*, 3, 215–231.

Bellman, R., R. Kalaba, and B. Kotkin, 1963, "Polynomial Approximation – A New Computational Technique in Dynamic Programming: Allocation Processes," *Mathematics of Computation*, 17, 155–161.

Benartzi, S., and R. H. Thaler, 1995, "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 110, 73–92.

Black, F., and R. B. Litterman, 1992, "Global Portfolio Optimization," *Financial Analyst Journal*, 48, 24–43.

Brandt, M. W., 1999, "Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach," *Journal of Finance*, 54, 1609–1645.

Brandt, M. W., 2003, "The Econometrics of Portfolio Choice Problems," in Y. Aït-Sahalia and L.P. Hansen, (eds.), *Handbook of Financial Econometrics*, Manuscript .

Brennan, M. J., forthcoming in, "The Role of Learning in Dynamic Portfolio Decisions," *European Finance Review*, 1, 295–306.

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

Brown, S. J., 1978, "The Portfolio Choice Problem: Comparison of Certainty Equivalence and Optimal Bayes Portfolios," *Communications in Statistics: Simulation and Computation*, 7, 321–334.

Browne, S., 1999, "Beating a Moving Target: Optimal Portfolio Strategies for Outperforming a Stochastic Benchmark," *Finance and Stochastics*, 3, 275–294.

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu, 1995, "A Limited Memory Algorithm for Bound Constrained Optimization," *SIAM Journal on Scientific Computing*, 16, 1190–1208.

Campbell, J. Y., and L. M. Viceira, 1999, "Consumption and Portfolio Decisions when Expected Returns are Time Varying," *Quarterly Journal of Economics*, 114, 433–495.

Campbell, J. Y., and L. M. Viceira, 2002, *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, Oxford University Press, New York.

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

Chacko, G., and L. M. Viceira, 2003, "Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets," Working paper, Harvard University.

Clément, E., D. Lamberton, and P. Potter, 2001, "An Analysis of the Longstaff-Schwartz Algorithm for American Option Pricing," Working paper, Université de Marne-la-Vallée.

Cochrane, J. H., 1989, "The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives," *American Economic Review*, 79, 319–337.

Conn, A. R., N. I. M. Gould, and P. Toint, 1988, "Global Convergence for a Class of Trust-Region Algorithms for Optimization with Simple Bounds," *SIAM Journal of Numerical Analysis*, 25, 433–460.

Conn, A. R., N. I. M. Gould, and P. Toint, 1994, "Large-Scale Nonlinear Constrained Optimization: A Current Survey in Algorithms for Continuous Optimization: The State of the Art," in E. Spedicato, (ed.), *Mathematical Physical Sciences*, Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 287–332.

Connor, G., 1997, "Sensible Return Forecasting for Portfolio Management," *Financial Analyst Journal*, 53, 44–51.

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

Dammon, R. M., C. S. Spatt, and H. H. Zhang, 2001, "Optimal Consumption and Investment with Capital Gains Taxes," *Review of Financial Studies*, 14, 583–616.

Daniel, J. W., 1976, "Splines and Efficiency in Dynamic Programming," *Journal of Mathematical Analysis and Applications*, 54, 402–407.

Das, S., and R. K. Sundaram, 2000, "A Numerical Algorithm for Optimal Consumption-Investment Problems," Working paper, Harvard University.

den Haan, W. J., and A. Marcet, 1990, "Solving the Stochastic Growth Model by Parameterized Expectations," *Journal of Business and Economic Statistics*, 8, 31–34.

Detemple, J. B., 1986, "Asset Pricing in a Production Economy with Incomplete Information," *Journal of Finance*, 41, 383–391.

Detemple, J. B., R. Garcia, and M. Rindisbacher, 2003, "A Monte Carlo Method for Optimal Portfolios," *Journal of Finance*, 58, 401–446.

Dothan, M. U., and D. Feldman, 1986, "Equilibrium Interest Rates and Multiperiod Bonds in a Partially Observable Economy," *Journal of Finance*, 41, 369–382.

Dybvig, P. H., 1995, "Duesenberry's Racheting of Consumption: Optimal Dynamic Allocation and Investment given Intolerance for any Decline in Standard of Living," *Review of Economic Studies*, 62, 287–313.

Epstein, L. G., and S. E. Zin, 1989, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

Fama, E. F., 1970, "Multiperiod Consumption-Investment Decisions," *American Economic Review*, 60, 163–174.

Gennotte, G., 1986, "Optimal Portfolio Choice under Incomplete Information," *Journal of Finance*, 41, 733–746.

Gill, P. E., W. Murray, and M. A. Saunders, 2002, "SNOPT: An SQP Algorithm for Large-Scale Constrained Optimization," *SIAM Journal of Optimization*, 12, 979–1006.

Gomes, F. J., 2002, "Portfolio Choice and Trading Volume with Loss-Averse Investors," Working paper, London Business School, forthcoming in *Journal of Business*.

Grauer, R. R., 1981, "A Comparison of Growth Optimal and Mean-Variance Investment Policies," *Journal of Financial and Quantitative Analysis*, 16, 1–21.

Grauer, R. R., and N. H. Hakansson, 1993, "On the Use of Mean-Variance and Quadratic Approximations in Implementing Dynamic Investment Strategies: A Comparison of Returns and Investment Policies," Working paper, University of California, Berkeley.

Grossman, S. J., and J. Vila, 1989, "Portfolio Insurance in Complete Markets: A Note," *Journal of Business*, 62, 473–476.

Hakansson, N. H., 1971, "Capital Growth and the Mean-Variance Approach to Portfolio Selection," *Journal of Financial and Quantitative Analysis*, 6, 517–557.

Jagannathan, R., and N. R. Kocherlakota, 1996, "Why Should Older People Invest Less in Stocks than Younger People," *Federal Reserve Bank of Minneapolis Quarterly Review*, 20, 11–23.

Judd, K. J., 1998, *Numerical Methods in Economics*, MIT Press, Cambridge, MA.

Kandel, S., and R. F. Stambaugh, 1996, "On the Predictability of Stock Returns: An Asset-Allocation Perspective," *Journal of Finance*, 51, 385–424.

Kim, T. S., and E. Omberg, 1996, "Dynamic Nonmyopic Portfolio Behavior," *Review of Financial Studies*, 9, 141–161.

Kogan, L., and R. Uppal, 2000, "Risk Aversion and Optimal Portfolio Policies in Partial and General Equilibrium Economies," Working paper, University of Pennsylvania.

Kroll, Y., H. Levy, and H. M. Markowitz, 1984, "Mean-Variance versus Direct Utility Maximization," *Journal of Finance*, 39, 47–75.

Liu, J., 1999a, "Portfolio Selection in Stochastic Environments," Working paper, UCLA.

Liu, W., 1999b, "Saving, Portfolio, and Investment Decisions: Observable Implications with Knightian Uncertainty," Working paper, University of Washington.

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

Merton, R. C., 1969, "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case," *Review of Economics and Statistics*, 51, 247–257.

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

Merton, R. C., 1980, "On Estimating the Expected return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361.

Moré, J. J., and G. Toraldo, 1989, "Algorithms for Bound Constrained Quadratic Programming Problems," *Numerical Mathematics*, 55, 377–400.

Newey, W. K., and K. D. West, 1987, "A Simple, Positive Semi-Definite, Heteroscedasticity and Serial Correlation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Pulley, L. B., 1981, "A General Mean-Variance Approximation to Expected Utility for Short Holding Periods," *Journal of Financial and Quantitative Analysis*, 16, 361–373.

Roy, A. D., 1952, "Safety First and the Holding of Assets," *Econometrica*, 20, 431–439.

Samuelson, P. A., 1969, "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics*, 51, 239–246.

Samuelson, P. A., 1970, "The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances, and Moments," *Review of Economic Studies*, 37, 537–542.

Samuelson, P. A., 1989, "A Case at Last for Age-Phased Reduction in Equity," *Proceedings of the National Academy of Sciences*, 86, 9048–9051.

Schroder, M., and C. Skiadas, 1999, "Optimal Consumption and Portfolio Selection with Stochastic Differential Utility," *Journal of Economic Theory*, 89, 68–126.

Tepla, L., 2001, "Optimal Investment with Minimal Performance Constraints," *Journal of Economic Dynamics and Control*, 25, 1629–1645.

Tsitsiklis, J. N., and B. V. Roy, 2001, "Regression Methods for Pricing Complex American-Style Options," Working paper, MIT.

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.

Williams, J. C., and B. D. Wright, 1984, "The Welfare Effects of the Introduction of Storage," *Quarterly Journal of Economics*, 99, 169–92.

Xia, Y., 2001, "Learning about Predictability: The Effect of Parameter Uncertainty on Optimal Dynamic Asset Allocation," *Journal of Finance*, 56, 205–247.

Zellner, A., 1971, *An Introduction to Bayesian Inference in Econometrics*, John Wiley and Sons, New York.

Zellner, A., and V. K. Chetty, 1965, "Prediction and Decision Problems in Regression Models from the Bayesian Point of View," *Journal of the American Statistical Association*, 60, 608–616.

Zhao, Y., and W. T. Ziemba, 2000, "Mean-Variance versus Expected Utility in Dynamic Investment Analysis," Working paper, University of British Columbia.

Copyright of Review of Financial Studies is the property of Society for Financial Studies and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.