

Dynamic Portfolio Choice with Bayesian Learning*

Georgios Skoulakis[†]

University of Maryland

March 15, 2007

Abstract

This paper examines the importance of parameter uncertainty and learning in the context of dynamic portfolio choice. In a discrete time setting, we consider a Bayesian investor who faces parameter uncertainty and solves her portfolio choice problem while updating her beliefs about the parameters. For different return data generating processes, including i.i.d. returns, autoregressive returns, and exogenous predictability, we show how the investor makes dynamic portfolio choices, taking into account that she will learn from future data. We find that, in general, learning introduces negative horizon effects and that ignoring parameter uncertainty may lead to significant losses in certainty equivalent return on wealth. However, the significance of learning is reduced when the investor uses more past data in her estimation and/or when her risk aversion increases. Learning about unconditional expected returns appears to be the most important aspect of the learning process. Using the earnings-to-price ratio as a predictor and an empirical Bayes prior, we find that learning reduces, but does not necessarily eliminate, the positive hedging demands induced by predictability and correlation between the return and predictor innovations.

*I wish to thank Patricia Ledesma, Jacob Sagi, Caroline Sasseville, Arne Staal, Jeremy Staum, and Kostas Zachariadis for helpful discussions and comments and especially to Ravi Jagannathan and Costis Skiadas for their guidance and support. I am indebted to Bruce Foster, Jim Guo, Patricia Ledesma, and Sotos Traftaris for their help with computational issues. Any errors are my sole responsibility. Comments and suggestions are welcome at gskoulak@rhsmith.umd.edu.

[†]Department of Finance, Robert H. Smith School of Business, University of Maryland, College Park, MD 20742.

Dynamic Portfolio Choice with Bayesian Learning

ABSTRACT

This paper examines the importance of parameter uncertainty and learning in the context of dynamic portfolio choice. In a discrete time setting, we consider a Bayesian investor who faces parameter uncertainty and solves her portfolio choice problem while updating her beliefs about the parameters. For different return data generating processes, including i.i.d. returns, autoregressive returns, and exogenous predictability, we show how the investor makes dynamic portfolio choices, taking into account that she will learn from future data. We find that, in general, learning introduces negative horizon effects and that ignoring parameter uncertainty may lead to significant losses in certainty equivalent return on wealth. However, the significance of learning is reduced when the investor uses more past data in her estimation and/or when her risk aversion increases. Learning about unconditional expected returns appears to be the most important aspect of the learning process. Using the earnings-to-price ratio as a predictor and an empirical Bayes prior, we find that learning reduces, but does not necessarily eliminate, the positive hedging demands induced by predictability and correlation between the return and predictor innovations.

1 Introduction

Dynamic portfolio choice is among the most fundamental problems in financial economics and asset pricing in particular. Following the seminal contributions of Samuelson (1969) and Merton (1969, 1971), a vast literature has emerged addressing various aspects of the problem. The recent reviews by Campbell and Viceira (2002),¹ Brandt (2004),² and Skiadas (2005)³ provide extensive surveys of the existing literature. Important issues in the context of portfolio selection include return predictability; frictions, such as transactions costs and taxes; and background risks, such as labor income and housing, among others.

Another major aspect of the portfolio choice problem currently receiving attention is parameter and model uncertainty. Kandel and Stambaugh (1996), Brennan(1998), Barberis (2000), Xia (2001), Boudry and Gray (2003), and Brandt, Goyal, Santa-Clara, and Stroud (2005) are recent papers that address the issue of parameter uncertainty. Avramov (2002) and Tu and Zhou (2003), among others, examine the role of model uncertainty in a static portfolio choice context. In this paper, we address the issue of parameter uncertainty and learning in the context of dynamic portfolio choice.

The starting point in a standard portfolio allocation problem is to postulate a model for the data generating process describing the joint stochastic evolution of all variables of interest, including asset returns. Certain aspects of such a model, e.g. parameters, are inherently unknown, although they are indispensable ingredients of the portfolio problem solution. The majority of models used in the literature involve a parametric specification. Given a model specification, the investor has to use the available data in order to make inferences about the unknown quantities and then use this information to solve her portfolio choice problem. Such an investor might be inclined to discard a certain feature of the data generating process, such as return predictability, if there is no statistically significant evidence to support it. Another possibility is that the investor is convinced about the importance of the specific feature through a formal statistical test or a heuristic argument. Then, if the investor subscribes to the so-called plug-in approach (see Brandt (2004)), she proceeds by making the assumption that the parameters are known precisely and equal to their estimates obtained from the available data. However, point estimates are subject to estimation error which

¹Campbell and Viceira (2002) focus on the dynamic portfolio choice of a long-term investor in the context of recursive preferences of Epstein-Zin (1989).

²Brandt (2004) analyzes the econometrics of portfolio choice problems with emphasis placed on the link between the theoretical developments and the econometric treatment of data.

³Skiadas (2005) studies portfolio-consumption choice problems in a continuous-time setting with a focus on the aspect of modelling risk aversion.

is passed on to the portfolio allocation solution. Hence, unless there is no estimation error, the resulting allocation will be suboptimal. This gives rise to estimation risk. We consider both of the above approaches to be extreme. As discussed in detail by Brandt (2004), the Bayesian decision-theoretic approach provides a more attractive alternative. In contrast to the plug-in approach, the Bayesian paradigm allows the investor to treat the parameters as unknown and incorporate parameter uncertainty into her portfolio choice problem in a natural fashion. Instead of operating under the assumption of fixed parameters, the investor integrates out the unknown parameters to obtain the predictive distribution which depends only on observed data.

Financial researchers have long recognized the importance of estimation risk in the context of portfolio choice. Zellner and Chetty (1965), Klein and Bawa (1976), Brown (1976), and Bawa, Brown, and Klein (1979) illustrate how to use the predictive return distribution in order to incorporate parameter uncertainty and estimation risk in the context of static portfolio choice problems. In an important contribution, Kandel and Stambaugh (1996) were the first to stress the significance of parameter uncertainty for the portfolio choices of a short-horizon investor when returns are predictable.

When the portfolio allocation problem is intertemporal, the investor who wishes to take parameter uncertainty into account faces an additional challenge. She has to acknowledge that future data will provide more information about the unknown parameters and integrate this aspect into her decision making. This feature of the problem introduces the notion of Bayesian updating or learning. A number of papers are devoted to the study of learning in a dynamic setting. Detemple (1986) analyzes an economy in which agents learn about the state variables that determine production through filtering of the noisy signals they receive. Gennotte (1986) studies optimal portfolio choices in an economy where investors learn about the unknown expected returns on investment. In a continuous time setting, Brennan (1998) examines the effect of learning about the unknown drift of the risky asset price process which is assumed to follow a geometric Brownian motion. Xia (2001) extends the analysis of Brennan (1998) to the case of predictable returns and considers the effect of learning about the slope in the predictive regression. Barberis (2000) studies the problem of learning about the expected risk premium in a discrete time framework with independent and identically distributed (i.i.d.) risky asset returns. In a more recent paper, Brandt et al. (2005) consider the case of predictable returns and analyze the problem of learning about the parameters in the predictive vector autoregressive (VAR) system.

The majority of papers on learning focus on particular parameters in the model describing the data generating process while they assume that the remaining parameters are known to the investor.

For instance, Brennan (1998) and Barberis (2000) assume that return volatility is known in their i.i.d. setting. Using a continuous time model for predictable returns, Xia (2001) focuses on learning about the predictive regression slope while assuming that the additional model parameters are known. It can be argued, however, that focusing on an individual parameter restricts the analysis and does not shed light on the significance of learning in a more realistic setting that treats *all* parameters as unknown.

We contribute to the literature by providing a comprehensive study of the role of learning in the context of dynamic portfolio choice. To contain the computational complexity at a manageable level, we consider a simple investment opportunity set consisting of a risky asset and a risk-free asset with constant return and study the problem of an investor who maximizes the expected utility of terminal wealth. To explore the importance of learning across various scenarios, we assume different data generating processes including i.i.d., autoregressive, and predictable risky asset returns.

To the best of our knowledge, the only existing paper that addresses learning about all model parameters is Brandt et al. (2005). While our analysis also covers the case of predictable returns studied by Brandt et al. (2005), there are important differences between the two papers that are briefly summarized as follows. First, we identify eight variables that characterize the investment opportunity set in the investor's perception and thus can serve as state variables for the investor's portfolio choice problem. The eight state variables consist of the current value of the predictor and seven variables determining the posterior distribution of the parameters. In contrast, Brandt et al. (2005) use 11 state variables in their analysis. Our view is that using a more parsimonious representation of the state will result in a more accurate approximation of the value function. Second, this paper uses a different numerical technique to solve the Bellman equation. Brandt et al. (2005) develop a novel method that uses a Taylor approximation to the value function and simulation to evaluate the conditional expectations in the Bellman equation. They illustrate the effectiveness of their method for a simple portfolio choice problem with a single state variable. However, it is not clear how accurate their solution technique is when the state variable is high dimensional. In contrast, we solve the portfolio choice problem following standard backward induction but using a powerful approximation to the value function based on feedforward neural networks. While this type of approximation has been used in the context of other dynamic programming problems (see Bertsekas and Tsitsiklis (1996)), this paper presents, as far as we know, its first application for solving a portfolio choice problem. The advantage of our approach is that we can assess the quality of the approximation by examining the neural network fit in each step of the backward induction. Finally, although the empirical findings of the two papers are not directly comparable since we use

different data sets and different priors, our results suggest that the effect of learning might not be as dramatic as shown by the analysis in Brandt et al. (2005).

Our analysis is carried out within a discrete-time framework. For each model under consideration, we describe in detail the learning process of the Bayesian investor who solves her intertemporal portfolio allocation problem taking into account that she will use future data to update her beliefs about the parameters. Building on the analysis of Barberis (2000) who treats the case of i.i.d. returns with known volatility, we identify the sufficient statistics that summarize the posterior parameter distribution and form part of the state variable for the investor's portfolio allocation problem. As shown by Barberis (2000), when returns are assumed to be i.i.d. with unknown mean but known volatility, the investor uses the return sample average to learn about the population mean. Here, the return sample mean, alone, is enough to determine the investment opportunity set according to the investor's perception and therefore serves as the single state variable. When both return mean and volatility are treated as unknown, the investor also uses the return sample variance to learn about volatility. As a result, the investor uses the first two sample moments of returns as state variables to solve her portfolio choice problem. To assess the role of potential serial correlation, we also consider an autoregressive model for returns. In this case, the investor learns about the intercept, the slope, and the conditional variance in the autoregression using the return sample mean, variance, and autocorrelation. Thus, the investor uses these sample statistics to learn about the unknown parameters, which combined with the current return level form the four state variables for the investor's portfolio problem.

When returns are assumed to be predictable, the Bayesian investor has to learn about the intercepts, the slopes, and the covariance matrix parameters in the predictive VAR system, a total of seven parameters. In this case, we identify seven sample statistics that characterize the posterior parameter distribution: the first two sample moments of the return and the predictor, the sample contemporaneous return-predictor correlation, the sample correlation between return and lagged predictor realizations, and the sample predictor serial correlation. Combining these variables with the current value of the predictor, we obtain the eight state variables used to solve the investor's optimization problem. Given the proliferation of state variables introduced by learning, the solution of the dynamic portfolio allocation problem poses a serious computational challenge. We employ a powerful approximation scheme based on feedforward neural networks to approximate the value function and solve the intertemporal optimization problem by backward induction.

Our main empirical findings are summarized as follows. We find that, in general, parameter uncertainty and learning induce negative horizon effects. This evidence is consistent across all

different models. In addition, we find that an investor who chooses to ignore learning might suffer significant losses as measured by the certainty equivalent return on wealth. The extent of these losses, however, appears to be highly dependent on the amount of data used by the investor in the initial estimation and the investor's risk aversion. The optimal allocation is most sensitive to the state variable that the investor uses to learn about unconditional expected returns, namely the return sample mean. We conclude, not surprisingly, that learning about the unconditional risk premium is the most important part of the learning process. In contrast, learning about volatility is not as important and the losses incurred by the investor who treats volatility as known are not substantial. When returns are predictable and the parameters are assumed to be known, the optimal allocation to the risky asset is horizon-dependent if return and predictor innovations are contemporaneously correlated. When the earnings-to-price ratio is used as the predictor, the horizon effect is positive and the investor allocates more heavily on the risky asset for longer horizons. Using an empirical Bayes prior, we find that learning reduces, but does not necessarily eliminate, the positive hedging demands due to predictability and negative correlation between shocks to return and earnings-to-price ratio.

The remainder of the paper is structured as follows. In Section 2, we describe a general discrete-time dynamic portfolio choice framework for a Bayesian investor who wishes to incorporate parameter uncertainty explicitly into her decision making. In Section 3, this framework is applied to four different data generating processes: i.i.d. returns with known and unknown volatility, autoregressive returns, and predictable returns with an exogenous predictor. In the final section, we offer some concluding remarks. Additional technical material is relegated to the Appendices.

2 Portfolio allocation with learning: General framework

In this section, we develop the framework for the portfolio choice problem of a Bayesian investor. From the standpoint of such an investor, the parameters of the model employed are unknown. The uncertainty about the unknown parameters is summarized by the posterior distribution of the parameters given the observed data up to a specific point in time. Hence, the investor is involved in a dynamic learning process and must take this feature of the decision process explicitly into account. It follows that the investor's posterior beliefs about the parameters will be part of the state variable that determines the investment opportunity set according to the investor's perception. We next describe in detail the portfolio choice faced by the Bayesian investor. The development builds on and generalizes the framework advanced by Barberis (2000).

For tractability reasons, we consider a simple investment opportunity set in which there are two assets: a riskless asset with constant continuously compounded return r_f and a risky asset, such as a stock index, with continuously compounded excess return r_t over period t . We assume that the investor has a CRRA power utility function and maximizes terminal wealth. The investor observes data up to time K and has an investment horizon of T periods. The time horizon is divided into N intervals of equal length (equal to $L = T/N$ periods) denoted by $[t_0, t_1], \dots, [t_{N-1}, t_N]$ where $t_0 = K$ and $t_N = K + T$. In general, we have $t_j = K + jL$ for $j = 0, 1, \dots, N$. Although the investor observes the data in every period, she only adjusts the portfolio every L periods, at points (t_0, \dots, t_{N-1}) . Denoting by W_t the wealth of the investor at time t , the evolution of wealth from time t_j to time t_{j+1} is described by

$$W_{t_{j+1}} = W_{t_j} R_{p,j+1},$$

where the portfolio return $R_{p,j+1}$ is given by

$$R_{p,j+1} = (1 - \omega_j) \exp(r_f L) + \omega_j \exp(r_f L + R_{1,j+1}), \quad (1)$$

ω_j is the allocation to the risky asset at time t_j , and $R_{1,j+1}$ is the cumulative excess return on the risky asset over the period from t_j to t_{j+1} :

$$R_{1,j+1} = r_{t_{j+1}} + \dots + r_{t_j}. \quad (2)$$

In this paper, we consider an investor who seeks to maximize utility of terminal wealth. Formally, the investor's problem is expressed as

$$\max_{\omega_0, \dots, \omega_{N-1}} E_K \left[\frac{W_{K+T}^{1-\gamma}}{1-\gamma} \right],$$

where $\omega_0, \dots, \omega_{N-1}$ are the allocations to the risky asset at times t_0, \dots, t_{N-1} , respectively, and γ is the coefficient of relative risk aversion. The state variable at time t is denoted by $s(t)$. As explained above, the state variable should include the investor's posterior beliefs about the parameters. Typically, the state variable contains a sufficient statistic that characterizes the posterior distribution of the parameters and hence the predictive distribution of future returns given the available data. In the context of normally distributed i.i.d. risky asset returns with unknown mean and known volatility, Barberis (2000) shows that the state variable is the historical average of past returns. If the volatility is unknown or the i.i.d. assumption is relaxed, additional sample statistics are needed to determine the risky return predictive distribution. If the data generating process postulated

by the investor is more complicated, the state variable $s(t)$ may even be infinite dimensional, rendering the solution of the dynamic portfolio choice problem computationally infeasible. The data generating processes we consider in this paper are based on Gaussian disturbances, ensuring that the state variable for the investor's problem stays finite dimensional. Let $D_{t_j, t_{j+1}}$ denote all the relevant data obtained after time t_j and up to and including time t_{j+1} . Clearly, $D_{t_j, t_{j+1}}$ includes all return data $r_{t_{j+1}}, \dots, r_{t_{j+1}}$. When the data generating process incorporates a predictive variable x , the data set $D_{t_j, t_{j+1}}$ also includes realizations of the variable x . Let the law of motion of the state variable be given by

$$s(t_{j+1}) = \Phi_t(s(t_j), D_{t_j, t_{j+1}}), \quad (3)$$

where Φ_t is a suitably chosen function that depends on the hypothesized data generating process. The derived utility of the investor at time t_j is given by

$$J(W_{t_j}, s(t_j), t_j) = \max_{\omega_j, \dots, \omega_{N-1}} E_{t_j} \left[\frac{W_{t_N}^{1-\gamma}}{1-\gamma} \right],$$

and the Bellman equation is

$$J(W_{t_j}, s(t_j), t_j) = \max_{\omega_j} E_{t_j} [J(W_{t_{j+1}}, s(t_{j+1}), t_{j+1})].$$

Note that the investor's expectation $E_{t_j}[\cdot]$ is taken with respect to the joint predictive distribution of returns and state variables $(R_{1,j+1}, s(t_{j+1}))$, given all available data up to time t_j or, equivalently, given the state variable $s(t_j)$. Utilizing the homotheticity of the utility function, we can write

$$J(W_j, s(t_j), t_j) = \frac{W_j^{1-\gamma}}{1-\gamma} V(s(t_j), t_j).$$

We observe that, under the commonly made assumption $\gamma > 1$, the Bellman equation becomes

$$V(s(t_j), t_j) = \min_{\omega_j} E_{t_j} [R_{p,j+1}^{1-\gamma} V(s(t_{j+1}), t_{j+1})],$$

where the portfolio return $R_{p,j+1}$ is given by equation (1). The Bellman equation is solved numerically by backward induction. Typically, the conditional expectation $E_{t_j}[\cdot]$ cannot be evaluated analytically. The alternative is to resort to Monte Carlo simulation in order to generate samples from the desired predictive distribution. Let ξ denote the vector of parameters used by the data generating process. To obtain the predictive distribution of $D_{t_j, t_{j+1}}$ given the data D_{t_j} up to time t_j , we employ the standard procedure of conditioning on the parameters and then integrating out. This yields the expression

$$p(D_{t_j, t_{j+1}} | D_{t_j}) = \int p(D_{t_j, t_{j+1}} | \xi, D_{t_j}) p(\xi | D_{t_j}) d\xi.$$

Note that since the state variable $s(t_j)$ includes a sufficient statistic summarizing the posterior distribution $p(\xi|D_{t_j})$, we actually have $p(\xi|D_{t_j}) = p(\xi|s(t_j))$. Furthermore, it is typically the case that all information contained in D_{t_j} that is relevant for the distribution $p(D_{t_j,t_{j+1}}|\xi, D_{t_j})$ is contained in the state variable $s(t_j)$. That is, $s(t_j)$ serves as a sufficient statistic and so $p(D_{t_j,t_{j+1}}|\xi, D_{t_j}) = p(D_{t_j,t_{j+1}}|\xi, s(t_j))$. Summarizing, we obtain

$$p(D_{t_j,t_{j+1}}|D_{t_j}) = \int p(D_{t_j,t_{j+1}}|\xi, s(t_j))p(\xi|s(t_j))d\xi.$$

Therefore we can use the following algorithm to generate samples from the predictive distribution $p(R_{1,j+1}, s(t_{j+1})|s(t_j))$:

1. Use the given value of the state variable $s(t_j)$ to simulate ξ^* from the posterior distribution $p(\xi|s(t_j))$.
2. Use the obtained parameter value ξ^* and the given value of the state variable $s(t_j)$ to simulate $D_{t_j,t_{j+1}}^*$ from $p(D_{t_j,t_{j+1}}|\xi^*, s(t_j))$.
3. Use the simulated data $D_{t_j,t_{j+1}}^*$, equation (2) and the law of motion (3) to generate draws $R_{1,j+1}^*$ and $s^*(t_{j+1})$ from the predictive distribution $p(R_{1,j+1}, s(t_{j+1})|s(t_j))$.

Repeating steps 1 through 3 sufficiently many times will produce a sample from the desired predictive distribution. The actual implementation differs on a case-by-case basis, as we will illustrate in the next section. Let us denote by $\{(R_{1,j+1}^{(i)}, s^{(i)}(t_{j+1})) : i = 1, \dots, I\}$ the sample from the predictive distribution. Then, in the typical backward induction step, for a given point $s(t_j)$ in the state space, we can compute (an approximation to) the function V at time t_j by solving

$$V(s(t_j), t_j) = \min_{\omega_j} \left[\frac{1}{I} \sum_{i=1}^I \left[R_{p,j+1}^{(i)} \right]^{1-\gamma} V_{j+1}(s^{(i)}(t_{j+1}), t_{j+1}) \right] \quad (4)$$

where

$$R_{p,j+1}^{(i)} = (1 - \omega_j) \exp(r_f L) + \omega_j \exp(r_f L + R_{1,j+1}^{(i)}).$$

The function $V(\cdot, t_j)$ is computed at a large number of suitably chosen points evenly distributed on the state space and then an approximation to $V(\cdot, t_j)$ is obtained using feedforward neural networks. Details about the implementation of the neural network approximation are deferred to Appendix A.

3 Data generating process

We consider four data generating processes: i.i.d. returns with known and unknown volatility, autoregressive returns, and predictable returns with an exogenous predictor. We assume throughout

that the investor perceives all random shocks to be normally distributed. The simple case of i.i.d. returns with known volatility has been analyzed by Brennan (1998) in a continuous-time diffusion setting and by Barberis (2000) in a discrete-time framework using a different numerical technique for solving the Bellman equation. However, we include this case in our analysis since it serves as the natural starting point and helps build the necessary intuition for the study of the more complex problems with learning that we consider later. Moreover, we provide evidence on the performance of alternative portfolio rules in terms of certainty equivalent return on wealth and illustrate the sensitivity of the results to the amount of data used in the estimation, as well as to the risk aversion of the investor.

3.1 IID returns with known volatility

3.1.1 Bayesian updating framework

In this section, we examine the effect of learning on the portfolio decisions made by the investor when returns are assumed to be i.i.d. normal with known variance but unknown mean. The continuously compounded excess returns follow the data generating process

$$r_t = \mu + \varepsilon_t \text{ with } \varepsilon_t \sim \text{i.i.d. } N(0, v),$$

where the variance v is assumed to be known. Instead of the improper diffuse prior used by Barberis (2000), we employ a proper prior on μ given by

$$\mu \sim N(\mu_0, v_0).$$

In our empirical application, we set the precision $1/v_0$ equal to ε , a small positive constant. This guarantees a proper but non-informative prior. The investor updates her beliefs about the unknown expected return μ using Bayes rule. It follows that, given return data up to and including time t , $D_t = (r_1, \dots, r_t)$, the posterior distribution of μ is also normal. Specifically, it is shown in part B7 of Appendix B that the posterior distribution of μ is given by

$$\mu|D_t \sim N(\tilde{\mu}_t, \tilde{v}_t), \tag{5}$$

where

$$\tilde{\mu}_t = \left(\frac{v_t}{v_0 + v_t} \right) \mu_0 + \left(\frac{v_0}{v_0 + v_t} \right) \bar{r}_t, \tag{6}$$

$$\tilde{v}_t = v_t \left(\frac{v_0}{v_0 + v_t} \right), \tag{7}$$

and

$$\bar{r}_t = \frac{1}{t} \sum_{\tau=1}^t r_\tau, \quad v_t = \frac{v}{t}.$$

Since the variance v is assumed to be known, it follows from equations (6) and (7) that the posterior distribution of μ is summarized by the sufficient statistic \bar{r}_t . For each t , we define the variable

$$s(t) = \bar{r}_t$$

which will serve as the state variable for the portfolio allocation problem. It follows from the i.i.d. assumption that the conditional distribution of the cumulative return $R_{1,j+1}$ is given by

$$R_{1,j+1} | \mu, D_{t_j} \sim N(L\mu, Lv).$$

Recalling that $\mu | D_t \sim N(\tilde{\mu}_t, \tilde{v}_t)$ and using a standard calculation to integrate out μ , we obtain the following predictive distribution

$$R_{1,j+1} | D_{t_j} \sim N(L\tilde{\mu}_{t_j}, L(v + L\tilde{v}_{t_j})). \quad (8)$$

The derivation of this fact can be found in part B7 of Appendix B. As in the analysis of Barberis (2000), the law of motion of $s(t)$ is given by⁴

$$s(t_{j+1}) = \frac{1}{t_{j+1}} \sum_{\tau=1}^{t_{j+1}} r_\tau = \frac{t_j}{t_{j+1}} s(t_j) + \frac{1}{t_{j+1}} R_{1,j+1}. \quad (9)$$

Given the state variable $s(t_j)$, we can use (8) and (9) to obtain a large number of random draws $R_{1,j+1}^*$ and $s^*(t_{j+1})$ from the predictive distribution $p(R_{1,j+1}, s(t_{j+1}) | s(t_j))$ that can be used to solve the approximate Bellman equation (4). The numerical solution proceeds as described in section 2 using backward induction.

3.1.2 Empirical results

In this subsection, we examine the empirical implications of learning about the return mean using the framework developed in the previous subsection. We consider an investor who has access to quarterly return data but only updates her portfolio holdings at the yearly frequency ($L = 4$ quarters). The underlying assumption is that the returns data are normal i.i.d. with known variance but unknown mean. The risky asset return is the return on the S&P 500 index. The source of our data set, which includes data up to 2003, is the recent paper by Goyal and Welch (2005).⁵ The

⁴See also the related discussion of recursive mean estimation given in Example 3.4, Section 3.2 in Bertsekas and Tsitsiklis (1996).

⁵The data can be found at <http://www.bus.emory.edu/AGoyal/>. We thank Professor Goyal for making the data publicly available.

investor uses a non-informative prior distribution for the return mean μ given by $\mu \sim N(\mu_0, v_0)$ with mean $\mu_0 = 0$ and precision $1/v_0 = \varepsilon = 10^{-4}$. The investment horizon is $N = 10$ years, so that T is equal to 40 quarters, and the (known) return variance is set equal to the sample variance using all the data available at the beginning of the investment period. Throughout, we use $I = 300,000$ Monte Carlo repetitions to evaluate the expectations in the (approximate) Bellman equation (4). We approximate the value function on the entire state space using a grid of 200 evenly distributed points and a feedforward neural network with two hidden layers. To provide a comprehensive picture of the effect of learning, we examine eight different scenarios in which the coefficient of relative risk aversion (γ) takes the commonly used values 5 or 10, the initial estimation is based on 15 or 30 years of past data, and the annualized risk-free rate (R_f) is assumed to be equal to 5 or 6 percent.

In Table 1, we present the optimal allocation to the risky asset for all eight scenarios. Figure 1 displays graphically the optimal allocation for the case in which $\gamma = 5$, $R_f = 6$ percent, and the initial estimation is based on 15 years of data with the sample ending in 2003. Consistent with the results of Brennan (1998) and Barberis (2000), we find that learning about the mean induces negative horizon effects. In all cases, the investor allocates less to the risky asset the longer the investment horizon. Compared to the policy that assumes the mean to be a known constant and equal to the sample mean calculated at the beginning of the investment period (denoted by C in the table), there is less investment in the risky asset according to the optimal learning policy (denoted by OL in the table). As the investor gets closer to the final date, the OL allocation rule converges to the C allocation rule. Overall, the importance of learning about the mean appears to depend on two factors: the level of risk aversion and the amount of data used by the investor at the beginning of the investment period. The effect of learning is more pronounced for the low level of risk aversion ($\gamma = 5$). Uncertainty about the mean makes the risky asset appear riskier in the investor's perception as compared to the case of known mean. For high levels of risk aversion, the difference becomes less important.

Furthermore, we expect the impact of learning to diminish when more data are used in the initial estimation. This is what we observe in Table 1. When the investor uses 30 years of data, as opposed to just 15, the negative hedging demands are reduced. The final point to notice from this table is that the optimal allocation heavily depends on the estimation period itself. For instance, according to the C allocation rule that ignores learning, when $\gamma = 5$, $R_f = 6$ percent, and using 15 years of data, the allocation to the risky asset equals 90.7 percent when the sample ends in 2000 while it equals 55.7 percent when the sample ends in 2003. For the OL allocation rule and investment

horizon of 10 years, the corresponding allocations are 52.8 and 35.9 percent, respectively. Overall, we observe that taking estimation risk and learning into account results in more robust portfolio allocations. In Table 2, we present evidence on the performance of the OL allocation policy; the C allocation policy; and a policy, denoted by MU in the table, based on myopic updating. That is, at each given time, the policy MU uses all available data to update the estimate of μ and then, myopically and ignoring future learning, solves the one-period-ahead portfolio allocation problem. Table 2 reveals that the investor can suffer significant losses if she uses the C allocation rule instead of the optimal OL rule. The losses associated with the suboptimal myopic rule MU that dynamically updates the mean estimates are much smaller. However, the losses are significantly reduced for high risk aversion ($\gamma = 10$) and/or when the initial estimation is based on 30, instead of only 15, years of data.

3.2 IID returns and learning with unknown volatility

In this subsection, we relax the assumption that the volatility is known. The investor solves the portfolio allocation problem taking into account that she will learn about both return mean and volatility from future data.

3.2.1 Bayesian updating framework

As in subsection 3.1, the continuously compounded excess returns are assumed to be i.i.d. and normally distributed

$$r_t = \mu + \varepsilon_t \text{ with } \varepsilon_t \sim \text{i.i.d. } N(0, v).$$

In the present context, however, in addition to the return mean μ , the variance v is also treated as unknown. To keep the problem tractable, we assume that the investor has a conjugate proper prior on (μ, v) given by

$$\mu|v \sim N(\mu_0, v/\pi_0) \text{ and } v \sim IG(\alpha_0, \beta_0)$$

where $\mu_0 \in \mathbb{R}$ and π_0 , α_0 , and β_0 are positive parameters. This is known as the normal inverse gamma (NIG) prior (see B2 in Appendix B for the definition). In our empirical application, we set $\pi_0 = \alpha_0 = 1/\beta_0 = \varepsilon$ where ε is a small positive number to ensure a proper but non-informative prior. Recall from the analysis in subsection 3.1.1 that when the expected return μ is the single unknown parameter, the investor uses the return sample mean as the single sufficient statistic in

her Bayesian updating. When the return variance is also unknown, the investor's updating process requires two sufficient statistics. Next, we show how this intuition is formalized.

Given return data up to and including time t , $D_t = (r_1, \dots, r_t)$, the posterior distribution of (μ, v) is also of the NIG type. This follows as a special case of the statement in part B8 of Appendix B. Specifically, the posterior distribution of (μ, v) is described by

$$v|D_t \sim IG(\tilde{\alpha}_t, \tilde{\beta}_t), \quad \mu|D_t, v \sim N(\tilde{\mu}_t, v/\tilde{\pi}_t),$$

where

$$\tilde{\alpha}_t = \alpha_0 + \frac{t}{2}, \tag{10}$$

$$\tilde{\beta}_t^{-1} = \beta_0^{-1} + \frac{\pi_0 t (\bar{r}_t - \mu_0)^2}{2(t + \pi_0)} + \frac{t \hat{v}_t}{2}, \tag{11}$$

$$\tilde{\mu}_t = \frac{t}{t + \pi_0} \bar{r}_t + \frac{\pi_0}{t + \pi_0} \mu_0, \tag{12}$$

$$\tilde{\pi}_t = \pi_0 + t, \tag{13}$$

and

$$\bar{r}_t = \frac{1}{t} \sum_{\tau=1}^t r_\tau, \quad \hat{v}_t = \frac{1}{t} \sum_{\tau=1}^t (r_\tau - \bar{r}_t)^2 = \frac{1}{t} \sum_{\tau=1}^t r_\tau^2 - \bar{r}_t^2.$$

For each t , we define the variables

$$m_1(t) = \frac{1}{t} \sum_{\tau=1}^t r_\tau, \quad m_2(t) = \frac{1}{t} \sum_{\tau=1}^t r_\tau^2,$$

and observe that $\bar{r}_t = m_1(t)$ and $\hat{v}_t = m_2(t) - m_1(t)^2$. Examination of equations (10)-(13) shows that, as far as the posterior distribution of the parameters (μ, v) is concerned, the information contained in the observed data D_t can be summarized by the two variables $m_1(t)$ and $m_2(t)$. Hence, the variables $m_1(t)$ and $m_2(t)$ can serve as state variables for the investor's portfolio choice problem. The natural extension of equation (9) in the present context with unknown volatility provides the law of motion of the state variables $(m_1(t), m_2(t))$:

$$m_k(t_{j+1}) = \frac{t_j}{t_{j+1}} m_k(t_j) + \frac{1}{t_{j+1}} R_{k,j+1}, \quad k = 1, 2,$$

where

$$R_{k,j+1} = r_{t_{j+1}}^k + \dots + r_{t_j}^k, \quad k = 1, 2. \tag{14}$$

Although we could proceed and use $m(t) = (m_1(t), m_2(t))$ as the state variable vector, our results are more easily interpretable if we transform $m(t)$ to $s(t) = (s_1(t), s_2(t))$ where

$$\begin{aligned} s_1(t) &= m_1(t) = \bar{r}_t, \\ s_2(t) &= m_2(t) - m_1(t)^2 = \hat{v}_t. \end{aligned}$$

Clearly, $m(t)$ and $s(t)$ contain the same information. Indeed, we can recover $m(t)$ from $s(t)$ using the identities $m_1(t) = s_1(t)$ and $m_2(t) = s_2(t) + s_1(t)^2$. Using the definition of $s(t)$ and the law of motion of $m(t)$, we can deduce the law of motion of the state variable $s(t)$:

$$s_1(t_{j+1}) = \frac{t_j}{t_{j+1}} s_1(t_j) + \frac{1}{t_{j+1}} R_{1,j+1}, \quad (15)$$

$$s_2(t_{j+1}) = \frac{t_j}{t_{j+1}} [s_2(t_j) + s_1(t_j)^2] + \frac{1}{t_{j+1}} R_{2,j+1} - s_1(t_{j+1})^2. \quad (16)$$

The subsequent analysis uses the state variable $s(t)$. Solving the Bellman equation requires knowledge of the predictive distribution of $(R_{1,j+1}, s(t_{j+1}))$, given the state variable vector $s(t_j)$. Since this is not available in analytic closed form, we resort to Monte Carlo simulation in order to generate samples from the desired predictive distribution. Conditioning on the parameters and integrating out, we obtain the expression

$$p(D_{t_j, t_{j+1}} | s(t_j)) = \int p(D_{t_j, t_{j+1}} | \mu, v, s(t_j)) p(\mu, v | s(t_j)) d\mu dv.$$

Since returns are assumed to be i.i.d., the distribution $p(D_{t_j, t_{j+1}} | \mu, v, s(t_j))$ does not depend on $s(t_j)$. On the other hand, we know that the posterior distribution $p(\mu, v | s(t_j))$ is normal inverse gamma, as described above. Therefore, we can devise the following algorithm to simulate from the predictive distribution:

1a. Generate v^* from $v | s(t_j) \sim IG(\tilde{\alpha}_{t_j}, \tilde{\beta}_{t_j})$, where $\tilde{\alpha}_{t_j}$ and $\tilde{\beta}_{t_j}$ are obtained using $s(t_j)$ and expressions (10) and (11), respectively.

1b. Given v^* from step 1a, generate μ^* from $\mu | (v^*, s(t_j)) \sim N(\tilde{\mu}_{t_j}, v^* / \tilde{\pi}_{t_j})$, where $\tilde{\mu}_{t_j}$ and $\tilde{\pi}_{t_j}$ are obtained using $s(t_j)$ and expressions (12) and (13), respectively.

2. Given (μ^*, v^*) from step 1, generate an i.i.d. sample $D^* = (r_{*,1}, \dots, r_{*,L})$ from $N(\mu^*, v^*)$.

3. Given the generated sample $(r_{*,1}, \dots, r_{*,L})$, use equation (14) to obtain $(R_{1,j+1}^*, R_{2,j+1}^*)$ and then use $s(t_j)$ and the law of motion (equations (15) and (16)) to obtain a draw $s^*(t_{j+1})$ from the predictive distribution of the state variable in the next period.

Repeating the above procedure sufficiently many times provides a large sample from the predictive distribution $p(R_{1,j+1}, s(t_{j+1}) | s(t_j))$, which we use to solve the approximate Bellman equation

(4). We then follow the procedure described in section 2 to solve the portfolio allocation problem using backward induction in the standard fashion.

3.2.2 Empirical results

In this subsection, employing the framework described in the preceding subsection, we provide empirical evidence on the effect of learning about the return mean and volatility on dynamic portfolio choices. We use the same data and assumptions as in subsection 3.1.2, except that in this case the investor is learning about both return mean μ and variance v . The prior distribution of (μ, v) is given by $\mu|v \sim N(\mu_0, v/\pi_0)$ and $v \sim IG(\alpha_0, \beta_0)$ with $\mu_0 = 0$ and $\pi_0 = \alpha_0 = 1/\beta_0 = \varepsilon = 10^{-4}$. This specification provides a proper but non-informative prior. As in subsection 3.1.2, the investor uses quarterly return data in her estimation, updates her portfolio holdings yearly, and has an investment horizon of 10 years. The expectations in the (approximate) Bellman equation (4) are computed using $I = 300,000$. We approximate the value function on the entire state space using a grid of 600 evenly distributed points and a feedforward neural network with two hidden layers. We consider eight different scenarios as in subsection 3.1.2. Table 3 presents the optimal allocation to the risky asset in all eight scenarios. The upper and lower panels correspond to the two cases when the initial sample the investor uses in her estimation ends in 2000 and 2003, respectively. In general, we observe the same patterns as in the case of unknown mean and known volatility; for comparison, see subsection 3.1.2 and Table 1. Learning induces negative hedging demands and the investor, following the optimal rule with learning (OL), allocates less in the risky asset compared to the C policy that treats both the mean μ and the variance v as known and equal to their sample counterparts obtained using all data at the beginning of the investment period. The OL allocation rule converges to the C allocation rule as the investor approaches the final date. The additional effect of learning about the variance v appears to be limited. In particular, when the initial sample ends in 2003, there is essentially no difference between the cases with known and unknown volatility. Note that for the samples ending in 2003, the return sample means are much lower than their counterparts using the samples ending in 2000. The underlying intuition is that when the estimates about expected returns are low the uncertainty about volatility becomes less important.

This point can also be seen in Figures 2, 3, and 4 which display graphically the optimal allocation for the case in which $\gamma = 5$, $R_f = 6$ percent, and the initial estimation is based on 15 years of data with the sample ending in 2003. In the upper graph of Figure 2, we set the state variable s_2 equal to its 25th, 50th, and 75th percentile and plot the optimal allocation as a function of the state variable

s_1 . In the lower graph, the roles of s_1 and s_2 are reversed. As expected, everything else equal, higher levels of the mean estimate s_1 and lower levels of the variance estimate s_2 correspond to higher risky asset allocations. Moreover, we observe that learning about volatility makes a difference only when return mean estimates are high. In Figures 3 and 4, we present the optimal allocation as a function of the investment horizon. Figure 3 shows that uncertainty about the return mean is as important as in the case with known volatility. Figure 3 shows that the importance of learning about volatility depends on the level of the mean estimate s_1 . For example, the upper graph in Figure 4 shows that when the state variable s_1 is equal to its 25th percentile, uncertainty about return volatility plays practically no role. Overall, the evidence suggests that learning about the return mean is more important than learning about the return volatility.

In Table 4, we present evidence on the performance of the OL allocation policy in comparison with two suboptimal alternatives. The first, denoted by OL-KV in the table, treats the return variance v as known and equal to the return sample variance \hat{v} at the beginning of the investment period but is based on optimal learning about the return mean μ . The second, denoted by MU in the table, is based on myopic updating. More specifically, at each given time, the policy MU uses all available data to update the estimates of μ and v and then, myopically and ignoring future learning, solves the one-period-ahead portfolio allocation problem. The losses associated with the policy OL-KV, in terms of certainty equivalent return, are insignificant: from 1 to 4 basis points when the initial sample ends in 2000 and from 0 to 1 basis points when the initial sample ends in 2003. This evidence agrees with the point made earlier that learning about volatility is much less important than learning about the mean. The losses associated with the MU allocation rule are more significant. For example, when $\gamma = 5$, $R_f = 5$ percent, and the initial estimation is based on 15 years of data with the sample ending in 2000, the investor incurs a loss of 28 basis points if the MU policy is used as opposed to the optimal policy OL. However, when $\gamma = 10$, $R_f = 6$ percent, and the initial estimation is based on 30 years of data with the sample ending in 2003, the loss is only 2 basis points. In general, these losses are reduced when more data are used in the initial estimation and/or when the investor is more risk averse.

3.3 Autoregressive returns

In this section, we relax the assumption of i.i.d. returns and examine the role of potential return serial correlation. For tractability reasons, we use an autoregressive model of order one to describe the return dynamics. In this context, the investor is uncertain about the intercept, the slope, and the conditional variance in the return autoregression and learning about these parameters is

explicitly incorporated in the solution of the dynamic portfolio choice problem.

3.3.1 Bayesian updating framework

We assume that the continuously compounded excess returns follow the autoregressive model

$$r_\tau = a + br_{\tau-1} + \varepsilon_\tau, \quad \tau = 1, \dots, t, \quad \varepsilon_\tau \sim \text{i.i.d. } N(0, v). \quad (17)$$

The model can be rewritten in concise vector-matrix form as

$$Y_t = X_t \theta + E_t$$

where

$$\theta = [a \quad b]'$$

$$Y_t = [r_1 \quad \dots \quad r_t]', \quad X_t = \begin{bmatrix} 1 & \dots & 1 \\ r_0 & \dots & r_{t-1} \end{bmatrix}', \quad E_t = [\varepsilon_1 \quad \dots \quad \varepsilon_t]'$$

We assume that the prior of the parameter (θ, v) is of the conjugate NIG type (see B2 in Appendix B for the definition). Specifically, the prior of v is $IG(\alpha_0, \beta_0)$ and the prior of θ given v is $N(\theta_0, v\Pi_0^{-1})$ where α_0 and β_0 are positive, $\theta_0 \in \mathbb{R}^2$, and Π_0 is a 2×2 symmetric and positive definite matrix. Given data $D_t = (Y_t, X_t)$, the posterior of (θ, v) is also of the NIG type. Specifically, it follows from part B8 in Appendix B that

$$\begin{aligned} v|D_t &\sim IG(\tilde{\alpha}_t, \tilde{\beta}_t), \\ \theta|v, D_t &\sim N(\tilde{\theta}_t, v\tilde{\Pi}_t^{-1}) \end{aligned}$$

where

$$\tilde{\alpha}_t = \alpha_0 + \frac{t}{2} \quad (18)$$

$$\tilde{\beta}_t^{-1} = \beta_0^{-1} + \frac{t\hat{v}_t}{2} + \frac{(\hat{\theta}_t - \theta_0)' [\Pi_0^{-1} + (X_t' X_t)^{-1}]^{-1} (\hat{\theta}_t - \theta_0)}{2}, \quad (19)$$

$$\tilde{\theta}_t = \tilde{\Pi}_t^{-1} [\Pi_0 \theta_0 + (X_t' X_t) \hat{\theta}_t], \quad (20)$$

$$\tilde{\Pi}_t = \Pi_0 + X_t' X_t \quad (21)$$

with

$$\hat{\theta}_t = (X_t' X_t)^{-1} X_t' Y_t, \quad (22)$$

$$\hat{v}_t = \frac{1}{t}(Y_t - X_t\hat{\theta}_t)'(Y_t - X_t\hat{\theta}_t) = \frac{1}{t} [Y_t'Y_t - (X_t'Y_t)'(X_t'X_t)^{-1}(X_t'Y_t)]. \quad (23)$$

Examination of equations (18)-(21) reveals that the Bayesian updating process is based on the quantities $\hat{\theta}_t$, \hat{v}_t and $X_t'X_t$. In light of equations (22) and (23), it becomes evident that the updating process is completely characterized by the matrices $X_t'X_t$, $Y_t'Y_t$, and $X_t'Y_t$. To obtain the vector of state variables, we need to express these matrices in an appropriate form. We observe that

$$\begin{aligned} X_t'Y_t &= \begin{bmatrix} \sum_{\tau=1}^t r_\tau \\ \sum_{\tau=1}^t r_{\tau-1}r_\tau \end{bmatrix} = \begin{bmatrix} tm_1(t) \\ tm_3(t) \end{bmatrix}, \\ Y_t'Y_t &= \sum_{\tau=1}^t r_\tau^2 = tm_2(t), \\ X_t'X_t &= \begin{bmatrix} t & \sum_{\tau=1}^t r_{\tau-1} \\ \sum_{\tau=1}^t r_{\tau-1} & \sum_{\tau=1}^t r_{\tau-1}^2 \end{bmatrix} \\ &= \begin{bmatrix} t & r_0 + tm_1(t) - m_4(t) \\ r_0 + tm_1(t) - m_4(t) & r_0^2 + tm_2(t) - m_4(t)^2 \end{bmatrix}, \end{aligned}$$

where the quantities $m_k(t)$, $k = 1, 2, 3, 4$ are defined by

$$m_1(t) = \frac{1}{t} \sum_{\tau=1}^t r_\tau, \quad m_2(t) = \frac{1}{t} \sum_{\tau=1}^t r_\tau^2, \quad m_3(t) = \frac{1}{t} \sum_{\tau=1}^t r_{\tau-1}r_\tau, \quad m_4(t) = r_t.$$

Therefore, all of the information that determines the posterior parameter distribution and therefore the predictive return distribution given past data is summarized by the four variables $m_k(t)$, $k = 1, 2, 3, 4$. Therefore, these four variables can serve as state variables for the portfolio choice problem of the Bayesian investor. The issue we address next is the time evolution of the state variable vector $m(t)$. Given the value of the state variable m at time t_j and the observed return data from time $t_j + 1$ to time t_{j+1} , namely $(r_{t_j+1}, \dots, r_{t_{j+1}})$, the state variable m at time t_{j+1} is updated according to the following equations

$$\begin{aligned} m_1(t_{j+1}) &= \frac{t_j}{t_{j+1}}m_1(t_j) + \frac{1}{t_{j+1}}R_{1,j+1}, \\ m_2(t_{j+1}) &= \frac{t_j}{t_{j+1}}m_2(t_j) + \frac{1}{t_{j+1}}R_{2,j+1}, \\ m_3(t_{j+1}) &= \frac{t_j}{t_{j+1}}m_3(t_j) + \frac{1}{t_{j+1}}[m_4(t_j)r_{t_j+1} + S_{j+1}], \\ m_4(t_{j+1}) &= r_{t_{j+1}} \end{aligned}$$

where

$$R_{k,j+1} = r_{t_{j+1}}^k + \dots + r_{t_{j+1}}^k, \quad k = 1, 2, \quad (24)$$

and

$$S_{j+1} = r_{t_j+1}r_{t_j+2} + \cdots + r_{t_{j+1}-1}r_{t_{j+1}}. \quad (25)$$

As in the case of i.i.d. returns with unknown volatility studied in the previous section, we can transform the state variable vector $m(t)$ to make our results easier to interpret and computational analysis more stable. Specifically, we transform the state variable $m(t)$ to $s(t) = (s_1(t_j), \dots, s_4(t_j))$ where

$$\begin{aligned} s_1(t) &= \frac{1}{t} \sum_{\tau=1}^t r_\tau = m_1(t), \\ s_2(t) &= \frac{1}{t} \sum_{\tau=1}^t (r_\tau - \bar{r}_t)^2 = m_2(t) - m_1(t)^2, \\ s_3(t) &= \frac{\frac{1}{t} \sum_{\tau=1}^t (r_{\tau-1} - \bar{r}_t)(r_\tau - \bar{r}_t)}{\frac{1}{t} \sum_{\tau=1}^t (r_\tau - \bar{r}_t)^2} = \frac{m_3(t) - m_1(t) \left[m_1(t) + \frac{r_0 - m_4(t)}{t} \right]}{s_2(t)}, \\ s_4(t) &= r_t = m_4(t). \end{aligned}$$

The interpretation of the state variables s_1 , s_2 , and s_3 is clear: s_1 is an estimate of the unconditional expected return, s_2 is an estimate of the unconditional return variance, and s_3 is an estimate of the return serial correlation. Note that the two variables $m(t)$ and $s(t)$ contain identical information. Specifically, we can recover $m(t)$ from $s(t)$ through the identities $m_1(t) = s_1(t)$, $m_2(t) = s_2(t) + s_1(t)^2$, $m_3(t) = s_2(t)s_3(t) + s_1(t)[s_1(t) + \frac{r_0 - s_4(t)}{t}]$, and $m_4(t) = s_4(t)$. Using the definition of $s(t)$ and the law of motion of $m(t)$, it is a matter of straightforward algebra to obtain the law of motion of $s(t)$. Given the state variable vector $s(t_j)$ and data $(r_{t_j+1}, \dots, r_{t_{j+1}})$, the state variable vector $s(t_{j+1})$ is given by

$$s_1(t_{j+1}) = \frac{t_j}{t_{j+1}} s_1(t_j) + \frac{1}{t_{j+1}} R_{1,j+1}, \quad (26)$$

$$s_2(t_{j+1}) = \frac{t_j}{t_{j+1}} [s_2(t_j) + s_1(t_j)^2] + \frac{1}{t_{j+1}} R_{2,j+1} - s_1(t_{j+1})^2, \quad (27)$$

$$s_3(t_{j+1}) = \frac{\frac{t_j}{t_{j+1}} [s_2(t_j)s_3(t_j) + s_1(t_j) [s_1(t_j) + \frac{r_0 - s_4(t_j)}{t_j}]]}{s_2(t_{j+1})} \quad (28)$$

$$+ \frac{\frac{1}{t_{j+1}} [s_4(t_j)r_{t_j+1} + S_{j+1}] - s_1(t_{j+1}) [s_1(t_{j+1}) + \frac{r_0 - r_{t_j+1}}{t_{j+1}}]}{s_2(t_{j+1})},$$

$$s_4(t_{j+1}) = r_{t_{j+1}}. \quad (29)$$

The predictive distribution of $D_{t_j, t_{j+1}} = (r_{t_j+1}, \dots, r_{t_{j+1}})$ given the state variable $s(t_j)$ is required for the solution to the Bellman equation. We use Monte Carlo simulation to generate samples from

the desired predictive distribution. Conditioning on the parameters and integrating out, we obtain the expression

$$p(D_{t_j, t_{j+1}} | s(t_j)) = \int p(D_{t_j, t_{j+1}} | \theta, v, s(t_j)) p(\theta, v | s(t_j)) d\theta dv.$$

Under the autoregressive model, it follows that the distribution $p(D_{t_j, t_{j+1}} | \theta, v, s(t_j))$ only depends on θ, v , and $s_4(t_j) = r_{t_j}$. As described above, the posterior distribution $p(\theta, v | s(t_j))$ is of the NIG type. Thus, the following algorithm can be used to simulate from the predictive distribution:

1a. Generate v^* from $v | (s(t_j)) \sim IG(\tilde{\alpha}_{t_j}, \tilde{\beta}_{t_j})$, where $\tilde{\alpha}_{t_j}$ and $\tilde{\beta}_{t_j}$ are obtained using $s(t_j)$ and expressions (18) and (19), respectively.

1b. Given v^* from step 1a, generate θ^* from $\theta | (v^*, s(t_j)) \sim N(\tilde{\theta}_{t_j}, v^* \tilde{\Pi}_{t_j}^{-1})$, where $\tilde{\theta}_{t_j}$ and $\tilde{\Pi}_{t_j}$ are obtained using $s(t_j)$ and expressions (20) and (21), respectively.

2. Given (θ^*, v^*) from step 1, generate a sample $(r_{*,1}, \dots, r_{*,L})$ as follows

$$r_{*,l} = a^* + b^* r_{*,l-1} + \varepsilon_l, \quad l = 1, \dots, L,$$

where $r_{*,0} = s_4(t_j) = r_{t_j}$, $\theta^* = [a^* \quad b^*]'$, and ε_l are i.i.d. $N(0, v^*)$.

3. Given the generated sample $(r_{*,1}, \dots, r_{*,L})$, use equations (24) and (25) to obtain $R_{1,j+1}^*$, $R_{2,j+1}^*$, and S_{j+1}^* and then use $s(t_j)$ and the law of motion (equations (26)-(29)) to obtain a draw $s^*(t_{j+1})$ from the predictive distribution of the state variable in the next period.

Repeating the above procedure sufficiently many times provides a large sample from the predictive distribution $p(R_{1,j+1}, s(t_{j+1}) | s(t_j))$, which is used to solve the approximate Bellman equation. We then follow the approach detailed in section 2 to solve the portfolio allocation problem using backward induction.

3.3.2 Empirical results

In this subsection, we report our empirical findings on the effect of learning about the unknown parameters on dynamic portfolio choices when returns are assumed to follow an autoregressive process. We use the same data and assumptions as in subsections 3.1.2 and 3.2.2, except that in this case the investor is learning about the intercept a , the slope b , and the conditional variance v in the autoregression (17). The prior on (θ, v) is given by $v \sim IG(\alpha_0, \beta_0)$ and $\theta | v \sim N(\theta_0, v \Pi_0^{-1})$ with $\alpha_0 = 1/\beta_0 = \varepsilon$, $\theta_0 = (0, 0)$, and $\Pi_0 = \varepsilon I_2$, where $\varepsilon = 10^{-4}$ and I_2 is the 2×2 identity matrix. This specification ensures a proper but non-informative prior. The investor uses quarterly return data in her estimation, updates her portfolio holdings yearly, and has an investment horizon of 10 years.

The solution to the (approximate) Bellman equation uses $I = 300,000$ Monte Carlo repetitions to evaluate the conditional expectations. We approximate the value function on the entire state space using a grid of 1,500 evenly distributed points and a feedforward neural network with two hidden layers. We consider eight different scenarios as in subsections 3.1.2 and 3.2.2.

Table 5 presents the optimal allocation to the risky asset in all eight scenarios. The upper and lower panels correspond to the two cases when the initial sample the investor uses in her estimation ends in 2000 and 2003, respectively. In general, we observe the same patterns as in the two previous cases with i.i.d. returns. Learning induces negative hedging demands and the investor, following the optimal rule with learning (OL), allocates less in the risky asset for longer horizons. For comparison, see Tables 1 and 3. The lower panel in Table 5 shows that the optimal allocations are essentially the same as in the case with i.i.d. returns (with either known or unknown volatility) when the initial sample ends in 2003. This can be seen as an indication that the serial correlation present in the data is not strong enough to allow for market timing. However, in the upper panel, we see that the optimal allocations according to the autoregressive model differ from the optimal allocations according to the i.i.d. models when the initial sample ends in 2000. Specifically, they are higher (lower) when 15 (30) years are used in the initial estimation, respectively. To understand this effect, consider the first two rows in the upper panel. The last return observation (-9.35 percent) is rather low. The autocorrelation estimate (s_3) is negative (positive) when 15 (30) years of data are used. Therefore, conditional expected returns appear higher when the investor uses 15 years of data and that explains why the risky asset allocations are higher.

Figures 5 and 6 display the optimal allocation to the risky asset as a function of the investment horizon. In each subgraph, which corresponds to a specific state variable, s_k , $k = 1, 2, 3, 4$, we plot three lines which correspond to the 25th, 50th, and 75th percentiles of s_k , while we set the rest of the state variables equal to their values at the end of 2000 (Figure 5) or 2003 (Figure 6). We observe that the serial correlation in the data is so weak that changes in the final return observation (state variable s_4) do not affect the optimal allocations. Moreover, learning about unconditional mean return (state variable s_1) appears to be the most important part of the Bayesian learning process. In Table 6, we compare the optimal allocation policy OL with two suboptimal alternatives. The first, denoted by OL-IID in the table, ignores return serial correlation and treats the return data as i.i.d.. However, according to this policy, the investor treats both return mean and volatility as unknown parameters and learning about these parameters is incorporated into the portfolio choice problem as in subsection 3.2. The second policy, denoted by MU in the table, uses all available data to update the estimates of a , b , and v , but in each period, myopically and ignoring future

learning, solves the one-period portfolio allocation problem. The losses associated with using the policy OL-IID are somewhat significant, ranging from 3 to 7 basis points when the initial sample ends in 2000 and from 1 to 4 when the initial sample ends in 2003. The losses associated with using the policy MU are more significant, ranging from 3 to 31 basis points. When only 15 years of data are used in the initial estimation, the parameter estimates are rather noisy and give rise to more suboptimal allocations. Consistent with the previous results from the i.i.d. models, we observe that the losses are reduced when the investor is more risk averse ($\gamma = 10$).

3.4 Predictable returns

In this section, we extend our analysis to the case of predictable risky asset returns using an exogenous predictor such as the dividend yield or the earnings-to-price ratio. The dynamic portfolio choice problem when returns are predictable has been the subject of a number of recent papers, including Brennan, Schwartz, and Lagnado (1997), Balduzzi and Lynch (1999), Campbell and Viceira (1999), Barberis (2000), Lynch and Balduzzi (2000), Xia (2001), and Brandt et al. (2005). With the exceptions of Xia (2001) and Brandt et al. (2005), the literature has not addressed the role of parameter uncertainty and learning in the context of dynamic portfolio choice with predictable returns. Xia (2001) develops a framework in which the investor is uncertain and learns about the slope in the predictive regression while all the other parameters are assumed to be known. Nevertheless, as pointed out by Brandt et al. (2005), learning about an individual parameter is of limited scope and does not address the importance of learning in a more realistic situation in which all parameters are treated as unknown. Indeed, our empirical analysis suggests that learning about the unconditional risk premium is a very important part of the learning process.

As in Brandt et al. (2005), our analysis incorporates simultaneous learning about all parameters in the restricted VAR system. However, the two papers differ in a number of aspects. We build on the intuition developed in the previous cases of i.i.d. and autoregressive returns to determine the appropriate state variables for the dynamic portfolio choice problem. We identify 7 state variables that characterize the posterior distribution of the parameters which, combined with the current value of the predictor, result in 8 state variables. In contrast, Brandt et al. (2005) base their analysis on 11 state variables which can be expressed as functions of the state variables that we use. We believe that using a parsimonious state variable vector will help us obtain a more accurate approximation to the value function. This point seems important in the context of the numerical solution used in Brandt et al. (2005), which utilizes a Taylor's approximation to the value function. Function approximation based on Taylor's expansion typically becomes less accurate in higher

dimensions. Secondly, the two papers differ in the numerical solution technique used to solve the Bellman equations. Our solution proceeds through backward induction in the standard fashion but uses a powerful approximation to the value function based on feedforward neural networks. Bertsekas and Tsitsiklis (1996) provide a comprehensive overview of the applications of neural networks in the context of dynamic programming. Their applications, however, do not include portfolio choice problems. To the best of our knowledge, this paper is the first to incorporate the neural network technology in a portfolio choice application. One of the advantages of our approach is that it provides an easy way to monitor the accuracy of the solution. Brandt et al. (2005) introduce a new method for solving dynamic portfolio choice problems, which uses a Taylor’s approximation to the value function and simulation to evaluate the conditional expectations involved in the Bellman equation. They demonstrate the effectiveness of their approach in the simple example in which returns are assumed predictable but all parameters are treated as known. This problem is computationally simple since it only involves a single state variable. However, it is not clear how accurate their simulated-based approach is for higher dimensional problems. Finally, while Brandt et al. (2005) use an improper diffuse prior for the parameters, we use a proper prior to deal with the fact that some predictors are highly persistent. In our empirical application, we use the term spread and the earnings-to-price ratio as predictors and empirical Bayes proper priors.

3.4.1 Bayesian updating framework

We assume that continuously compounded excess returns r_t are predictable by a variable denoted by x . Following Barberis (2000) and Brandt et al. (2005), we postulate that the joint return-predictor dynamics are described by the restricted VAR model

$$y_\tau = \Theta' z_{\tau-1} + \varepsilon_\tau, \quad \tau = 1, \dots, t \tag{30}$$

where $y_\tau = [r_\tau \quad x_\tau]'$, $z_\tau = [1 \quad x_\tau]'$, $x_{\tau-1}$ is the predictor realization at the end of period $\tau - 1$, and $\varepsilon_t \sim N(0_2, \Sigma)$ is a 2×1 vector of regression disturbances. The system uses the parameters Θ , a 2×2 matrix of intercepts and slopes, and Σ , a 2×2 symmetric and positive definite covariance matrix. We can rewrite the model in a more convenient form as

$$Y_t = Z_t \Theta + E_t$$

where

$$Y_t = [R_t \quad X_t], \quad R_t = [r_1 \quad \dots \quad r_t]', \quad X_t = [x_1 \quad \dots \quad x_t]',$$

$$Z_t = [1_t \quad X_t^l], \quad X_t^l = [x_0 \quad \cdots \quad x_{t-1}]', \quad E_t = [\varepsilon_1 \quad \cdots \quad \varepsilon_t]',$$

with the superscript l standing for lagged and 1_t denoting a $t \times 1$ vector of ones. The distribution of the data D_t given the parameters Θ and Σ is matrixvariate normal. Specifically, the likelihood of the data D_t is given by

$$p(D_t|\Theta, \Sigma) = \frac{1}{C_{MN}(I_t, \Sigma; t, 2)} \exp\left(-\frac{1}{2}\text{tr}[(Y_t - Z_t\Theta)'(Y_t - Z_t\Theta)\Sigma^{-1}]\right)$$

where the normalizing constant $C_{MN}(I_t, \Sigma; t, 2)$ is defined by equation (58) in Appendix B.

To obtain predictive distributions and posterior densities, one has to specify the prior of the parameter $\xi = (\Theta, \Sigma)$. For tractability reasons, we use a conjugate matrixvariate normal inverse Wishart prior. We assume that the prior of Σ is an inverse Wishart distribution $IW_2(\Lambda_0, \nu_0)$ with hyperparameters $\nu_0 > 1$ and Λ_0 , a symmetric and positive definite 2×2 matrix. The prior density of Σ is given by

$$p(\Sigma) = \frac{1}{C_{IW}(\Lambda_0, \nu_0; 2)} |\Sigma|^{-\frac{\nu_0+3}{2}} \exp\left(-\frac{1}{2}\text{tr}[\Lambda_0\Sigma^{-1}]\right)$$

where the normalizing constant $C_{IW}(\Lambda_0, \nu_0; 2)$ is defined by equation (59) in Appendix B. Then, given Σ , the prior on Θ is a matrixvariate normal distribution $MN_{2,2}(\Theta_0, \Sigma \otimes \Pi_0^{-1})$ with density

$$p(\Theta|\Sigma) = \frac{1}{C_{MN}(\Pi_0^{-1}, \Sigma; 2, 2)} \exp\left(-\frac{1}{2}\text{tr}[(\Theta - \Theta_0)' \Pi_0 (\Theta - \Theta_0) \Sigma^{-1}]\right)$$

with hyperparameters $\Theta_0 \in \mathbb{R}^{2 \times 2}$ and Π_0 , a symmetric and positive definite 2×2 matrix.

One can specify the prior parameters to guarantee a non-informative prior. For instance, one possibility is to specify $\nu_0 = 2$, $\Lambda_0 = \varepsilon I_2$, $\Theta_0 = 0_{2 \times 2}$, and $\Pi_0 = \varepsilon I_2$, where ε is a small positive number, I_2 is the 2×2 identity matrix, and $0_{2 \times 2}$ is the 2×2 zero matrix. Such a prior will have a minimal effect on the posterior distribution of (Θ, Σ) . Another possibility is to specify an informative prior based on some economic reasoning. One prominent example of such a prior is the empirical Bayes prior in Kandel and Stambaugh (1996), also employed by Avramov (2002), which is based on the assumption of no predictability.

Regardless of how the prior hyperparameters are specified, the posterior of $\xi = (\Theta, \Sigma)$, that is $p(\xi|D_t) = p(\Sigma|D_t)p(\Theta|\Sigma, D_t)$, is also of the matrixvariate normal inverse Wishart form as shown in part B9 of Appendix B. Specifically, the posterior of Σ , $p(\Sigma|D_t)$, is an inverse Wishart $IW_2(\tilde{\Lambda}_t, \tilde{\nu}_t)$ where

$$\tilde{\Lambda}_t = \Lambda_0 + t\hat{\Sigma}_t + (\hat{\Theta}_t - \Theta_0)' [\Pi_0^{-1} + (Z_t'Z_t)^{-1}]^{-1} (\hat{\Theta}_t - \Theta_0) \quad (31)$$

$$\tilde{\nu}_t = \nu_0 + t, \quad (32)$$

with

$$\hat{\Theta}_t = (Z_t' Z_t)^{-1} Z_t' Y_t \quad (33)$$

and

$$\hat{\Sigma}_t = \frac{1}{t} (Y_t - Z_t \hat{\Theta}_t)' (Y_t - Z_t \hat{\Theta}_t) = \frac{1}{t} [Y_t' Y_t - (Z_t' Y_t)' (Z_t' Z_t)^{-1} (Z_t' Y_t)]. \quad (34)$$

In addition, the posterior of Θ given D_t and Σ , is a matrixvariate normal $MN_{2,2}(\tilde{\Theta}_t, \Sigma \otimes \tilde{\Pi}_t^{-1})$ where

$$\tilde{\Pi}_t = \Pi_0 + Z_t' Z_t \quad (35)$$

and

$$\tilde{\Theta}_t = \tilde{\Pi}_t^{-1} [\Pi_0 \Theta_0 + (Z_t' Z_t) \hat{\Theta}_t]. \quad (36)$$

It follows from equations (31), (32), (35), and (36) that the posterior distribution of $\xi = (\Theta, \Sigma)$ is determined by the quantities $Z_t' Z_t$, $\hat{\Theta}_t$, and $\hat{\Sigma}_t$. Combining this observation with equations (33) and (34) yields that the Bayesian updating process is completely characterized by the matrices $Z_t' Y_t$, $Y_t' Y_t$, and $Z_t' Z_t$. Given the data generating process (30), the predictive distribution of returns given data up to time t is determined by the value of the predictor x_t and the posterior parameter distribution. This implies that the state variable for the investor's dynamic portfolio allocation problem consists of x_t plus additional variables that determine the matrices $Z_t' Y_t$, $Y_t' Y_t$, and $Z_t' Z_t$. To obtain a parsimonious representation of the state variable we express these matrices in the following convenient form:

$$\begin{aligned} Z_t' Y_t &= \begin{bmatrix} \sum_{\tau=1}^t r_\tau & \sum_{\tau=1}^t x_\tau \\ \sum_{\tau=1}^t x_{\tau-1} r_\tau & \sum_{\tau=1}^t x_{\tau-1} x_\tau \end{bmatrix} = \begin{bmatrix} tm_1(t) & tm_3(t) - x_0 + m_8(t) \\ tm_7(t) & tm_5(t) \end{bmatrix}, \\ Y_t' Y_t &= \begin{bmatrix} \sum_{\tau=1}^t r_\tau^2 & \sum_{\tau=1}^t r_\tau x_\tau \\ \sum_{\tau=1}^t r_\tau x_\tau & \sum_{\tau=1}^t x_\tau^2 \end{bmatrix} = \begin{bmatrix} tm_2(t) & tm_6(t) \\ tm_6(t) & tm_4(t) - x_0^2 + m_8(t)^2 \end{bmatrix}, \\ Z_t' Z_t &= \begin{bmatrix} t & \sum_{\tau=0}^{t-1} x_\tau \\ \sum_{\tau=0}^{t-1} x_\tau & \sum_{\tau=0}^{t-1} x_\tau^2 \end{bmatrix} = \begin{bmatrix} t & tm_3(t) \\ tm_3(t) & tm_4(t) \end{bmatrix}, \end{aligned}$$

where

$$\begin{aligned} m_1(t) &= \bar{r}_t, \quad m_2(t) = \frac{1}{t} \sum_{\tau=1}^t r_\tau^2, \quad m_3(t) = \frac{1}{t} \sum_{\tau=0}^{t-1} x_\tau = \bar{x}_t, \quad m_4(t) = \frac{1}{t} \sum_{\tau=0}^{t-1} x_\tau^2, \\ m_5(t) &= \frac{1}{t} \sum_{\tau=1}^t x_{\tau-1} x_\tau, \quad m_6(t) = \frac{1}{t} \sum_{\tau=1}^t r_\tau x_\tau, \quad m_7(t) = \frac{1}{t} \sum_{\tau=1}^t x_{\tau-1} r_\tau, \quad m_8(t) = x_t. \end{aligned}$$

The preceding analysis shows that the variables $m_k(t)$, $k = 1, \dots, 8$ characterize the predictive distribution of returns and thus can serve as state variables for the investor's portfolio choice problem. Our next step is to describe the law of motion of the state variable $m(t) = (m_1(t), \dots, m_8(t))$. Given the value of the state variable m at time t_j and the observed data (r_τ, x_τ) , $\tau = t_j + 1, \dots, t_{j+1}$, the state variable m at time t_{j+1} is updated according to the following equations

$$\begin{aligned} m_k(t_{j+1}) &= \frac{t_j}{t_{j+1}} m_k(t_j) + \frac{1}{t_{j+1}} R_{k,j+1}, \quad k = 1, 2, \\ m_3(t_{j+1}) &= \frac{t_j}{t_{j+1}} m_3(t_j) + \frac{1}{t_{j+1}} [m_8(t_j) + Q_{1,j+1}], \\ m_4(t_{j+1}) &= \frac{t_j}{t_{j+1}} m_4(t_j) + \frac{1}{t_{j+1}} [m_8(t_j)^2 + Q_{2,j+1}], \\ m_5(t_{j+1}) &= \frac{t_j}{t_{j+1}} m_5(t_j) + \frac{1}{t_{j+1}} [m_8(t_j) x_{t_{j+1}} + F_{j+1}], \\ m_6(t_{j+1}) &= \frac{t_j}{t_{j+1}} m_6(t_j) + \frac{1}{t_{j+1}} G_{j+1}, \\ m_7(t_{j+1}) &= \frac{t_j}{t_{j+1}} m_7(t_j) + \frac{1}{t_{j+1}} [m_8(t_j) r_{t_{j+1}} + H_{j+1}], \\ m_8(t_{j+1}) &= x_{t_{j+1}}, \end{aligned}$$

where

$$R_{k,j+1} = r_{t_{j+1}}^k + \dots + r_{t_{j+1}}^k, \quad k = 1, 2, \tag{37}$$

$$Q_{k,j+1} = x_{t_{j+1}}^k + \dots + x_{t_{j+1}-1}^k, \quad k = 1, 2, \tag{38}$$

$$F_{j+1} = x_{t_{j+1}} x_{t_{j+2}} + \dots + x_{t_{j+1}-1} x_{t_{j+1}}, \tag{39}$$

$$G_{j+1} = r_{t_{j+1}} x_{t_{j+1}} + \dots + r_{t_{j+1}} x_{t_{j+1}}, \tag{40}$$

$$H_{j+1} = x_{t_{j+1}} r_{t_{j+2}} + \dots + x_{t_{j+1}-1} r_{t_{j+1}}. \tag{41}$$

As in the case of autoregressive returns (subsection 3.3.1), it is useful to work with a transformation of the state variable $m(t)$. In the present context, we define

$$s_1(t) = \bar{r}_t = m_1(t), \tag{42}$$

$$s_2(t) = \frac{1}{t} \sum_{\tau=1}^t (r_\tau - \bar{r}_t)^2 = m_2(t) - m_1(t)^2, \tag{43}$$

$$s_3(t) = \bar{x}_t = m_3(t), \tag{44}$$

$$s_4(t) = \frac{1}{t} \sum_{\tau=0}^{t-1} (x_\tau - \bar{x}_t)^2 = m_4(t) - m_3(t)^2, \tag{45}$$

$$s_5(t) = \frac{\frac{1}{t} \sum_{\tau=1}^t (x_{\tau-1} - \bar{x}_t) (x_\tau - \bar{x}_t)}{\frac{1}{t} \sum_{\tau=0}^{t-1} (x_\tau - \bar{x}_t)^2} = \frac{m_5(t) - m_3(t) \left[m_3(t) - \frac{x_0 - m_8(t)}{t} \right]}{s_4(t)}, \quad (46)$$

$$s_6(t) = \frac{\frac{1}{t} \sum_{\tau=1}^t (r_\tau - \bar{r}_t) (x_\tau - \bar{x}_t)}{\sqrt{\frac{1}{t} \sum_{\tau=1}^t (r_\tau - \bar{r}_t)^2} \sqrt{\frac{1}{t} \sum_{\tau=0}^{t-1} (x_\tau - \bar{x}_t)^2}} = \frac{m_6(t) - m_1(t) \left[m_3(t) - \frac{x_0 - m_8(t)}{t} \right]}{\sqrt{s_2(t) s_4(t)}}, \quad (47)$$

$$s_7(t) = \frac{\frac{1}{t} \sum_{\tau=1}^t (x_{\tau-1} - \bar{x}_t) (r_\tau - \bar{r}_t)}{\sqrt{\frac{1}{t} \sum_{\tau=0}^{t-1} (x_\tau - \bar{x}_t)^2} \sqrt{\frac{1}{t} \sum_{\tau=1}^t (r_\tau - \bar{r}_t)^2}} = \frac{m_7(t) - m_1(t) m_3(t)}{\sqrt{s_2(t) s_4(t)}}, \quad (48)$$

$$s_8(t) = m_8(t). \quad (49)$$

The two vectors $m(t)$ and $s(t)$ contain the same information. Knowing $s(t)$ we can obtain $m(t)$ by

$$m_1(t) = s_1(t),$$

$$m_2(t) = s_2(t) + s_1(t)^2,$$

$$m_3(t) = s_3(t),$$

$$m_4(t) = s_4(t) + s_3(t)^2,$$

$$m_5(t) = s_4(t) s_5(t) + s_3(t) \left[s_3(t) - \frac{x_0 - s_8(t)}{t} \right],$$

$$m_6(t) = \sqrt{s_2(t) s_4(t)} s_6(t) + s_1(t) \left[s_3(t) - \frac{x_0 - s_8(t)}{t} \right],$$

$$m_7(t) = \sqrt{s_2(t) s_4(t)} s_7(t) + s_1(t) s_3(t),$$

$$m_8(t) = s_8(t).$$

A few lines of algebra show that the law of motion of the state variable $s(t)$ is described by

$$s_1(t_{j+1}) = \frac{t_j}{t_{j+1}} s_1(t_j) + \frac{1}{t_{j+1}} R_{1,j+1}, \quad (50)$$

$$s_2(t_{j+1}) = \frac{t_j}{t_{j+1}} [s_2(t_j) + s_1(t_j)^2] + \frac{1}{t_{j+1}} R_{2,j+1} - s_1(t_{j+1})^2, \quad (51)$$

$$s_3(t_{j+1}) = \frac{t_j}{t_{j+1}} s_3(t_j) + \frac{1}{t_{j+1}} [s_8(t_j) + Q_{1,j+1}], \quad (52)$$

$$s_4(t_{j+1}) = \frac{t_j}{t_{j+1}} [s_4(t_j) + s_3(t_j)^2] + \frac{1}{t_{j+1}} [s_8(t_j)^2 + Q_{2,j+1}] - s_3(t_{j+1})^2, \quad (53)$$

$$s_5(t_{j+1}) = \frac{\frac{t_j}{t_{j+1}} [s_4(t_j) s_5(t_j) + s_3(t_j) [s_3(t_j) - [x_0 - s_8(t_j)]/t_j]]}{s_4(t_{j+1})} \quad (54)$$

$$+ \frac{\frac{1}{t_{j+1}} [s_8(t_j) x_{t_{j+1}} + F_{j+1}] - s_3(t_{j+1}) [s_3(t_{j+1}) - (x_0 - x_{t_{j+1}})/t_{j+1}]}{s_4(t_{j+1})},$$

$$s_6(t_{j+1}) = \frac{\frac{t_j}{t_{j+1}} [\sqrt{s_2(t_j)s_4(t_j)}s_6(t_j) + s_1(t_j) [s_3(t_j) - [x_0 - s_8(t_j)]/t_j]]}{\sqrt{s_2(t_{j+1})s_4(t_{j+1})}} \quad (55)$$

$$+ \frac{\frac{1}{t_{j+1}}G_{j+1} - s_1(t_{j+1}) [s_3(t_{j+1}) - (x_0 - x_{t_{j+1}})/t_{j+1}]}{\sqrt{s_2(t_{j+1})s_4(t_{j+1})}},$$

$$s_7(t_{j+1}) = \frac{\frac{t_j}{t_{j+1}} [\sqrt{s_2(t_j)s_4(t_j)}s_7(t_j) + s_1(t_j)s_3(t_j)]}{\sqrt{s_2(t_{j+1})s_4(t_{j+1})}} \quad (56)$$

$$+ \frac{\frac{1}{t_{j+1}} [s_8(t_j)r_{t_{j+1}} + H_{j+1}] - s_1(t_{j+1})s_3(t_{j+1})}{\sqrt{s_2(t_{j+1})s_4(t_{j+1})}},$$

$$s_8(t_{j+1}) = x_{t_{j+1}}. \quad (57)$$

Following the framework described in section 2, we use Monte Carlo simulation to produce samples from the predictive distribution of $D_{t_j, t_{j+1}} = (r_{t_j+1}, x_{t_j+1}, \dots, r_{t_{j+1}}, x_{t_{j+1}})$ given the state variable $s(t_j)$. Conditioning on the parameters and integrating out, we obtain the expression

$$p(D_{t_j, t_{j+1}} | s(t_j)) = \int p(D_{t_j, t_{j+1}} | \Theta, \Sigma, s(t_j)) p(\Theta, \Sigma | s(t_j)) d\Theta d\Sigma.$$

Given the data generating process (30), it follows that the distribution $p(D_{t_j, t_{j+1}} | \Theta, \Sigma, s(t_j))$ only depends on Θ , Σ , and $s_8(t_j) = x_{t_j}$. The following algorithm shows how to simulate from the predictive distribution using the fact that the posterior distribution $p(\Theta, \Sigma | s(t_j))$ is multivariate normal inverse Wishart:

1a. Generate Σ^* from $\Sigma | s(t_j) \sim IW(\tilde{\Lambda}_{t_j}, \tilde{\nu}_{t_j})$ where $\tilde{\Lambda}_{t_j}$ and $\tilde{\nu}_{t_j}$ are obtained using $s(t_j)$ and expressions (31) and (32), respectively.

1b. Given Σ^* from step 1a, generate Θ^* from $\Theta | (\Sigma^*, s(t_j)) \sim MN_{2 \times 2}(\tilde{\Theta}_{t_j}, \Sigma^* \otimes \tilde{\Pi}_{t_j}^{-1})$ where $\tilde{\Pi}_{t_j}$ and $\tilde{\Theta}_{t_j}$ are obtained using $s(t_j)$ and expressions (35) and (36), respectively.

2. Given (Θ^*, Σ^*) from step 1, generate a sample $(r_{*,1}, x_{*,1}, \dots, r_{*,L}, x_{*,L})$ as follows

$$y_{*,l} = \Theta^{*l} z_{*,l-1} + \varepsilon_l, \quad l = 1, \dots, L$$

where $y_{*,l} = [r_{*,l} \quad x_{*,l}]'$, $z_{*,l-1} = [1 \quad x_{*,l-1}]'$, $x_{*,0} = s_8(t_j) = x_{t_j}$, and ε_l are i.i.d. $N(0_2, \Sigma^*)$.

3. Given the generated sample $(r_{*,1}, x_{*,1}, \dots, r_{*,L}, x_{*,L})$, use equations (37)-(41) to obtain $R_{k,j+1}^*$, $Q_{k,j+1}^*$, $k = 1, 2$, F_{j+1}^* , G_{j+1}^* , H_{j+1}^* and then use $s(t_j)$ and the law of motion (equations (50)-(57)) to obtain a draw $s^*(t_{j+1})$ from the predictive distribution of the state variable in the next period.

Part B6 in Appendix B explains how one can simulate from the inverse Wishart distribution as needed in step 1a above. We can obtain a large sample from the predictive distribution

$p(R_{1,j+1}, s(t_{j+1})|s(t_j))$ by repeating the above steps a large number of times. The sample is used to evaluate the conditional expectations in the approximate Bellman equation, and the numerical solution proceeds by backward induction as described in section 2.

3.4.2 Empirical results

In this subsection we present the empirical implications of parameter uncertainty and learning in the context of predictable risky asset returns. The analysis applies the framework developed in the preceding subsection. The risky asset returns are the returns on the S&P 500 index. We use two commonly used economic variables as predictors: the term spread and the earnings-to-price ratio. The source of the data is the paper by Goyal and Welch (2005).⁶ The continuously compounded excess returns are assumed to be predictable according to the restricted VAR model (30). The investor is uncertain about the parameters Θ and Σ , which consist of seven unique elements. The portfolio choice problem is solved using the eight state variables defined in equations (42)-(49). The initial estimation is based on quarterly data from 1984 to 2003 and the portfolio holdings are updated yearly. Due to the computational intensity of the solution, we restrict the investment horizon to 5 years. Throughout we use $I = 300,000$ Monte Carlo repetitions to evaluate the conditional expectations in the approximate Bellman equation (4). We approximate the value function on the entire state space using a grid of 3,000 evenly distributed points and a feedforward neural network with two hidden layers.

Variables commonly used to predict future returns, such as dividend yield and earnings-to-price ratio, are highly persistent. As a result, using a diffuse conjugate matricvariate normal inverse Wishart prior gives rise to problematic posterior distributions in which the slope in the predictor regression exceeds one with high probability. In such a case, the VAR system exhibits nonstationary explosive behavior. In order to avoid this undesired feature, we resort to an empirical Bayes prior. Specifically, we assume that, at each time t , the prior of (Θ, Σ) is determined by setting $\nu_0 = t$, $\Lambda_0 = (\nu_0 - 3)\hat{\Sigma}_t$, $\Theta_0 = \hat{\Theta}_t$, and $\Pi_0 = Z_t'Z_t$. This choice ensures that the prior of Θ is centered around the estimate $\hat{\Theta}_t$, and the prior of Σ is centered around the estimate $\hat{\Sigma}_t$.

Table 7 presents the allocations to the risky asset according to the optimal policy with learning (OL) and the policy C which assumes that all parameters are known and equal to the estimates obtained using data from 1984 to 2003. The upper (lower) panel includes the results for the case in which returns are assumed to be predictable by the term spread (earnings-to-price ratio). In the

⁶The data were obtained from Prof. Goyal's web site: <http://www.bus.emory.edu/AGoyal/>

case of the term spread, we observe that when the parameters are assumed to be known (policy C), there are no hedging demands, reflecting the fact that there is no strong correlation between return and term spread innovations. However, as one would expect based on our previous results with i.i.d. and autoregressive returns, we observe that learning induces negative hedging demands and the investor allocates less in the risky asset for longer horizons.

The results are quite different when the earnings-to-price ratio is used as the predictor. First, under the assumption of known parameters, we observe that there are positive horizon effects due to the fact that the return and earnings-to-price innovations are highly negatively correlated. Second, we observe that the positive hedging demands are reduced but not eliminated when learning is incorporated in the portfolio choice problem. The results are consistent throughout all four different scenarios covering two levels of risk aversion ($\gamma = 5$ and 10) and two levels of risk free rate (5 and 6 percent in annual terms). Examining the differences between the two panels, we observe that the optimal allocation, under both policies OL and C, is much more sensitive to changes in the current predictor value when the earnings-to-price ratio is used as the predictor. This demonstrates that the predictive relation is stronger when we use the earnings-to-price ratio, rather than the term spread, to predict future returns.

In Figures 7 and Figures 8, we present the optimal allocation to the risky asset as a function of the investment horizon when the term spread and earnings-to-price ratio are used as predictors, respectively. We display the optimal allocation for different levels of the state variables s_1 , s_2 , s_3 , and s_4 in order to illustrate the importance of learning about the unconditional first and second moments of the VAR system. Note that the investor uses these state variables to learn about the unconditional return mean (s_1) and variance (s_2) and the unconditional predictor mean (s_3) and variance (s_4). In Figure 7 we observe that, when the term spread is used as the predictor, the optimal allocation is very sensitive to changes in the state variable s_1 and somewhat sensitive to changes in the state variable s_2 . We conclude that learning about the unconditional risk premium remains of major importance even in the context of predictable returns. In contrast, the optimal allocation is not sensitive to changes in the state variables s_3 and s_4 , and therefore learning about the unconditional first two moments of the predictor is of limited importance.

When the earnings-to-price ratio is used as the predictor, the situation is different. Figure 8 shows that the optimal allocation is very sensitive to changes in the state variable s_1 , as well as the state variable s_3 . In this case, learning about both the return and predictor unconditional means is very important. This is a manifestation of the fact that the predictive relationship is stronger when we use the earnings-to-price ratio rather than the term spread as predictor. Figure 8 also

illustrates that learning does not eliminate the positive horizon effect induced by predictability and the negative contemporaneous correlation between shocks to the return and the earnings-to-price ratio.

4 Conclusion

This paper examines the importance of parameter uncertainty and learning in the context of dynamic portfolio choice. We consider a Bayesian investor who maximizes utility of terminal wealth and recognizes the uncertainty surrounding the model parameters. The investor makes allocation decisions, taking into account that the data observed in subsequent periods will provide additional information about the unknown parameters. A sufficient statistic is identified that summarizes all the information in the data relevant for determining the posterior parameter distribution and therefore the predictive distribution of future returns. This sufficient statistic is part of the state variable vector that determines the investment opportunity set from the investor's perspective. As a result, we have an abundance of state variables that makes the numerical solution of the multi-period portfolio choice problem computationally challenging. We overcome the computational difficulties by using a powerful approximation to the value function based on feedforward neural networks.

To provide a comprehensive description of the effects of learning on dynamic portfolio choices, we examine sequentially four different data generating processes for risky asset returns: i.i.d. returns with known and unknown volatility; autoregressive returns; and returns predictable by an exogenous variable, such as the term spread or the earnings-to-price ratio. We find, in general, that parameter uncertainty and learning induce negative horizon effects: the risk-averse investor allocates less in the risky asset for longer investment horizons as compared to an investor that ignores learning. When returns are assumed to be i.i.d., we find that, not surprisingly, learning about return mean is much more important than learning about volatility.

Overall, learning about unconditional expected returns appears the most important component of the dynamic learning process. When we use the earnings-to-price ratio to predict future returns, we find that learning reduces, but does not necessarily reverse, the positive hedging demands induced by predictability and the negative correlation between return and predictor innovations. Ignoring learning can result in significant losses in certainty equivalent return, but these losses are reduced by a sizable amount when the investor uses more past data in her estimation and/or when her risk aversion increases.

A number of important issues related to learning remain to be analyzed. One can assess the

significance of learning in the presence of realistic transaction costs and background risks, such as uncertain labor income. The asset allocation framework can be adapted to allow investment in multiple risky assets instead of the single index portfolio. For example, an investor could invest in a small number of factor portfolios in the US market or a small number of international index portfolios. Finally, in an important direction, one can discard the assumption that the data generating process is known and incorporate model uncertainty in the allocation process. All of these issues are challenging subjects for further research.

A Value function approximation using neural networks

In this paper, we solve the dynamic programming problem by the standard method of backward recursion. A traditional numerical method for carrying out the backward recursion is to discretize the state space and solve the Bellman equation at each point on the discrete grid. In the next step, the value function at each point is approximated by the value function evaluated at the closest point on the grid. To achieve enough precision, one has to use a large number of grid points. It is well known that the state space discretization approach cannot be easily implemented when the state space is high dimensional. In such a case, the Bellman equation has to be solved at a very large number of points, rendering the technique highly inefficient. This problem is rather severe in our context since we work with up to eight state variables, and the expectations in the Bellman equation have to be evaluated through simulation. A more efficient alternative is to calculate the value function at a number of grid points and then use this information to obtain an approximation to the entire value function. Standard function approximation schemes use classes of basis functions, such as polynomials and splines. Judd (1998) provides an excellent discussion of approximation methods. Miranda and Fackler (2002) offer an overview of the function interpolation problem using Chebychev polynomial and polynomial spline approximations. Most of the commonly used approximation schemes are efficient when the state space is low dimensional. However, for problems with a large number of state variables, the computational requirements in terms of time and memory tend to grow exponentially. This fact has been referred to as the *curse of dimensionality*.

An approximation scheme is based on an approximation architecture, that is, a functional form of a certain type that depends on a small number of parameters that need to be tuned. In selecting an approximation architecture, we are primarily concerned with two issues. First, the approximation architecture has to be flexible and rich enough so that it can provide adequate approximation of the functions of interest. Second, we require efficient algorithms for selecting (or tuning) the

parameters of the approximation architecture. One powerful and efficient approximation scheme is based on neural networks which consist of flexible nonlinear functional forms. Comprehensive reviews of neural networks and their applications can be found in the texts by White(1992), Bishop (1995), Hassoun (1995), Haykin (1998), and Reed and Marks (1999). Bertsekas and Tsitsiklis (1996) provide an excellent treatment of the applications of neural networks in the context of dynamic programming. Although neural networks have several diverse applications, it is their function approximation capability that is relevant for our purposes. Cybenko (1989), Hornik et al. (1989), and Funahashi (1989) have shown that one-hidden-layer feedforward neural networks (FFNNs) can approximate uniformly any continuous multivariate function to any desired degree of accuracy. To be specific, a FFNN with one hidden layer defined on \mathbb{R}^d is a function of the form

$$F(x; \alpha, B, \theta, \gamma) = \sum_{m=1}^M \alpha_m g(\beta'_m x + \theta_m) + \gamma,$$

where $x \in \mathbb{R}^d$, $g(\cdot)$ is a scalar function referred to as the *activation* (or *transfer*) function, $\alpha \in \mathbb{R}^M$, $B = [\beta_1 \ \cdots \ \beta_M]' \in \mathbb{R}^{M \times d}$, $\theta \in \mathbb{R}^M$, and $\gamma \in \mathbb{R}$. M is referred to as the number of nodes in the hidden layer. In general, the activation function $g(\cdot)$ is chosen to be differentiable, monotonically increasing, and bounded. Commonly used choices for g include the *hyperbolic tangent* function

$$g(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}},$$

and the *logistic* (or *sigmoid*) function

$$g(x) = \frac{1}{1 + e^{-x}}.$$

Using the same idea, we can define FFNNs with multiple hidden layers. A two-hidden layer FFNN, with N and M nodes in the first and second layer, respectively, is defined as

$$\begin{aligned} F(x; \alpha, B, \theta, \Delta, \lambda, \gamma) &= \sum_{m=1}^M \alpha_m g(\beta'_m z + \theta_m) + \gamma, \\ z &= [z_1 \ \cdots \ z_N]', \\ z_n &= h(\delta'_n x + \lambda_n), \quad n = 1, \dots, N, \end{aligned}$$

where $g(\cdot)$ and $h(\cdot)$ are activation functions, $\alpha \in \mathbb{R}^M$, $B = [\beta_1 \ \cdots \ \beta_M]' \in \mathbb{R}^{M \times N}$, $\theta \in \mathbb{R}^M$, $\Delta = [\delta_1 \ \cdots \ \delta_N]' \in \mathbb{R}^{N \times d}$, $\lambda \in \mathbb{R}^N$, and $\gamma \in \mathbb{R}$. Although one-hidden layer FFFNs are theoretically justified as universal approximators, the use of two hidden layers may provide additional flexibility and power in practice. In our portfolio choice application, we find that one layer is typically enough for low-dimensional problems (up to two state variables), but two layers are required for adequate approximation for high-dimensional problems.

Suppose we wish to approximate a function f defined on \mathbb{R}^d for which the values $y_j = f(x_j)$ on a number of points $\{x_j\}_{j=1}^J$ are available. Let $F(x; \xi)$ be the approximating architecture of choice. The parameter ξ can be tuned so as to minimize some measure of distance between the function $f(\cdot)$ and its approximant $F(\cdot; \xi)$. Specifically, we can obtain an approximation by finding the parameter ξ that minimizes the criterion (or error function)

$$\sum_{j=1}^J d(f(x_j) - F(x_j; \xi))$$

where $d(\cdot)$ is an appropriate scalar distance function such as $d(r) = r^2$ or $d(r) = |r|$.

The solution of the above minimization problem can be a rather tedious computational exercise. However, a number of very efficient algorithms have been developed for this purpose. In particular, a procedure known as *back-propagation* exploits the neural network structure and allows efficient calculation of the derivatives of the error function with respect to the parameter ξ .⁷ Gradient descent, or other general purpose optimization methods, in conjunction with back-propagation can be used to solve the minimization problem. As pointed out by Judd (1998), from experience we know that the error function may have multiple local minima. However, experience also shows that different local minima provide comparable approximation accuracy. A simple way to deal with this issue is to solve the minimization problem with different starting points and compare the resulting fit. In addition, it is good practice to fit neural networks with a varying number of nodes in the hidden layer(s) and examine the quality of the fit. In our application, we used two-hidden-layer FFNNs with the number of hidden nodes ranging from 3 (for one-dimensional problems) to 25 (for eight-dimensional problems). All computations in the paper were carried out using the MATLAB Neural Network toolbox.

B Definitions, properties, and facts

B1. Inverse Gamma Distribution. The random variable $X \in \mathbb{R}$ follows the inverse gamma distribution, denoted by $IG(\alpha, \beta)$, where $\alpha, \beta > 0$, if its probability density function is expressed as

$$f_{IG}(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \frac{1}{x^{\alpha+1}} \exp\left(-\frac{1}{\beta x}\right), \quad x > 0.$$

As $\alpha \rightarrow 0$ and $\beta \rightarrow \infty$, the IG density approaches the improper prior $p(x) \propto \frac{1}{x}$ commonly used as a variance prior distribution.

⁷See Chapter 5 in Reed and Marks (1999) for a detailed discussion of back-propagation.

Let C_n denote the set of positive definite and symmetric $n \times n$ matrices.

B2. Normal Inverse Gamma Distribution. The random variables $X \in \mathbb{R}^m$ and $v \in \mathbb{R}$ follow a joint normal inverse gamma distribution with parameters $(\theta, \Pi, \alpha, \beta)$, denoted by $NIG(\theta, \Pi, \alpha, \beta)$, where $\theta \in \mathbb{R}^m$, $\Pi \in C_m$, $\alpha > 0$, and $\beta > 0$, if v follows the $IG(\alpha, \beta)$ distribution and X given v follows the multivariate normal distribution $N_p(\theta, v\Pi^{-1})$.

B3. Matricvariate Normal Distribution. The $p \times q$ random matrix X is said to follow the matricvariate normal distribution with parameters $M \in \mathbb{R}^{p \times q}$, $P \in C_p$, and $Q \in C_q$, denoted by $X \sim MN_{p,q}(M, Q \otimes P)$, if and only if $vec(X)$ follows the multivariate normal distribution $N_{pq}(vec(M), Q \otimes P)$ where $vec(\cdot)$ is the column-stacking operator. The probability density function of X is then given by

$$f_{MN}^{p,q}(X|M, Q \otimes P) = \frac{\exp\left\{-\frac{1}{2}\text{tr}\left[Q^{-1}(X-M)'P^{-1}(X-M)\right]\right\}}{C_{MN}(P, Q; p, q)}$$

where the normalizing constant is given by

$$C_{MN}(P, Q; p, q) = \sqrt{(2\pi)^{pq}|P|^q|Q|^p}. \quad (58)$$

B4. Wishart Distribution. Let X_i be i.i.d. $N_p(0_p, \Omega)$, $i = 1, \dots, m$ where $\Omega \in C_p$ and define $S = \sum_{i=1}^m X_i X_i'$. Then S is said to follow the Wishart distribution with scale matrix Ω and m degrees of freedom and is denoted by $S \sim W_p(\Omega, m)$. In general, if $\nu > p - 1$, with ν not necessarily an integer, we define $S \sim W_p(\Omega, \nu)$ if its density is given by

$$f_W^p(S|\Omega, \nu) = \frac{|S|^{(\nu-p-1)/2} \exp\left(-\frac{1}{2}\text{tr}(\Omega^{-1}S)\right)}{C_W(\Omega, \nu; p)}$$

where the normalizing constant is given by

$$C_W(\Omega, \nu; p) = 2^{\nu p/2} \pi^{p(p-1)/4} |\Omega|^{\nu/2} \prod_{i=1}^p \Gamma((\nu + 1 - i)/2).$$

In the one-dimensional case ($p = 1$), the Wishart distribution reduces to (a multiple of) a χ^2 distribution. Specifically, if $s \sim W_1(\omega, \nu)$ then $s \sim \omega G\left(\frac{\nu}{2}, 2\right) = \omega \chi^2(\nu)$.

B5. Inverse Wishart Distribution. If $S \sim W_p(\Omega, \nu)$ where $\nu > p - 1$ and $\Omega \in C_p$, then $U = S^{-1}$ is said to have an inverse Wishart distribution denoted by $IW_p(\Omega^{-1}, \nu)$. Thus, by definition, $U \sim IW_p(\Lambda, \nu)$ if and only if $U^{-1} \sim W_p(\Lambda^{-1}, \nu)$ (see B4 above). The density of $U \sim IW_p(\Lambda, \nu)$ is given by

$$f_{IW}^p(U|\Lambda, \nu) = \frac{|U|^{-(\nu+p+1)/2} \exp\left(-\frac{1}{2}\text{tr}(U^{-1}\Lambda)\right)}{C_{IW}(\Lambda, \nu; p)}$$

where the normalizing constant is given by

$$C_{IW}(\Lambda, \nu; p) = 2^{\nu p/2} \pi^{p(p-1)/4} |\Lambda|^{-\nu/2} \prod_{i=1}^p \Gamma((\nu + 1 - i)/2). \quad (59)$$

B6. Simulating from the Inverse Wishart Distribution. Let $m \geq p$ be an integer and $\Lambda \in C_p$. To simulate U from the inverse Wishart distribution $IW_p(\Lambda, m)$, we need to generate S from the Wishart distribution $W_p(\Lambda^{-1}, m)$ and then take $U = S^{-1}$. Let $\Lambda = \Xi \Xi'$ be the Cholesky decomposition of Λ so that $\Lambda^{-1} = (\Xi^{-1})' \Xi^{-1}$. The Wishart distribution has the following property: if $S \sim W_p(\Omega, m)$ and B is a $p \times q$ matrix, then $B' S B \sim W_q(B' \Omega B, m)$. Therefore, if $S_0 \sim W_p(I_p, m)$, the above property implies $S = (\Xi^{-1})' S_0 \Xi^{-1} \sim W_p(\Lambda^{-1}, m)$. Notice that we can obtain $S_0 \sim W_p(I_p, m)$ by letting $S_0 = \sum_{i=1}^m Z_i Z_i' \sim W_p(I_p, m)$ where Z_i are i.i.d. $N_p(0_p, I_p)$, $i = 1, \dots, m$. Thus,

$$U = S^{-1} = [(\Xi^{-1})' S_0 \Xi^{-1}]^{-1} = \Xi S_0^{-1} \Xi' = \Xi [\sum_{i=1}^m Z_i Z_i']^{-1} \Xi' \sim IW_p(\Lambda, m).$$

B7. Posterior and predictive distributions for i.i.d. returns with known volatility.

Assume that the returns r_t follow

$$r_t = \mu + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } N(0, v)$$

with v assumed to be a known constant and the informative prior of μ is $N(\mu_0, v_0)$. It follows from the analysis in section 4.2 in Gill (2002) that the posterior of μ given return data $D_t = (r_1, \dots, r_t)$ is $N(\tilde{\mu}_t, \tilde{v}_t)$ where

$$\tilde{\mu}_t = \left(\frac{v_t}{v_0 + v_t} \right) \mu_0 + \left(\frac{v_0}{v_0 + v_t} \right) \bar{r}_t, \quad \tilde{v}_t = v_t \left(\frac{v_0}{v_0 + v_t} \right), \quad v_t = \frac{v}{t}.$$

The predictive distribution of $R_{1,j+1} = r_{t_j+1} + \dots + r_{t_{j+1}}$ given return data D_{t_j} is obtained by integrating out μ as follows

$$p(R_{1,j+1} | D_{t_j}) = \int p(R_{1,j+1} | \mu, D_{t_j}) p(\mu | D_{t_j}) d\mu.$$

Combining terms we have

$$p(R_{1,j+1} | \mu, D_{t_j}) p(\mu | D_{t_j}) = \frac{1}{2\pi \sqrt{Lv \tilde{v}_{t_j}}} \exp \left(-\frac{(R_{1,j+1} - L\mu)^2}{2Lv} - \frac{(\mu - \tilde{\mu}_{t_j})^2}{2\tilde{v}_{t_j}} \right).$$

The expression in the exponent equals

$$\begin{aligned}
& - \frac{\tilde{v}_{t_j} \left(R_{1,j+1}^2 + L^2 \mu^2 - 2R_{1,j+1} L \mu \right) + Lv \left(\mu^2 + \tilde{\mu}_{t_j}^2 - 2\tilde{\mu}_{t_j} \mu \right)}{2Lv\tilde{v}_{t_j}} \\
= & - \frac{(v + L\tilde{v}_{t_j}) L \mu^2 - 2L \left(\tilde{v}_{t_j} R_{1,j+1} + v\tilde{\mu}_{t_j} \right) \mu + \tilde{v}_{t_j} R_{1,j+1}^2 + Lv\tilde{\mu}_{t_j}^2}{2Lv\tilde{v}_{t_j}} \\
= & - \frac{\mu^2 - 2 \left(\frac{\tilde{v}_{t_j} R_{1,j+1} + v\tilde{\mu}_{t_j}}{v + L\tilde{v}_{t_j}} \right) \mu + \frac{\tilde{v}_{t_j} R_{1,j+1}^2 + Lv\tilde{\mu}_{t_j}^2}{(v + L\tilde{v}_{t_j})L}}{2 \frac{v\tilde{v}_{t_j}}{v + L\tilde{v}_{t_j}}} \\
= & - \frac{\left(\mu - \frac{\tilde{v}_{t_j} R_{1,j+1} + v\tilde{\mu}_{t_j}}{v + L\tilde{v}_{t_j}} \right)^2}{2 \frac{v\tilde{v}_{t_j}}{v + L\tilde{v}_{t_j}}} - \frac{\frac{\tilde{v}_{t_j} R_{1,j+1}^2 + Lv\tilde{\mu}_{t_j}^2}{(v + L\tilde{v}_{t_j})L} - \left(\frac{\tilde{v}_{t_j} R_{1,j+1} + v\tilde{\mu}_{t_j}}{v + L\tilde{v}_{t_j}} \right)^2}{2 \frac{v\tilde{v}_{t_j}}{v + L\tilde{v}_{t_j}}}.
\end{aligned}$$

Finally, calculating the integral over μ and simplifying yields

$$p(R_{1,j+1} | D_{t_j}) = \frac{1}{\sqrt{2\pi L (v + L\tilde{v}_{t_j})}} \exp \left(- \frac{\left(R_{1,j+1} - L\tilde{\mu}_{t_j} \right)^2}{2L (v + L\tilde{v}_{t_j})} \right)$$

and thus the predictive distribution of $R_{1,j+1}$ given D_{t_j} is normal with mean $L\tilde{\mu}_{t_j}$ and variance $L(v + L\tilde{v}_{t_j})$.

B8. Posterior distribution for the linear model with NIG prior. Consider the linear model

$$y_t = x_t' \theta + \varepsilon_t, \quad t = 1, \dots, T$$

where $y_t \in \mathbb{R}$, $x_t \in \mathbb{R}^m$, $\theta \in \mathbb{R}^m$, and ε_t are i.i.d. $N(0, v)$. Assume that the prior of (θ, v) is $NIG(\theta_0, \Pi_0, \alpha_0, \beta_0)$. Let

$$\begin{aligned}
Y &= [y_1 \quad \cdots \quad y_T]', \quad (T \times 1) \\
X &= [x_1 \quad \cdots \quad x_T]', \quad (T \times m) \\
E &= [\varepsilon_1 \quad \cdots \quad \varepsilon_T]', \quad (T \times 1)
\end{aligned}$$

so that the model is concisely written as

$$Y = X\theta + E.$$

Then, it follows from Theorem 2.24 in Bauwens et al. (1999) that the posterior of (θ, v) given the

data $D = (Y, X)$ is $NIG(\theta_*, \Pi_*, \alpha_*, \beta_*)$ where

$$\begin{aligned}\Pi_* &= \Pi_0 + X'X \\ \theta_* &= \Pi_*^{-1} \left[\Pi_0 \theta_0 + (X'X) \hat{\theta} \right] \\ \alpha_* &= \alpha_0 + \frac{T}{2} \\ \beta_*^{-1} &= \beta_0^{-1} + \frac{T \hat{v}}{2} + \frac{(\hat{\theta} - \theta_0)' [\Pi_0^{-1} + (X'X)^{-1}]^{-1} (\hat{\theta} - \theta_0)}{2}\end{aligned}$$

with

$$\begin{aligned}\hat{\theta} &= (X'X)^{-1} X'Y, \\ \hat{v} &= \frac{1}{T} (Y - X\hat{\theta})'(Y - X\hat{\theta}).\end{aligned}$$

B9. Multivariate Linear Regression: Bayesian Analysis with Proper Priors. Consider the linear regression model

$$y_t = \Theta' w_t + \varepsilon_t, \quad t = 1, \dots, T$$

where $y_t \in \mathbb{R}^p$, $w_t \in \mathbb{R}^q$, $\Theta \in \mathbb{R}^{q \times p}$, ε_t are i.i.d. $N_p(0_p, \Sigma)$, and $\Sigma \in \mathbb{R}^{p \times p}$ is symmetric and positive definite. Denoting

$$Y_T = [y_1 \quad \cdots \quad y_T]',$$

$$W_T = [w_1 \quad \cdots \quad w_T]',$$

$$E_T = [\varepsilon_1 \quad \cdots \quad \varepsilon_T]'$$

we can write the model concretely as

$$Y_T = W_T \Theta + E_T.$$

The matrix of disturbances E_T follows a matrixvariate normal distribution $MN_{T,p}(0_{T \times p}, \Sigma \otimes I_T)$ (see B3 in Appendix B for the definition) and therefore the likelihood of the data $D_T = (Y_T, W_T)$ is given

$$p(D_T | \Theta, \Sigma) = \frac{\exp \left\{ -\frac{1}{2} \text{tr} \left[(Y_T - W_T \Theta) \Sigma^{-1} (Y_T - W_T \Theta)' \right] \right\}}{C_{MN}(I_T, \Sigma; T, p)}.$$

The prior of the covariance matrix Σ is specified as inverse Wishart $IW_p(\Lambda_0, \nu_0)$ with density

$$p(\Sigma) = \frac{|\Sigma|^{-(\nu_0 + p + 1)/2} \exp \left\{ -\frac{1}{2} \text{tr}(\Sigma^{-1} \Lambda_0) \right\}}{C_{IW}(\Lambda_0, \nu_0; p)}.$$

Given Σ , the prior of Θ is specified as matricvariate normal $MN_{q,p}(\Theta_0, \Sigma \otimes P_0^{-1})$ with density

$$p(\Theta|\Sigma) = \frac{\exp\left\{-\frac{1}{2}\text{tr}\left[(\Theta - \Theta_0)'P_0(\Theta - \Theta_0)\Sigma^{-1}\right]\right\}}{C_{MN}(P_0^{-1}, \Sigma; q, p)}.$$

By Bayes rule, the posterior density of the parameter $\xi = (\Theta, \Sigma)$ satisfies

$$\begin{aligned} p(\xi|D_T) &\propto p(D_T|\xi)p(\xi) = p(D_T|\Theta, \Sigma)p(\Theta|\Sigma)p(\Sigma) \\ &\propto |\Sigma|^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\text{tr}\left[(Y_T - W_T\Theta)'(Y_T - W_T\Theta)\Sigma^{-1}\right]\right\} \\ &\times |\Sigma|^{-\frac{\nu_0+p+1}{2}} \exp\left\{-\frac{1}{2}\text{tr}(\Lambda_0\Sigma^{-1})\right\} \\ &\times |\Sigma|^{-\frac{q}{2}} \exp\left\{-\frac{1}{2}\text{tr}\left[(\Theta - \Theta_0)'P_0(\Theta - \Theta_0)\Sigma^{-1}\right]\right\} = \\ &= |\Sigma|^{-\frac{(\nu_0+T)+p+1}{2}} |\Sigma|^{-\frac{q}{2}} \exp\left\{-\frac{1}{2}\text{tr}\left[G\Sigma^{-1}\right]\right\} \end{aligned}$$

where the matrix G is given by

$$G = \Lambda_0 + (Y_T - W_T\Theta)'(Y_T - W_T\Theta) + (\Theta - \Theta_0)'P_0(\Theta - \Theta_0).$$

Let

$$\hat{\Theta}_T = (W_T'W_T)^{-1}W_T'Y_T$$

and note that $W_T'(Y_T - W_T\hat{\Theta}_T) = 0_{q \times p}$ and so

$$\begin{aligned} &(Y_T - W_T\Theta)'(Y_T - W_T\Theta) \\ &= (Y_T - W_T\hat{\Theta}_T)'(Y_T - W_T\hat{\Theta}_T) + (\Theta - \hat{\Theta}_T)'W_T'W_T(\Theta - \hat{\Theta}_T) \end{aligned}$$

Define

$$\tilde{P}_T = P_0 + W_T'W_T$$

and

$$\tilde{\Theta}_T = \tilde{P}_T^{-1} \left[P_0\Theta_0 + (W_T'W_T)\hat{\Theta}_T \right]$$

and observe that

$$\begin{aligned}
& (\Theta - \hat{\Theta}_T)' W_T' W_T (\Theta - \hat{\Theta}_T) + (\Theta - \Theta_0)' P_0 (\Theta - \Theta_0) \\
= & \Theta' (W_T' W_T) \Theta - \hat{\Theta}_T' (W_T' W_T) \Theta - \Theta' (W_T' W_T) \hat{\Theta}_T + \hat{\Theta}_T' (W_T' W_T) \hat{\Theta}_T \\
& + \Theta' P_0 \Theta - \Theta_0' P_0 \Theta - \Theta' P_0 \Theta_0 + \Theta_0' P_0 \Theta_0 \\
= & \Theta' \tilde{P}_T \Theta - \left[(W_T' W_T) \hat{\Theta}_T + P_0 \Theta_0 \right]' \Theta - \Theta' \left[(W_T' W_T) \hat{\Theta}_T + P_0 \Theta_0 \right] \\
& + \hat{\Theta}_T' (W_T' W_T) \hat{\Theta}_T + \Theta_0' P_0 \Theta_0 \\
= & \Theta' \tilde{P}_T \Theta - (\tilde{P}_T \tilde{\Theta}_T)' \Theta - \Theta' (\tilde{P}_T \tilde{\Theta}_T) + \tilde{\Theta}_T' \tilde{P}_T \tilde{\Theta}_T \\
& - \tilde{\Theta}_T' \tilde{P}_T \tilde{\Theta}_T + \hat{\Theta}_T' (W_T' W_T) \hat{\Theta}_T + \Theta_0' P_0 \Theta_0 \\
= & (\Theta - \tilde{\Theta}_T)' \tilde{P}_T (\Theta - \tilde{\Theta}_T) + H
\end{aligned}$$

where

$$\begin{aligned}
H &= - \left[P_0 \Theta_0 + (W_T' W_T) \hat{\Theta}_T \right]' \tilde{P}_T^{-1} \left[P_0 \Theta_0 + (W_T' W_T) \hat{\Theta}_T \right] \\
&+ \hat{\Theta}_T' (W_T' W_T) \hat{\Theta}_T + \Theta_0' P_0 \Theta_0 \\
&= \Theta_0' \left(P_0 - P_0 \tilde{P}_T^{-1} P_0 \right) \Theta_0 - \hat{\Theta}_T' (W_T' W_T) \tilde{P}_T^{-1} P_0 \Theta_0 \\
&- \Theta_0' P_0 \tilde{P}_T^{-1} (W_T' W_T) \hat{\Theta}_T \\
&+ \hat{\Theta}_T' \left[(W_T' W_T) - (W_T' W_T) \tilde{P}_T^{-1} (W_T' W_T) \right] \hat{\Theta}_T.
\end{aligned}$$

Using the property $(A + B)^{-1} = A^{-1} - A^{-1}(B^{-1} + A^{-1})^{-1}A^{-1}$ we conclude

$$\begin{aligned}
P_0 - P_0 \tilde{P}_T^{-1} P_0 &= [P_0^{-1} + (W_T' W_T)^{-1}]^{-1}, \\
(W_T' W_T) - (W_T' W_T) \tilde{P}_T^{-1} (W_T' W_T) &= [P_0^{-1} + (W_T' W_T)^{-1}]^{-1}.
\end{aligned}$$

Moreover,

$$\begin{aligned}
(W_T' W_T) \tilde{P}_T^{-1} P_0 &= [P_0 + (W_T' W_T)] \tilde{P}_T^{-1} P_0 - P_0 \tilde{P}_T^{-1} P_0 \\
&= P_0 - P_0 \tilde{P}_T^{-1} P_0 = [P_0^{-1} + (W_T' W_T)^{-1}]^{-1}.
\end{aligned}$$

Collecting terms yields

$$H = (\hat{\Theta}_T - \Theta_0)' [P_0^{-1} + (W_T' W_T)^{-1}]^{-1} (\hat{\Theta}_T - \Theta_0).$$

Therefore,

$$\begin{aligned}
G &= \Lambda_0 + (Y_T - W_T \hat{\Theta}_T)' (Y_T - W_T \hat{\Theta}_T) \\
&+ (\Theta - \tilde{\Theta}_T)' \tilde{P}_T (\Theta - \tilde{\Theta}_T) \\
&+ (\hat{\Theta}_T - \Theta_0)' [P_0^{-1} + (W_T' W_T)^{-1}]^{-1} (\hat{\Theta}_T - \Theta_0).
\end{aligned}$$

and so

$$\begin{aligned}
p(\xi|D_T) &\propto |\Sigma|^{-\frac{\tilde{\nu}_T+p+1}{2}} \exp\left\{-\frac{1}{2}\text{tr}[\tilde{\Lambda}_T\Sigma^{-1}]\right\} \\
&\quad \times |\Sigma|^{-\frac{q}{2}} \exp\left\{-\frac{1}{2}\text{tr}[(\Theta - \tilde{\Theta}_T)' \tilde{P}_T(\Theta - \tilde{\Theta}_T)\Sigma^{-1}]\right\}
\end{aligned}$$

where

$$\tilde{\nu}_T = \nu_0 + T,$$

$$\tilde{\Lambda}_T = \Lambda_0 + T\hat{\Sigma}_T + (\hat{\Theta}_T - \Theta_0)' [P_0^{-1} + (W_T'W_T)^{-1}]^{-1} (\hat{\Theta}_T - \Theta_0),$$

and

$$\hat{\Sigma}_T = \frac{1}{T}(Y_T - W_T\hat{\Theta}_T)'(Y_T - W_T\hat{\Theta}_T).$$

Thus, the posterior of Σ is $IW_p(\tilde{\Lambda}_T, \tilde{\nu}_T)$ and the posterior of Θ given Σ is $MN_{q,p}(\tilde{\Theta}_T, \Sigma \otimes \tilde{P}_T^{-1})$.

References

- [1] Avramov, Doron, 2002, Stock return predictability and model uncertainty, *Journal of Financial Economics* 64, 423-458.
- [2] Balduzzi, Pierluigi, and Anthony Lynch, 1999, Transaction costs and predictability: some utility costs calculations, *Journal of Financial Economics* 52, 47-78.
- [3] Barberis, Nicholas, 2000, Investing in the long run when returns are predictable, *Journal of Finance* 55, 225-264.
- [4] Bawa, Vijay, Stephen Brown, and Roger Klein, 1979, *Estimation Risk and Optimal Portfolio Choice*, North Holland, Amsterdam.
- [5] Bauwens, Luc, Michel Lubrano, and Jean-Francois Richard, 1999, *Bayesian Inference in Dynamic Econometric Models*, Oxford University Press, New York, NY.
- [6] Bertsekas, Dimitri, and John Tsitsiklis, 1996, *Neuro-Dynamic Programming*, Athena Scientific, Nashua, NH.
- [7] Bishop, Christopher, 1995, *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford.
- [8] Boudry, Walter, and Philip Gray, 2003, Assessing the economic significance of return predictability: A research note, *Journal of Business Finance and Accounting* 30, 1305-1326.
- [9] Brandt, Michael, 2004, Portfolio Choice Problems, in Yasin Ait-Sahalia and Lars Hansen (Eds.), *Handbook of Financial Econometrics*, forthcoming.
- [10] Brandt, Michael, Amit Goyal, Pedro Santa-Clara, and Jonathan Stroud, 2005, A simulation approach to dynamic portfolio choice with an application to learning about return predictability, *Review of Financial Studies* 18, 831-873.
- [11] Brennan Michael, 1998, The role of learning in dynamic portfolio decisions, *European Finance Review* 1, 295-306.
- [12] Brennan Michael, Eduardo Schwartz, and Ronald Lagnado, 1997, Strategic asset allocation, *Journal of Economic Dynamics and Control* 21, 1377-1403.
- [13] Brown, Stephen, 1978, The portfolio choice problem: Comparison of certainty equivalent and optimal Bayes portfolios, *Communications in Statistics: Simulation and Computation* B7, 321-334.

- [14] Campbell, John, and Luis Viceira, 1999, Consumption and portfolio decisions when expected returns are time varying, *Quarterly Journal of Economics* 114, 433-495.
- [15] Campbell, John, and Luis Viceira, 2002, *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, Oxford University Press, New York, NY.
- [16] Cybenko, George, 1989, Approximation by superpositions of a sigmoidal function, *Mathematical Control, Signals, and Systems* 2, 303-314.
- [17] Detemple, Jerome, Asset pricing in a production economy with incomplete information, *Journal of Finance* 41, 383-391.
- [18] Dothan, Michael, and David Feldman, 1986, Equilibrium interest rates and multiperiod bonds in a partially observable economy, *Journal of Finance* 41, 369-382.
- [19] Epstein, Larry, and Stanley Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937-969.
- [20] Funahashi, Ken-ichi, 1989, On the approximate realization of continuous mappings by neural networks, *Neural Networks* 2, 183-192.
- [21] Gennotte, Gerard, 1986, Optimal portfolio choice under incomplete information, *Journal of Finance* 41, 733-746.
- [22] Gill, Jeff, 2002, *Bayesian Methods: A Social and Behavioral Sciences Approach*, Chapman and Hall/CRC, Boca Raton, FL.
- [23] Goyal, Amit, and Ivo Welch, 2005, A comprehensive look at the empirical performance of equity premium prediction, working paper, Emory University.
- [24] Hassoun, Mohamad, 1995, *Fundamentals of Artificial Neural Networks*, MIT Press, Cambridge, MA.
- [25] Haykin, Simon, 1998, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, Upper Saddle River, NJ.
- [26] Hornik, Kurt, Maxwell Stinchcombe, and Halbert White, 1989, Multilayer feedforward networks are universal approximators, *Neural Networks* 2, 359-366.
- [27] Judd, Kenneth, 1998, *Numerical Methods in Economics*, MIT Press, Cambridge, MA.

- [28] Kandel, Shmuel, and Robert Stambaugh, 1996, On the predictability of stock returns: An asset-allocation perspective, *Journal of Finance* 51, 385-424.
- [29] Klein, Roger, and Vijay Bawa, 1976, The effect of estimation risk on optimal portfolio choice, *Journal of Financial Economics* 3, 215-231.
- [30] Lynch, Anthony, and Pierluigi Balduzzi, 2000, Predictability and transaction costs: The impact on rebalancing rules and behavior, *Journal of Finance* 55, 2285-2309.
- [31] Miranda, Mario, and Paul Fackler, 2002, *Applied Computational Economics and Finance*, MIT Press, Cambridge, MA.
- [32] Merton, Robert, 1969, Lifetime portfolio choice: The continuous-time case, *Review of Economics and Statistics* 51, 247-257.
- [33] Merton, Robert, 1971, Optimum consumption and portfolio rules in a continuous-time model, *Journal of Economic Theory* 3, 373-413.
- [34] Neiderreiter, Harald, 1988, Low-discrepancy and low-dispersion sequences, *Journal of Number Theory* 30, 51-70.
- [35] Reed, Russell, and Robert Marks, 1999, *Neural Smoothing: Supervised Learning in Feedforward Artificial Neural Networks*, MIT Press, Cambridge, MA.
- [36] Samuelson, Paul, 1969, Lifetime portfolio selection by dynamic stochastic programming, *Review of Economics and Statistics* 51, 239-246.
- [37] Skiadas, Costis, 2005, Dynamic portfolio choice and risk aversion, *Handbook of Financial Engineering*, forthcoming.
- [38] Tu, Jun, and Guofu Zhou, Data-generating process uncertainty: What difference does it make in portfolio decisions? *Journal of Financial Economics* 72, 385-421.
- [39] White, Halbert, 1992, *Artificial Neural Networks: Approximation and Learning Theory*, Blackwell, Oxford.
- [40] Xia, Yihong, 2001, Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation, *Journal of Finance* 56, 205-246.
- [41] Zellner, Arlond, and Karrupan Chetty, 1965, Prediction and decision problems in regression models from the Bayesian point of view, *Journal of the American Statistical Association* 60, 608-616.

Table 1: IID returns with known volatility - Optimal allocation

This table presents the optimal allocation to the risky asset when the continuously compounded excess returns are assumed to be i.i.d. with known volatility. The upper panel corresponds to the case in which the investment period starts in 2000 and the lower panel corresponds to the case in which the investment period starts in 2003. The state variable is set equal to its value at the end of the corresponding sample. For each case, we consider eight different scenarios in which the coefficient of relative risk aversion (γ) takes the values 5 and 10; the annualized risk-free rate (R_f) takes the values 5 and 6 percent; and the sample size (K) takes the values 15 and 30 years. The investor uses quarterly data, updates the portfolio yearly, and has an investment horizon of 10 years. The volatility is assumed to be known and equal to the sample return standard deviation. For reference, we report the corresponding return sample average (state variable s) and return sample standard deviation (σ) at the beginning of the investment period. The optimal policy with learning about the mean return μ is denoted by OL. The policy denoted C treats the mean return μ as constant and equal to the return sample average at the beginning of the investment period.

Investment period starts in 2000

γ	R_f	K	s	σ	OL / Years to final date				C
					10	7	3	1	
5	5	15	2.48	7.47	57.9	67.6	79.8	85.6	99.2
5	5	30	1.88	7.90	48.0	52.1	57.2	59.7	70.4
5	6	15	2.25	7.47	52.8	61.8	72.9	78.4	90.7
5	6	30	1.64	7.90	43.2	46.6	51.1	53.4	62.7
10	5	15	2.48	7.47	27.6	32.8	39.4	42.7	49.6
10	5	30	1.88	7.90	23.3	25.5	28.4	29.7	35.1
10	6	15	2.25	7.47	25.3	30.0	36.2	39.1	45.3
10	6	30	1.64	7.90	20.9	22.9	25.3	26.5	31.2

Investment period starts in 2003

γ	R_f	K	s	σ	OL / Years to final date				C
					10	7	3	1	
5	5	15	1.66	7.88	40.7	47.5	56.3	60.5	63.4
5	5	30	1.65	8.56	43.2	46.6	51.4	53.7	55.2
5	6	15	1.42	7.88	35.9	42.0	49.5	53.2	55.7
5	6	30	1.42	8.56	38.3	41.4	45.4	47.4	48.7
10	5	15	1.66	7.88	19.5	23.0	27.8	30.1	31.6
10	5	30	1.65	8.56	21.0	22.9	25.4	26.7	27.4
10	6	15	1.42	7.88	17.2	20.3	24.4	26.5	27.7
10	6	30	1.42	8.56	18.6	20.3	22.4	23.5	24.2

Table 2: IID returns with known volatility - Certainty equivalent return

This table presents the certainty equivalent return for three portfolio policies when the continuously compounded excess returns are assumed to be i.i.d. with known volatility. The upper panel corresponds to the case in which the investment period starts in 2000 and the lower panel corresponds to the case in which the investment period starts in 2003. The state variable is set equal to its value at the end of the corresponding sample. For each case, we consider eight different scenarios in which the coefficient of relative risk aversion (γ) takes the values 5 and 10; the annualized risk-free rate (R_f) takes the values 5 and 6 percent; and the sample size (K) takes the values 15 and 30 years. The investor uses quarterly data, updates the portfolio yearly, and has an investment horizon of 10 years. The volatility is assumed to be known and equal to the sample return standard deviation. For reference, we report the corresponding return sample average (state variable s) and return sample standard deviation (σ) at the beginning of the investment period. The optimal policy with learning about the mean return μ is denoted by OL. The policy denoted by C treats the mean return μ as constant and equal to the return sample average at the beginning of the investment period. The policy denoted by MU uses all available data to update the estimate of μ but in each period, myopically and ignoring future learning, solves the one-period portfolio allocation problem.

Investment period starts in 2000

γ	R_f	K	s	σ	OL	C	PUNL	MU
5	5	15	2.48	7.47	8.92	7.79	8.11	8.74
5	5	30	1.88	7.90	7.59	7.41	7.46	7.58
5	6	15	2.25	7.47	9.32	8.31	8.59	9.17
5	6	30	1.64	7.90	8.08	7.93	7.97	8.07
10	5	15	2.48	7.47	6.86	6.24	6.40	6.75
10	5	30	1.88	7.90	6.25	6.15	6.17	6.24
10	6	15	2.25	7.47	7.58	7.04	7.18	7.49
10	6	30	1.64	7.90	7.01	6.93	6.94	7.00

Investment period starts in 2003

γ	R_f	K	s	σ	OL	C	PUNL	MU
5	5	15	1.66	7.88	6.84	6.24	6.38	6.75
5	5	30	1.65	8.56	6.91	6.74	6.77	6.87
5	6	15	1.42	7.88	7.47	6.99	7.09	7.39
5	6	30	1.42	8.56	7.50	7.37	7.39	7.47
10	5	15	1.66	7.88	5.88	5.57	5.64	5.83
10	5	30	1.65	8.56	5.92	5.83	5.85	5.90
10	6	15	1.42	7.88	6.70	6.44	6.50	6.66
10	6	30	1.42	8.56	6.73	6.65	6.66	6.71

Table 3: IID returns with unknown volatility - Optimal allocation

This table presents the optimal allocation to the risky asset when the continuously compounded excess returns are assumed to be i.i.d. with unknown mean and volatility. The upper panel corresponds to the case in which the investment period starts in 2000 and the lower panel corresponds to the case in which the investment period starts in 2003. The state variable is set equal to its value at the end of the corresponding sample. For each case, we consider eight different scenarios in which the coefficient of relative risk aversion (γ) takes the values 5 and 10; the annualized risk-free rate (R_f) takes the values 5 and 6 percent; and the sample size (K) takes the values 15 and 30 years. The investor uses quarterly data, updates the portfolio yearly, and has an investment horizon of 10 years. For reference, we report the corresponding return sample average (s_1) and return sample standard deviation ($\sqrt{s_2}$) at the beginning of the investment period. The optimal policy with learning about the return mean μ and the return variance v is denoted by OL. The policy denoted by C treats the return mean μ and variance v as constant and equal to the return sample average and the return sample variance at the beginning of the investment period, respectively.

Investment period starts in 2000

γ	R_f	K	s_1	$\sqrt{s_2}$	OL / Years to final date				C
					10	7	3	1	
5	5	15	2.48	7.47	61.2	72.4	86.7	93.5	99.2
5	5	30	1.88	7.90	54.3	59.1	65.0	68.0	70.4
5	6	15	2.25	7.47	56.2	66.4	79.5	85.6	90.7
5	6	30	1.64	7.90	48.3	53.0	58.2	60.7	62.7
10	5	15	2.48	7.47	29.3	35.1	42.8	46.8	49.6
10	5	30	1.88	7.90	26.5	28.9	32.2	33.8	35.1
10	6	15	2.25	7.47	26.9	32.3	39.4	42.8	45.3
10	6	30	1.64	7.90	23.6	25.9	28.8	30.2	31.2

Investment period starts in 2003

γ	R_f	K	s_1	$\sqrt{s_2}$	OL / Years to final date				C
					10	7	3	1	
5	5	15	1.66	7.88	39.9	47.0	55.8	60.2	63.4
5	5	30	1.65	8.56	42.8	46.6	51.1	53.4	55.2
5	6	15	1.42	7.88	35.3	41.4	49.3	53.0	55.7
5	6	30	1.42	8.56	37.9	40.9	45.2	47.2	48.7
10	5	15	1.66	7.88	19.1	22.8	27.6	29.9	31.6
10	5	30	1.65	8.56	20.7	22.7	25.3	26.5	27.4
10	6	15	1.42	7.88	16.9	20.2	24.4	26.3	27.7
10	6	30	1.42	8.56	18.3	20.2	22.4	23.5	24.2

Table 4: IID returns with unknown volatility - Certainty equivalent return

This table presents the certainty equivalent return for three portfolio policies when the continuously compounded excess returns are assumed to be i.i.d. with unknown mean and volatility. The upper panel corresponds to the case in which the investment period starts in 2000 and the lower panel corresponds to the case in which the investment period starts in 2003. The state variable is set equal to its value at the end of the corresponding sample. For each case, we consider eight different scenarios in which the coefficient of relative risk aversion (γ) takes the values 5 and 10; the annualized risk-free rate (R_f) takes the values 5 and 6 percent; and the sample size (K) takes the values 15 and 30 years. The investor uses quarterly data, updates the portfolio yearly, and has an investment horizon of 10 years. For reference, we report the corresponding return sample average (s_1) and return sample standard deviation ($\sqrt{s_2}$) at the beginning of the investment period. The optimal policy with learning about the return mean μ and the return variance v is denoted by OL. The policy denoted by OL-KV treats the return variance v as known and equal to the return sample variance at the beginning of the investment period, but is based on optimal learning about the return mean μ . The policy denoted by MU uses all available data to update the estimates of μ and v but in each period, myopically and ignoring future learning, solves the one-period portfolio allocation problem.

Investment period starts in 2000

γ	R_f	K	s_1	$\sqrt{s_2}$	OL	C	OLKV	PUNL	MU
5	5	15	2.48	7.47	8.84	7.12	8.81	7.76	8.57
5	5	30	1.88	7.90	7.60	7.32	7.57	7.40	7.54
5	6	15	2.25	7.47	9.25	7.78	9.23	8.34	9.02
5	6	30	1.64	7.90	8.09	7.86	8.07	7.92	8.05
10	5	15	2.48	7.47	6.82	5.86	6.81	6.21	6.67
10	5	30	1.88	7.90	6.26	6.11	6.24	6.15	6.22
10	6	15	2.25	7.47	7.55	6.77	7.54	7.04	7.42
10	6	30	1.64	7.90	7.01	6.89	7.00	6.92	6.99

Investment period starts in 2003

γ	R_f	K	s_1	$\sqrt{s_2}$	OL	C	OLKV	PUNL	MU
5	5	15	1.66	7.88	6.81	6.07	6.80	6.32	6.69
5	5	30	1.65	8.56	6.88	6.69	6.88	6.74	6.84
5	6	15	1.42	7.88	7.45	6.87	7.44	7.04	7.36
5	6	30	1.42	8.56	7.49	7.32	7.49	7.36	7.46
10	5	15	1.66	7.88	5.87	5.48	5.86	5.60	5.80
10	5	30	1.65	8.56	5.91	5.82	5.91	5.84	5.89
10	6	15	1.42	7.88	6.69	6.38	6.69	6.47	6.64
10	6	30	1.42	8.56	6.72	6.64	6.72	6.66	6.71

Table 5: Autoregressive returns - Optimal allocation

This table presents the optimal allocation to the risky asset when the continuously compounded excess returns are assumed to be autoregressive. The upper panel corresponds to the case in which the investment period starts in 2000 and the lower panel corresponds to the case in which the investment period starts in 2003. The state variable is set equal to its value at the end of the corresponding sample. For each case, we consider eight different scenarios in which the coefficient of relative risk aversion (γ) takes the values 5 and 10; the annualized risk-free rate (R_f) takes the values 5 and 6 percent; and the sample size (K) takes the values 15 and 30 years. The investor uses quarterly data, updates the portfolio yearly, and has an investment horizon of 10 years. For reference, we report the values of the state variables at the beginning of the investment period: the return sample average (s_1), return sample standard deviation ($\sqrt{s_2}$), autocorrelation estimate (s_3) and final return observation (s_4). The optimal policy with learning about the unknown parameters is denoted by OL.

Investment period starts in 2000

γ	R_f	K	s_1	$\sqrt{s_2}$	s_3	s_4	OL / Years to final date			
							10	7	3	1
5	5	15	2.48	7.47	-0.08	-9.35	66.4	82.2	100.0	100.0
5	5	30	1.88	7.90	0.05	-9.35	44.6	49.3	55.2	59.0
5	6	15	2.25	7.47	-0.08	-9.59	63.1	76.6	93.2	99.6
5	6	30	1.64	7.90	0.05	-9.59	39.8	43.5	49.0	52.2
10	5	15	2.48	7.47	-0.08	-9.35	31.6	39.7	50.0	54.4
10	5	30	1.88	7.90	0.05	-9.35	21.6	24.1	27.2	29.2
10	6	15	2.25	7.47	-0.08	-9.59	30.1	36.8	46.3	50.1
10	6	30	1.64	7.90	0.05	-9.59	19.2	21.4	24.2	26.0

Investment period starts in 2003

γ	R_f	K	s_1	$\sqrt{s_2}$	s_3	s_4	OL / Years to final date			
							10	7	3	1
5	5	15	1.66	7.88	-0.03	10.24	41.3	46.9	56.2	60.4
5	5	30	1.65	8.56	0.05	10.24	43.2	46.5	50.6	53.2
5	6	15	1.42	7.88	-0.03	10.00	36.7	41.8	49.6	52.7
5	6	30	1.42	8.56	0.05	10.00	38.4	41.5	45.1	47.5
10	5	15	1.66	7.88	-0.03	10.24	19.8	23.5	27.7	29.9
10	5	30	1.65	8.56	0.05	10.24	21.0	22.8	25.0	26.4
10	6	15	1.42	7.88	-0.03	10.00	17.5	20.3	24.5	26.3
10	6	30	1.42	8.56	0.05	10.00	18.8	20.3	22.4	23.4

Table 6: Autoregressive returns - Certainty equivalent return

This table presents the certainty equivalent return for three portfolio policies when the continuously compounded excess returns are assumed to be autoregressive. The upper panel corresponds to the case in which the investment period starts in 2000 and the lower panel corresponds to the case in which the investment period starts in 2003. The state variable is set equal to its value at the end of the corresponding sample. For each case, we consider eight different scenarios in which the coefficient of relative risk aversion (γ) takes the values 5 and 10; the annualized risk-free rate (R_f) takes the values 5 and 6 percent; and the sample size (K) takes the values 15 and 30 years. The investor uses quarterly data, updates the portfolio yearly, and has an investment horizon of 10 years. For reference, we report the values of the state variables at the beginning of the investment period: the return sample average (s_1), return sample standard deviation ($\sqrt{s_2}$), autocorrelation estimate (s_3) and final return observation (s_4). The optimal policy with learning about the unknown parameters is denoted by OL. The policy denoted by OL-IID treats returns as i.i.d., but is based on optimal learning about the return mean and variance. The policy denoted by MU uses all available data to update the estimates of all parameters but in each period, myopically and ignoring future learning, solves the one-period portfolio allocation problem.

Investment period starts in 2000

γ	R_f	K	s_1	$\sqrt{s_2}$	s_3	s_4	OL	OL-IID	MU
5	5	15	2.48	7.47	-0.08	-9.35	9.24	9.19	8.93
5	5	30	1.88	7.90	0.05	-9.35	7.35	7.28	7.27
5	6	15	2.25	7.47	-0.08	-9.59	9.63	9.59	9.32
5	6	30	1.64	7.90	0.05	-9.59	7.90	7.84	7.84
10	5	15	2.48	7.47	-0.08	-9.35	7.01	6.98	6.75
10	5	30	1.88	7.90	0.05	-9.35	6.14	6.10	6.09
10	6	15	2.25	7.47	-0.08	-9.59	7.73	7.70	7.52
10	6	30	1.64	7.90	0.05	-9.59	6.92	6.89	6.89

Investment period starts in 2003

γ	R_f	K	s_1	$\sqrt{s_2}$	s_3	s_4	OL	OL-IID	MU
5	5	15	1.66	7.88	-0.03	10.24	6.85	6.83	6.71
5	5	30	1.65	8.56	0.05	10.24	6.81	6.77	6.76
5	6	15	1.42	7.88	-0.03	10.00	7.47	7.45	7.37
5	6	30	1.42	8.56	0.05	10.00	7.45	7.41	7.41
10	5	15	1.66	7.88	-0.03	10.24	5.89	5.87	5.81
10	5	30	1.65	8.56	0.05	10.24	5.87	5.85	5.85
10	6	15	1.42	7.88	-0.03	10.00	6.70	6.69	6.65
10	6	30	1.42	8.56	0.05	10.00	6.70	6.68	6.68

Table 7: Predictable returns - Optimal Allocation

This table presents the optimal allocation to the risky asset when the continuously compounded excess returns are assumed to be predictable. The upper (lower) panel corresponds to the case in which the investor uses the term spread (earnings-to-price ratio) as the predictor. The initial estimation is based on quarterly data from the sample period 1984-2003. The state variables s_k , $k = 1, \dots, 7$ are set equal to their values at the end of the sample. For each predictor, we consider four different scenarios in which the coefficient of relative risk aversion (γ) takes the values 5 and 10; and the annualized risk-free rate (R_f) takes the values 5 and 6 percent. For each scenario, we present the optimal allocations that correspond to the predictor being equal to x_{25} , x_{50} , and x_{75} , respectively, where x_p denotes the p -th percentile of the sample distribution of the predictor. The investor uses quarterly data, updates the portfolio yearly, and has an investment horizon of 5 years. The optimal policy with learning about the unknown parameters is denoted by OL and the policy that treats the parameters as known constants is denoted by C. Under policy C the parameters are set equal to their estimates obtained using data from 1984 to 2003.

		Years to final date										
γ	R_f	5		4		3		2		1		
		OL	C	OL	C	OL	C	OL	C	OL	C	
Predictor : Earnings-to-price ratio												
5	5	x_{25}	24.7	31.2	22.2	26.8	19.1	22.0	16.1	17.2	12.7	12.5
		x_{50}	100.0	100.0	98.3	100.0	92.6	94.5	87.9	90.8	79.3	79.9
		x_{75}	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
5	6	x_{25}	14.4	18.9	12.2	15.1	9.5	11.2	7.0	7.2	4.1	3.6
		x_{50}	92.0	100.0	88.6	96.7	83.2	88.7	78.3	81.6	70.6	70.9
		x_{75}	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
10	5	x_{25}	12.9	18.1	11.7	15.0	10.0	12.0	8.3	8.9	6.4	6.2
		x_{50}	53.2	66.9	51.5	61.3	47.9	54.3	44.7	47.6	39.7	39.9
		x_{75}	71.7	100.0	72.2	95.0	69.9	85.5	68.5	76.9	63.7	65.0
10	6	x_{25}	7.7	11.4	6.4	8.8	5.0	6.4	3.6	3.9	2.2	1.8
		x_{50}	48.0	59.1	46.0	54.2	42.7	48.0	39.8	42.2	35.3	35.3
		x_{75}	68.4	95.5	68.0	89.0	65.7	79.9	63.9	71.5	59.4	60.6

Table 8: Predictable returns using term spread - Optimal Allocation

This table presents the optimal allocation to the risky asset when the continuously compounded excess returns are assumed to be predictable by the term spread. The investor uses 20 years of data in the initial estimation. The three panels correspond to the cases in which the investment period starts in 2000, 2001, and 2002, respectively. The state variables s_k , $k = 1, \dots, 7$ are set equal to their values at the end of the corresponding sample. For each predictor, we consider four different scenarios in which the coefficient of relative risk aversion (γ) takes the values 5 and 10; and the annualized risk-free rate (R_f) takes the values 5 and 6 percent. For each scenario, we present the optimal allocations that correspond to the predictor being equal to x_{25} , x_{50} , and x_{75} , respectively, where x_p denotes the p -th percentile of the sample distribution of the predictor. The investor uses quarterly data, updates the portfolio yearly, and has an investment horizon of 5 years.

γ	R_f		Years to final date				
			5	4	3	2	1
Investment period starts in 2000							
5	5	x_{25}	77.5	80.9	84.2	87.8	91.2
		x_{50}	87.2	90.5	93.8	96.9	100.0
		x_{75}	98.4	99.3	100.0	100.0	100.0
5	6	x_{25}	69.6	72.5	75.6	79.3	83.2
		x_{50}	78.9	81.9	85.0	88.9	92.8
		x_{75}	90.1	93.1	96.2	98.3	100.0
10	5	x_{25}	37.5	39.4	41.3	42.9	44.5
		x_{50}	42.1	44.1	46.1	47.8	49.5
		x_{75}	47.5	49.6	51.8	53.9	56.1
10	6	x_{25}	34.1	35.9	37.7	39.4	41.1
		x_{50}	38.6	40.5	42.4	44.0	45.7
		x_{75}	44.1	46.1	48.1	49.9	51.7
Investment period starts in 2001							
5	5	x_{25}	68.2	71.2	74.3	77.1	79.9
		x_{50}	74.5	77.6	80.7	83.8	86.8
		x_{75}	79.8	83.6	87.3	91.2	95.0
5	6	x_{25}	61.7	64.3	66.9	69.4	71.9
		x_{50}	68.5	71.0	73.6	76.3	79.0
		x_{75}	74.3	77.3	80.4	83.9	87.4
10	5	x_{25}	33.5	35.1	36.7	38.3	39.8
		x_{50}	36.7	38.4	40.0	41.7	43.3
		x_{75}	39.3	41.2	43.1	45.2	47.4
10	6	x_{25}	30.4	31.7	33.1	34.5	35.9
		x_{50}	33.6	35.1	36.5	37.9	39.4
		x_{75}	36.4	38.0	39.7	41.6	43.4
Investment period starts in 2002							
5	5	x_{25}	57.1	59.5	62.0	65.0	68.0
		x_{50}	50.0	52.2	54.5	57.1	59.8
		x_{75}	38.3	40.5	42.7	45.3	48.0
5	6	x_{25}	51.5	53.6	55.8	58.5	61.2
		x_{50}	44.5	46.4	48.3	50.6	53.0
		x_{75}	33.0	34.8	36.6	38.9	41.3
10	5	x_{25}	28.0	29.2	30.4	32.3	34.0
		x_{50}	24.4	25.6	26.8	28.3	29.8
		x_{75}	18.6	19.8	20.9	22.4	23.8
10	6	x_{25}	25.3	26.4	27.6	29.1	30.5
		x_{50}	21.8	22.9	23.9	25.1	26.4
		x_{75}	16.2	17.1	18.1	19.3	20.6

Table 9: Predictable returns using earnings-to-price ratio - Optimal Allocation

This table presents the optimal allocation to the risky asset when the continuously compounded excess returns are assumed to be predictable by the earnings-to-price ratio. The investor uses 20 years of data in the initial estimation. The three panels correspond to the cases in which the investment period starts in 2000, 2001, and 2002, respectively. The state variables s_k , $k = 1, \dots, 7$ are set equal to their values at the end of the corresponding sample. For each predictor, we consider four different scenarios in which the coefficient of relative risk aversion (γ) takes the values 5 and 10; and the annualized risk-free rate (R_f) takes the values 5 and 6 percent. For each scenario, we present the optimal allocations that correspond to the predictor being equal to x_{25} , x_{50} , and x_{75} , respectively, where x_p denotes the p -th percentile of the sample distribution of the predictor. The investor uses quarterly data, updates the portfolio yearly, and has an investment horizon of 5 years.

γ	R_f		Years to final date				
			5	4	3	2	1
Investment period starts in 2000							
5	5	x_{25}	71.0	71.6	72.3	73.0	73.7
		x_{50}	89.7	90.7	91.5	91.0	90.4
		x_{75}	96.5	98.4	100.0	100.0	100.0
5	6	x_{25}	62.7	63.4	64.0	64.6	65.1
		x_{50}	81.0	82.0	82.8	81.9	81.1
		x_{75}	88.9	90.9	93.0	92.3	92.2
10	5	x_{25}	35.6	35.9	36.2	36.4	36.7
		x_{50}	45.0	45.3	45.6	45.5	45.3
		x_{75}	47.5	48.8	50.1	50.7	51.3
10	6	x_{25}	30.9	31.2	31.6	31.9	32.3
		x_{50}	40.4	40.8	41.1	41.1	41.0
		x_{75}	44.5	45.3	46.1	46.7	47.3
Investment period starts in 2001							
5	5	x_{25}	57.4	53.5	49.5	45.6	41.8
		x_{50}	100.0	98.3	94.9	88.5	81.9
		x_{75}	100.0	100.0	100.0	100.0	100.0
5	6	x_{25}	45.8	42.5	39.3	36.4	33.4
		x_{50}	91.8	88.4	84.9	79.1	73.0
		x_{75}	100.0	100.0	100.0	99.7	99.5
10	5	x_{25}	29.6	27.2	24.9	23.0	21.1
		x_{50}	54.8	52.0	49.0	45.1	41.2
		x_{75}	62.7	62.7	62.3	59.0	55.7
10	6	x_{25}	23.5	21.6	19.6	18.0	16.5
		x_{50}	48.8	45.8	42.8	39.7	36.5
		x_{75}	59.3	58.0	56.5	53.8	50.9
Investment period starts in 2002							
5	5	x_{25}	24.0	19.4	14.6	9.9	5.2
		x_{50}	100.0	98.0	93.8	83.8	73.3
		x_{75}	100.0	100.0	100.0	100.0	100.0
5	6	x_{25}	11.3	7.3	3.4	1.2	0.0
		x_{50}	96.7	89.7	82.6	73.3	63.9
		x_{75}	100.0	100.0	100.0	100.0	100.0
10	5	x_{25}	13.8	11.0	8.2	5.5	2.8
		x_{50}	61.2	56.0	50.5	43.6	36.7
		x_{75}	83.5	80.3	76.6	69.1	61.2
10	6	x_{25}	7.2	4.8	2.4	1.0	0.0
		x_{50}	53.9	48.9	43.8	37.8	31.7
		x_{75}	78.9	74.9	70.6	63.6	56.3

Table 10: Predictable returns using term spread - Certainty equivalent return

This table presents the certainty equivalent return for three portfolio policies when the continuously compounded excess returns are assumed to be i.i.d. with unknown mean and volatility. The upper panel corresponds to the case in which the investment period starts in 2000 and the lower panel corresponds to the case in which the investment period starts in 2003. The state variable is set equal to its value at the end of the corresponding sample. For each case, we consider eight different scenarios in which the coefficient of relative risk aversion (γ) takes the values 5 and 10; the annualized risk-free rate (R_f) takes the values 5 and 6 percent; and the sample size (K) takes the values 15 and 30 years. The investor uses quarterly data, updates the portfolio yearly, and has an investment horizon of 10 years. For reference, we report the corresponding return sample average (s_1) and return sample standard deviation ($\sqrt{s_2}$) at the beginning of the investment period. The optimal policy with learning about the return mean μ and the return variance v is denoted by OL. The policy denoted by OL-KV treats the return variance v as known and equal to the return sample variance at the beginning of the investment period, but is based on optimal learning about the return mean μ . The policy denoted by MU uses all available data to update the estimates of μ and v but in each period, myopically and ignoring future learning, solves the one-period portfolio allocation problem.

Investment period starts in 2003

γ	R_f	K	s_1	$\sqrt{s_2}$	OL	C	IID	PUNL	MU
5	5	20	1.82	8.25	7.40	7.16	7.31	7.27	7.34
5	6	20	1.59	8.25	7.94	7.74	7.86	7.81	7.89
10	5	20	1.82	8.25	6.16	6.04	6.11	6.09	6.13
10	6	20	1.59	8.25	6.94	6.83	6.90	6.88	6.92

Figure 1: IID returns with known volatility - Optimal allocation

These graphs present the optimal allocation to the risky asset for the case of i.i.d. returns with known volatility. In the upper graph, the optimal allocation is plotted as a function of the state variable s for investment horizon equal to 1, 3, 7, and 10 years respectively. In the lower graph, the optimal allocation is plotted as a function of the investment horizon. The three lines correspond to the state variable s being equal to s_{25} , s_{50} , and s_{75} where s_p denotes the p -th percentile of the distribution of s . The coefficient of relative risk aversion equals 5, the annualized risk-free rate equals 6 percent, the estimation is based on 15 years of data, and the sample ends in 2003.

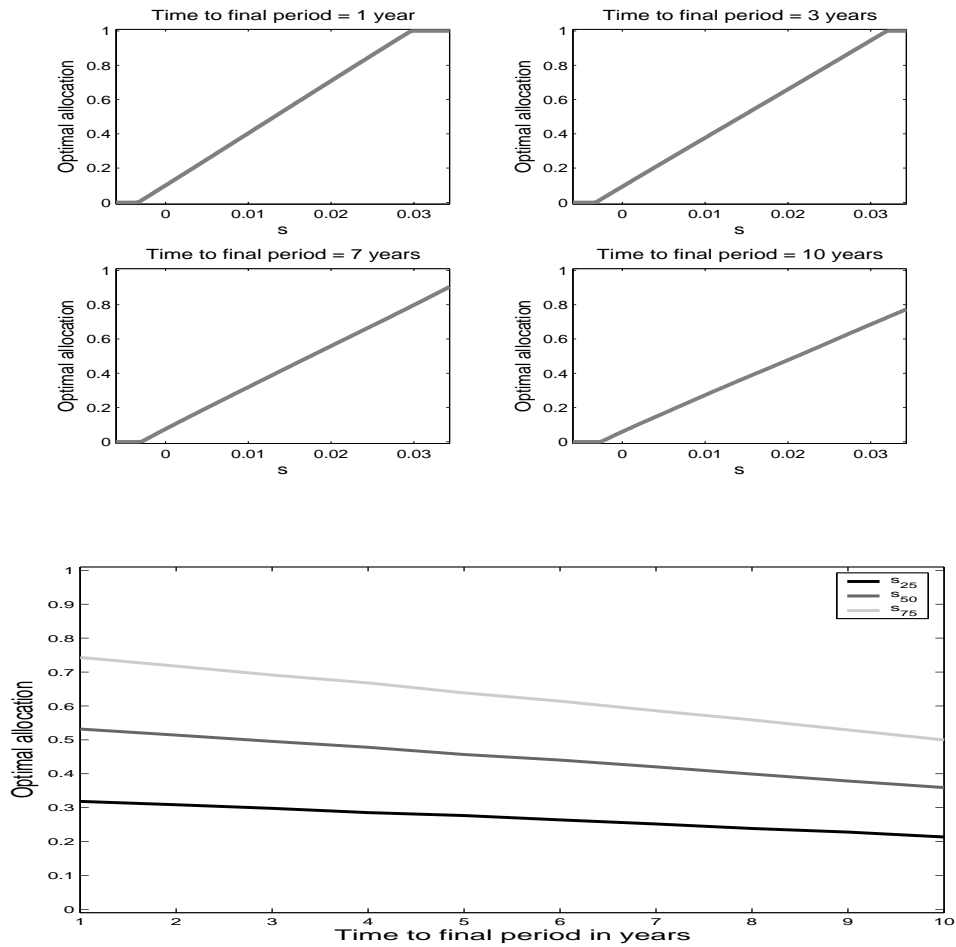


Figure 2: IID returns with unknown volatility - Optimal allocation

These graphs present the optimal allocation to the risky asset for the case of i.i.d. returns with unknown volatility. In the upper graph, the optimal allocation is plotted as a function of the state variable s_1 . The three lines correspond to the state variable s_2 being equal to $s_{2,25}$, $s_{2,50}$, and $s_{2,75}$ where $s_{2,p}$ denotes the p -th percentile of the distribution of s_2 . In the lower graph, the optimal allocation is plotted as a function of the state variable s_2 . The three lines correspond to the state variable s_1 being equal to $s_{1,25}$, $s_{1,50}$, and $s_{1,75}$ where $s_{1,p}$ denotes the p -th percentile of the distribution of s_1 . In both graphs, the optimal allocation is presented for investment horizon equal to 1, 3, 7, and 10 years respectively. The coefficient of relative risk aversion equals 5, the annualized risk-free rate equals 6 percent, the estimation is based on 15 years of data, and the sample ends in 2003.

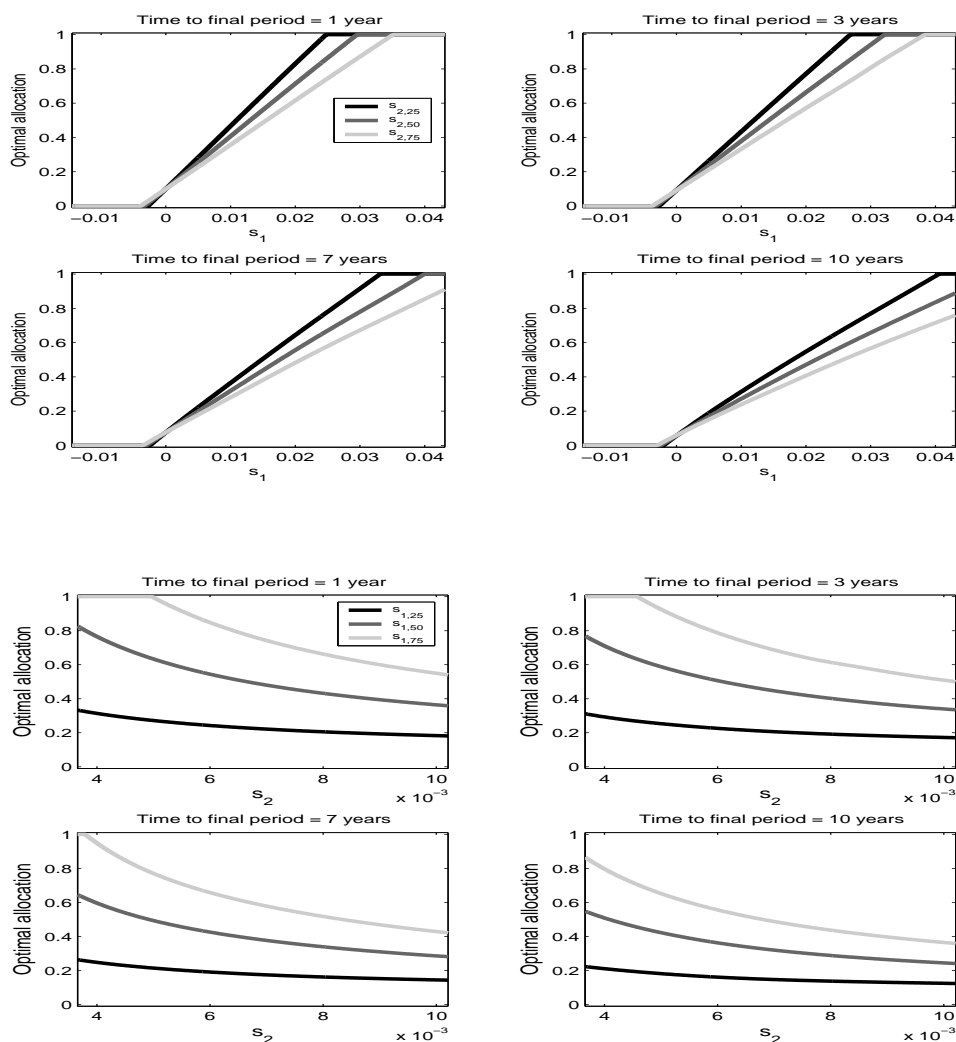


Figure 3: IID returns with unknown volatility - Optimal allocation

These graphs present the optimal allocation to the risky asset for the case of i.i.d. returns with unknown volatility. The optimal allocation is plotted as a function of the investment horizon. In the upper, middle, and lower graphs the state variable s_2 is set equal to the 25th, 50th, and 75th percentile of its distribution, respectively. In each graph, the three lines correspond to the state variable s_1 being equal to $s_{1,25}$, $s_{1,50}$, and $s_{1,75}$ where $s_{1,p}$ denotes the p -th percentile of the distribution of s_1 . The coefficient of relative risk aversion equals 5, the annualized risk-free rate equals 6 percent, the estimation is based on 15 years of data, and the sample ends in 2003.

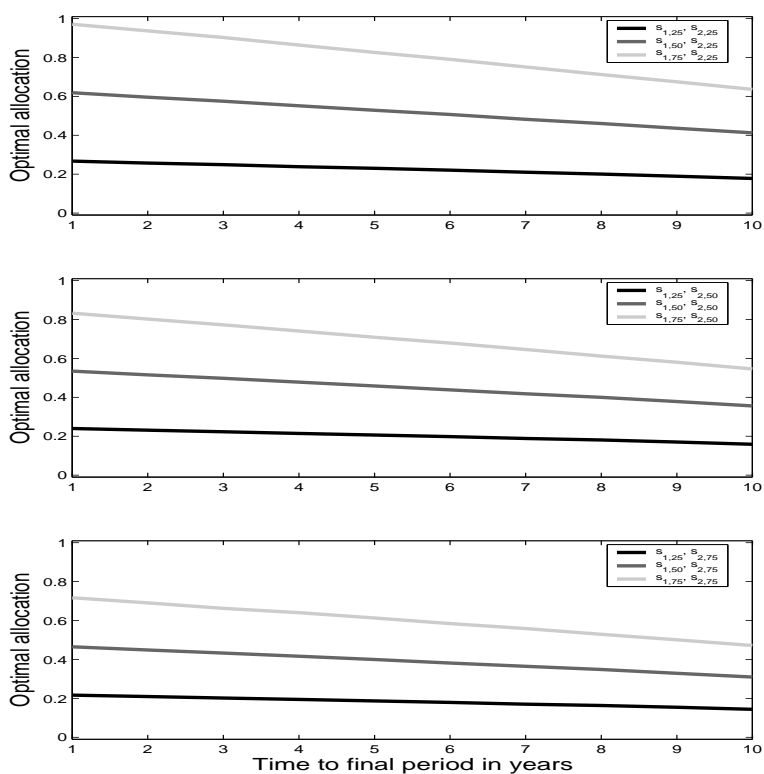


Figure 4: IID returns with unknown volatility - Optimal allocation

These graphs present the optimal allocation to the risky asset for the case of i.i.d. returns with unknown volatility. The optimal allocation is plotted as a function of the investment horizon. In the upper, middle, and lower graphs the state variable s_1 is set equal to the 25th, 50th, and 75th percentile of its distribution, respectively. In each graph, the three lines correspond to the state variable s_2 being equal to $s_{2,25}$, $s_{2,50}$, and $s_{2,75}$ where $s_{2,p}$ denotes the p -th percentile of the distribution of s_2 . The coefficient of relative risk aversion equals 5, the annualized risk-free rate equals 6 percent, the estimation is based on 15 years of data, and the sample ends in 2003.

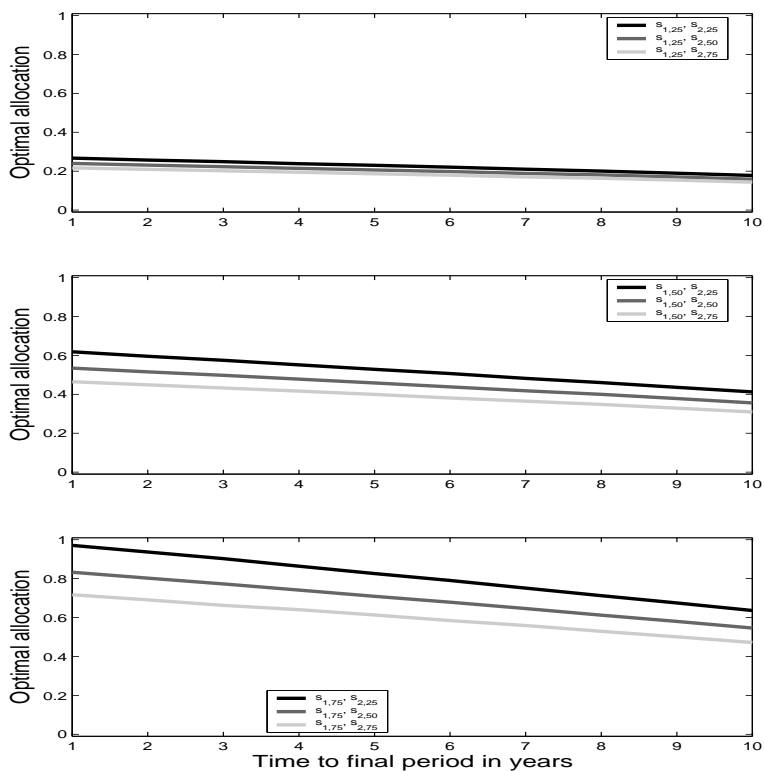


Figure 5: Autoregressive returns - Optimal allocation

These graphs present the optimal allocation to the risky asset for the case of autoregressive returns. For $k = 1, 2, 3, 4$, the k th graph displays the optimal allocation as a function of the investment horizon with the three lines corresponding to the state variable s_k being equal to $s_{k,25}$, $s_{k,50}$, and $s_{k,75}$ where $s_{k,p}$ denotes the p -th percentile of the distribution of s_k . The coefficient of relative risk aversion equals 5, the annualized risk-free rate equals 6 percent, the estimation is based on 30 years of data, and the initial sample ends in 2000. In each graph, the rest of the state variables are set equal to their values at the end of 2000.

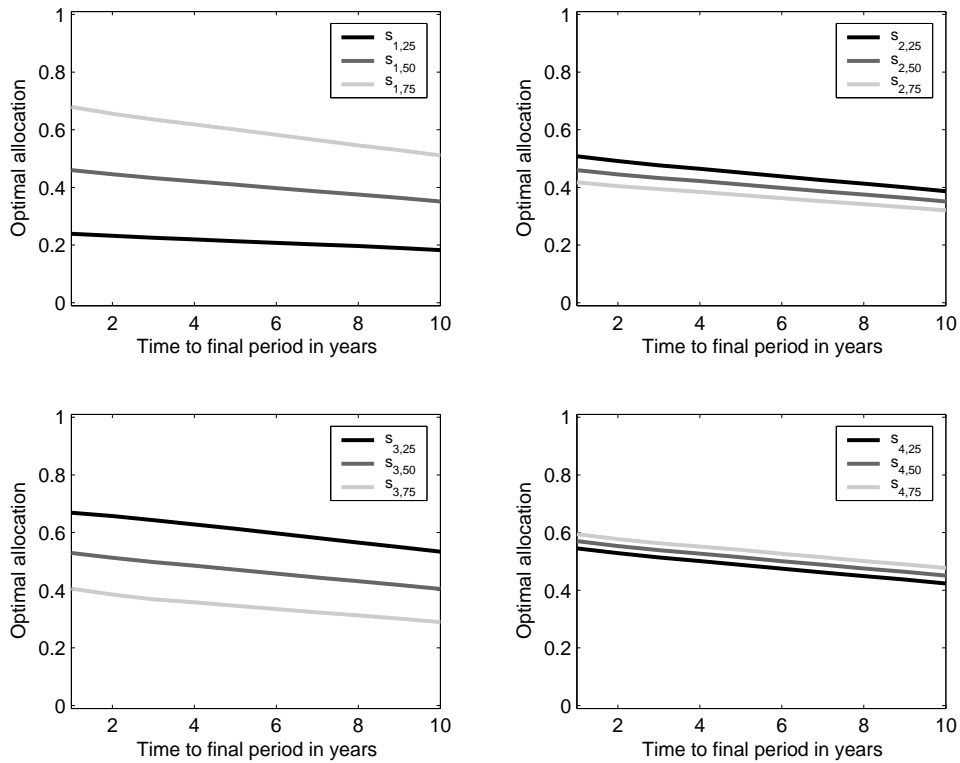


Figure 6: Autoregressive returns - Optimal allocation

These graphs present the optimal allocation to the risky asset for the case of autoregressive returns. For $k = 1, 2, 3, 4$, the k th graph displays the optimal allocation as a function of the investment horizon with the three lines corresponding to the state variable s_k being equal to $s_{k,25}$, $s_{k,50}$, and $s_{k,75}$ where $s_{k,p}$ denotes the p -th percentile of the distribution of s_k . The coefficient of relative risk aversion equals 5, the annualized risk-free rate equals 6 percent, the estimation is based on 30 years of data, and the initial sample ends in 2003. In each graph, the rest of the state variables are set equal to their values at the end of 2003.

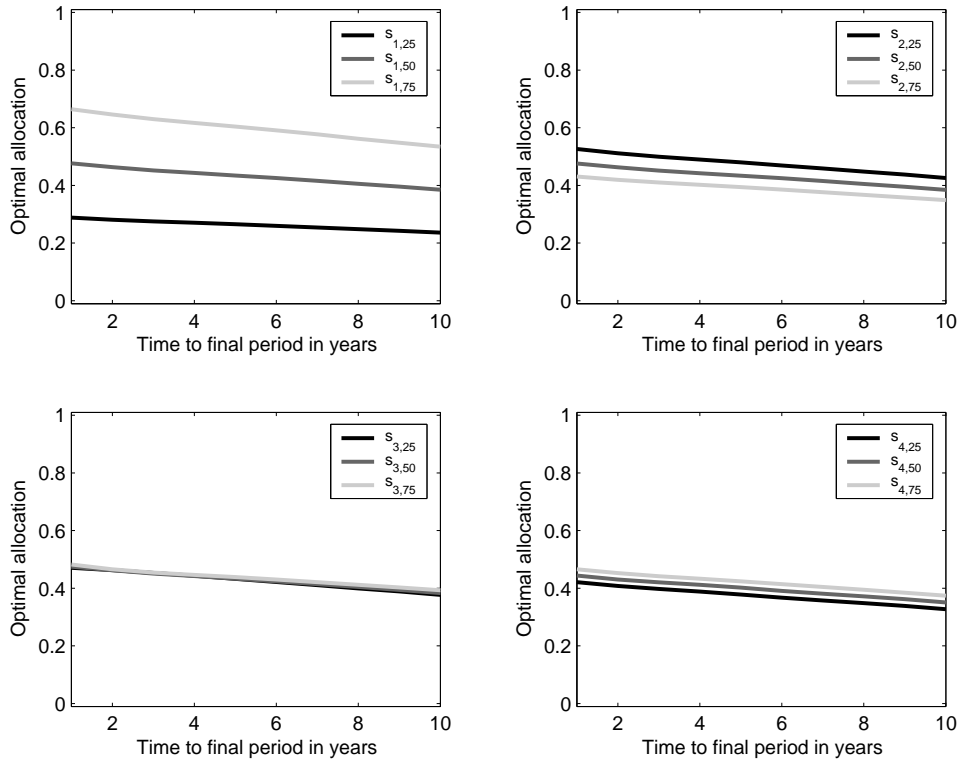


Figure 7: Predictable returns using term spread - Optimal allocation

These graphs present the optimal allocation to the risky asset when returns are assumed to be predictable by the term spread. For $k = 1, 2, 3, 4$, the k th graph displays the optimal allocation as a function of the investment horizon with the three lines corresponding to the state variable s_k being equal to $s_{k,25}$, $s_{k,50}$, and $s_{k,75}$ where $s_{k,p}$ denotes the p -th percentile of the distribution of s_k . The coefficient of relative risk aversion equals 5, the annualized risk-free rate equals 6 percent, the estimation is based on 20 years of data, and the initial sample ends in 2003. In each graph, the rest of the state variables are set equal to their median values.

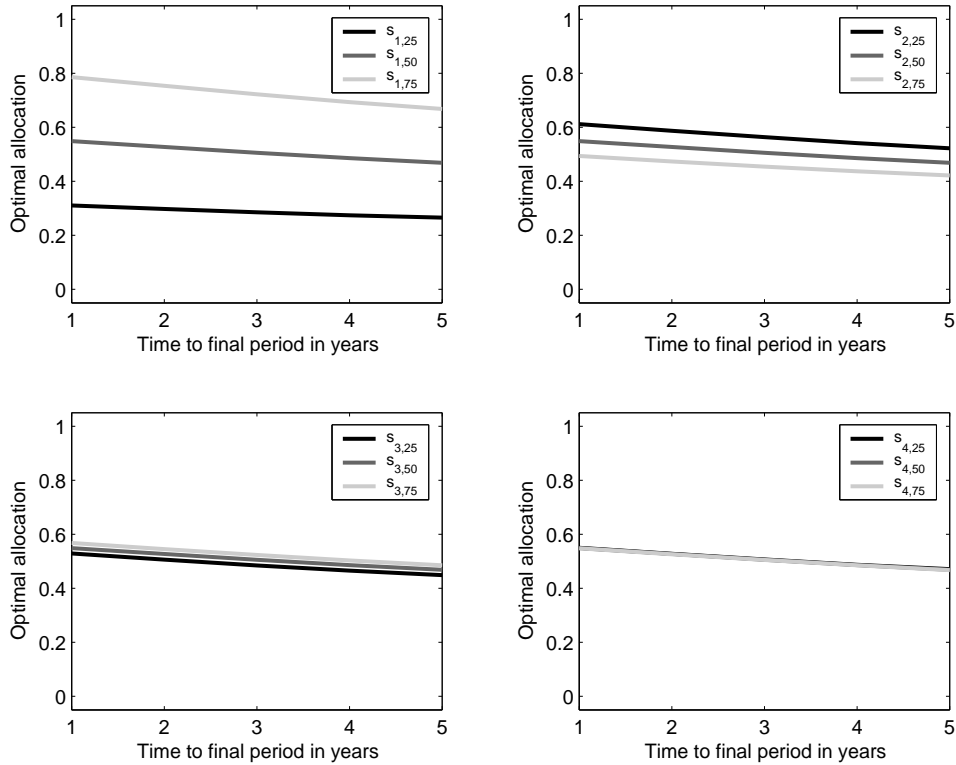


Figure 8: Predictable returns using earnings-to-price ratio - Optimal allocation

These graphs present the optimal allocation to the risky asset when returns are assumed to be predictable by the earnings-to-price ratio. For $k = 1, 2, 3, 4$, the k th graph displays the optimal allocation as a function of the investment horizon with the three lines corresponding to the state variable s_k being equal to $s_{k,25}$, $s_{k,50}$, and $s_{k,75}$ where $s_{k,p}$ denotes the p -th percentile of the distribution of s_k . The coefficient of relative risk aversion equals 5, the annualized risk-free rate equals 6 percent, the estimation is based on 20 years of data, and the initial sample ends in 2003. In each graph, the rest of the state variables are set equal to their median values.

