Long-term strategic asset allocation: An out-of-sample evaluation

Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Long-Term Strategic Asset Allocation: An Out-of-Sample Evaluation

Bart Diris, Franz Palm, Peter Schotman

To cite this article:
Bart Diris, Franz Palm, Peter Schotman (2015) Long-Term Strategic Asset Allocation: An Out-of-Sample Evaluation. Management Science 61(9):2185-2202. https://doi.org/10.1287/mnsc.2014.1924

MANAGEMENT SCIENCE
Vol. 61, No. 9, September 2015, pp. 2185–2202
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2014.1924

Long-Term Strategic Asset Allocation: An Out-of-Sample Evaluation

Bart Diris
Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, 3000 DR, Rotterdam, The Netherlands, diris@ese.eur.nl

Franz Palm
Quantitative Economics Department, Maastricht University, 6200 MD, Maastricht, The Netherlands, f.palm@maastrichtuniversity.nl

Peter Schotman
Finance Department, Maastricht University, 6200 MD, Maastricht, The Netherlands, p.schotman@maastrichtuniversity.nl

We evaluate the out-of-sample performance of a long-term investor who follows an optimized dynamic trading strategy. Although the dynamic strategy is able to benefit from predictability out-of-sample, a short-term investor using a single-period market timing strategy would have realized an almost identical performance. The value of intertemporal hedge demands in strategic asset allocation appears negligible. The result is caused by the estimation error in predicting the predictors. A myopic investor only needs to predict one-period-ahead expected returns, but hedge demands also require accurate predictions of the predictor variables. To reduce the problem of errors in optimized portfolio weights, we consider Bayesian procedures. Myopic and dynamic portfolios are similarly affected by such modifications, and differences in performance become even smaller.

Data, as supplemental material, are available at http://dx.doi.org/10.1287/mnsc.2014.1924.

Keywords: strategic asset allocation; out-of-sample analysis; performance evaluation; finance; portfolio
History: Received December 19, 2011; accepted January 21, 2014, by Brad Barber, finance. Published online in *Articles in Advance* July 11, 2014.

1. Introduction

We compare the long-term out-of-sample performance of two types of investors. Both investors can choose between equity, long-term nominal bonds and short-term nominally risk-free Treasury bills (T-bills). The first investor follows the dynamic portfolio strategy derived from optimizing a long-term objective function. The second investor solves the much simpler single-period optimization problem and follows a myopic strategy each period. For a moderate level of relative risk aversion, we find that the two types of investors perform almost equally well out-of-sample when we evaluate them using the long-term objective function.

This seems a puzzling result in the light of the literature on strategic asset allocation. An important element in the optimal portfolios of long-term investors is the hedge demand that anticipates future changes in investment opportunities. The theoretical foundation for these hedge demands dates back to Merton (1969, 1971). The empirical importance has been established by Brennan et al. (1997), Campbell et al. (2003), and many others who show that long-term investors should hold sizable hedge portfolios using an empirical model with time-varying expected returns on all three asset classes. The hedge portfolios have two components. The first relies on predictability of excess returns on stocks and bonds relative to the instantaneous risk-free rate. The other element reflects the exposure of long-term real wealth to inflation and interest rate risk from rolling over a short-term nominally risk-free asset. Subsequent studies, like Sangvinatsos and Wachter (2005), Jurek and Viceira (2011), and Larsen and Munk (2012), have pointed at large certainty equivalent utility gains from both types of hedge demands.

The portfolio weights and utility gains in these and similar studies of strategic asset allocation are in-sample, based on estimated models. In the portfolio choice literature it has been recognized that the theoretical gains may be hard to realize out-of-sample because parameter estimation errors are exacerbated by the optimization. In general, DeMiguel et al. (2009) find that in many cases a naive equally weighted portfolio of all asset classes outperforms an optimized portfolio. More specifically, for the case of predicting equity returns, Goyal and Welch (2008) and others find that out-of-sample predictability of stock returns is minimal. Even so, Campbell and Thompson (2008), Pesaran and Timmermann (1995), Wachter and Warusawitharana

(2009), and Dangl and Halling (2012) all find that trading on this predictability adds value if sensible restrictions on portfolio weights are imposed or parameter uncertainty and instability are accounted for. Although these studies find evidence of useful out-of-sample predictability, they consider only myopic trading strategies aimed at optimizing a single-period performance criterion, ignoring hedge demands. Our results pertain to the added value of hedge demands.

Quantifying the hedge portfolios requires much from an empirical model. Just as for the myopic demands, one needs an estimate of the time variation in expected returns. On top of that, the optimal hedge demands depend on the covariance of innovations in predictor variables with the future variation in expected returns. In a vector autoregressive model (VAR), this means that one needs the full VAR, i.e., not just the predictive equations for asset returns, but also the equations for the state variables that predict returns. This may be asking too much from a model. Our results seem consistent, therefore, with the recent work of Pastor and Stambaugh (2012), who argue that misspecification of the VAR for stock returns causes equity to be even more risky for long-term investors despite significant evidence of predictability.

Although there are many studies on the out-of-sample performance of strategies that maximize a single-period objective, little evidence exists for dynamic strategies based on intertemporal optimization. The seminal Brennan et al. (1997) study looks at a single realization of following a dynamic strategy with a 20 year investment horizon for 20 years. Another example is the study by DeMiguel et al. (2005),[1] who evaluate the Campbell et al. (2003) model. They note that this model has an infinite horizon and also includes an optimal consumption plan, which makes the evaluation over a finite horizon somewhat problematical. With that caveat, they find that strategic asset allocation based on plug-in parameter estimates and unrestricted portfolio weights does not outperform naive diversification. Johannes et al. (2014) study the out-of-sample performance of dynamic portfolio choice with a two-year investment horizon in a model with equity as the only risky asset. In this setting they conclude that dynamic strategies improve over a static benchmark once time-varying volatility is included. However, they do not separate the gains in a component due to myopic market timing and a component due to hedge demands.

In designing our out-of-sample evaluation we make a number of choices. First, we use the asset menu of stocks, nominal bonds and cash considered by Brennan et al. (1997) and many others.[2] With these three asset classes, long-term investors are subject to real interest rate risk. This distinguishes our setting from studies that assume the existence of a real risk-free rate,[3] and hence only concentrate on the hedge demands from predictable movements in expected excess returns.

A second choice is the objective function. We maximize the expected constant relative risk-aversion (CRRA) utility of real end-of-period wealth for a finite horizon as in Brennan et al. (1997), Jurek and Viceira (2011), Brandt et al. (2005), and others. This means that we do not include intermediate consumption as in Campbell et al. (2003), Rapach and Wohar (2009), Engsted and Pedersen (2012) and others. Both objectives generate intertemporal hedge demands, but we opt for the finite horizon expected utility of terminal wealth $\mathrm{E}_t[U(W_{t+K})]$ criterion. This allows us to construct time series of out-of-sample realized utilities based on returns from different portfolio strategies. With these we can compare the dynamic optimization results with myopic portfolios. We thus sidestep modeling consumption smoothing and labor income in a realistic life-cycle model. Since the importance of hedge demands increases with the investment horizon, we would like to take the investment horizon K as long as possible. But the longer the investment horizon, the fewer independent observations we will have on realized utilities. As a compromise we set K to five years for our main results.

Third, an important lesson from much of the single-period portfolio choice literature is that it is important to use some form of shrinkage estimators to reduce the effects of parameter estimation error. In Campbell et al. (2003) unrestricted optimized portfolios for long-term investors show wildly fluctuating portfolio weights that are even more extreme than for short-term investors. Campbell et al. (2003) acknowledge the phenomenon and suggest Bayesian priors and short-sell restrictions for future work.[4] In this respect, we follow Wachter and Warusawitharana (2009) and specify a skeptical Bayesian prior that shrinks slope coefficients in the predictive regressions for excess returns to zero, and shrinks the coefficients of the state variables to a random walk. For our six-dimensional VAR, we use a general Bayesian shrinkage prior for time-series models advocated by Ni and Sun (2003). With the shrinkage prior optimized portfolio strategies become less aggressive compared to

[1] DeMiguel et al. (2005) is an early draft of DeMiguel et al. (2009) that contains evidence on intertemporal portfolio optimization that is not in the final version.

[2] See, e.g., Brennan and Xia (2002), Koijen et al. (2010), Sangvinatsos and Wachter (2005), and Campbell et al. (2003).

[3] See, e.g., Barberis (2000), Chacko and Viceira (2005), Lynch (2001), Campbell et al. (2001), Pettenuzzo and Timmermann (2011), and Branger et al. (2013).

[4] Examples from the growing Bayesian literature include Merton (1980), Cremers (2002), Jorion (1986), Black and Littermann (1992), Avramov (2002), and Pastor and Stambaugh (2000).

those for a uniform prior. We find that our results are robust to the choice of prior. It affects the out-of-sample performance of dynamic and myopic strategies, but does not change their relative performance.

A fourth implementation issue is the solution to the optimization problem. The "plug-in" method conditions on a given set of estimated parameters (the posterior mean). For the log-normal distribution and unrestricted portfolio weights, Jurek and Viceira (2011) derive an approximate closed-form solution for the portfolio weights. Much of the recent portfolio choice literature, e.g., Barberis (2000) and Johannes et al. (2014), advocates the use of Bayesian decision theory to explicitly account for parameter uncertainty in the optimization. In this case, and also when short-sell constraints are imposed, the optimization must be done numerically. Our performance analysis requires a fast and stable numerical algorithm, since we need to repeat the optimization many times. To make our extensive out-of-sample analysis feasible, we adapt the method of Brandt et al. (2005) by introducing a quadratic interpolation step that dramatically reduces the number of portfolios that must be evaluated. We describe the method in the online appendix (available at http://ssrn.com/abstract=1973648).

For our main results, we use monthly returns for the period 1954–2012 for a full VAR and an investment horizon of five years. Using a wide range of Bayesian priors and implementations of the portfolio optimization method, our main result is that repeated myopic portfolios perform as well as strategic portfolios. On a five-year horizon we do not find that hedge demands add value out-of-sample. We offer two explanations for the negligible value of hedge demands. First, we conduct a Monte Carlo study where we generate data according to the estimated VAR with predictable variation in expected returns. In the simulated data the estimated myopic rules slightly outperform the strategic rules. Both rules suffer from estimation error, but the strategic rule is hurt more by estimation error than the repeated myopic rule. The additional exposure to estimation risk offsets the potential gains from hedge demands.

When we impose short-sell constraints a second mechanism explains the similarity in performance. Portfolio weights are often extreme because of error maximization if weights are unrestricted and are therefore often equal to upper and lower bound constraints if weight restrictions are imposed. Mean reversion induces a positive hedge demand for equity. But when the myopic portfolio is already 100% of wealth, increasing the share of equity in the portfolio beyond 100% for hedging purposes is infeasible. Analogously, when myopic bond demand is zero, negative hedge demands are infeasible, and therefore the myopic and dynamic portfolios will coincide.

Our choices in the design of the out-of-sample tests imply a few aspects that are hard to address. We do not incorporate hedge demands due to learning in our setting with three risky asset classes. Our investors learn about the parameters, since we update parameter estimates every period. But parameter updates are also changes in the investment opportunity set, and therefore a long-term investor may want to hedge against future parameter updates. Such hedge demands are studied in Barberis (2000), Xia (2001) and Brandt et al. (2005). In a vector autoregressive model with multiple asset classes it is, however, computationally infeasible to compute the hedge demands from learning.[5] This means that we can only account for learning in a very parsimoniously parameterized model with two assets and a single predictor variable. In our empirical work we evaluate such models and find that also in this case the hedging demands do not add value. The learning effects are limited in our case, since, as noted by Skoulakis (2008), the benefits of hedging against learning are small when the investment horizon is much smaller than the estimation sample. In our setup the estimation sample is at least 20 years. In addition, the results in Xia (2001) for a model with equity as the only risky asset suggest that optimal hedge demands for learning may offset the hedge demands from mean reversion of equity returns, which would imply that the myopic allocation is very close to the true optimal strategic asset allocation.

A second potential limitation is our assumption of homoskedastic returns. This is the common assumption in the strategic asset allocation literature following Chacko and Viceira (2005). In their model time-varying risk is an important feature of equity returns, but it does not generate large hedge demands, i.e., the effect on myopic and dynamic portfolios is similar.

For robustness, we explore various alternatives: 10- and 15-year investment horizons, an alternative predictor, an extended data set, rolling window versus expanding window estimation and alternative specifications with smaller asset menus. Many of the robustness checks are available in the online appendix.

2. Methodology

2.1. General Setup

Define the n-vector y_t as

$$y_t' = \begin{pmatrix} r_t & x_t' & s_t' \end{pmatrix}, \tag{1}$$

where r_t is the real return on a nominal T-bill, x_t is the vector of excess returns on stocks $x_{s,t}$ and bonds $x_{b,t}$,

[5] The computational problem is well recognized in the cited references and arises because the number of state variables grows quadratically with the dimension of the VAR. Johannes et al. (2014) report that the intertemporal optimization problem is still beyond current computing technology.

and s_t is a vector of predictor variables. As state variables we use the nominal three-month T-bill yield, the price–earnings (PE) ratio, and the yield spread between long-term bonds and the nominal T-bill. Hence, $n = 6$.

We consider investors who choose portfolios of the three asset classes to maximize expected utility over terminal wealth W_{t+K} with an investment horizon of K periods. Formally, the investor chooses portfolio weights $w_t, \ldots, w_{t+K-1}$ such that the value function at time t maximizes the objective

$$V(K, y_t, W_t) = \max_{\{w_t,\ldots,w_{t+K-1}\}} \mathrm{E}[U(W_{t+K}) \mid y_t], \qquad (2)$$

where final wealth from the portfolio strategy equals

$$\frac{W_{t+K}}{W_t} = \prod_{s=t}^{t+K-1} w_s' R_{s+1}, \qquad (3)$$

with $R_t = e^{r_t}(1, e^{x_{s,t}}, e^{x_{b,t}})'$ the vector of gross returns on the assets in period t. Portfolio weights add up to 1. In the optimization we normalize initial wealth to $W_t = 1$. We assume power utility preferences $U(x) = x^{1-\gamma}/(1-\gamma)$ with relative risk aversion $\gamma > 1$.

An econometric model is needed to evaluate the conditional expectation in (2). Following Campbell et al. (2003), Jurek and Viceira (2011) and many others, the dynamics of asset returns and state variables are assumed to follow the first order vector autoregression

$$y_{t+1} = B_0 + B_1 y_t + \epsilon_{t+1}, \qquad (4)$$

where B_0 is an n-vector of intercepts, B_1 an $(n \times n)$ matrix of slope coefficients, and ϵ_{t+1} an n-vector of errors for which we make the common assumption that they are normally distributed with mean zero and covariance matrix Σ.[6] We collect all parameters in the vector θ. The investor estimates the parameters using different Bayesian priors that we explain in §2.5.

We compare two types of strategies: a dynamic strategy and a myopic strategy. The myopic strategy is the optimal solution for $K = 1$. The dynamic strategy is the optimal solution to a long-horizon problem with $K > 1$. The hedge demand at time t is defined as the difference between the dynamic and the myopic strategy. Our main question is whether investors should take the hedge component into account. To answer this question we compare the performance of the dynamic strategy with the repeated myopic strategy out-of-sample. If we would know the process that generates asset returns and state variables perfectly and had an infinitely long sample, this would be a trivial question to answer. A dynamic strategy would be superior to repeated myopic strategies, since the former strategy encompasses the latter (for the same investment horizon). Since we do not know the true data generating process, the model we select for portfolio construction will be subject to misspecification and parameter estimation error. For the myopic portfolio, the errors are only related to estimation error in the single-period expected returns. The hedge component, however, is also sensitive to the long-run predictions of returns and their covariance with current returns. Out-of-sample, it is therefore far from trivial which strategy works best.

In the out-of-sample analysis, our first investor has an investment horizon of K periods and uses all data available until period T_0 to estimate the model and choose her initial portfolio weights w_{T_0}. In the next period, $T_0 + 1$, her investment horizon is $K - 1$, and she updates her information set to choose portfolio weights w_{T_0+1}, etc. In period $T_0 + K - 1$, her investment horizon is 1 period, and she uses all data until that period to choose her last portfolio weights w_{T_0+K-1}. This sequence of K portfolio weights results in exactly one terminal wealth value W_{T_0+K}. The next investor follows a similar strategy but she starts in period $T_0 + 1$ and ends in period $T_0 + 1 + K$ with again exactly one terminal wealth value W_{T_0+1+K}. We repeat this analysis for many investors, all with horizon K, who start their strategies one period after each other. The last investor starts in $T - K$ and ends in T, the end of the sample. In this way, we obtain time series of terminal wealth values W_{t+K} and realized utility values $U(W_{t+K})$. These provide a measure of out-of-sample performance of the investors, using only information that is available to investors in real time. We use the sample means of realized utilities to analyze whether dynamic and repeated myopic strategies provide the same expected utility.

For reporting purposes, we express average realized utility as an annual certainty equivalent return (CER). The CER is the risk-free return that would make investors indifferent between following a strategy or accepting this risk-free real return. The CER is a monotone transformation of the average out-of-sample realized utility and is given by

$$1 + \mathrm{CER} = ((1-\gamma)\bar{U})^{1/(1-\gamma)}, \qquad (5)$$

with $\bar{U} = (1/(T - K - T_0 + 1)) \sum_{t=T_0}^{T-K} (W_{t+K}^{1-\gamma}/(1-\gamma))$. With monthly data, we report the annualized certainty equivalent returns $(1 + \mathrm{CER})^{12/K} - 1$.

We choose our starting date T_0 such that we have 20 years of initial observations to estimate a model. For our main results, the investment horizon is five years with a monthly decision interval; hence

[6] The VAR(1) is restrictive in two ways. First, it is unlikely that all dynamics in the data are accounted for by using only one lag. Second, it is unlikely that the error term is homoskedastic, i.e., that risk is constant over time. There is a trade-off between model size and estimation error. Much of the forecasting literature prefers smaller models and therefore we stick to the VAR(1).

$K = 60$. The investment horizon of 60 months is a medium to long-term horizon. With monthly postwar U.S. data, it gives us almost eight nonoverlapping out-of-sample investment periods.

Every month we allow investors to use all available information up to this month to update their portfolio holdings. This means that we reestimate our models every month to include the newest observations using an expanding data window. The period t investor makes a plan at t, but reoptimizes it using new parameter values at $t+1$.

We compare the myopic and dynamic strategies for a number of alternatives regarding estimation and portfolio optimization:

- three levels of risk aversion, $\gamma \in \{2, 5, 10\}$;
- *unrestricted* portfolio weights or weights that are *restricted* by short-sell constraints;
- optimization conditional on parameter estimates (*plug-in*) or incorporating Bayesian parameter uncertainty (*decision theoretic*);
- three different priors: *uniform*, *shrinkage* or *no predictability*.

We perform the out-of-sample analysis for three levels of risk aversion, but will focus our discussion mainly on $\gamma = 5$. An investor with $\gamma = 2$ is fairly aggressive and hedge demands for both stocks and bonds are small. For $\gamma = 10$, the hedge demand for stocks is small, since such investors hardly invest in stocks.

As a benchmark for all strategies we consider four model-free strategies. Three strategies invest fully in a single asset, either stocks, bonds, or T-bills. The fourth is the $1/N$ portfolio, which invests an equal amount in every asset.

2.2. Unrestricted Weights

Solving the optimization problem (2) requires numerical techniques. Even if all returns in R_{t+1} are lognormally distributed (conditional on the parameters and the states y_t), the portfolio return $w_t' R_{t+1}$ in (3) is not lognormal, and the expectation in (2) is not available in closed form. To obtain an approximate analytical solution, a large part of the literature solves a slightly modified optimization problem. Some studies (e.g., Larsen and Munk 2012, Sangvinatsos and Wachter 2005) consider a continuous time version, whereas others (e.g., Campbell et al. 2003, Jurek and Viceira 2011) take a loglinear approximation to the wealth evolution (3). Both approximations lead to solutions that allow for short selling and leverage, and indeed applications often exhibit portfolios with negative weights for one or more asset classes. In discrete time negative portfolio weights can, however, lead to negative wealth and hence a realized utility of minus infinity. Hence negative weights are formally inadmissable in discrete time under CRRA utility and lognormally distributed returns.[7] Even so, many papers (e.g., Brennan et al. 1997, Campbell et al. 2003, Jurek and Viceira 2011) still consider short selling in a discrete time setting. With real data, wealth will remain positive for moderate levels of leverage as long as actual excess returns x_s and x_b are believed to be bounded from becoming too negative, i.e., if the actual distribution differs from the assumed lognormal. The more extreme the portfolio weights are, the larger the risk is that wealth becomes negative. The probability that wealth does become negative is deemed to be small enough that it can be ignored in most applications.[8] Since strategies with unrestricted weights are so common, we evaluate the out-of-sample performance of the approximate analytical solutions using the "plug-in" method under the label *unrestricted weights*. These results should be interpreted with some care, though.

2.3. Plug-In vs. Decision-Theoretic Method

In the portfolio choice literature, there are two methods to use the estimation results. The plug-in method substitutes parameter estimates for the true parameters, ignoring any form of parameter uncertainty. This approach is adopted by, e.g., Campbell and Viceira (2002), Jurek and Viceira (2011) and Koijen et al. (2009). In this case the conditional distribution of future values y_{t+1} for asset returns and state variables given their current values is

$$p(y_{t+1} \mid \hat{\theta}, y_t) = \mathrm{N}(\hat{B}_0 + \hat{B}_1 y_t, \hat{\Sigma}), \tag{6}$$

where $\hat{B}_0$, $\hat{B}_1$, and $\hat{\Sigma}$ are estimates for the unknown parameters in θ based on a sample Y_t of observations up to period t. The current values of asset returns and state variables summarize the conditioning space (next to the parameter estimates). For point estimates, we use posterior means.

The alternative is the decision-theoretic method, which takes parameter uncertainty into account through the predictive density. References for this method are Barberis (2000) and Brandt et al. (2005). The predictive density for asset returns and state variables, given the current state y_t, is

$$p(y_{t+1} \mid Y_t, y_t) = \int p(y_{t+1} \mid \theta, y_t) p(\theta \mid Y_t)\, d\theta, \tag{7}$$

[7] The problem is well known in the literature. For a textbook discussion, see Campbell and Viceira (2002, pp. 28–29) or Brandt (2010, pp. 283–284). The problem does not occur in continuous time when the investor can continuously rebalance.

[8] Brandt (2010) uses the example of a leveraged portfolio with 220% in equity and −120% in a risk-free asset. For this highly leveraged portfolio, he estimates the probability of obtaining negative wealth in one quarter to be of the order 10^{-9}. Branger et al. (2010) analyze the naive implementation of a continuous time strategy in discrete time and conclude that it is viable as long as derivatives are not part of the asset menu.

where $p(\theta \mid Y_t)$ is the posterior density of the parameters. Parameter uncertainty is taken into account by integrating the conditional density $p(y_{t+1} \mid \theta, y_t)$ with respect to the posterior distribution of the parameters. The predictive distribution of returns in (7) is not lognormal anymore, but will usually have fatter tails than the normal distribution. Since all assets, even the T-bill, are risky in real terms, the expected utility of all portfolio strategies is minus infinity unless we make a slight modification. We impose a lower bound of -20% on the monthly real return on the T-bill, which guarantees that portfolios with a positive investment in the T-bill and nonnegative weights for the other asset classes have finite expected utility under the decision-theoretic approach. In all numerical optimizations, we impose these short-sell constraints.

When making long-term decisions, the investor can anticipate that she learns more about parameter values as time passes. For our main results, we ignore such effects since it is computationally infeasible to account for them given the size of the problem. We follow Barberis (2000) and assume that investors take parameter uncertainty into account, but ignore the impact of future learning on today's asset allocation. Under this assumption, y_t summarizes the conditioning space at time t (next to the time t posterior distribution). In a parsimoniously parameterized model with only one risky asset that is predictable by only one predictor variable, we are able to analyze the impact of learning.

2.4. Solution Method

Except for the approximate analytical solutions for the plug-in methods with unrestricted weights, all portfolio strategies require numerical optimization. Since we recalculate dynamic strategies many times, computation time is an important issue for the numerical methods. Therefore, we propose a refinement of the method of Brandt et al. (2005) and van Binsbergen and Brandt (2007). Relying on an important observation made by Koijen et al. (2010), we parameterize the regression coefficients in regressions that approximate conditional utility by a quadratic function of portfolio weights. This local quadratic approximation of the conditional expectation of the objective function speeds up calculations and makes our out-of-sample analysis feasible. Technical details are reported in the online appendix.

Our numerical solutions solve the discrete time problem with monthly decision intervals and never exhibit negative weights. We report these numerical solutions under the label *restricted weights*.

2.5. Priors and Estimation

We use Bayesian methods to estimate the VAR in (4). Our first prior is a uniform prior on $\beta = \mathrm{vec}([B_0, B_1]')$ and a Jeffreys prior on Σ,

$$p(\beta, \Sigma) \propto |\Sigma|^{-(n+1)/2} I(B_1), \tag{8}$$

where $I(B_1)$ is an indicator function that is 1 if the matrix B_1 implies a stationary model and 0 otherwise. It is the most commonly used prior for VARs. The properties of its posterior distribution are well known (see, e.g., Zellner 1971). The fact that the predictor variables are close to having a unit root does not pose a problem for inference, because (possible) nonstationarity does not require specific Bayesian methods (see Sims and Uhlig 1991). Nevertheless, since it is common in the strategic asset allocation literature to impose the assumption of stationarity (see, e.g., Campbell and Viceira 2002, Stambaugh 1999), we impose this assumption.

When many parameters need to be estimated, as in a VAR, a uniform prior may result in very noisy results. To reduce estimation noise, a shrinkage prior is commonly adopted. In the absence of much data information such a prior pulls the posterior density toward a baseline value β_0 for the parameters. As our second prior we therefore consider the shrinkage prior specifically proposed in the context of a VAR by Ni and Sun (2003):

$$p(\beta, \Sigma) \propto ((\beta - \beta_0)'(\beta - \beta_0))^{(-(n(n+1)-2))/2} \cdot |\Sigma|^{-(n+1)/2} I(B_1). \tag{9}$$

The prior itself is not proper, but Ni and Sun (2003) show that the posterior is proper in a VAR. The negative exponent on the distance between β and β_0 shrinks the coefficients toward β_0. We set all elements of β_0 equal to zero except for the three diagonal elements in B_1 that correspond to the first own lag of the state variables s_t. Since the predictor variables s_t exhibit near unit root behavior, we set these elements in β_0 equal to one. The prior on Σ is the noninformative Jeffreys prior.

As in Wachter and Warusawitharana (2009), the shrinkage prior reflects the beliefs of an investor who is skeptical about predictability of asset returns. Such an investor downplays the predictability that is found in the data, but does not dogmatically ignore predictability. If there is sufficient evidence in the data that asset returns are predictable, this investor will take (some) asset return predictability into account. Shrinking the first lags of near unit root variables to one is a well-known procedure dating back to the Minnesota prior of Doan et al. (1984). This means that we shrink the dynamics of the three state variables toward a random walk.

As a third prior we dogmatically impose that excess returns of stocks and bonds are not predictable. In this case the second and third rows of B_1 are both equal to zero. In the prior we restrict these parameters to zero and use the uniform prior for all other elements of β. We refer to this specification as the no-predictability prior. This prior serves as a benchmark to evaluate the utility value of predictability. Investors that follow these

rules do not actively time the stock and bond markets. The dynamic and myopic strategies still differ, however, since the expected real T-bill rate is assumed to vary over time. To limit the total number of specifications, we only combine the no-predictability prior with the plug-in method.

For estimation, we use the conditional likelihood function that conditions on the first observation. Combining the likelihood with each of the priors we obtain the posterior densities. Drawing from the posterior density is fairly standard. We therefore refer to the online appendix for technical details on the posterior distributions and the Markov chain Monte Carlo algorithm.

2.6. Additional Specifications

We consider several alternative specifications to test the robustness of our results. First, we consider simplified models in which an investor can only invest in the T-bill rate and in one risky asset (either stocks or bonds). The risky asset is only predictable by one predictor variable (either the price–earnings ratio or the yield spread) using a VAR(1) with only the lagged predictor as right-hand-side variable, whereas the T-bill rate is set equal to its average in the estimation window. Such a tighter specification increases estimation precision at the cost of potential model misspecification. This setting is similar to the model in Barberis (2000) and Brandt et al. (2005) and allows us to analyze the impact of learning.[9] Second, we consider strategies with investment horizons of 120 and 180 months. Since one would expect that hedging is more important for long horizons, these specifications should give additional insight in the importance of the hedge demand. The downside is that we now have fewer than four nonoverlapping investments periods such that the probability that results are driven by chance is much larger. Third, we replace the price–earnings ratio by the dividend–price (DP) ratio to check whether our results are sensitive to a slightly different model specification. Fourth, we consider an extended data set that starts in 1927. However, we should keep in mind that bond returns and T-bill rates are very different in the early data. Finally, we estimate our models using a rolling window of 20 years instead of an expanding window. This method should allow us more easily to incorporate time-variation in model parameters, but could lead to increased estimation error.

[9] Since the lagged asset return itself is not a predictor in the restricted VAR(1), there are only seven unique parameters in total: two intercepts, two slope coefficients on the lagged predictor variable, two error variances, and one error correlation. This leads to eight conditioning variables in total, since the current value of the predictor variable is also a conditioning variable. This makes the incorporation of the learning hedge term feasible. The online appendix explains how we define the conditioning variables.

The parsimonious model is analyzed in §4.3, and the other alternative specifications are considered in §4.4 and the online appendix.

3. Data and Estimation Results

Our empirical analysis is based on monthly data for the U.S. stock and bond market for the period February 1954 to December 2012. The first variable is the ex post real T-bill rate, which is the difference between the log return (or lagged yield) on the nominal three-month T-bill obtained from the Federal Reserve Economic Data (FRED) website,[10] and log inflation, obtained from the Center for Research in Security Prices (CRSP). The second variable is the excess log stock return, defined as the difference between the value weighted log total return on the NYSE, NASDAQ, and AMEX markets and the log return on the three-month T-bill. The third variable, the excess log bond return, is defined in a similar way, but it uses the five-year bond return from CRSP.

To make our out-of-sample analysis credible, we choose three predictor variables that have been used as predictors for decades. We use the nominal yield, the yield spread, and the price–earnings ratio. The log yield spread is defined as the difference between the log yield on a five-year bond obtained from the FRED site and the log yield on the 90-day T-bill. The log of the price–earnings ratio is obtained from the *Irrational Exuberance* data, available from the website of Professor Robert Shiller.[11] It is defined as the log of the ratio of the current price over the lagged mean of earnings over the past 10 years.

Pesaran and Timmermann (1995) discuss the history of stock return predictors, whereas Fama (1976) and Fama and Schwert (1977) have references to early predictors of bond returns. The asset return and predictor variables are also commonly used in the strategic asset allocation literature; see, e.g., Campbell et al. (2003) and Jurek and Viceira (2011). Table 1 provides summary statistics of our monthly data.

In the robustness section we analyze an extended data set starting in February 1927. We use the shorter data set in our main analysis for two reasons. First, the Federal Reserve kept the short-term nominal rate basically fixed before the Treasury accord of 1952, which makes it very hard to model interest rates accurately. Second, because of a lack of suitable bond issues before 1941, the early bond return and bond yield data is less reliable.

Table 2 reports posterior moments for B and Σ for the shrinkage prior. We obtain the standard result that the state variables are highly autocorrelated. The nominal

[10] See http://research.stlouisfed.org/fred2.

[11] See http://www.econ.yale.edu/~shiller/data.htm.

Table 1 Summary Statistics

	r	x_s	x_b	s_{nom}	s_{PE}	s_{spr}
Avg.	0.0008	0.0045	0.0014	0.0460	2.8762	0.0113
Std. dev.	0.0032	0.0442	0.0147	0.0280	0.4003	0.0089
Min	−0.0112	−0.2607	−0.0692	0.0001	1.8929	−0.0160
Max	0.0197	0.1483	0.0898	0.1443	3.7887	0.0421
AR(1)	0.4424	0.1003	0.1149	0.9889	0.9954	0.9197

Notes. This table reports the average, standard deviation, minimum, maximum. and first order autocorrelation for the ex post real return on the nominal T-bill (r), the excess stock return (x_s), the excess bond return (x_b), the nominal yield (s_{nom}), the price–earnings ratio (s_{PE}), and the yield spread (s_{spr}). The data set starts in February 1954 and ends in December 2012. Returns are in monthly units.

yield and the price–earnings ratio have a negative effect on stock return predictions, whereas the yield spread has a positive impact on expected bond returns. There is also a large positive correlation between shocks to the price–earnings ratio and excess stock returns, which means that unexpected positive shocks to stock returns lead to negative future investment opportunities. This result implies that there is mean reversion in stock returns. The predictability of the excess returns is the main motivation for considering hedge demands in the literature.

Table 2 Parameter Estimates—Shrinkage Prior

Parameter estimates

	r	x_s	x_b	s_{nom}	s_{PE}	s_{spr}	R^2
r	0.3432	−0.0037	0.0129	0.0280	0.0010	0.0399	0.2298
	0.0347	*0.0024*	*0.0075*	*0.0052*	*0.0003*	*0.0136*	
x_s	−0.0473	0.0647	0.1333	−0.1662	−0.0125	−0.0055	0.0334
	0.1233	*0.0358*	*0.0844*	*0.0617*	*0.0046*	*0.1027*	
x_b	0.1568	−0.0610	0.0805	0.0232	0.0014	0.2537	0.0805
	0.1061	*0.0121*	*0.0360*	*0.0239*	*0.0017*	*0.0603*	
s_{nom}	−0.0082	0.0123	−0.0615	0.9892	0.0000	0.0302	0.9780
	0.0414	*0.0036*	*0.0109*	*0.0072*	*0.0005*	*0.0190*	
s_{PE}	0.1749	0.4289	0.2045	−0.0557	0.9933	0.1323	0.9946
	0.1204	*0.0246*	*0.0612*	*0.0437*	*0.0032*	*0.0823*	
s_{spr}	−0.0164	−0.0039	−0.0661	0.0036	−0.0003	0.9464	0.8581
	0.0379	*0.0029*	*0.0090*	*0.0061*	*0.0004*	*0.0161*	

Error standard deviations and correlations

	r	x_s	x_b	s_{nom}	s_{PE}	s_{spr}
r	0.2845	0.0612	0.0784	−0.0565	0.1372	0.0309
x_s		4.3853	0.0695	−0.0271	0.7761	−0.0380
x_b			1.4193	−0.6119	0.0214	0.1997
s_{nom}				0.4192	−0.0142	−0.8447
s_{PE}					2.9710	−0.0295
s_{spr}						0.3400

Notes. This table reports estimates for the VAR(1) using the shrinkage prior. Columns 2–7 show the posterior mean and standard deviations (in italic) of the slope coefficients. The last column shows the R^2 implied by the posterior mean. The elements on the diagonal of the error correlation matrix are the standard deviations (×100) of the error terms, the off-diagonal elements are the correlations. The sample period is February 1954–December 2012.

Parameter estimates using the uniform prior are reported in the online appendix. Comparing estimation results, we find that the shrinkage estimator reduces the predictability of asset returns, while at the same time increasing persistence of the state variables. The lower predictability is reflected in lower R^2 values (calculated using the posterior mean of the coefficients) under the shrinkage prior, especially for excess stock returns. The increased persistence is reflected in the increase of the maximum eigenvalue from 0.9917 for the uniform prior to 0.9928 for the shrinkage prior (evaluated at the posterior means).

In the out-of-sample analysis, we estimate our models on data sets of increasing length by including the newest observations. Figure 1 presents time-series plots of the estimates of the important predictive coefficients (x_s, s_{PE}) and (x_b, s_{spr}) for both priors. The figure shows that there is a lot of uncertainty about the posterior parameter values, since parameters vary considerably over time and posterior probability intervals are large. However, the estimated values for the shrinkage estimator are less variable and closer to 0. For equity, predictability becomes weaker over time, whereas for bonds there is stronger predictability over time. The effect of the shrinkage prior diminishes over time, since the likelihood dominates when sample size grows.

4. Out-of-Sample Performance

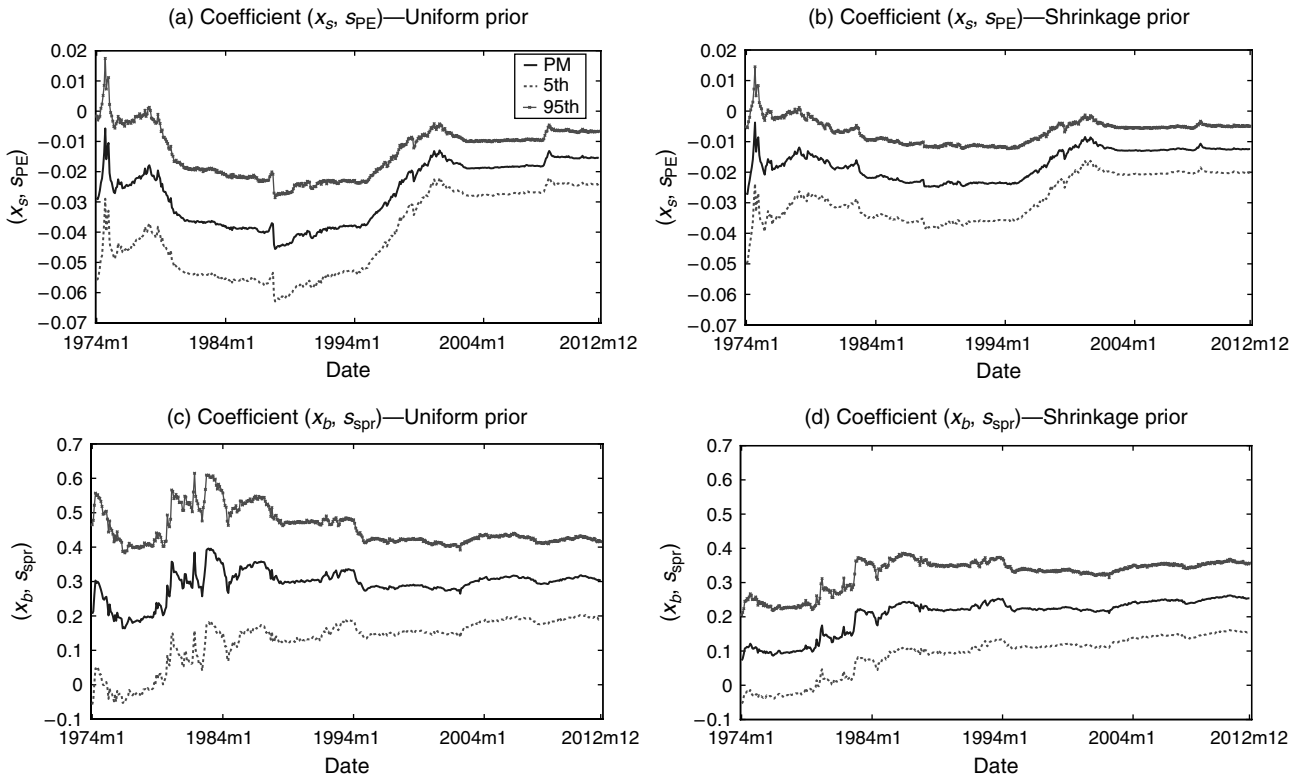
4.1. Main Results

The starting date for our out-of-sample experiment is January 1974. Table 3 reports benchmark results for model-free strategies. The table shows that moderate ($\gamma = 5$) and conservative investors ($\gamma = 10$) would have preferred the $1/N$ portfolio, whereas more aggressive investors ($\gamma = 2$) have a preference for full investment in the risky stock market. We use these specifications as benchmarks for the model-dependent specifications.

Table 4 presents results for the model-dependent specifications. Each line of the table compares the performance of the dynamic strategy in a particular specification to the repeated myopic strategy in the same specification. The specifications differ by weight restriction (long only or unrestricted), method (plug-in or decision-theoretic method), prior distribution (no-predictability, uniform, or shrinkage prior), and risk preferences ($\gamma \in \{2, 5, 10\}$).

Most striking in the results is the similar economic performance of dynamic and myopic strategies. Differences in prior and optimization technique lead to much larger performance differences than different strategies. In most cases the hedge component hardly adds value, and in some cases even deteriorates the certainty equivalent return. For the uniform prior, the repeated myopic strategies often outperform dynamic strategies. For example, for $\gamma = 5$, the hedge component leads

Figure 1 Time Series for Predictive Coefficients



Notes. This figure plots time series for the mean and the 5th and 95th percentiles of the posterior distribution of the coefficients (x_s, s_{PE}) and (x_b, s_{spr}) in the VAR for the uniform and shrinkage priors. The model is estimated from February 1954 until the date on the x-axis.

to a loss in CER of 0.24% per year for a specification that uses restricted weights and the plug-in method. The hedge component performs better if estimation error is controlled using the shrinkage prior or if one uses the no-predictability prior. But even here the gains are small: with $\gamma = 5$, the largest gain with restricted weights is 0.15% CER (shrinkage prior with plug-in method). It appears that hedging changes in future investment opportunities for stock and bond returns can destroy value. Even though the average terminal wealth is always higher for dynamic strategies, the risk more than proportionally increases.

The lack of value of the hedge portfolio is not due to a general failure of the predictive models, since most model-based strategies outperform the model-free benchmarks. Even long-term investors who do not believe in predictability should select a diversified portfolio based on an estimated model without predictability instead of following simple rules of thumb. It is only for $\gamma = 2$ and $\gamma = 10$ that some model-based strategies underperform compared to model-free benchmarks. For $\gamma = 2$, a portfolio that fully invests in stocks beats the restricted no-predictability strategies, whereas for $\gamma = 10$ the repeated myopic strategy underperforms compared to the $1/N$ strategy. The latter finding is consistent with results of DeMiguel et al. (2009), who find that the $1/N$ strategy is tough to beat out-of-sample. Investors who believe in predictability outperform investors that ignore predictability for nearly all specifications. It pays off to time the stock and bond markets. The only exception is an aggressive investor ($\gamma = 2$) who uses unrestricted weights, the uniform prior, and the plug-in method. With all modeling options at their most aggressive setting, the strategy sometimes leads to negative wealth and hence a CER of −100%.

The shrinkage prior boosts performance for all types of investors because it leads to less aggressive portfolios. The use of the shrinkage estimator reduces the standard deviation of terminal wealth compared to the uniform prior, which more than offsets the (for some specifications) reduction in average terminal wealth.

Table 3 Benchmark Results

	$\gamma = 2$ CER	$\gamma = 5$ CER	$\gamma = 10$ CER	$\overline{TW}$	$\sigma(TW)$
$1/N$	0.0410	0.0337	0.0243	1.2539	0.2042
T-bill	0.0130	0.0108	0.0074	1.0750	0.0935
Stock	0.0586	0.0309	0.0018	1.4699	0.4749
Bond	0.0353	0.0237	0.0034	1.2295	0.2217

Notes. This table reports the annualized CER, average terminal wealth ($\overline{TW}$), and the standard deviation of terminal wealth ($\sigma(TW)$) for three different risk-aversion levels γ for four strategies: either the $1/N$ strategy, or else full investment in the T-bill, the stock index, or the government bond. The out-of-sample period is February 1974–December 2012.

Table 4 Dynamic vs. Myopic Strategies—Main Data Set

		Dynamic			Repeated myopic		
		CER	$\overline{TW}$	$\sigma(TW)$	CER	$\overline{TW}$	$\sigma(TW)$
$\gamma = 2$							
Restricted							
PlugIn	Nopr	0.0578	1.4673	0.4774	0.0580	1.4680	0.4769
PlugIn	Unif	0.0714	1.5132	0.4454	0.0714	1.5068	0.4297
PlugIn	Shrink	0.0781	1.5312	0.3751	0.0779	1.5249	0.3624
DecTh	Unif	0.0712	1.5063	0.4330	0.0710	1.5021	0.4237
DecTh	Shrink	0.0773	1.5210	0.3611	0.0773	1.5199	0.3584
Unrestricted							
PlugIn	Nopr	0.0876	1.7692	0.7635	0.0846	1.7516	0.7713
PlugIn	Unif	−1	21.5171	70.6508	−1	19.0981	59.6174
PlugIn	Shrink	0.3066	12.2410	11.5037	0.3086	12.3057	13.2611
$\gamma = 5$							
Restricted							
PlugIn	Nopr	0.0397	1.3211	0.2538	0.0390	1.3207	0.2648
PlugIn	Unif	0.0535	1.4747	0.4093	0.0559	1.4651	0.3797
PlugIn	Shrink	0.0651	1.4936	0.3323	0.0636	1.4695	0.3094
DecTh	Unif	0.0542	1.4707	0.4025	0.0556	1.4593	0.3732
DecTh	Shrink	0.0642	1.4845	0.3309	0.0631	1.4647	0.3062
Unrestricted							
PlugIn	Nopr	0.0456	1.3445	0.2390	0.0423	1.3238	0.2439
PlugIn	Unif	0.0794	5.9031	6.5341	0.0865	4.6443	4.3208
PlugIn	Shrink	0.1233	3.6141	1.9235	0.1214	3.2918	1.7511
$\gamma = 10$							
Restricted							
PlugIn	Nopr	0.0258	1.2192	0.1442	0.0241	1.2047	0.1488
PlugIn	Unif	0.0428	1.4425	0.3564	0.0418	1.4074	0.3118
PlugIn	Shrink	0.0521	1.4436	0.2898	0.0478	1.3936	0.2442
DecTh	Unif	0.0432	1.4238	0.3260	0.0415	1.4020	0.3061
DecTh	Shrink	0.0491	1.4041	0.2449	0.0474	1.3893	0.2413
Unrestricted							
PlugIn	Nopr	0.0249	1.2146	0.1421	0.0230	1.1940	0.1418
PlugIn	Unif	0.0440	2.8393	1.9480	0.0574	2.3863	1.2507
PlugIn	Shrink	0.0656	2.0585	0.6719	0.0645	1.9046	0.5768

Notes. This table reports the annualized CER, average terminal wealth ($\overline{TW}$), and the standard deviation of terminal wealth ($\sigma(TW)$) for the dynamic and repeated myopic strategies for specifications that differ in the weight restrictions (restricted or unrestricted), the method (plug-in (PlugIn) or decision theoretic (DecTh)), the prior (no-predictability (Nopr), uniform (Unif), or shrinkage (Shrink)), and the risk-aversion level. Results are based on the general VAR(1) estimated using an expanding window. The out-of-sample period runs from February 1974 until December 2012.

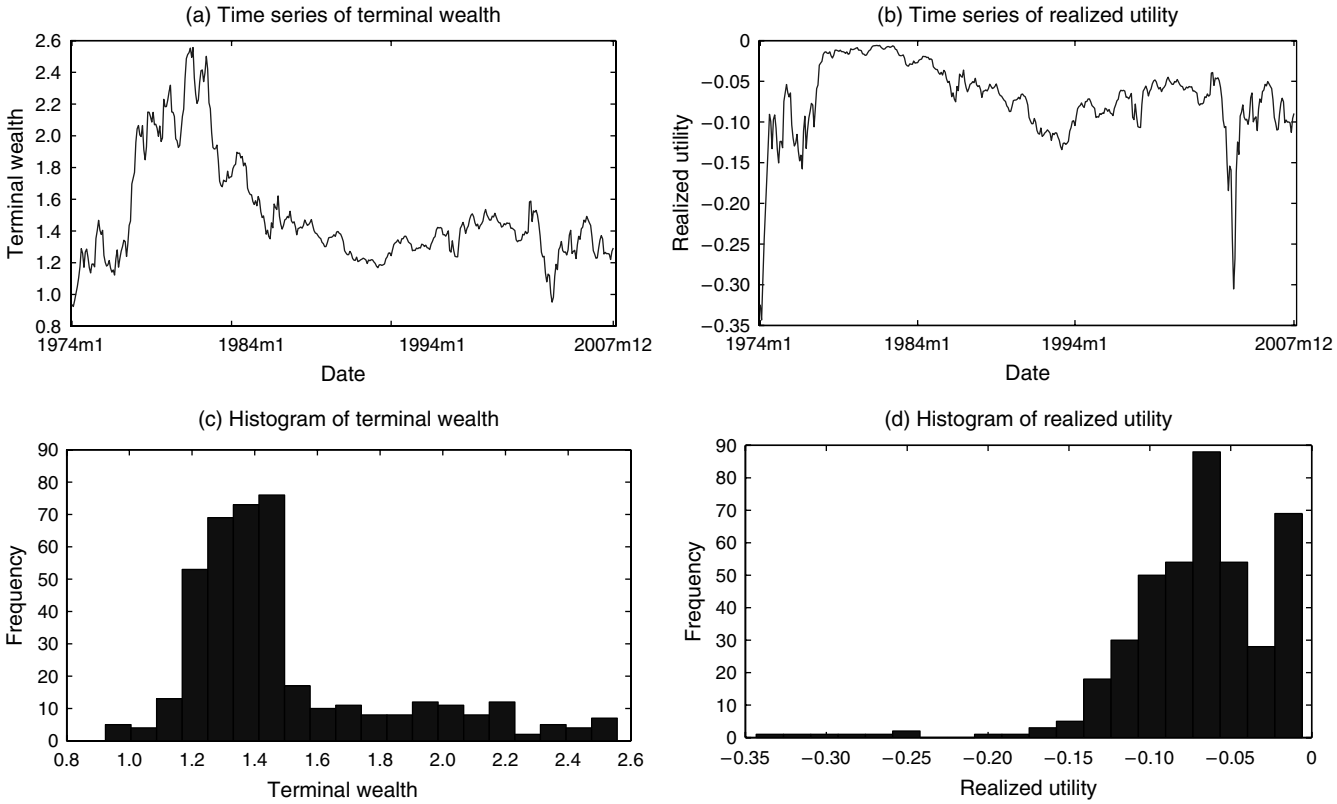
Differences are large for unrestricted weights, which are the most sensitive to estimation risk. However, as indicated above, the difference between dynamic and repeated myopic strategies remains small, because CERs for myopic and dynamic strategies increase together when using shrinkage.

Accounting for parameter uncertainty with the decision-theoretic optimization does not lead to better performance. Brandt et al. (2005) show that accounting for parameter uncertainty in the portfolio optimization problem mainly has an impact on the hedging component of a dynamic strategy. Therefore, it is not surprising that there is no performance difference between the plug-in and decision-theoretic methods, since the hedging component does not have an important impact on performance.

Imposing weight restrictions leads to a substantial reduction in risk and average terminal wealth for investors using either the uniform or shrinkage prior. It avoids CERs of -100% for $\gamma = 2$. In theory, investors with $\gamma = 5$ or $\gamma = 10$ can also become insolvent if short selling is allowed. However, empirically we find that allowing for short selling in discrete time does not lead to negative terminal wealth for such investors, since portfolio weights are not as extreme as the portfolio weights for $\gamma = 2$.

For a more detailed analysis, we focus on the specification with a shrinkage prior, restricted weights, the plug-in method, and moderate risk aversion $\gamma = 5$. Figure 2 plots time series and histograms of terminal wealth and realized utility for the dynamic strategy. Even though the distribution of terminal wealth is somewhat right skewed, the distribution of realized utility is left skewed. Most utility values are near zero, but there are a few negative outliers corresponding to low terminal wealth values. Realized utility is a metric

Figure 2 Performance Realizations

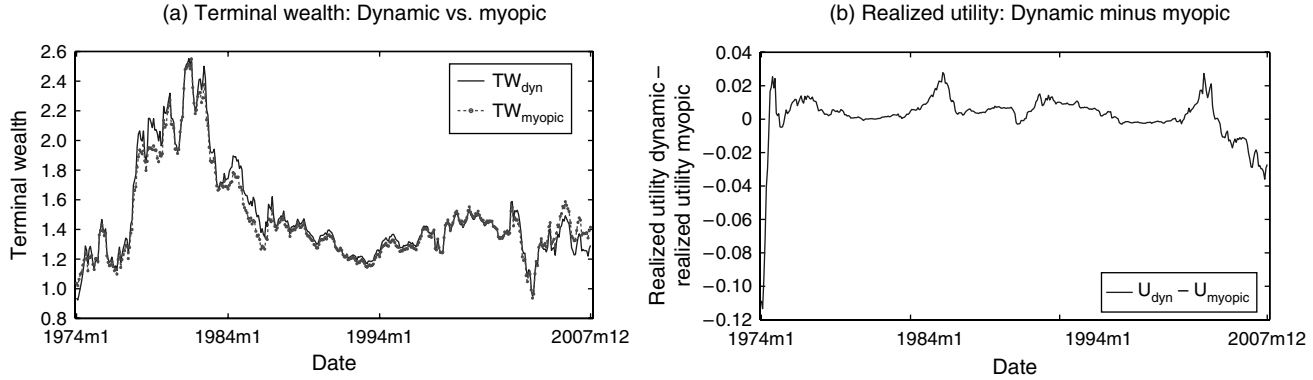


Notes. Panels (a) and (b) plot time series and panels (c) and (d) plot histograms of terminal wealth and realized utility for the restricted dynamic strategy that uses the plug-in method and the shrinkage prior for $\gamma = 5$. Dates refer to the start dates of the investment periods in panels (a) and (b).

that heavily penalizes bad realizations. These outliers strongly affect the CER values in Table 4. This suggests that specifications that limit the number and/or size of negative events are the ones with the highest average realized utility.

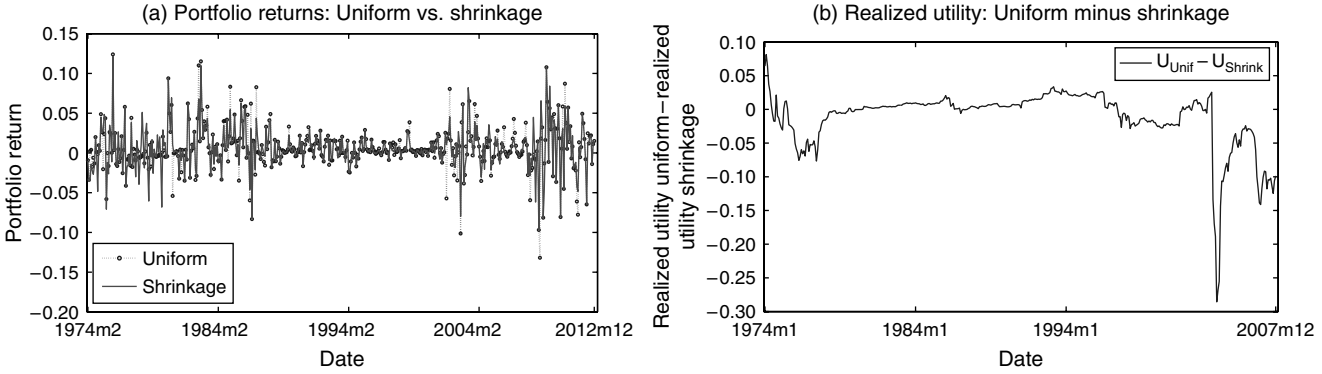
Even though hedging has a small positive effect for this specification, the left panel of Figure 3 shows that the realized terminal wealth values for the dynamic and myopic strategies are very similar. The right panel shows that differences in realized utility are also mostly small. The difference in realized utility can be both positive and negative, which suggests that neither strategy substantially outperforms the other. Since there is much overlap in the evaluation periods, the figures show prolonged periods where one strategy is doing slightly better than the other. The most recent four years of data, all of which include the financial crisis in the out-of-sample period, show negative values

Figure 3 Realized Hedging Performance



Notes. This figure compares the dynamic to the myopic strategy for a specification with the plug-in method, shrinkage prior, restricted weights, and $\gamma = 5$. Panel (a) plots time series of terminal wealth values, whereas panel (b) plots the difference in realized utility values.

Figure 4 Realized Myopic Performance



Notes. This figure compares the uniform to the shrinkage prior for a myopic strategy in a specification with the plug-in method, restricted weights, and $\gamma = 5$. Panel (a) plots the portfolio returns, whereas panel (b) plots the difference in realized utility values. Dates refer to the start dates of the investment periods.

for hedge demands. The dynamic strategy does not succeed in reducing risk by hedging future changes in investment opportunities. It leads to both the highest and lowest terminal wealth values and therefore amplifies (negative) events. Since risk-averse investors dislike negative events, the hedge component hardly adds value.

The shrinkage prior generally provides better performance than the uniform prior. Figure 4 plots the portfolio returns and the difference in realized utility for the uniform and shrinkage priors for the same specification with restricted weights and $\gamma = 5$. The results are shown for the myopic strategy.[12] Portfolio returns and realized utilities are very similar in most periods. However, if there is a large negative event, the shrinkage prior manages to reduce these losses compared to the uniform prior. This is the reason why risk-averse investors value this specification the most. Panel (a) shows the relatively large negative portfolio returns for the uniform prior, and panel (b) reports the corresponding relatively large negative utility differences, which imply outperformance by the shrinkage prior in these particular periods.

4.2. Weights

Table 5 reports portfolio weight statistics for $\gamma = 5$ for the different specifications. Panel A shows results for a dynamic investor with a remaining horizon of 60 months, panel B reports statistics for a myopic investor and, panel C shows results for the hedge demand, defined as the difference between the dynamic portfolio with $K = 60$ and the myopic portfolio.[13]

[12] At every point in time, there are 60 dynamic investors active with different remaining horizons and different strategies, and therefore different portfolio returns. However, the 60 repeated myopic investors choose the same portfolio weight in every period and hence have the same portfolio returns in every period.

[13] Results for other risk-aversion levels can be found in the online appendix.

First consider strategies with restricted weights. The upper and lower bound constraints for stock and bond weights are often binding (except when using the no-predictability prior). For example, for the uniform prior and plug-in optimization, we find that the portfolio weight of stocks is at either the upper or lower limit 76% of the time, whereas the bond weight is at the boundary in 79% of cases. A dynamic investor who uses a uniform prior invests on average 40% in stocks and 38% in bonds. The standard deviation of weights is 45% for both stocks and bonds, meaning that the investor changes the portfolio composition aggressively. The estimated portfolio weights are extremely sensitive to small changes in predictor variables and parameter estimates. The shrinkage prior or the decision-theoretic optimization both lower the variability of weights, but not by much. By construction the no-predictability prior greatly reduces the variability of the stock portfolio weights, since it shuts off all forms of market timing in the stock market. Bond weights still vary a lot under the no-predictability prior because of changes in parameter estimates.

Hedge demands are small, since the weights are often equal to upper or lower bound constraints. The hedge demand for stocks cannot be positive if a myopic investor already invests 100% in the stock market. Similarly, the hedge demand for bonds cannot be negative if the myopic bond weight is 0%. The average hedge demand varies between 4% and 11% for stocks and −4% and −8% for bonds for specifications with predictability. However, if we only consider periods with hedge demands that differ from 0%, the average hedge demand for stocks varies between 7% and 20%, and the average hedge demand for bonds varies between −9% and −20%. The hedge term mainly adds noise though, since performance in general does not improve when hedging is taken into account.

With unrestricted weights, both average weights as well as standard deviations are a lot more extreme

Table 5 Portfolio Weights for $\gamma = 5$

	Restricted weights					Unrestricted weights		
	PlugIn Nopr	PlugIn Unif	PlugIn Shrink	DecTh Unif	DecTh Shrink	PlugIn Nopr	PlugIn Unif	PlugIn Shrink
				Panel A: Dynamic weights				
$\bar{w}_s$	0.4813	0.3993	0.4750	0.3705	0.3959	0.5082	0.3500	0.8048
$\sigma(w_s)$	0.0713	0.4461	0.4202	0.4357	0.3923	0.0485	2.0455	1.2324
$\bar{w}_b$	0.3756	0.3813	0.3422	0.3628	0.3768	0.5862	3.0556	1.8186
$\sigma(w_b)$	0.2402	0.4482	0.4097	0.4375	0.4104	0.6762	4.7704	3.5117
$\rho(w_s, w_b)$	−0.4607	−0.6191	−0.6511	−0.5671	−0.5830	−0.2631	−0.1478	−0.0611
$w_s = UB$	0	0.2869	0.2677	0.2484	0.1734			
$w_s = LB$	0	0.4711	0.2827	0.4861	0.3212			
$w_b = UB$	0	0.2719	0.1606	0.2484	0.1820			
$w_b = LB$	0.2420	0.5225	0.4904	0.5161	0.4240			
				Panel B: Myopic weights				
$\bar{w}_s$	0.5549	0.3334	0.3607	0.3293	0.3572	0.5716	0.1069	0.3992
$\sigma(w_s)$	0.0606	0.4269	0.3893	0.4218	0.3850	0.0507	1.5499	0.6964
$\bar{w}_b$	0.2937	0.4275	0.4218	0.4236	0.4176	0.3974	2.2160	1.7669
$\sigma(w_b)$	0.1998	0.4572	0.4238	0.4520	0.4191	0.4816	4.1714	3.3731
$\rho(w_s, w_b)$	−0.4290	−0.5853	−0.5760	−0.5840	−0.5755	−0.3454	0.1217	−0.0812
$w_s = UB$	0	0.2013	0.1456	0.2034	0.1413			
$w_s = LB$	0	0.5396	0.3683	0.5375	0.3683			
$w_b = UB$	0	0.3212	0.2184	0.3169	0.2184			
$w_b = LB$	0.2163	0.4582	0.3876	0.4604	0.3883			
				Panel C: Hedge weights				
$\bar{w}_s$	−0.0736	0.0659	0.1143	0.0411	0.0387	−0.0634	0.2430	0.4056
$\sigma(w_s)$	0.0385	0.1200	0.1421	0.0794	0.0773	0.0102	0.8169	0.6921
$\bar{w}_b$	0.0819	−0.0462	−0.0796	−0.0609	−0.0409	0.1888	0.8397	0.0517
$\sigma(w_b)$	0.1305	0.1340	0.1652	0.1726	0.1214	0.2760	2.2333	1.1404
$\rho(w_s, w_b)$	−0.2901	−0.8350	−0.8307	−0.4054	−0.4019	−0.0134	−0.4638	−0.5284
$\bar{w}_s \neq 0$	−0.0736	0.2013	0.1999	0.1325	0.0705			
$\bar{w}_b \neq 0$	0.0983	−0.1599	−0.1667	−0.2017	−0.0892			

Notes. This table reports portfolio weight statistics for the dynamic strategy with a remaining horizon of 60 months (panel A), the myopic strategy (panel B), and the hedge demands with a remaining horizon of 60 months (panel C) for specifications that differ in the weight restrictions (restricted or unrestricted); the method (plug-in (PlugIn) or decision theoretic (DecTh)); and the prior (no-predictability (Nopr), uniform (Unif), or shrinkage (Shrink)) for risk aversion $\gamma = 5$. Statistics $\bar{w}_s$, $\bar{w}_b$, $\sigma(w_s)$, $\sigma(w_b)$, and $\rho(w_s, w_b)$ give the average stock and bond weights, their standard deviations, and the correlation between stock and bond weights in the out-of-sample analysis. Quantities $w_s = UB$, $w_b = UB$, $w_s = LB$, and $w_b = LB$ show the fraction of months that the stock and bond weights are equal to their upper bounds (1 for the plug-in method, 0.99 for the decision-theoretic method) or equal to their lower bounds (0 for both methods). Statistics $\bar{w}_s \neq 0$ and $\bar{w}_b \neq 0$ give the average stock and bond hedge demands when these differ from 0.

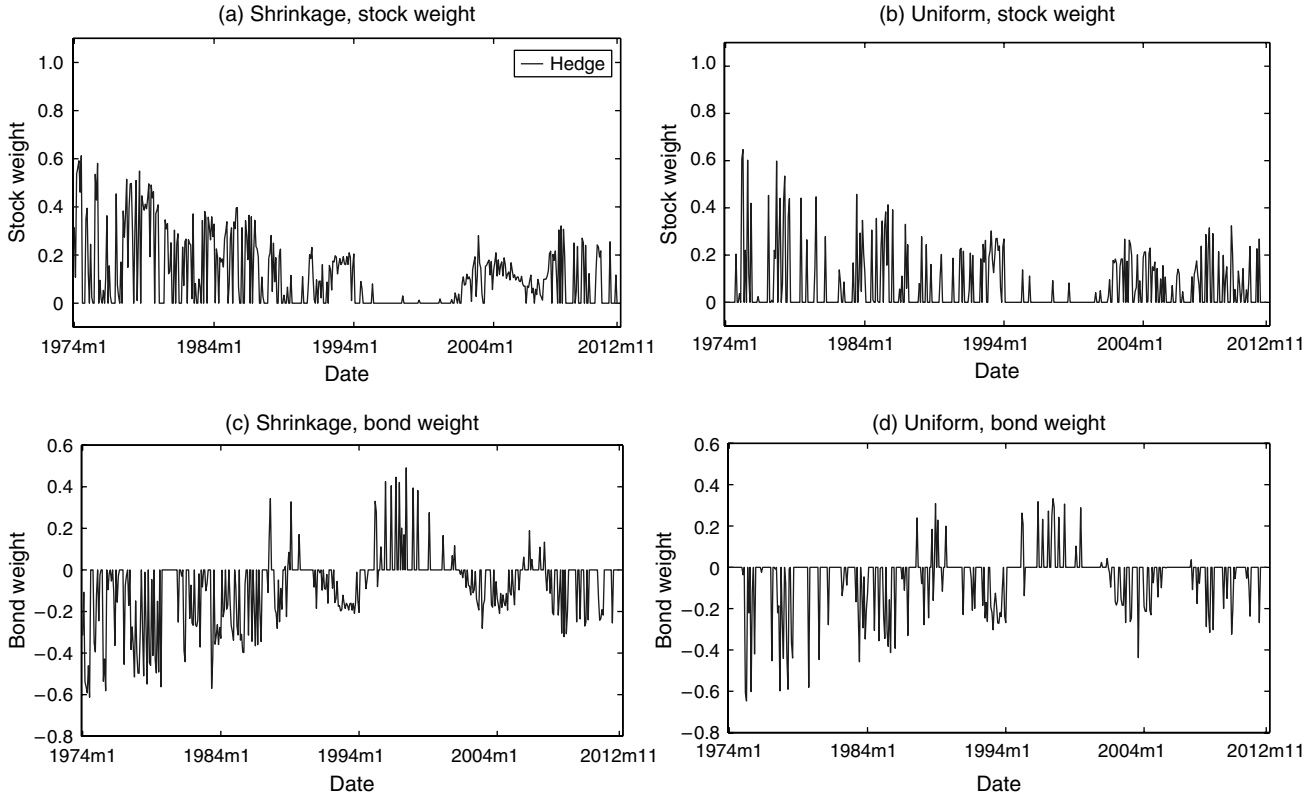
because of error maximization. The unrestricted stock weights have a standard deviation of over 200%, whereas for bonds it is even larger at 477% of wealth. Given these results, it is not surprising that the portfolio weights are often equal to the upper or lower bound constraints when short-sell restrictions are imposed. Hedge demands are large with unrestricted weights, but also very sensitive to the prior. For example, with a uniform prior, the average hedge demand for bonds is 84%, whereas it is just 5% under the shrinkage prior. The unrestricted hedge demand for stocks is always larger than for the restricted weights. These demands make dynamic strategies riskier than myopic strategies due to the large standard deviation of hedge demands.

Figure 5 plots time series of hedge demands for specifications with restricted weights, the plug-in method, $\gamma = 5$, and using either a uniform or shrinkage prior. The figure shows that the hedge demand for stocks is, in general, positive (never negative), whereas the hedge demand for bonds is, in general, negative, as expected (see, e.g., Campbell et al. 2003). Even though the hedge demands can be as large as 60% for stocks and as small as −60% for bonds, the hedge demands are often 0% and change a lot over time. This illustrates that estimated portfolio weights are very sensitive to changes in predictor variables and parameter estimates, and that myopic weights are often equal to their upper and lower bounds. Nevertheless, no matter whether hedge demands are only different from zero when weights are not too extreme or are essentially unrestricted, the hedge demands hardly add any value and mainly add noise.

4.3. A Parsimonious Model

One possible explanation why the hedge demand does not add value may be the large number of estimated parameters and the resulting estimation noise. Therefore, we consider a simplified setting with an investor

Figure 5 Time Series of Hedge Demands



Notes. The first row plots time series of the stock hedge demand (with remaining horizon $K = 60$) for the shrinkage (panel (a)) and uniform priors (panel (b)), whereas the second row plots time series of the bond hedge demand for the shrinkage (panel (c)) and uniform priors (panel (d)). Results are for specifications with restricted weights, the plug-in method, and $\gamma = 5$.

who can only invest in the stock index and a T-bill which is assumed constant. We also assume that the PE ratio is the only predictor variable. This simplified model structure is analogous to the specification considered by Barberis (2000), who uses the dividend yield instead of the price–earnings ratio. The results are in Table 6.

First, the CERs are a lot lower than for the full model, which indicates that the combined predictability of stocks and bonds is what leads to high CERs.[14] Second, the hedge demands hardly have an impact on performance. Even in a setting with much fewer parameters and less estimation uncertainty, the impact of the hedge demand is minor. Third, the shrinkage prior is not effective here. This is not surprising since shrinkage estimators mainly improve performance when the number of parameters is large. It appears that the shrinkage prior removes almost all predictability of stock returns, which leads to very small hedge demands and CERs that are close to the CERs for the no-predictability strategies.

In this parsimonious model we can also account for the hedge demands from learning about the parameters.

[14] The online appendix shows alternative specifications with T-bills and bonds with similar low CERs.

Table 6 Dynamic Strategies vs. Myopic Strategies—Restricted Model for $\gamma = 5$

		Dynamic strategy			Repeated myopic		
		CER	$\overline{TW}$	$\sigma(TW)$	CER	$\overline{TW}$	$\sigma(TW)$
Restricted							
PlugIn	Nopred	0.0279	1.2845	0.2890	0.0279	1.2845	0.2890
PlugIn	Unif	0.0321	1.2836	0.2899	0.0313	1.2687	0.2751
PlugIn	Shrink	0.0273	1.2396	0.2257	0.0266	1.2352	0.2262
DecTh	Unif	0.0310	1.2629	0.2572	0.0310	1.2648	0.2755
DecTh	Shrink	0.0251	1.2201	0.2180	0.0264	1.2331	0.2251
Learn	Unif	0.0312	1.2655	0.2689	0.0310	1.2648	0.2755
Unrestricted							
PlugIn	Nopred	0.0278	1.2838	0.2891	0.0278	1.2838	0.2891
PlugIn	Unif	0.0079	1.3034	0.3616	0.0285	1.2820	0.3044
PlugIn	Shrink	0.0256	1.2372	0.2319	0.0261	1.2347	0.2277

Notes. This table reports the annualized CER, average terminal wealth ($\overline{TW}$), and standard deviation of terminal wealth ($\sigma(TW)$) for the dynamic and the repeated myopic strategies for specifications that differ in the weight restrictions (restricted or unrestricted), the method (plug-in (PlugIn), decision theoretic (DecTh), or learning (Learn)), and the prior (no-predictability (Nopred), uniform (Unif), or shrinkage (Shrink) prior). The decision-theoretic method incorporates parameter uncertainty but ignores learning, whereas the learning method incorporates both. Results are for $\gamma = 5$ and based on the simplified VAR(1) with the stock return and the price–earnings ratio estimated using an expanding window. The risk-free rate is assumed constant. The out-of-sample period runs from February 1974 until December 2012.

The results show that the hedge demand due to learning does not improve performance. It lowers the investment in the stock market and therefore leads to lower average terminal wealth and lower standard deviations of terminal wealth as expected, but does not improve performance.[15] We also consider alternative specifications in the online appendix that use a rolling window of 20 years instead, since one might expect that learning is more important in settings with smaller estimation windows. Nevertheless, even for these settings we do not find that incorporating learning has a positive impact on out-of-sample performance. We find very similar results for the decision-theoretic specifications that incorporate parameter uncertainty but ignore learning. We conclude that the simplified setting confirms the findings for the general model with multiple risky asset classes.

4.4. Alternative Specifications

Table 7 reports results on a number of robustness checks for $\gamma = 5$. More robustness checks are reported in the online appendix. First, we consider a longer investment horizon of 120 months. We expect that the difference between myopic and dynamic portfolios becomes larger, since longer investment horizons should lead to more important hedge demands. A close look at the average terminal wealth and standard deviation of terminal wealth suggests that this is indeed the case. The difference between myopic and dynamic strategies becomes larger. This does not always translate into a gain in CER though. In pairwise comparisons, the shrinkage estimator leads to higher CERs for dynamic strategies, whereas the myopic strategy is still the best for the uniform prior. For both unrestricted and restricted weights, the best performing strategies are dynamic strategies. For unrestricted weights, the CER of the best dynamic strategy (CER of 16.77%) is only slightly higher than the CER for the best myopic strategy (CER of 16.37%). The online appendix shows additional results for a 15 years horizon with similar conclusions.

Second, we analyze specifications that use DP as a predictor instead of PE. The CERs with DP as the predictor are quite a bit lower than the CERs for our main specification and can even be negative if one combines the uniform prior with unrestricted weights. It appears that the model with DP is less stable and more sensitive to estimation error. The hedge demand does not add much value.

Third, we consider the extended data set. The CERs are a lot lower and sometimes negative (even for the no-predictability strategies). This is most likely caused by model instability, since the early data for the T-bill rate, the bond return, and the yield spread are hardly comparable to more recent data. The performance of the dynamic and myopic strategies are very close to each other for restricted weights, whereas differences are larger (both positive and negative) for unrestricted weights. The online appendix shows that the outperformance of the dynamic strategy for the shrinkage prior and unrestricted weights is unique to $\gamma = 5$.

Fourth, we use a rolling window instead of an expanding window to estimate our model. Remarkably, the hedge component destroys value for specifications with restricted weights, whereas it adds value for

[15] Since we already concluded that the shrinkage estimator does not work well in this setting, we only analyze learning for the uniform prior.

Table 7 Dynamic Strategies vs. Myopic Strategies—Alternative Specifications

		Dynamic strategy			Repeated myopic		
		CER	$\overline{TW}$	$\sigma(TW)$	CER	$\overline{TW}$	$\sigma(TW)$
Panel A: Horizon = 120 months							
Restricted							
PlugIn	Nopred	0.0446	1.8364	0.4343	0.0447	1.8217	0.4145
PlugIn	Unif	0.0626	2.3317	0.8442	0.0653	2.2478	0.7020
PlugIn	Shrink	0.0730	2.3677	0.7174	0.0686	2.2074	0.5877
Unrestricted							
PlugIn	Nopred	0.0551	1.9370	0.4061	0.0488	1.8250	0.3878
PlugIn	Unif	0.1510	48.1012	64.2794	0.1637	20.4874	22.4607
PlugIn	Shrink	0.1677	15.2246	13.3397	0.1577	9.7265	6.4691
Panel B: DP as predictor							
Restricted							
PlugIn	Nopred	0.0398	1.3260	0.2605	0.0390	1.3208	0.2651
PlugIn	Unif	0.0606	1.4713	0.3294	0.0596	1.4584	0.3339
PlugIn	Shrink	0.0590	1.4405	0.2766	0.0557	1.4008	0.2465
Unrestricted							
PlugIn	Nopred	0.0454	1.3423	0.2397	0.0423	1.3240	0.2442
PlugIn	Unif	−0.0858	5.4109	5.3886	0.0106	4.2374	3.5769
PlugIn	Shrink	0.0846	2.5822	1.4741	0.0839	2.4064	1.5134
Panel C: Extended data set							
Restricted							
PlugIn	Nopred	0.0261	1.2282	0.2239	0.0262	1.2289	0.2285
PlugIn	Unif	0.0349	1.3280	0.2918	0.0341	1.2855	0.2556
PlugIn	Shrink	0.0329	1.3071	0.2621	0.0320	1.2578	0.2255
Unrestricted							
PlugIn	Nopred	−0.0009	1.2755	0.3802	0.0056	1.2574	0.3413
PlugIn	Unif	−0.0103	2.7430	2.3822	−0.0079	2.4642	2.2409
PlugIn	Shrink	0.0250	2.5004	1.9296	0.0181	2.2818	1.8354
Panel D: Rolling window							
Restricted							
PlugIn	Nopred	0.0317	1.2613	0.2325	0.0306	1.2584	0.2522
PlugIn	Unif	0.0515	1.4826	0.4227	0.0535	1.4716	0.3962
PlugIn	Shrink	0.0362	1.3150	0.2994	0.0379	1.3107	0.2829
Unrestricted							
PlugIn	Nopred	0.0521	1.4496	0.3217	0.0474	1.4068	0.3111
PlugIn	Unif	0.0778	4.9793	5.3891	0.0755	4.0642	3.7248
PlugIn	Shrink	0.1112	2.4863	1.3177	0.0933	2.2275	1.2095

Notes. This table reports the annualized CER, average terminal wealth ($\overline{TW}$), and standard deviation of terminal wealth ($\sigma(TW)$) for specifications that differ in one dimension from the main specification of Table 4 for $\gamma = 5$. Panel A extends the investment horizon to 120 months. Panel B uses DP as the predictor. Panel C uses the extended data set from January 1927 until December 2012. Panel D uses a rolling window of 240 months. Results are based on the general VAR(1).

specifications with unrestricted weights. It is important to note though that these results are more affected by estimation noise, since these are based on smaller estimation windows.

4.5. Statistical Tests

To evaluate the statistical significance of the results on hedge demands, we test whether the difference in average realized utility between dynamic strategies and myopic strategies is equal to zero. As a test statistic we use the Diebold and Mariano (1995) test on the utility series. Diebold and Mariano (1995) test whether the performance difference between two forecasts has zero expectation by means of a standard t-ratio. They show that this test statistic has an asymptotically standard normal distribution. Estimating the standard error of the average difference in realized utility is somewhat problematic because of the large number of overlapping data points in the 60-month cumulative returns. We estimate the covariance matrix of average realized utility nonparametrically by means of the Newey and West (1987) estimator and choose the lag length using the Newey and West (1994) selection criterion.

Results are reported in Table 8. Almost all t-statistics are small. The difference in value between the myopic and dynamic strategies is statistically insignificant. Among the restricted weight strategies, the only exceptions are for the shrinkage prior with $\gamma = 10$. For the unrestricted weights, we find significant hedging value in the model without predictability. Hence, hedging interest rate risk pays off, but recall from Table 4 that the CER values for the no-predictability prior are much lower than for models with predictability.[16]

The t-statistics for the alternative strategies of Tables 6 and 7 show similar results. The hedge demand adds some value for restricted strategies that use the plug-in method and shrinkage. The gain is statistically significant for the parsimonious model (t-statistic, 2.22), for $K = 120$ (t-statistic, 3.28), and with DP as the predictor (t-statistic, 4.84). However, the economic gains are still very small, as shown in the previous sections.

When using unrestricted weights, we do not find any significant results for $K = 120$. For this case, we found the largest gains in CER for the dynamic strategy, but at the same time were faced with a much smaller effective sample period due to the large overlap in the different 10-year utilities.

4.6. Monte Carlo Simulation

We complete our analysis of the value of hedge demands by a small Monte Carlo experiment, where we generate data under the null of predictability. In particular, we assume that the posterior means of the VAR parameters obtained using the uniform prior on the full data set define the true data generating process (DGP). For this model, we simulate 100 samples of 58 years of data, and for each sample we perform the same out-of-sample analysis as on the real data set. We calculate portfolios using the estimated parameters based on either the uniform or shrinkage prior. Because these simulations are computationally intensive, we consider only the approximate solutions with unrestricted portfolio weights in the Monte Carlo analysis.

Table 8 Significance Tests: Dynamic vs. Myopic Strategies

		$\gamma = 2$	$\gamma = 5$	$\gamma = 10$
Restricted				
PlugIn	Nopred	−1.64	0.74	1.19
PlugIn	Unif	0.08	−1.24	0.63
PlugIn	Shrink	0.20	1.10	3.24
DecTh	Unif	0.59	−1.58	1.12
DecTh	Shrink	−0.12	1.27	2.79
Unrestricted				
PlugIn	Nopred	4.10	2.57	1.09
PlugIn	Unif	—	−0.49	−1.43
PlugIn	Shrink	−0.15	0.59	0.41

Notes. This table presents t-statistics to test whether average realized utility for the dynamic and repeated myopic strategies are statistically different from each other. Specifications differ by weight restrictions (restricted or unrestricted), method (plug-in (PlugIn) or decision theoretic (DecTh)), and prior (no predictability (Nopred), uniform (Unif), or shrinkage (shrink)). Results are based on the general VAR(1) estimated using an expanding window.

The differences in CERs between a dynamic and a myopic strategy are shown in Table 9. A positive difference implies that the dynamic strategy outperforms the repeated myopic strategy. In the Monte Carlo simulations, we find that when parameters need to be estimated using only the available data, the dynamic strategy is often inferior to the repeated myopic strategy over a 58-year sample period. At the median, the dynamic strategy is slightly worse than the myopic strategy for the uniform prior, whereas the two are almost equal under the shrinkage prior. The median certainty equivalent return is very close to what we find in the actual data (see Table 4), except for the aggressive investor with $\gamma = 2$. The sampling distribution of the CER differences is also fairly negatively skewed. In the tails, negative differences over a 58-year sample tend to be larger than positive ones.

The median difference in the Monte Carlo simulations is negligible (or even negative) despite the fact that a dynamic strategy based on the true (unknown) model parameters would have been optimal given the DGP with predictability. As an (infeasible) benchmark, we also report the gains from hedging if the investor would know the full sample parameter estimates. In these ideal circumstances, investors with a γ of 5 or 10 would have benefited from hedging. Irrespective of the

[16] As an alternative, we also calculate t-statistics using only nonoverlapping observations and find that these t-statistics are on average 2.5 times smaller than the ones reported in Table 8. We again find that the added value of the hedge demand is not significant. Note however that this alternative test hardly has any power because of the small number of nonoverlapping observations.

Table 9 Monte Carlo Simulation with Predictability

	Uniform prior			Shrinkage prior		
	$\gamma = 2$	$\gamma = 5$	$\gamma = 10$	$\gamma = 2$	$\gamma = 5$	$\gamma = 10$
Median	−0.0109	−0.0049	0.0000	−0.0032	0.0006	0.0009
Min	−1.6968	−1.2137	−0.0922	−1.5752	−0.1434	−0.0840
5th	−1.2450	−0.1347	−0.0594	−0.1123	−0.0534	−0.0286
95th	0.0339	0.0371	0.0239	0.0435	0.0332	0.0191
Max	0.1739	0.0564	0.0473	0.0651	0.0514	0.0404
Data	0	−0.0071	−0.0134	−0.0020	0.0019	0.0011
In-sample	0	0.0204	0.0109	0	0.0196	0.0103

Notes. This table gives results for 100 Monte Carlo simulations assuming predictability. The table reports quantiles of the CER differences ($CER_{dyn} - CER_{myopic}$) for specifications based on either the uniform or shrinkage prior. Specifications differ in their risk-aversion level. Median, Min, Max, 5th, and 95th show the median, minimum, maximum, and percentiles, respectively, in the Monte Carlo simulations. Results are based on the VAR(1) estimated using an expanding window. Entries in the row labeled "Data" are the out-of-sample results found in the actual data set and are based on Table 4. "In-sample" gives the results for the actual data set when using the full in-sample estimates instead of expanding window estimates.

prior, an investor with $\gamma = 5$ would have made a 2% certainty equivalent gain from following the dynamic strategy. With the larger $\gamma = 10$, the gains from hedging are reduced, but are still positive at 1% and better than the median in the Monte Carlo simulation of the out-of-sample performance. This illustrates the effect of estimation error. For $\gamma = 2$, the CERs for both the dynamic and myopic strategies are −100% for the in-sample specification. Hence, no matter whether an investor knows the in-sample estimates or not, she takes too much risk and becomes insolvent in some periods. This result once again illustrates that allowing for short selling in a discrete time setting can lead to disastrous results for investors with very low risk aversion.

5. Conclusion

The hedge portfolio is the difference between the optimal portfolios of a long-term and a short-term investor. In theory, long-term investors should hold hedge portfolios to hedge future changes in investment opportunities. However, in an out-of-sample test, we find that this hedge component hardly adds value. Using a wide range of Bayesian priors and portfolio optimization methods, we find that repeated myopic portfolios perform as well as dynamic portfolios. Both dynamic and myopic investors benefit from predictability out-of-sample, but the certainty equivalent returns for both strategies remain very close to each other across alternative model specifications.

The negligible impact of the hedge term is due to parameter estimation error. First, Monte Carlo simulations show that even in simulated data that contain predictability, estimated dynamic strategies underperform compared to repeated myopic strategies. Second, portfolio weights are extreme, since weights are susceptible to error maximization. When short-selling restrictions are imposed, estimated myopic weights are often equal to the upper or lower bound, and therefore there is no room for hedge demands. The hedge component is more susceptible to estimation error than the myopic component, since the hedge demands are also affected by estimation error in state variable predictions. The additional estimation error offsets the theoretical gains from hedging changes in future investment opportunities.

For our main results, we consider an investor with a horizon of 60 months who can choose between equity, long-term nominal bonds and a short-term nominally risk-free rate. There are some aspects that we do not address with such a setup. First, hedge demands might have a larger impact when investment horizons are longer. In general, we do not find consistent outperformance of the dynamic strategy for longer horizons, but these results are statistically less reliable because we simply do not have enough nonoverlapping observations for reliable inference for very long investment horizons. Second, hedging might be more important for settings with more assets and for models with more complicated dynamics. One would expect, though, that estimation error plays an even more important role in such settings. Third, we cannot incorporate the hedge demand due to learning in our most general setting. Our results for learning in a restricted setting suggest that such hedge demands do not have a large impact on performance, but this might be different in more general settings.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2014.1924.

Acknowledgments

The authors are grateful to Brad Barber (department editor), the anonymous associate editor, two anonymous referees, Michael Brennan, Herman van Dijk, Roy Hoevenaars, Frank de Jong, Ronald Mahieu, Yuliya Plyakha, Jeffrey Pontiff, Marno Verbeek, Luis Viceira, and Peter Vlaar, as well as (seminar) participants at the Second International Workshop on Computational and Financial Econometrics in Neuchâtel, the 3M in Finance Workshop in Rotterdam, the 14th International Conference on Computing in Economics and Finance in Paris, Aarhus School of Business, APG Investments, the 2008 Econometric Society European Meeting in Milan, the 2010 European Finance Association Meeting in Frankfurt, the Long Term Asset Management Conference in Lausanne, Maastricht University, Manchester Business School, the NBER Time-Series Conference in Aarhus, the Netspar Pension Workshop, the Stockholm School of Economics, and Universität Münster. Bart Diris acknowledges financial support from the Netherlands Organisation for Scientific Research (NWO). Research by Peter Schotman was funded by Netspar.

References

Avramov D (2002) Stock return predictability and model uncertainty. *J. Financial Econom.* 64:423–458.

Barberis N (2000) Investing for the long run when returns are predictable. *J. Finance* 55:225–264.

Black F, Littermann R (1992) Global portfolio optimization. *Financial Analyst J.* 48:28–43.

Brandt M (2010) Portfolio choice problems. Aït-Sahalia Y, Hansen LP, eds. *Handbook of Financial Econometrics, Volume* 1: *Tools and Techniques* (North-Holland, Amsterdam), 269–336.

Brandt M, Goyal A, Santa-Clara P, Stroud J (2005) A simulation approach to dynamic portfolio choice with an application to learning about return predictability. *Rev. Financial Stud.* 18: 831–873.

Branger N, Breuer B, Schlag C (2010) Discrete-time implementation of continuous-time portfolio strategies. *Eur. J. Finance* 16:137–152.

Branger N, Larsen L, Munk C (2013) Robust portfolio choice with ambiguity and learning about return predictability. *J. Banking Finance* 37:1397–1411.

Brennan M, Xia Y (2002) Dynamic asset allocation under inflation. *J. Finance* 57:1201–1238.

Brennan M, Schwartz E, Lagnado R (1997) Strategic asset allocation. *J. Econom. Dynam. Control* 21:1377–1403.

Campbell J, Thompson S (2008) Predicting excess stock returns out of sample: Can anything beat the historical average? *Rev. Financial Stud.* 21:1509–1531.

Campbell J, Viceira L (2002) *Strategic Asset Allocation* (Oxford University Press, New York).

Campbell J, Chan Y, Viceira L (2003) A multivariate model of strategic asset allocation. *J. Financial Econom.* 67:41–80.

Campbell J, Cocco J, Gomes F, Maenhout P, Viceira L (2001) Stock market mean reversion and the optimal equity allocation of a long-lived investor. *Eur. Finance Rev.* 5:269–292.

Chacko G, Viceira L (2005) Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. *Rev. Financial Stud.* 18:1369–1402.

Cremers M (2002) Stock return predictability: A Bayesian model selection perspective. *Rev. Financial Stud.* 15:1223–1249.

Dangl T, Halling M (2012) Predictive regressions with time-varying coefficients. *J. Financial Econom.* 106:157–181.

DeMiguel V, Garlappi L, Uppal R (2005) How inefficient is the $1/N$ asset-allocation strategy? Working paper, London Business School, London. http://ssrn.com/abstract=785164.

DeMiguel V, Garlappi L, Uppal R (2009) Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? *Rev. Financial Stud.* 22:1915–1953.

Diebold F, Mariano R (1995) Comparing predictive accuracy. *J. Bus. Econom. Statist.* 13:253–263.

Doan T, Litterman R, Sims C (1984) Forecasting and conditional projection using realistic prior distributions. *Econometric Rev.* 3:1–100.

Engsted T, Pedersen T (2012) Return predictability and intertemporal asset allocation: Evidence from a bias-adjusted VAR model. *J. Empirical Finance* 19:241–253.

Fama E (1976) Forward rates as predictors of future spot rates. *J. Financial Econom.* 3:361–377.

Fama E, Schwert W (1977) Asset returns and inflation. *J. Financial Econom.* 5:115–146.

Goyal A, Welch I (2008) A comprehensive look at the empirical performance of equity premium prediction. *Rev. Financial Stud.* 21:1455–1508.

Johannes M, Korteweg A, Polson N (2014) Sequential learning, predictability, and optimal portfolio returns. *J. Finance* 69:611–644.

Jorion P (1986) Bayes-Stein estimation for portfolio analysis. *J. Financial Quant. Anal.* 21:279–292.

Jurek J, Viceira L (2011) Optimal value and growth tilts in long-horizon portfolios. *Rev. Finance* 15:29–74.

Koijen R, Nijman T, Werker B (2010) When can life-cycle investors benefit from time-varying bond risk premia. *Rev. Financial Stud.* 23:741–780.

Koijen R, Rodriguez JC, Sbuelz A (2009) Momentum and mean-reversion in strategic asset allocation. *Management Sci.* 55: 1199–1213.

Larsen L, Munk C (2012) The costs of suboptimal dynamic asset allocation: General results and applications to interest rate risk, stock volatility risk, and growth/value tilts. *J. Econom. Dynam. Control* 36:266–293.

Lynch A (2001) Portfolio choice and equity characteristics: Characterizing the hedging demands induced by return predictability. *J. Financial Econom.* 62:67–130.

Merton R (1969) Lifetime portfolio selection under uncertainty: The continuous time case. *Rev. Econom. Statist.* 51:247–257.

Merton R (1971) Optimal consumption and portfolio rules in a continuous-time model. *J. Econom. Theory* 3:373–413.

Merton R (1980) On estimating the expected return on the market: An exploratory investigation. *J. Financial Econom.* 8: 323–361.

Newey W, West K (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–708.

Newey W, West K (1994) Automatic lag selection in covariance matrix estimation. *Rev. Econom. Stud.* 61:631–653.

Ni S, Sun D (2003) Noninformative priors and frequentist risks of Bayesian estimators of vector-autoregressive models. *J. Econometrics* 115:159–197.

Pastor L, Stambaugh R (2000) Comparing asset pricing models: An investment perspective. *J. Financial Econom.* 56:335–381.

Pastor L, Stambaugh R (2012) Are stocks really less volatile in the long run? *J. Finance* 67:431–478.

Pesaran H, Timmermann A (1995) Predictability of stock returns: Robustness and economic significance. *J. Finance* 50: 1201–1228.

Pettenuzzo D, Timmermann A (2011) Predictability of stock returns and asset allocation under structural breaks. *J. Econometrics* 164:60–78.

Rapach D, Wohar M (2009) Multi-period portfolio choice and the intertemporal hedging demands for stocks and bonds: International evidence. *J. Internat. Money Finance* 28:427–453.

Sangvinatsos A, Wachter J (2005) Does the failure of the expectations hypothesis matter for long-term investors? *J. Finance* 60:179–230.

Sims C, Uhlig H (1991) Understanding unit rooters: A helicopter view. *Econometrica* 59:1591–1599.

Skoulakis G (2008) Dynamic portfolio choice with Bayesian learning. Working paper, University of Maryland, College Park.

Stambaugh R (1999) Predictive regressions. *J. Financial Econom.* 54:375–421.

van Binsbergen J, Brandt M (2007) Solving dynamic portfolio choice problems by recursing on optimized portfolio weights or on the value function. *Computational Econom.* 29:355–367.

Wachter J, Warusawitharana M (2009) Predictable returns and asset allocation: Should a skeptical investor time the market? *J. Econometrics* 148:162–178.

Xia Y (2001) Learning about predictability: The effect of parameter uncertainty on dynamic asset allocation. *J. Finance* 56:205–246.

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