Netspar Network for Studies on Pensions, Aging and Retirement

Frank Lutgens

Peter Schotman

Predictability-Robust Dynamic Portfolio Choice

Discussion Paper 2007 - 054

March 6, 2007

Predictability-Robust Dynamic Portfolio Choice

Frank Lutgens
Maastricht University and NETSPAR
and
Peter C. Schotman
Maastricht University, CEPR and NETSPAR

March 6, 2007

Abstract: Dynamic portfolio choice crucially depends on the predictability of returns. In contrast to its importance, confirmation of the existence of predictability is lacking. We consider a robust investor who arms herself against the adverse effects of uncertainty about predictability but also exploits the benefits of predictability insofar it is significant. We show that robust dynamic portfolio choice does not exhibit the extreme risky investment of traditional and Bayesian portfolios, is less volatile and features intertemporal hedging demands. We also show the worst plausible form of predictability for the robust portfolio result. This worst case depends on the horizon, initial state, the portfolio, and the relations among the predictor variables. For long term investment the worst case features a low long term average of the predictor variable and a high exposure to this low average.

JEL classifications: C11, C44, D80

Keywords: Robustness, Dynamic portfolio choice, model uncertainty, estimation uncertainty

Corresponding author: F. Lutgens, Department of Finance, Maastricht University, P.O. Box 616, 6200 MD Maastricht, Netherlands, email: `f.lutgens@finance.unimaas.nl`

Acknowledgement:

1 Introduction

If returns are predictable, then larger investment returns can be obtained by anticipating the predictable changes in the returns over time. If predictability turns out to be an illusion, then the investment that builds on predictability will proof disappointing, more disappointing than a portfolio that disregards predictability. We assume the perspective of a prudent investor who wants to avoid disillusion but also does not want to miss out on the benefits to the extent that returns are indeed predictable. How should the investor build her portfolio?

Recommendations differ. Predictability has supported a long tradition of critics. Classic studies by Samuelson (1973) and Shiller (1981) build on the economic rationale that excess returns must be constant over time. Variations in excess returns will merely reflect variations in risk. This support static/myopic portfolio choice. Dynamic investment would merely lead to extra transactions costs and be more prone to parameter uncertainty. Empirical findings on predictability are ambiguous. Literature seems to agree on the presence of long run predictability. Fama and French (1988) and Lo and MacKinlay (1988) find that weak short-term predictability accumulates to significant levels of return predictability in the long run. Cochrane (2006) finds this confirmed by the absence of dividend growth predictability that is irreconcilable with observed time varying dividend price ratios if returns would be unpredictable. In contrast, supporters of short term predictability are few and far between. Goetzman and Jorion (1993) find reduced evidence for return predictability. Goyal and Welch (2003), Goyal and Welch (2005) predict the next period's return with regression estimates based on all available data but conclude that the sample mean produces better out of sample prediction than do return forecasting estimates. Others do find use of short term predictability. Ang and Bekaert (2005) find short run predictive power of the short rate and Shanken and Tamayo (1964) miss support for the hypothesis that expected return predictability related to yield is explained by changes in the conditional return variance. Yet these results lack robustness. But if long-term predictability is a fact, why do we not find short term predictability; Do long-horizon regressions capture any information that is not present in one-period regressions? Cochrane (2004), and Cochrane (2006) argue that, given the persistence of dividend yields and related forecasting variables, by and large, they do not. Large short term return variances obscure predictability. Moreover, predictive variables are typical persistent, complicating accurate estimation. So, does weak evidence for predictability render dynamic strategies obsolete?

The problem is that changes in predictability change the optimal portfolio choice dramatically. The portfolio effect of uncertainty in predictability is stronger than uncertainty about a static return model. Indeed, in a model that features predictability, inaccuracies accumulate as predictors propagate through the model. Incorrect assumptions lead to suboptimal (extreme) portfolio choice and extra regret as the investor realizes that her delusive expectations prove non-existent. Barberis (2000), Campbell and Viceira (2002), Brennan and Xia (2005) and Kandel and Stambaugh (1996) empirically illustrate this. They report that risky investment may change by as much as 60% when the economic state variable changes by one standard deviation. In the high economic states, risky investment approaches 100%; in low economic states, the investor refrains from risky investment. Compare this to $30\% - 50\%$ (state independent) risky investment for myopic investors. Ignoring this implied time variation in expected returns would notably reduce portfolio performance. Does this means that predictability supports dynamic investment strategies? Long run predictability only finds no use for short run dynamic strategies[1]. Although Barberis (2000), and Campbell and Viceira (2002) show that time variation in the investment opportunities has significant effects on the optimal portfolio strategies of long lived investors, it does not justify short term dynamic strategies. Short-run predictability is a pre-requisite for dynamic strategies. Brennan and Xia (2005) show that this is hard to detect from short run returns using standard econometric methods, despite that it would imply major deviations of valuation ratios from their long run 'equilibrium' averages. Yet, Wachter and Warusawitharana (2006), and Kandel and Stambaugh (1996) spark hope. They show that even when predictive relationship is statistically weak, predictability is still significant measured by the optimal investment's sensitivity to the state predictive variables.

Summarizing, dynamic strategies come recommended only in the case of demonstrable short run predictability. But the latter is hard to detect if present at all. What should a skeptical investor pursue when confronted with these conflicting recommendations. How does the investor deal with uncertainty about predictability? We enter the field of decision making under uncertainty. A range of approaches have been proposed. The traditional *naive* approach (e.g. Merton (1973)) simply ignores uncertainty. This would would not satisfy the prudent investor that we consider.

[1]Unless, of course, long run expectations change on the short run.

Jagannathan and Ma (2003) suggest to counter the undesirable effects of uncertainty by imposing *additional constraints* on the portfolio. This is a treatment of symptoms and a priori it is difficult to determine what valid constraints need to be imposed: should we restrict to static investment? But that implies that we miss out (if not incur significant losses) under potential predictability. *Bayesian* approaches as in Kandel and Stambaugh (1996), and Barberis (2000) explicitly model the uncertainty about predictability. In this case the uncertainty about predictability is expressed as parameter uncertainty in the excess return model. More specifically uncertainty about the parameters that model the inter-temporal relationships between the excess return and the predictor variables. The Bayesian decision maker combines her subjective prior belief on predictability with sample information to derive the predictive moments of the excess returns. The expected investment result follows from integrating the portfolio performance over the parameter distribution. A tight prior featuring predictability will lead to little parameter uncertainty and thus little dispersion in the portfolio results when integrating over the parameter values. Such a prior will trigger dynamic portfolios. Alternatively, a tight prior featuring no predictability will trigger static portfolios. In the case of a loose or uninformative prior centered at no predictability, the sample information will be leading. The prior will merely serve to specify the uncertainty in the parameters. The dynamics of the investment choice will depend on the empirical evidence on predictability. Clearly, Bayesian investment crucially depends on the choice of subjective prior.

The above non-robust *dynamic* portfolios feature *extreme* and *volatile* investments. The *naive buy-and-hold* investor who invests once and awaits the result of that portfolio at investment horizon, increases risky investment with the time horizon. This is due to the inter-temporal hedging demand triggered by the positive effect of concurrent low returns on future returns implied by the predictive relationship. *Bayesian buy-and-hold* investors invest more prudent as uncertainty about predictability accumulates over time. The Bayesian *dynamic* investor can regularly update her portfolio and will take recourse actions before losses accumulate, also invests up to 100% in risk assets.

The *robust* investor, either following the axioms of Gilboa and Schmeidler (1989) or building on behavioral motivations, pursues to maximize the guaranteed performance over all reasonable assumptions on predictability. The predicate reasonable indicates that the investor does not consider every possible parameter combination when determining the worst case. Some parameter combinations are forthright implausible and would render the investor unnecessarily conservative. Instead the investor restricts to combinations that she deems plausible. She is confident that the selection contains the true parameter combination. Therefore, by construction, her robust portfolio guarantees the performance for the least favorable (but plausible) parameter combination and thus also for the true combination. While the robust solution may not achieve the performance of the tweaked portfolio in the case of strong predictability, it always provides a reasonable performance whatever the extent of predictability. Lutgens and Schotman (2006) show that the robust portfolio is not necessarily the most conservative portfolio. At times the robust solution features a larger risky investment than the pessimistic recommendations would. In the context of long-run dynamic investment, Anderson, Hansen and Sargent (1999), Maenhout (1999), Uppal and Wang (2003) and Cavadini, Sbuelz and Trojani (2001), use robust control theory to deduce the optimal portfolio choice. One could interpret these approaches as specific parsimonious quantifications of the 'multiple priors' uncertainty set in Gilboa and Schmeidler (1989). Kogan and Wang (2002) and Uppal and Wang (2003) elaborate on these equivalent approaches. For a critical discussion of this literature, we refer to Sims (2001) and Pathak (2002).

We compare the robust investor to the naive and Bayesian investors as described by Barberis (2000) and show that the robust portfolio does not exhibit the volatile and extreme risky investment. If predictability stands indisputably, then the robust portfolio features the inter-temporal hedging demand of risky assets. Typically, only a weak predictive relationship stands indisputably. In that case risky investment is reduced accordingly. Weaker predictability also reduces the sensitivity of the portfolio choice on the economic state. This leads to less volatile portfolios. The robust portfolio is directed by the worst plausible form of predictability. This worst case depends on the horizon, initial state, the portfolio, and the relations among the predictor variables. For long term investment this results in low long term average of the predictor variable and a high exposure to this low average. The latter varies with the time horizon and initial economic state. For short term investment the worst case depends on the initial economic state. We will graphically illustrate the sensitivity of the worst case to horizon and initial state effects.

The next two sections introduce the investor, her objective, investment opportunity set and her uncertainty about predictability. We also show the expected and worst case long-run performance of the risky market. Section 4 compares the the traditional, Bayesian and robust buy-and-hold investors and quantifies the worst case form of predictability. In section 5 we derive the robust dynamic portfolio and repeat the comparison. Section 6 concludes with some annotations and recommendations.

2 Predictability, uncertainty and robustness

We study an agent who considers investment in a risky stock index and a riskless Treasury Bill over a finite horizon of T years, possibly rebalancing her portfolio yearly. Her preferences are summarized by a utility function $u(\cdot)$. The investor does not derive utility from intermediate consumption, instead utility depends exclusively on the terminal wealth W_T accumulated over T years. She adopts the power utility function over terminal wealth with constant relative risk aversion $\gamma > 1$:

$$u(W_T) = \frac{W_T^{1-\gamma}}{1-\gamma}. \tag{1}$$

Wealth accumulates through portfolio returns over time according to

$$W_{t+1} = (1 + R_{t+1}(w_t))W_t \tag{2}$$

where w_t denotes the investment in the risky asset with an implicit investment of $(1 - w_t)$ in the riskfree asset and $R_{t+1}(w_t)$ denotes the portfolio return on asset allocation w_t over period $[t, t + 1)$. We stylize the problem by assuming that the continuously compounded riskless monthly return on Treasury bills is a constant r_f. The returns for the stock index are given by a model for the monthly continuously compounded excess stock index returns y_t:

$$y_t = \alpha_1 + \beta_1 x_{t-1} + \varepsilon_{1t}. \tag{3}$$

with the dividend yield x_t serving as a predictor variable. This predictor variable is assumed to follow the stochastic process

$$x_t = \alpha_2 + \beta_2 x_{t-1} + \varepsilon_{2t}. \tag{4}$$

The error terms $(\varepsilon_{1t}, \varepsilon_{2t})$ have a joint normal distribution with covariance matrix Σ. The stochastic portfolio return over the period $[t, t + 1)$ on asset allocation w_t is

$$1 + R_{t+1}(w_t) = (1 - w_t)e^{r_f} + w_t e^{r_f + y_{t+1}}. \tag{5}$$

Note that the return model family (3)-(4) includes the static model by setting $\beta_1 = 0$. The distribution of multi-period returns follows from the return model (3)-(4)

which features predictability. Among others, we are interested in the cumulative continuously compounded excess return over T months. This long-term return y_T is normally distributed with[2]

$$\begin{aligned}\mu_T &= T(\alpha_1 + \tfrac{\beta_1}{1-\beta_2}\alpha_2) + (x_0 - \nu)\beta_1\chi(T)\\ \sigma_T^2 &= T\sigma_1^2 + 2\beta_1\rho(T)\sigma_{12} + \beta_1^2\phi(T)\sigma_2^2.\end{aligned}$$

The definitions of ξ, ρ and ϕ can be found in the appendix. Note that the expected log excess returns (33) are linear in x_0 and the variance does not depend on x_0. Cumulative returns $R_T = e^{Tr_f + y_T} - 1$ are lognormally distributed with parameters $\mu_T + Tr_f$ and σ_T^2.

The portfolio choice problem comes in two modes. A *buy-and-hold* investor chooses an allocation at the beginning of the horizon and passively awaits the results of her decision at the end of the horizon. Although this excludes dynamic portfolios, the optimal investment strategy anticipates the dynamics of the return model. The investor solves

$$\begin{aligned}&\max_{w\in\mathcal{W}} \mathsf{E}_0[u(W_T(w))]\\ &\text{subject to } W_T \leq (1 + R_T(w))W_0.\end{aligned} \tag{6}$$

A *dynamic* investor rebalances her portfolio at certain time points, say once a year. In this case, the investor chooses her current allocation, with the prospect that she may rebalance at the beginning of every year to come, taking into consideration the information that will be available at that time. This allows her to react timely to new developments with so-called recourse actions. With the knowledge that she can alter her strategy to adapt to the return model dynamics and can possibly correct her decisions later if the economy turns against her, the investor may wish to choose an initial asset allocation that differs from a buy-and-hold strategy. The dynamic investor solves

$$\begin{aligned}&\max_{\mathbf{w}\in\mathcal{W},\ W} \mathsf{E}_0[u(W_T(\mathbf{w}))]\\ &\text{subject to } W_{t+1} \leq (1 + R_{t+1}(w_t))W_t,\ \forall t \in \{1, .., T-1\}.\end{aligned} \tag{7}$$

2.1 Uncertainty

The return model (3)-(4) is estimated to predict future excess returns. If the return model is correctly specified, then all deviations from the *true* model are exclusively due to estimation errors in the estimators for the parameters $\xi = [\alpha_1\ \beta_1\ \alpha_2\ \beta_2]'$

[2] Given $\beta_2 < 1$.

and Σ. To simplify the analysis we focus on estimation errors (uncertainty) in the parameters[3] ξ. Possible estimation errors induce uncertainty about all quantities that depend on the parameters. Note that the expectation operator and the portfolio return (5) are functions of the parameter ξ. If we use an estimate for ξ which may, due to estimation uncertainty, deviate from the true parameter value, we introduce uncertainty in the expectation operator and the portfolio return. When required for clarity, we will explicitly show the functional dependence of the expectation operator and portfolio return on the parameter by including ξ in the notation.

If the investor adopts the traditional *naive* approach to uncertainty, she conditions on a single parameter estimate and solves

$$\begin{array}{ll} \max_{w \in \mathcal{W}} & \mathsf{E}_0[u(W_T) |\ \hat{\xi}] \\ \text{subject to} & W_{t+1} = (1 + R_{t+1}(w_t, \hat{\xi}))W_t. \end{array} \tag{8}$$

A *Bayesian* investor recognizes that an estimated model is only approximately true. Therefore she does not condition on $\hat{\xi}$, but integrates over all possible parameter values using the posterior parameter distribution. Given $\tilde{T}$ observations on monthly dividend yield $x_{-1}, .., x_{-\tilde{T}}$ ($X = [\iota \ x]$), monthly continuously compounded excess returns $Y = [y_{-1}, .., y_{-\tilde{T}}]'$ and the posterior parameter distribution $p(\xi |\ Y,\ X)$, the Bayesian investor maximizes

$$\mathsf{E}_{B,0}\left[u(W_T)\right] = \int \mathsf{E}_0\left[u(W_T) \mid \xi\right] p(\xi \mid X, Y) d\xi \tag{9}$$

with wealth accumulation following from $W_{t+1} = (1 + R_{t+1}(w_t, \xi))W_t$. We will follow Barberis (2000) and consider a Bayesian investor who uses a standard non-informative prior.

Also a *robust* investor is aware that an estimated model can only be approximately true. The investor wishes to account for this uncertainty. Unlike a Bayesian investor, however, she does not (know how to) assign probabilities to alternative estimates. Instead she considers the least favorable alternative ξ from within some uncertainty set $\mathcal{U}$ of conceivable parameter values and maximizes,

$$\min_{\xi \in \mathcal{U}} \mathsf{E}_0\left[u(W_T) \mid \xi\right]. \tag{10}$$

The investor is confident that the true model parameters lie within the uncertainty set. Therefore the investor trusts that an asset allocation with good performance

[3] Σ can be estimated much more precisely than ξ, cf. Chopra and Ziemba (1993).

for the least favorable model in the set will also perform well for the true model. The critical determinant of robust decision making is the set of alternative models considered by the investor. If this set is large, the worst case return model and consequently the investor's behavior will be rather prudent and robustness will result in passive portfolio strategies. On the other hand, if the investor has a strong belief in some reference model and the set of alternatives is small, her portfolio will be almost equal to the portfolio of a naive investor who relies on the reference model.

We use the return model (3)-(4) estimated on the historical observations as reference model denoted with parameter (estimate) $\hat{\xi}$. We form the uncertainty set in a similar way as a confidence region for the true parameter values and consider the set of parameter values with highest posterior density with a cumulative density of, say, 95%, given the observed data,

$$\mathcal{U}: \int_{\xi \in \mathcal{U}} p(\xi | X,\ Y) d\xi = 0.95 \tag{11}$$

Assume Σ is known, and note that the conditional distribution of ξ is multivariate normal, hence a member of the class of ellipsoidal distributions. The equi-probable parameter values form an ellipsoid and the uncertainty can be expressed as

$$\mathcal{U} = \{\xi :\ (\xi - \hat{\xi})' \Omega^{-1} (\xi - \hat{\xi}) \leq \theta^2\}. \tag{12}$$

We calibrate the uncertainty set on the posterior distribution of Barberis (2000) which is based on an uninformative prior. In this case the covariance matrix $\Omega = \Sigma \otimes (X'X)^{-1}$ and $\theta = \sqrt{\chi^2_{inv}(4)}$. Note that if the static model ($\beta_1 = 0$) is part of the uncertainty set $\mathcal{U}$, return predictability is not significant according to the robust investor. In this case the robust portfolio choice must consider predictable and non-predictable returns.

We append the uncertainty set definition with an economic consideration that mean reversion for dividend yields. Hereto we impose that the long term mean $\nu = \frac{\alpha_2}{1-\beta_2}$ is finite. This certainly is the case if we append the constraint $\beta_2 \leq 1 - \epsilon$ ($\epsilon = 0.01$) to the uncertainty set,

$$\mathcal{U} \rightarrow \mathcal{U} \cap \{\xi :\ \beta_2 < 1 - \epsilon\} \tag{13}$$

Figure 2 depicts the uncertainty set based on empirical data that will be introduced in the next section. The strong correlations between the parameters cause 'narrow' ellipsoidal forms; large values for α_1 comply with small values for β_1 and α_2

and large values for β_2 and vice versa. Observe that the restriction $\beta_2 < 1$ also implies a truncation of the other parameters due to the relations associated with the ellipsoid. The uncertainty set for the short data set featuring larger uncertainty contains the return model without predictability $\beta_1 = 0$. Note that by considering relations between parameters the uncertainty set is significantly smaller that the Carthesian product of the confidence intervals of the individual parameters. The prevents unnecessary conservativeness. conservativeness.

Given the uncertainty set, a robust investor chooses an asset allocation that is (a) *feasible* for all parameters, i.e. an investment strategy that satisfies the budget constraint for all parameters $\xi \in \mathcal{U}$, and (b) has best robust performance, which means that the worst performance over the parameter set is maximal. The *robust buy-and-hold portfolio choice problem* is

$$\begin{aligned} &\max_{w \in \mathcal{W}} \min_{\xi \in \mathcal{U}} \; \mathsf{E}_0[u(W_T(w)|\, \xi)] \\ &\text{subject to } W_T \leq (1 + R_T(w, \xi))W_0, \quad \forall \xi \in \mathcal{U}. \end{aligned} \tag{14}$$

A solution to (14) is an asset allocation w with the highest level of expected utility which is robust to the parameter values in the uncertainty set.

The robust model of the dynamic multi-period problem is more intricate. Nilim and El Ghaoui (2002) propose a model and solution method to the robust multi-period dynamic programming problem. If we apply their model to problem (8), we obtain the robust optimization problem

$$\begin{aligned} &\max_{w,\ W} \min_{\xi \in \mathcal{U}} \; \mathsf{E}_0[u(W_T(w)|\, \xi)] \\ &\text{subject to } W_{t+1} \leq (1 + R_{t+1}(w_t, \xi_t))W_t, \quad \forall \xi_t \in \mathcal{U},\ t \in \{1, .., T-1\}. \end{aligned} \tag{15}$$

A feasible solution to (15) consists of portfolio allocations $w = \{w_t\}_{t=1}^{T-1}$ and wealth variables $W = \{W_t\}_{t=2}^{T}$ that satisfy the constraints for each period and for all parameter values. However, we deal with a special problem for which model (15) leads to conservative solutions. Indeed, the worst case in model (15) may feature different worst case parameters at different times and states. However nature (the bogeyman) has one single shot to make the investor's life miserable: Given the investor's portfolio choice, nature may yield up its physical constants ξ possibly ruining the investor's performance. However, nature cannot reconsider its physical constants at a later state. Problem (15) allows nature to change its physical constants over time, portraying an over conservative performance

The crucial difference between our problem and model (15) is the function of the variables W_t. In our problem definition, these merely serve as 'book-keepers' and are not subjected to other constraints such that only terminal wealth W_T is relevant for the problem. However in model (15), the W_t serve as decision variables that must be guaranteed, also at intermediate time periods. As seen by the example where robust return was zero instead of 10%, this leads to a 'leak' in the wealth accumulation. Intermediate outcomes would be relevant if the utility function would contain consumption in every period as in Campbell and Viceira (2002). This would be a more difficult problem to solve.

We resolve this modelling issue by substituting for the redundant state variable wealth and express terminal wealth as a function of the asset allocations,

$$W_T(w, \xi) = \prod_{t=1}^{T} (1 + R_t(w_{t-1}|\ \xi))\, W_0. \tag{16}$$

We substitute (16) in (8) and omit the variables W_t in our problem definition. Accordingly, we define the *robust multi-period portfolio choice problem* as:

$$\max_{w \in \mathcal{W}} \min_{\xi \in \mathcal{U}} \ \mathsf{E}_0 \left[u(W_T(w))|\ \xi \right], \tag{17}$$

with W_T defined in (16). A robust solution is a plan which specifies the portfolio allocation for every time period and (future) economic state and this plan maximizes the worst case expected utility.

3 Investment opportunities

In this section we analyze the data that we will use as input for portfolio choice in the next sections. We consider the uncertainty in the data and its potential to sustain a positive Sharpe ratio, an indicator for the performance potential of the investment opportunity set.

3.1 Data

The estimation results by Barberis (2000), for convenience repeated in table 1, serve as data input. Barberis (2000) considers two samples of different length. The long dataset consists of 523 monthly observations of stock index returns, dividend yields and Treasury bill returns from June 1952 to December 1995. The short dataset consists of 120 observations from January 1985 to December 1995. The stock index return and Treasury bill returns are used to calculate the excess stock index return. The continuously compounded annual risk-free rate is set to 4.41%.

Insert table 1 Parameter estimates about here.

The estimation results of the two samples differ fundamentally in parameter uncertainty. The long dataset produces a significant mean excess return, the short sample produces a mean excess return which is not significantly different from zero. Consequently, only the first set of estimators suggest a sustainable positive Sharpe ratio which is an indicator for the profitability of active participation in the stock market.

Yet, if predictability in the asset returns is considered, this conclusion may change. Barberis (2000, p.243-p.245) and Campbell and Viceira (2002), among others, note that the conditional variance of cumulative stock returns may grow slower than linearly with the investor's horizon. The economic intuition follows from the negative contemporaneous correlation between dividend yield and stock returns. A decrease in dividend yield is likely to be accompanied by a contemporaneous positive shock to returns. However, since the dividend yield is lower, stock returns are forecast to be lower in the future ($\hat{\beta}_1 > 0$). The rise, followed by a fall in returns, generates a component of negative serial correlation in returns which slows the evolution of the variance of cumulative returns as the horizon grows.

A naive buy-and-hold investor who relies entirely on the estimated parameter values, interprets this effect of predictability as a reduction in risk and hence as a justification for active participation in the risky market for long- term asset allocation. Investors who account for parameter uncertainty do not necessarily invest more in the risky asset for longer horizons. Barberis (2000) shows that an investor who adopts a Bayesian approach to uncertainty believes that conditional variances grow more quickly as the horizon grows and doubts whether the predictive power of the dividend yield is large enough to slow the evolution of conditional variances.

Also a robust investor is reluctant to interpret (estimated) predictability entirely as a reduction in risk. Based on the uncertainty of the parameters, the robust investor also considers alternative conceivable parameter configurations that may not feature predictability. As portfolio choice must be robust to such alternative parameter configurations, the robust investor may choose to not exploit the predictability featured by the parameter (point) estimates.

3.2 A detour - Effect of uncertainty on the Sharpe ratio

The relevant summary statistic for the dataset is its implied performance potential of risky investment. The intended performance measure is this article is given by (1). Yet, for illustrative reasons we take a short detour in this section and consider the Sharpe ratio to describe the potential performance. More specifically, we consider the maximal (geometric) average monthly Sharpe ratio for investment horizons ranging from one month to 100 years:

$$\mathrm{Sh} \approx \frac{\mu_T}{\sigma_T} + \frac{1}{2}\sigma_T/T.$$

The uncertainty set for the robust approach is based on (12) with 95% preference for robustness ($\theta \approx 3.1$). The *robust* average monthly Sharpe ratio is defined by

$$\min\nolimits_{\xi \in \mathcal{U}} \mathrm{Sh}(\xi). \tag{18}$$

Note that the Sharpe ratio may take negative values as it assumes a forced 100% in risky assets. Indeed, it has direct implications for the maximal portfolio performance. The robust Sharpe ratio serves as an upperbound on the robust mean-variance utility function $Q(w, \xi) = \mu_T(\xi, w) - \frac{1}{2}\gamma\sigma_T^2(\xi, w)$ where γ denotes the investor's risk aversion, $\mu_T(\xi, w)$ is portfolio return and $\sigma_T(\xi, w)$ is the portfolio variance. Indeed,

$$\max_w \min_{\xi \in \mathcal{U}} f(w, \xi) \leq \min_{\xi \in \mathcal{U}} \max_w f(w, \xi) = \min_{\xi \in \mathcal{U}} \frac{1}{2\gamma}\mathrm{Sh}(\xi).$$

Hence for mean-variance portfolio choice with multiple risky assets, a positive robust Sharpe ratio is not indisputable evidence for existence of a robust asset allocation with positive expected performance.

Problem (18) is a non-convex optimization problem. Fortunately it has only four variables and we solve it with a conjugate gradient method as provided by Matlab version 6.1 (R12.1) with multiple starting solutions $\mathcal{U}_f$ to enhance global optimality. The set $\mathcal{U}_f$ is a grid of parameter values centered at $\hat{\xi}$ and with grid distances $\alpha\hat{\sigma}_\xi$ with the scalar α set such that $P = 120$ grid points are contained in the original set $\mathcal{U}$. In theory, the non-convex inner optimization problem may produce local optima for the worst case parameter value which leads to a decline in robustness of the solution. Numerical experiments[4] do not indicate that this is a problem.

[4] The computed solution to the inner minimization problem does not improve when we increase the number of starting solutions to $P = 2.500$.

Empirical illustration

Do traditional and robust performance evaluations differ regarding their dependence on the initial economic state and the investment horizon? Figure 1 shows the traditional and robust average monthly Sharpe ratio for alternative investment horizons and different initial states. The traditional and robust performance estimates react fundamentally different to changes in the initial state and investment horizon.

INSERT FIGURE 1 EXPECTED PERFORMANCE.

Without considering uncertainty, the market appears lucrative. Only low initial dividend yields and short investment horizons feature negative Sharpe ratios. The variance of the cumulative (log) stock return grows slower than the expected (log) stock return. Consequently performance increases with the horizon. The long term expected performance under predictability is, even for the low initial states, significantly higher than a static model would imply (with an expected $12,1\%$ and $14,9\%$ performance). Note that the differences in performances corresponding to different initial states initially rise but later decline as the investment horizon lengthens. The initial increase is due to *persistence* of the dividend yield. The consequences of a high initial dividend yield are still felt several years hence. Ultimately the average dividend yield will converge to its long term average dividend yield, implying long-term *mean reversion*. The first signs of this convergence are best observed from the short dataset that features less persistent dividend yields (lower β_2).

A robust evaluation of the long dataset only shows significant investment opportunities for high initial states and long horizons. The uncertainty set contains exclusively $\beta_1 > 0$ and $\alpha_2,\ \beta_2 > 0$. Hence predictability is a robust property of the return model albeit in a weakened form. It also implies that the long term average dividend yield $\nu = \alpha_2/(1 - \beta_2)$ is positive. Hence the positive investment performance for long horizons. The short dataset does not feature significant investment opportunities. Return predictability is not a robust characteristic of this dataset. The uncertainty set also contains the static model and even a return model with a negative dependence on the state variable ($\beta_1 < 0$ and $\alpha_2,\ \beta_2 > 0$). This leads to a negative robust performance irrespective of the initial dividend yield or investment horizon.

INSERT FIGURE 2 WORST CASE CONFIGURATION ABOUT HERE.

Also observe that the robust performance does not monotonically in- or decrease with the investment horizon. This is due to different worst cases for different investment horizons. Indeed, the worst case form of predictability ξ depends on the investment horizon, the initial state and the portfolio. Figure 2 shows the worst case parameter configurations for alternative initial dividend yields and investment horizons. For a one-year investment horizon and small dividend yield, the worst case features small α_1 and β_2. This directly implies a small return and small dividend yields in the short-term. The high correlations between the parameters imply that small α_1 and β_2 coincide with large β_1 and α_2. The importance of β_1 reduces when initial dividend yields and α_2 are low. For large (persistent) initial dividend yields, β_1 and β_2 are reduced to form the worst case. For extremely long horizons, e.g. 100 years, the initial dividend yield does not affect the average log returns. In that case worst case features low long-term average dividend yield $\nu = \alpha_2/(1-\beta_2)$ and β_1. The worst case parameter configurations for a 5-year investment horizon combine the aforementioned effects: In the low states, the worst case parameter configurations concentrate on small α_1 and β_2. When the initial dividend yield increases, β_1 is swiftly reduced. The worst cases for the short dataset are slightly different. For short term horizons, the worst case parameter configurations follow a similar pattern as for the long dataset. But as the larger uncertainty set contains negative β_1 the robust performance remains negative for large initial dividend yields. Moreover as the uncertainty set contains negative α_2 and consequently negative long term average dividend yields ν, the robust performance for long-term horizons may be negative for positive β_1. In this case a robust evaluation of the performance does not favor long-term over short-term investment.

4 Buy-and-hold portfolio choice

The previous section described the investment opportunities in terms of the Sharpe ratio. We continue with studying optimal portfolio choice under the intended power utility performance measure (1). We start with the buy-and-hold investor who was introduced in section 2. She solves (6). The Power utility prevents further simplification of expected utility as a function of the first return moments as was possible in the previous section. Instead we solve (6) numerically. For given ξ, we approximate the conditional expectation by a finite approximation,

$$\sum_{j=1}^{J} p_j(\xi) \frac{W_T(w, y_{T,j}(\xi))^{1-\gamma}}{1-\gamma} \tag{19}$$

with

$$W_T(w, y_{T,j}(\xi)) = \left(we^{Tr_f + y_{T,j}(\xi)} + (1-w)e^{Tr_f}\right) W_0.$$

and $\{p_j(\xi), y_{T,j}(\xi)\}_{j=1}^{J}$ is a stratified sample conditional on the parameter ξ. We stratify a sample of residuals ε from the standard normal distribution in $J = 1000$ equispaced intervals with means ε_j and cumulative probabilities p_j. Conditional on x_0, we compute $y_{T,j}(\xi) = \mu_T(\xi) + \sigma_T(\xi)\varepsilon_j$ with $\mu_T(\xi)$ and $\sigma_T^2(\xi)$ given by (33) and the first diagonal element of Σ_T respectively. We keep the stratified sample fixed across different T and ξ. This enhances comparison between alternative horizons and parameter values, even if the fixed distribution deviates slightly from the original distribution. For the reader's comfort we note that the the stratified sample leads to the same results (see figures (4) and (6)) as reported by Barberis (2000) who has verified that his approximation is accurate. The stratified sample is computationally faster than a large random sample, which is needed to calculate robust solutions[5].

A naive investor assumes $\xi = \hat{\xi}$ and maximizes (19). A Bayesian investor prepares an unconditional approximate return distribution. She samples $\tilde{K} = 1.000.000$ parameter values from the posterior parameter distribution and for each sampled parameter $\tilde{\xi}_{\tilde{j}}$, $\tilde{j} = 1, .., \tilde{J}$, she draws a random residual $\tilde{\varepsilon}$ from the stratified sample $\{p_k(\tilde{\xi}), y_{T,j}(\tilde{\xi})\}_{j=1}^{J}$ and computes the corresponding conditional excess return $y_{T,\tilde{j}} = \mu_T(\tilde{\xi}_{\tilde{j}}) + \sigma_T(\tilde{\xi}_{\tilde{j}})\tilde{\varepsilon}$. The Bayesian investor solves

$$\sum_{\tilde{j}=1}^{\tilde{J}} \frac{1}{\tilde{J}} \frac{W_T(w, y_{T,\tilde{j}})^{1-\gamma}}{1-\gamma}. \tag{20}$$

The robust investor solves the approximation to (14):

$$\max_{w \in \mathcal{W}} \min_{\xi \in \mathcal{U}} \sum_{j=1}^{J} p_j(\xi) \frac{W_{T,j}(\xi)^{1-\gamma}}{1-\gamma}. \tag{21}$$

The objective function (19) is convex on $w \in [0, 1]$ for each parameter ξ and consequently the portfolio choice problems for a naive and a Bayesian investor are convex

[5]The computations for figure 7 take 72 hours CPU time on an AMD Athlon 2400 Mhz XP+/1GB RAM running Matlab version 6.1.0.450 (R12.1) under Debian BGU/Linux 3.0.

optimization problems on $w \in [0,1]$ and can be solved to (global) optimality. Also the minimum of a set of convex functions is a convex function; hence the outer maximization (21) is a convex optimization problem. However, each function evaluation of the maximization problem is computationally costly as it involves solving the inner minimization problem. Moreover, the minimization problem is not a convex optimization problem on the domain $\xi \in \mathcal{U}$. We adopt sequential optimization to solve iteratively the outer maximization of (21) with iterations that involve solving the inner minimization problem. Convergence to a global optimal solution of the non-convex inner minimization problem is enhanced by considering multiple starting solutions, given by the set $\mathcal{U}_f$ with $P = 120$ as explained in the previous section.

Results

We adopt the same setting as Barberis (2000) with $\gamma = 10$. This enables us to validate and compare the results depicted in figures 4 and 6.

Insert figure 4 Buy and Hold (52) about here.

The *naive* investor monotonically increases risky asset allocation with the horizon and initial state. Without estimation uncertainty, predictability reduces the risk at long horizons; the optimal risky investment is increased accordingly. The difference in asset allocation for alternative states decreases as the investment horizon lengthens and the expected dividend yield converges to its long term average. The *Bayesian* investor finds that the conditional variance of the stock return increases due to uncertainty about predictaiblity. The resulting risky investment follows from balancing predictability which makes stocks look less risky at long horizons and estimation uncertainty which makes stocks look more risky. The optimal investment seizes to be a monotonic function of the investment horizon and dividend yield. Uncertainty may even reduce long-term risky investment below short term risky investment.

The robust investor evaluates the long dataset as supporting significant yet weakened predictability. Accordingly the investor increases, albeit with reduced slope, risky investment monotonically over the investment horizon and initial dividend yield. Risky investment for different initial states converges as the investment horizon lengthens and the expected dividend yield converges to its long term average. The worst case forms of predictability are depicted in figure 5 and follow a similar pattern as described in the previous section. For short term investment and low initial dividend yields, the worst case parameter configuration features small α_1 and α_2. For above average dividend yields, the 'predictability parameter' β_1 is reduced. For long term investment the average dividend yield converges to ν and the worst parameter configuration features small β_1. The short dataset does not feature significant predictability. The uncertainty set supports negative α_1, β_1 and α_2. Negative α_1s lead to negative returns on the short term horizons and low initial states while negative β_1 lead to negative returns for short term horizons with high initial dividend yields and for long term horizons (with $\alpha_2 > 0$ and $\beta_2 \approx 0.99$).

Insert figure 5 WC Buy and Hold (52) about here.

Insert figure 6 Buy and Hold (85) about here.

The Bayesian and robust investors arrive at fundamentally different portfolio strategies although they both address uncertainty. One explanation follows from the slightly different priors that the Bayesian and robust investors adopt. The robust investor, unlike the Bayesian investor, uses economic rationale to limit $\beta_2 < 1$. We may improve the basis for comparison by aligning the prior information of the two approaches. The lower panel in figure 4 shows the Bayesian strategy if the economic rationale $\beta_2 < 1$ is considered. A fundamentally different strategy results. A Bayesian still invests more in the risky assets than a robust investor but the portfolios show similar development over time horizons. It also illustrates that the optimal decision of a Bayesian investor is sensitive to the prior.

5 Dynamic portfolio choice

A dynamic investor may revise her portfolio every year to adapt to newly obtained information, in this case the dividend yield, at the end of each year. More specifically, assuming a K year horizon, the investor may revise her portfolio at the beginning of each year $k = \{0, 1, .., K-1\}$ whereupon the portfolio is invested for the period $[t_k, t_{k+1})$. We denote the value of the dividend yield at the beginning of year k by x_k and the cumulative continuously compounded excess return over the period $[t_k, t_{k+1})$ conditional on x_k is denoted y_{k+1,x_k}. A dynamic portfolio allocation w is a plan that indicates the investment w_{k,x_k} for each decision moment t_k, $k = 0,..,K-1$ and concurrent dividend yield x_k. We denote wealth accumulated at t_k in state x_k by W_{k,x_k}. Moreover, we use W_K to refer to the wealth at the end of the horizon.

Conditional on the parameter ξ, which we omit in the notation for now, the dynamic portfolio choice problem is

$$\max_{w \in \mathcal{W}} \mathsf{E}_0 \left[\frac{{W_K}^{1-\gamma}}{1-\gamma} \right]. \tag{22}$$

with wealth transitions

$$\begin{aligned} W_{k+1,x_{k+1}} &= (1 + R_{k+1,x_k}(w))\, W_{k,x_k} \\ W_K &= \big(1 + R_{K,x_{K-1}}(w)\big)\, W_{K-1,x_{K-1}}. \end{aligned}$$

and portfolio return

$$1 + R_{k+1,x_k}(w) = (w_{k,x_k} e^{r_f + y_{k+1,x_k}} + (1 - w_{k,x_k}) e^{r_f}).$$

and $\mathcal{W}$ presents the set of portfolios w with $w_{k,x_k} \in [0,1]$ for all t_k and x_k. Note that the expectation operator in (22) is a function of the returns y_{k,x_k} and the transition probabilities between states $[x_k, x_{k+1}]$ in subsequent time periods $[t_k, t_{k+1})$. Problem (22) presents a dynamic optimization problem suitable for applications of the Bellman optimality principle. The derived utility of wealth at any state (t_k, x_k) is

$$J(W_{k,x_k}, t_k, x_k) = \max_{w} \mathsf{E}_{t_k} \left[\frac{W_K^{1-\gamma}}{1-\gamma} \mid\; W_{k,x_k}, t_k, x_k \right] \tag{23}$$

which is the maximal expected (final) utility as expected at time t_k in state x_k. The Bellman principle of optimality states

$$J(W_{k,x_k}, t_k, x_k) = \max_{w_{k,x_k}} \mathsf{E}_{t_k} \left[J(W_{k+1,x_{k+1}}, t_{k+1}, x_{k+1}) \mid\; W_{k,x_k}, t_k, x_k \right]. \tag{24}$$

For a more convenient expression, define

$$Q(t_k, x_k) = \max_{w_{k,x_k}} \mathsf{E}_{t_k} \left[\frac{W_K^{1-\gamma}}{W_{k,x_k}^{1-\gamma}} \mid\; t_k, x_k \right] \tag{25}$$

and note that (by homogeneity in W of degree $1-\gamma$),

$$J(W_{k,x_k}, t_k, x_k) = \frac{W_{k,x_k}^{1-\gamma}}{1-\gamma} Q(t_k, x_k). \tag{26}$$

Substitution of (26) in the Bellman equation (24),

$$Q(t_k, x_k) = \max_{w_{k,x_k} \in [0,1]} \mathsf{E}_{t_k} \left[(1 + R_{k+1,x_k}(w_{k,x_k}))^{1-\gamma} Q(t_{k+1}, x_{k+1}) \middle|\; t_k, x_k \right]. \tag{27}$$

Note that (25) implies $Q(t_K, \cdot) = 1$ and, due to the investor's constant relative risk aversion, W does not enter expression (27)[6]. Given $Q(0, x_0)$ and initial wealth W_0, the expected terminal utility of portfolio choice w is $W_0/Q(0, x_0)$. Hence solving (22) is equivalent to computing $Q(0, x_0)$. Moreover, if we follow Barberis (2000) and discretize the state space in 25 equally spaced grid points on the interval ranging from three standard deviations below the historical mean to three standard deviations above, we can solve for $Q(0, x_0)$ by backward induction using (27). To calculate the expectation in (27) we rely on a sample $\{p_j(\xi), x_{k+1,x_k,j}, y_{k+1,x_k,j}(\xi)\}_{j=1}^{J}$ based on a stratified sample of two-variate residuals $\varepsilon \sim N(0, I_2)$ in $J = 1000$ equi-spaced intervals with means ε_j and cumulative returns p_j. Conditional on x_k, we compute

$$\begin{pmatrix} y_{k+1,x_k,j}(\xi) \\ x_{k+1,j}(\xi) \end{pmatrix} = \begin{pmatrix} \mu_T(\xi) \\ \nu_T(\xi) \end{pmatrix} + A_T(\xi)\varepsilon_j \tag{28}$$

with $A_T(\xi)$ the lower triangular part of the Choleski decomposition of $\Sigma_T(\xi)$ and $(\mu_T(\xi),\ \nu_T(\xi))$ and $\Sigma_T(\xi)$ given by (33). As we assume yearly portfolio revisions, we consider $T = 12$-monthly returns. To fit the finite state space, we match each $x_{k+1,x_k,j}(\xi)$ to the closest dividend yield in the discretized state space.

$$Q(t_k, x_k) = \max_{w_{k,x_k} \in [0,1]} \sum_{j=1}^{J} p_j(\xi) \left((1 + R_{k+1,x_k,j}(w_{k,x_k}, \xi))^{1-\gamma} Q(t_{k+1}, x_{k+1,x_k,j}) \right). \tag{29}$$

with

$$R_{k+1,x_k,j}(w, \xi) = (w_{k,x_k} e^{r_f + y_{k+1,x_k,j}(\xi)} + (1 - w_{k,x_k}) e^{r_f}) - 1.$$

We keep the stratified sample of residuals constant to fix the return distribution, albeit an approximation to the original distribution, to improve the comparison among alternative horizons, parameter values and investors.

The naive investor solves the backward recursion (29) conditional on $\xi = \hat{\xi}$. The Bayesian investor samples from the unconditional approximate return distribution as described in the previous section. The robust investor solves

$$\max_{w \in \mathcal{W}} \min_{\xi \in \mathcal{U}} \mathsf{E}_0 \left[\frac{{W_K}^{1-\gamma}}{1-\gamma} | \xi \right]. \tag{30}$$

To calculate the expectation, we rely on the stratified sample. In this case, a straightforward implementation of (30) demands the construction of scenario paths for each

[6] Without constant relative risk aversion, W is a state variable and enters the state space definition. The size of an accurate approximation of this multi-dimensional state space grows exponentially with the number of state variables and complicates the computations.

evaluated parameter configuration ξ. As the scenarios paths recombine (thanks to the finite state space), a backward recursion is a fast way to evaluate the objective of the inner minimization problem for w given. Therefore we define

$$Q(t_k, x_k, w, \xi) = \sum_{j=1}^{J} p_j(\xi) \left((1 + R_{k+1,x_k,j}(w,\xi))^{1-\gamma} Q(t_{k+1}, x_{k+1,j}(\xi), w, \xi) \middle| \, x_k, \xi \right). \tag{31}$$

Given a portfolio choice w and a parameter ξ, $W_0/Q(0, x_0, w, \xi)$ is equal to the objective value $\mathsf{E}_0[\frac{W_K^{1-\gamma}}{1-\gamma} | \, \xi]$ associated with portfolio choice w. Therefore problem (30) is equivalent to

$$\max_{w \in \mathcal{W}} \min_{\xi \in \mathcal{U}} W_0/Q(0, x_0, w, \xi). \tag{32}$$

Observe that for given w and ξ, $Q(0, x_0, w, \xi)$ is readily computed by (31). We use sequential optimization as explained in the previous section to solve (30) to reasonable accuracy.

Results

The *traditional* investor increases investment with the investment horizon and economic state. This horizon dependence is due to the intertemporal hedging demand described by described by Merton (1973). The return model features a negative contemporaneous correlation between realized return on the one side and current dividend yield and future expected return on the other side. Increases in the dividend yield are likely to decrease the realized return but also lead to an increase in the expected return. Hence risky investment provides higher wealth precisely when investment opportunities worsen. Observe that the dynamic investor invest less in the low economic states and more in the high economic states than a buy-and-hold investor. Obviously, the dynamic investor can adapt the portfolio as the expected dividend yield converges to its long term average, i.e. the investor may forego loss giving investment in the first year in the case of a low initial state and become active when the dividend yield increases. The sensitivity to the state implies volatile investments as the economy propagates through the economic states. The *Bayesian* investor keeps less risky assets than the traditional investor but invests considerably more than the buy-and-hold investor for long horizons and high initial states and less in the low initial states. The aforementioned logic applies. The investor will seize or await opportunities as they arise. Observe that also the Bayesian strategy is highly volatile.

Insert figure 7 Dynamic (52) about here.

Insert figure 9 Buy and Hold (85) about here.

The robust dynamic investor is less radical than the traditional and Bayesian investor. Investments vary moderately over time and initial states. Also the robust investor exploits the possibility to take corrective actions. For long horizons, she invests less than her buy-and-hold counterpart for low initial states and more in high initial states. Yet the differences as not as pronounced as for the non-robust investors. The portfolio remains therefore more stable over time and states. Again the worst case form of predictability (cf. figure 8) shows a similar structure as in the previous sections. [Not so sensitive to initial economic state? Confirm!]

Insert figure 8 WC Buy and Hold (52) about here (if at all).

6 Concluding remarks and further research

The robust buy-and-hold and dynamic portfolios are fundamentally different than traditional or Bayesian portfolios. The robust investor is only active if the investment opportunities are significant. Even then the robust investor is more prudent as the worst case may feature a weakened form of predictability. This also reduces the sensitivity on the economic state. Recognizing that predictability is uncertain, the robust investor is reluctant to fundamentally revise the portfolio when the economic state changes.

We can also identify which (plausible) forms of predictability are most detrimental to the portfolio performance. For short-term investment and low initial states, the worst case features small α_1 and β_2 that keep the short term dividend yield and return low. For high states, the exposure to the dividend yield is decisive and reduced accordingly. For long-term investment, the long-term average dividend yield and the exposure to the dividend yield are crucial and reduced accordingly. The worst case configurations only differ little for buy-and-hold and dynamic investors. If predictability turns out to be insignificant, then the worst case may look fundamentally different, e.g. a combination of a negative exposure to predictability combined with a large long-term average dividend yield.

Barberis (2000) also studies learning in the presence of estimation uncertainty. However learning about the true parameter value is incompatible with a robust approach to estimation uncertainty. A dynamic investment strategy which adapts to learning of the uncertain parameters should include the investor's belief about the parameters as state variables. The investor's belief is typically a function of the newly realized returns which depend on the true parameters. Hence a portfolio which is robust to parameter uncertainty it must also be robust to learning. References for learning are Brennan (1998), Xia (2001) and Kandel and Stambaugh (1996).

We confine to a normative analysis of robust investment. Next step would be to compare the empirical performance of traditional, Bayesian and robust portfolios. Moreover our study is based on a specific combination of return model and utility function. Although the basic results presumably hold for alternative return models, other state variables (see Barberis (2000)) and other utility functions that feature a risk-return tradeoff, generalizations to multi-period portfolio choice with multiple assets and utility functions that depend on intermediate consumption would be interesting and computationally challenging generalizations.

A Appendix

We are interested in the cumulative continuously compounded excess return over T months, $y^a_T = y_1 + y_2 + .. + y_T$, denoted by y_T henceforth. Propagating through the return model (3)-(4) and recognizing the arithmetic series, we can rewrite the return and the future state as:

$$\begin{aligned}
\begin{pmatrix} y_T \\ x_T \end{pmatrix} \mid x_0 \quad &\sim \quad N\left(\begin{pmatrix} \mu_T \\ \nu_T \end{pmatrix}, \Sigma_T\right) \qquad (33)\\
\begin{pmatrix} \mu_T \\ \nu_T \end{pmatrix} \quad &= \quad \begin{pmatrix} T\alpha_1 + \beta_1\rho\alpha_2 + \beta_1\chi x_0 \\ \chi\alpha_2 + \beta_2^T x_0 \end{pmatrix} \\
\Sigma_T \quad &= \quad \begin{pmatrix} T\sigma_1^2 + 2\beta_1\rho\sigma_{12} + \beta_1^2\phi\sigma_2^2 & \chi\sigma_{12} + (\chi_2 - \chi)/(\beta_2 - 1)\sigma_2^2 \\ \chi\sigma_{12} + (\chi_2 - \chi)/(\beta_2 - 1)\sigma_2^2 & \chi_2\sigma_2^2 \end{pmatrix}
\end{aligned}$$

given the arithmetic series

$$\begin{aligned}
\chi(T) &= \frac{\beta_2^T - 1}{\beta_2 - 1} \\
\chi_2(T) &= \frac{\beta_2^{2T} - 1}{\beta_2^2 - 1} \\
\rho(T) &= \sum_{t=1}^{T-1} (T - t)\beta_2^{t-1} = \frac{\beta_2^T - 1 + T(1 - \beta_2)}{(\beta_2 - 1)^2} \qquad (34)\\
\phi(T) &= \sum_{t=1}^{T-1} (\frac{\beta_2^{T-t} - 1}{\beta_2 - 1})^2 = \frac{T}{(\beta_2 - 1)^2} - 2\frac{\beta_2^T - 1}{(\beta_2 - 1)^3} + \frac{\beta_2^{2T} - 1}{(\beta_2^2 - 1)^3}
\end{aligned}$$

The geometric average cumulative T-month return is $\sqrt[T]{R_T}$ and is lognormally distributed with parameters[7] μ_T/T and σ_T^2/T^2 and, using the moments of the lognormal distribution, we derive the geometric average monthly Sharpe ratio

$$\text{Sh} = \frac{e^{r_f + \mu_T/T + 0.5\sigma_T^2/T^2} - e^{r_f}}{e^{r_f + \mu_T/T + 0.5\sigma_T^2/T^2}\sqrt{e^{\sigma_T^2/T^2} - 1}}$$

We use $e^x \approx 1 + x$ for small x to approximate the geometric average monthly Sharpe ratio by

$$\begin{aligned}
\text{Sh} &= \frac{(e^{\mu_T/T + 0.5\sigma_T^2/T^2} - 1)e^{r_f}}{e^{\mu_T/T + 0.5\sigma_T^2/T^2}e^{r_f}\sqrt{e^{\sigma_T^2/T^2} - 1}} \\
&= \frac{1 - e^{-\mu_T/T - 0.5\sigma_T^2/T^2}}{\sqrt{e^{\sigma_T^2/T^2} - 1}} \qquad (35)\\
&\approx \frac{\mu_T}{\sigma_T} + \frac{1}{2}\sigma_T/T.
\end{aligned}$$

[7] If $R_T \sim LOGN(\mu_T, \sigma_T^2)$, then $\sqrt[T]{R_T} \sim LOGN(\mu_T/T, \sigma_T^2/T^2)$.

If $\beta_2 < 1$, $\rho(T) = \frac{T-\chi(T)}{1-\beta_2}$ and

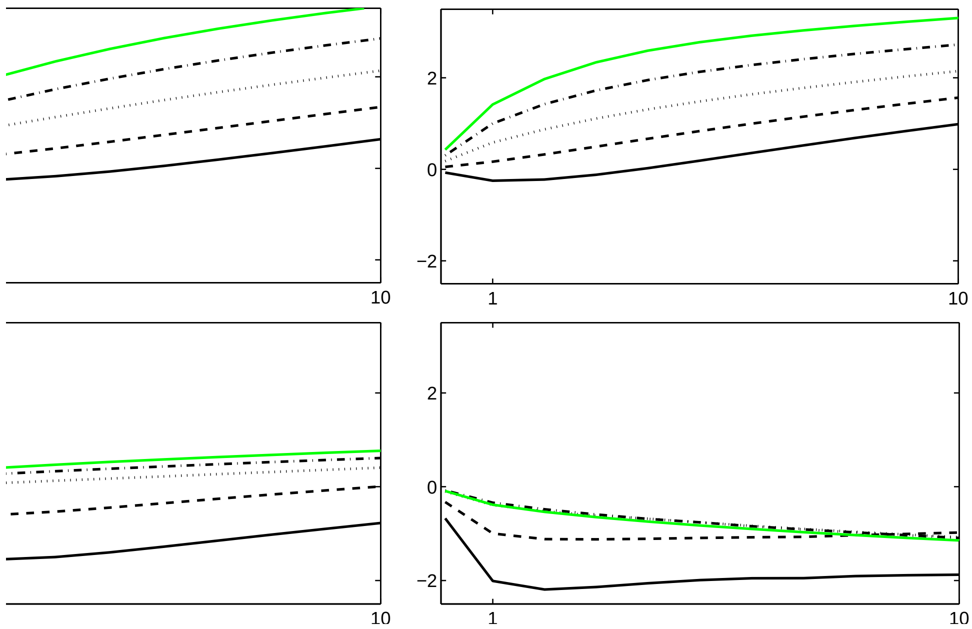
$$\begin{aligned}\mu_T &= T\alpha_1 + \beta_1\rho(T)\alpha_2 + (x_0 - \nu + \nu)\beta_1\chi(T) \\ &= T\alpha_1 + \beta_1\frac{T-\chi}{1-\beta_2}\alpha_2 + (x_0 - \nu + \frac{\alpha_2}{1-\beta_2})\beta_1\chi(T) \\ &= T(\alpha_1 + \frac{\beta_1}{1-\beta_2}\alpha_2) + (x_0 - \nu)\beta_1\chi(T).\end{aligned}$$

and

$$\sigma_T^2 = T\sigma_1^2 + 2\beta_1\rho(T)\sigma_{12} + \beta_1^2\phi(T)\sigma_2^2.$$

B Figures and tables

Figure 1: Expected performance as a function of the investment horizon



Notes: The figure describes the average monthly Sharpe ratio as a function of the investment horizon. The left and right panels correspond to the long (1952-1995) and short (1985-1995) datasets respectively. The upper and lower panels correspond to the traditional and robust ($\theta = 3.1$) evaluation of the Sharpe ratio. The curves within each graph correspond to different dividend yields: the long term mean dividend yield (dotted), the latter minus two standard deviation (solid black), minus one standard deviation (dashed), plus one standard deviation (dash-dot), plus two standard deviations (solid grey). dataset.

Table 1: Parameter estimates for a VAR model of stock returns

	1952-1995		1985-1995	
	μ (bps)	σ^2 (bps)	μ (bps)	σ^2 (bps)
	50	17	65	19
	(18)	(1)	(39)	(3)

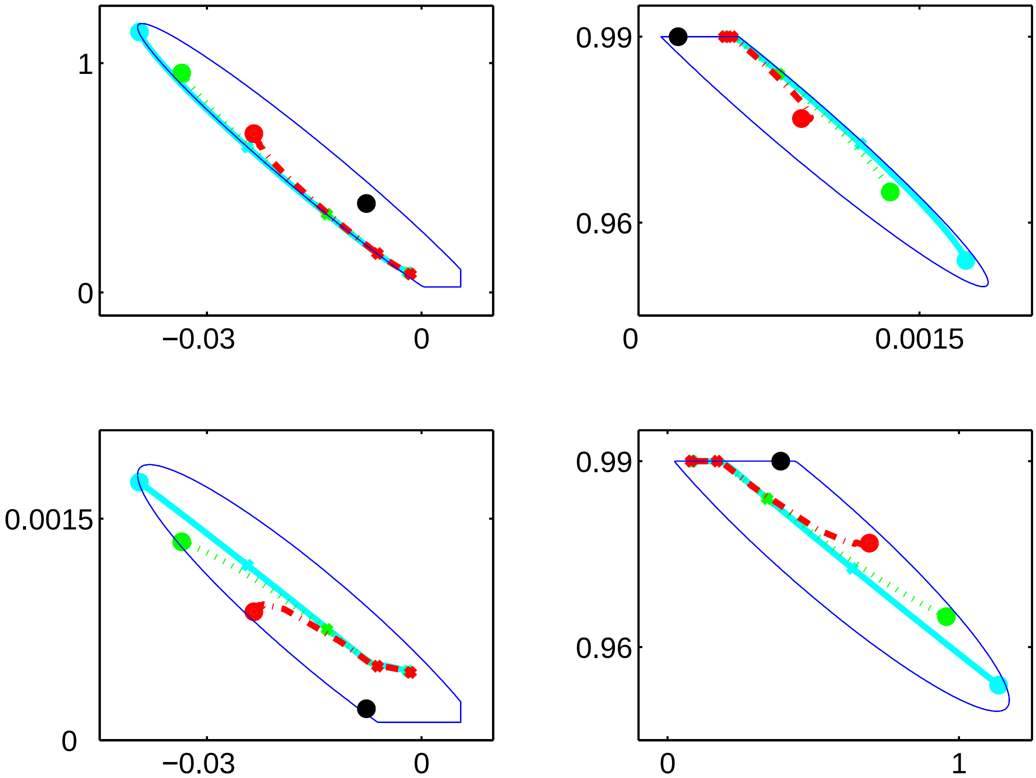
α (bps)	β	α (bps)	β
-143	0.52	-303	1.09
(81)	(0.21)	(281)	(0.83)
8	0.98	13	0.96
(3)	(0.01)	(10)	(0.03)

Σ (1952-1995)		Σ (1985-1995)	
0.0017	$\mathbf{-0.9351}$	0.0019	$\mathbf{-0.9323}$
(0.0001)	(0.0055)	(0.0003)	(0.0122)
	3.0E-6		2.6E-6
	(1.9E-7)		(3.4E-7)

	α_1	β_1	α_2	β_2	α_1	β_1	α_2	β_2
α_1	0.0081	$\mathbf{-0.98}$	$\mathbf{-0.94}$	**0.91**	0.0281	$\mathbf{-0.99}$	$\mathbf{-0.93}$	**0.92**
β_1		0.2129	**0.91**	$\mathbf{-0.94}$		0.8265	**0.92**	$\mathbf{-0.93}$
α_2			0.0003	$\mathbf{-0.98}$			0.0010	$\mathbf{-0.99}$
β_2				0.0091				0.0305

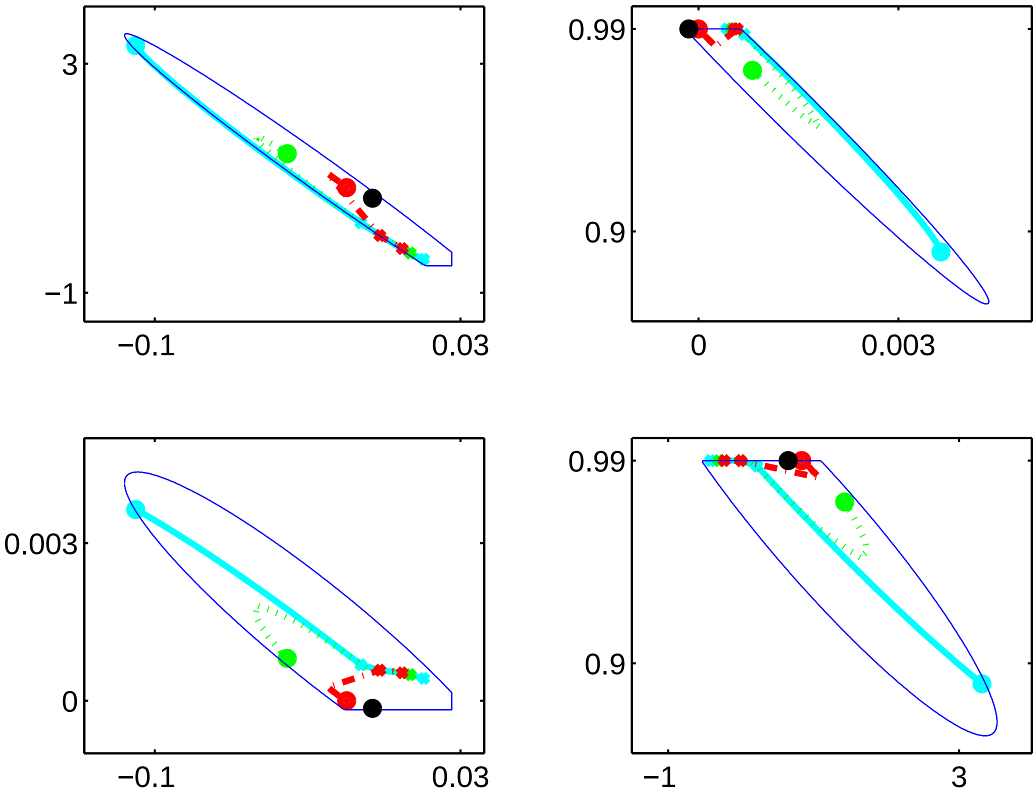
Source Barberis (2000). The results in the upper panel are based on the model $r_t = \mu + \varepsilon_t$, where r_t is the continuously compounded excess stock index return in month t and $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$. The results in the middle panels are based on the return model (3)-(4). The lower panel describes the parameter uncertainty by the uncertainty matrix Ω corresponding to (12). The values in bold above the diagonal denote correlations. The left and right panels correspond to the long and short datasets respectively.

Figure 2: Worst case parameter configurations



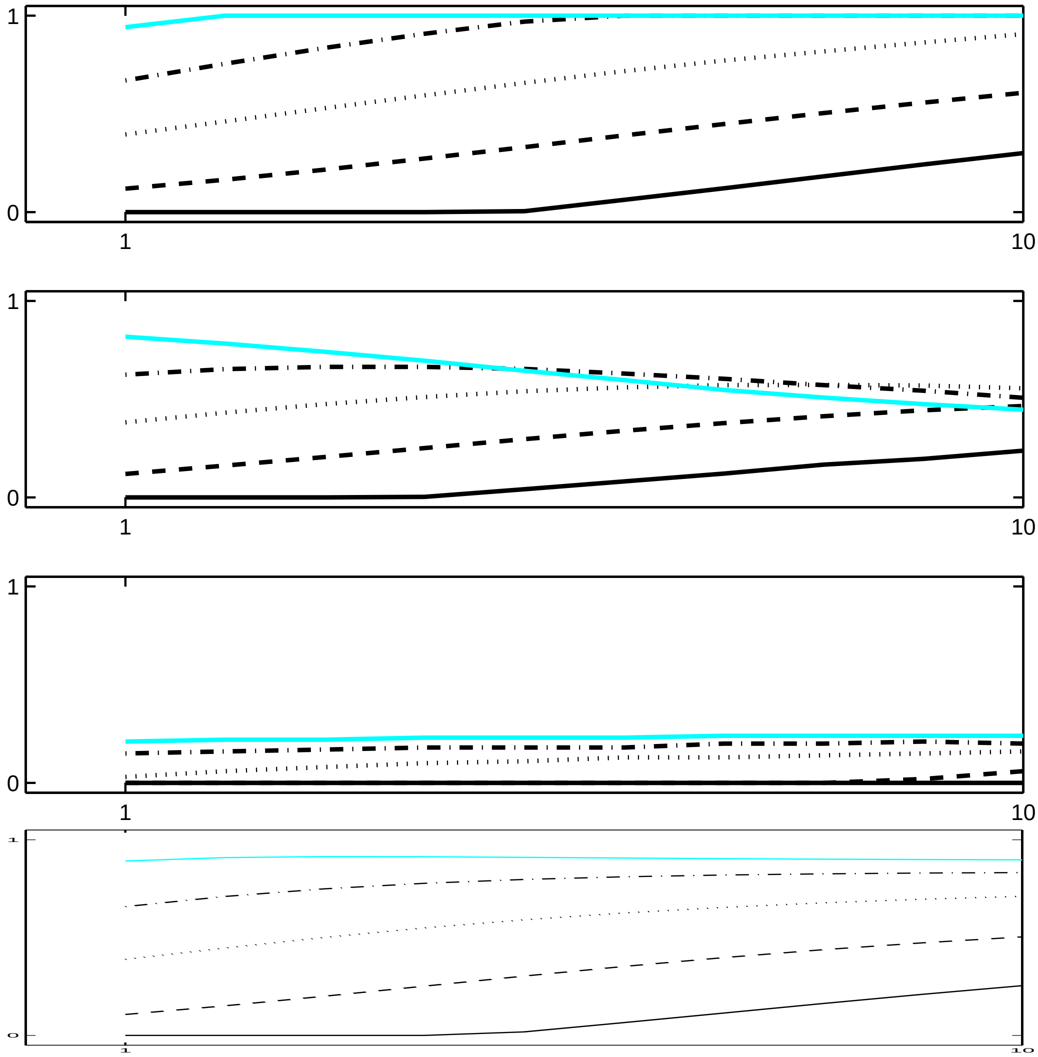
Notes: The figure shows a parametric plot of the worst parameter configurations for the average monthly Sharpe ratio. Results are based on return model (3)-(4) estimated on the 1952-1995 dataset. The panels correspond to projections of the parameter space on two dimensional subsets of parameters: (α_1, β_1) (upper-left), (α_2, β_2) (upper-right), (α_1, α_2) (lower-left) and (β_1, β_2) (lower-right). The truncated ellipsoid within each figure presents the projection of the uncertainty set (13) for $\theta = 3.1$ on the relevant parameter space. Each curve connects the worst case parameter configurations for values of the initial dividend yield between $[0.01, 0.06]$ and corresponds to a given investment horizon: $T = 12$ months (solid grey), $T = 60$ months (dotted grey), $T = 120$ months (dash-dot) and $T = 1200$ (black dot) months. The worst case parameter configuration associated with the smallest initial dividend yield is indicated by a dot.

Figure 3: Worst case parameter configurations



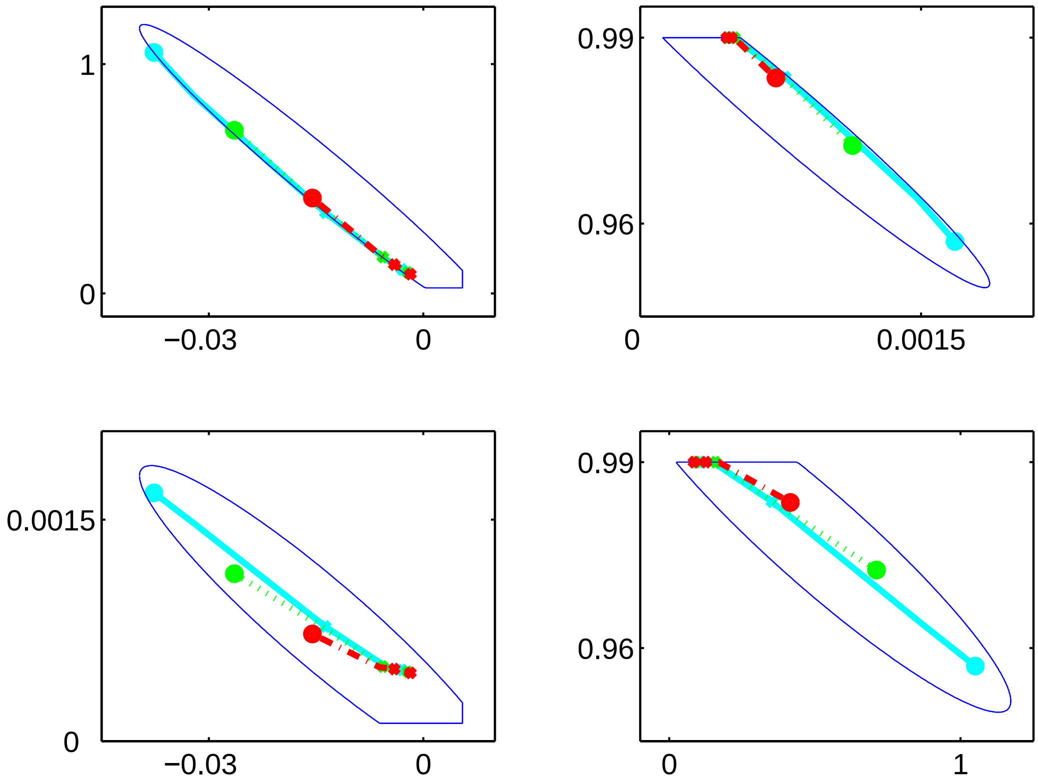
Notes: The figure shows a parametric plot of the worst parameter configurations for the average monthly Sharpe ratio. Results are based on return model (3)-(4) estimated on the 1985-1995 dataset. The panels correspond to projections of the parameter space on two dimensional subsets of parameters: (α_1, β_1) (upper-left), (α_2, β_2) (upper-right), (α_1, α_2) (lower-left) and (β_1, β_2) (lower-right). The truncated ellipsoid within each figure presents the projection of the uncertainty set (13) for $\theta = 3.1$ on the relevant parameter space. Each curve connects the worst case parameter configurations for values of the initial dividend yield between $[0.01, 0.06]$ and corresponds to a given investment horizon: $T = 12$ months (solid grey), $T = 60$ months (dotted grey), $T = 120$ months (dash-dot) and $T = 1200$ (black dot) months. The worst case parameter configuration associated with the smallest initial dividend yield is indicated by a dot.

Figure 4: Optimal investment as a function of the investment horizon in years



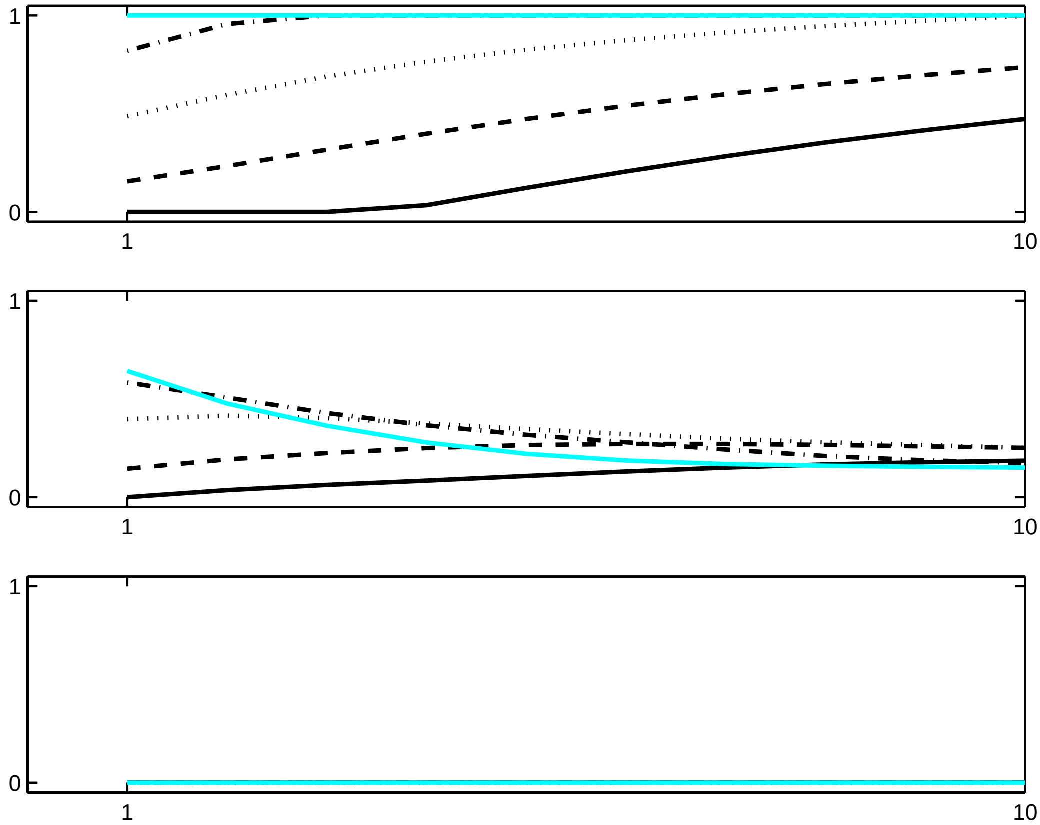
Notes: The figure describes the optimal buy–and–hold strategy of an investor who maximizes power utility function (1) over terminal wealth. The returns are estimated and predicted using (3)-(4) on the 1952-1995 dataset. The three upper panels describe the initial asset allocation as a function of the investment horizon in years, for a naive investor, a Bayesian investor and a robust investor respectively. The curves within each graph correspond to alternative values of the initial dividend yield x_0: $x_0 = 2.06\%$ (solid), $x_0 = 2.91\%$ (dashed), $x_0 = 3.75\%$ (dotted), $x_0 = 4.59\%$ (dash-dot) and $x_0 = 5.43\%$ (solid grey). The lower panel depicts the Bayesian buy-and-hold strategy assuming a posterior distribution that dogmatically sets $\beta_2 < 1$.

Figure 5: Worst case parameter configuration



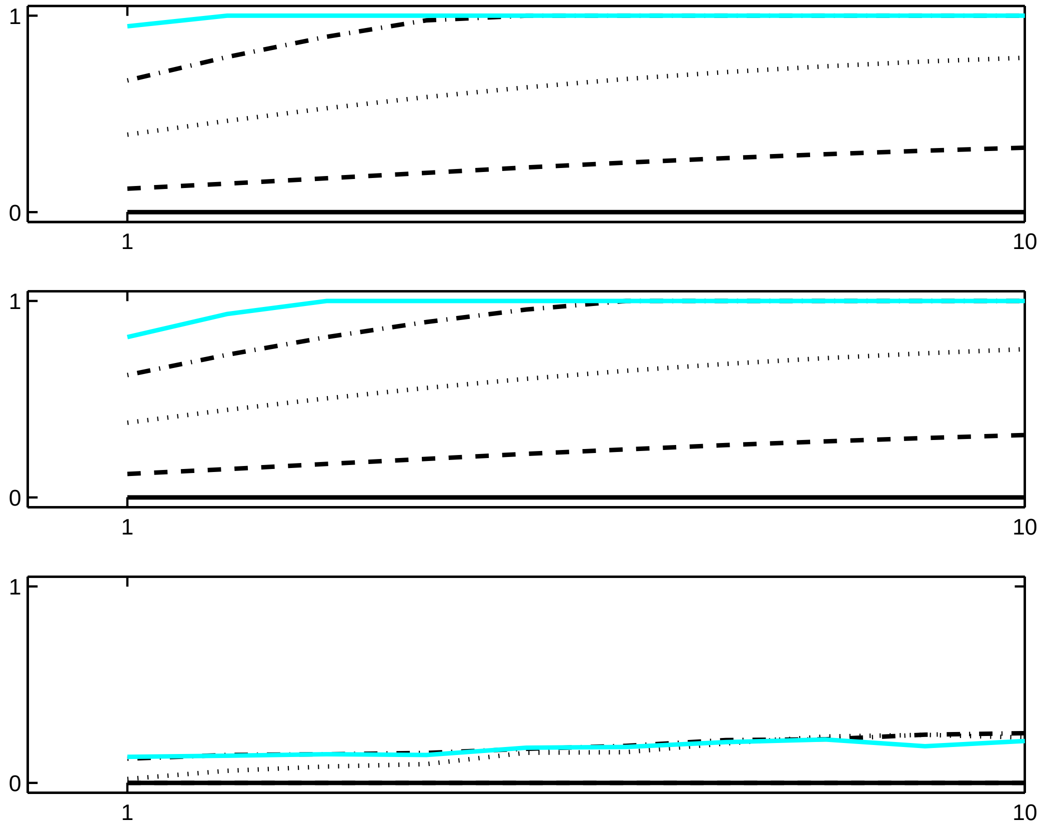
Notes: The figure describes the worst case parameter configurations for a buy–and–hold investor who maximizes power utility function (1) over terminal wealth. The returns are estimated and predicted using model (3)-(4) on the 1952-1995 dataset. The panels correspond to projections of the parameter space on two dimensional subsets of parameters: (α_1, β_1) (upper-left), (α_2, β_2) (upper-right), (α_1, α_2) (lower-left) and (β_1, β_2) (lower-right). The truncated ellipsoid within each figure presents the projection of the uncertainty set (13) for $\theta = 3.1$ on the relevant parameter space. Each curve connects the worst case parameter configurations for values of the initial dividend yield between 2.06% and 5.43% and corresponds to a given investment horizon: $T = 12$ months (solid grey), $T = 60$ months (dotted grey), $T = 120$ months (dash-dot). The worst case parameter configuration associated with the smallest initial dividend yield is indicated by a dot.

Figure 6: Optimal asset allocation as a function of the investment horizon in years



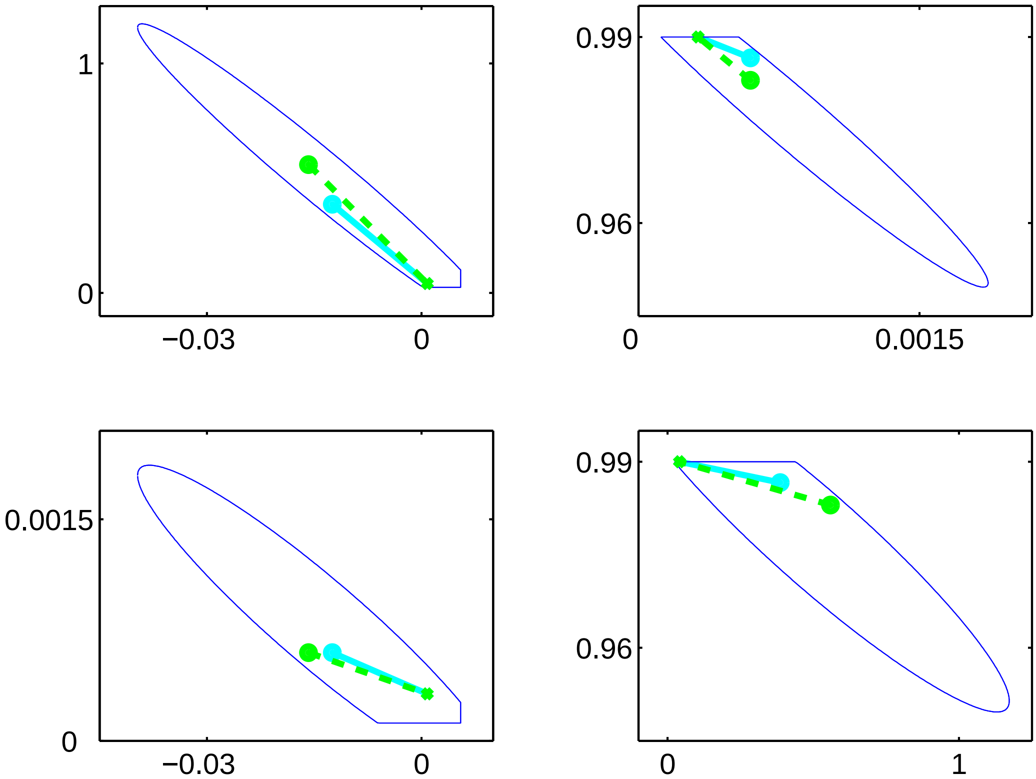
Notes: The figure describes the optimal buy–and–hold strategy of an investor who maximizes power utility function (1) over terminal wealth. The returns are estimated and predicted using model (3)-(4) on the 1985-1995 dataset. The three panels describe the initial asset allocation as a function of the investment horizon in years, for a naive investor (upper panel), a Bayesian investor (middle panel) and a robust investor (lower panel). The curves within each graph correspond to alternative values of the initial dividend yield x_0:$x_0 = 2.36\%$ (solid), $x_0 = 2.86\%$ (dashed), $x_0 = 3.36\%$ (dotted), $x_0 = 3.86\%$ (dash-dot) and $x_0 = 4.36\%$ (solid grey).

Figure 7: Optimal asset allocation as a function of the investment horizon in years



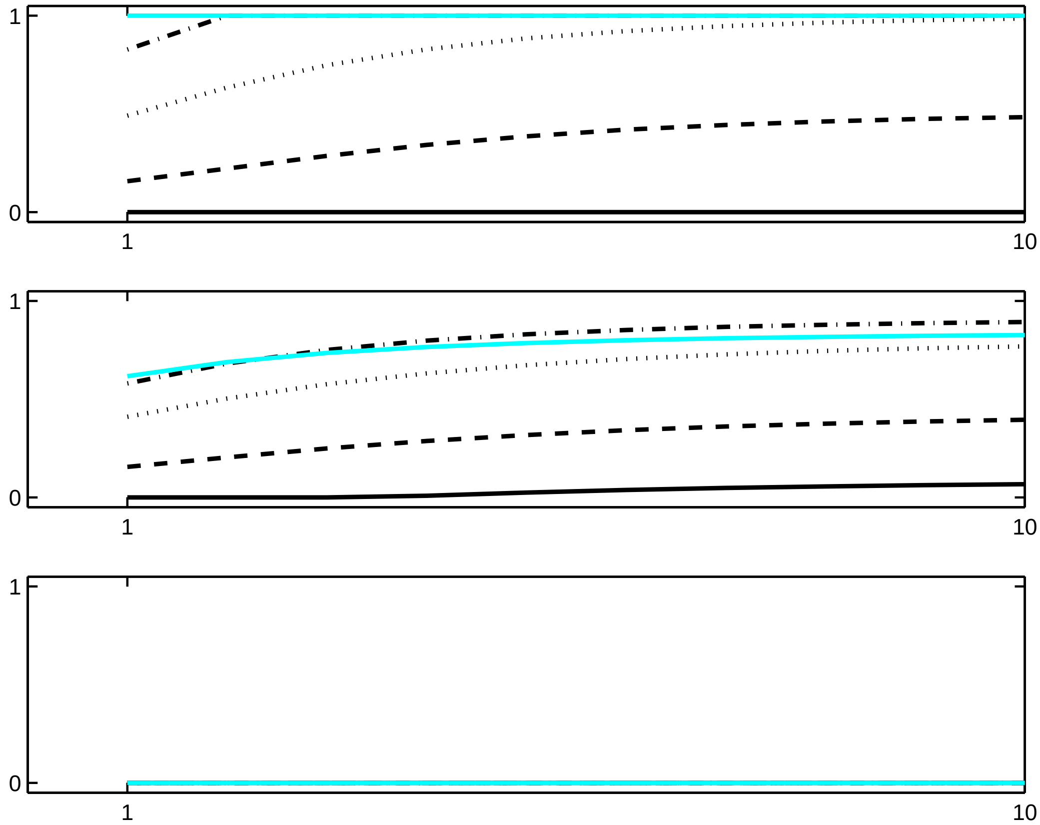
Notes: The figure shows the optimal strategy of an investor who rebalances once a year and who maximizes power utility function (1) over terminal wealth. Returns are estimated and predicted using model (3)-(4) on the 1952-1995 dataset. The three panels describe the initial asset allocation as a function of the investment horizon in years, for a naive investor (upper panel), a Bayesian investor (middle panel) and a robust investor (lower panel). The curves within each graph correspond to alternative values of the initial dividend yield x_0: $x_0 = 2.06\%$ (solid), $x_0 = 2.91\%$ (dashed), $x_0 = 3.75\%$ (dotted), $x_0 = 4.59\%$ (dash-dot) and $x_0 = 5.43\%$ (solid grey).

Figure 8: Worst case parameter configuration



Notes: The figure describes the worst case parameter configurations for a dynamic investor who maximizes power utility function (1) over terminal wealth. The returns are estimated and predicted using model (3)-(4) on the 1952-1995 dataset. The panels correspond to projections of the parameter space on two dimensional subsets of parameters: (α_1, β_1) (upper-left), (α_2, β_2) (upper-right), (α_1, α_2) (lower-left) and (β_1, β_2) (lower-right). The truncated ellipsoid within each figure presents the projection of the uncertainty set (13) for $\theta = 3.1$ on the relevant parameter space. Each curve connects the worst case parameter configurations for values of the initial dividend yield between 2.06% and 5.43% and corresponds to a given investment horizon: $T = 60$ months (dashed grey), $T = 120$ months (solid). The worst case parameter configuration associated with the smallest initial dividend yield is indicated by a dot.

Figure 9: Optimal asset allocation as a function of the investment horizon in years



Notes: The figure shows the optimal strategy of an investor who rebalances once a year and who maximizes power utility function (1) over terminal wealth. Returns are estimated and predicted using model (3)-(4) on the 1985-1995 dataset. The three panels describe the initial asset allocation as a function of the investment horizon in years, for a naive investor (upper panel), a Bayesian investor (middle panel) and a robust investor (lower panel). The curves within each graph correspond to alternative values of the initial dividend yield x_0: $x_0 = 2.36\%$ (solid), $x_0 = 2.86\%$ (dashed), $x_0 = 3.36\%$ (dotted), $x_0 = 3.86\%$ (dash-dot) and $x_0 = 4.36\%$ (solid grey).

References

Anderson, E., Hansen, L. and Sargent, T.: 1999. Robustness, detection, and the price of risk. Working paper, Stanford University.

Ang, A. and Bekaert, G.: 2005. Stock return predictability: Is it there?. Manuscript, Colombia University.

Barberis, N. 2000. Investing for the long run when returns are predictable. *The Journal of Finance* **55**(1), pp. 225–264.

Brennan, M. 1998. The role of learning in dynamic portfolio decisions. *European Financial Review* **1**, pp. 295–306.

Brennan, M. J. and Xia, Y.: 2005. Persistence, predictability and portfolio planning. Working Paper.

Campbell, J. and Viceira, L.: 2002. *Strategic asset allocation: portfolio choice for long term investors*. Oxford University Press.

Cavadini, F., Sbuelz, A. and Trojani, F.: 2001. A simplified way of incorporating model risk, estimation risk and robustness in mean variance portfolio management. Working Paper CentER, Tilburg University, The Netherlands.

Chopra, V. and Ziemba, W. 1993. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management* **19**(2), pp. 6–12.

Cochrane, J.: 2006. The dog that did not bark: A defense of return predictability. Working Paper.

Cochrane, J. H.: 2004. *Asset Pricing*. Princeton University Press.

Fama, E. and French, K. 1988. Dividend yields and expected stock returns. *Journal of Financial Economics* **23**, pp. 3–25.

Gilboa, I. and Schmeidler, D. 1989. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* **18**(2), pp. 141–153.

Goetzman, W. N. and Jorion, P. 1993. Testing the predictive power of dividend yields. *Journal of Finance* **48**, pp. 663–679.

Goyal, A. and Welch, I. 2003. Predicting the equity premium with dividend ratios. *Management Science* **49**, 639–654.

Goyal, A. and Welch, I.: 2005. A comprehensive look at the empricial performance of equity premium prediction. Manuscript.

Jagannathan, R. and Ma, T. 2003. Risk reduction in large portfolios – why imposing wrong constraints helps. *Journal of Finance* **58**(4), pp. 1651–1684.

Kandel, S. and Stambaugh, R. 1996. On the predictability of stoch returns: an asset-allocation perspective. *The Journal of Finance* **51**, pp. 385–424.

Kogan, L. and Wang, T.: 2002. A simple theory of asset pricing under model uncertainty. Working paper, University of British Columbia.

Lo, A. W. and MacKinlay, C. A. 1988. Stock market prices do not follow random walks: evidence from a simple specification test.. *Review of Financial Studies* pp. 41–66.

Lutgens, F. and Schotman, P. C.: 2006. Robust portfolio optimization with multiple experts. METEOR working paper, Maastricht University.

Maenhout, P.: 1999. *Robust portfolio rules and asset pricing*. PhD thesis. Harvard University.

Merton, R. 1973. An intertemporal asset pricing model. *Econometrica* **41**, pp. 867–887.

Nilim, A. and El Ghaoui, L.: 2002. Robust solutions to Markov decision problems. Presentation at NSF-ADP Workshop.

Pathak, P. 2002. Notes on robust portfolio choice. *Discussion paper, Harvard University.*

Samuelson, P. 1973. Proof that properly discounted present values of asset vibrate randomly. *Bell Journal of Economics and Management Science* **4**, 369–374.

Shanken, J. and Tamayo, A. 1964. Risk, mispricing and asset allocation: conditioning on divided yield. *NBER Working Paper.*

Shiller, R. J. 1981. Do stock prices move too much to be justified by subsequent chages in dividends?. *American Economic Review* **71**, 421–436.

Sims, C. 2001. Pitfalls of minimax approaches to model uncertainty. *American Economic Review* **91**, pp. 51–54.

Uppal, R. and Wang, T. 2003. Model misspecification and underdiversification. *Journal of Finance* **58**(6), pp. 2465–2486.

Wachter, J. A. and Warusawitharana, M.: 2006. Predictable returns and asset allocation: Should a sceptical investor time the market?. Working Paper.

Xia, Y. 2001. Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation.. *Journal of Finance* **56**(1), pp. 205–246.