Network for Studies on Pensions, Aging and Retirement

Netspar DISCUSSION PAPERS

Roy Hoevenaars, Roderick Molenaar, Peter Schotman and Tom Steenkamp

Strategic Asset Allocation for Long-Term Investors

Parameter Uncertainty and Prior Information

Discussion Paper 04/2011-038

Strategic Asset Allocation for Long-Term Investors: Parameter Uncertainty and Prior Information

Roy P.P.M. Hoevenaars
APG Asset Management

Roderick Molenaar
APG Asset Management

Peter C. Schotman
Maastricht University and Netspar

Tom Steenkamp
Free University Amsterdam and Robeco

This version: April 4, 2011

Abstract: We study the effect of parameter uncertainty on the long-run risk of three alternative asset classes: equity, nominal bonds and short-term T-Bills. We estimate the long-run risk as the annualized predictive variance of returns at different horizons implied by a vector autoregression using alternative Bayesian priors. Under an uninformative prior we conclude that not only equity becomes more risky relative to estimates that are conditional on known parameter values. The long-run risk of long- and short-term bonds increases proportionally with the same factor. Correlations among returns appear robust against parameter uncertainty.

Alternative informative priors imply large differences in expected returns, which lead to different optimal portfolios. To limit the effect of a single prior we derive a robust portfolio rule that associates a portfolio with the worst prior for that portfolio. The optimal robust portfolio appears well-diversified and stable with respect to the investment horizon.

Keywords: strategic asset allocation, Bayesian vector autoregression, parameter uncertainty, robust portfolio choice

JEL codes: C32, G11, C11

Corresponding author: Peter Schotman, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands, tel: (+31) 43 3883838, fax: (+31) 43 3884875, email: `p.schotman@maastrichtuniversity.nl`

* We thank Michael Brandt, Ralph Koijen, Franz Palm, and Luis Viceira for constructive comments on an earlier version, which was presented at the Econometric Society European Meeting, European Quantitative Forum in London, a Netspar pension workshop, and the CREATES workshop on Dynamic Asset Allocation in Aarhus. We thank participants for their comments.

** The views expressed in this paper are those of the authors and do not necessarily reflect those of our employer and our colleagues.

1 Introduction

One of the main results in the empirical literature on strategic asset allocation is that equity is less risky for long-term investors because of its mean reverting dynamics.[1] An important counterargument is that parameter estimates in the empirical models are subject to substantial uncertainty. Bayesian studies by Barberis (2000) and Pastor and Stambaugh (2010) show that accounting for parameter uncertainty raises the long-run risk of equity and this may be quantitatively more important than mean reversion. They explain that uncertainty about the expected return of equity is particularly important for long-run investors, since the estimation error is perfectly correlated over time and therefore quickly increases with the investment horizon.

The implications for portfolio choice are not clear, however. Both studies, as well as Avramov (2002) and Brandt *et al* (2005), among others, assume that long-term investors can also invest in a real riskfree asset. When a real long-term bond is not available, a long-term investor's alternative to equity consists mainly of nominal bonds of various maturities. Since the real returns on nominal bonds are subject to inflation and interest rate risk, parameter uncertainty will also increase their long-run risk. For portfolio choice we need to know the impact of parameter uncertainty on all variances and covariances.

To compare the impact of parameter uncertainty on different asset classes we use the term structure of risk introduced by Campbell and Viceira (2005) as a useful summary statistic. The term structure of risk is defined as the per period variance of cumulative log returns: $\frac{1}{k}\mathrm{Var}(R_{t+k}^{(k)})$. Campbell and Viceira (2005) estimate the risk of stocks, bonds and bills at different horizons using a vector autoregression with returns and predictor state variables. Their risk estimates are conditional on the parameter estimates of the VAR. Instead of conditioning on the estimated parameters we take a Bayesian approach and explicitly account for parameter uncertainty. Using the same VAR specification as Campbell and Viceira (2005) we find that parameter uncertainty adds substantially to the real long-run risk of all three asset classes. The effects are strongest at longer horizons and affect variances as well as covariances. With uninformative priors, the correlations among returns are hardly affected by the parameter uncertainty.

[1] See, for example, Campbell *et al* (2003).

As informative priors we consider alternative views on the unconditional means of returns and state variables in the VAR. Not surprisingly, more informative priors mostly imply smaller effects of parameter uncertainty. But informative priors that move expected returns away from the sample mean also increase the term structure of risk relative to maximum likelihood estimates of the VAR model of returns. For example, when the prior view is informative about a low long-run dividend yield, this also affects the vector autoregressive dynamics by shifting more probability mass to near unit root behavior of the dividend yield, which in turn increases the long-term predictive variance of equity returns.

Optimal portfolios are even more sensitive to the prior. At the 15 years horizon the conditional portfolio is almost fully in equity, reflecting both the large historical equity premium and strong mean reversion. Optimal Bayesian portfolios imply a much reduced weight for equity, but the differences across alternative priors are large. An extreme case is a pessimistic prior on long-term equity returns, under which the correlation between bond and equity returns becomes negative and which thus induces a large allocation to nominal bonds instead of equity.

The differences are also large in certainty equivalent utility terms. For example, we find that the optimal portfolio of a moderately optimistic investor entails a cost of 1.6% per year in certainty equivalent wealth over the 15 years horizon from the perspective of a moderately pessimistic investor. In other words, the pessimistic investor would be willing to give up 1.6% of his wealth per year to avoid the risky equity position of the optimistic investor. Conversely, the optimistic investor attaches a cost of 1.0% per year to the pessimistic portfolio. Costs are not symmetric because the utility function is concave. The costs for the optimist are therefore less than for the pessimist.

As the sensitivity to the prior appears economically important, we consider a robust portfolio strategy that can deal with multiple priors and limits the effects of any single prior. We assume that portfolio decisions are made by an investment committee. The members of the committee have different prior beliefs about the long-run outlook for different assets. Combining the different priors with the data evidence each committee member obtains a different predictive distribution of returns. The optimal portfolio for each individual committee member will be suboptimal according to the views of the other members and therefore induces an expected utility loss. To reach consensus the final optimal portfolio minimizes the maximum expected loss for all committee members. We call this the robust portfolio.

In our case the robust portfolio is based on a prior with moderately pessimistic views about equity. The robust portfolio has a substantial allocation to long-term nominal bonds. This is in contrast with most of the optimal portfolios based on individual priors, in which nominal bonds are usually dominated by equity for its high expected returns and short-term T-Bills for their low risk. Equity is costly, however, for investors with a more pessimistic view on the equity premium. Due to the asymmetry of the costs long-term nominal bonds are less costly when evaluated under alternative priors.

The portfolio selection model is motivated by the robust portfolio choice models of Rustem *et al* (2000), Wang (2005) and Lutgens and Schotman (2010), but differs in one crucial element. The existing robust methods assume there is a single investor with ambiguity aversion. Such an investor selects the portfolio that maximizes the worst case expected utility. Our committee members are subject to heterogeneous beliefs and wish to minimize the expected loss relative to their personal optimal portfolio. As a result the optimal portfolio is not based on the worst case over all priors, but on the worst case relative to each committee member's beliefs.

The robust approach to portfolio selection is Bayesian but also distinct from the Bayesian model averaging methodology used in Avramov (2002). Avramov (2002) first obtains the overall predictive distribution of returns over different investment horizons by averaging over all possible models according to the posterior weight of each model. This requires prior probabilities for different informative priors for the model parameters. Instead the robust approach relies on the expected utility losses related to alternative priors.

Our work is related to various strands in the literature on Bayesian portfolio choice. In their literature survey Avramov and Zhou (2010) point at three distinct attractive features of the Bayesian approach. First, Bayesian procedures explicitly account for the parameter uncertainty, which increases the riskiness of returns. Second, informative Bayesian priors shrink sample moments to theoretically plausible targets and this improves the out-of-sample accuracy of predictions and the performance of portfolios. Third, numerical algorithms for Bayesian analysis and decision making are fast and intuitive. Above we already discussed the effects of parameter uncertainty on risk estimates. The other two points are strongly related to the choice of priors.

We formulate informative priors about the unconditional means of returns and state variables in the vector autoregression. Maximum likelihood estimates of the parameters result in optimal portfolios that are heavily tilted towards equity because of the high equity premium and the strong mean reversion in the historical data. Yet historical time series estimates of these parameters are very imprecise.[2] Formal Bayesian models can add subjective prior views to support the weak sample evidence. Priors on the expected returns have been applied successfully by, among others, Pastor and Stambaugh (2000), Avramov (2004), Jorion (1986) and Black and Litterman (1992). These studies put much prior structure on the cross-section of expected returns. Since our cross-section only contains three asset classes, it does not lend itself to further structure.

Villani (2005, 2009) provides an econometric motivation for an informative prior on the long-run means in a vector autoregression. A conjugate informative prior is tractable for simulation by Gibbs sampling. More importantly, an informative prior on the unconditional mean of very persistent state variables is shown to reduce the probability of drawing explosive roots for the VAR. Another motivation for a prior on the unconditional means stems the common practice for long-term investors to use historical data to estimate volatilities and correlations of the long-term future return distribution, and use economic theory and current market circumstances to form their view about the long-term mean.

Other Bayesian studies have proposed a prior on the amount of predictability, like the R^2 in the return prediction regressions.[3] While providing valuable prior information, these priors are difficult to generalize to a full VAR model.[4] Important differences between our work and many of the earlier studies is our focus on long-horizon returns and the riskiness of the benchmark asset. The long-term investor in our model can not invest in a long-term real riskfree asset and thus faces uncertainty about the returns of all asset classes.

The remainder of this paper is organized as follows. Section 2 analyzes the effect of

[2] With 50 years of excess returns the sample mean is about 5% with a standard deviation of 18%. That implies a standard error of the mean equal to $18/\sqrt{50} = 2.5\%$ and a classical 95% confidence interval ranging from 0 to 10%. The predictability is highly debated by Goyal and Welch (2008).

[3] See, among others, Wachter and Warusawitharana (2009), Cremers (2002), Shanken and Tamayo (2004), Avramov (2004), Xia (2001), Kandel and Stambaugh (1996), Johannes, Korteweg and Polson (2008).

[4] The closest to a full VAR is Wachter and Warusawitharana (2009). In their most general specification they allow multiple asset classes and multiple predictor variables, but they do not allow feedback from returns to the predictor variables.

Bayesian parameter uncertainty on measures of long-run risk and the composition of long-term portfolios. We then introduce multiple priors and analyze decision making within an investment committee. In section 4 we describe the vector autoregressive model of return dynamics and the details of the Bayesian estimation and simulation of the model. Section 5 contains the details of the data and the most salient estimation results. The implications for long-term risk are discussed in 6. We then come to the implications for robust long-term portfolio choice in section 7. Finally, section 8 concludes.

2 Term structure of risk

Let r_t be a vector of log real returns, more specifically the (3×1) vector with real returns on a short term T-Bill, long-term bonds, and stocks, respectively,

$$r_t = \begin{pmatrix} r_{tb,t} \\ r_{b,t} \\ r_{s,t} \end{pmatrix} \tag{1}$$

For much of the analysis we will be interested in cumulative returns over k periods,

$$R^{(k)}_{i,t+k} = \sum_{j=1}^{k} r_{i,t+j} \tag{2}$$

Return dynamics will be modeled through a first order vector autoregression

$$y_t = \mu + \mathbf{B}(y_{t-1} - \mu) + \epsilon_t \tag{3}$$

containing the variables

$$y_t = \begin{pmatrix} r_t \\ s_t \end{pmatrix},$$

with s_t a vector of state variables that help predict returns, and ϵ_t a vector of shocks with zero mean and covariance matrix Σ. The VAR has parameter vector $\theta' = (\mu' \ \mathrm{vec}(\mathbf{B})' \ \mathrm{vech}(\Sigma)')$ and implies conditional return distributions

$$p(R^{(k)}_{t+k} | y_t, \theta), \tag{4}$$

and more specifically $p(r_{t+1} | y_t, \theta)$ for $k = 1$.

For inference we use Bayesian methods and define a prior $p(\theta)$. Adding historical sample data $\mathbf{Y}$ we obtain the posterior $p(\theta|\mathbf{Y})$ and the predictive distributions

$$p(R_{t+k}^{(k)}|y_t) = \int p(R_{t+k}^{(k)}|y_t, \theta)\ p(\theta|\mathbf{Y})\ \mathrm{d}\theta, \tag{5}$$

We use these quantities to estimate long-run risk and construct portfolios for long-run investors.

Our measure of long-run risk is the per period variance of returns as in Campbell and Viceira (2005). As a function of the investment horizon k it is called the term structure of risk, which for each asset class i is defined as

$$V_i(k) = \frac{1}{k}\mathsf{Var}_t\left[R_{i,t+k}^{(k)}\right] \tag{6}$$

In Campbell and Viceira (2005) this variance is computed conditional on the estimated VAR parameters $\hat{\theta}$ using the conditional densities $p(R_{t+k}^{(k)}|y_t, \theta = \hat{\theta})$. In the Bayesian framework we account for parameter uncertainty and compute the variances using the predictive densities $p(R_{t+k}^{(k)}|y_t)$ in (5).

In practice the term structure of risk is computed numerically. We draw parameters $\theta^{(\ell)}$ from the posterior density $p(\theta|\mathbf{Y})$. For each draw $\theta^{(\ell)}$ ($\ell = 1, \dots, L$) we compute the conditional means and variances

$$E_{i\ell}(k) = \mathsf{E}_t\left[R_{i,t+k}^{(k)}|\theta^{(\ell)}, y\right], \tag{7}$$

$$V_{i\ell}(k) = \mathsf{Var}_t\left[R_{i,t+k}^{(k)}|\theta^{(\ell)}\right], \tag{8}$$

where the conditional mean is evaluated at some fixed value $y_t = y$.[5] The term structure of risk is obtained by averaging over the parameters using the standard formula

$$\bar{V}_i(k) = \frac{1}{L}\sum_{\ell} V_{i\ell}(k) + \frac{1}{L}\sum_{\ell}\left(E_{i\ell}(k) - \bar{E}_i(k)\right)^2, \tag{9}$$

where

$$\bar{E}_i(k) = \frac{1}{L}\sum_{\ell} E_{i\ell}(k).$$

The second term in (9) is the variance of the conditional mean and adds positively to the overall risk of each asset class. The Bayesian term structure of risk will therefore generally be larger than the conditional variance reported by Campbell and Viceira (2005). The term

[5] In a linear vector autoregression the conditional variance $V_{i\ell}(k)$ does not depend on y.

structure of risk also depends on the prior. Different informative priors affect estimates of the first term. In addition more informative priors lead to less mean uncertainty and a smaller second term.

To obtain intuition on the magnitude of effects of parameter uncertainty we analyze three stylized examples that each highlight one particular aspect of the parameter uncertainty. The empirical sections deal with a Bayesian analysis of the full VAR.

2.1 Uncertainty about the unconditional mean

The uncertainty about the conditional mean will generally increase with k and eventually become the dominating component. This can be seen most easily in the simple case of a single risky asset with return

$$r_t = \mu + \epsilon_t \tag{10}$$

with ϵ_t white noise with zero mean and known variance σ^2. Parameter uncertainty in μ is summarized by the posterior

$$p(\mu|\mathbf{Y}) \sim \mathsf{N}(\hat{\mu}, \omega^2) \tag{11}$$

Conditional on $\hat{\mu}$ the term structure is flat at

$$V(k|\hat{\mu}) = \sigma^2 \tag{12}$$

Since cumulative returns

$$R_{t+k}^{(k)} = k\mu + \sum_{j=1}^{k} \epsilon_{t+j} = k\hat{\mu} + k(\mu - \hat{\mu}) + \sum_{j=1}^{k} \epsilon_{t+j} \tag{13}$$

also contain the cumulative parameter uncertainty $k(\mu - \hat{\mu})$, the unconditional term structure is

$$\bar{V}(k) = \sigma^2 + k\omega^2, \tag{14}$$

which is linearly increasing in k. The long-run risk increases, since estimation errors are perfectly correlated over time and thus accumulate quickly. As a benchmark case the term structure is evaluated at a horizon of $k = 15$ years with expected returns estimated from a sample of $T = 60$ years. In that case the second term is of the order $\frac{k}{T}\sigma^2 = \frac{1}{4}\sigma^2$ and raises the long run risk by about 25%. Since uncertainty in the unconditional mean has such a profound effect on long term risk and since this uncertainty could also be substantial, it is the focus of our econometric approach.

For assets with mean reversion the term structure will be decreasing for small k, exactly as in Campbell and Viceira (2005), since the effect of parameter uncertainty is usually very small for short horizons. Once there is parameter uncertainty, however, every term structure of risk will eventually be upward sloping beyond a certain horizon.

2.2 Uncertainty about persistence

A second important parameter is the persistence of predictor variables. As an example consider a bivariate VAR for the interactions between nominal and real interest rates. Using 57 years of US quarterly data we obtain the estimated model (standard errors in parentheses),

$$\begin{pmatrix} i_t \\ r_{tb,t} \end{pmatrix} = \begin{pmatrix} \underset{(0.026)}{0.043} \\ \underset{(0.075)}{-0.072} \end{pmatrix} + \begin{pmatrix} \underset{(0.018)}{0.967} & \underset{(0.020)}{-0.015} \\ \underset{(0.055)}{0.202} & \underset{(0.058)}{0.433} \end{pmatrix} \begin{pmatrix} i_{t-1} \\ r_{tb,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{i,t} \\ \epsilon_{r,t} \end{pmatrix}, \tag{15}$$

where i_t is the nominal 3-months T-Bill rate and $r_{tb,t} = i_{t-1} - \pi_t$ is the ex-post real interest rate (π_t is inflation). Since the feedback from the real to the nominal rate is not significant, the nominal rate reduces to a univariate AR(1) process. The standard error on the lagged nominal rate indicates that the AR(1) coefficient is not significantly different from one. Formal test are also not able to reject the unit root null hypothesis for the nominal interest rate.

As a stylized model we analyze

$$\begin{pmatrix} i_t - \mu_i \\ r_{tb,t} - \mu_r \end{pmatrix} = \begin{pmatrix} \rho_i & 0 \\ \beta & \rho_r \end{pmatrix} \begin{pmatrix} i_{t-1} - \mu_i \\ r_{tb,t-1} - \mu_r \end{pmatrix} + \begin{pmatrix} \epsilon_{i,t} \\ \epsilon_{r,t} \end{pmatrix}, \tag{16}$$

with $\mathsf{Var}[\epsilon_t] \equiv \Sigma$ having elements σ_{ii}, σ_{rr} and σ_{ir}. We will investigate the effect of uncertainty in ρ_i on the term structure of risk of the real return from rolling over nominal T-Bills for k periods.

The term structure of risk for the ex-post real rate in this particular model follows as (see appendix)

$$\begin{aligned} V(k) = \frac{1}{k} \sum_{\ell=1}^{k} \Bigg\{ \left(\frac{\beta(\psi_\ell(\rho_i) - \psi_\ell(\rho_r))}{\rho_i - \rho_r} \right)^2 \sigma_{ii} \quad + \\ + 2\beta\psi_\ell(\rho_r) \frac{\psi_\ell(\rho_i) - \psi_\ell(\rho_r)}{\rho_i - \rho_r} \sigma_{ir} + \psi_\ell(\rho_r)^2 \sigma_{rr} \Bigg\} \end{aligned} \tag{17}$$

where $\Psi_\ell(x) = \sum_{j=0}^{\ell-1} x^j$. For ρ_i close to one and large k the quadratic cumulative impulse responses $\psi_\ell(\rho_i)^2$ in the first term of (17) will dominate. In the limit, as $\rho_i \to 1$, we have $\psi_k(1) = k$, leading to $V(k)$ being of order k^2, and thus increasing rapidly with the horizon. For $\rho_i < 1$, the term structure of risk $V(k)$ will converge to a constant.

In practice we are not in the limiting situation of either $\rho_i = 1$ or $k \to \infty$, but the different limiting effects are already quantitatively important at shorter horizons and $\rho_i < 1$. Figure 1 shows the term structure of risk for real T-Bill returns as a function of ρ_i for $0.9 \leq \rho_i \leq 1$, conditional on estimated values for β, ρ_r and Σ. For relatively short horizons, like 1 and 5 years, the term structure hardly depends on the value for ρ_i. The two lines in the figures are almost flat. But at longer horizons, like the 15 years in the figure, the relation between $V(k)$ and ρ_i is upward sloping and concave. Risk estimates range from 2.1% when $\rho_i = 0.90$ to 3.9% at $\rho_i = 0.99$. Close to the unit root the term structure of risk is extremely sensitive to small changes in ρ_i. Due to the concavity of $V(k)$,

$$\mathsf{E}[V(k)] > V(k|\hat{\rho}_i) \tag{18}$$

In this example the average variance over different values of ρ is much larger than the conditional variance at the point estimate $\hat{\rho}_i$. In the empirical sections we will obtain a posterior distribution for ρ_i and quantify the increase in volatility due to uncertainty about ρ_i.

Uncertainty about the persistence is also an important determinant of the long-run unconditional term structure of risk for other asset classes. For equity the predictor variable is the log dividend yield, and again this is a time series exhibiting with near unit root behavior. The effects will be weaker in this case, since it takes a much larger k before the dominant upward sloping term in (17) becomes important.

In a complete VAR model all asset classes are affected by the persistence of state variables and all term structures will eventually increase if the largest eigenvalue of the system is close enough to unity.

2.3 State dependence

For a linear vector autoregression the conditional term structure $V_{i\ell}(k)$ in (8) is independent of the current state $y_t = y$. The conditional mean in (7), however, is a linear function of the state y_t. The second term in (9), the variance of the conditional mean, will therefore be a function of y_t as well and the term structure of risk will in general be state dependent. The term is larger the further the initial condition y_t is away from the unconditional mean $\mathsf{E}[y_t]$. The quantitative effect of the state dependence will mostly be very limited, however, as we will illustrate below.

As an illustration we consider a stylized model with real equity returns predicted by the log dividend yield d_t,

$$\begin{aligned} r_{s,t+1} &= \mu_s + \beta(d_t - \mu_d) + \epsilon_{r,t+1} \\ d_{t+1} &= \mu_d + \rho_d(d_t - \mu_d) + \epsilon_{d,t+1} \end{aligned} \tag{19}$$

We earlier discussed the importance of uncertainty in the unconditional mean μ_s. In this example we concentrate on ρ_d and β, since these parameters determine the state dependence of expected returns. Conditional on the parameters the expected cumulative return on equity is

$$\mathsf{E}_t\left[R^{(k)}_{s,t+k}\right] = k\mu_s + \beta\psi_k(\rho_d)(d_t - \mu_d) \tag{20}$$

For US quarterly data we find the parameter estimates (standard errors) equal to $\hat{\beta} = 2.543$ (1.22) and $\hat{\rho}_d = 0.977$ (0.012). At the fifteen years horizon ($k = 60$) this implies $\psi_{60}(\hat{\rho}_d) = 32.7$ (9.3) with the asymptotic standard error obtained through the delta method. Furthermore, at the end of 2007 and early 2008 the dividend yield was 1.8%, at which point $d_t - \mu_d = -0.005$ and thus

$$\frac{1}{k}\mathsf{E}_t\left[R^{(k)}_{s,t+k}\right] - \mu_s = \frac{2.543 \times 32.7 \times -0.005}{60/4} = -0.0277, \tag{21}$$

meaning that for the next 15 years the return on equity is expected to be 2.7% per year below its unconditional mean μ_s. The dividend yield thus has a persistently large effect on long-run returns, depending on parameters that are subject to substantial estimation error.

The variance of the conditional mean is

$$\mathsf{Var}\left[\mathsf{E}_t\left[R^{(k)}_{s,t+k}\right]\right] = (d_t - \mu_d)^2\mathsf{Var}\left[\hat{\beta}\psi_k(\hat{\rho}_d)\right] \tag{22}$$

Table 1 provides an estimate of the total variance and a breakdown in the components due to the variance of $\hat{\beta}$, the variance of $\hat{\rho}_d$ and their covariance. The total variance is negligibly small. Even with a low estimate of the annual variance of stock returns equal to $0.15^2 = 0.0225$, the sum of the components in table 1 less than 2% of this. The total variance is so small, because the covariance between $\hat{\beta}$ and $\hat{\rho}_d$ is large. When estimating the system (19) by least squares the correlation between $\hat{\beta}$ and $\hat{\rho}_d$ will be equal to the correlation between $\epsilon_{r,t}$ and $\epsilon_{d,t}$. Since the correlation between shocks to equity and shocks to the dividend yield is -0.98, the same also holds for the terms in $\mathsf{Var}\left[\hat{\beta}\psi_k(\hat{\rho}_d)\right]$. Therefore, the same strong covariance that is the key to mean reversion in equity returns, also limits the effects of parameter estimation error on the long-term expected returns.

3 Asset allocation

The Bayesian analysis will increase risk estimates for all asset classes relative to a model conditional on θ. It depends on covariances whether this will have implications for asset allocation.

We derive the optimal portfolio for a power-utility investor who maximizes end-of-period wealth. The investor plans to hold constant proportions of his wealth in each of the asset classes until the end of the period. Fixed or stable portfolio weights appear closely connected to the industry practice of large institutional investors. Pension funds commonly plan their strategic portfolio on a constant mix basis and then allow various tactical bets depending on short term market views. At time t the investor allocates wealth to real T-bills, stocks and bonds with portfolio weights $w = (w_r, w_s, w_b)$ and intends to keep the weights fixed until the end of the planning horizon at time $t+k$. Assuming power utility with risk aversion γ, the investor solves

$$\max_w \mathsf{E}_t\left[U(W_{t+k})\right] = \max_w \mathsf{E}_t\left[\frac{W_{t+k}^{1-\gamma}}{1-\gamma}\right] \tag{23}$$

When investment allocations are rebalanced at the end of each period back to the initial weights, final wealth is given by

$$W_{t+k} = \prod_{j=1}^{k}\left(\sum_{i=r,s,b} w_{it} e^{r_{i,t+j}}\right), \tag{24}$$

where initial wealth is normalized at $W_t = 1$.

In order to calculate the maximum expected utility in (23) we need the distribution of future asset returns. As for the term structure of risk we first draw a set of parameters from the posterior distribution. Conditional on parameters $\theta^{(\ell)}$ we simulate a scenario of future returns $r_{t+j}^{(\ell)}$ from the VAR. Letting $W_{t+k}^{(\ell)}(w)$ denote the final wealth in scenario ℓ using portfolio weights w, we estimate the expected utility of a portfolio as the average realized utility over all scenarios,

$$\bar{U}_k(w) = \frac{1}{L} \sum_{\ell} U\left(W_{t+k}^{(\ell)}(w)\right) \tag{25}$$

The optimal portfolio is the w that maximizes expected utility in (25). We assume that short-sell constraints restrict the weights to be non-negative for all three asset classes, and that the weights sum to one. In the numerical optimization we use a fixed grid for the weights with stepsize 0.01.

The expectation in (23) will not always exist. Although we assume that the innovations in the VAR are normally distributed conditional on the parameters, the predictive densities (5) will have fatter tails than the normal distribution. If a random variable X has a fat-tailed distribution, the expectation $\mathsf{E}[e^X]$ required in (24) does not exist. For the expected utility problem (23) this implies that expected utility will be negative infinity for portfolio returns that have fat left (negative) tails. We can guarantee that at least some portfolios have finite utility if we assume that the real return on the riskfree asset is always above a lower limit r_{min}. In simulating paths of asset returns we reject draws that would violate the constraint $r_{tb,t} > r_{min}$.[6] If the T-bill has a positive weight in the portfolio, wealth will never go to zero and expected utility is bounded away from minus infinity. This approach ensures that the maximization problem (23) is well-defined.

3.1 Uncertainty about the unconditional mean

To gauge the order of magnitude of effects of parameter uncertainty on asset allocation, we continue the example of section 2.1. Consider an investor who can choose between a single risky asset with return r_t and a riskfree asset with constant return r_0. To avoid the numerical problems noted above, also consider the loglinear approximation of Campbell and Viceira (2002) for the portfolio returns,

$$r_{p,t+1} = r_0 + w(r_{t+1} - r_0) + \tfrac{1}{2}w(1-w)\sigma_\epsilon^2, \tag{26}$$

[6] In practice we set the lower bound to twelve standard deviations below the sample mean, meaning that the probability of ever hitting this lower bound is virtually zero.

Assuming log-normality, maximizing expected utility becomes equivalent to

$$\max_w \frac{\ln\left((1-\gamma)\mathsf{E}[U(W_{t+k})]\right)}{1-\gamma} = \\ k\left(r_0 + w(\hat{\mu}-r_0) + \tfrac{1}{2}w(1-w)\sigma_\epsilon^2\right) + \tfrac{1}{2}k(1-\gamma)w^2\left(\sigma_\epsilon^2 + k\omega^2\right), \qquad (27)$$

which leads to the optimal portfolio

$$w_k = \frac{\hat{\mu} - r_0 + \frac{1}{2}\sigma_\epsilon^2}{\gamma\sigma_\epsilon^2 - (1-\gamma)k\omega^2} \qquad (28)$$

With estimation error the portfolio weights are horizon-dependent even though the investment opportunity set is constant. It is only for $\omega^2 = 0$ that portfolio weights do not depend on the investment horizon. With the same typical parameters as before, *i.e.* $k/T = \frac{15}{60}$, $\omega^2 = \sigma_\epsilon^2/T$ and $\gamma = 5$ the denominator in (28) increases by 20% and reduces the allocation to the risky asset by the same amount. The large reduction in the allocation to equity arises here because equity is the only risky asset and the only asset subject to estimation error. In the full VAR model returns on equity, long term bonds and T-Bills are all subject to estimation error. All three asset classes thus become more risky for long-term investors.

3.2 Multiple Priors and Expected Loss

One additional layer of parameter uncertainty are differences in opinion on the prior. The portfolio allocation of the fund is typically decided on by an investment committee. We will assume that different members of the investment committee have different priors indexed by a subscript i. Member i combines his prior $p_i(\theta)$ with the common likelihood $L(\mathbf{Y}|\theta)$ to form his posterior $p_i(\theta|\mathbf{Y})$ and predictive distribution

$$p_i(R_{t+k}^{(k)}|y_t) = \int_\theta p(R_{t+k}^{(k)}|y_t,\theta) \times p_i(\theta|\mathbf{Y})\ \mathrm{d}\theta, \qquad (29)$$

which differs from (5) by the subscript i.

Let $Q_{it}(w)$ be the expected utility of committee member i who uses a prior $p_i(\theta)$, when investing in portfolio w,

$$Q_{it}(w) = \mathsf{E}_{it}\left[U(W_{t+k}(w))\right] \qquad (30)$$

The expectation operator has the additional subscript i to indicate that the expectation is taken relative to the predictive distribution of returns based on prior i, *i.e.*, with respect to $p_i(R_{t+k}^{(k)}|y_t)$ in (29).

According to committee member j, who has prior views $p_j(\theta)$, the expected utility of the same portfolio is

$$Q_{jt}(w) = \mathsf{E}_{jt}\left[U(W_{t+k}(w))\right] \tag{31}$$

The optimal portfolio according to committee member i need not be optimal in the eyes of member j. Let w_i^* and w_j^* be the optimal portfolios of members i and j, respectively. Define the expected loss of the optimal portfolio of committee member j relative to the views of member i as

$$\ell_i(w_j) = Q_{it}(w_i^*) - Q_{it}(w_j^*) \tag{32}$$

Since by construction $Q_{it}(w_i^*)$ is optimal under prior i, the cost is always positive. Conversely, the optimal portfolio of committee member j will be evaluated as sub-optimal by member i and entail the cost $\ell_j(w_i)$.

For a meaningful quantitative comparison of the expected loss, we express it in terms of a certainty equivalent. The economic loss is defined as the annualized percentage with which the initial asset value must increase to compensate an investor for suboptimal investing. For the CRRA utility function we use throughout the paper the costs are computed as

$$C_{ij} = \frac{1}{k}\frac{1}{1-\gamma}\ln\left(\frac{\mathsf{E}_{it}\left[W_{t+k}(w_i^*)^{1-\gamma}\right]}{\mathsf{E}_{it}\left[W_{t+k}(w_j^*)^{1-\gamma}\right]}\right), \tag{33}$$

The economic loss C_{ij} is the percentage risk free return investor i would need to be compensated for being forced to choose asset allocation w_j^*. This way we construct the entire matrix C with elements giving the costs of portfolio w_j^* under prior i.

Since all priors are deemed reasonable, we assume it is impossible to put a probability on the validity of each prior. This also leads to the impossibility to define an overall weighted average prior and weighted average predictive density to evaluate the utility and costs of all portfolios as in Avramov (2002). There will thus remain some ambiguity in how to define the optimal portfolio. This situation with multiple priors has been analyzed in a variety of recent papers, for example Uppal and Wang (2003) and Goldfarb and Iyengar (2003).

In the setting with multiple decision makers at the fund level, we assume that committee member i dislikes any portfolio w_j^* that has a high cost C_{ij}. Since this holds for all committee members, a robust portfolio is defined as the best worst case or minimax solution. For each portfolio the prior that gives the maximum cost is selected. The robust portfolio is the best among all these worst case evaluations. With our limited set of portfolios w_j^* and limited set of priors, the robust portfolio is selected as the minimax solution within the matrix C. As each column of C represents a portfolio, we select the worst case of portfolio w_i^* as the maximum element in column i. The portfolio with the lowest maximum is the minimax portfolio.

4 Econometric specification

Following Campbell *et al* (2003), among others, we describe the return dynamics by the first-order vector autoregression (VAR) for returns and state variables,

$$y_{t+1} = c + \mathbf{B}y_t + \epsilon_{t+1}, \tag{34}$$

where ϵ_t is normally distributed with zero mean and covariance matrix Σ. Our main interest is in how long-term risk and portfolio choice are affected by parameter and model uncertainty. Parameter uncertainty is accounted for by the Bayesian analysis of the VAR. Model uncertainty is represented by a series of alternative priors on the unconditional mean of the asset returns and state variables.

As a benchmark prior we consider the flat prior on c and $\mathbf{B}$ and the non-informative Wishart prior on Σ,

$$p(c, \mathbf{B}|\Sigma) \propto I(\mathbf{B}) \tag{35}$$

$$p(\Sigma) \propto |\Sigma|^{-(n+1)/2}, \tag{36}$$

where the indicator function $I(\mathbf{B})$ is equal to one if the maximum eigenvalue of $\mathbf{B}$ is less than one, and zero otherwise. The flat prior has previously been used by Barberis (2000) in his analysis of the effects of parameter uncertainty on long-term portfolio decisions between the riskfree asset and equity. Like Barberis (2000) and Hollifield *et al* (2003) we exclude non-stationary models, since unconditional means can not be defined in such models and because explosive roots also lead to explosive behavior of the term structure of risk.[7] When the likelihood attaches negligible probability mass to the nonstationary region, the posterior mode coincides with the least squares estimate of the parameters c and $\mathbf{B}$.

[7] Even very small probabilities of explosive roots will dominate the results. If there is substantial evidence against stationarity of the VAR we would rather include VECM specifications or structural breaks. Adding these features will only add to the long-run risk and further strengthen our conclusion that parameter uncertainty is an important component of long-run risk.

In order to impose alternative prior views, we reparameterize the VAR model as in (3) and use the unconditional means

$$\mu = (I - \mathbf{B})^{-1} c \tag{37}$$

instead of the constant terms c. For our econometric analysis we formulate informative priors on μ as

$$\mu \sim \mathsf{N}\left(\mu_0, \frac{1}{\kappa}\Omega_0\right), \tag{38}$$

where both μ_0 and Ω_0 are exogenously specified. The scalar parameter κ is a shrinkage factor. It represents the investor's degree of confidence in the prior information. A shrinkage factor close to zero corresponds to a dispersed prior on μ. A large shrinkage factor gives much weight to the prior information, while a precision factor equal to infinity imposes the mean. As in the benchmark prior we specify a flat prior on $\mathbf{B}$ in the stationary region and an uninformative inverted Wishart on Σ. The different priors for μ (and implicitly c) affect the estimates of the term structure of risk in two ways. First, uncertainty in the expected returns is a separate source of risk, analyzed in section 2.1. Second, different priors on μ or c also affect inference on the dynamic properties $\mathbf{B}$ and therefore the long-run risk estimates.

Since the VAR is nonlinear in the parameters μ and $\mathbf{B}$, the setting is very different from the benchmark prior in the linear representation (34). To analyze the difference, the prior on μ can be transformed back to a prior on the reduced form parameters $(c, \mathbf{B})$. As a shorthand notation, let $\mathbf{A} = \mathbf{I} - \mathbf{B}$. The transformation $c = \mathbf{A}\mu$ induces a Jacobian $|\mathbf{A}|^{-1}$, leading to the implied joint prior density of c and $\mathbf{B}$,

$$p(c, \mathbf{B}) \propto |\mathbf{A}|^{-1} \exp\left(-\tfrac{1}{2}\kappa(c - \mathbf{A}\mu_0)'(\mathbf{A}\Omega_0\mathbf{A}')^{-1}(c - \mathbf{A}\mu_0)\right) \tag{39}$$

Since the matrix $\mathbf{A}$ is singular at points where $\mathbf{B}$ has a unit root, the prior forces a singularity on the constant terms c when the dynamics of the system move towards the unit root. Conditional on a small constant term, in the sense that the term in the exponent in (39) is small, the Jacobian term $|\mathbf{A}|^{-1}$ will be important.

The effects of the prior are best explained with the univariate AR(1) model

$$y_{t+1} = \mu + \rho(y_t - \mu) + \epsilon_{t+1} \tag{40}$$

As an illustration of the prior, figure 2 shows the marginal prior distribution of c in the univariate AR(1) case when μ is normally distributed with mean $\mu_0 = 1$ and variance $\omega_0^2 =$

$\Omega_0/\kappa = 4$, while the first order autocorrelation coefficient ρ has a uniform prior distribution on $\rho \in (0, 1)$. The density has a distinct spike at $c = 0$, since the parameterization implies that $c \to 0$ when $\rho \to 1$. Even more telling are the conditional densities $p(\rho|c)$ for alternative values of c. Conditional on $c = 0.05$ the density is highly concentrated on the unit root, whereas for $c = 0.5$, the density drops to zero for $\rho > 0.98$. Despite its large peak close to $\rho = 1$, the conditional density for $c = 0.05$ still puts zero mass at the unit root itself, since from (39) we have the limit

$$\lim_{\rho \to 1} p(\rho|c) = \lim_{\rho \to 1} \frac{1}{1-\rho} \exp\left(-\frac{(c-(1-\rho)\mu_0)^2}{2(1-\rho)^2\omega_0^2}\right) = 0 \tag{41}$$

when $c \neq 0$. Inference on the autocorrelation is strongly correlated with inference on the constant term in the model. Schotman (1994) shows that the combination of a normal prior on μ and a flat prior on ρ leads to posterior densities for ρ that are more concentrated towards the unit root than a flat prior on both ρ and $c = (1-\rho)\mu$. Intuitively, once we fix the unconditional mean far away from the sample mean, either above or below, the sample path of the time series will cross this unconditional mean less often. This induces a higher estimate of the first order autocorrelation of the series. With an informative prior on μ we will therefore obtain a posterior for ρ closer to the unit root.

We focus on priors for the unconditional mean for various reasons. Technically it mitigates the econometric inconvenience of the censoring of the prior on the stationary region $I(\mathbf{B})$ in the benchmark flat prior on c and $\mathbf{B}$. Villani (2005, 2009) explains that even a weakly informative prior on the unconditional mean will reduce the probability of drawing explosive parameters $\mathbf{B}$.

Another inconvenience of the benchmark prior is that with a flat prior on (c, $\mathbf{B}$) the implied posterior density of μ will not have first or second moments. In the univariate case, the ratio $\mu = c/(1-\rho)$ will have a posterior density that is the ratio of two normal densities, *i.e.* a Cauchy density, which does not have finite moments.

The parameterization with μ and $\mathbf{B}$ in (37) allows efficient numerical computation of the posterior density. Villani (2009) shows that a simple three block Gibbs sampler, partitioning the parameters in $\theta = (\mathbf{B}, \mu, \Sigma)$, works very well. For given μ the model is linear in $\mathbf{B}$, while conditional on $\mathbf{B}$ the model is linear in μ.

A second motivation for adopting informative priors on μ is that priors on the level of future returns are already in place for some time in asset allocation at many institutional investors and asset managers (see Black and Litterman (1992)). Although historical data can provide robust estimates of future volatility and correlation, historical average returns are very sensitive to the choice of the data period. The speculative part of optimal portfolio choice is extremely sensitive to small changes in the expected returns. Chopra and Ziemba (1993) argue that the primary emphasis in portfolio choice should be on obtaining superior estimates for means. It is therefore common practice for long-term investors to base their future return expectations not only on historical data, but also on current market circumstances, economic theory, human judgement and experiences in other countries.[8] For our model this means a flat prior on $\mathbf{B}$ and Σ, and an informative prior on the long-term economic outlook.

To investigate the sensitivity of our long-run risk estimates and to check the working of the robust asset allocation rule we implement both optimistic outlooks and more negative views on the future. We want our alternative priors to be sufficiently distinct so as to generate meaningful model uncertainty. Specifically, we consider two different prior means μ_0. For our application to a postwar US sample we split the historical data in two parts: NBER expansion periods and NBER contraction periods. Averages in the expansion periods represent a positive outlook for all asset returns and state variables in y_t, whereas the contraction period averages define a pessimistic outlook for long-term means. The alternative vectors of prior means will be labeled *optimistic* and *pessimistic*.

We also distinguish between very confident views and highly dispersed priors on the long run expected returns. For the prior precision κ in (38) we consider three different values. To calibrate κ we take an empirical perspective as well and estimate Ω_0 as the long-run variance of our historical data,

$$\Omega_0 = \frac{1}{T}\sum_{\ell=-L}^{L} \nu_\ell \left(\frac{1}{T}\sum_t (y_t - \bar{y})(y_{t+\ell} - \bar{y})' \right) \tag{42}$$

with Newey-West weights $\nu_\ell = \left(1 - \frac{|\ell|}{L+1}\right)$. With this choice of Ω_0 we can interpret κ as the weight of the prior. We vary κ to increase or decrease the precision of the prior while

[8] Forward rates are an example of the current market view of about interest rates. The macro-finance literature on the equity premium starting with Mehra and Prescott (1985) is an example of theory motivated priors for the equity premium. Welch (2001) is an example of survey evidence of expected returns. Dimson *et al* (2002) compare average bond and equity premiums of sixteen countries over more than a century.

keeping Ω_0 fixed. The prior precision parameter varies between $\kappa = 0.01$ (*uninformative*), $\kappa = 1$ (*moderate*) and $\kappa = 100$ (*dogmatic*). For $\kappa = 100$ we almost impose the unconditional means and all uncertainty is in the dynamic aspects of the term structure of risk and the hedge demands in asset allocation. If the prior precision factor $\kappa = 0.01$, the prior hardly carries any weight and all evidence will be data based.

5 Estimation Results

We consider returns on three asset classes (stocks, bonds and T-Bills) and three state variables that help predict asset returns (inflation, dividend yield and term spread). For our empirical analysis we use quarterly US data. All series start in 1952:I and end in 2008:IV. The 90-days T-bill and the 10-years constant maturity yield are from the FRED website.[9] In order to generate the yield spread we obtain the zero yield data from Duffee (2002).[10] As these data are only available until 1998:IV, we have extended the series using the data from Gürkaynak *et al* (2006).[11] For inflation we use the non-seasonally adjusted consumer price index for all urban consumers and all items also from the FRED website. Data on stock returns and the dividend price ratio are based on the S&P Composite and are from the "Irrational Exuberance" data of Shiller.[12] We construct the gross bond return series from 10 year constant maturity yields on US bonds using a log-linear approximation approach as in Campbell *et al* (1997).

Table 2 provides summary statistics. Even with the financial crisis year 2008 included, the sample equity premium of around 6% is larger than most recent studies on the prospective equity premium suggest. For example, Claus and Thomas (2001) suggest a forward looking equity premium of about 3.5%; the same number that emerges from the average of 10 years of survey evidence in Graham and Harvey (2010). Since our post-second-world-war sample period is rather short, the sample means of the equity and bond premia may be very poor estimates of the long-run expected returns. The standard errors of the annualized sample means are 1% and 2% for the bond and equity premium, respectively.

The final columns in table 2 give the summary statistics for the prior means. We as-

[9] `http://research.stlouisfed.org/fred2/`
[10] `http://faculty.haas.berkeley.edu/duffee/affine.htm`
[11] `http://www.federalreserve.gov/econresdata/researchdata.htm`
[12] `http://aida.econ.yale.edu/~shiller/data.htm`

sign each observation in our sample period to NBER contraction or expansion periods. A contraction starts at the peak of a business cycle and ends at the trough, and the expansion vice versa. Nine contraction periods exist in our data sample, which have a duration between two and six quarters. Ten expansion periods exist in our sample period with a duration ranging from four to 40 quarters. Contractions are on average much shorter than expansions, and consequently 191 out of the 228 observations are assigned to expansions, and the remaining 37 observations are contractions. We choose the closest quarter to end the contraction or expansion whenever a through occurs during a quarter.

The prior mean is set at either the average over the contraction and expansion period, respectively. The contraction and expansion sub-sample averages support a wide range of prior expectations. Bonds seem less attractive than stocks during expansions, whereas they seem more attractive during contractions. Even though negative expected returns may be unreasonable from an asset pricing perspective, we leave the averages as they are, since they generate much dispersion in the priors. For our results on the term structure of risk we will look at each of the priors in isolation for their long-run risk properties. We can thus study the robustness of the dynamic properties with respect to the choice of the prior. For our portfolio results the dogmatic prior provides the opportunity to check the effect of extreme priors on the robust portfolio choice.

The averages of state variables as the short interest rate, dividend yield, and term spread also differ between the two periods. Fama and French (1989) link the dividend yield and yield spread to the business cycle. They argue that the risk premia are high in contraction periods and low in expansion periods. The opposite applies to the dividend price ratio which is high in expansion periods and low in contraction periods.

The VAR system is estimated on the entire sample. Tables 3 and 4 summarize the OLS parameter estimates together with the correlations and standard deviations of the residuals. Since our model is similar to much of the literature, except for the sample period, we only highlight the most important results. First, the three state variables (nominal interest rate, dividend yield, term spread) are almost univariate AR(1) processes. Second, as in Ang and Bekaert (2007) the nominal interest rate and the dividend price ratio jointly predict excess stocks returns. As in Campbell and Viceira (2005) the combination of a negative correlation of shocks to the dividend price ratio and stocks, and the positive predictive coefficient of the dividend price ratio imply mean reversion in stocks returns. The excess return on bonds is related to the yield spread, the nominal interest rate and stock returns. Third, bond returns are also mean-reverting. The nominal interest rate is a predictor of excess bond returns, which has the required opposite signs of the predictive coefficient and residual correlation. The term spread leads to a mean aversion part. The R^2 around 8% for both stocks and bond returns is of the magnitude typically found for quarterly data. It is low enough to be credible for efficient financial markets and large enough to be economically meaningful at longer horizons (see Campbell and Thompson (2005)).

The priors influence the persistence of state variables and the predictability of stock and bond returns. Table 5 indicates that the posterior mean of the autocorrelation parameter of the two most persistent state variables, dividend yield and nominal T-Bill rate, increases when we specify an uninformative prior on the long-run means μ instead of a flat prior on the constant term c. The difference is in line with the analysis in section 4. The differences are small, but potentially meaningful so close to the unit root. The largest differences are for the predictability parameter of the dividend price ratio and excess stocks returns: predictability varies from large and very significant under the *flat* prior to small and insignificant for the *dogmatic pessimist* prior. Predictability of bonds is less affected by the choice of prior.

An important parameter for the long-term risk estimates is the maximum eigenvalue of the VAR system. From the OLS estimates of the coefficient matrix we have the point estimate $\lambda_{max} = 0.976$, implying that the VAR is stationary with a halflife of 28.5 quarters, a little more than seven years. Figure 3 shows the posterior distribution of λ_{max} for the different priors. Even though the priors are only formulated on the unconditional means, they have a strong impact on the persistence of the system. For the *flat* prior the posterior mean of the largest eigenvalue is almost identical to the OLS estimate.[13] For all the other priors the persistence is larger than the OLS estimate. The posterior for an *uninformative* prior on the unconditional mean is extremely skewed: the mode of the distribution is 0.996, almost at the unit root. As we argued in the example in section 2.2 such strong persistence leads to steep initial mean reversion in the term structure of risk, but also to a quick reversal

[13] For a flat prior, the posterior of the VAR parameters $\mathbf{B}$ is a matrixvariate Student-t distribution with mean equal to the OLS estimates. Since $\lambda_{max} = f(\mathbf{B})$ is a nonlinear transformation of the OLS estimates, it is not necessarily true that $\bar{\lambda}_{max} = f(\bar{\mathbf{B}})$. In practice we reject about 10% of draws because of $\lambda_{max} > 1$. This truncates the right tail of the posterior of the maximum eigenvalue, but nevertheless leads to a posterior mean and mode that are (slightly) above the OLS point estimate.

towards an upward sloping term structure.

Table 6 shows the different estimates for the unconditional means μ. For the OLS estimates we can compute the implied long-run means $\hat{\mu} = \hat{\mathbf{A}}^{-1}\hat{c}$ and its asymptotic standard errors. These asymptotic standard errors are all much smaller than the posterior standard deviations obtained with an *uninformative* prior.[14] It seems as if the asymptotic OLS standard errors underestimate the true estimation uncertainty in expected returns.

6 Long-term risks

Figure 4 shows different estimates of the term structure of risk. Conditional on the *OLS* estimates of the VAR the results are very similar to the figures in Campbell and Viceira (2005) and others. Bills become more risky with the investment horizon due to roll-over risk and inflation risk. Stocks exhibit strong mean reversion that reduces the long-term risk by a factor two. Real returns on nominal bonds show some mean reversion as well, but far less than equity.

Adding parameter uncertainty increases the risk for all three asset classes. The figure shows the term structure of risk for the *flat* prior and the *uninformative* priors for the long-term mean. As expected the term structure is hardly affected by parameter uncertainty at very short horizons. At longer horizons parameter uncertainty becomes an important risk component. For equity the mean reversion almost disappears. The annualized standard deviation moves from 15% at the 1 quarter horizon to 13.5% at the 15 years horizon and after a long flat region it starts to increase at very long horizons. Risk estimates for the *uninformative* prior are generally above the *flat* prior estimates. This is related to the different posteriors for the largest eigenvalue of the VAR system.

Most striking is the almost uniform effect on all three asset classes. At the 15 years horizon a *flat* prior implies a 25% higher annualized standard deviation for bills, bonds as well as equity. Not only equity becomes more risky, but all asset classes, and all risk estimates increase proportionally. Stretching the horizon to 50 years the increase in risk remains almost proportional for all three asset classes. Using the *uninformative* prior on

[14] The table does not contain posterior moments for μ under the *flat* prior, since these moments do not exist. If both c and B are approximately normally distributed the posterior density of $\mu = \mathbf{A}^{-1}c$ has fat tails.

the unconditional mean, either *pessimist* or *optimist*, the increase in risk is also very similar for all three asset classes.

Informative priors can lead to very different term structures of risk. Figure 5 zooms in on the equity term structure. The *uninformative* term structure is the same as in figure 4, while the lines for *dogmatic optimist* and *dogmatic pessimist* show the effect of fixing the unconditional means, thereby leaving out one component of the risk estimates. For the *optimist* prior this indeed leads to a reduction of the risk estimates compared to the *uninformative* prior, albeit that the effect is very small. For the *pessimist* prior the effect is much stronger. Here the risk increases, because the prior not only affects inference on the long-run mean, but also the inference on the dynamics. In section 5 we already discussed the effect on the coefficient relating the predicted equity return to the dividend yield. This coefficient is smaller in the *dogmatic pessimist* prior and hence the mean reversion disappears. The *dogmatic* priors have less effect on the risk estimates for the other asset classes (not shown), since for these asset classes the prior has very little impact on the most important parameters for their return dynamics.

Continuing on the equity term structure, figure 6 offers a closer look at the source of the increase in risk. The figure shows a decomposition of the total equity risk in the average conditional variance and the variance of the conditional mean, as in (9). The first term, the average conditional variance, is the biggest component. The variance of the expected returns is negligible at short horizons; even at 50 years it is still much smaller than the first term. The big risk increase from parameter uncertainty is therefore related to the concavity of the relation between the term structure of risk and the persistence of the system (consistent with the analysis in section 2.2). Evaluating the term structure at the point estimates is very different from the average term structure across many different draws from the posterior density of the parameters. The rise in annualized volatilities for investment horizons longer than 25 years indicates that parameter uncertainty dominates the mean reverting dynamics in the long run due to the uncertainty about the unconditional expected return.

In contrast to volatilities, correlations are robust against parameter uncertainty. Figure 7 shows the pairwise correlation among the real returns. The correlations exhibit interesting patterns as a function of the investment horizon, but they are hardly affected by parameter uncertainty. For the *flat* and *uninformative* priors the correlations are hardly distinguishable from the *OLS* results. Again, the impact of parameter uncertainty on the covariance is proportional to the impact on the variance, and cancels out in the correlations. The only exception is the *dogmatic pessimist* prior, which leads to a very different correlation between equity and bonds at all horizons. In this case the prior mean is far from the sample mean and as we noted before this has an impact on the dynamic properties of the system. The *dogmatic pessimist* prior leads to slightly negative correlation between stocks and bonds, and hence implies much more diversification benefits than the other priors.

As a descriptive measure for the inflation hedge properties of the three asset classes we examine the correlation of nominal returns and inflation over different horizons. Results in figure 8 show that the inflation hedge qualities of T-bills, stocks and a constant maturity treasury portfolio are robust to incorporating parameter uncertainty. Apparently, the effect of parameter uncertainty on the covariance between the nominal returns and inflation is proportional to the effect on the corresponding variances. Even the *dogmatic pessimist* prior does not lead to really different estimates. The horizon effects themselves are in line with Hoevenaars *et al* (2008) for fixed parameter estimates. Rolling over 3-month T-bills ensures that the lagged inflation is incorporated, and consequently the T-bill is the best inflation hedge among the asset classes we consider at all investment horizons. At long horizons constant maturity bonds become an inflation hedge as well. However, due to the inverse relationship between yield changes and bond prices, the short-term inflation hedging properties are poor. For stocks there is hardly any correlation with inflation at all on horizons up to 15 years as shown in the figure. For a positive correlation we would have to extend the horizon to more than 25 years. For the *dogmatic pessimist* prior the correlation remains negative at all horizons. Once again it shows that results can be sensitive to the location of the prior on the unconditional mean of returns.

7 Asset Allocation

Our main focus has been the effect of parameter uncertainty on risk. For the effects on optimal portfolio choice we first compare the conditional portfolios based on the *OLS* estimates with the Bayesian portfolios implied by the *flat* and *uninformative* priors. Obviously, the alternative *dogmatic* priors shift the optimal portfolios in a pre-determined direction as they condition on widely different long-term expected returns. The costs and effects of including these informative priors will be assessed using the robust portfolio choice in the second part of this section.

7.1 Parameter uncertainty

Table 7 contains the results for a benchmark case with risk aversion $\gamma = 5$ and various horizons ranging from one to fifteen years. For the *OLS* estimates we obtain the standard result that the weight of stocks increases with the investment horizon. This result is a combination of the high historical equity premium and the strong mean reversion conditional on the estimated VAR models. For horizons longer than 15 years the investor would prefer to be fully invested in stocks. T-bills have a high weight for short investment horizons, but reinvestment risk makes them less attractive at longer horizons. Due to the low bond premium in the data bonds are not in the optimal portfolio.

As anticipated, parameter uncertainty does not matter much at the short one year horizon. For long horizons, however, the optimal investment in equity falls substantially. Whether we use the *flat* prior or either the *optimist* or *pessimist uninformative* priors, the optimal allocation to equity is always in the 50-60% range. Consequently, the horizon effect also diminishes: the difference in allocation to equity between the one year and fifteen years horizon is never more than 15%. Whereas the share of equity decreases, the allocation to T-Bills increases. T-Bills are a good risk diversifier due to the very low correlation with stocks at longer horizons. Bonds are not in the optimal portfolio, because their risk rises in the same proportion as the risk of stocks, while the bond premium is much lower. Although ignoring parameter uncertainty leads to an overallocation to stocks at all horizons, the investor should still increase the weight of equity for investment horizons up to 15 years. Parameter uncertainty causes mean reversion in equity to largely disappear, but since it also increases the risk of the other asset classes, equity remains more attractive for long-term investors.

Obviously, the optimal asset allocation varies substantially with the view and confidence the investor has about the level of the future asset returns. Under the *dogmatic pessimist* view long-term bonds have a high premium and therefore dominate the portfolio. Under the *dogmatic optimist* prior the short-term allocation to equity is much more than conditional on the *OLS* estimates. This is the optimistic mean effect. At long horizons the allocation is much lower than in the *OLS* case, since parameter uncertainty about the dynamics increases the risk.

Figure 9 shows how the risk attitude of the investor influences the long-term (15 years) optimal portfolio choice under the different priors. The pattern is the same in all cases, except for the *dogmatic pessimist* prior. Beyond a certain risk aversion, the allocation to stocks becomes less than 100% and negatively related to the risk aversion (γ). The trade-off is between equity and T-Bills. Nominal bonds are only attractive under the *dogmatic pessimist* prior.

7.2 Robust asset allocation

The optimal portfolio according to one of the priors is sub-optimal when evaluated under one of the other priors. Table 8 reports the annualized certainty equivalent costs (see (33)) at the different investment horizons with $\gamma = 5$. The opportunity costs of alternative priors are economically meaningful and increase with the investment horizon.

Independent of the investment horizon we find that all portfolios are considered as very costly from the perspective of the *dogmatic pessimist* (*P100* in the table) prior. Equity is unattractive under this prior. The more equity in the portfolio the larger the cost according to the *dogmatic pessimist*. The cost function is asymmetric, however. For example, at the 15 years horizon, the *dogmatic optimist* portfolio (*O100*) incurs a certainty equivalent cost of 12.7% per year when evaluated by the *dogmatic pessimist*. For the reverse cost, when the *dogmatic optimist* evaluates the optimal portfolio of the *dogmatic pessimist*, we find the much lower 4.5%. The difference is due to the concavity of the cost function. Pessimists attach larger cost to optimistic portfolios than vice versa. In general, the cost of aggressive portfolios with more equity is more sensitive to the prior than more conservative portfolios.

At all horizons the minimax solution is the optimal portfolio according to the *moderately pessimistic* view. The *dogmatic optimistic* investor dislikes this portfolio most and incurs a of 1.9% at the one year horizon to 2.6% at the 15 years horizon. Table 7 shows that this robust portfolio has a mix of about 20% T-Bills, 20% equity and 60% long-term bonds and hardly depends on the investment horizon. Nominal bonds are thus important in this portfolio. The risk and returns properties of bonds make them less attractive than alternatives under most priors, but holding them is also not considered as very costly either. This distinguishes bonds from equity, which are perceived as very risky and costly from the perspective of investors with a pessimist view on the equity premium.

8 Conclusion

We considered the impact of parameter uncertainty on long-term risk and asset allocation of long-term investors who can invest in stocks, bonds and T-bills. When using uninformative priors, we find that real returns on stocks and bonds exhibit limited mean reversion at short and medium investment horizons. Mean reversion disappears, however, at long horizons. Parameter uncertainty affects all three asset classes, stocks, bonds and bills, in proportionally the same way, with annualized standard deviations of returns increasing by a factor 1.25 at the fifteen years horizon relative to estimates conditional on maximum likelihood parameter estimates. Asset return correlations are stable under weakly and moderately informative priors. Similarly, the correlation between inflation and nominal asset returns are also robust to incorporating parameter uncertainty.

We also considered informative priors on the long-run mean of returns and predictor variables and find that for highly persistent time series (like dividend yield and nominal interest rate) even weak prior information changes the estimated persistence of shocks and the predictability of excess returns. Both effects increase the term structure of annualized volatility. Alternative strong prior views on long-run stock and bond returns also affect the correlation between returns on stocks and bonds. It is positive under uninformative priors, but turns negative under a pessimistic long-run outlook for the stock market.

Parameter uncertainty alters optimal portfolio allocations. Obviously, assets with high prior expected returns get a large portfolio weight. But more interesting is the effect of an uninformative prior. Even though the risk of all asset classes increases, the long-term portfolio weight of equity decreases substantially, and the weight of short-term bonds increases. Moreover, the effect of the investment horizon on the optimal portfolio diminishes. Combining widely different priors in a robust portfolio, we find that the optimal weight of equity is further reduced to about 20%, whereas long-term nominal bonds obtain a large weight. The reason is that large equity exposure is very costly in the opinion of investors with a more pessimistic prior opinion on the equity premium.

In this study we considered a limited number of alternative priors. We have not exhaustively looked at the many different dynamic models of asset returns and priors that exist in the literature, but rather focused on the mechanism by which parameter uncertainty increases risk at different horizons for different asset classes. Alternative priors may come up with lower risk, because they reduce parameter uncertainty by specifying an informative prior on the VAR dynamics. Cremers (2002) and Wachter and Warusawitharana (2009), for example, specify a prior on the R^2 of a return prediction equation, while Pastor and Stambaugh (2009) add information on mean reversion through an informative prior on the correlation between shocks to expected and unexpected returns. On the other hand, risk estimates may also increase by extending the class of models. For example, Pettenuzzo and Timmermann (2010) introduce structural breaks and the risk of future structural breaks, while Pastor and Stambaugh (2010) add an unobserved state variable to the model.

More in line with the construction of the robust portfolio would be to develop alternative priors from scenario descriptions that are often used in practice.

Appendix A Term structure of risk

Start from the first order VAR in (3), repeated here without constant terms

$$y_{t+1} = \mathbf{B} y_t + \epsilon_{t+1} \tag{A1}$$

Conditional expectations follow directly as $\mathsf{E}_t[y_{t+j}] = \mathbf{B}^j y_t$ and give the decomposition

$$y_{t+j} - \mathsf{E}_t[y_{t+j}] = \sum_{i=0}^{j} \mathbf{B}^{j-i} \epsilon_{t+i} \tag{A2}$$

Cumulating over k periods we have

$$Y_{t+k}^{(k)} - \mathsf{E}_t\left[Y_{t+k}^{(k)}\right] \equiv \sum_{j=1}^{k} (y_{t+j} - \mathsf{E}_t[y_{t+j}]) = \sum_{i=1}^{k} \left(\sum_{j=i}^{k} \mathbf{B}^{j-i} \right) \epsilon_{t+i} \tag{A3}$$

The conditional variance of $Y_{t+k}^{(k)}$ is

$$\begin{aligned}
\mathsf{Var}\left[Y_{t+k}^{(k)}\right] &= \sum_{i=1}^{k}\left(\sum_{j=i}^{k}\mathbf{B}^{j-i}\right)\mathbf{\Sigma}\left(\sum_{j=i}^{k}\mathbf{B}^{j-i}\right)' \\
&= \sum_{\ell=1}^{k}\left(\sum_{j=0}^{\ell-1}\mathbf{B}^{j}\right)\mathbf{\Sigma}\left(\sum_{j=0}^{\ell-1}\mathbf{B}^{j}\right)' \\
&\equiv \sum_{\ell=1}^{k}\mathbf{\Psi}_{\ell}\mathbf{\Sigma}\mathbf{\Psi}_{\ell}' \qquad \text{(A4)}
\end{aligned}$$

The last line of (A4) defines the matrix sums $\mathbf{\Psi}_{\ell} = \sum_{j=0}^{\ell-1}\mathbf{B}^{j}$. By selecting the relevant elements from the covariance matrix and dividing by k, we obtain the term structure of risk for real asset returns of alternative asset classes.

The general formulas simplify in specific cases. In section 2.2 we consider a bivariate VAR for real and nominal interest rates with the structure

$$\mathbf{B} = \begin{pmatrix} \rho_i & 0 \\ \beta & \rho_r \end{pmatrix}, \qquad \text{(A5)}$$

which implies that powers of $\mathbf{B}$ are given by

$$\mathbf{B}^{j} = \begin{pmatrix} \rho_i^{j} & 0 \\ \beta\frac{\rho_i^{j}-\rho_r^{j}}{\rho_i-\rho_r} & \rho_r^{j} \end{pmatrix}, \qquad j > 0, \qquad \text{(A6)}$$

and by definition $\mathbf{B}^{0} = I$. Define the scalar sums

$$\psi_k(x) = \sum_{i=0}^{k-1} x^{i} = \frac{1-x^{k}}{1-x}, \qquad \text{(A7)}$$

with obviously $\psi_k(1) = k$. Using this definition and (A6) we can now write the matrix sum in the conditional variance (A4) as

$$\mathbf{\Psi}_k = \begin{pmatrix} \psi_k(\rho_i) & 0 \\ \beta\frac{\psi_k(\rho_i)-\psi_k(\rho_r)}{\rho_i-\rho_r} & \psi_k(\rho_r) \end{pmatrix} \qquad \text{(A8)}$$

Since the cumulative real return on the T-Bill is the second element in the VAR, we only need the $(2,2)$ element of the covariance matrix $\mathsf{Var}\left[Y_{t+k}^{(k)}\right]$. Using the structure of $\mathbf{\Psi}_{\ell}$ we evaluate the terms in (A4) as

$$
\begin{aligned}
\left(\mathbf{\Psi}_\ell \Sigma \mathbf{\Psi}_\ell'\right)_{2,2} &= \begin{pmatrix} \beta \frac{\psi_\ell(\rho_i)-\psi_\ell(\rho_r)}{\rho_i-\rho_r} & \psi_\ell(\rho_r) \end{pmatrix} \begin{pmatrix} \sigma_{ii} & \sigma_{ir} \\ \sigma_{ii} & \sigma_{rr} \end{pmatrix} \begin{pmatrix} \beta \frac{\psi_\ell(\rho_i)-\psi_\ell(\rho_r)}{\rho_i-\rho_r} \\ \psi_\ell(\rho_r) \end{pmatrix} \\
&= \left(\frac{\beta(\psi_\ell(\rho_i)-\psi_\ell(\rho_r))}{\rho_i-\rho_r}\right)^2 \sigma_{ii} + 2\beta\psi_\ell(\rho_r)\frac{\psi_\ell(\rho_i)-\psi_\ell(\rho_r)}{\rho_i-\rho_r}\sigma_{ir} + \psi_\ell(\rho_r)^2\sigma_{rr}
\end{aligned}
\tag{A9}
$$

Summing over ℓ and dividing by k leads to the formula given in (17) in the text.

Appendix B Simulation

We derive the posterior for the different prior specifications in the text. The first prior is the flat prior on the parameters of the reduced form VAR in (34). Defining $\tilde{\mathbf{B}} = (c\,\mathbf{B})$ and $x_t = (1\; y_{t-1}')'$ the VAR model can be written as the system of linear regression equations

$$y_t = \tilde{\mathbf{B}} x_{t-1} + e_t \tag{B1}$$

The stationarity restriction on $\mathbf{B}$ implies that exact analytical integration over $\mathbf{B}$ is not possible. An obvious simulation algorithm is available, since the conditional posterior of $\tilde{\mathbf{B}}$ will be truncated normal and therefore easily sampled from. To generate draws for the parameters we consider a two-block Gibbs sampler.

1. The conditional posterior of Σ is the inverted Wishart with T degrees of freedom and moment matrix $\hat{S} = \sum_t e_t e_t'$.

2. Ignoring the stationarity constraint the conditional posterior of $\tilde{\mathbf{B}}$ is the matricvariate normal $\mathsf{N}\left(\hat{\tilde{\mathbf{B}}}, \Sigma \otimes \mathbf{X}^{-1}\right)$ with $\hat{\tilde{\mathbf{B}}}$ the matrix of OLS regression estimates and $\mathbf{X} = \sum_t x_{t-1}x_{t-1}'$. Given Σ we draw until we obtain a valid $\mathbf{B}$ inside the stationary range.

For the structural VAR,

$$y_t = \mu + \mathbf{B}(y_{t-1} - \mu) + e_t, \tag{B2}$$

with the informative prior (38) on μ the simulation proceeds in three blocks as in Villani (2009).

1. As before, the posterior of Σ, conditional on $\mathbf{B}$ and μ, is the inverted Wishart distribution $iW(\hat{S}, T)$.

2. Since the prior on $\mathbf{B}$ is still flat, the conditional posterior, given μ and Σ, is the matricvariate normal $\mathsf{N}\left(\hat{\mathbf{B}}, \Sigma \otimes \mathbf{X}_\mu^{-1}\right)$, where $\hat{\mathbf{B}}$ are the OLS estimates of the regression system (B2) estimated conditional on μ, and the matrix $\mathbf{X}_\mu = \sum_t (y_{t-1} - \mu)(y_{t-1} - \mu)'$. We draw until we obtain a valid $\mathbf{B}$ in the stationary region.

3. For the posterior of μ conditional on $(\mathbf{B}, \Sigma)$, define $\mathbf{A} = I - \mathbf{B}$ and the system

$$z_t \equiv y_t - \mathbf{B}y_{t-1} = \mathbf{A}\mu + e_t \tag{B3}$$

Then the conditional posterior follows as $\mu \sim N(\hat{\mu}, \mathbf{V})$, where the conditional moments are defined as

$$\mathbf{V} = \left(T\mathbf{A}'\Sigma^{-1}\mathbf{A} + \kappa\Omega_0^{-1}\right)^{-1} \tag{B4}$$

$$\hat{\mu} = \mathbf{V}\left(\mathbf{A}'\Sigma^{-1}\bar{z} + \kappa\Omega_0^{-1}\mu_0\right), \tag{B5}$$

with $\bar{z}$ the sample average of z_t conditional on $\mathbf{B}$.

We simulate from the joint posterior by iterating over the sequence of conditional posteriors. For the initialization of the Gibbs sampler we use the OLS estimates. We start the Gibbs sampler using 2500 draws that we discard. For the risk and portfolio analysis we use the subsequent sample of 20,000 parameter draws. Conditional on each draw for the parameters we simulate two antithetic scenarios of future returns. In this way we create 40,000 scenarios of future returns.

The burn-in phase is chosen by visual inspection of the posterior draws and supported by the convergence tool of Yu and Mykland (1998). We use a standardized version of their cumsum statistic as suggested by Bauwens *et al* (2003). For all priors the plot of the standardized version of the cumsum statistic converges smoothly and quickly to zero, especially after the burn-in phase, which indicates the convergence of the Monte Carlo chain. Sufficient conditions for convergence of the Gibbs sampler are given in Geweke (1996). Plots of the autocorrelation function suggest that the draws do not suffer from serious autocorrelation.

References

Ang, A. and G. Bekaert (2007), Stock Return Predictability: Is it There?, *Review of Financial Studies*, **20**, 651–707.

Avramov, D. (2002), Stock Return Predictability and Model Uncertainty, *Journal of Financial Economics*, **64**, 423–458.

Avramov, D. (2004), Stock Return Predictability and Asset Pricing Models, *Review of Financial Studies*, **17**, 699–738.

Avramov, D. and G. Zhou (2010), Bayesian Portfolio Analysis, *Annual Review of Financial Economics*, **2**, 25–47.

Barberis, N. (2000), Investing for the Long Run when Returns are Predictable, *Journal of Finance*, **55**, 389–406.

Bauwens, L., L.M. Lubrano, and J.F. Richard (2003), *Bayesian Inference in Dynamic Econometric Models*, Oxford University Press, New York.

Black, F. and Litterman (1992), Global Portfolio Optimization, *Financial Analysts Journal*, **48**, 28–43.

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

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

Campbell, J.Y., A.W. Lo, and A.C. MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton University Press.

Campbell, J.Y. and S. Thompson (2007), Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?, *Review of Financial Studies*, forthcoming.

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

Campbell, J.Y. and L.M. Viceira (2005), The Term Structure of the Risk-Return Tradeoff, *Financial Analysts Journal*, **61**, 34–44.

Chopra, V.K. and W.T. Ziemba (1993), The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice, *Journal of Portfolio Management*, 6–11, winter.

Claus, J. and J. Thomas (2001), Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stocks Markets, *Journal of Finance*, **56**, 1629–1666.

Cremers, K.J.M. (2002), Stock Return Predictability: A Bayesian Model Selection Perspective, *Review of Financial Studies*, **15**, 1223–1249.

Dimson, E., P. Marsh, and M. Staunton (2002), *Triumph of the Optimists*, Princeton University Press (Princeton).

Duffee, G.R. (2002), Term Premia and Interest Rate Forecasts in Affine Models, *Journal of Finance*, **57**, 405–443.

Fama, E.F. and K.R. French (1989), Business Conditions and the Expected Returns on Stocks and bonds, *Journal of Financial Economics*, **25**, 23–49.

Geweke, J. (1996), Monte Carlo Simulation and Numerical Integration, in H. Amman, D. Kendrick, and J. Rust, (eds.), *Handbook of Computational Economics*, North-Holland.

Goldfarb, D. and G. Iyengar (2003), Robust Portfolio Selection Problems, *Mathematics of Operations Research*, **28**, 1–38.

Goyal, A. and I. Welch (2008), A Comprehensive Look at The Empirical Performance of Equity Premium Prediction, *Review of Financial Studies*, **21**, 1455–1508.

Graham, J.R. and C.R. Harvey (2010), The Equity Risk Premium in 2010, SSRN working paper 1654026.

Gürkaynak, R.S., B. Sack, and J.H. Wright (2007), The U.S. Treasury Yield Curve: 1961 to the Present, *Journal of Monetary Economics*, **54**, 2291–2304.

Hoevenaars, R.P.M.M., R.D.J. Molenaar, P.C. Schotman, and T.B.M. Steenkamp (2008), Strategic Asset Allocation with Liabilities: Beyond Stocks and Bonds, *Journal of Economic Dynamics and Control*, **32**, 2939–2970.

Hollifield, B., G. Koop, and K. Li (2003), A Bayesian Analysis of a Variance Decomposition for Stock Returns, *Journal of Empirical Finance*, **10**, 583–601.

Johannes, M., A. Korteweg, and N. Polson (2008), Sequential Learning, Predictive Regressions, and Optimal Portfolio Returns, SSRN working paper 1108905.

Jorion, P. (1986), Bayes-Stein Estimation for Portfolio Analysis, *Journal of Financial and Quantitative Analysis*, **21**, 279–292.

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

Lutgens, F. and P.C. Schotman (2010), Robust Portfolio Optimisation with Multiple Experts, *Review of Finance*, **14**, 343–383.

Mehra, R. and E.C. Prescott (1985), The Equity Premium Puzzle, *Journal of Monetary Economics*, **15**, 145–161.

Pastor, L. and R.F. Stambaugh (2000), Comparing Asset Pricing Models: An Investment Perspective, *Journal of Financial Economics*, **56**, 335–381.

Pastor, L. and R.F. Stambaugh (2010), Are Stocks Really Less Volatile in the Long Run?, SSRN working paper 1349590.

Pettenuzzo, D. and A. Timmermann (2010), Predictability of Stock Returns and Asset Allocation under Structural Breaks, *Journal of Econometrics*, forthcoming.

Rustem, B., R. Becker, and W. Marty (2000), Robust Min-max Portfolio Strategies for Rival Forecast and Risk Scenarios, *Journal of Economic Dynamics and Control*, **24**, 1591–1621.

Schotman, P.C. (1994), Priors for the AR(1) Model: Parameterization Issues and Time Series Considerations, *Econometric Theory*, **10**, 579–595.

Shanken, J.A. and A. Tamayo (2004), Dividend Yield, Risk, and Mispricing: A Bayesian Analysis, SSRN working paper 548364.

Uppal, R. and T. Wang (2003), Model Misspecification and Underdiversification, *Journal of Finance*, **58**, 2465–2486.

Villani, M. (2005), Inference in Vector Autoregressive Models with an Informative Prior on the Steady State, Sveriges Riksbank working paper 181.

Villani, M. (2009), Steady-State Priors for Vector Autoregressions, *Journal of Applied Econometrics*, **24**, 630–650.

WACHTER, J.A. AND M. WARUSAWITHARANA (2009), Predictable returns and Asset Allocation: Should a Skeptical Investor Time the Market?, *Journal of Econometrics*, **148**, 162–178.

WANG, Z. (2005), A Shrinkage Approach to Model Uncertainty and Asset Allocation, *Review of Financial Studies*, **18**, 673–705.

WELCH, I. (2001), The Equity Premium Consensus Forecast Revisited, Cowles foundation discussion paper 1325.

XIA, Y. (2001), Learning about Predictability: the Effect of Parameter Uncertainty on Dynamic Asset Allocation, *Journal of Finance*, **56**, 205–246.

YU, B. AND P. MYKLAND (1998), Looking at Markov samplers through cusum path plots: A simple diagnostic idea, *Statistics and Computing*, **8**, 275–286.

Table 1: Variance of the conditional mean

The entries report the asymptotic variances and covariance that sum up to the total variance of the conditional mean in (22).

component	variance
$(d_t - \mu_d)^2 \hat{\psi}_k^2 \mathsf{Var}[\hat{\beta}]$	0.0026
$2(d_t - \mu_d)^2 \hat{\beta} \hat{\psi}_k \mathsf{Cov}[\hat{\beta}, \Psi_k(\hat{\rho}_d)]$	-0.0030
$(d_t - \mu_d)^2 \hat{\beta}^2 \mathsf{Var}[\hat{\psi}_k]$	0.0009
sum	0.0005

Table 2: Summary Statistics

The first four columns report means ("ave"), standard deviations ("stdev") and autocorrelations ("AR(1)") for the entire sample (1952:II - 2008:IV). Summary statistics are based on logarithmic returns in percent per quarter. Standard errors of the mean ("se") are computed using the Newey-West estimate of the long-run variance.
The final three columns report the implied annualized means for simple returns by adding half the variance and multiplying by four. The means are reported for both the full sample as well as for the two subsamples based on NBER business cycles that are used as optimist and pessimist priors.
Variables are real 3-months T-Bill return (r_{tb}), excess stock returns (x_s), excess bond returns (x_b), nominal Treasury Bill return (i), log dividend yield (d) and term spread (S).

	Full sample, quarterly				Annualized, simple		
	ave	se	stdev	AR(1)	full	pessimist	optimist
i	1.23	(0.28)	1.35	0.93	4.92	5.23	4.86
r_{tb}	0.32	(0.15)	1.33	0.46	1.26	1.47	1.22
d	-3.49	(0.20)	0.41	0.98	-3.49	-3.27	-3.53
S	0.30	(0.05)	0.59	0.80	1.21	1.33	1.19
x_s	1.18	(0.58)	14.91	0.13	5.83	-1.88	7.31
x_b	0.16	(0.24)	7.83	-0.02	0.96	6.66	-0.15

Table 3: Parameter estimates

The table reports OLS parameter estimates of the VAR $y_{t+1} = c + \mathbf{B}y_t + \epsilon_{t+1}$ with variables: real 3-months T-Bill return (r_{tb}), excess stock returns (x_s), excess bond returns (x_b), nominal Treasury Bill return (i), dividend yield (d) and term spread (S). Standard errors are in parentheses. The last column contains the R^2.

	i_t	$r_{tb,t}$	d_t	S_t	$x_{s,t}$	$x_{b,t}$	R^2
i_{t+1}	0.96	0.00	0.04	0.08	0.00	0.00	0.89
	(0.03)	(0.03)	(0.04)	(0.06)	(0.00)	(0.00)	
$r_{tb,t+1}$	0.21	0.46	-0.03	0.27	0.00	-0.01	0.27
	(0.07)	(0.07)	(0.10)	(0.15)	(0.01)	(0.01)	
d_{t+1}	0.01	-0.01	0.97	0.00	0.00	0.00	0.97
	(0.01)	(0.01)	(0.01)	(0.02)	(0.00)	(0.00)	
S_{t+1}	0.02	-0.01	-0.03	0.82	0.00	0.00	0.66
	(0.02)	(0.02)	(0.03)	(0.05)	(0.00)	(0.00)	
$x_{s,t+1}$	-1.58	0.95	3.64	0.10	0.11	0.25	0.09
	(0.91)	(0.88)	(1.28)	(1.88)	(0.07)	(0.14)	
$x_{b,t+1}$	0.46	0.39	-0.42	3.32	-0.08	-0.09	0.07
	(0.48)	(0.47)	(0.68)	(0.99)	(0.04)	(0.07)	

Table 4: Residual correlation matrix

The table reports the residual correlation matrix $\hat{\Sigma}$ for the OLS estimates of the VAR $y_{t+1} = c + \mathbf{B}y_t + \epsilon_{t+1}$. Diagonal entries are standard deviations; off-diagonal entries are correlations.

	i	r_{tb}	d	S	x_s	x_b
Nominal T-Bill rate (i)	0.23	—	—	—	—	—
Real T-Bill rate (r_{tb})	-0.34	0.58	—	—	—	—
Log dividend yield (d)	0.05	-0.08	0.07	—	—	—
Term spread (S_t)	-0.83	0.16	-0.03	0.18	—	—
Excess stock returns ($x_{s,t}$)	-0.03	0.06	-0.98	0.02	7.23	—
Excess bond returns ($x_{b,t}$)	-0.65	0.40	0.01	0.14	-0.02	3.83

Table 5: Posterior means of selected VAR parameters

This table shows the effect of different priors on selected important VAR parameters $\mathbf{B}_{ij}$. *OLS* are the least squares estimates with standard errors in parentheses. All other entries denote posterior means with posterior standard deviations in parentheses. (O = *Optimist*, P = *Pessimist*, 100 = *Dogmatic*, 01 = *Uninformative*)

parameter	*OLS*	*Flat*	*P01*	*P1*	*P100*	*O01*	*O1*	*O100*
i_{t+1}, i_t	0.961	0.958	0.967	0.957	0.957	0.967	0.960	0.958
	(0.028)	(0.027)	(0.026)	(0.026)	(0.026)	(0.026)	(0.027)	(0.027)
d_{t+1}, d_t	0.968	0.968	0.972	0.986	0.979	0.972	0.969	0.968
	(0.013)	(0.013)	(0.013)	(0.009)	(0.010)	(0.013)	(0.012)	(0.013)
$x_{s,t+1}$, i_t	-1.58	-1.53	-1.49	-1.11	-1.39	-1.50	-1.51	-1.55
	(0.91)	(0.91)	(0.88)	(0.94)	(0.97)	(0.88)	(0.90)	(0.91)
$x_{s,t+1}$, d_t	3.64	3.71	3.30	1.67	0.82	3.28	3.60	3.53
	(1.28)	(1.27)	(1.28)	(0.90)	(1.15)	(1.25)	(1.21)	(1.24)
$x_{b,t+1}$, i_t	0.46	0.52	0.35	0.35	0.47	0.37	0.48	0.52
	(0.48)	(0.46)	(0.45)	(0.47)	(0.48)	(0.45)	(0.46)	(0.46)
$x_{b,t+1}$, S_t	3.32	3.36	3.17	2.97	3.73	3.19	3.29	3.39
	(0.99)	(0.99)	(1.00)	(0.99)	(1.02)	(1.00)	(0.99)	(1.00)

Table 6: Unconditional means

This table shows the effect of different priors on the unconditional means. *sample* denotes the sample average with the Newey-west standard error in parentheses; *VAR* reports the implied means $\hat{\mu} = (I - \hat{\mathbf{B}})^{-1}\hat{c}$ for the OLS estimates with asymptotic standard errors in parentheses. The other entries denote posterior means (standard deviations) for the uninformative priors (O = *Optimist*, P = *Pessimist*, 01 = *Uninformative*, 100 = *Dogmatic*). Units for returns are percent per quarter and refer to logarithmic returns. No results are shown for the *Flat* prior, since posterior moments for the unconditional means implied by the *Flat* prior do not exist.

	sample	*VAR*	*P01*	*P100*	*O01*	*O100*
Nominal rate	1.23	1.06	1.15	1.32	1.07	1.21
	(0.28)	(0.45)	(0.73)		(0.73)	
Real T-Bill	0.31	0.27	0.29	0.37	0.28	0.29
	(0.15)	(0.14)	(0.23)		(0.24)	
Dividend Yield	-3.49	-3.62	-3.52	-3.25	-3.63	-3.54
	(0.20)	(0.23)	(0.39)		(0.38)	
Term Spread	0.30	0.32	0.30	0.33	0.32	0.30
	(0.05)	(0.06)	(0.10)		(0.09)	
Equity premium	1.18	0.93	0.93	-1.11	0.92	1.52
	(0.58)	(0.26)	(0.45)		(0.45)	
Bond Premium	0.16	0.21	0.19	1.48	0.21	-0.08
	(0.24)	(0.09)	(0.14)		(0.15)	

Table 7: Portfolio choice, investment horizon and priors

Optimal portfolio choice under power utility for an investor with a k-period investment horizon and rebalancing to a constant mix periodically with risk aversion $\gamma = 5$ and for different investment horizons and different priors (*O = optimistic. P = pessimistic. 01 = uninformative. 100 = dogmatic*). Short-selling restrictions have been imposed. All results are based on simulations using 40.000 scenarios.

	OLS	*flat*	*P01*	*P1*	*P100*	*O01*	*O1*	*O100*
1 year horizon								
Bills	0.28	0.35	0.35	0.20	0.01	0.35	0.45	0.37
Bonds	0.24	0.21	0.24	0.61	0.99	0.21	0.10	0.00
Equity	0.48	0.44	0.41	0.19	0.00	0.44	0.45	0.63
5 years horizon								
Bills	0.34	0.47	0.48	0.30	0.01	0.47	0.47	0.29
Bonds	0.03	0.01	0.09	0.55	0.99	0.07	0.00	0.00
Equity	0.63	0.52	0.43	0.15	0.00	0.46	0.53	0.71
10 years horizon								
Bills	0.20	0.43	0.48	0.24	0.01	0.47	0.39	0.22
Bonds	0.00	0.00	0.03	0.59	0.99	0.00	0.00	0.00
Equity	0.80	0.57	0.49	0.17	0.00	0.53	0.61	0.78
15 years horizon								
Bills	0.05	0.41	0.50	0.20	0.01	0.45	0.36	0.23
Bonds	0.00	0.00	0.00	0.61	0.99	0.00	0.00	0.00
Equity	0.95	0.59	0.50	0.19	0.00	0.55	0.64	0.77

Table 8: Certainty equivalent costs and robust portfolio choice

Certainty equivalent costs (in percentages of initial wealth) if the strategic asset allocation decision is based on prior $p_j(\mu)$ while the investor has prior $p_i(\mu)$. The rows refer to the prior used in the evaluation, while the columns indicate the prior used for constructing an optimal portfolio. Comparisons of $p_i(\mu)$ with itself have zero costs and are empty. The maximum costs in each column (for each horizon) are in **boldface** and the robust portfolio is in ***boldface italics***.

	Portfolio							
Prior	OLS	*Flat*	*P01*	*P1*	*P100*	*O01*	*O1*	*O100*
1 year horizon								
Pessimist *P01*	0.03	0.01		0.48	1.80	0.01	0.04	0.38
Pessimist *P1*	0.76	0.65	0.53		0.47	0.65	0.84	1.83
Pessimist *P100*	**4.89**	**4.61**	**4.19**	1.32		**4.61**	**5.23**	**7.87**
Optimist *O01*	0.01	0.00	0.00	0.59	2.00		0.02	0.29
Optimist *O1*	0.04	0.02	0.04	0.79	2.35	0.02		0.20
Optimist *O100*	0.30	0.35	0.45	***1.88***	**4.10**	0.35	0.24	
5 years horizon								
Pessimist *P01*	0.21	0.04		0.59	1.79	0.00	0.05	0.41
Pessimist *P1*	1.75	1.19	0.73		0.41	0.87	1.25	2.32
Pessimist *P100*	**8.60**	**7.24**	**5.73**	1.48		**6.17**	**7.43**	**10.01**
Optimist *O01*	0.15	0.02	0.01	0.71	1.99		0.02	0.33
Optimist *O1*	0.07	0.00	0.05	0.94	2.27	0.03		0.18
Optimist *O100*	0.07	0.20	0.50	***2.32***	**4.33**	0.40	0.17	
10 years horizon								
Pessimist *P01*	0.52	0.03		0.67	1.77	0.01	0.07	0.45
Pessimist *P1*	2.72	1.26	0.90		0.34	1.08	1.45	2.56
Pessimist *P100*	**12.79**	**8.41**	**7.11**	1.51		**7.84**	**9.02**	**12.31**
Optimist *O01*	0.43	0.01	0.01	0.76	1.93		0.04	0.37
Optimist *O1*	0.18	0.01	0.07	1.01	2.19	0.03		0.14
Optimist *O100*	0.00	0.25	0.49	***2.56***	**4.44**	0.35	0.17	
15 years horizon								
Pessimist *P01*	3.24	0.08		0.69	1.79	0.02	0.19	0.89
Pessimist *P1*	6.74	1.33	0.90		0.37	1.11	1.66	3.09
Pessimist *P100*	**17.32**	**9.05**	**7.63**	1.37		**8.39**	**9.96**	**12.70**
Optimist *O01*	1.00	0.01	0.01	0.77	1.93		0.05	0.30
Optimist *O1*	0.55	0.01	0.10	1.01	2.18	0.04		0.09
Optimist *O100*	0.31	0.22	0.48	***2.57***	**4.47**	0.33	0.12	

Figure 1: Persistence and T-Bill risk

The figure shows the annualized volatility of real T-Bill returns as a function of the persistence of the nominal interest rate dynamics for three different returns horizons (1 year. 5 years and 15 years). Estimates are implied by model (16) for $\gamma = 0.42$ and $\beta = 0.21$. The dashed line is at the point estimate $\hat{\rho} = 0.967$

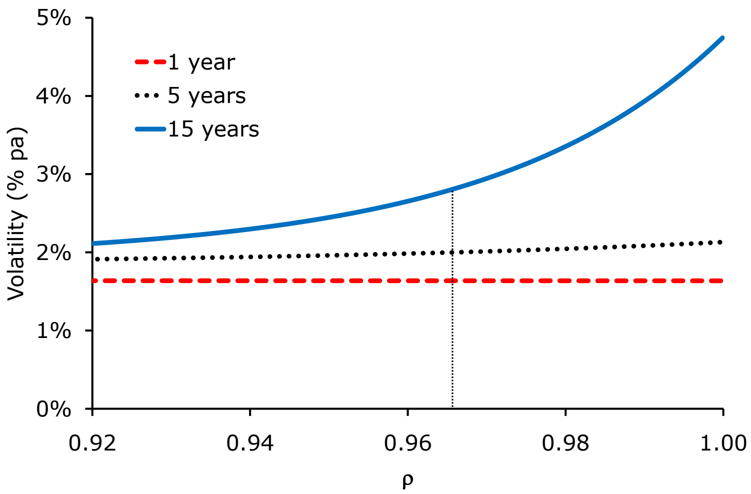


Figure 2: Informative prior on unconditional mean

The left panel of the figure shows the implied marginal prior density on $c = \mu(1-\rho)$ with a flat prior on $\rho \in [0.1]$ and a normal prior $\mu \sim N(1.2^2)$. The right panel shows two conditional prior densities on ρ implied by the same prior. Densities shown are for $c = 0.05$ and $c = 0.5$.

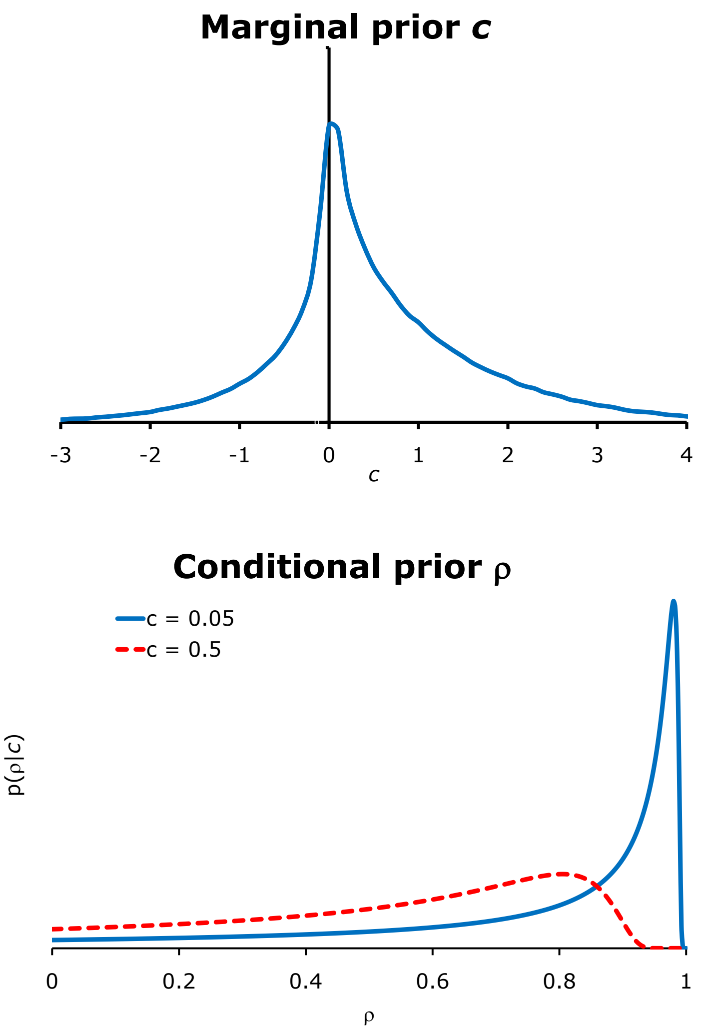


Figure 3: Posterior density of maximum eigenvalue

The figure shows the posterior density of the maximum eigenvalue λ_{max} in the VAR system. *Flat* is the uniform benchmark prior; *Uninformed* refers to the priors with $\kappa = 0.01$ (optimist and pessimist are indistinguishable). *Optimist* and *Pessimist* refer to the dogmatic priors with $\kappa = 100$.

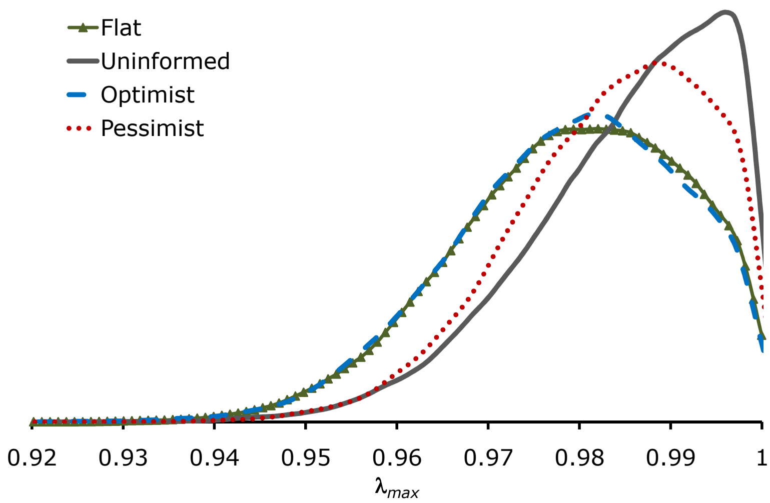
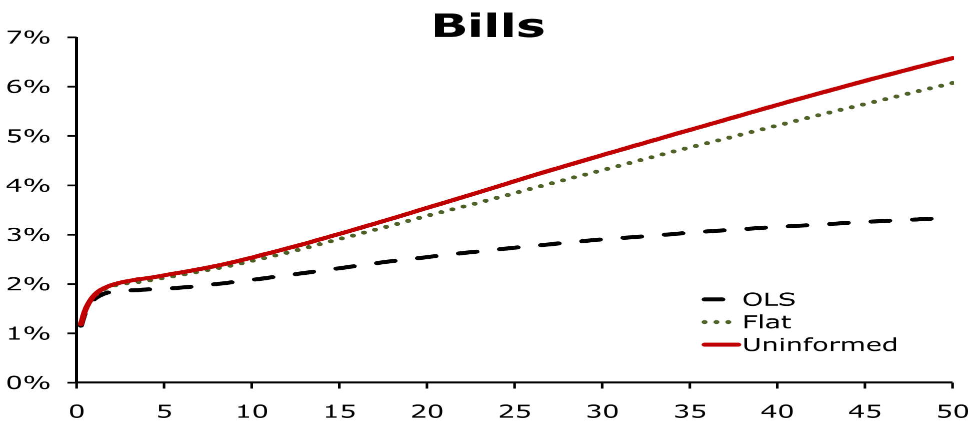
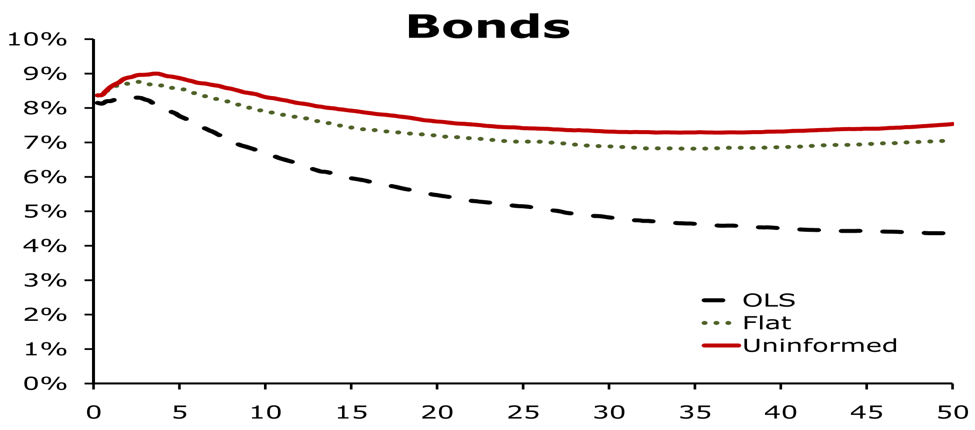


Figure 4: Term structure of risk with parameter uncertainty

The figure shows the estimated term structures of risk, defined as the annualized predictive standard deviations, for Bills, Bonds and Equity. *OLS* is the standard deviation conditional on the OLS estimates of the VAR. The other lines are Bayesian predictive standard deviations, based on either the *flat* prior, or the *uninformative* ($\kappa = 0.01$) prior on the unconditional mean. Lines for *optimist* and *pessimist* are not shown individually as the two are indistinguishable. The horizontal axis denotes the horizon in years.





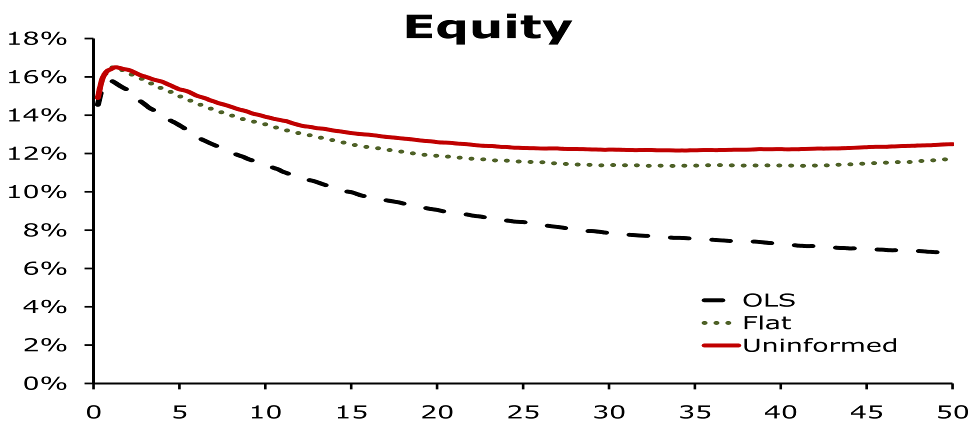


Figure 5: Term structure of risk and informative priors

The figure shows the estimated term structures of risk, defined as the Bayesian annualized predictive standard deviations, for Bills, Bonds and Equity. The estimates use either an *uninformative* ($\kappa = 0.01$) or *dogmatic* ($\kappa = 100$) prior on the unconditional mean. For the *dogmatic* priors the figure shows risk estimates for both the *optimist* and the *pessimist* prior. For the *uninformative* priors these are not shown individually as the two are indistinguishable. The horizontal axis denotes the horizon in years.

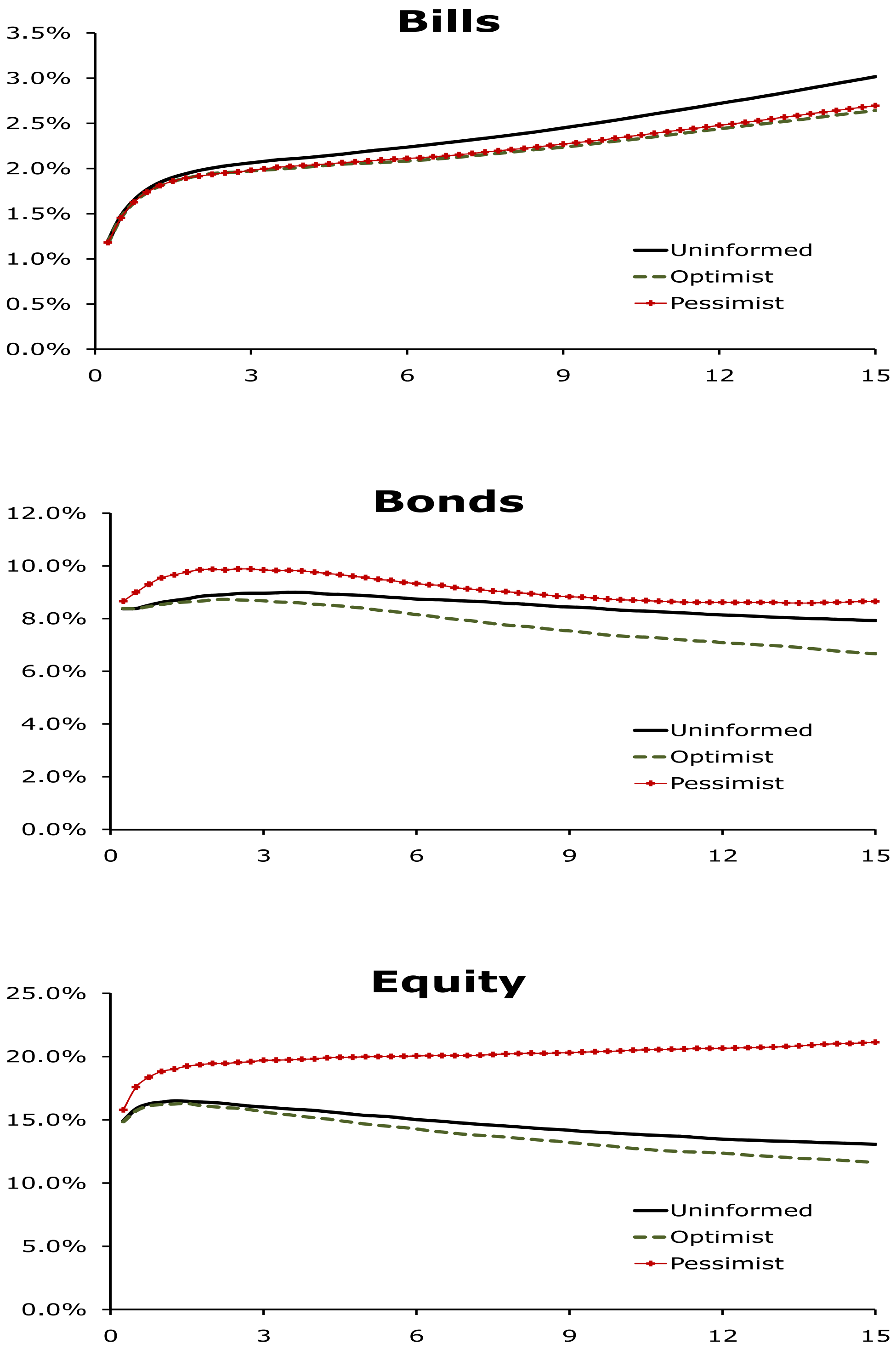


Figure 6: Decomposition of equity risk

The figure shows the term structure of risk for real equity returns. *OLS* denotes the term structure conditional on the *OLS* estimates (identical to the OLS curve in the lower panel of figure 4. The other lines are derived from the predictive scenarios based on the *uninformative* prior (*optimist* and *pessimist* are almost identical).

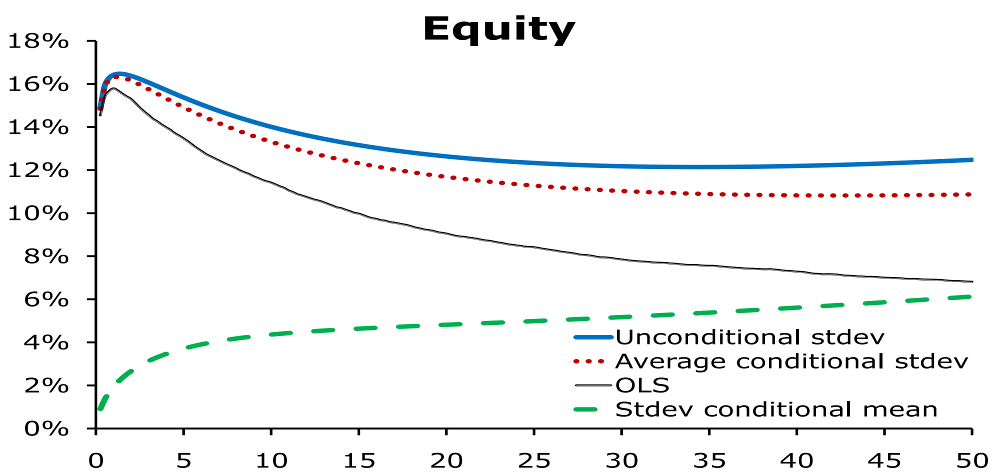


Figure 7: Correlations and prior information

The figure shows the effect of prior information on the correlation among real holding period returns of stocks, bonds and bills. *OLS* are the correlations conditional on the OLS estimates of the VAR. The other lines are Bayesian predictive correlations, based on either the *flat* prior, or the *dogmatic* ($\kappa = 100$) *optimist* and *pessimist* priors on the unconditional means. Not shown is the line for the *uninformative* priors, since it is indistinguishable from the *flat* prior. The horizontal axis denotes the horizon in years.

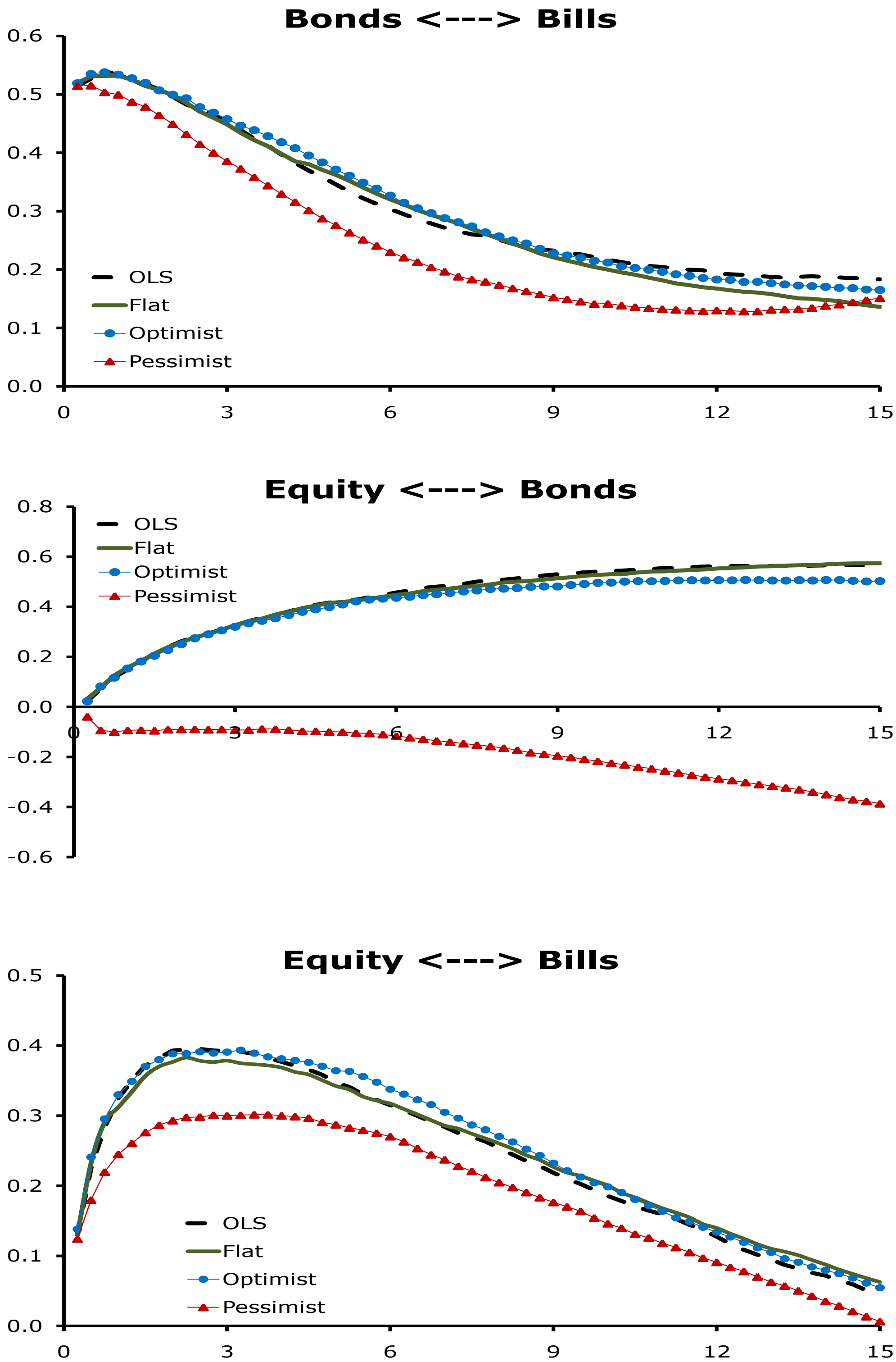


Figure 8: Inflation hedge

The figures show the correlation between shocks to cumulative nominal returns and shocks to cumulative inflation over different horizons. *OLS* is the correlation conditional on the parameter estimates. The Bayesian estimates are based on the *flat* prior and the *optimist* and *pessimist* priors for the *dogmatic* design ($\kappa = 100$). On the horizontal axis is the horizon in years; on the vertical axis the correlation.

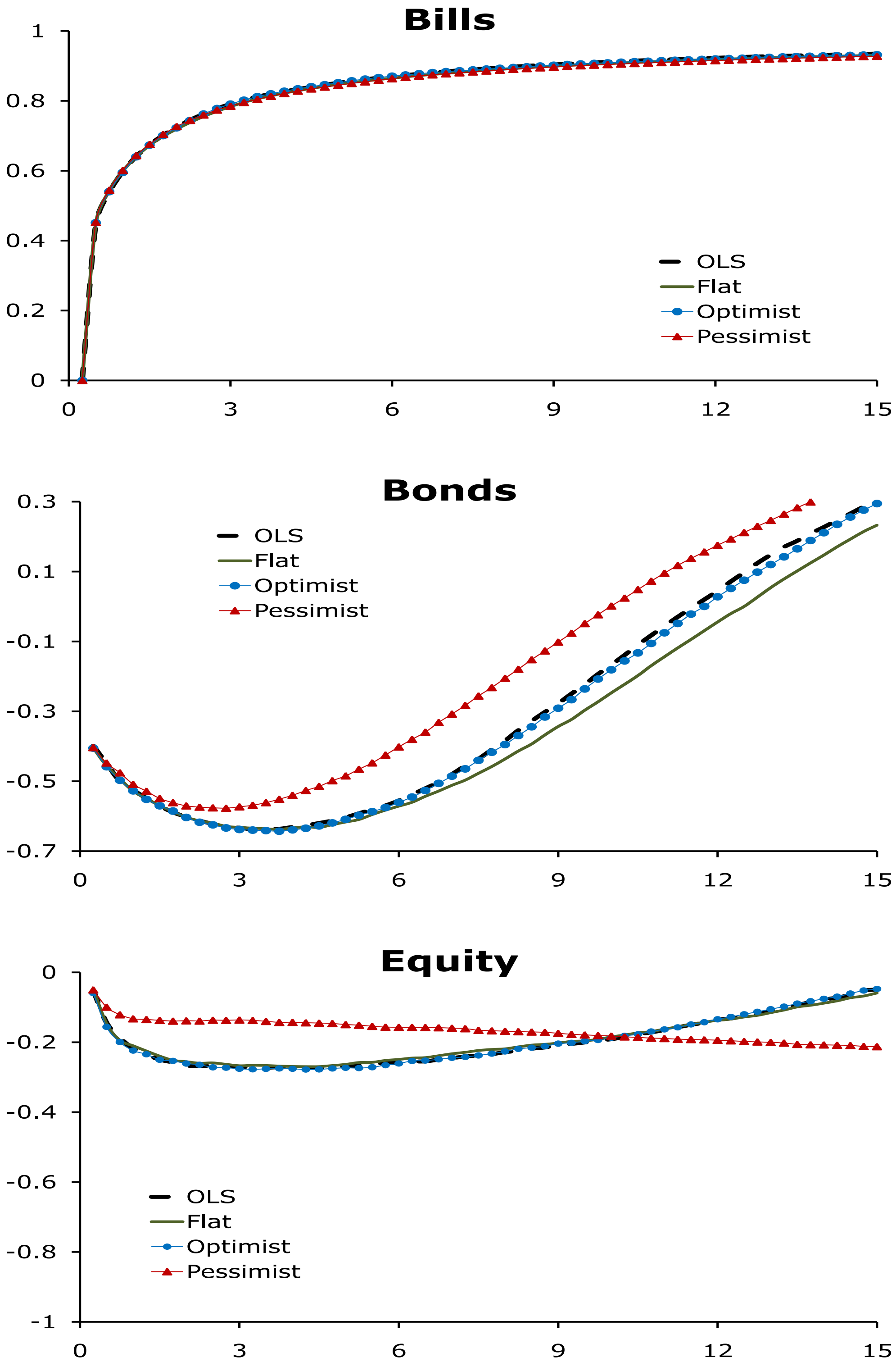


Figure 9: Portfolio holdings

The figure shows the optimal asset allocation for the *OLS* estimates and for different Bayesian priors as a function of the risk aversion parameter γ. The investment horizon is 15 years.