Portfolio Allocation for European markets with Predictability and Parameter Uncertainty

Eric Jondeau
University of Lausanne
Swiss Finance Institute
Eric.Jondeau@unil.ch

Michael Rockinger
University of Lausanne
Swiss Finance Institute
CEPR
Michael.Rockinger@unil.ch

August 2009

Abstract

This paper addresses the portfolio choice problem of an investor with a long investment horizon when stock returns are partly predictable. We begin with a discussion of the predictability of stock returns, both in-sample and out-of-sample, which allows us to confirm, for several European markets, by using several macroeconomic variables as predictors, that there is evidence for in-sample predictability of the dividend-price ratio as well as of inflation. Out-of-sample prediction is poor. To address the question of the economic relevance of this weak predictability, we consider a buy-and-hold as well as a dynamic multi-period investment strategy that follows Barberis (2000) and which incorporates estimation risk. It is demonstrated that the taking into account of even weak predictability may lead to significant different allocations. In particular, for the UK, we demonstrate strong horizon effects and hedging demand. In the multiperiod allocation, parameter uncertainty is dominated by the variation in the explanatory variable.

Keywords: Stock returns, Predictability, Estimation risk, Portfolio choice.

JEL classification: C11, C22, C32, C51, C61, G11.

We are most grateful to Ioana Covrig for exceptional research assistance. Neither the Swiss Finance Institute nor the University of Lausanne necessarily share the opinion expressed in this paper. This paper was written as a tender by the Observatoire de l'Epargne Européenne. We are most grateful to the Observatoire de l'Epargne Européenne for their support.

1 Introduction

With an ever-increasing life span, there is more and more social responsibility on institutional investor and pension funds to allocate assets, which have been entrusted to them, in an optimal manner. Also, the allocation of wealth needs to be made on a long-term basis, where the institutional investor needs to take into account that at times markets may boom and at others go through a trough. From an academic point of view, there is some guidance on how such an allocation may look. Early work by Samuelson (1969) demonstrates that if asset returns are independent and identically distributed and if investors preferences do not change over time, then a simple buy-an-hold strategy is optimal. Merton (1969, 1971, 1973) discusses the multi-period portfolio allocation under the assumption that returns' distribution fluctuates over time. Merton demonstrates that in such a case an investor may have an incentive to hedge against adverse movements of returns. He also shows that, if preferences of an investor depend on certain state variables, which may move in adverse ways, the investor may consider offsetting portfolio allocations. From an empirical point of view, Campbell and Viceira (1999), Kim and Omberg (1996), and Watcher (2002) demonstrate the economic relevance of such hedging demands. Balduzzi and Lynch (1999) incorporate in addition transaction costs and confirm that hedging demand subsists.

This paper contributes to this literature by considering European markets and by showing that there is some evidence that asset returns are predictable. This preliminary research does, however, not yield uniform findings. The UK market is rather well predicted by the dividend-price ratio (a variable that appears to play a role also for other markets). For other countries, such as Spain, inflation turns out to be an additional predictor. Also, as we will establish, out-of-sample predictability does not remain statistically significant. Thus, at that stage, our work corroborates the one of Bossaerts and Hillion (1999) and Goyal and Welch (2004), which demonstrates that there is little predictability in stock returns out-of-sample.

Recent theoretical and empirical work by Cochrane (2001) argues, however, that based on theoretical grounds returns are predictable. His reasoning is based on Gordon's formula relating current prices with future dividends. An algebraic manipulation reveals then that there exists a causal link between dividend-price ratios and future returns. Cochrane (2001) also demonstrates that econometric tests of predictability lack statistical power. Thus, it is not because statistical tests fail to detect evidence for predictability by dividend-price ratios that there is indeed no predictability.[1]

[1] Many other contributions establish that asset returns are predictable, such as Campbell (1987, 1996), Campbell and Schiller (1988), Fama and French (1988, 1989), Poterba and Summers (1988), Keim and Stambaugh (1986), Barberis (2000), Lettau and Ludvigson (2001), Ferreira and Santa Clara (2008), Campbell and Thompson (2008), or Bec and Gollier (2009).

A related issue that compounds the difficulty to forecast future stock returns is the uncertainty that surrounds the parameter estimates which are required to calibrate the portfolio choice models. Intuitively, estimation risk represents an additional source of uncertainty, which, if incorporated into the portfolio choice model, should yield more precautious allocations. This problem was already studied by Williams (1977), Gennotte (1986), and Pastor and Stambaugh (2008). To account for estimation risk, a Bayesian approach may prove useful. Indeed, Bayesian analysis provides guidance how information contained in actual data may be combined with weak prior information to yield better information on the distributional properties of parameters.

The question that we want to address in this contribution is: what is the economic relevance of weak predictability of stock returns, in the presence of parameter uncertainty and in a long run perspective? Should institutional managers become passive investors whose objective is to duplicate indices and whose only task is to decide on static buy-and-hold allocations? If predictability is weak, it may still be of economic relevance, especially for the long run where one may expect small effects to conjugate. In the literature, Brennan, Schwartz, and Lagnado (1997) analyze numerically the impact of myopic versus dynamic portfolio choice. They find that, due to mean reversion in stock and bond returns, the investor with a long horizon will place a larger fraction of her wealth in stocks and bonds than an investor with a short horizon will do. Kandel and Stambaugh (1996) consider the effects of predictability within a myopic setting, where the investor rebalances monthly and where parameter uncertainty is taken into account by using a Bayesian setting. Barberis (2000) studies these effects for the multi-period portfolio choice, with annual rebalancing in discrete time. All these contributions tend to show that the various effects are important.

In this work, we reconsider Barberis' (2000) model, which provides a simple setting within which the effects of predictability under parameter uncertainty may be investigated. By applying this setting to several European markets, we confirm that weak predictability affects rather significantly long-run portfolio allocations. We also show that, by using the Bayesian learning model of Barberis (2000), parameter uncertainty is important but may be in certain cases dominated by the predictability of returns. For instance, if we consider a buy-and-hold strategy for the long-run, an investor who takes parameter uncertainty into account under predictable returns will allocate more to stocks than a naive investor who ignores parameter uncertainty while assuming i.i.d stock returns.

The framework proposed by Barberis (2000) is moreover sufficiently rich to allow the investigation of dynamic trading strategies where an investor gets the possibility to modify the portfolio at regular periods. Such a strategy may be particularly interesting if there is predictability. In addition, it allows us to investigate the role played by parameter uncertainty. Indeed, if the state variables that predict returns change sufficiently over time, one may expect that, even if there is parameter uncertainty, this uncertainty is dominated by the first-order effects which stem from time changes of the predictors. Indeed, we are able to confirm in certain cases that parameter uncertainty is dominated by fluctuations of the predictors. This implies that the role of parameter uncertainty may be diminished in a dynamic setting.

The rest of the paper is organized as follows. In Section 2, we evaluate the predictability in European stock markets, with a particular focus on out-of-sample predictability and a distinction between frequentist and Bayesian parameter uncertainty accounting. In Section 3, we consider the consequences of stock return predictability on the optimal asset allocation from a long-term perspective. The final conclusions are presented in Section 4.

2 Stock Return Predictability

In this first part of the paper, we wish to review stock return predictability. We will perform this investigation first by focusing on an univariate setting and then on a multivariate one. The univariate setting is of interest on its own, since it allows us to investigate the time-series properties of returns for a short horizon. Within this setting, we examine the ability to predict stock returns, but only for a one-period ahead horizon. Later on, we will investigate multi-period stock allocations. In such a setting, it is necessary to also predict future explanatory variables, which will in turn predict future one-period ahead stock returns. We will then consider return predictability in a multivariate setting. Finally, in order to account for parameter uncertainty, we will follow Barberis (2000) and cast the multivariate model into a Bayesian setting.

2.1 Methodology to Test Predictability

2.1.1 Univariate and Multivariate Models

We consider the following univariate regression model:

$$r_t = a + BX_{t-1} + e_t, \tag{1}$$

where

$$X_t = \begin{pmatrix} r_t & r_{f,t} & r_{b,t} & dp_t & spr_t & \pi_t \end{pmatrix}'$$

is the set of possible explanatory variables and e_t is the vector of innovations, assumed to be i.i.d., normally distributed with mean 0 and variance σ_ε^2. The vector of constant parameters B contains sensitivities of stock returns to predictive variables. The set of explanatory variables consists of lags of: log excess return (r), log excess return on bonds (r_b), log short term real interest rate (r_f), log dividend-to-price ratio (dp), the yield spread (spr) between the long-term bond yield and short-term interest rate, and the log yearly inflation (π). The actual construction of the variables is discussed in Appendix 1.

The multivariate regression model which we borrow from Barberis (2000) is given by

$$\begin{pmatrix} r_t \\ x_t \end{pmatrix} = \begin{pmatrix} a_r \\ a_x \end{pmatrix} + \begin{pmatrix} B_r \\ B_x \end{pmatrix} x_{t-1} + \begin{pmatrix} \varepsilon_{r,t} \\ \varepsilon_{x,t} \end{pmatrix}, \tag{2}$$

where x_t represents a list of n explanatory variables. Typically, this list will represent a subset of X_t, say $x_t = (dp_t, \pi_t)'$. $\varepsilon_{r,t}$ and $\varepsilon_{x,t}$ denote error terms. The parameter a_r represents a real constant, B_r a $(1 \times n)$ vector of constants. By a_x and B_x we denote a $(n \times 1)$ vector and a $(n \times n)$ matrix, respectively. Clearly, equation (1) allows for a one-step ahead only forecast, whereas equation (2) will allow us to consider the performance in multi-step ahead forecasts.[2] This first-order vector autoregressive process type of dynamics (2) has been extensively used by Campbell and Shiller (1988, 1998), Barberis (2000), Bec and Gollier (2009) and others. As seen above, this specification is very general, since any type of autoregression can be rewritten as a VAR(1) model, by expanding the explanatory variables.

Parameter uncertainty will also play an important role for portfolio allocation. Presently, we want to discuss how a Bayesian setting allows us to take into account this phenomenon. We start with the case where returns are i.i.d. normal. This simple setting has pedagogical value and will allow us to demonstrate the main steps involved in Bayesian analysis.[3] So, assume we observe a vector r of T i.i.d. observations from a univariate normal distribution, $r_t \sim N(\mu, \sigma^2)$. A traditional investor who does not take parameter uncertainty into account would use the T observations and estimate μ with $\bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t$, and σ^2 with $s^2 = \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \bar{r})^2$. These parameters and the distributional assumption would also characterize future, out-of-sample returns, say form $T+1, \cdots, T+\hat{T}$. However, the stance taken in a Bayesian setting is different. The idea is that the investor may have a *prior* about returns, before observing the data. This prior could be very *informative*, for instance that returns will be between -2% and $+2\%$. It is, however, usual to assume that the observer has as little information as possible. In such an instance, one obtains *noninformative priors*. The main reason for using noninformative priors is that one lets the data speak for themselves. For the normal distribution. One may assume that μ and $\log \sigma$ are each uniformly distributed with $-\infty < \mu < \infty$ and $0 < \sigma < \infty$. It appears also reasonable to assume independence between the

[2]Careful inspection of the model reveals that, here, the model consists of a VAR(1) for x_t and a linear regression reminiscent of equation (1) to describe returns, just as in Barberis (2000). We experimented at some point with a VAR(1) model including returns. Since returns had little predictive power, we decided to give a differential treatment to the regression yielding returns.

[3]The discussion here and many more results may be found in Zellner (1971).

priors of the location and scale parameters, meaning that the joint distribution of μ and σ splits into a product of distributions. Then the joint prior distribution can be shown to be proportional to the precision ($1/\sigma$) of the returns:[4]

$$p(\mu, \sigma) \propto \frac{1}{\sigma}.$$

In Bayesian analysis, it is also customary to assume prior distributions for the parameters, the so-called *hyperdistributions*, so that after learning about the data, the distributions update in a natural manner the main distribution of the data. As has been shown by Zellner (1971), the marginal posterior density for σ is an inverted gamma and the conditional posterior density for μ given σ and the sample information is a univariate normal:

$$\sigma | r \sim IG\left(\frac{T-1}{2}, 2/\sum_{t=1}^{T}(r_t - \bar{r})^2\right), \tag{3}$$

and

$$\mu | \sigma, r \sim N\left(\bar{r}, \frac{\sigma^2}{T}\right). \tag{4}$$

Now, an investor who needs to perform a forecast of future returns could integrate out parameter uncertainty. This would lead to the so-called *predictive density.* For instance, if we focus on r_{T+1}, we have

$$\begin{aligned} p(r_{T+1}|r) &= \int_{\mu=-\infty}^{+\infty} \int_{\sigma=0}^{+\infty} p(r_{T+1}, \mu, \sigma | r) d\mu d\sigma \\ &= \int_{\mu=-\infty}^{+\infty} \int_{\sigma=0}^{+\infty} p(r_{T+1}|\mu, \sigma, r) p(\mu|\sigma, r) p(\sigma|r) d\mu d\sigma. \end{aligned}$$

This predictive density does no longer depend upon μ and σ. As we will see later on, in a more general case, for this simple situation, the predictive density is shown to be a Student-t distribution. In more complex situations, it will not be possible to obtain an explicit formula for the predictive density. In such cases, it is still possible to obtain the predictive density numerically by invoking Monte Carlo simulation. For the simple case at hand, this would boil down to performing a draw from the inverse gamma (3). Then, one would draw from equation (4), and eventually, draw r_{T+1} from $N(\mu, \sigma^2)$. By using a kernel-smoother, one could then obtain a histogram from the eventual distribution. One would find, for this case, that the distribution is a Student-t.

Let us now discuss how the predictive density of the regression model (2) could be obtained. We start by rewriting the model in a way that is more amenable for Bayesian analysis. The investor will have at disposal a sample $t = 1, \cdots, T$. This

[4]The symbol $\propto$ denotes proportionality.

sample will allow to update possible priors. Introducing $z_t' = (r_t \; x_t')'$, we may rewrite the multivariate regression model (2) as

$$\begin{pmatrix} z_2' \\ \ldots \\ z_T' \end{pmatrix} = \begin{pmatrix} 1 & x_1' \\ \cdots & \ldots \\ 1 & x_{T-1}' \end{pmatrix} \times \begin{pmatrix} a' \\ B_0' \end{pmatrix} + \begin{pmatrix} \varepsilon_2' \\ \ldots \\ \varepsilon_T' \end{pmatrix}. \tag{5}$$

In a more compact form, we obtain

$$Z = XC + E, \tag{6}$$

where Z, X, and E are $(T-1 \times n+1)$, $(T-1 \times n+1)$, and $(T-1 \times n+1)$ matrices of endogenous variables, regressors, and unobserved error terms, respectively. The rows of E are independently distributed, each with a $n+1$ dimensional normal distribution with vector mean zero and $(n+1 \times n+1)$ covariance matrix Σ. C is a $(n+1 \times n+1)$ matrix.

In order to account for parameter uncertainty, we need to determine the posterior predictive distribution, which is the distribution of future z. Following again Zellner (1971), we assume that the elements of C and of Σ are independently distributed and we assume again uninformative priors

$$\begin{aligned} p(C, \Sigma) &= p(C)p(\Sigma) \\ p(C) &= \text{constant} \\ p(\Sigma) &\propto |\Sigma|^{-(n+2)/2} \\ p(C, \Sigma) &\propto |\Sigma|^{-(n+2)/2}. \end{aligned}$$

Then, the posterior $p(C, \Sigma^{-1}|Z)$ is given by:

$$\Sigma^{-1} \; |Z \sim W(T-n-2, S^{-1}) \quad \text{or} \quad \Sigma \mid Z \sim IW(T-n-2, S),$$

and

$$C \mid \Sigma, Z \sim MN(\hat{C}, \Sigma \otimes (X'X)^{-1}),$$

where W (IW) denotes the Wishart distribution (inverse Wishart distribution) and MN denotes the matricvariate normal distribution,[5] with $S = (Z - X\hat{C})'(Z - X\hat{C})$ and $\hat{C} = (X'X)^{-1}X'Z$.

To sum up, if one had to obtain the predictive density for z_{T+1}, one would start estimating $\hat{C}$ and S. Next, one would simulate from the inverse Wishart $IW(T-n-2, S)$ and then from the $MN(\hat{C}, \Sigma \otimes (X'X)^{-1})$. Again, repeating these simulations many times, one would obtain the predictive density of z_{T+1}.

[5]Essentially, this distribution states that each of the the elements of a matrix are normal.

2.1.2 Measuring Predictability

From a traditional econometric point of view, both in-sample and out-of-sample predictability matters. For portfolio allocation, which is naturally oriented towards the future, in-sample performance of regression (1) or (2) plays a less important role than out-of-sample performance. In sample analysis only provides a first idea about the relative performance of various explanatory variables. As we will shown in the empirical section, even a model with high in-sample performance may have low out-of-sample forecasting power. For this reason, for the out-of-sample analysis, we report not only the traditional root mean squared error but several other performance measures. Overall, for each regression, we will consider the following measures, whose detailed construction is provided in Appendix 2:

- the root mean square error (**RMSE**), which measures the average of the squared residuals from the regression of the actual return on its estimated value.
- the t-test of the coefficient of the predicted returns and the R-squared of the regression of the actual returns on their out-of-sample predicted values (denoted by **tstat2** and R^2**)**.
- the percentage of correct predictions of the sign of returns (**corrs**). This measure indicates if, based on a statistical analysis, one may realize gains by purchasing indices subsequent to the prediction of increasing prices and shorting if the sign is negative.
- the **corrs** measure alone, is, however, not sufficient to measure if statistical arbitrage is possible, since the magnitude of non-predicted negative returns may undo the gains that are realized by predicted returns. We, therefore, introduce another measure, denoted (**corrupd**), giving the percentage of correct predictions of the changes in returns with respect to the current level of returns and this independently of the sign of the returns. In other words, this measure captures if returns will increase or decrease with respect to the current level of returns.
- the Akaike information criterion (**AIC**) and the Schwarz criterion, which is also called Bayesian information criterion (**BIC**).

For the out-of-sample analysis, we used an expanding rolling window. This method assumes that, at each out-of-sample date, we obtain a new set of regression coefficients (according to the data available at that date), which are used to predict next period returns only.[6]

[6]We also consider the case where we keep the parameters fixed for the out-of sample period. In this case, we estimate the parameters for a given in-sample date, and then keep these parameters

2.2 Data

We consider monthly stock market returns for a large set of European countries (France, Germany, United Kingdom, Italy, Switzerland, Sweden, Denmark, Netherlands, Norway, and Spain). The data is in local currency. End-of-month data was used to construct monthly return series. We also use a large set of macro-economic variables, available at the monthly frequency: the consumer price index, used to construct the inflation rate, the long-term rate (10-year government bond yield), the short-term rate, and the dividend-to-price ratio.

The sample starts in January 1975 for all the countries, except Switzerland (September 1975) and Spain (March 1975). The sample ends in April 2008. The period January 2005 to April 2008 is used for out-of-sample investigation. As already mentioned, a complete description of the data is provided in Appendix 1.

2.3 Analysis of Predictability in Stock Returns

2.3.1 Univariate Extended Model

Table 1 presents in-sample results for the entire sample used, 1.1975–12.2004, for all the countries. First, contemplating the F-test for the various countries, which represents a joint test of significance of the explanatory variables, we notice that the p-value of the test is smaller than the 5% level for UK, Netherlands, Norway and Spain. For Sweden, the test is marginally significant. Thus, for 5 out of 10 countries, we find statistical evidence for in-sample predictability. However, the interpretation of these p-values, even popular in assessing predictability in stock returns, must be done with care, since as will be shown, there is evidence of predictability in stock returns even for the countries which had high F-test p-values. The reason for this is the relatively high value of the coefficient on the dividend-price variable which will yield to economically significant different allocations, especially for the long run.

[Insert Table 1 here]

Next, we wish to discuss the parameter estimates. The first explanatory variable, lagged excess returns, is statistically significant for Switzerland, Sweden, and Norway. These markets could be considered small and relatively illiquid markets. The positive autocorrelation in stock market returns could therefore be attributed to staggered information incorporation. Indeed, as discussed in Campbell, Lo, and MacKinlay (1997), in illiquid markets, new information which arrives on day t may be incorporated into stocks at day $t+1$ only if there was no or little trading in these stocks on day t. The risk-free return is nowhere statistically significant and its sign alternate. Similarly for long-term bond returns. The dividend-price ratio,

constant over time.

a popular explanatory variable in many of the papers dealing with predictability of stock returns and which should be statistically significant based on theoretical grounds (see Cochrane, 2009), is so in four cases. The predictability is particularly high in the UK. Interestingly, the dividend-price ratio has been found to be statistically significant in the US (see Barberis, 2000). The term spread is insignificant for all countries except for Sweden and its sign changes across the various regressions. Last, the inflation rate has a significant effect for three out of the ten regressions. Interestingly, the sign of this variable is always negative, the unique exception being Germany. Fear of higher inflation appears to depress subsequent stock market returns. It is therefore difficult to find explanatory variables which are universally significant across countries, samples, and sampling frequencies.[7,8]

Next, we consider the goodness-of-fit of the regression of the actual out-of-sample return on its predicted value. Contemplating the R^2 and the mean squared errors, we observe that the model has a poor linear fit for all the countries. The country with the highest R^2 is the UK, with a 11.91% value. The lowest R^2 is 1.1%, for Germany.

Overall, contemplating the individual parameter estimates and the goodness-of-fit statistics in-sample is not very encouraging. The question is therefore, how does the model perform in its predicting ability. Using the parameter estimates for the overall sample 1.1975–12.2004, we ask how well could one predict the actual returns. This measure will be obtained by regressing actual returns on their predicted values and by reporting the corresponding t-statistics, named **tstat2** in the tables. As Table 1 reveals, for the extended model, the t-statistics are in all cases significant at the 95% level.[9] This means that even if the model does not have high statistical power, at least, in sample, the model has predictive power. If one considers the RMSE of a prediction, the forecast is poor. For instance, focusing on Switzerland where the RMSE is smallest, we obtain a prediction error of 4.74% on average.

Finally, the table reports the statistics **corrs**, which corresponds to the number of times that the predicted returns were having the same sign as the actual returns, and **corrupd**, which reflects the number of times when the model could predict the correct evolution of returns in comparison with the current level of returns. Since for all these countries these numbers are higher than 50%, we conjecture that there is some evidence of predictability.

[7] We also performed an analysis of the data sampled at weekly and quarterly frequencies. For such data, the results were quite similar. No uniform pattern of predictability emerges.

[8] One could try to find the ideal predictor variables to be used for certain countries at certain frequencies in order to obtain a better fit and could take for instance the approach presented in Bossaerts and Hillion (1999) and find the best prediction models. Nevertheless, this is not the purpose of the present paper and we will keep the focus on the models presented below.

[9] Since we are regressing returns on predicted returns, both variables are stationary. Co-integration is, therefore, not an issue here.

At this stage, one may conclude that in-sample, there appears evidence for predictability of future stock returns. There is, however, no clear pattern of some particular variable being able to explain stock returns in the cross section. For this reason, it appears interesting to consider a reduced model where one retains only a few variables to investigate the forecasting capability of returns.[10]

2.3.2 Univariate Reduced Model

Table 2 presents the results of an in-sample regression involving the time period 1.1975–12.2004, where we focus only on the dividend-price ratio and the inflation rate. This model will be referred to as the *reduced model*.

[Insert Table 2 here]

The results in Table 2 show that for the UK, Netherlands, Norway, and Spain, the coefficients of dividend-price ratio and inflation remain significant, but decrease in level for the UK and Netherlands. The UK has the highest dividend-price ratio coefficient and Spain the smallest, as for the case of the extended model. The inflation coefficient for Germany has presently become negative, whereas the one for Denmark has reduced to zero.[11]

The UK has still the highest R^2 (10.65%), followed by Spain with (3.61%). Contemplating **tstat2**, the return predicted by this reduced model is significant at the 95% confidence level for the UK, Netherlands, Norway, and Spain. The RMSE and the **corrs** and **corrupd** measures remain close to their levels obtained from the model including all the predictors. Thus, one may conjecture that most of the predictive power for the UK, Netherlands, Norway, and Spain comes from the dividend-price ratio and inflation variables. For other countries such as Germany, France, or Switzerland to just name a few, based on statistical grounds, there is much less evidence for predictability. Interestingly, even though the parameters may not be statistically significant, the magnitude of the parameters is not small in all cases. This seems to suggest that an investigation using this little available information may be of relevance to see if there is some economic relevance to our findings.

2.3.3 Parameter Stability of the Extended Model

At this stage, we wish to analyze the stability of the parameter estimates. To do so, we present results for the extended model, including all the explanatory variables,

[10]There is also a technical reason for focusing on a subset of the explanatory variables. Indeed, when we consider later the allocation for a multi-step ahead forecast, it would be difficult, from a computational point of view, to perform the allocation for a large set of predictive variables.

[11]The coefficient is taking a very small value. We did not perform an estimation with positivity constraint where the estimate would have hit a boundary.

by considering two subsamples. The subsamples which we consider consist of data stemming from the beginning of the sample (the period 1.1975–12.1984) and from the end of the sample (the period 1.1995–12.2004). Each subsample corresponds to one third of the full sample. To save space, we did not represent the estimates obtained with the data from the middle of the sample since our findings obtained from the subsamples will suffice to make the point that parameters are varying through time. The results of this estimation are presented in **Table 3**.

[Insert Table 3 here]

The results are shown for the three different countries namely Germany, the UK, and Spain. These countries were selected because of their differences regarding in-sample performance over the entire sample: Germany has no significant explanatory variable and the model has the poorest fit in terms of R^2 and p-value of the F-test; the UK has the highest in-sample performance with the highest coefficient for the dividend-price ratio, Spain is the second in the top of in-sample performance, but unlike the UK, it has a bigger coefficient for inflation and a smaller coefficient for the dividend-price ratio.

If one takes subsamples, one would expect that there is an increase of the standard errors of the estimates and a deterioration of the goodness-of-fit. Parameters should, however, be relatively constant. Instead, by inspecting Table 3, some coefficients change their level and significance. For instance, for Germany, the coefficient of the dividend-price ratio is not significant for the full sample (estimate of 0.131) but becomes so for the first subsample (estimate of 1.682) and again not significant for the second subsample (estimate of 0.311). Besides this, for the UK and Spain the dividend-price ratio remains significant in both subsamples. However, inflation is not significant for the UK over the recent period with a positive parameter estimate. Thus, for further analysis, we include just the dividend-price ratio in the set of the UK predictors for excess stock returns. For Spain, it seems that both predictors are significant over all the samples considered. For Germany, even if the parameter estimates are not significant, they keep the same signs over the subsamples.

The UK has over the first sample a large R^2 of 34%, followed by Spain with 15%. Germany has again the smallest R^2 and no general significance of the model at 95% confidence level. Concerning the second sample, the R^2 drops for all regressions. The best fit is for Spain with 9.15%. The other performance measures (the **tstat2** statistics and the measures of predictability of the sign of returns and relative return variations) remain similar to the overall sample. Those results are encouraging from the point of view of the economic relevance of the return predictability. They reveal, however, the importance to take into account parameter uncertainty.

2.3.4 Out-of-Sample Performance

Until now, only the in-sample predictability has been analyzed. The empirical literature, see for instance Bossaerts and Hillion (1999), has shown that out-of-sample performance of most of the models considered is often worse than in-sample performance. Before turning to the estimates, it is necessary to carefully describe how we obtained the out-of-sample regression estimates. First of all, the out-of-sample period starts in January 2005 and ends in April 2008. We consider two methods to assess the out-of-sample performance. In the first one, we consider parameter estimates obtained for data covering 1975–2004 then held constant. In the second one, we consider an increasing sample which starts in January 1975 and incorporates increasingly more data ending with April 2008.

In **Table 4**, we report the out-of-sample performance measures for these two methods. The first row of each Panel represents the R^2 of the regression of actual returns on the predicted returns. We notice small numbers for the goodness-of-fit statistics. The R^2 ranges between 2.36% for the UK and as little as 0.15% for Germany. Similarly, the p-value of the t-statistics (**tstat2**) for the regression of the actual return on a constant and the predicted return is deteriorated. Actually, both for the constant parameter regression as well as for the one with updated parameters, the out-of-sample performance seems to be poor.

[Insert Table 4 here]

Not everything is gloomy, however. Inspection of the **RMSE** reveals a decrease for all the countries compared to in-sample **RMSE**. This decrease in the **RMSE** for the out-of-sample period is counterintuitive but may be explained by the fact that over the in-sample period the fit turns out to be particularly bad for certain time period. Moreover, the indicator **corrs** is again higher than 50%, which means that even out-of-sample, the model predicts more than 50% of time the correct sign of returns. Since this result holds for all the countries, it appears implausible that this result is due to pure chance. In contrast, the indicator **corrupd** has decreased under 50% for France, Switzerland, and Norway, which indicates that in making optimal portfolio choice, the out-of-sample performance could be poor, the model being unable to predict the direction of return for these countries. From the point of view of **AIC** and **BIC** criteria, the UK has again the best out-of-sample performance, followed by Spain.

For the forecasts based on rolling regression, results seem to be close or better than the ones for the model which uses constant parameters. For instance, the UK remains in the top, the model being able to determine 61% of times the direction of returns and 56% of times its sign. **AIC** and **BIC** have decreased for all the countries. This same or even better out-of-sample performance of the reduced model gives another indication that learning of the parameters, over time, could be an important issue.

We conclude this section by noticing that we continue to expect economic relevance for the out-of-sample predictability and this in particular for a model with dynamic parameters.

2.4 Multivariate Regression Model Results

As already mentioned, a multivariate regression model is needed to study multi-horizon portfolio choice. In such a model, the covariance matrix of error terms plays an important role. To be consistent with our ensuing analysis of the buy-and-hold and optimal rebalancing strategies, we focus on three different countries only (Germany, the UK, and Spain). However, for comparison, we will also discuss the differences with the other countries and between the expanded model and the reduced model, which uses only the dividend-price ratio for the UK, and both the dividend-price ratio and the inflation rate for Germany and Spain.

When the extended model is considered, shocks in returns are in general negatively correlated with shocks in dividend-price ratio and shocks in inflation. However, shocks in the real risk-free rate, in excess return on bonds and in spread do not have such a general result over countries. For some countries, they present negative correlation with return shocks and for others they present positive correlation. This result was partly expected from the univariate regression coefficients, which were unstable across countries and samples. However, their absolute correlation with shocks in stock returns is relatively small compared to the correlation of shocks in the dividend-price ratio, which takes values from 90% (for the UK) to 74% for Italy and Denmark. This result remains unchanged if the reduced model is being used instead of the extended model. The same correlation in the case of inflation takes much lower values, which range from 1% to 11%.

Table 5 presents in the upper part the results of the multivariate regressions for the three countries, while the lower part presents the mean and standard deviation (in parenthesis) of each parameter's posterior distribution.[12] Under the heading Σ we present along the diagonal the variances of the errors and above the diagonal correlations.

We notice that for the chosen countries all parameter estimates are significant except for Germany. The highest predictability for returns, in terms of in-sample indicators like the R^2 and F-test p-value is again obtained for the UK, followed by Spain. For Germany, as discussed in the univariate regressions, there is no sign of in-sample predictability if one relies on standard tests.

As expected, the mean of the posterior distribution of each parameter is equal to the estimate reported in Panel A and the highest uncertainty in parameters is

[12]In order to obtain the posterior distribution, 1'000'000 simulations were used.

found in the return equation. The predictive power of the dividend-price ratio in the return equation is important for the UK, this predictor having a coefficient with mean 1.34 and standard deviation 0.24. For the other two countries, the coefficient is lower, and its standard deviation is higher, especially for Germany. Inflation has predictive power for returns in Spain but none for Germany even though for the latter, the sign is negative as expected.

The dynamics of the dividend-price ratio and the inflation rate are close to an univariate AR(1) process. Their lagged values have parameter estimates higher than 0.9 for all countries.

Regarding the covariance matrix Σ of the innovations in the VAR system, we notice that unexpected excess stock returns (innovations or shocks in excess stock return) are highly negatively correlated with shocks to the dividend-price ratio, this result being consistent with results presented in the previous literature. Interestingly, the highest correlation is in the case of the UK, where it takes a value of -0.9. Unexpected excess stock returns also present negative correlation with shocks in inflation but the coefficients are rather small with -4% for Germany and nearly -8% for Spain.

[Insert Table 5 here]

We turn now to the posterior distribution of the parameters in the i.i.d. returns case, which corresponds to our benchmark (Panel C of Table 5). We notice that the monthly expected returns are highest for the UK, followed by Germany and Spain. The associated standard errors are taking similar values for the three countries with differences in the fourth decimal. When we turn to the variability of returns measured by their annualized monthly standard deviation, $\sqrt{12}\sigma$, we obtain 20.5%, 19.6% and 21.9% for Germany, the UK, and Spain. These are usual order of magnitudes.

If one takes the last available monthly risk-free rate in the database, 0.33%, 0.44%, 0.33% for Germany, the UK, and Spain respectively, the posterior mean and standard deviation imply a Sharpe Ratio of 40.36% for the UK, 20% for Germany, and 8.62% for Spain. Thus, one would expect that the allocation to stocks is the highest for the UK and the lowest for Spain.

3 Portfolio Allocation

3.1 Generalities

The framework upon which we build follows closely Barberis (2000). The objective is to investigate the influence of predictability and parameter uncertainty on multiperiod portfolio allocations for different countries. We consider an investor who has gathered a sample of T observations and who wishes to invest optimally for the next $\hat{T}$ months. To keep the problem as simple as possible, we consider an elementary investment between the risk-free asset and a stock index. The investor is not allowed to leverage stock nor is she allowed to sell short. For simplicity, we take the continuously compounded risk-free return as a constant, r_f. We consider that the initial wealth at time T is $W_T = 1$ and denote by ω the allocation to the stock index. Then, the end-of-horizon wealth is:

$$W_{T+\hat{T}} = (1-\omega)\exp(r_f\hat{T}) + \omega\exp(r_f\hat{T} + r_{T+1} + ... + r_{T+\hat{T}}).$$

The investor maximizes her utility, $\nu(W)$, which belongs to the CRRA utility function class:

$$\nu(W) = \frac{W^{1-A}}{1-A},$$

where A is the coefficient of risk aversion. Now, if we write the cumulative stock return over $\hat{T}$ periods as $R_{T+\hat{T}} = r_{T+1} + r_{T+2} + \cdots + r_{T+\hat{T}}$, the investor has to solve:

$$\max_{\omega} E_T\left[\frac{\{(1-\omega)\exp(r_f\hat{T}) + \omega\exp(r_f\hat{T} + R_{T+\hat{T}})\}^{1-A}}{1-A}\right]. \tag{7}$$

The investor calculates the expectation conditional on her information set at time T. There are two approaches to solve this maximization problem depending on whether parameter uncertainty is taken into account or not.

I. Ignoring parameter uncertainty. In this case, the distribution for future stock returns is conditioned on a set of parameter estimates ($\hat{\theta}$) and is written as $p(R_{T+\hat{T}}|z,\hat{\theta})$, where $z = (z_1', \cdots, z_T')$ and the data are observed by the investor until the start of her investment horizon. Thus, the investor solves:

$$\max_{\omega} \int \nu(W_{T+\hat{T}})\ p(R_{T+\hat{T}}|z,\hat{\theta})\ dR_{T+\hat{T}}. \tag{8}$$

II. Integrating out parameter uncertainty. In this Bayesian approach, parameter uncertainty is taken into account and one needs to use the predictive distribution for long-horizon returns. In this case, the investor solves:

$$\max_{\omega} \int \nu(W_{T+\hat{T}})\ p(R_{T+\hat{T}}|z)\ dR_{T+\hat{T}}, \tag{9}$$

where $p(R_{T+\hat{T}}|z) = \int p(R_{T+\hat{T}}|z,\theta)\ p(\theta|z)\ d\theta$. The predictive distribution is conditioned only on the observed data and not on the parameter value θ. In order to sample from the predictive distribution, it is necessary to first sample from the posterior distribution $p(\theta|z)$ and then from the conditional distribution $p(R_{T+\hat{T}}|z,\theta)$. Thus, the optimization problem (9) can also be written as:

$$\max_{\omega} \int\int \nu(W_{T+\hat{T}})\ p(R_{T+\hat{T}}|z,\theta)\ p(\theta|z)\ dR_{T+\hat{T}}\ d\theta.$$

These two approaches will help understanding how parameter uncertainty affects the portfolio choice, by comparing the solutions to problems I and II. The maximization of problems I and II is solved as in Barberis (2000), who evaluates the integrals over a fine grid and then takes the largest expected utility. The grid is given by $\omega = 0,\ 0.01,\ 0.02,\ \cdots,\ 0.99,\ 1$, that is 100 finely spaced values. The integration is solved by different methods, depending on the investment strategy adopted by the investor. In some cases, it is possible to perform a numerical quadrature, while in other cases it is necessary to rely on Monte-Carlo integration. In order to investigate how predictability in asset returns and parameter uncertainty affect the portfolio choice, two main strategies will be discussed: the buy-and-hold strategy and the optimal rebalancing strategy. Within each of these strategies, we will considered the case of i.i.d. returns and of predictable returns. We now turn to the discussion of these alternative strategies.

3.2 The Buy-and-Hold Strategy

This strategy assumes that the investor chooses an allocation at the beginning of the first year, using the data known until that moment, and keeps her portfolio weights constant until the end of the investment horizon. We now consider the special case of i.i.d. stock returns and the case when returns are predictable by a set of variables.

3.2.1 Case 1: i.i.d. normal returns ignoring parameter uncertainty

In this situation, returns are assumed to be distributed as a simple normal $r \sim N(\mu, \sigma^2)$. An investor who does not take into account parameter uncertainty would estimate μ and σ^2 by $\bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t$ and $s^2 = \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \bar{r})^2$. The cumulative stock return, $R_{T+\hat{T}}$, is then also normally distributed with mean $\hat{T}\mu$ and variance $\hat{T}\sigma^2$. In this case, a quadrature method can be used to solve this expectation. Numerical quadrature is described in Appendix 3.

3.2.2 Case 2: i.i.d. normal returns with parameter uncertainty

An investor who wishes to take parameter uncertainty into account needs to know the posterior predictive distribution of the cumulative return over the next $\hat{T}$ time periods, $R_{T+\hat{T}}$. Under the assumption that the $\hat{T}$ future returns are all normal, $R_{T+\hat{T}}$ will also be normal:

$$R_{T+\hat{T}}|\sigma, \mu \sim N\left(\hat{T}\mu, \sigma^2\hat{T}\right).$$

Assuming that the prior is noninformative as discussed in Section 2.1.1, the marginal posterior pdf for σ is an inverted gamma and the conditional posterior for μ given

σ and the sample information is a univariate normal:

$$\sigma|r \sim IG\left(\frac{T-1}{2}, 2/\sum_{t=1}^{T}(r_t-\bar{r})^2\right),$$

where $\bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t$, and

$$\mu|\sigma, r \sim N\left(\bar{r}, \frac{\sigma^2}{T}\right),$$

In this case, the predictive distribution of $R_{T+\hat{T}}$ can be determined analytically by integration

$$\begin{aligned} p(R_{T+\hat{T}}|r) &= \int_0^\infty \int_{-\infty}^\infty p(R_{T+\hat{T}}|r,\mu,\sigma^2)p(\sigma^2|r)p(\mu|\sigma^2,r)d\mu d\sigma \\ &= \frac{\Gamma(\frac{T}{2})\ \sqrt{T}(s^2(T-1))^{(T-1)/2}}{\Gamma(\frac{T-1}{2})\ \sqrt{\pi}\sqrt{\hat{T}(T+\hat{T})}}\left(\frac{T(R_{T+\hat{T}}-\bar{r}\hat{T})^2}{\hat{T}(T+\hat{T})}+s^2(T-1)\right)^{-T/2}, \end{aligned}$$

which shows that the density function is a univariate Student-t,

$$R_{T+\hat{T}}|r \sim t\left(\bar{r}\hat{T}, \sum_{i=1}^{T}(r_t-\bar{r})^2, \frac{T}{\hat{T}(T+\hat{T})}, T-1\right), \tag{10}$$

with degrees of freedom $v = T-1$ and scale factor: $s^2(T-1) = \sum_{t=1}^{T}(r_t-\bar{r})^2$.[13] Since the Student-t distribution tends to have fatter tails than the normal distribution, one would expect that an investor will be more precautious under parameter uncertainty.

Even though an analytic expression is available for the predictive density, namely equation (10), we will rely on Monte-Carlo integration to evaluate the expected utility (7). The procedure to do so is to obtain N draws from the $p(\sigma^2|r)$ density, here an inverse gamma distribution. Next, for each σ^2 drawn, one obtains one draw from the normal $N(\bar{r}, \sigma^2/T)$, a normal distribution. The last step, yielding the distribution of $R_{T+\hat{T}}$, is easy. Since $R_{T+\hat{T}} \sim N\left(\hat{T}\mu, \sigma^2\hat{T}\right)$, one only needs to simulate from the normal density, using the μ and σ^2 that were just obtained. Let us denote by $R^i_{T+\hat{T}}$, $i = 1, \cdots, N$ the resulting sample. The integral corresponding to the expected utility is approximated by

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\{(1-\omega)\exp(r_f\hat{T})+\omega\exp(r_f\hat{T}+R^i_{T+\hat{T}})\}^{1-A}}{1-A}. \tag{11}$$

The convergence towards the expected value is guaranteed by the law of large numbers.

[13] Appendix 5 summarizes the expressions for the densities involved in this Bayesian context.

3.2.3 Case 3: Predictable returns ignoring parameter uncertainty

In the case of predictable returns, it is necessary to condition the future realization of returns on the current level of the observable exogenous variables as well as on the parameters involved in the estimation of the dynamic of the system. To do so, it is convenient to follow Barberis (2000) and to write $z_t = a + B_0 x_{t-1} + \varepsilon_t$ as a VAR(1) system

$$z_t = a + Bz_{t-1} + \varepsilon_t$$

where $B = ((0 \cdots 0)' \; B_0)$. As is shown in Appendix 4, conditional on a, B, and Σ, the sum $Z_{T+\hat{T}} = z_{T+1} + z_{T+2} + \cdots + z_{T+\hat{T}}$ is normally distributed with mean μ_{sum} and variance Σ_{sum}, where:

$$\mu_{sum} = \hat{T}a + (\hat{T}-1)Ba + (\hat{T}-2)B^2a + \cdots + B^{\hat{T}-1}a + (B + B^2 + \cdots + B^{\hat{T}})z_T, \quad (12)$$

and

$$\begin{aligned}\Sigma_{sum} = {} & \Sigma + (I+B)\Sigma(I+B)' + (I+B+B^2)\Sigma(I+B+B^2)' + \cdots \\ & + (I+B+B^2+\cdots+B^{\hat{T}-1})\Sigma(I+B+B^2+\cdots+B^{\hat{T}-1})'. \quad (13)\end{aligned}$$

Obviously, $R_{T+\hat{T}}$ is given by the first element of $Z_{T+\hat{T}}$. The investor who does not take parameter uncertainty into account may use the usual non-Bayesian parameter estimates and then approximate expected utility (8) by using numerical quadrature as discussed in Case 1 above or in Appendix 3.

3.2.4 Case 4: Predictable returns with parameter uncertainty

The investor who wishes to take parameter uncertainty into account needs to consider the predictive density of $R_{T+\hat{T}}$. Considering the algebra of the previous case, we notice that the problem boils down to obtaining the predictive density of $Z_{T+\hat{T}}$. As in Case 2, the predictive density will be obtained by simulation. The procedure is to obtain N draws (where N is large) from the marginal density $p(\Sigma|z)$, here an inverse Wishart, $IW\ (T-n-2,\ S)$, as discussed in Section 2.1.1. Next, for each Σ, one obtains one draw from the density $p(C|\Sigma, z)$, here this is the matricvariate density, $MN\ (\hat{C},\ \Sigma \otimes (X'X)^{-1})$. Last, one obtains one draw from the posterior distribution $p(Z_{T+\hat{T}}|a, C, \Sigma, z)$, a normal distribution. This procedure will generate a set of N draws which can be used to approximate the expected utility (11) by an average.

3.3 Dynamic Investment Strategy

So far, the investor were assumed to compute at time T the portfolio allocation which then remains constant till time $\hat{T}$. However, when returns are predictable, one would expect that, as time passes by, a new state occurs (given by the realization of a new exogenous variable). The investor should then take into account the possibility of rebalancing the portfolio. This is the strategy that we wish to discuss in this section. This strategy is naturally associated with multi-period portfolio choice.

The model used for this strategy is the one presented in equation (5), with $x_t = dp_t$ or $x_t = (dp_t, \pi_t)'$. This reduced number of predictive variables is used since extra variables proved to be insignificant. There is also the issue of reducing the dimensionality of the state space due to programming limitations, which render the dynamic programming problem difficult to solve by using available technology.

To solve this problem, we follow again Barberis (2000) and discretize the state space and use backward induction. The investment horizon of $\hat{T}$ months is divided into K intervals of equal length, such that the investor adjusts her portfolio each year, K times over the horizon, at the dates $(t_0, t_1, ..., t_{K-1})$. Since we have monthly data, we have the relation $K = \hat{T}/12$. The control variables in the optimization are $(\omega_{t_0}, \omega_{t_1}, \cdots, \omega_{t_{K-1}})$, the allocations to the stock index at each rebalancing time. The investor has to solve the problem:

$$\max_{\omega_{t_0}} E_{t_0} \left(\frac{W_{t_K}^{1-A}}{1-A} \right).$$

Denote $R_{t_{k+1}} = r_{t_k+1} + r_{t_k+2} + \cdots + r_{t_{k+1}}$, the cumulative return between rebalancing points, from t_k to t_{k+1} (this corresponds to the cumulative return over each year as the investor rebalances yearly). Then, one obtains that wealth evolves from time t_k to time t_{k+1} as

$$W_{t_{k+1}} = W_{t_k} \left[(1-\omega_{t_k}) \exp\left(r_f \frac{\hat{T}}{K} \right) + \omega_{t_k} \exp\left(r_f \frac{\hat{T}}{K} + R_{t_{k+1}} \right) \right],$$

for $k = 0, \cdots, K-1$. Considering that the investor at each time t_k maximizes over all remaining decisions from time t_k on, we define the value function:

$$J(W_{t_k}, x_{t_k}, t_k) = \max_{\omega_{t_k}} E_{t_k} \left(\frac{W_{t_K}^{1-A}}{1-A} \right).$$

The value function J does not depend on r_{t_k}, the stock return over month t_k, since in this model the current values of the predictive variable(s) characterize the investment opportunity set. It can be shown, following the discussion in Barberis (2000), that the Bellman equation can be written as:

$$\begin{aligned} Q(x_{t_k}, t_k) = \max_{\omega_{t_k}} E_{t_k} & \left[(1-\omega_{t_k}) \exp\left(r_f \frac{\hat{T}}{K} \right) + \omega_{t_k} \exp\left(r_f \frac{\hat{T}}{K} + R_{t_{k+1}} \right) \right]^{1-A} \\ & \times Q(x_{t_{k+1}}, t_{k+1}), \end{aligned} \tag{14}$$

where $Q(x_{t_k}, t_k)$ can be thought of as the maximized expected utility at time t_k.

It is important to note here that, if the cumulative excess return $R_{t_{k+1}}$ were independent from the next period state variable $x_{t_{k+1}}$, the expectation would reduce to a product of two expectations. Since $Q(x_{t_{k+1}}, t_{k+1})$ is not a function of portfolio weights ω_{t_k}, this multi-period maximization problem would simply reduce to a single-period maximization problem. Therefore, the joint distribution $p(R_{t_{k+1}}, x_{t_{k+1}}|\theta, x_{t_k})$ is needed to solve for ω_{t_k}, as it is explained in the sequel.

The way to solve equation (14) is the following: the investor starts at date t_{K-1} and performs the usual maximization problem explained in the previous section, but for each state of the world considered at time t_{K-1}. By doing this, she obtains a set of optimal weights at t_{K-1} and a set of maximized expected utilities at year $t_K = T+\hat{T}$, conditional on the states defined for the predictive variables at year t_{K-1}. Thus, we obtain in this way $Q(x_{t_K}, t_K)$. Going backward, each year t_k for $k = 0, \cdots, K-2$, the optimization problem changes, the investor taking into account the expected utility in the next period, as shown in the Bellman equation above.

3.3.1 One state variable

We consider first the case where the dividend-price ratio is used alone as predictive variable. The dividend-price ratio range is discretized into M equally spaced grid points, denoted by $\{x^j_{t_k}\}_{j=1,\cdots,M}$.[14] Then, at time t_k, one needs to compute $Q(x_{t_{k+1}}, t_{k+1})$ for all possible values $x_{t_{k+1}} = x^j_{t_{k+1}}$. Then, using the Bellman equation (14), at time t_k we obtain $Q(x^j_{t_k}, t_k)$. More precisely, for each $x^j_{t_k}$, we draw a large sample $\{R^{(i)}_{t_{k+1}}, x^{(i)}_{t_{k+1}}\}^N_{i=1}$ from the joint distribution $p(R_{t_{k+1}}, x_{t_{k+1}}|\theta, x^j_{t_k})$, which is multivariate normal. The methodology used to obtain this distribution is detailed in Appendix 4.

Then, simulations are used to determine $Q(x^j_{t_k}, t_k)$ as:[15]

$$\begin{aligned} Q(x^j_{t_k}, t_k) = \max_{\omega_{t_k}} \frac{1}{N} \sum_{i=1}^{N} & \left[(1-\omega_{t_k}) \exp\left(r_f \frac{\hat{T}}{K} \right) + \omega_{t_k} \exp\left(r_f \frac{\hat{T}}{K} + R^{(i)}_{t_{k+1}} \right) \right]^{1-A} \\ & \times Q(x^{(i)}_{t_{k+1}}, t_{k+1}). \end{aligned}$$

Since $Q(x_{t_{k+1}}, t_{k+1})$ is known only for all $x_{t_{k+1}} = \{x^j_{t_{k+1}}\}_{j=1,\cdots,M}$, the approximation of $Q(x^{(i)}_{t_{k+1}}, t_{k+1})$ by $Q(x^j_{t_{k+1}}, t_{k+1})$ is needed. It is done by taking $x^j_{t_{k+1}}$ as the closest element of the discretized space to $x^{(i)}_{t_{k+1}}$, which was simulated. Backward induction through all K rebalancing points finally gives $Q(x^j_{t_0}, t_0)$ for all $j = 1, \cdots, M$, thus the optimal allocation at t_0, ω_{t_0}.

[14] For the case where we will have only one state variable as here, we will chose $M = 25$. In the next section, where we consider two state variables, we use $M = 11$. The discretization is taken over the interval starting from the historical mean plus and minus three standard deviations. The choice of M is dictated by the speed of execution and accuracy. We performed tests for stability and noticed that small changes of M did not substantially affect the allocation.

[15] We used 10'000 simulations for this strategy.

3.3.2 Two state variables

We consider now the case of two state variables, denoted here by $x1$ and $x2$. Expected utility needs to be computed recursively for each pair $(x1^{j1}_{t_k}, x2^{j2}_{t_k})$ where $j1 = 1, \cdots, M$ and $j2 = 1, \cdots, M$. Again, for each pair, we need to draw a large sample $(R^{j1,j2}_{t_{k+1}}, x1^{j1}_{t_{k+1}}, x2^{j2}_{t_{k+1}})^N_{j=1}$ from the joint conditional distribution $p(R_{t_{k+1}}, x1_{t_{k+1}}, x2_{t_{k+1}} | \theta, x1^{j1}_{t_k}, x2^{j2}_{t_k})$, which is a normal distribution. Note that in this case, $R^{j1,j2}_{t_{k+1}}$ depends on both states $j1$ and $j2$ of $x1$ and $x2$. The methodology of backward induction is the same as in the case of just one state variable, the higher dimensionality becomes more time consuming of course.

It is worth emphasizing that, if the joint distribution of $R_{t_{k+1}}$ and the next period state variables are only weakly correlated, the impact of the value of next period utility is irrelevant for the backward computation. Thus, when going backward in time the optimal portfolio weights remain identical to those computed in an earlier stage. This means that, in this case, even when we consider different investment horizons, the allocation lines will be constant over time and less sensitive to the investment horizon. In this case, the so-called horizon effects are relatively small.

3.3.3 Parameter uncertainty

Incorporating parameter uncertainty in a multi-period setting is more complicated due to the following issues. First, the expectation in equation (14) should be taken by using the unconditional distribution of cumulative excess returns. However, parameter uncertainty can change over time, as new information is available to the investor. The level of uncertainty then becomes a new state variable which increases dramatically the complexity of the model. This is the effect of "learning" about the true parameter values on portfolio holdings, which may change itself the portfolio choice. This problem was studied by Williams (1977) and Gennotte (1986).

In order to incorporate at least partially parameter uncertainty, Barberis (2000) suggested the following simplification. The investor solves the dynamic problem assuming that beliefs about the parameters will not change over time, and are set at the beginning on the investment horizon according to the information available at that date. Thus, these beliefs summarized by the posterior distribution will give the posterior predictive distribution of cumulative excess returns, and the expectation will be solved as explained in Section 3.3 by using the predictive density.

To conclude this section, for an investor who cares about parameter uncertainty, we expect the allocation to stocks to be smaller, since the investor becomes uncertain about the predictive power of the state variables, and thus about the fact that the investment opportunity set is really changing over time as the state variables change.

3.4 Results for the Buy-and-Hold Strategy

Having already estimated the VAR(1) model required for the computation of the multi-horizon portfolio allocation, as a preliminary step, we investigate the shape of standard deviation for selected countries as a function of the time horizon of the allocation. To gain some intuition of what to expect, let us compare the simple model of i.i.d returns where $r_{t+1} = \varepsilon_{t+1}$ and $Var(\varepsilon_{t+1}) = \sigma_\varepsilon^2$ with the situation where returns are predictable as in

$$\begin{aligned} r_{t+1} &= x_{t+1} + \varepsilon_{t+1}, \qquad & Var(\varepsilon_{t+1}) &= \sigma_\varepsilon^2, \\ x_{t+1} &= x_t + \eta_{t+1}, \qquad & Var(\eta_{t+1}) &= \sigma_\eta^2. \end{aligned}$$

It follows that $r_{t+1} = x_t + \eta_{t+1} + \varepsilon_{t+1}$ and $Var(r_{t+1}) = \sigma_\varepsilon^2 + \sigma_\eta^2 + \sigma_\varepsilon\sigma_\eta\rho$, with ρ the correlation between the two sources of uncertainty. These computations reveal that if returns are predictable and the innovations to the source of predictability are positively correlated with innovations to returns, the overall uncertainty may be increased. In the case where one deals with even longer time periods, this feature conjugates and the longer the time period of an allocation, the greater the uncertainty. Inversely, if the correlation is negative, and this will be the general case for the dividend-price ratio, one may expect a reduction in uncertainty if returns are predictable. In the case of multiple predictors, there will be different correlations involved and the overall impact will depend on the various correlations among the predictors.

This type of result has already been reported in the literature. For instance Campbell and Viceira (2002), and Bec and Gollier (2009) have shown, for US and French quarterly data, that mean reversion in stock returns reduces significantly their standard deviation as the horizon increases. From the point of view of the investment strategy, this would imply that stock returns are safer at longer horizons, meaning that a longer-horizon investor should allocate more to stocks.

Figure 1 displays for the three selected countries the annualized standard deviation if one incorporates parameter uncertainty or not in the model with the investment horizon increasing from 1 to 10 years. Those curves with continuous line correspond to the case where parameter uncertainty is not taken account for and the dashed line incorporates parameter uncertainty.[16] We observe that for countries like Germany and the UK, the standard deviation decreases monotonically with the time horizon for both cases. Interestingly, for a country like Germany, the preliminary econometric analysis revealed that there was little evidence for predictability

[16]For the case where there is no uncertainty, we simply take equation (13), where Σ and B result from the OLS estimations. The case which incorporates parameter uncertainty is more complicated. Here, one uses the Bayesian approach. First Σ gets simulated according to the inverse Wishart and C is drawn according to the matricvariate normal. Since C contains the matrix B, all the required elements for the computation of Σ_{Sum}, as an average across the simulations, are consistently simulated.

by using the R^2 and the F-test. However, as one contemplates the figures, there is definitely a temporal effect whereby the standard deviation is 2% lower for an investor with a 10-year horizon than for a 1-year horizon only.

[**Insert Figure 1 here**]

Campbell and Viceira (2002) and Bec and Gollier (2009) have found a similar negative trend of the conditional standard deviation as a function of the investment horizon. Even though we do not report the figure, we would like to mention that the decrease of standard deviation that we obtained for France is comparable to the one obtained by Bec and Gollier (2009). The ratio obtained by dividing the 10-year standard deviation by the 1-year conditional standard deviation is around 0.94 in their case and 0.84 in our case for French monthly data. This relatively small difference may be attributed to the fact that they use data sampled at quarterly frequency whereas we use monthly data. They also use different additional predictor variables.

By contemplating the plots presented in Figure 1, we observe that UK conditional standard deviation has the highest decrease over the investment horizon presented here. The ratio of 10-year to 1-year standard deviation is 0.63. The same ratio is 0.90 for Germany and 0.98 for Spain. If we interpret this ratio as a measure of the speed of mean reversion in excess stock returns, we expect that the allocation to stocks to increase with investment horizon only for the UK, the country with the highest speed of mean reversion.

The pattern of the temporal evolution of the standard deviation is more intricate for Spain. After briefly decreasing, the curve becomes hump shaped with a concave shape. This pattern may be explained by the fact that we have used here two explanatory variables. Even though both have errors that are negatively correlated with the errors of the return regression, the correlation is positive among themselves. The computation leading to Σ_{Sum} thus becomes highly non-linear and complex patterns may occur as is the case at hand.

Taking into account of parameter uncertainty is expected to lead to increased uncertainty. Remember that we found, for instance, that the distribution of returns in the i.i.d. case was no longer normal but t-distributed after one takes parameter uncertainty into account. As inspection of the Figure 1 shows (dashed curves), this is indeed the case. For the case of the UK, the modification of the standard deviation is relatively small. For the other countries, the impact of parameter uncertainty is more pronounced. For Germany and Spain, the long-run risk differential is about 1%.

Having established that, due to predictability, the standard deviations will differ according to the time horizon, we now investigate the actual relevance of return predictability on portfolio allocations.

3.4.1 Optimal Portfolio Choice

In **Figure 2**, we present the results of the portfolio allocation for the four cases discussed in Section 3.2. We choose as levels of risk aversion 3 and 5 for Germany and Spain as well as 5 and 10 for the UK. The overall choice of such parameters is given by the literature. Most contributions would deem a level of risk aversion between 5 and 15 as standard values. The choice of slightly different parameters for the UK than for Germany and Spain is only for presentation purposes. Had one chosen a level of risk aversion of 3 for the UK, then some of the allocations would have saturated the 100% in stock allocation.

Figure 2 presents in its left half those allocations that are obtained for the lower level of risk aversion. In each plot, the dashed line represents the i.i.d. allocations that do not take parameter uncertainty into account (Case 1). The dotted curve takes in addition parameter uncertainty into account (Case 2). The continuous curve is obtained for the model where predictability is taken care of, but where the investor ignores parameter uncertainty (Case 3). Last, Case 4 also allows return predictability but takes parameter uncertainty into account (the continuous line with crosses). All the allocations that presume predictability of stock returns require in their calibration an assumption of the level of the explanatory variables. We set, in this section, the explanatory variables to their sample mean, leaving the discussion of changes in these values to the next section.

[**Insert Figure 2 here**]

Inspection of Figure 2 reveals that, as expected, the country with the highest Sharpe ratio has also the largest allocation to stocks, that is the UK where the i.i.d. allocation of Case 1 implies about 55% in stocks for $A = 5$ and about 25% if $A = 10$. These allocations are followed by Germany where the investor puts about 30% in the index (for $A = 5$) and Spain where the investor only risks 20% (still for $A = 5$). Comparison between the left side of Figure 2 with the right part suggests that less wealth will be invested into stocks when the risk aversion increases. Even if the optimal weights vary with the level of risk aversion (comparing the left and right hand sides of the figure), the relative patterns between the strategies remain similar. For this reason, we focus from now on the discussion of the left hand side plots of Figure 2.

The next consideration is the introduction of parameter uncertainty into the case where returns are i.i.d. As the comparison between the dashed and dotted line reveals, parameter uncertainty may play strong effects. For instance, for the UK, the long term investor would decrease her allocation to stocks from 55% to 40%. For Germany the impact would also be important with an allocation decreasing from 50% down to 40%. For Spain the effect would be somewhat less but still significant.

A dramatic change in the allocation occurs however for some of the countries if one takes into account predictability. For the UK, the investor would put all of his wealth into the stock index, already with a 5-year horizon. For longer horizons, the investor would be willing to purchase leveraged stock. Even for the short horizon of one year, the effect is dramatic. Indeed, instead of investing 55% into stocks, the investors who uses predictability in stocks would invest up to 75%. For Spain the impact is also important, less so for the German investor.

The last case to discuss is the one where predictability and parameter uncertainty are both accounted for. As already seen, the introduction of predictability decreases the allocation to stocks. If one considers the case of $A = 5$, now the UK investor no longer invests 100% into stocks but only about 85% for a horizon of 10 years. The impact of parameter uncertainty undoes, therefore, some of the dramatic effect of predictability. However, for the UK and Spain, for the 10-year allocation, the investor would put significantly more wealth into stocks than for the i.i.d. case. Interestingly, for Germany, the picture is somewhat different. There, predictability is not so important, hence, parameter uncertainty leads to allocations similar to the case without predictability.

3.4.2 Sensitivity of Portfolio Weights to the Explanatory Variables

In the case of predictable returns, the optimal allocation depends on the initial levels of the state variables, here the dividend-price ratio and the inflation rate. As we discussed analytically, at the beginning of Section 3.4, the correlation and magnitude of the shocks to returns and the explanatory variables matter for the predictability of returns. Thus, the effect of the dividend-price ratio is expected to be higher given the level of correlation of shocks between the dividend-price ratio with shocks in returns. Moreover, as equation (12) shows, the level of the explanatory variables will matter through a direct effect on the expected return of the stock.

We will discuss now, for the various countries and the two explanatory variables, how the allocations change if one changes the level of the explanatory variables, which have been set so far at their average level. In all the cases, we will consider plots where one takes the average, the average plus/minus 1 standard deviation, and the average plus/minus 2 standard deviations. This choice of initial levels of the explanatory variables appears to describe in a satisfactory manner the possible range of those parameters.

Figure 3 displays the sensitivity of optimal portfolio weight to the initial level of dividend-price ratio for the UK. The left part of the figure corresponds to the plots of allocation where no parameter uncertainty is taken into account. In the right part the allocations assume that the investor accounted for parameter risk. The continuous line represents the allocations obtained for the average level of a given explanatory variable. The dashed line represent the allocation for minus one standard deviation and the dashed line with pluses, the allocation for minus two standard deviations. The dotted lines with crosses represent the allocation for plus one standard deviation and the dotted line the one if one adds two standard deviations to the average level. This set of conventions will also hold for the coming figures.

[**Insert Figure 3 here**]

Figure 3 shows that there is significant sensitivity of the optimal allocation to the initial values of the dividend-price ratio. Given the positive regression coefficient of 1.34 in the regression of returns on lagged dividend-price ratio, one expects that, when the dividend-price ratio is low relative to its historical mean, then future expected returns are also low. In such a case, one would choose a more conservative portfolio allocating less wealth to the stock market. Since the regression is linear and the same for equation (12), we expect the opposite to hold for a high dividend-price ratio, namely a high allocation to stocks. This is exactly the finding in Figure 3.

The next issue to be discussed is the convexity of the curves. First, it should be noticed that, for low risk aversion ($A = 5$), the investor chooses extreme portfolio weights, placing either 100% in stocks or 100% in the risk-free asset, for large deviations of the dividend-price ratio from its mean. Only for the case of an investor with a high risk aversion ($A = 20$), one does not systematically saturate the upper bound of 100% in stocks.

Next, we notice that the convexity of the curves changes with the level of the explanatory variables. If one chooses a very high level for the explanatory variable, it will have an important impact for an allocation where the horizon is just one year. An investor, who focuses on the long term, such a level will not be sustainable. The investor will know that the dividend-price ratio will converge towards its long-run mean. In that case, the lasting impact will be weaker than for a short-term allocation. It matters here, obviously, that the portfolio is based on a buy-and-hold strategy. Thus, the allocation should be made under the assumption of an average level of the explanatory variable. For instance, if the explanatory variable takes a very low level, the investor will be very conservative if her horizon is short, but the allocation will be based on some average level of the explanatory variable for a long-term horizon.

For Germany and Spain, results for portfolio weights depending on initial levels of dividend-price ratio are presented in **Figures 4 and 5**. We notice the same sensitivity of portfolio weights as in the case of the UK for both Germany and Spain: for higher than average dividend-price ratios, the allocation decreases with the time horizon. However, when the estimation risk is incorporated, the sensitivity is reduced significantly at longer horizons due to higher parameter uncertainty.

Considering an investment horizon of 10 years, it can be noticed that the portfolio weights tend to converge regardless of the initial level of the dividend-price ratio or inflation.

[**Insert Figure 4 here**]
[**Insert Figure 5 here**]

Figures 6 and 7 show, for Germany and Spain respectively, the sensitivity of portfolio weights to initial levels of inflation. We observe uniformly that the higher the initial inflation, the lower the allocation to stocks. This result was expected, since inflation negatively affects future stock returns in both countries. Thus, for a lower than average inflation, the investor expects higher cumulative predicted returns, and thus allocates more to stocks. The slope of the curve is negative for very low initial inflation levels, alike the case when the initial level of the dividend-price ratio is high. For average and higher than average inflation levels, the allocation to stocks increases with the investment horizon.

[**Insert Figure 6 here**]
[**Insert Figure 7 here**]

When estimation risk is taken into account, the allocation to stocks decreases. However, as Figure 3 reveals, since UK estimation risk is low, this effect is barely noticeable over our sample. There is however a tendency to have less sensitivity to initial conditioning values for cases with greater levels of uncertainty as time horizon increases, as confirmed by Barberis (2000) for the US. This result is easier to detect in the case of Spain and is very strong in the case of Germany. For Germany, as noticed in Figures 5 and 7, because of the high uncertainty, the allocation to stocks converges faster to the one obtained for the average level of the explanatory variable for long investment horizons. Another important effect of estimation risk is observed for Germany: there, the portfolio weights are not monotonic in the initial values of the predictor variables. For high initial values of dividend-price ratio, the investor allocates less than for lower values (or in the case of inflation, for lower initial values of inflation the investor allocates less than for higher values). The intuition behind this result may be found in Stambaugh (1999) or Barberis (2000). These authors argue that the presence of a time-changing parameter in the regression implies for returns a non-normal distribution. More precisely, if one focuses on the dividend-price ratio, one knows that, because of the negative regression coefficient, a high level will lead to low future returns. The explanatory variable thus induces like a change in the mean while the error terms remain Gaussian. From a time-series point of view, this means that the distribution of returns alone becomes a mixture of distributions and the skewness thereof will be negative because returns will be on the low (negative) side. Clearly, given that one optimizes the portfolio using a CRRA utility function, as shown by Jondeau and Rockinger (2006), one expects a decreased allocation to stocks for negatively skewed returns. The same negative skewness could be introduced for low levels of inflation, explaining the non-monotonicity in this case.

At this point, we now discuss the tendencies for the other countries. We first notice that all countries present mean reversion in stock returns in the sense that the conditional standard deviation decreases with the investment horizon.[17]

The countries with the highest speed of mean reversion, after the UK, are Netherlands and Norway with a ratio of 10-year to 1-year conditional standard deviations of 0.74 for Netherlands and 0.72 for Norway. The next group is France, Switzerland, and Sweden with ratios of 0.84, 0.86, and 0.84. The lowest speed of mean reversion in stock returns is recorded for Germany, Italy, Denmark, and Spain with ratios of 0.90, 0.91, 0.95, and 0.98. All these countries have the tendencies presented in Figure 2, with some countries allocating less to stocks at long horizons in the case of predictability and parameter uncertainty than in the i.i.d. case (the cases where the estimation risk is higher), with the extreme case of Germany, which was presented above.

It should be noticed that the result which states that predictability of stocks leads to increased allocations in stocks is mainly driven by the inclusion of the dividend-price ratio in the set of predictors. This leading role played by the dividend-price ratio comes from the highly negative correlation between shocks in dividend-price ratios, shocks in returns, and the positive coefficient in the VAR system. This may explain why in the literature dealing with portfolio choice under predictable returns, the dividend-price ratio is always included in the set of explanatory variables. Obviously, the argument by Cochrane (2009) that testing for statistical significance of the dividend-price ratio is a difficult task due low power in the test is also of relevance. The results presented in this section tend to suggest that even if parameters have small point estimates and are possibly statistically insignificant, from an economic point of view, they may nonetheless play an important role. These findings suggest that less emphasis should be put on statistical tests if the dividend-price ratio matters and more on the economic relevance of this variable for actual allocations.

3.5 Results for the Dynamic Investment Strategy

We now turn to the results regarding the dynamic allocation problem. As in the previous sections, we will focus now on Germany, the UK, and Spain and will present the results for the reduced model, the one that used dividend-price ratio (for the UK) and both the dividend-price ratio and the inflation rate (for Germany and

[17]For Denmark and Spain, there exists mean reversion in stock returns, but in the sense that the line of conditional standard deviation decreases for investment horizons higher than 2 years in the case of Denmark and higher than 6 years in the case of Spain.

Spain).

3.5.1 Optimal Portfolio Choice

Figure 8 shows the optimal weight for the UK investor in the cases without parameter uncertainty (in the left part) and with parameter uncertainty (in the right part) for different initial values of dividend-price ratio.[18] The continuous line shows the allocation for an initial level of the dividend-price ratio equal to its historical average.

[Insert Figure 8 here]

The figure shows that there is a pattern similar to the one observed in the buy-and-hold case when the initial value of the dividend-price ratio is set at its long-run average (see also Figure 3): the allocation to stocks increases with the time horizon. This pattern is obtained whether parameter uncertainty is taken into account or not. The reason for this is again the positive coefficient of the dividend-price ratio in the VAR system and the negative correlation between the shocks. However, contrary to the buy-and-hold strategy, the monotone increasing shape of the allocation holds for all initial values of the dividend-price ratio. Clearly, for lower than average dividend-price ratios, the investor allocates less to stocks than in the case of an initial dividend-price ratio held at its historical average. The systematic increase in optimal portfolio weights with time horizon is due to the so-called intertemporal hedging demand first discussed by Merton (1973). Imagine a strong drop in stock returns. Because of the negative and strong correlation of -0.9 between the return shocks and the dividend-price ratio shocks, it is very likely that the dividend-price ratio will increase. Now, if the divided-price ratio increases, this signals higher future returns (due to the positive coefficient in the VAR). The implication is that the investor can at that moment consider investing more in stocks and thereby hedge the suffered drop in stocks. Thus, a more aggressive strategy may be adapted at the inception of the portfolio, if the investor faces a long time horizon since future drops in the stock may be hedged by future reallocations. If one compares carefully Figure 3 with Figure 8 and focuses on the long-horizon allocation of 10 years (taking for instance the case without parameter uncertainty), one notices that for the buy-and-hold investor the allocation to stocks would be about 50% ($A = 10$). In the case of the investor who dynamically allocates her wealth, this amount increases up to 85% (still $A = 10$). This represents a huge difference and indeed, this would be consistent with the ability of the investor to time the market.

[18] As previously, the five levels of the explanatory variable are the average, the average plus/minus 1 standard deviation, and the average plus/minus 2 standard deviations. Careful observation of this and the coming figures shows that there are slight kinks in the curve. These come from the discretization of the state space.

When estimation risk is considered, as displayed in the right part of Figure 8, the allocation to stocks decreases compared to the case when it is not taken into account and the allocation curves are flatter since the investor is more uncertain about the predictive power of the dividend-price ratio for future stock returns. The argument for hedging demand is therefore still present but not as strongly as before.

The results for Spain and Germany are presented next. Figures 9 and 10 display the optimal allocation when the dividend-price ratio and inflation rate are kept at their historical average.

[**Insert Figure 9 here**]
[**Insert Figure 10 here**]

First, we compare **Figure 9** with its buy-and-hold counterpart contained in Figure 2. Whereas the buy-and-hold allocation for Spain varies in an inverse U shape between 40% and 50% for an investor with $A = 3$ and between 20% and 30% with $A = 5$, in Figure 9 the allocations (taking parameter uncertainty into account) increase monotonically from 65% to 77% if $A = 3$ and from 42% to 50% if $A = 5$. The allocations to stocks are therefore between some 40% ($A = 3$) and 20% ($A = 5$) higher in the case of Spain. This represents clearly a significant change.

If we compare **Figure 10**, for Germany with its counterpart in Figure 2, we notice for the short term a much smaller difference than for Spain or for the UK. The main difference is that the allocation increases now monotonically with the time horizon for the dynamic allocation whereas it decreased previously.

Next, if we compare the left part (parameter uncertainty is ignored) with the right part (parameter uncertainty is integrated out) of Figures 9 and 10, we notice that the allocations are rather similar except for Germany and the case of an investor with a high level of risk aversion. The figures thus show that, unlike the buy-and-hold case, where parameter uncertainty played a role, in the dynamic allocation case, it is too small to play any sizeable role. One possible reason is that in the dynamic setting the investor has the possibility to rebalance the portfolio every year. Now, if the parameter gets rebalanced, it means that the level of the state variable has changed quite significantly from last year. Therefore, the allocation becomes similar to an allocation with a very short time horizon for which we have seen, even for the buy-and-hold strategy, that the impact of parameter uncertainty is small. In other words, it appears that the changes in the state variable and in the resulting allocation dominate the effect of parameter uncertainty. Whereas changes in the state variable appears to play a first order effect, parameter uncertainty then plays a second order effect.

Still, we need to remind that in this case we do not fully incorporate learning about parameters and about their uncertainty. We expect that estimation risk might have a bigger impact if this full learning was incorporated.

3.5.2 Sensitivity of Portfolio Weights to the Explanatory Variables

Figures 11 and 12 show the sensitivity of portfolio weights to initial values of the predictor variables for Germany. We only present the results for Germany since the ones for the UK and Spain are qualitatively similar.

[**Insert Figure 11 here**]
[**Insert Figure 12 here**]

The results for Germany are similar to the buy-and-hold case: the allocation to stocks increases with the time horizon, but again this happens for all different initial levels of predictors, unlike in the buy-and-hold case. The allocation is monotonic in the initial values of the predictor variables and the same trends as in the buy-and-hold case are observed: for higher initial dividend-price ratio, the investor allocates more to stocks than for an initial value equal to the historical average of the dividend-price ratio; and for lower than average initial inflation, the investor allocates more to stocks than in the case of average initial inflation. To conclude, the hedging demands are very small and decrease as the level of risk aversion increases.

As discussed earlier, in the dynamic allocation case, a high correlation between the cumulative excess return and next-period state variables and a high next-period utility are important to yield horizon effects. Without such a high correlation, the allocation is essentially independent from the investment horizon, since the next-period utility has no impact on a given period portfolio weight. Therefore, the higher the correlation between the cumulative excess return and next-period state variables, the higher the horizon effects. This correlation depends on the VAR coefficients and variance-covariance matrix of error terms and reaches the highest level for UK data.

It is interesting to identify the factors which render hedging demands and horizon effects for Spain and Germany significantly smaller than for the UK. As mentioned in a previous section, the Sharpe ratio for stocks is the highest for the UK market, and the coefficient of dividend-price ratio in the VAR system is also the highest (as seen in Table 5) amongst these countries. Last but not least, the UK has the highest initial correlation between unexpected excess returns and shocks in the dividend-price ratio. All these characteristics have an important impact in dynamic allocation. Since the risk-free rate has comparable levels across countries, but the UK has the highest Sharpe ratio, the UK investor allocates more to stocks in the case of both buy-and-hold and dynamic investment strategies. Together with the high coefficient on the dividend-price ratio and high negative correlation, UK stocks become a very good hedge against variation in their future returns. These factors together have a weaker impact for the other countries, leading to smaller hedging demands.

4 Conclusion

This paper investigates the role of predictability and parameter uncertainty for long-term allocations with a focus on European countries. From a methodological point of view, we consider linear relations and use a Bayesian setting with diffuse priors to integrate uncertainty out. We also allow for both a buy-and-hold strategy as well as a dynamic rebalancing strategy. All the reported investigations are calibrated with data sampled at a monthly frequency.

A first investigation considers predictability both in and out of sample. We confirm the findings already reported in the literature, for different datasets, that there is empirical evidence for predictability of returns by the dividend-price ratio. Predictability of the dividend-price ratio is particularly strong for the UK market. For several countries we also obtain that inflation predicts returns. As one turns to the out-of sample analysis, predictability appears to be rather weak.

Theoretical arguments of Cochrane (2001) suggest, however, that the dividend-price ratio should have predictive power for returns. To investigate the economic relevance of this weak predictability, we focus on just three countries that have different characteristics in terms of parameter uncertainty, predictive variables, and correlation structure of the innovations, namely the UK, Germany, and Spain.

Considering the buy-and-hold strategy, we obtain that predictability and parameter uncertainty plays a significant role especially for long time horizons. In certain cases, such as Germany, parameter uncertainty may undo the effect of predictability. However, for the UK and Spain, the tendency is that predictability is so strong that, even if one takes parameter uncertainty into account, the resulting allocations differ from the ones of an investor who would neglect predictability and parameter uncertainty.

When we turn to the dynamic allocation, where portfolios may be rebalanced each year (still maintaining the monthly calibration), we obtain quite dramatic effects in comparison with the buy-and-hold strategy. Intuitively, an allocation that cannot be modified over the next ten years should be more conservative than one where the position can be modified regularly. Thus, regular rebalancing allows the investor to carry more risk. Another important feature we obtain is that parameter uncertainty may play only a second-order effect in comparison with the first-order effects that stem from variations in the explanatory variables. This observation is consistent with the one concerning the buy-and-hold strategy where for the short-term uncertainty plays a relatively small role.

We conclude that overall, even though we obtain little statistical evidence for out-of-sample predictability, once one implements a dynamic strategy based on the point estimates, there is indeed economic relevance. Such observations may be of relevance for the ongoing debate on pension funds and the type of strategy that they should use.

5 Appendix

5.1 Appendix 1: Description of the data

We follow in our construction Campbell and Viceira (1999), and Bec and Gollier (2009). The data for equity price indices (PI) and return indices (RI) at monthly frequency comes from the Morgan Stanley Capital International (MSCI) database available on Datastream. This data starts in December 1969.

The macro data is also collected at monthly frequency. The monthly Consumer Price Indices (CPI) comes from the OECD statistics. The interest rates are from the International Monetary Fund (IMF) - International Financial Statistics (IFS) CD-ROM. For the long-term interest rate ($ltir$), we use the monthly series of 10-year Government bond yields. For the short-term interest rates ($stir$), we use different series of rates according to their availability: the call money rate for France, Germany, Denmark, Netherlands, and Spain; the money market rate for Italy, Switzerland, and Norway; the overnight interbank rate for the UK; the 3-month Treasury discount note rate for Sweden.

The sample of macro data starts in January 1975 for all the countries, except Switzerland (September 1975) and Spain (March 1975). The sample ends in April 2008. The period January 2005 to April 2008 is used for the out-of-sample investigation.

The monthly total return series and dividend series are obtained using the following methodology: The dividend yield for each month is calculated as

$$DY_t = (RI_t/RI_{t-1})/(PI_t/PI_{t-1}) - 1.$$

The dividend for each month is calculated as

$$D_t = DY_t \times PI_t.$$

The dividend-price ratio, calculated as the ratio of the dividends over the past year to the current price, is defined as

$$dp_t = \frac{1}{PI_t} \sum_{i=0}^{11} D_{t-i}.$$

The log inflation over the last year (π) is computed from the monthly CPI for each month, as:

$$\pi = \log(CPI_t/CPI_{t-12}).$$

The log inflation over the month or period log inflation (πp) is computed from the monthly CPI for each month, as:

$$\pi p = \log(CPI_t/CPI_{t-1}).$$

The dependent variable is the log excess return on equities (r) defined as the difference between the log return on equities (Rt) and the log of short term nominal interest rate (r_0) (which is a proxy for the risk free rate)

$$r_t = \log(PI_{t+1} + D_t) - \log(PI_t) - r_0,$$
$$r_0 = \log(1 + stir)/12.$$

The set of explanatory variables used to explain excess return is given by:

- log excess return on bonds (r_b) - computed using the long-term bond return ($LTret$) based on the log-linear approximation technique described in Campbell, Lo, and MacKinlay (1997, Chapter 14):

$$r_b = LTret - r_0.$$

- log short-term real interest rates (r_f)

$$r_f = r_0 - \pi p.$$

- log dividend-to-price ratio (dp), computed as above.
- the spread (spr) between the long-term bond yield and short-term interest rate:

$$spr = ltirp - r_0,$$

  where $ltirp = \log(1 + ltir)/12$.
- the log yearly inflation (π).

5.2 Appendix 2: Out-of-sample Performance Measures

We computed the following performance measures:

- the root mean squared error (**RMSE**). If one denotes by Y_t the observed return and by $\hat{Y}_t$ its predictor, the forecast error is $\varepsilon_t = Y_t - \hat{Y}_t$. The **RMSE** measures the square root of the average of the squared difference between the estimator and the variable to be estimated,

$$\mathbf{RMSE} = \sqrt{\frac{1}{df}\sum_{t=1}^{T}\varepsilon_t^2},$$

  where df is the degree of freedom of the residual term, computed as the difference between the number of observations (T) and the number of regressors (n) used in regression (1).

- we follow Bossaerts and Hillion (1999) and consider the out-of-sample regression $Y_t = a + b\hat{Y}_t + \varepsilon_t$, for which we present the t-test of the coefficient of the predicted returns (denoted by **tstat2**) and the R^2. Ideally, high values for **tstat2** and R^2 are desired.

- the percentage of correct predictions of the sign of returns (**corrs**). This indicator measures the percentage of occasions when we were able to predict whether one would make a profit by investing in equities if stock markets are predicted to go up:
$$\mathbf{corrs} = \frac{100}{T} \times \sum_{t=1}^{T} 1_{\{Y_t \times \hat{Y}_t > 0\}}.$$
The preferred model is the one with the highest value of the **corrs** indicator.

- the percentage of correct predictions of variations of returns as compared with the current return and this independently of the sign of returns (**corrupd**). This indicator displays the frequency of correct prediction of future returns with respect to the current return:
$$\mathbf{corrupd} = \frac{100}{T} \times \sum_{t=1}^{T-1} 1_{\{((\frac{Y_{t+1}}{Y_t} >= 1) \wedge (\frac{\hat{Y}_{t+1}}{\hat{Y}_t} >= 1)) \vee ((\frac{Y_{t+1}}{Y_t} < 1) \wedge (\frac{\hat{Y}_{t+1}}{\hat{Y}_t} < 1))\}},$$
where $\wedge$ is the logical and, the symbol $\vee$ denotes the logical or. According to this indicator, the preferred model is the one with the highest **corrupd** value.

- Akaike's information criterion (**AIC**):
$$\mathbf{AIC} = T \log \frac{SSE}{T} + 2n,$$
where SSE is the sum of squared errors. The preferred model using this criterion is the one with the lowest **AIC** value.

- Schwarz's Bayesian information criterion (**BIC**):
$$\mathbf{BIC} = T \log \frac{SSE}{T} + n \log T.$$
According to this criterion, we prefer the model with the lowest **BIC** value.

5.3 Appendix 3: Numerical integration

Whenever possible, for reasons of speed, we compute integrals by Gaussian Quadrature. The Gaussian Quadrature is a class of integration techniques used to compute an approximation to the integral
$$\int_a^b F(x)\rho(x)dx \simeq \sum_{j=1}^{M} w_j F(x_j),$$

where $F(x)$ is a function to be integrated, $\rho(x)$ is some weighting scheme. The Gaussian integration, thus, proposes a weighing scheme w_i and abscissa x_i, i.e. points where the function $F(x)$ should be evaluated. Typically, the approximation of the integral will be exact if $F(x)$ is a polynomial of degree lower than $2M$, where M is the number of quadrature points used for the approximation. In this paper, we have to deal with integrals of the type

$$\int_{-\infty}^{+\infty} F(x)e^{-x^2}dx. \tag{15}$$

For such cases, that is when $\rho(x) = e^{-x^2}$, $a = -\infty$, and $b = +\infty$, the Gauss-Hermite Quadrature rule is optimal. For instance, if one has to integrate

$$I = \int_{-\infty}^{\infty} F(x)\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx, \tag{16}$$

one considers the change of variable:

$$y = \frac{x-\mu}{\sigma\sqrt{2}} \quad \text{or} \quad x = \sigma\sqrt{2}y + \mu,$$

so that the integral in Equation (16) becomes:

$$I = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} F(\sigma\sqrt{2}y + \mu)e^{-y^2}dy.$$

Now, I can be approximated using Gauss-Hermite weights and abscissa

$$I \simeq \frac{1}{\sqrt{\pi}}\sum_{i=1}^{n} w_i F(\sigma\sqrt{2}x_i + \mu).$$

For the optimization problem used in this paper, we have

$$F(x) = F(R_{T+\hat{T}}) = \frac{\{(1-\omega)\exp(r_f\hat{T}) + \omega\exp(r_f\hat{T} + R_{T+\hat{T}})\}^{1-A}}{1-A},$$

and

$$\rho(x) = p(R_{T+\hat{T}}|z, \hat{\theta}).$$

By using just a few points, it is possible to evaluate the integrals with a high degree of accuracy.

5.4 Appendix 4: Moments of Cumulative Returns

We describe in this appendix how to compute the covariance matrix between the cumulative return R_{t+k} and the state variables x_{t+k}. We begin with the case of one state variable only. For this purpose, we consider the following VAR(1) system for $z_t = (r_t \; x_t)'$:

$$
\begin{aligned}
z_t &= a + Bz_{t-1} + \varepsilon_t, \\
\varepsilon_t &= \begin{pmatrix} \varepsilon_{r,t} \\ \varepsilon_{x,t} \end{pmatrix} \overset{i.i.d.}{\sim} N(0, \Sigma), \\
\text{and } \Sigma &= \begin{pmatrix} Var(\varepsilon_r) & Cov(\varepsilon_r, \varepsilon_x) \\ Cov(\varepsilon_r, \varepsilon_x') & Var(\varepsilon_x) \end{pmatrix}.
\end{aligned}
$$

Iterating forwards gives:

$$
\begin{aligned}
z_{t+k} &= a + Ba + B^2 a + ... + B^{k-1} a + B^k z_t \\
&\quad + B^{k-1}\varepsilon_{t+1} + B^{k-2}\varepsilon_{t+2} + ... + B\varepsilon_{t+k-1} + \varepsilon_{t+k},
\end{aligned} \tag{17}
$$

and

$$
\begin{aligned}
Z_{t+k} &= z_{t+1} + z_{t+2} + ... + z_{t+k} \\
&= [k + (k-1)B + (k-2)B^2 + ... + B^{k-1}]a + (B^k + B^{k-1} + ... + B)z_t \\
&\quad + (I + B + ... + B^{k-1})\varepsilon_{t+1} + (I + B + ... + B^{k-2})\varepsilon_{t+2} + \cdots \\
&\quad + (I + B)\varepsilon_{t+k-1} + \varepsilon_{t+k},
\end{aligned} \tag{18}
$$

where $z_{t+k} = \begin{pmatrix} r_{t+k} & x_{t+k} \end{pmatrix}'$ and $Z_{t+k} = \begin{pmatrix} R_{t+k} & X_{t+k} \end{pmatrix}'$. Thus, z_{t+k} contains the k-period ahead variables, while Z_{t+k} is the vector of k-period ahead cumulative variables.

Conditional on a, B, and Σ, we observe that the distributions of z_{t+k} and Z_{t+k} are given by $z_{t+k} \sim N(\mu_z, \Sigma_z)$ and $Z_{t+k} \sim N(\mu_Z, \Sigma_Z)$ respectively, where:

$$
\begin{aligned}
\mu_z &= a + Ba + B^2 a + ... + B^{k-1} a + B^k z_t \\
\Sigma_z &= B^{k-1}\Sigma B^{k-1\prime} + B^{k-2}\Sigma B^{k-2\prime} + ... + B\Sigma B' + \Sigma,
\end{aligned}
$$

and

$$
\begin{aligned}
\mu_Z &= [k + (k-1)B + (k-2)B^2 + ... + B^{k-1}]a + (B^k + B^{k-1} + ... + B)z_t, \\
\Sigma_Z &= \Sigma + (I + B)\Sigma(I + B)' + (I + B + B^2)\Sigma(I + B + B^2)' + \cdots \\
&\quad + (I + B + B^2 + ... + B^{k-1})\Sigma(I + B + B^2 + ... + B^{k-1})'.
\end{aligned}
$$

Then, we easily obtain the mean and variance of R_{t+k} and x_{t+k}:

$$
\begin{aligned}
\mu_R &= \mu_Z(1), & \sigma_R^2 &= \Sigma_Z(1,1), \\
\mu_x &= \mu_z(2), & \sigma_x^2 &= \Sigma_z(2,2).
\end{aligned}
$$

The covariance $Cov(x_{t+k}, R_{t+k})$, can be obtained by noticing that both z_{t+k} and Z_{t+k} are functions of the same error terms ε_{t+i}, with $i = 1, \cdots, k$. Since ε_t is i.i.d., we have $Cov(\varepsilon_{t+i}, \varepsilon_{t+j}) = 0$ for $i \neq j$, and $Cov(\varepsilon_{t+i}, \varepsilon_{t+i}) = \Sigma$. We denote the coefficient multiplying ε_{t+i} in the formula of z_{t+k} by cz^i (Equation (17)) and the coefficient of ε_{t+i} in the formula of Z_{t+k} by cZ^i (Equation (18)). Then, since these coefficients are matrices, we write the i^{th} component of the formula of z_{t+k} and Z_{t+k} as:

$$\begin{pmatrix} cz_{11}^i & cz_{12}^i \\ cz_{21}^i & cz_{22}^i \end{pmatrix} \begin{pmatrix} \varepsilon_{r,t+i} \\ \varepsilon_{x,t+i} \end{pmatrix} = \begin{pmatrix} cz_{11}^i \ \varepsilon_{r,t+i} + cz_{12}^i \ \varepsilon_{x,t+i} \\ cz_{21}^i \ \varepsilon_{r,t+i} + cz_{22}^i \ \varepsilon_{x,t+i} \end{pmatrix},$$

and

$$\begin{pmatrix} cZ_{11}^i & cZ_{12}^i \\ cZ_{21}^i & cZ_{22}^i \end{pmatrix} \begin{pmatrix} \varepsilon_{r,t+i} \\ \varepsilon_{x,t+i} \end{pmatrix} = \begin{pmatrix} cZ_{11}^i \ \varepsilon_{r,t+i} + cZ_{12}^i \ \varepsilon_{x,t+i} \\ cZ_{21}^i \ \varepsilon_{r,t+i} + cZ_{22}^i \ \varepsilon_{x,t+i} \end{pmatrix}.$$

Since z_{t+k} and Z_{t+k} involve a sum of these elements from $i = 1, \cdots, k$ and given that R_{t+k} is the first element of Z_{t+k} and x_{t+k} is the second element of z_{t+k}, we have

$$\begin{aligned} \sigma_{xR} = Cov(x_{t+k}, R_{t+k}) &= \sum_{i=1}^{k} Cov(cz_{21}^i \ \varepsilon_{r,t+i} + cz_{22}^i \ \varepsilon_{x,t+i}, cZ_{11}^i \ \varepsilon_{r,t+i} + cZ_{12}^i \ \varepsilon_{x,t+i}) \\ &= \sum_{i=1}^{k} [cz_{21}^i \ cZ_{11}^i \ Var(\varepsilon_r) + cz_{21}^i \ cZ_{12}^i \ Cov(\varepsilon_r, \varepsilon_x) \\ &\quad + cz_{22}^i \ cZ_{11}^i \ Cov(\varepsilon_r, \varepsilon_x) + cz_{22}^i \ cZ_{12}^i \ Var(\varepsilon_x)]. \end{aligned}$$

Now, all the elements to determine the distribution of $Rx_{t+k} = \begin{pmatrix} R_{t+k} & x_{t+k} \end{pmatrix}'$ are known:

$$Rx_{t+k} \sim N\left(\mu_{Rx}, \Sigma_{Rx}\right),$$

with

$$\begin{aligned} \mu_{Rx} &= \begin{pmatrix} \mu_R & \mu_x \end{pmatrix}', \\ \Sigma_{Rx} &= \begin{pmatrix} \sigma_R^2 & \sigma_{xR} \\ \sigma_{xR} & \sigma_x^2 \end{pmatrix}. \end{aligned}$$

In the case of two state variables, denoted by $x1_t$ and $x2_t$, we need the distribution of $Rx1x2_{t+k} \equiv \begin{pmatrix} R_{t+k} & x1_{t+k} & x2_{t+k} \end{pmatrix}'$. We set

$$\begin{aligned} \mu_{Rx1x2} &= \begin{pmatrix} \mu_R & \mu_{x1} & \mu_{x2} \end{pmatrix}', \\ \Sigma_{Rx1x2} &= \begin{pmatrix} \sigma_R^2 & \sigma_{x1R} & \sigma_{x2R} \\ \sigma_{x1R} & \sigma_{x1}^2 & \sigma_{x1x2} \\ \sigma_{x2R} & \sigma_{x1x2} & \sigma_{x2}^2 \end{pmatrix}. \end{aligned}$$

The method explained above for computing the covariances can then be implemented for this case or even for the general case involving n state variables.

5.5 Appendix 5

In this appendix, we summarize the pdfs used in specifying the distribution of model parameters.

- **Inverted Gamma pdf for the case of** σ

$$IG(\sigma, v, s) = \frac{2}{\Gamma(v/2)} \left(\frac{vs^2}{2}\right)^{v/2} \frac{1}{\sigma^{v+1}} \exp\left(-\frac{vs^2}{2\sigma^2}\right).$$

- **Inverted Wishart distribution:** $G \sim IW(v, H)$ if its density function is given by

$$IW(G \mid v, H) = k1 \frac{|H|^{v/2}}{|G|^{(v+m+1)/2}} \exp\left(-\frac{1}{2} tr\ G^{-1}H\right),\ |G| > 0,$$

  with

$$k1 = \{2^{vm/2}\pi^{m(m-1)/4}\Pi_{i=1}^{m}\Gamma[(v+1-i)/2]\}^{-1},\ v \geq m.$$

- **Matricvariate Normal distribution** (MN): If X denotes a $p \times q$ random matrix, $X \sim MN(vec(M), Q \otimes P)$ if its density function is given by:

$$p(X \mid M, Q \otimes P) = K \exp\left\{-\frac{1}{2} tr[Q^{-1}(X-M)'P^{-1}(X-M)]\right\},$$

  with

$$K = [(2\pi)^{pq}|P|^{q}|Q|^{p}]^{-1/2}.$$

- **Univariate Student Distribution pdf:** $X \sim t(\mu, s, m, v)$ if its density function is given by:

$$f_t(x \mid \mu, s, m, v) = C^{-1}[s + m(x-\mu)^2]^{-(v+1)/2},$$

  with

$$C = \frac{\Gamma(\frac{v}{2})}{\Gamma(\frac{v+1}{2})}\pi^{1/2}s^{-v/2}m^{-1/2}.$$

  Clearly, this definition differs from the one used traditionally in statistics.

References

[1] Balduzzi, P., and A.W. Lynch (1999), Transaction Costs and Predictability: Some Utility Cost Calculations, Journal of Financial Economics 52, 47–78.

[2] Barberis, N. (2000), Investing for the Long-run when Returns Are Predictable, Journal of Finance 55, 225–264.

[3] Bec, F., and C. Gollier (2009), Assets Returns Volatility and Investment Horizon: The French Case. CESifo Working Paper Series 2622.

[4] Bossaerts, P., and P. Hillion (1999), Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn, Review of Financial Studies 12, 405–428.

[5] Brennan, M.J., E.S. Schwartz, and R. Lagnado (1997), Strategic Asset Allocation, Journal of Economic Dynamics and Control 21, 1377–1403.

[6] Campbell, J.Y. (1987), Stock Returns and the Term Structure, Journal of Financial Economics 18, 373-400.

[7] Campbell, J.Y. (1996), Understanding Risk and Return, Journal of Political Economy 104, 298–345.

[8] Campbell, J.Y., A. Lo, A.C. MacKinlay (1997), The Econometrics of Financial Markets, Princeton University Press.

[9] Campbell, J.Y., and R.J. Shiller (1988a), The Dividend Price Ratio and Expectations of Future Dividends and Discount Factors, Review of Financial Studies 1, 195–228.

[10] Campbell, J.Y., and R.J. Shiller (1988b), Stock Prices, Earnings, and Expected Dividends, Journal of Finance 43, 661–676.

[11] Campbell, J.Y., and S. Thompson (2008), Predicting the Equity Premium Out Of Sample: Can Anything Beat the Historical Average?, Review of Financial Studies 21, 1509–1531.

[12] Campbell, J.Y., and L.M. Viceira (1999), Consumption and Portfolio Decisions when Expected Returns Are Time Varying, Quarterly Journal of Economics 114, 433–495.

[13] Cochrane, J.H. (1999), New Facts in Finance, Economic Perspectives 33, 36–58.

[14] Cochrane, J.H. (2001), Asset Pricing. Princeton University Press.

[15] Cochrane, J.H. (2008), The Dog That Did Not Bark: A Defense of Return Predictability, Review of Financial Studies 21, 1533–1575.

[16] Dow, J., and S.R. da Costa Werlang (1992), Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolios, Econometrica 60, 197–204.

[17] Fama, E.F., and K.R. French (1988), Dividend Yields and Expected Stock Returns, Journal of Financial Economics 22, 3–25.

[18] Fama, E.F., and K.R. French (1989), Business Conditions and Expected Returns on Stocks and Bonds, Journal of Financial Economics 25, 23–49.

[19] Ferreira, M.A., and P. Santa-Clara, 2008, Forecasting Stock Market Returns: The Sum Of The Parts Is More Than The Whole, NBER, WP 14571.

[20] Gennotte, G. (1986), Optimal Portfolio Choice under Incomplete Infirmation, Journal of Finance 41, 733–746.

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

[22] Lettau, M., and S. Ludvigson, 2001, Consumption, Aggregate Wealth and Expected Stock Returns, Journal of Finance 56, 815–849.

[23] Jondeau, E., and M. Rockinger (2006), Optimal Portfolio Allocation under Higher Moments, European Financial Management 12, 29–55.

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

[25] Keim, D.B., and R.F. Stambaugh (1986), Predicting Returns in the Stock and Bond Markets, Journal of Financial Economics 17, 357–390.

[26] Kim, S., and E. Omberg (1996), Dynamic Nonmyopic Portfolio Behavior, Review of Financial Studies 9, 141–161.

[27] Lynch, A.W. (2001), Portfolio Choice and Equity Characteristics: Characterizing the Hedging Demands Induced by Return Predictability, Journal of Financial Economics 62, 67–130.

[28] Merton, R.C. (1969), Lifetime Portfolio Selection under Uncertainty: The Continuous-time Case, Review of Economics and Statistics 51, 247–257.

[29] Merton, R.C. (1971), Optimum Consumption and Portfolio Rules in a Continuous Time Model, Journal of Economic Theory 3, 373–413.

[30] Merton, R.C. (1973), An Intertemporal Capital Asset Pricing Model, Econometrica 41, 867–887

[31] Pastor, L., and R. Stambaugh, 2008, Predictive Systems: Living with Imperfect Predictors, Journal of Finance, forthcoming.

[32] Poterba, J., and L.H. Summers (1988), Mean Reversion in Stock Returns: Evidence and Implications, Journal of Financial Economics 22, 27–60.

[33] Samuelson, P.A. (1969), Lifetime Portfolio Selection by Dynamic Stochastic Programming, Review of Economics and Statistics 51, 238–246.

[34] Wachter, J. (2002), Portfolio and Consumption Decisions under Mean-reverting Returns: an Exact Solution for Complete Markets, Journal of Financial and Quantitative Analysis 37, 69-91.

[35] Williams, J. (1977), Capital Asset Prices with Heterogeneous Beliefs, Journal of Financial Economics 5, 219–239.

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

Captions

Table 1: This table reports for each country the extended regression, which includes a lag of the market return, of the risk-free rate, of the bond yield, of the dividend-price ratio, of the spread between the long-term and short-term rates, and of the inflation rate. For each regressor, the first row corresponds to the parameter estimates, the second row to the p-value of the estimate. The table also reports several summary statistics about the regression, described in the text and in Appendix 2. The sample period is 1.1975–12.2004.

Table 2: This table reports for each country the reduced regression, which includes a lag of the market return, of the dividend-price ratio, and of the inflation rate. For each regressor, the first row corresponds to the parameter estimates, the second row to the p-value of the estimate. The table also reports several summary statistics about the regression, described in the text and in Appendix 2. The sample period is 1.1975–12.2004.

Table 3: This table reports for Germany, the UK, and Spain the extended regression over two subsamples. Sample 1 refers to the period 1.1975–12.1984. Sample 2 refers to the period 1.1995–12.2004. The table also reports several summary statistics about the regression, described in the text and in Appendix 2.

Table 4: This table reports for each country summary statistics about the out-of-sample performance of the extended regression. The summary statistics are described in the text and in Appendix 2. Panel A refers to the out-of-sample performance measures if parameters are estimated between 1.1975–12.2004 and then kept constant. Panel B refers to the regressions where the parameters are obtained for increasing samples, the first observation being January 1975. The out-of-sample period is 1.2005–4.2008.

Table 5: This table reports for Germany, the UK, and Spain results of the multivariate regression. Panel A presents for each regressor the parameter estimate with the p-values in parenthesis below. Panel B summarizes the posterior distribution of the parameters (matrix C) and the covariance matrix of the error terms (matrix Σ): For matrix C, the table reports for each predictor the mean of posterior distribution with the standard deviation in parenthesis below. For matrix Σ, the table reports the correlations above the diagonal (in bold), and the standard deviations are presented below the means in parenthesis. Panel C presents the posterior distribution results in the case of i.i.d. returns.

Figure 1: The figure displays for Germany, the UK, and Spain the conditional and posterior values for annualized standard deviations for various horizons. The results refer to the multivariate regression reported in Table 5. The full line represents the

case where parameter uncertainty is ignored whereas the dashed line corresponds to the case where parameter uncertainty is accounted for.

Figure 2: The figure displays for Germany, the UK, and Spain the optimal portfolio weights for various values of the risk aversion A and several time horizons of the allocation. The thin lines correspond to the case where predictability is not accounted for (i.i.d. returns): the dashed line corresponds to the case where uncertainty is ignored, and the dotted one to the case where parameter uncertainty is accounted for. The thick lines correspond to the case where future returns are predictable: the line without marks corresponds to the case where parameter uncertainty is ignored and the line with crosses corresponds to the case where parameter uncertainty is integrated out.

Figure 3: The figure displays the sensitivity of portfolio weights to the initial dividend-price ratio for the UK for various values of the risk aversion A and time horizons. The five lines correspond to different initial values of dividend-price ratio: 1.45% (dotted), 2.65% (dotted with crosses), 3.84% (full, historical average), 5.03% (dashed), 6.22% (dashed with pluses). These levels are the historical average plus/minus one and two standard deviations. The plots in the left do not take estimation risk into account, while those in the right do.

Figure 4: The figure displays the sensitivity of portfolio weights to initial dividend-price ratio for Spain for various values of the risk aversion A. The five lines correspond to different initial values of dividend-price ratio: 0% (dotted), 1.95% (dotted with crosses), 5.42% (full, historical average), 8.89% (dashed), 12.36% (dashed with pluses). These levels are the historical average plus/minus one and two standard deviations. The plots in the left do not take estimation risk into account, while those in the right do. levels.

Figure 5: The figure displays the sensitivity of portfolio weights to initial dividend-price ratio level for Germany for various values of the risk aversion A. The five lines correspond to different initial values of dividend-price ratio: -0.7% (dotted), 1.5% (dotted with crosses), 2.45% (full, historical average), 3.32% (dashed), 4.2% (dashed with pluses). These levels are the historical average plus/minus one and two standard deviations. The plots in the left do not take estimation risk into account, while those in the right do.

Figure 6: The figure displays the sensitivity of portfolio weights to initial inflation level for Spain for various values of the risk aversion A. The five lines correspond to different initial values of inflation: -3.14% (dotted), 2.4% (dotted with crosses), 7.9% (full, historical average), 13.5% (dashed), 19% (dashed with pluses). These levels are the historical average plus/minus one and two standard deviations. The

plots in the left do not take estimation risk into account, while those in the right do.

Figure 7: The figure displays the sensitivity of portfolio weights to initial inflation level for Germany for various values of the risk aversion A. The five lines correspond to different initial values of inflation: -0.78% (dotted), 0.96% (dotted with crosses), 2.7% (full, historical average), 4.45% (dashed), 6.19% (dashed with pluses). These levels are the historical average plus/minus one and two standard deviations. The plots in the left do not take estimation risk into account, while those in the right do.

Figure 8: The figure displays the optimal rebalancing for a UK investor for various values of the risk aversion A. The lines represent different initial dividend-price ratios: 2.65% (dashed with pluses), 3.24% (dashed), 3.84% (continuous - historical average), 4.44% (dotted with crosses), 5.03% (dotted). The plots in the left do not take estimation risk into account, while those in the right do.

Figure 9: The figure displays the optimal rebalancing for a Spanish investor for various values of the risk aversion A. The results are presented for initial values of predictor variables equal to their historical averages. The plots in the left do not take estimation risk into account, while those in the right do.

Figure 10: The figure displays the optimal rebalancing for a German investor for various values of the risk aversion A. The results are presented for initial values of predictor variables equal to their historical averages. The plots in the left do not take estimation risk into account, while those in the right do.

Figure 11: The figure displays the sensitivity of portfolio weights to initial dividend-price ratio level for Germany for various values of the risk aversion A. The five lines correspond to different initial values of dividend-price ratio: 1.4% (dashed with crosses), 1.92% (dashed), 2.45% (continuous, historical average), 2.97% (dotted with pluses), 3.5% (dotted). The plots in the left do not take estimation risk into account, while those in the right do.

Figure 12: The figure displays the sensitivity of portfolio weights to initial inflation level for Germany for various values of the risk aversion A. The five lines correspond to different initial values of inflation: 0.6% (dashed with crosses), 1.65% (dashed), 2.71% (continuous, historical average), 3.75% (dotted with pluses), 4.79% (dotted). The plots in the left do not take estimation risk into account, while those in the right do.

Table 1: Univariate regression–Extended model

	FRA	GER	UK	ITA	SWI	SWE	DEN	NET	NOR	SPA
cst	0.003	-0.007	-0.069	-0.008	-0.008	0.003	-0.001	-0.013	-0.013	-0.002
	(0.737)	(0.526)	(0.000)**	(0.580)	(0.408)	(0.765)	(0.850)	(0.129)	(0.216)	(0.761)
r_{t-1}	0.085	0.049	0.079	0.063	0.137	0.132	0.035	0.046	0.129	0.065
	(0.114)	(0.363)	(0.119)	(0.239)	(0.013)**	(0.013)**	(0.524)	(0.381)	(0.014)**	(0.216)
$r_{f,t-1}$	-0.826	0.065	0.573	1.051	0.212	-0.225	0.278	-0.102	0.482	0.598
	(0.494)	(0.952)	(0.275)	(0.265)	(0.798)	(0.743)	(0.592)	(0.885)	(0.522)	(0.330)
$r_{b,t-1}$	-1.259	1.087	-0.145	1.978	0.745	1.327	0.484	2.573	0.040	1.783
	(0.274)	(0.346)	(0.875)	(0.165)	(0.321)	(0.308)	(0.361)	(0.011)**	(0.970)	(0.124)
dp_{t-1}	0.520	0.131	2.319	0.305	0.922	0.618	0.047	0.533	1.036	0.381
	(0.201)	(0.804)	(0.000)**	(0.486)	(0.117)	(0.104)	(0.895)	(0.015)**	(0.012)**	(0.008)**
spr_{t-1}	1.061	2.950	2.199	-4.753	-1.113	-6.200	2.117	0.384	4.008	1.362
	(0.756)	(0.368)	(0.173)	(0.199)	(0.682)	(0.047)**	(0.140)	(0.868)	(0.165)	(0.442)
π_{t-1}	-0.255	0.072	-0.295	-0.005	-0.249	-0.181	-0.016	-0.240	-0.309	-0.239
	(0.202)	(0.798)	(0.002)**	(0.949)	(0.250)	(0.185)	(0.921)	(0.125)	(0.039)**	(0.016)**
$R^2 \times 100$	1.49	1.10	11.92	1.82	2.86	3.15	1.36	4.07	3.94	6.39
$tstat2$	2.33	2.00	6.97	2.58	3.15	3.42	2.23	3.90	3.84	4.94
$F-test$	0.89	0.66	7.98	1.09	1.63	1.92	0.82	2.50	2.42	4.01
$p-val$	0.50	0.69	0.00	0.37	0.14	0.08	0.56	0.02	0.03	0.00
$RMSE \times 100$	6.08	5.94	5.36	6.96	4.74	6.83	5.11	5.17	7.16	6.17
$Corrs$	56.8	58.7	59.8	52.9	58.4	52.4	58.2	57.3	57.6	61.0
$Corrupd$	56.2	56.2	53.5	59.3	56.9	57.3	50.1	52.6	55.1	58.5
T	361	361	361	361	339	361	361	361	361	359

Table 2: Univariate regression–Reduced model

	FRA	GER	UK	ITA	SWI	SWE	DEN	NET	NOR	SPA
cst	0.001	-0.002	-0.068	0.001	-0.004	0.001	0.002	-0.008	-0.007	0.006
	(0.932)	(0.871)	(0.000)**	(0.939)	(0.661)	(0.883)	(0.716)	(0.295)	(0.459)	(0.365)
dp_{t-1}	0.431	0.279	2.372	0.078	0.732	0.403	0.051	0.477	0.904	0.377
	(0.279)	(0.572)	(0.000)**	(0.854)	(0.170)	(0.278)	(0.884)	(0.025)**	(0.027)**	(0.006)**
π_{t-1}	-0.183	-0.069	-0.281	-0.011	-0.225	-0.099	0.000	-0.178	-0.333	-0.307
	(0.349)	(0.782)	(0.001)**	(0.875)	(0.225)	(0.452)	(1.000)	(0.225)	(0.023)**	(0.000)**
$R^2 \times 100$	0.33	0.10	10.65	0.02	0.61	0.33	0.03	1.40	1.67	3.61
$tstat2$	1.09	0.59	6.54	0.23	1.44	1.09	0.31	2.26	2.47	3.66
$F - test$	0.59	0.17	21.35	0.03	1.03	0.59	0.05	2.53	3.04	6.67
$p - val$	0.55	0.84	0.00	0.97	0.36	0.56	0.95	0.08	0.05	0.00
$RMSE \times 100$	6.08	5.93	5.37	6.98	4.77	6.88	5.11	5.22	7.20	6.23
$Corrs$	56.8	56.8	57.9	49.9	61.4	54.9	54.3	59.3	56.8	59.6
$Corrupd$	50.7	49.0	51.3	50.7	50.2	51.0	47.1	50.1	52.4	52.9
T	361	361	361	361	339	361	361	361	361	359

Table 3: Univariate regression–Extended model with subperiods

	Jan 1975–Dec 1984			Jan 1995–Dec 2004		
	GER	UK	SPA	GER	UK	SPA
cst	-0.054	-0.253	-0.024	-0.015	-0.079	0.010
	(0.058)*	(0.000)**	(0.508)	(0.644)	(0.012)**	(0.672)
r_{t-1}	0.032	0.147	-0.136	-0.032	0.024	0.016
	(0.730)	(0.063)*	(0.166)	(0.739)	(0.799)	(0.862)
$r_{f,t-1}$	2.497	1.314	0.823	-1.627	0.543	0.927
	(0.046)**	(0.093)*	(0.265)	(0.473)	(0.591)	(0.583)
$r_{b,t-1}$	1.108	-0.311	0.073	2.430	1.041	2.486
	(0.386)	(0.850)	(0.971)	(0.290)	(0.423)	(0.273)
dp_{t-1}	1.682	5.255	0.407	0.311	2.338	2.169
	(0.080)*	(0.000)**	(0.024)**	(0.799)	(0.029)**	(0.043)**
spr_{t-1}	5.276	5.893	4.028	19.268	-5.740	-2.534
	(0.115)	(0.019)**	(0.091)*	(0.063)*	(0.198)	(0.803)
π_{t-1}	-0.314	-0.223	-0.128	-1.157	0.899	-1.889
	(0.412)	(0.075)*	(0.478)	(0.362)	(0.208)	(0.021)**
$R^2 \times 100$	7.06	34.08	15.07	4.91	7.59	9.15
$tstat2$	3.01	7.84	4.56	2.48	3.13	3.46
$F-test$	1.44	9.83	3.31	0.98	1.56	1.91
$p-val$	0.20	0.00	0.01	0.44	0.17	0.09
$RMSE \times 100$	3.73	5.81	5.16	7.06	3.90	6.31
$Corrs$	58.7	61.2	65.6	56.2	65.3	63.6
$Corrupd$	61.2	54.6	61.3	63.6	57.0	55.4
T	121	121	119	121	121	121

Table 4: Out-of-sample performance (1.2005–4.2008)

	FRA	GER	UK	ITA	SWI	SWE	DEN	NET	NOR	SPA
Panel A: Constant out-of-sample parameters										
$R^2 \times 100$	0.19	0.15	2.36	1.06	0.84	0.24	0.17	0.96	0.71	2.19
$tstat2$	0.26	0.24	0.95	-0.63	0.56	0.30	0.25	0.60	0.51	0.91
$RMSE \times 100$	4.28	4.81	3.27	4.16	3.85	5.26	4.54	4.56	7.11	4.52
$Corrs$	64.1	51.3	59.0	53.9	61.5	53.9	61.5	56.4	61.5	56.4
$Corrupd$	35.9	56.4	51.3	53.9	41.0	53.9	53.9	53.9	46.2	51.3
AIC	-239.54	-230.40	-260.47	-241.85	-247.94	-223.43	-235.02	-234.52	-200.00	-235.32
BIC	-227.89	-218.76	-248.82	-230.21	-236.29	-211.79	-223.37	-222.88	-188.35	-223.68
Panel B: Rolling regression										
$R^2 \times 100$	0.39	0.45	3.92	0.58	1.35	0.36	1.89	1.28	1.20	3.25
$tstat2$	0.38	0.41	1.23	-0.47	0.71	0.37	0.85	0.69	0.67	1.12
$RMSE \times 100$	4.27	4.84	3.26	4.12	3.82	5.24	4.56	4.56	7.06	4.51
$Corrs$	61.5	51.3	56.4	61.5	64.1	53.9	66.7	53.9	61.5	53.9
$Corrupd$	41.0	51.3	61.5	46.2	41.0	56.4	48.7	53.9	43.6	53.9
AIC	-239.70	-230.01	-260.77	-242.37	-248.29	-223.67	-234.58	-234.52	-200.40	-235.45
BIC	-228.06	-218.36	-249.13	-230.72	-236.65	-212.03	-222.94	-222.87	-188.75	-223.80

Table 5: Multivariate regression results

	GER			UK		SPA		
	Panel A: Parameter estimates							
	r	dp	π	r	dp	r	dp	π
cst	-0.0015	0.0005	-0.0002	-0.0451	0.0027	0.0058	-0.0001	0.0005
	(0.871)	(0.045)**	(0.607)	(0.000)**	(0.000)**	(0.365)	(0.737)	(0.458)
dp_{t-1}	0.2791	0.9750	0.0413	1.3381	0.9258	0.3778	0.9714	-0.0059
	(0.572)	(0.000)**	(0.101)	(0.000)**	(0.000)**	(0.006)**	(0.000)**	(0.651)
π_{t-1}	-0.0687	0.0029	0.9674	-	-	-0.3079	0.0208	0.9930
	(0.782)	(0.632)	(0.000)**			(0.000)**	(0.000)**	(0.000)**
$R^2 \times 100$	0.10	97.24	96.99	7.98	93.94	3.61	99.03	98.76
$p-val$	0.84	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$RMSE \times 100$	5.93	0.30	0.32	5.43	0.35	6.23	0.30	0.63
	Panel B: Posterior distribution of parameters							
		C		C			C	
	-0.0015	0.0005	-0.0002	-0.0451	0.0027	0.0058	-0.0001	0.0005
	(0.0010)	(0.0002)	(0.0005)	(0.0097)	(0.0005)	(0.0064)	(0.0003)	(0.0006)
	0.2791	0.9750	0.0413	1.3381	0.9258	0.3778	0.9714	-0.0059
	(0.4941)	(0.0121)	(0.0252)	(0.2392)	(0.0124)	(0.1346)	(0.0074)	(0.0135)
	-0.0687	0.0029	0.9674	-	-	-0.3079	0.0208	0.9930
	(0.2486)	(0.0061)	(0.0127)			(0.0839)	(0.0047)	(0.0084)
		Σ		Σ			Σ	
	0.0035	**-0.8446**	**-0.0360**	0.0030	**-0.9058**	0.0039	**-0.8044**	**-0.0772**
	(0.0003)	**(0.0152)**	**(0.0526)**	(0.0002)	**(0.0095)**	(0.0003)	**(0.0188)**	**(0.0526)**
	-	2.1E-05	**0.1330**	-	8.0E-06	-	1.2E-05	**0.0825**
		(0.0000)	**(0.0518)**		(E-06)		(E-06)	**(0.0524)**
	-	-	9.2E-06	-	-	-	-	3.8E-05
			(E-06)					(E-06)
	Panel C: Posterior distribution (i.i.d. case)							
	μ	σ^2		μ	σ^2		μ	σ^2
	0.0035	0.0035		0.0066	0.0032		0.0015	0.0040
	(0.0031)	(0.0003)		(0.0030)	(0.0002)		(0.0033)	(0.0003)

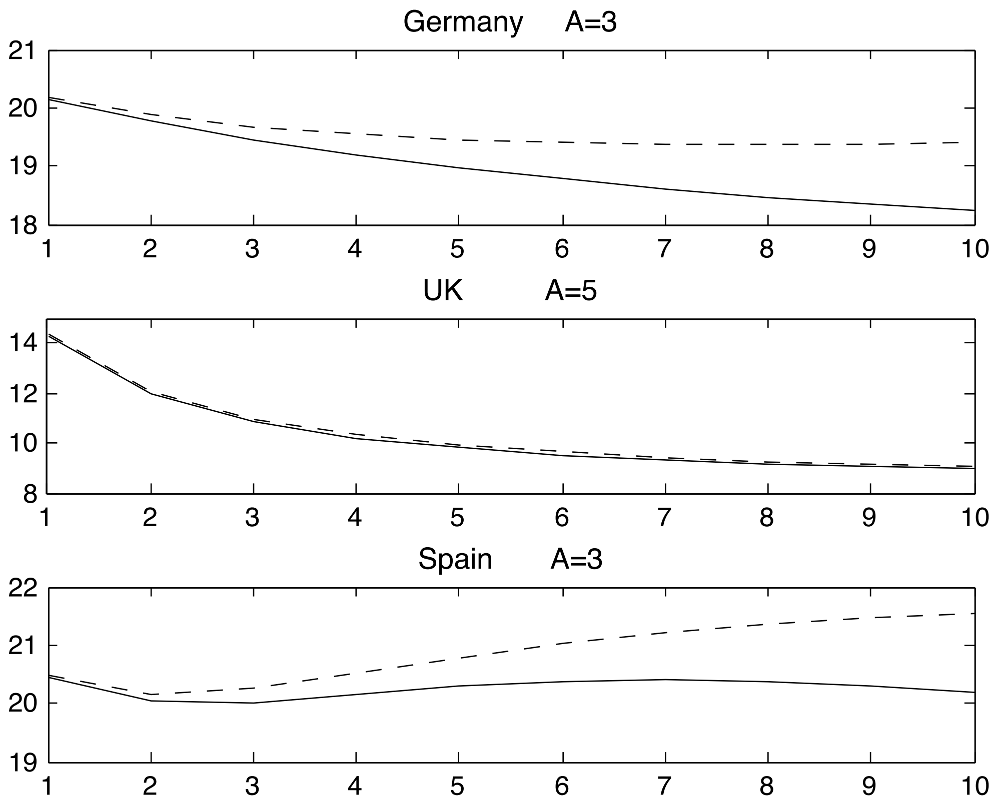

Figure 1: Conditional and posterior values for annualized standard deviations.

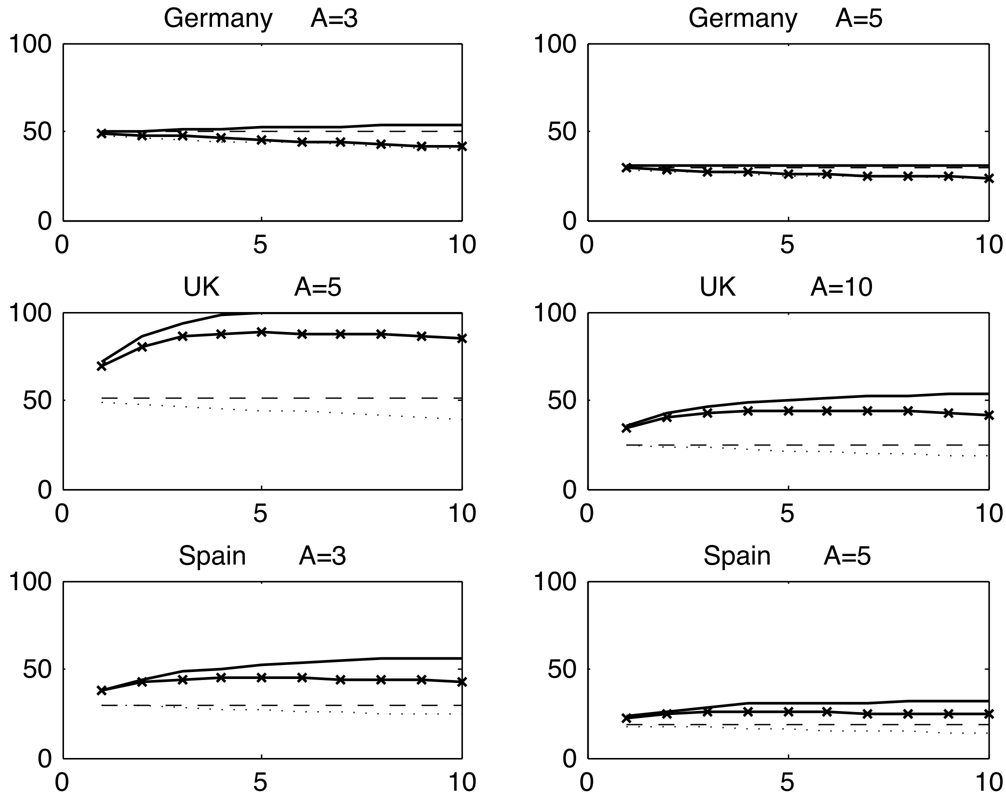

Figure 2: Buy-and-hold portfolio choice.

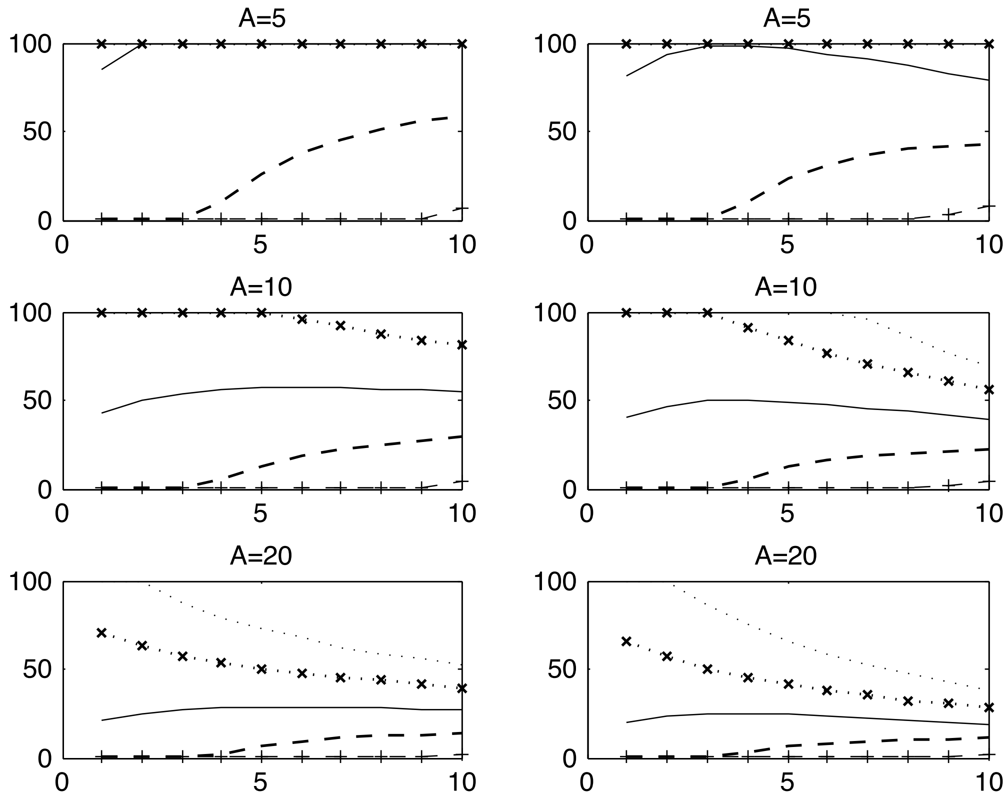

Figure 3: Sensitivity of portfolio weights to the initial dividend-price ratio for UK.

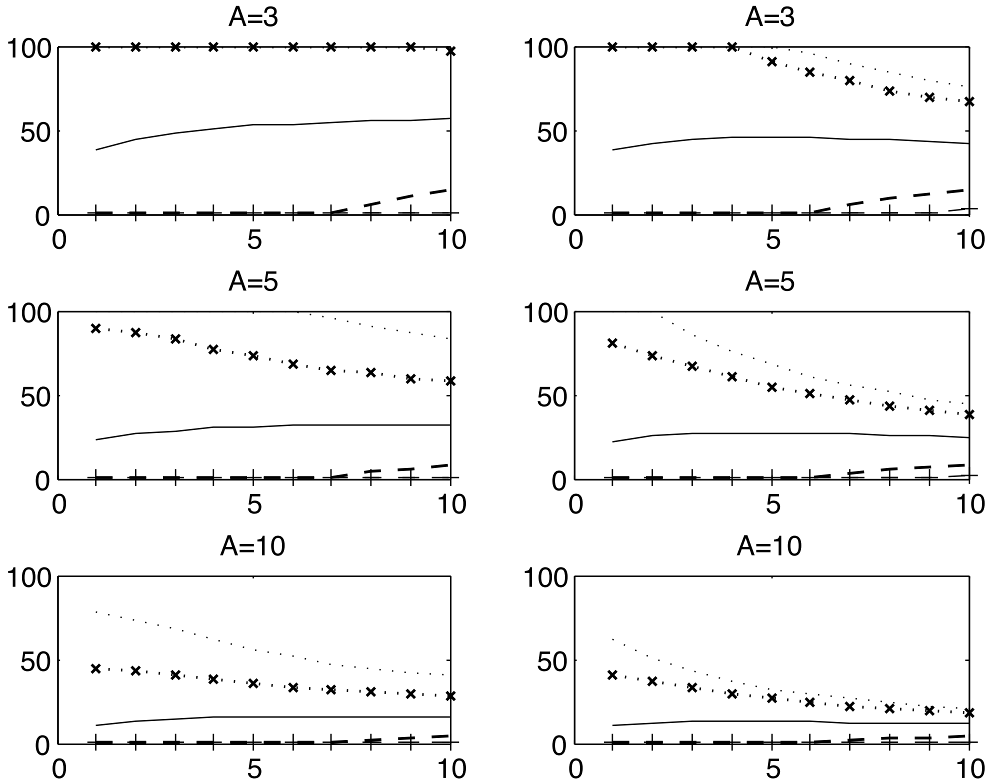

Figure 4: Sensitivity of portfolio weights to the initial dividend-price ratio for Spain.

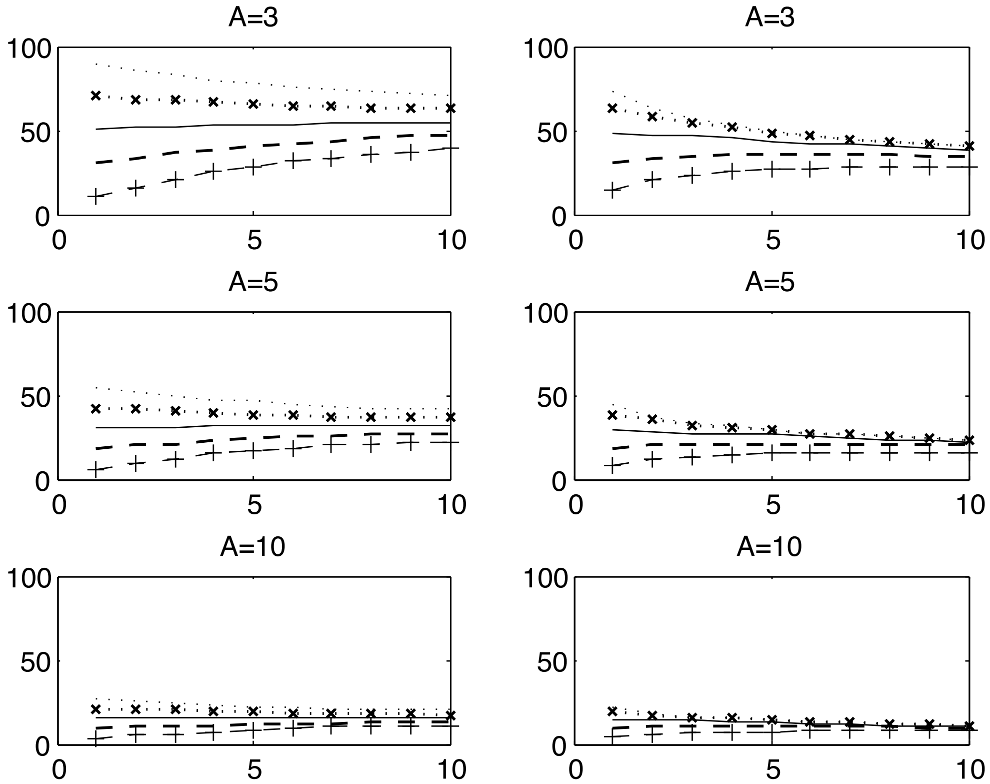

Figure 5: Sensitivity of portfolio weights to initial dividend-price ratio level for Germany.

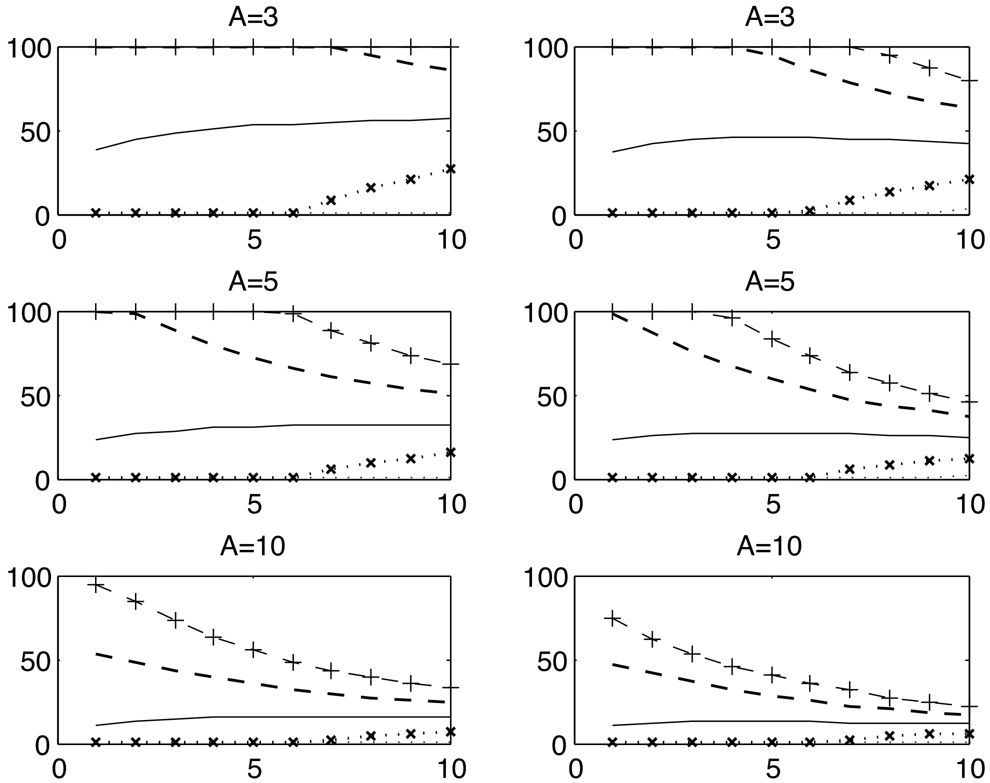

Figure 6: Sensitivity of portfolio weights to the initial inflation level for Spain.

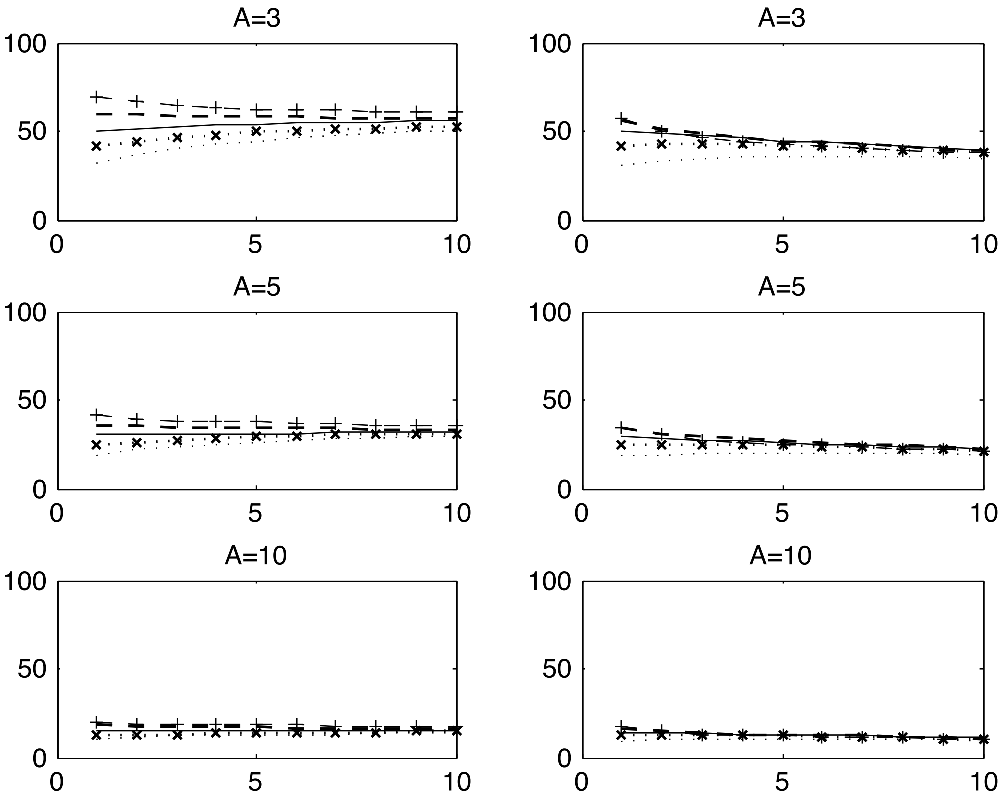

Figure 7: Sensitivity of portfolio weights to the initial inflation level for Germany.

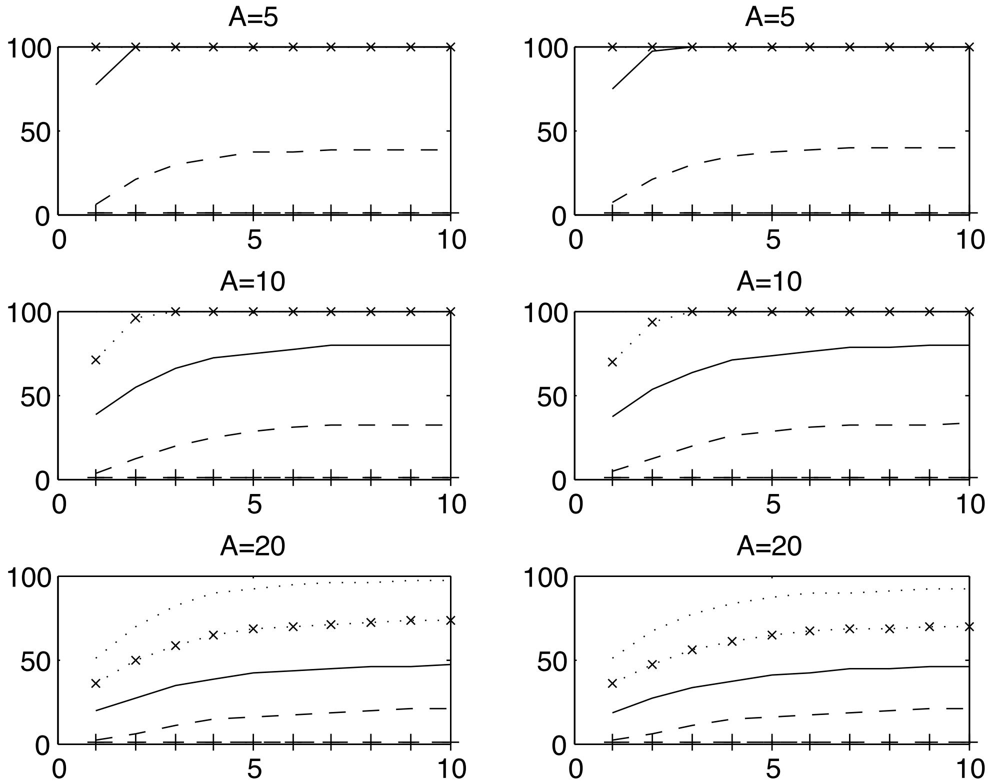

Figure 8: Optimal rebalancing for the UK investor.

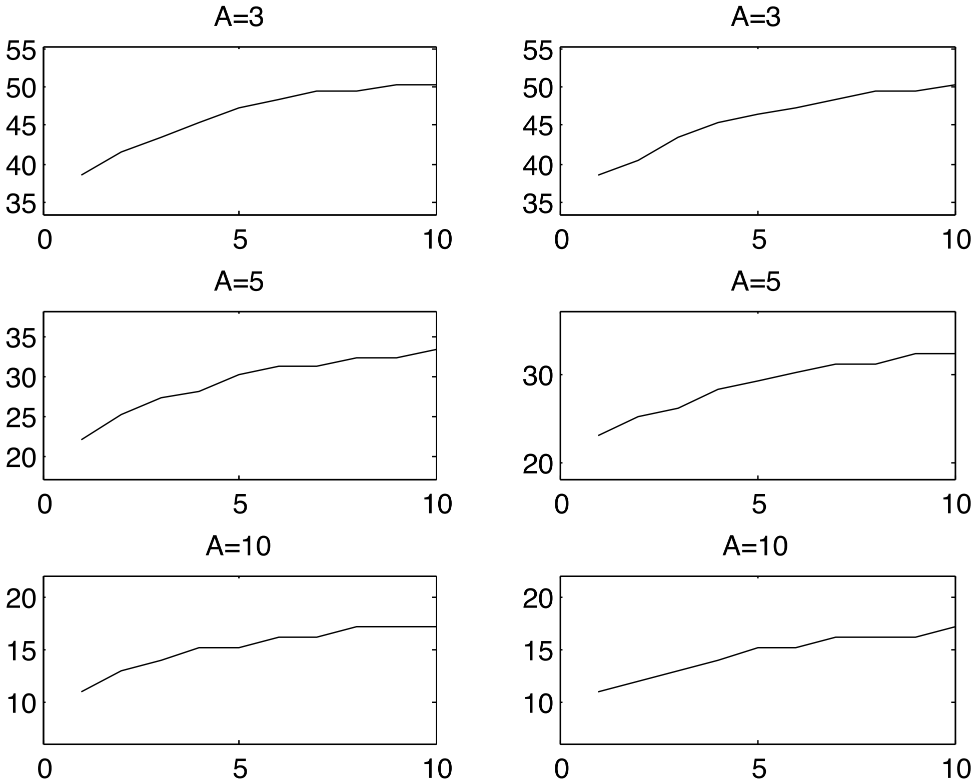

Figure 9: Optimal rebalancing for the Spanish investor.

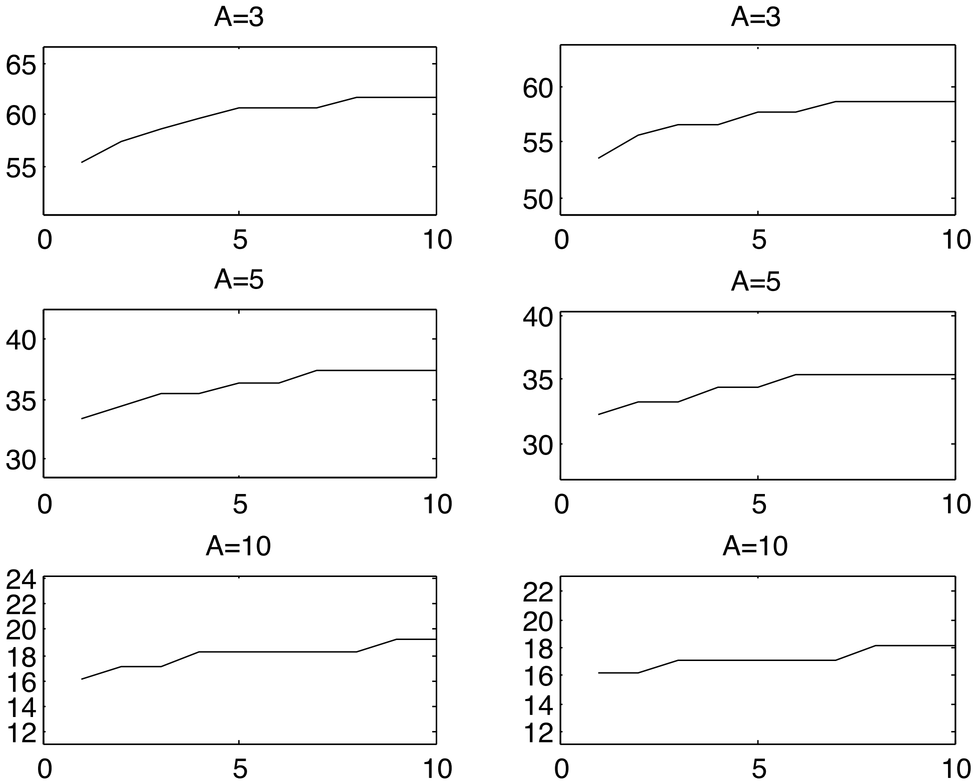

Figure 10: Optimal rebalancing for a German investor.

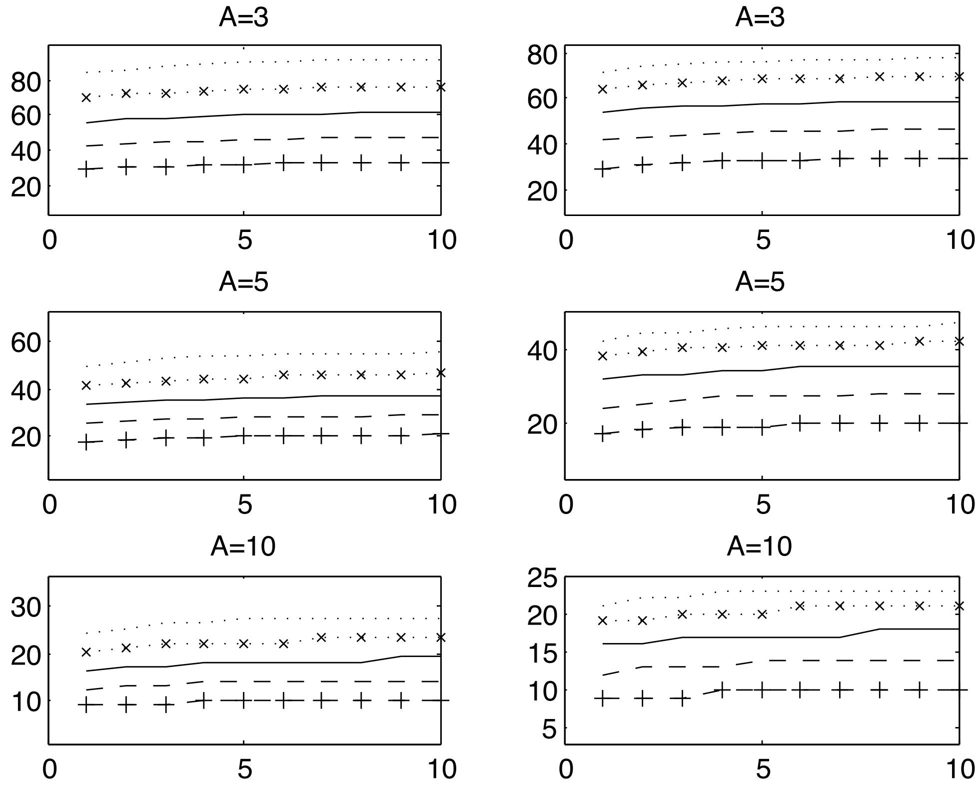

Figure 11: Sensitivity of portfolio weights to the initial dividend-price ratio level for Germany.

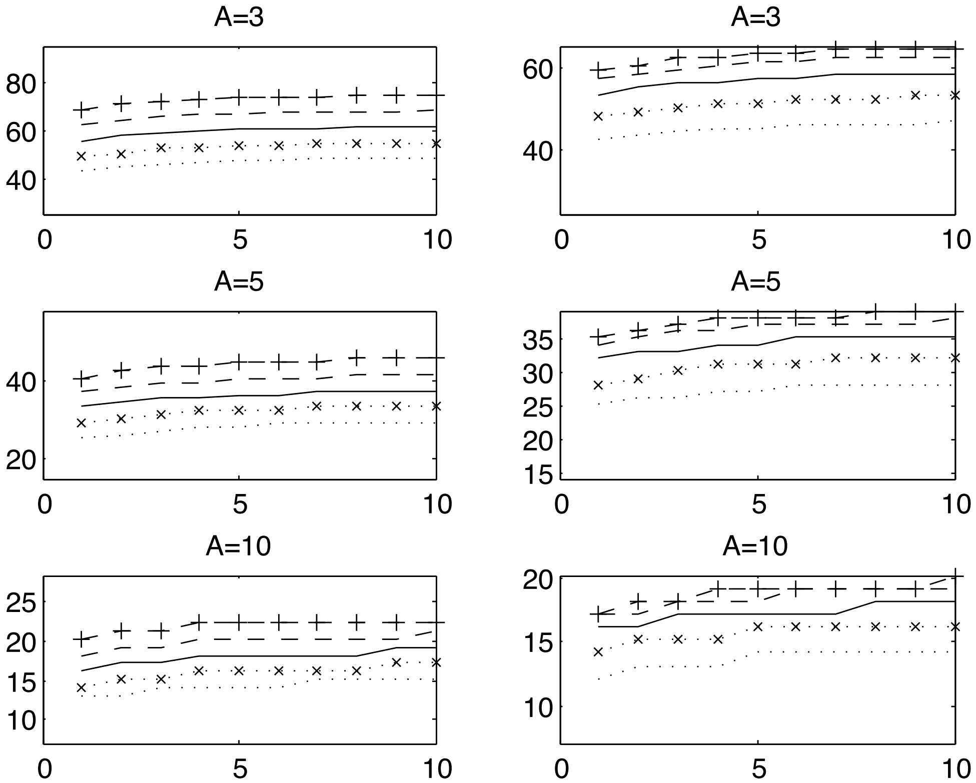

Figure 12: Sensitivity of portfolio weights to the initial inflation level for Germany.