EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1575 / AUGUST 2013

ASSET RETURNS UNDER MODEL UNCERTAINTY

EVIDENCE FROM THE EURO AREA, THE U.S. AND THE U.K.

João Sousa and Ricardo M. Sousa

In 2013 all ECB publications feature a motif taken from the €5 banknote.

MACROPRUDENTIAL RESEARCH NETWORK

NOTE: This Working Paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the authors and do not necessarily reflect those of the ECB.

Macroprudential Research Network

This paper presents research conducted within the Macroprudential Research Network (MaRs). The network is composed of economists from the European System of Central Banks (ESCB), i.e. the national central banks of the 27 European Union (EU) Member States and the European Central Bank. The objective of MaRs is to develop core conceptual frameworks, models and/or tools supporting macro-prudential supervision in the EU.

The research is carried out in three work streams: 1) Macro-financial models linking financial stability and the performance of the economy; 2) Early warning systems and systemic risk indicators; 3) Assessing contagion risks.

MaRs is chaired by Philipp Hartmann (ECB). Paolo Angelini (Banca d'Italia), Laurent Clerc (Banque de France), Carsten Detken (ECB), Simone Manganelli (ECB) and Katerina Šmídková (Czech National Bank) are workstream coordinators. Javier Suarez (Center for Monetary and Financial Studies) and Hans Degryse (Katholieke Universiteit Leuven and Tilburg University) act as external consultants. Fiorella De Fiore (ECB) and Kalin Nikolov (ECB) share responsibility for the MaRs Secretariat.

The refereeing process of this paper has been coordinated by a team composed of Gerhard Rünstler, Kalin Nikolov and Bernd Schwaab (all ECB).

The paper is released in order to make the research of MaRs generally available, in preliminary form, to encourage comments and suggestions prior to final publication. The views expressed in the paper are the ones of the author(s) and do not necessarily reflect those of the ECB or of the ESCB.

João Sousa
Banco de Portugal and Portuguese Ministry of Finance; e-mail: joao.miguel.sousa@bportugal.pt; jmsousa@mf.gov.pt

Ricardo M. Sousa
University of Minho - NIPE; London School of Economics and LSE - Financial Markets Group; e-mail: rjsousa@eeg.uminho.pt; rjsousa@alumni.lse.ac.uk

Address	Kaiserstrasse 29, 60311 Frankfurt am Main, Germany
Postal address	Postfach 16 03 19, 60066 Frankfurt am Main, Germany
Telephone	+49 69 1344 0
Internet	http://www.ecb.europa.eu
Fax	+49 69 1344 6000

ISSN	1725-2806 (online)
EU Catalogue No	QB-AR-13-072-EN-N (online)

This paper can be downloaded without charge from http://www.ecb.europa.eu or from the Social Science Research Network electronic library at http://ssrn.com/abstract_id=2305938.

Information on all of the papers published in the ECB Working Paper Series can be found on the ECB's website, http://www.ecb.europa.eu/pub/scientific/wps/date/html/index.en.html

Abstract

The goal of this paper is to analyze predictability of future asset returns in the context of model uncertainty. Using data for the euro area, the US and the U.K., we show that one can improve the forecasts of stock returns using a model averaging approach, and there is a large amount of model uncertainty. The empirical evidence for the euro area suggests that several macroeconomic, financial and macro-financial variables are consistently among the most prominent determinants of the risk premium. As for the US, only a few predictors play an important role. In the case of the UK, future stock returns are better forecast by financial variables.

Keywords: Stock returns, model uncertainty, Bayesian Model Averaging.

JEL classification: E21, G11, E44.

Non-Technical Summary

The Great Recession drew the attention of many academics, central banks and policy makers to the destabilizing power of financial markets. Financial instability became a prominent concern, thus, leading to an intense research effort to better understand the behaviour of asset markets and the roots and causes of financial imbalances.

From a policy perspective, this line of investigation should help developing tools that can provide early warning indicators and which will contribute to the design of macroprudential policies aimed at preventing that such imbalances arise and mitigating their effects. Those indicators are important as asset fluctuations can degenerate in bubbles or price misalignments that may ultimately lead to financial disruptions. This stream of research is also valuable for monetary policy purposes, given that it helps assessing the uncertainty faced by monetary authorities in relation to financial markets. As such, it can provide evidence regarding where macroprudential policy should act (an important example being the findings that link excessive credit growth to asset price booms).

Our work analyses stock market returns. One difficulty in studying their behaviour is that there is no agreement on how they should be modelled. There are several competing theoretical frameworks, which makes it particularly difficult to choose among the "best" ones. Therefore, model uncertainty highlights the drawbacks of assessing the forecasting properties of a single indicator (e.g. credit growth), as different models would predictably lead to different results. In addition, the models' parameters may change over time, thus, further increasing the difficulty of the exercise.

One way to get around this problem is to use Bayesian Model Averaging (BMA). With this simulation technique, it is possible to deal with both model and parameter uncertainty. The method estimates the weights associated with each indicator/model and allows the construction of a weighted-average model, thereby, optimizing the forecasts made with a large number of models. This way, one can take into account that different predictors or model specifications may be more adequate than others over time.

We apply BMA to data for the euro area, the US and the UK, and show that it helps improving the predictability of stock returns. In particular, the empirical evidence for the euro area suggests that several macroeconomic (the inflation rate, the change in the inflation rate and the commodity price), financial (the lagged returns and the government bond yields) and macro-financial variables (the consumption-(dis)aggregate wealth ratio, the labour income-to-consumption ratio and the stock price index scaled by GDP) are valuable predictors of the future risk premium. In contrast, only a few factors (such as the change in the government bond yield, the change in the inflation rate and the consumption-(dis)aggregate wealth ratio) seem to display predictive content for the future stock returns in the US. As for the UK, the major predictors of the future risk premium are the financial variables,

in particular, the government bond yield, the change in the government bond yield and the dividend yield ratio. Not surprisingly, we find that the degree of model uncertainty is large in all countries.

The robustness of the results is then assessed by comparing the forecasting power of the BMA with an equally-weighted model and the autoregressive and the constant expected returns' benchmark models. We find that the predictive ability of the weighted-average model is stronger in the medium-term, especially, between 3 and 8 quarters ahead. We also confirm its superiority vis-à-vis the other models. Thus, accounting for model uncertainty clearly provides better forecasts of the risk premium.

"... the ECB has no intention of being the prisoner of a single system ... We highly praise robustness. There is no substitute for a comprehensive analysis of the risks to price stability."

- Jean-Claude Trichet, 2005.

"Uncertainty is not just an important feature of the monetary policy landscape; it is the defining characteristic of that landscape."

- Alan Greenspan, 2003.

"Self-confidence is infectious. It can also be dangerous. How often have we drawn false comfort from the apparent confidence of a professional advisor promising certain success only to be disappointed by subsequent performance? Uncertainty pervades almost all public policy questions. Economics and many other disciplines are united by a common need to grapple with complex systems."

- Mervin King, 2010.

1 Introduction

A major source of uncertainty in economics arises from disagreements over theoretical frameworks. Model uncertainty - i.e., the possibility that the theoretical model may be wrong - and not just parameter uncertainty means that models have become probability frameworks (Sims, 2007).

Despite being relevant per se, this question gains a renewed relevance in the context of asset return predictability for two main reasons. First, investors who fail to make asset allocation decisions based on predictions about the future returns may suffer important welfare losses (Campbell and Viceira, 2002). Second, understanding if returns are predictable is crucial for detecting the macroeconomic, financial and macro-financial risks for which investors demand a premium.

The empirical finance literature typically assumes that investors choose among a specific set of variables that exhibit forecasting power for future asset returns. However, given the large number of predictors that have been considered, there is an enormous amount of uncertainty about the variables that define the "true" model governing asset returns. As a result, taking model uncertainty into account when assessing stock return predictability is crucial and extremely useful.

Therefore, the purpose of this paper is to look at the predictability of asset returns through the lenses of the model uncertainty. Specifically, we use Bayesian Model Averaging (BMA) to analyze the role played by model uncertainty in the provision of indicators that track time-variation in future stock returns. This approach averages over all competing models in a given set, with weights given by their posterior probabilities. For any convex scoring rule, the average model outperforms any individual specification chosen via different model selection rules (Madigan and Raftery, 1994).

We investigate the performance of BMA under various settings using simulation approaches and looking at the estimated parameters of the average overall model. That is, we consider prediction when the researcher does not know the true model but has several candidate models. Therefore, we simultaneously deal with both model and parameter uncertainty, which represents a substantial improvement over commonly used methods that only take into account parameter uncertainty.

Using data for the euro area, the US and the UK, we show that one can improve the predictability of stock returns by making use of the BMA approach. The empirical evidence for the euro area suggests that several variables, in particular, macroeconomic (the inflation rate, the change in the inflation rate and the change in the commodity price), financial (the lagged returns and the government bond yields) and macro-financial ones (the consumption-(dis)aggregate wealth ratio, the labour income-to-consumption ratio and the stock price index scaled by GDP) are valuable predictors of the future risk premium. In contrast, only a few factors (such as the change in the government bond yield, the change in the inflation rate and the consumption-(dis)aggregate wealth ratio) seem to display predictive content for future stock returns in the US. As for the UK, the major predictors of future stock returns are financial variables, in particular, the government bond yield, the change in government bond yield and the dividend yield ratio.

These results are confirmed for both the case in which the true model is not in the model set and when the true model is in the model set. We call these frameworks the "agnostic approach" and the selection among models taken from the asset pricing literature.

The degree of model uncertainty is large in all countries: the estimated cumulative probability of the 10 "best" or "best-performing" models is around 46%, in the euro area, between 58% and 61% for the US, and lies in the interval 46%-61% in the case of the UK.

The robustness of the results is then assessed along several lines. First, we compare predictability at short run horizons vis-a-vis long run horizons. In principle, BMA may work better at shorter horizons, as model uncertainty is more important when less data is employed, while at longer horizons, averaging introduces noise in the predictive model. Another possible reason is that asset returns may be more accurately predicted in the short run due to phenomena such as momentum (Torous et al., 2005; Ang and Bekaert, 2007; Gomes, 2007). In contrast, the recent literature that developed economically motivated variables to capture time-variation in risk premium has shown that the underlying models exhibit stronger forecasting power at horizons from 3 to 8 quarters. Therefore, this would justify why BMA could perform better at longer horizons. Second, we consider a wide range of model priors and

Zellner's $g-$priors. Third, we compute recursive forecasts and provide sub-sample analysis. In this context, Chapman and Yan (2002) suggest that sub-samples, rather than the full sample, are more informative about the predictive regression parameters. Finally, we compare the predictability of the weighted-average model with the autoregressive and the constant expected returns' benchmark models, and also generate out-of-sample forecasts.

We show that the weighted-average model performs better at longer horizons, especially, between 3 and 8 quarters ahead. Interestingly, the predictive ability of the weighted-average model that is built with the probabilities estimated using BMA is stronger than the equally weighted-average model. In addition, the superiority of accounting for model uncertainty is clear when compared with the benchmark specifications, as suggested by the nested forecasts.

The rest of the paper is organized as follows. Section 2 briefly reviews the related literature, while Section 3 describes the econometric methodology. Section 4 presents the data and discusses the empirical results. Section 5 provides the robustness analysis. Section 6 concludes with the main findings.

2 A Brief Review of the Literature

The risk premium is generally considered as reflecting the ability of an asset to insure against consumption fluctuations. However, the empirical evidence has shown that the covariance of returns across portfolios and contemporaneous consumption growth is not sufficient to justify the differences in expected returns (Mankiw and Shapiro, 1986; Breeden et al., 1989; Campbell, 1996; Cochrane, 1996). On the one hand, the inefficiencies of the financial markets (Fama (1970, 1991, 1998), Fama and French (1996), Farmer and Lo (1999)), and the rational response of the agents to time-varying investment opportunities that is driven by the variation in the risk aversion (Sundaresan (1989), Constantinides (1990), Campbell and Cochrane (1999)) or in the joint distribution of consumption and asset returns (Duffee, 2005) can help justifying why expected excess returns on assets appear to vary with the business cycle. On the other hand, several macroeconomic variables, financial indicators and economically motivated macro-financial variables have also been referred as incorporating important informational content about the future business conditions and as being particularly useful at capturing time-variation in expected returns. Table 1 provides a summary of such variables.

[PLACE TABLE 1 HERE.]

Despite the abovementioned advances in the literature of asset pricing, the identification of the economic sources of risks remains an important issue. In recognition of the uncertainty associated with a specific predictive model, Kandell and Stambaugh (1996), Barberis (2000) and Xia (2001) use Bayesian methods to account for parameter uncertainty and find that predictability can be significantly improved. However, there is no consensus on what the true predictors are and what the exact forecasting model should be. Moreover, even if the model is correctly specified, it is not trivial that more structure can improve its performance. In fact, model uncertainty can outweigh parameter uncertainty. For instance, Avramov (2002) shows that this is so in the context of portfolio selection, while Pastor and Stambaugh (2000) reach the same conclusion in the case of portfolio constraints. Cremers (2002) demonstrates that the existence of predictability is reinforced for both skeptical and confident investors using Bayesian model averaging (BMA) and concludes that it provides better in-sample model fitting and out-of-sample forecasting ability for predictive models.[1]

The concept of model uncertainty which has received a lot of interest in the statistical literature refers to uncertainty about the number and the nature of the covariates to include in the model. It can be explicitly assessed by means of Bayesian statistical techniques, in particular, the Bayesian model averaging (BMA) methodology. In fact, it proposes averaging the parameter values over all (relevant) alternative models using posterior model probabilities as the corresponding weights for evaluating the relative importance of the different predictors (Raftery, 1995).

In addition, model uncertainty is a prominent feature of the literature on asset return predictability. Most studies concentrate on specific transmission mechanisms between macroeconomic developments and asset price dynamics or between financial valuation ratios and stock return fluctuations, but they do not tend to be mutually exclusive. The fact that the future risk premium may be explained by different theoretical models implies that the choice of a single specification underestimates the degree of uncertainty of the estimated parameters as it ignores model uncertainty. Consequently, the main goal of the current work is to revisit the various models of asset return predictability and to explicitly account for model uncertainty while predicting stock returns.

3 Accounting for Model Uncertainty: Bayesian Model Averaging

The predictive regression typically considered in the empirical finance literature is as follows

$$r_t = \alpha + X_{t-1}\beta + \epsilon_t \tag{1}$$

[1] For an interesting application of BMA to the analysis of the determinants of the horizontal spillovers from FDI, see Havránek and Irsová (2011).

where r_t is the asset return, X_{t-1} is a $K \times 1$ vector of K predictors, α is a constant, and ϵ_t denotes the disturbance term or forecasting error.

The basic step for building a linear predictive regression is to choose among a group of candidate predictors, $X = \{1, x_1, x_2, ..., x_K\}$, and decide which of these variables should enter equation (1). The goal is to find the "best" model $X^* \subset X$ for the linear predictive regression and to proceed as if X^* is the true model.[2]

While the abovementioned procedure is easy to implement, it ignores that uncertainty about the model may be a relevant feature, in particular, when there is limited data availability. Bayesian model averaging helps directly tackling this issue. The basic idea is to construct an overall model which is a weighted-average of the individual models in the model set, where the weights are given by the posterior probabilities (Raftery et al., 1997).[3]

Suppose that we observe data $\Omega = \{r_t, X_t\}$ generated from a set of competing models. For K potential predictors, there are 2^K competing models. Let Θ be the quantity of interest. Then, the posterior distribution of Θ is

$$\Pr(\Theta \mid \Omega) = \sum_{k=1}^{2^K} \Pr(\Theta \mid M_k, \Omega) \Pr(M_k \mid \Omega). \tag{2}$$

The posterior probability for model M_k is given by the Bayes rule

$$\Pr(M_k \mid \Omega) = \frac{\Pr(\Omega \mid M_k) \Pr(M_k)}{\sum_{l=1}^{2^K} \Pr(\Omega \mid M_l, \Omega) \Pr(M_l)}, \tag{3}$$

where

$$\Pr(\Omega \mid M_k) = \int \Pr(\Omega \mid \theta_k, M_k) \Pr(\theta_k \mid M_k) d\theta_k \tag{4}$$

is the integrated likelihood of model M_k, θ_k is the vector of parameters of model M_k, $Pr(\theta_k \mid M_k)$ is the prior density of θ_k under model M_k, and $\Pr(\Omega \mid \theta_k, M_k)$ is the likelihood of model M_k.The prior probability that model M_k is the true model for each competing model, $Pr(M_k), k = 1, 2...2^K$, is exogenously specified based on prior information. All probabilities are conditional on M, i.e., the set of all models under consideration (the so-called M-closed perspective).

When K is large, it is unfeasible to average over 2^K models, and there are two approaches to handle

[2] The "best" model may be the one with the best in-sample fitting in case one is interested in recovering historical data dynamics. It may also be the one with the best mean out-of-sample forecasting properties when the researcher is interested in the model's predictive power.

[3] Despite its strength in handling model uncertainty, BMA has some potential problems. First, it assumes that the true model lies in the model set and, as a result, the consequence of omitting some true predictors is unknown. Second, it makes the assumption that $\epsilon_t \sim N(0, \sigma^2)$, which may not be realistic.

this problem. The first approach consists in the use of the Occam's Window to filter out: (i) models with more complicated structure but smaller posterior probability compared to relatively simpler models; and (ii) models with very small posterior probability. The second approach is the Markov Chain Monte Carlo Model Composition (MC^3) method, which consists of four steps: 1) start with a model, M_k; 2) look at its neighbourhood models, $M_{k'}$, with some transition density $q(M_k \longrightarrow M_{k'})$ along the Markov chain; 3) switch to model $M_{k'}$ with probability $\min\left\{1, \frac{\Pr(M_{k'}|\Omega)}{\Pr(M_k|\Omega)}\right\}$, otherwise, stay at model M_k; and 4) average over the entire Markov chain.[4]

In order to be able to obtain the BMA posterior distribution, one needs to specify three components: the model prior, $\Pr(M_k)$, the model likelihood for a given model, $Pr(\Omega \mid \theta_k, M_k)$ and the prior distribution of the parameters given a model, $Pr(\theta_k \mid M_k)$.

The prior probability on model M_k can be specified as

$$\Pr(M_k) = \prod_{j=1}^{K} p_j^{\pi_{k_j}} (1 - p_j)^{1-\pi_{k_j}}, \qquad k = 1, 2, ..., 2^K, \tag{5}$$

where $p_j \in [0, 1]$ is the prior probability that $\beta_j \neq 0$ in a regression model and represents the researcher's prior confidence in the predictive power of the regressors, and π_{k_j} is an indicator of whether variable j is included in model M_k.

We take an "agnostic" position in that we assume that we do not have any special information on the relative predictive power of individual predictors. In fact, a potential drawback of the choice of the model's space prior is that it assigns a relatively large prior probability to models that may be considered "highly parameterized". In our paper, this is less of a problem given that: 1) in the M-open perspective, we only consider univariate forecasting regressions; and 2) in the M-closed perspective, the number of predictors included in the different asset pricing models is fairly similar. In this context, the g-priors of Zellner (1986) have been advocated for BMA (Fernandez et al., 2001a). We assume that the disturbance term, ϵ_t, in the forecasting regression follows a normal distribution

$$\epsilon_t \sim N(0, \sigma^2), \tag{6}$$

and the parameter priors are given by

$$\beta \mid \sigma^2 \sim N(\mu, \sigma^2 V), \tag{7}$$

$$\frac{\nu\lambda}{\sigma^2} \sim \chi^2_\nu \tag{8}$$

[4] The MC^3 approach has two major advantages. First, it averages over all models according to their posterior probabilities. Second, it simultaneously handles model and parameter uncertainty.

where $\mu = (\overset{\curlywedge}{\beta}_0, 0, 0, ..., 0)$ and $\overset{\curlywedge}{\beta}_0$ is the OLS estimate of β_0, $V = diag(s_r^2, gs_1^{-2}, ..., gs_k^{-2})$, with g being the parameter for the standard Zellner's g-priors and s^2 the sample variance, and ν and λ are hyperparameters. These priors are determined by the choice of g and, following Fernandez et al. (2001a), we use a g-prior with $g = \frac{1}{n}$, which consists of assigning the same amount of information to the conditional prior of β as the one that is contained in one observation. Interestingly, Ley and Steel (2009) propose the use of an hyper-prior on model size, which reflects the robustness of the inference when applying BMA. Ley and Steel (2011a, 2011b) analyze the effect of the prior on the posterior probability of including regressors and predictive performance. The authors combine a Binomial-Beta prior on the model size with a g-prior on the coefficients of each model, and propose a benchmark Beta prior that leads to consistent model selection. They also highlight that this technique does not necessarily perform better than a model with fixed values of g that are chosen to optimize a given criterion, because "optimal" values are not known. However, poor performance can be originated by getting those values wrong. In the current work, we rely on as little prior information as possible and this is motivated by the lack of consensus in the literature about what the major determinants of the risk premium are. In this context, our approach is in line with the works of Avramov (2002) and Cremers (2002) and is close in spirit to the "unit information priors" (Kass and Wasserman, 1995). The Bayes factor comparing two models, $B_{kl} = \frac{\Pr(\Theta|M_k)}{\Pr(\Theta|M_l)}$, can be approximated using the Bayesian information criterion (Schwarz, 1978) as

$$-2 \log B_{kl} \approx BIC_k - BIC_l, \qquad (9)$$

where BIC_i is the Bayesian Information Criterion of model i.

Another computational problem is caused by the cardinality of the model space, which can lead to the intractability of the expression (2). The Occam's window approach and the Markov Chain Monte Carlo Model Composite (MC^3) method are particularly helpful in setting bounds to the number of models (Raftery, 1995; Fernandez et al., 2001b; Koop, 2003).

4 Can BMA Improve Stock Return Predictability?

4.1 Data

This Section provides a summary of the data used in the estimations. A detailed version can be found in the Appendix. We consider a set of macroeconomic, financial and macro-financial variables, which are selected in accordance with the previous literature and the data availability. Among the set of predictors considered in the BMA analysis, we include:

- *Macroeconomic variables*: consumption growth, consumption growth over the last 12 quarters, output gap, inflation, change in inflation, change in the interest rate, growth rate of the monetary aggregate, growth rate of the housing price index, change in the real effective exchange rate, growth rate of the commodity price index, change in the unemployment rate, and growth rate of credit.

- *Financial variables*: lagged stock returns, real government bond yield rate, change in the real government bond yield rate, and dividend-yield ratio.

- *Macro-financial variables*: consumption-wealth ratio, change in the consumption-wealth ratio, consumption-(dis)aggregate wealth ratio, change in the consumption-(dis)aggregate wealth ratio, residential wealth-to-income ratio, aggregate wealth-to-income ratio, ratio of the stock price index scaled by the real GDP, ratio of durable to nondurable consumption, and leverage ratio of brokers and dealers' institutions.

For the euro area, the data sources are the European Central Bank (ECB), the International Financial Statistics (IFS) of the International Monetary Fund (IMF), and the Bank of International Settlements (BIS). In the case of the US, the data comes from the Flow of Funds Accounts (FoF) of the Board of Governors of the Federal Reserve System, the Bureau of Labor Statistics (BLS), the US Census, and the BIS. Finally, for the UK, the data sources are the Office for National Statistics (ONS), the Datastream, the Nationwide Building Society, the Halifax Plc, and the BIS. The data is available for: 1980:1-2007:4, in the case of the euro area; 1967:2-2008:4, in the US; and 1975:1-2007:4, in the UK.

4.2 An Agnostic Approach: The $\mathcal{M}$-Open Perspective

We start by considering the role of BMA in the context of the M-open perspective. First, we adopt an "ad-hoc" selection of potential determinants of asset returns. These include: (i) the lag of consumption growth (ΔC_{t-1}); (ii) the growth of consumption over the last 12 quarters ($\Delta C_{t-1,t-12}$); (iii) the lag of asset returns (r_{t-1}); (iv) the lag of the real government bond yield ($bond_{t-1}$); (v) the change in the lag of the real government bond yield ($\Delta bond_{t-1}$); (vi) the lag of the output gap (og_{t-1}); (vii) the lag of inflation (π_{t-1}); $(viii)$ the change in the lag of inflation ($\Delta\pi_{t-1}$); (ix) the lag of the change in the short-term interest rate (Δi_{t-1}); (x) the lag of the growth rate of the monetary aggregate (Δm_{t-1});(xi) the lag in the growth rate of the housing price index (Δhp_{t-1}); (xii) the lag in the change of the real effective exchange rate (Δe_{t-1}); $(xiii)$ the lag in the growth rate of the commodity price index (Δcp_{t-1}); (xiv) the lag of the consumption-(dis)aggregate wealth ratio ($cday_{t-1}, cay_{t-1}$); (xv) the change in the lag of the consumption-(dis)aggregate wealth ratio ($\Delta cday_{t-1}, \Delta cay_{t-1}$); (xvi) the lag of the labour income-consumption ratio (lc_{t-1}); ($xvii$) the lag in the residential wealth-to-income ratio (rwy_{t-1}); ($xviii$) the lag in the aggregate wealth-to-income ratio (wy_{t-1}); (xix) the lag in the dividend-yield ratio ($divyld_{t-1}$); (xx) the lag in the ratio of the stock price index scaled by the real GDP ($spgdp_{t-1}$); (xxi) the lag of the change in the unemployment rate (Δu_{t-1}); ($xxii$) the lag of the ratio of durable to nondurable consumption (φ_{t-1}); ($xxiii$) the lag of the growth rate of credit ($\Delta cred_{t-1}$); and ($xxiv$) the lag of the leverage ratio of the brokers and dealers' institutions ($SBRDLR_{t-1}$, only for US). In this section we do not impose a specific structure in the model, so that the algorithm looks for all possible combinations of regressors and the technique estimates their posterior probabilities. Finally, we consider "in-sample" one-period ahead forecasting regressions.

4.2.1 Euro area

The evidence for the euro area can be found in Tables 2 and 3. Table 2 provides a summary of the results for the 10 "best" models (i.e. the ones with the highest posterior probability) using the Occam's Window approach. In addition, the number of selected models, the cumulative posterior probability associated with the 10 "best" models, the posterior inclusion probability, and the mean and the standard deviation of the posterior distribution of each parameter, the number of variables included in each model and the corresponding adjusted-R^2 statistics are also reported. Table 3 describes the 10 "top-performing" specifications when we use the Monte Carlo Markov Chain Model Composition (MC^3) method and, for simplicity, the models are defined by inclusion (X) or exclusion (blank) of the specific variable. It also provides information about the number of the selected models, the cumulative posterior probability associated with the 10 "best" models, the posterior inclusion probability, and the number of variables included in each model.

As shown in Table 2, several variables seem to be valuable predictors of stock returns in the euro area. In particular, the posterior probability of inclusion is large (that is, above 25%) in the case of the lag of asset returns, the government bond yield, the inflation rate, the change in the inflation rate, the commodity price, the consumption-(dis)aggregate wealth ratio, the labour income-to-consumption ratio, and the stock price index scaled by GDP. In the case of other variables, such as the growth rate of the monetary aggregate, the housing price index, the change in the real effective exchange rate, the real estate wealth-to-income ratio, the aggregate wealth-to-income ratio and the growth rate of credit, the empirical results do not support their usefulness in predicting stock returns.

Among the selection of 64 models, the cumulative posterior probability of the 10 best-performing specifications is high (about 46.7%). Similarly, the adjusted-R^2 statistics associated to each model are also large, ranging between 24.8% and 40.5%. Interestingly, the majority of the 10 "best" models include a relatively large number of predictors, which highlights the predictability power of several macroeconomic, financial and macro-financial variables. Similarly, the coefficients associated to the predictors do not change substantially among the various specifications, which shows that they are consistently important drivers of variation in the future risk premium.

As for Table 3, it confirms the previous results. In particular, among the set of potential predictors, the posterior probability of inclusion is large for financial (the lag of asset returns, the government bond yield and the dividend yield ratio) and macroeconomic (the output gap and the commodity price), and for proxies that capture time-variation in expected returns (the stock price index scaled by GDP). The cumulative posterior probability of the top 10 models is also substantial (46.5%) from a total of 614 selected models, and their posterior probability ranges between 2.5% and 10.5%. In fact, the model with the highest posterior probability includes four predictors: the lag of stock returns, the government bond yield, the commodity price and the stock price scaled by GDP.

[PLACE TABLE 2 HERE.]

[PLACE TABLE 3 HERE.]

4.2.2 US

Tables 4 and 5 summarize the empirical evidence for the US. In contrast with the euro area, only a few variables seem to display predictive. These are the change in the government bond yield, the change in the inflation rate and the consumption-(dis)aggregate wealth ratio, which all have a posterior probability of inclusion above 25%. The aggregate wealth-to-income ratio also exhibits a posterior probability of inclusion above 10%. The model with the highest posterior probability (13.5%) includes both the change in government bond yield and the consumption-(dis)aggregate wealth ratio. The coefficients associated with the predictors are also in line with the theory: (*i*) an increase in the risk premium of government bonds forecasts a fall in stock returns, reflecting the flight to quality, i.e. a reallocation of wealth towards risk-free assets; (*ii*) an acceleration of inflation predicts an increase in the risk premium as it tends to be associated with higher economic risks. In addition: (*i*) as in Sousa (2010a), a rise in the consumption-(dis)aggregate wealth ratio forecasts an increase in stock returns, reflecting the increase in the wealth composition risk; (*ii*) when the aggregate wealth-to-income ratio increases, agents demand a lower stock return as they become less exposed to idiosyncratic risk (in line with the work of Sousa (2010b)). The cumulative posterior probability of the 10 "best" models reaches 61.4% from a total of 37 models selected by the Occam's Window method. The models explain between 4.3% and 10.4% of the next quarter stock returns as reflected by the adjusted-R^2 statistics. Interestingly, the constant expected returns benchmark model has the sixth highest posterior probability (3.1%), which suggests that some historical periods have been characterized by constancy in the risk premium.

Table 5 provides similar results in that the change in the government bond yield and the consumption-(dis)aggregate wealth remain as the most important predictors of the stock returns. In particular, the model with *cday* has the highest posterior probability 19.6%, reflecting the role played by the wealth composition risk. Again, the constant expected return benchmark model is relevant with a posterior probability of 13%, that is, the second highest among all models. The cumulative posterior probability of the top 10 models is also large (58.3%) from a total of 652 selected models. Their posterior probability ranges between 1% and 19.6%.

[PLACE TABLE 4 HERE.]

[PLACE TABLE 5 HERE.]

4.2.3 UK

The results for the UK are presented in Tables 6 and 7. In sharp contrast with the euro area and the US, the empirical evidence suggests that the major predictors of future stock returns in the UK are financial variables, in particular, the government bond yield, the change in the government bond yield and the dividend yield ratio. Only the aggregate wealth-to-income ratio seems to be another important predictor. In fact, the posterior probability of inclusion of these variables lie well above 25%. As a result, models with macroeconomic variables and/or empirical proxies developed to capture time-variation in risk premium do not seem to be relevant in explaining one quarter-ahead stock returns. The posterior probability associated with the 10 "best" specifications ranges between 2.7% and 20.1% and, in accordance with the findings for the euro area, a reasonable number of variables seem to consistently guide future returns. These models have a cumulative posterior probability of 60.9% from a total of 37 models selected by the Occam's Window method. The adjusted-R^2 statistics associated to best-performing models are also large and lie between 7.5% and 26.3%. Similarly, the magnitude of the coefficients associated to the various predictors does not change substantially between specifications.

Table 7 corroborates the previous findings: the posterior probability of inclusion of financial variables such as the government bond yield, the change in the government bond yield and the dividend yield ratio lie above or are close to 25%. In addition, some predictors capturing investors' expectations, such as the consumption-(dis)aggregate wealth ratio, the aggregate wealth-to-income ratio and the stock price index scaled by GDP have a posterior probability of inclusion above 10%. Among 653 models selected by the MC^3 method, the 10 "best" models represent a cumulative posterior probability of 45.9%. The model with the highest probability (10.6%) only includes one predictor (the dividend yield ratio), while the one with the lowest posterior probability (1.4%) has five regressors (the constant, the government bond yield, the change in the government bond yield, the wealth-to-income ratio and the dividend yield ratio).

[PLACE TABLE 6 HERE.]

[PLACE TABLE 7 HERE.]

4.3 A Focus on the Empirical Finance Literature: The $\mathcal{M}$-Closed Perspective

A more interesting case is when the true model is not in the model set (the M-closed perspective), that is, we assess the relevance of variable exclusion. In practice, we restrict the attention to a set of models developed in the empirical finance literature. These are based on the works of: (*i*) Chen et al. (1986); (*ii*) Campbell (1987) and Ferson (1990); (*iii*) Harvey (1989); (*iv*) Ferson and Harvey (1991); (*v*) Ferson and Harvey (1993); (*vi*) Whitelaw (1994), Pontiff and Schall (1998), and Ferson and Harvey (1999); (*vii*) Pesaran and Timmerman (1995); (*viii*) Julliard and Sousa (2007a); (*ix*) Julliard and Sousa (2007b); (*x*) Bossaerts and Hillion (1999); (*xi*) Rubinstein (1976) and Breeden (1979), that is, the Consumption-Capital Asset Pricing Model (C-CAPM); (*xii*) Sousa (2010b); (*xiii*) Lettau and Ludvigson (2001); (*xiv*) Sousa (2010a); (*xv*) Parker and Julliard (2005); (*xvi*) Lustig and van Nieuwerburgh (2005); (*xvii*) Santos and Veronesi (2006); (*xviii*) Yogo (2006) and Piazzesi et al. (2007); and (*xix*) and Adrian et al. (2010). The first 17 models are considered for both the euro area, the US and the UK. In addition, model (*xviii*) is taken into account for the US and the UK, while model (*xix*) is only analyzed in the case of the US. This is explained by the lack of data.

Using the restricted set of models, we apply BMA to estimate the posterior probability associated to each of them. For illustration, we present in Tables 8 and 9 an overview of the variables included in the 19 models taken from the empirical finance literature. One can see that each model focuses on a particular number of predictive variables, which are linked to stock returns in the context of forecasting. In addition, it is clear that there is no consensus regarding the appropriate model, as the set of predictors differs from one model to another. As a result, there is a large amount of uncertainty not only regarding the "true" model, but also in terms of the variables that explain risk premium.

[PLACE TABLE 8 HERE.]

[PLACE TABLE 9 HERE.]

4.3.1 Euro area

We start by analyzing the role of BMA in the context of the M-open perspective for the euro area. Table 10 summarizes the results using the Monte Carlo Markov Chain Model Composite (MC^3) method. It reports the root mean-squared error (RMSE), the ratio of the RMSE of the selected model and the RMSE of the constant expected return benchmark model, the ratio of the RMSE of the selected model and the RMSE of the autoregressive benchmark model, the adjusted-R^2 statistic of the selected model and its posterior probability.

The nested forecast comparisons show that, in general, the models perform better than the benchmark models. This is particularly important when the benchmark model is the constant expected returns benchmark, and, therefore, it supports the existence of time-variation in expected returns. In fact, the evidence is a bit mixed in what concerns the autoregressive benchmark model. In addition, the posterior probability associated with the models ranges between 0.0%-0.1% (Lustig and van Nieuwerburgh, 2005; Whitelaw, 1994; Pontiff and Schall, 1998; Ferson and Harvey, 1999; Pesaran and Timmerman, 1995; Sousa, 2010b) and 35.3% (Chen et al., 1986). This finding suggests that the lag of stock returns, the government bond yield, the change in the government bond yield, the output gap, the inflation rate and the growth in inflation are among the most prominent predictors of stock returns in the euro area. Consequently, financial variables and, above all, macroeconomic variables seem to play a major relevance in forecasting asset returns. In fact, this is also reflected in the adjusted-R^2 statistics of models (*i*), (*ii*), (*iii*), (*iv*), (*v*) and (*x*) which are the highest among all models, as they explain between 12.7% and 17.4% of the variation in the real stock returns of the next quarter. The C-CAPM model performs badly: both the posterior probability and the adjusted-R^2 statistic of the model are negligible. As for the models that include empirical proxies capturing time-variation in risk premium, they are typically associated with low posterior probabilities (in general, below 1%) and also low adjusted-R^2 statistics.

Table 11 provides the results for the weighted-average model. Specifically, it displays information about the root mean-squared error and the nested forecast comparisons. We consider four situations: (a) the equally weighted-average model using the Occam's Window approach; (b) the weighted-average model with the posterior probabilities computed with the Occam's Window approach; (c) the equally weighted-average model using the Monte Carlo Markov Chain Model Composite (MC^3) method; and (d) the weighted-average model with the posterior probabilities computed with the Monte Carlo Markov Chain Model Composite (MC^3) method.

The improvement in terms of forecasting ability of the weighted-average model is substantial: the RMSE of the average model is clearly below the one found for individual models; and the nested forecast comparisons show that the weighted-average model also outperforms the constant expected return and the autoregressive benchmark models. Therefore, this suggests that one can obtain better forecasts for future stock returns in the euro area while accounting for model uncertainty.

[PLACE TABLE 10 HERE.]

[PLACE TABLE 11 HERE.]

4.3.2 US

The empirical evidence concerning the US can be found in Table 12. All models improve upon the benchmark models as the ratio of the RMSEs suggest. In contrast with the euro area, it can be seen that the posterior probability associated with the models that capture expectations about future returns is the largest. This is particularly the case for the models developed by: (i) Sousa (2010a), with a probability of 51.7%; (ii) Julliard and Sousa (2007b), with a probability of 26.2%; (iii) Yogo (2006) and Piazzesi et al. (2007), with a probability of 5.4%; and (iv) Lettau and Ludvigson, with a probability of 4.5%. Therefore, the models which focus on the wealth composition risk, the long-run risk and the willingness to smooth consumption, and the composition risk are among the ones that better forecast time-variation in expected returns. The adjusted-R^2 statistics are also relevant. For instance, the consumption-(dis)aggregate wealth ratio explains 5% of the stock return in the next quarter.

Table 13 displays the information about the weighted-average model. Once again, the gains in terms of predictive ability are important: not only the weighted-average model outperforms the benchmark model, but it also has a RMSE that is smaller than any individual model.

[PLACE TABLE 12 HERE.]

[PLACE TABLE 13 HERE.]

4.3.3 UK

The empirical findings for the UK are similar to the ones described for the US and are displayed in Tables 14 and 15. Table 14 shows that the models with the largest posterior probability are the ones based on the works of Julliard and Sousa (2007b), Sousa (2010b), Yogo (2006) and Piazzesi et al. (2007), Julliard and Sousa (2007a) and Lettau and Ludvigson (2001). In fact, for these models, the posterior probability ranges between 8.7% and 25.8%. This suggests models that include macro-financially motivated variables developed to track changes in investors' expectations about future returns are more likely to forecast risk premium. Interestingly, the C-CAPM model has a posterior probability of 14.5%, that is, one of the highest among the selected models. However, when we look at the adjusted-R^2 statistics, it seems that models that include financial variables perform better. For instance, the models of Pesaran and Timmerman (1995), Bossaerts and Hillion (1999), Whitelaw (1994), Pontiff and Schall (1998) and Ferson and Harvey (1999), and Harvey (1989) explain 7.8%, 7.7%, 6.3%, and 5.4% of the variation of stock returns in the next quarter, respectively. Taken together, these findings show that while macro-financial models are more "likely" to be included in the "true" models, the models that only consider financial variables tend to have a higher predictive ability despite their lower posterior probability.

Table 15 corroborates the previous findings for the euro area and the US: the BMA helps improving the forecasting ability for stock returns as the weighted-average model delivers a much lower RMSE than the models taken individually. In addition, the weighted-average model also outperforms the benchmark models in terms of predictive properties.

[PLACE TABLE 14 HERE.]

[PLACE TABLE 15 HERE.]

5 Robustness Analysis

The results presented so far clearly show that the BMA improves the predictability of the future returns. In particular, the weighted-average model delivers superior forecasting power at the one quarter-ahead horizon.

We now assess the robustness of the previous findings in several directions, namely, in terms of: (*i*) different model priors; (*ii*) various Zellner's g-priors; (*iii*) long-run predictability; (*iv*) recursive forecasts; and (*v*) "out-of-sample" predictability.

5.1 Different Selection of Model Priors

We start by looking at a different selection of model priors. In particular, we consider: (*i*) a "fixed" model prior, which sets a fixed common prior inclusion probability for each regressor (Sala-i-Martin et al., 2004); (*ii*) a "random" model prior, i.e. the "random theta" prior by Ley and Steel (2008), who suggest a binomial-beta hyperprior on the ex-ante inclusion probability and give the same weight to models of different size; (*iii*) an "uniform" model prior, which means that models with average size ($K/2$) get more weight than, for instance, very parsimonious specifications; and (*iv*) a custom prior inclusion probability ("custom") model prior.

5.1.1 Euro area

Table 16 presents the evidence for the euro area. In particular, it summarizes, for each regressor, the posterior inclusion probability (PIP) associated with the abomentioned model priors. It reveals that no explanatory variable has a PIP equal to 1 and, therefore, there is substantial model uncertainty in forecasting future stock returns. Moreover and with the exception of the case where we employ the "random" model prior, all explanatory variables have a PIP that is larger than 0.1. As a result, they help predicting the future risk premium. In addition, the average number of regressors ranges from 2 to 7, suggesting that several variables display predictive content. In particular, the lagged asset returns (0.88-0.99), the stock price index scaled by GDP (0.17-0.74), the inflation rate (0.04-0.41), the commodity price (0.34-0.86), the change in the inflation rate (0.04-0.33), the labour income-to-consumption ratio (0.08-0.55), the wealth composition risk (0.03-0.39) and the dividend yield ratio (0.17-0.30) rank among the variables with the largest PIP.

[PLACE TABLE 16 HERE.]

In Figure 1, we also present graphically the results of the BMA estimation using an "uniform" model prior. The columns refer to individual models and these include the regressors for which the corresponding cells are not blank. The blue colour of the cell can be interpreted as a positive sign of the estimated coefficient associated with the variable included in the model. The red color denotes a negative sign. The horizontal axis displays the posterior model probabilities and a wider column is linked with a better model's fit. For instance, the best model only includes one explanatory variable (the lagged asset return) and its posterior probability is about 23%. Models with the commodity price, the output gap, the stock price index scaled by GDP and the dividend yield also display the largest posterior probabilities. Finally, one can see that the sign of the various coefficients is consistent for all regressors, thereby, displaying stability.

[PLACE FIGURE 1 HERE.]

5.1.2 US

In Table 17, we provide the US evidence on the PIP associated with the different model priors for each regressor. As in the case of the euro area, there is an important amount of uncertainty in terms of the predictive variables for future stock returns. However, there is also a reasonably large number of regressors that have a PIP larger than 0.1, such as the consumption-(dis)aggregate wealth ratio (0.15-0.68), the change in the real government bond yield (0.07-0.55), the change in the inflation rate (0.03-0.37), the wealth-to-income ratio (0.01-0.18), the stock price index scaled by GDP (0.01-0.19), the labour income-to-consumption ratio (0.01-0.19) and the dividend yield (0.01-0.15). The average number of regressors included in the visited models ranges between 0.4 and 4, highlighting the role played by several variables.

[PLACE TABLE 17 HERE.]

Figure 2 displays the evidence of the BMA estimation using an "uniform" model prior. It shows that the posterior probability associated with the consumption-(dis)aggregate wealth ratio is large and, therefore, this model provides a good description of the risk premium. However, several other predictors systematically appear in the different models. For instance, this happens with the change in the real government bond yield, the change in the inflation rate or the wealth-to-income ratio. The signs of the coefficients associated to these variables are also consistently the same, i.e., they do not change across the various models under consideration.

[PLACE FIGURE 2 HERE.]

5.1.3 UK

Table 18 summarizes the UK findings. Several financial indicators, such as the dividend yield (0.14-0.77), the change in the real government bond yield (0.02-0.56) and the real government bond yield (0.02-0.56), display the largest PIP. However, there is also a reasonable number of macro-financial variables for which the PIP is larger than 0.10. This is the case of the wealth-to-income ratio (0.03-0.45), the housing wealth-to-income ratio (0.04-0.32) and the consumption-(dis)aggregate wealth ratio (0.03-0.26).Similarly, the C-CAPM model seems relevant for predicting the future risk premium in the UK, as the PIP associated with the lagged consumption growth ranges between 0.02 and 0.28. As for the average number of regressors included in the different models, it ranges between 0.5 and 5, which shows that several variables consistently emerge as being able to track time-variation in the risk premium.

[PLACE TABLE 18 HERE.]

Figure 3 allows us to have a visual idea of the best models, their predictive variables and the signs of the estimated coefficients. In accordance with the previous findings, the dividend yield, the real government bond yield, the change in the real government bond yield and the wealth-to-income ratio are often present in the models with the largest posterior probabilities. In addition, the signs of their coefficients remain unchanged.

[PLACE FIGURE 3 HERE.]

5.2 Different Selection of Zellner's $g-$Priors

In this Section, we provide evidence based on a range of Zellner's g-priors. More specifically, we consider the following hyperparameters on the Zellner's g-prior for the regression coefficients: (*i*) an uniform information prior ("UIP"), which corresponds to $g = N$, where N is the number of observations; (*ii*) a benchmark risk inflation criterion ("BRIC"), which denotes the benchmark prior suggested by Fernandez et al. (2001b), i.e. $g = max(N; K^2)$, where K is the total number of covariates; (*iii*) a risk inflation criterion ("RIC"), which sets $g = K^2$ as in Foster and George (1994); (*iv*) a prior that asymptotically mimics the Hannan and Quinn (1979) criterion ("HQ") and which can be defined as

$g = \log(N)^3$ with $C_{HQ} = 3$; (v) a local Empirical Bayes prior ("EBL"), where $g = \max(0; F_\gamma - 1)$, $F \equiv \frac{R^2_\gamma(N-1-k_\gamma)}{(1-R^2_\gamma)k_\gamma}$ and R^2 is the R-squared of the corresponding model (as in Liang et al. (2008)); and (vi) a more complex hyper g-prior advocated by Feldkircher and Zeugner (2009).

5.2.1 Euro area

Table 19 summarizes the results for the euro area regarding the posterior inclusion probability (PIP) associated to each regressor and the abovementioned Zellner g-priors. It confirms that, despite the amount of uncertainty, some macroeconomic variables (such as the growth rate of the commodity price index (0.14-0.95), the inflation rate (0.04-0.77) and the change in the inflation rate (0.01-0.77)), financial indicators (such as the lagged stock returns (0.78-0.97) and the dividend yield ratio (0.08-0.67)) and macro-financial proxies (such as the stock price index scaled by the real GDP (0.07-0.89), the labour income-to-consumption ratio (0.02-0.87) and the wealth composition risk (0.01-0.84)) are more likely to capture the variation in the future risk premium, as reflected in their PIPs. Indeed, the average number of regressors ranges between 1 and 15, corroborating the predictive ability associated with a reasonably large number of forecasting variables.

[PLACE TABLE 19 HERE.]

5.2.2 US

The empirical findings for the US are presented in Table 19 and reveal that model uncertainty is an important feature for understanding risk premium. In fact, the largest PIPs are associated with the consumption-(dis)aggregate wealth ratio (0.08-0.55), the change in the real government bond yield (0.03-0.53), the change in the inflation rate (0.01-0.51), the labour income-to-consumption ratio (0.00-0.49), the dividend yield (0.01-0.46), the stock price index scaled by GDP (0.00-0.46) and the wealth-to-income ratio (0.01-0.44)). In addition, the average number of regressors varies between 0.2 and 11.

[PLACE TABLE 20 HERE.]

5.2.3 UK

Finally, in Table 21, we summarize the UK findings, which highlight the role played by financial indicators at predicting future real stock returns. In particular, the PIP of the dividend yield (0.06-0.75), the change in the real government bond yield (0.01-0.70) and the real government bond yield (0.01-0.66) rank among the largest ones. A number of macro-financial variables, such as the housing wealth-to-income ratio (0.01-0.59), the wealth-to-income ratio (0.01-0.58) and the consumption-(dis)aggregate wealth ratio (0.01-0.57) are also likely to capture the time-varying pattern of the risk premium. All in all, the average number of potential predictors of real stock returns lies between 0.2 and 13.

[PLACE TABLE 21 HERE.]

5.3 Long-Run Horizon Predictability

The asset pricing literature has documented long-term predictability of stock returns as Section 2 shows. In addition, the results provided in Section 4 display empirical evidence that is consistent with an improvement in terms of forecasting power when the model uncertainty is taken into account. However, it refers to short-run predictability, in that we only consider specifications at the one quarter-ahead horizon. Therefore, the issue of long run horizon predictability and whether BMA helps improving it remains an open question that we try to address in this sub-section.

We start by looking at the set of models taken from the empirical literature. In particular, we consider their "in-sample" predictive ability over different time horizons, H. Then, we account for model uncertainty, and use BMA to estimate the posterior probability associated to each model. Finally, we analyze the forecasting power of the weighted-average model, namely, by comparing it with the benchmark specifications.

In principle, BMA may deliver a better performance at shorter horizons, given that model uncertainty is more important when less data is employed. In fact, at longer horizons, averaging over the different models included in the information set may introduce noise in the predictive model. Similarly, the precision of the predictions about future stock returns may be larger in the short run due to, for example, momentum. On the other hand, the recent literature that developed economically motivated variables that are able to capture time-variation in the risk premium has shown that these models exhibit stronger forecasting power at horizons from 3 to 8 quarters. As a result, one cannot safely say whether the weighted-average model can do better in the short-run or in the long-run.

5.3.1 Euro area

Table 22 reports the results about the relative predictive ability of the weighted-average model vis-a-vis the constant expected returns and the autoregressive benchmark specification at different horizons and for the euro area. It provides a summary of the root mean-squared error and the nested forecast comparisons. We consider two situations: (*a*) the equally weighted-average model using BMA with the Monte Carlo Markov Chain Model Composite (MC^3) method; and (*b*) the weighted-average model built with the posterior probabilities computed by using BMA with the Monte Carlo Markov Chain Model Composite (MC^3) method.

It can be seen that the weighted-average model performs better at longer horizons. In fact, the RMSE strongly falls, in particular, for horizons between 3 and 8 quarters. The superiority of accounting for model uncertainty is also clear in comparison with the benchmark specifications, as suggested by the nested forecasts. Interestingly, the predictive ability of the weighted-average model that is built with the posterior probabilities estimated using BMA is stronger than the equally weighted-average model, both in terms of the RMSE and when analyzed in relation with the constant expected returns and the autoregressive models.

[PLACE TABLE 22 HERE.]

5.3.2 US

The empirical findings for the US can be found in Table 23. As in the case of the euro area, one concludes that the performance of BMA improves over longer horizons. This is highlighted not only by the RMSE of the weighted-average model, but also by the nested forecast comparisons. However, in contrast with the euro area, there is no substantial difference between the predictive ability of the weighted-average model that is built with the posterior probabilities estimated using BMA and the equally weighted-average model. The only exception lies at the horizon of 8 quarters.

[PLACE TABLE 23 HERE.]

5.3.3 UK

As for the UK, a summary of the results is reported in Table 24. As before, BMA delivers stronger forecasting ability at longer horizons, in line with the evidence for the euro area and the US. This is particularly important when the predictive power is assessed vis-a-vis the autoregressive benchmark model. However, in sharp contrast with the findings for the euro area and the US, the results for the UK suggest that the weighted-average model that is built with the posterior probabilities estimated using BMA performs worse than the equally weighted-average model.

In sum, BMA works better with data encompassing long run horizons, where uncertainty about the "true" model governing risk premium is larger. In the short run, it does not work as well, probably, reflecting the large amount of available data. These findings imply that one can better track return predictability at horizons between 3 to 8 quarters when using the BMA framework. In this context, it is in contrast with the works of Chapman and Yan (2002), Torous et al. (2005), Ang and Bekaert (2007) and Gomes (2007), who suggest that short periods, rather than long ones, may be more informative for predictability of asset returns. However, it is in accordance with the findings of Lettau and Ludvigson (2001), Lustig and van Nieuwerburgh (2005), Yogo (2006), Piazzesi et al. (2007) and Sousa (2010a) among others, who find that the asset returns can be better forecast at horizons between 3 to 8 quarters.

[PLACE TABLE 24 HERE.]

5.4 Recursive Forecasts

We now use BMA to track time-variation in the likelihood of the different models. In fact, one potential drawback of the previous findings is that the choice of the forecast period may have a substantial impact on the results, because the predictive ability may substantially vary over time (Goyal and Welch, 2008). In order to address this issue, we investigate the time-variation in BMA performance with the use of recursive forecasts. In practice, we start by considering a minimum number of observations, which we use to assess the posterior probability associated with each model. Then, we add one observation at time and account for model uncertainty by reestimating the posterior probabilities. We keep iterating until the full sample is used. This procedure allows us to build time-series for the estimated posterior probabilities associated with the different models, so that we can understand how the likelihood of a given model in terms of representing the "best-performing" specification for the future risk premium has evolved over time. In fact, in this way, we can infer how the performance of BMA (and, therefore, of the different models) evolves over time and where major forecast breakdowns take place.

5.4.1 Euro area

The recursive posterior probabilities associated with the different models for the euro area are plotted in Figures 4 and 5. The results are broadly consistent with the findings of Section 4. In fact, the models that largely dominate in terms of posterior probability are the ones based on purely financial or macroeconomic indicators. This is the case of the models by: Chen et al. (11986), with an estimated probability around 5% and 60%; Campbell (1987) and Ferson (1990), with an estimated probability of between 10% and 50%; Harvey (1989), with an estimated probability of between 5% and 20%; Ferson and Harvey (1991), where the estimated probability ranges between 5% and 20%; and Ferson and Harvey (1993), where the estimated probability lies between 5% and 30%. In this respect, they clearly reflect the periods of high or low inflation, government bond yields and dividend yield ratio. They also outperform the macro-financial models. In fact, in these cases, the majority of the specifications collects less than 10% of posterior probability (see, for instance, the models by Lettau and Luvigson (2001), Parker and Julliard (2005), Julliard and Sousa (2007a, 2007b) and Sousa (2010a)), despite its sharp increase around the late nineties, that is, a period of strong boom in the stock markets.

[PLACE FIGURE 4 HERE.]

[PLACE FIGURE 5 HERE.]

5.4.2 US

Figures 6 and 7 display the recursive posterior probabilities associated with the different models for the US. In general, the ones that display the highest posterior probability are: Lettau and Ludvigson (2001), with an estimated probability around 4% and 14%; Sousa (2010a), with an estimated probability of between 10% and 50%; Sousa (2010b), with an estimated probability of between 5% and 25%; Julliard and Sousa (2007a), where the estimated probability ranges between 5% and 25%; and Julliard and Sousa (2007b), where the estimated probability lies between 10% and 60%. This piece of evidence largely reflects the importance of episodes of strong financial wealth dynamics that were typically associated with periods of booms in the stock market. Interestingly, the C-CAPM model and the labour income-consumption ratio (Santos and Veronesi, 2006) seem to perform relatively well, although the posterior probabilities associated with these models have substantially declined around 2000. In fact, this represents an important forecast breakdown for these models. Consequently, one can interpret this result as providing support for important changes in the pattern of long-run equilibrium consumption among euro area countries due to the burst of the technological bubble. Another model with a reasonably high forecasting ability over time is the one based on the works of Campbell (1987) and Ferson (1990), which reflects the evolving dynamics of the government bond yields. The models linked with the behavior of the housing markets, such as the housing collateral ratio (Lustig and van Nieuwerburgh, 2005) and the composition risk (Yogo, 2006; Piazzesi et al., 2007) have a higher posterior probability in the first few years of the recursive forecasting period, which highlights the solid growth of the real estate markets in this sub-sample. Finally, the model by Adrian et al. (2010) exhibits a posterior probability that has dramatically increased after 2001. This is explained by the enormous growth of the wealth under dealers and brokers' activity. With the collapse of the financial system in 2007, the estimated posterior probability associated with this model has also strongly fallen.

[PLACE FIGURE 6 HERE.]

[PLACE FIGURE 7 HERE.]

5.4.3 UK

Figures 8 and 9 plot the recursive posterior probabilities of the several models under consideration for the UK. The evidence is similar to the US in that macro-financial models are generally associated with the highest posterior probabilities. For instance, the model by Lettau and Ludvigson (2001) has an estimated probability of between 6% and 12%; the one by Sousa (2010b) has an estimated probability of 6%-20%; the one based on the work of Julliard and Sousa (2007a) has an estimated probability ranging between 8% and 20%; and the model developed by Julliard and Sousa (2007b) has an estimated probability that lies between 10% and 25%. Interestingly, the C-CAPM model performed best among all models in the first 5 years of the sample, when the posterior probability ranged between 20% and 60%. Nevertheless, there is a clear downward trend in the recursive probability of this model, which explains its relatively poor forecasting power over the full sample. This is in line with the works of Paye and Timmermann (2006) and Ang and Bekaert (2007), who find a steady decline in predictability since the late eighties. As for the models that take into account the behavior of the housing markets, they have a higher posterior probability in the first half of the nineties, in correspondence with a period of long-lived fluctuations in housing prices.

[PLACE FIGURE 8 HERE.]

[PLACE FIGURE 9 HERE.]

5.5 Out-of-Sample Forecasts

As a final robustness check, we assess the forecasting power of BMA in an "out-of-sample" context. This exercise faces several econometric issues. First, Ferson et al. (2003) and Torous et al. (2005) argue that the results from the "in-sample" regressions could be spurious and the R^2 statistics and the statistical significance of the regressors might be biased upwards when both the expected returns and the predictive variable are highly persistent. Consequently, we perform an exercise based on "out-of-sample" forecasts, although (as pointed by Inoue and Kilian (2004)) the "in-sample" and "out-of-sample" tests are asymptotically equally reliable under the null of no predictability Similarly, Cochrane (2008) emphasizes the low power of the "out-of-sample" forecasting exercises. Second, Brennan and Xia (2005) show that a "look-ahead" bias could arise when the coefficients of the predictive variable are estimated using the full data sample. This is particularly important in the case of predictors built from the estimation of a fixed cointegrating vector, such as the consumption-wealth ratio (Lettau and Ludvigson, 2001; Julliard and Sousa, 2007a), the housing collateral ratio (Lustig and van Nieuwerburgh, 2005), the consumption-(dis)aggregate wealth ratio (Julliard and Sousa, 2007b; Sousa, 2010a) and the aggregate wealth-to-income ratio (Sousa, 2010b). As a result, we present the results from out-of-sample forecasts using only the data available at the time of the forecast. In particular, we consider the last 10 years of data as the forecasting period. The difficulty with this technique, as argued in Lettau and Ludvigson (2001), is that it could strongly understate the predictive power of the regressor, therefore, making it difficult to display forecasting power when the theory is true.

5.5.1 Euro area

Table 25 reports the results for the relative predictive ability of the weighted-average model vis-a-vis the constant expected returns and the autoregressive benchmark specification at different horizons and for the euro area. It summarizes the information about the root mean-squared error and the nested forecast comparisons. We consider two situations: (*a*) the equally weighted-average model using BMA with the Monte Carlo Markov Chain Model Composite (MC^3) method; and (*b*) the weighted-average model built with the posterior probabilities computed by using BMA with the Monte Carlo Markov Chain Model Composite (MC^3) method.

The empirical findings suggest that the weighted-average model has a stronger forecasting power at longer horizons. In fact, the RMSE strongly falls, in particular, 3 and 4 quarters-ahead. The superiority of BMA is also visible in the comparisons with the benchmark models. Interestingly, while the predictive ability of the weighted-average model built with the posterior probabilities estimated using BMA is larger than the equally weighted-average model at longer horizons, the last model delivers higher precision at short horizons.

[PLACE TABLE 25 HERE.]

5.5.2 US

The empirical findings for the US can be found in Table 26. We conclude that the performance of the equally weighted-average model is larger at longer horizons. In contrast, the weighted-average model based on the posterior probabilities delivers better forecasting properties in the short-run. When we compare the predictive ability of the weighted-average model and the benchmark specifications, we can see that the gains in terms of the precision of the forecasts are magnified vis-a-vis the autoregressive model.

[PLACE TABLE 26 HERE.]

5.5.3 UK

Table 27 provides a summary of the results for the UK. Similar to the evidence for the euro area and the US, the BMA delivers stronger forecasting ability at longer horizons, in particular, when assessed versus the autoregressive benchmark model. The findings suggest that the weighted-average model built with the posterior probabilities estimated using BMA performs worse than the equally weighted-average model, a feature that can also be found in the US. Therefore, the "out-of-sample" evidence largely confirms the "in-sample" findings.

[PLACE TABLE 27 HERE.]

6 Conclusion

The current financial crisis has demonstrated that the financial system, the housing market and the banking sector are strongly connected not only in domestic terms, but also when considering cross-country dimensions. These linkages can generate important wealth dynamics.

In this paper, we show that predicting asset returns in the euro area, the US and the UK faces a large amount of model uncertainty. We use a Bayesian Model Averaging (BMA) approach to account for such uncertainty, and find that it can deliver superior forecasting ability.

The empirical evidence for the euro area suggests that several macroeconomic, financial and macro-financial variables are consistently among the most prominent determinants of the future risk premium. As for the US, only a few factors play an important role. In the case of the UK, the major predictors of future stock returns are financial variables.

These results are corroborated for both the M-open and the M-closed perspectives, a different selection of model priors and Zellner's $g-$priors and in the context of "in-sample" and "out-of-sample" forecasting.

Moreover, we highlight that the predictive power of the weighted-average model is stronger at longer periods and clearly superior to the constant expected returns and autoregressive benchmark models.

From a policy perspective, the BMA can be a useful tool towards resolving the problem of model uncertainty. Most importantly, it can contribute towards the identification of a set of predictors that are able to track future stock returns and, therefore, time-variation in risk premium. Moreover, by assessing the likelihood that some macro-financial variables represent the "best-performing" specification for the expectations about future asset returns, it contributes towards the development of indicators of risk, which can prove helpful in the conduct of macro-prudential policy. Finally, given that the BMA is particularly accurate at predicting risk premium in the medium-term, it allows controlling for the uncertainty about a system of early warning models, as well as the establishment of a weighted-average model with superior forecasting ability.

References

Adrian, T., Moench, E., and H. S. Shin, 2010. Financial intermediation, asset prices, and macroeconomic dynamics. Federal Reserve Bank of New York, Staff Report No. 422.

Ang, A., and G. Bekaert, 2007, Stock return predictability: is it there? Review of Financial Studies, 20(3), 651-707.

Avramov, D., 2002. Stock return predictability and model uncertainty. Journal of Financial Economics, 64, 423-458.

Bansal, R., and A. Yaron, 2004. Risks for the long run: a potential resolution of asset pricing puzzles. Journal of Finance, 59(4), 1481-1509.

Barberis, N., 2000. Investing for the long run when returns are predictable. Journal of Finance, 55, 225-264.

Bossaerts, P. and P. Hillion, 1999. Implementing statistical criteria to select return forecasting models: What do we learn? Review of Financial Studies, 12(2), 405-428.

Breeden, D. T., 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of Financial Economics, 7, 265-296.

Breeden, D. T., Gibbons, M. R., and R. H. Litzenberger, 1989. Empirical tests of the consumption-oriented CAPM, Journal of Finance, 44, 231-62.

Brennan, M., and Y. Xia, 2005. tay's as good as cay. Finance Research Letters, 2, 1-14.

Campbell, J. Y., 1987. Stock returns and the term structure. Journal of Financial Economics, 18(2), 373-399.

Campbell, J. Y., 1996. Understanding risk and return. Journal of Political Economy, 104, 298-45.

Campbell, J. Y., and J. Cochrane, 1999. By force of habit: a consumption-based explanation of aggregate stock market behavior. Journal of Political Economy, 107, 205-251.

Campbell, J. Y., and R. Shiller, 1988. The dividend-price ratio, expectations of future dividends and discount factors. Review of Financial Studies, 1, 195-227.

Campbell, J.Y., and L. M. Viceira, 2002. Strategic asset allocation: Portfolio choice for long-term investors. Oxford University Press: Oxford.

Chapman, D. A., and H. Yan, 2002. A consistent confidence set for an asset return predictability problem. University of Texas at Austin, Working paper.

Chen, N., Roll, R., and S. Ross, 1986. Economic forces and the stock market. Journal of Business, 59, 383-403.

Cochrane, J. H., 1996. A cross-sectional test of an investment-based asset pricing model. Journal of Political Economy, 104, 572-621.

Cochrane, J. H., 2008. The dog that did not bark: a defense of return predictability. Review of Financial Studies, 21, 1533-1575.

Constantinides, G., 1990. Habit-formation: a resolution of the equity premium puzzle. Journal of Political Economy, 98, 519-543.

Cooper, I., and R. Priestley, 2009. Time-varying risk premiums and the output gap. Review of Financial Studies, 22(7), 2801-2833.

Cremers, K. J. M., 2002, Stock return predictability: a Bayesian model selection perspective, Review of Financial Studies, 15, 1223-1249.

Duffee, G., 2005. Time variation in the covariance between stock returns and consumption growth. Journal of Finance, 60, 1673-1712.

Fama, E., 1970. Efficient capital markets: a review of theory and empirical work. Journal of Finance, 25, 383-417.

Fama, E., 1991. Efficient capital markets: II. Journal of Finance, 46, 1575-1617.

Fama, E., 1998. Market efficiency, long term returns and behavioral finance. Journal of Financial Economics, 49, 283-306.

Fama, E., and K. French, 1996. Multifactor explanations of asset pricing anomalies. Journal of Financial Economics, 51, 55-84.

Fama, E., and K. French, 1989. Business conditions and expected returns on stocks and bonds. Journal of Financial Economics, 25(1), 23-49.

Fama, E., and G. W. Schwert, 1977. Asset returns and inflation. Journal of Financial Economics, 5, 115-146.

Farmer, J., and A. Lo, 1999. Frontiers of science: evolution and efficient markets. Proceedings of the National Academy of Sciences, 96, 9991-9992.

Feldkircher, M., and S. Zeugner, 2009. Benchmark priors revisited: On adaptive shrinkage and the supermodel effect in Bayesian Model Averaging. International Monetary Fund, IMF Working Paper No. 202.

Fernandez, C., Ley, E., and M. F. Steel, 2001a. Model uncertainty in cross-country growth regressions. Journal of Applied Econometrics, 16, 563–576.

Fernandez, C., Ley, E., and M. F. Steel, 2001b. Benchmark priors for Bayesian model averaging. Journal of Econometrics, 100, 381-427.

Fernandez-Corugedo, E., Price, S., and A. Blake, 2007. The dynamics of consumers' expenditure: the UK consumption ECM redux. Economic Modelling, 24, 453-469.

Ferson, W., 1990. Are the latent variables in time-varying expected returns compensation for consumption risk? Journal of Finance, 45, 397-429.

Ferson, W., and C. Harvey, 1991. The variation of economic risk premiums. Journal of Political Economy, 99, 385-415.

Ferson, W., and C. Harvey, 1993. The risk and predictability of international equity returns. Review of Financial Studies, 6, 527-566.

Ferson, W., and C. Harvey, 1999. Conditioning variables and the cross-section of stock returns. Journal of Finance, 54, 1325-1360.

Ferson, W., Sarkissian, S., and T. T. Simin, 2003. Spurious regressions in financial economics. Journal of Finance, 58, 1393-1413.

Foster, D., and E. George, 1994. The risk inflation criterion for multiple regression. Annals of Statistics 22, 1947-1975.

Gomes, F. J., 2007. Exploiting short-run predictability. Journal of Banking and Finance, 31, 1427-1440.

Goyal, A., and I. Welch, 2008. A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies, 21, 1456-1508.

Hannan, E. J., and B. G. Quinn, 1979. The determination of the order of an autoregression. Journal of the Royal Statistical Society, B, 41, 190-195.

Harvey, C. R., 1989. Time-varying conditional covariances in tests of asset pricing models. Journal of Financial Economics, 24, 289-317.

Havránek, T., and Z. Irsová, 2011. Determinants of horizontal spillovers from FDI: Evidence from a large meta-analysis. Czech National Bank, CNB Working Paper No. 7.

Hodrick, R., 1992. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. Review of Financial Studies, 3, 357-386.

Inoue, A., and L. Kilian, 2004. In-sample or out of sample tests of predictability: which one should we use? Econometric Reviews, 23(4), 371-402.

Julliard, C., 2004. Labor income risk and asset returns. Princeton University, Primary Job Market Paper.

Julliard, C., and R. M. Sousa, 2007. Risks for the long-run and the cross-section of asset returns. London School of Economics and Political Science, mimeo.

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

Kass, R. E., and L. Wasserman, 1995. A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion. Journal of the American Statistical Association, 90, 928-934.

Koop, G., 2003. Bayesian Econometrics. John Wiley and Sons: Chichester.

Lettau, M., and S. Ludvigson, 2001. Consumption, aggregate wealth, and expected stock returns. Journal of Finance, 56, 815-849.

Ley, E., and M. F. Steel, 2008. On the effect of prior assumptions in Bayesian Model Averaging with applications to growth regressions. Warwick University, Working Paper.

Ley, E., and M. F. Steel, 2009. On the effect of prior assumptions in bayesian model averaging with applications to growth regression. Journal of Applied Econometrics, 24, 651-674.

Ley, E., and M. F. Steel, 2011a. Mixtures of g-priors for Bayesian Model Averaging with economic applications. The World Bank, Policy Research Working Paper No. 5732.

Ley, E., and M. F. Steel, 2011b. Bayesian model averaging and forecasting. Bulletin of E.U. and U.S. Inflation and Macroeconomic Analysis, 30-41.

Liang, F., Paulo, R., Molina, G., Clyde, M. A., and J. O. Berger, 2008. Mixtures of g priors for bayesian variable selection. Journal of the American Statistical Association, 103, 410-423.

Lintner, J., 1965. The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. Review of Economics and Statistics, 47, 13-37.

Lucas, R. E., 1978. Asset prices in an exchange economy. Econometrica, 46, 1429-1445.

Lustig, H., and S. Van Nieuwerburgh, 2005. Housing collateral, consumption insurance, and risk premia: an empirical perspective. Journal of Finance, 60(3), 1167-1219.

Madigan, D., and A. E. Raftery, 1994, Model selection and accounting for model uncertainty in graphical models using Occam's window. Journal of the American Statistical Association, 89, 1535-1546.

Mankiw, N. G., and M. D. Shapiro, 1986. Risk and return: consumption beta versus market beta. Review of Economics and Statistics, 68, 452-59.

Parker, J. A., and C. Julliard, 2005. Consumption risk and the cross section of expected returns. Journal of Political Economy, 113(1), 185-222.

Pastor, L., and R. F. Stambaugh, 2000. Comparing asset pricing models: an investment perspective. Journal of Financial Economics, 56, 335-381.

Pesaran, M. H., and A. Timmermann, 1995. Predictability of stock returns: Robustness and economic significance. Journal of Finance, 50, 1201-1228.

Piazzesi, M., Schneider, M., and S. Tuzel, 2007. Housing, consumption and asset pricing. Journal of Financial Economics, 83, 531-569.

Pontiff, J., and L. D. Schall, 1998. Book-to-market ratios as predictors of market returns. Journal of Financial Economics, 49, 141-160.

Raftery, A. E., 1995. Bayesian model selection in social research. Sociological Methodology, 25, 111-196.

Raftery, A. E., Madigan, D., and J. A. Hoeting, 1997, Bayesian model averaging for linear regression models, Journal of the American Statistical Association, 92, 179-191.

Rangvid, J., 2006. Output and expected returns. Journal of Financial Economics, 81, 595-624.

Rubinstein, M., 1976. The valuation of uncertain income streams and the pricing of options. Bell Journal of Economics, 7(2), 407-425.

Sala-i-Martin, X., Doppelhofer, G., and R.I. Miller, 2004. Determinants of long-term growth: a Bayesian averaging of classical estimates (BACE) approach. American Economic Review, 94(4), 813–835.

Santos, T., and P. Veronesi, 2006. Labor income and predictable stock returns. Review of Financial Studies, 19, 1-44.

Schwarz, T. J., 1978. Estimating the dimension of a model. Annals of Statistics, 6, 416-464.

Sharpe, W., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. Journal of Finance, 19, 425-442.

Sims, C. A., 2007. Monetary policy models. Brookings Papers on Economic Activity, 38(2), 75-90.

Sousa, R. M., 2010a. Consumption, (dis)aggregate wealth, and asset returns. Journal of Empirical Finance, 17(4), 606-622.

Sousa, R. M., 2010b. Collateralizable wealth, asset returns and systemic risk: international evidence. In: Barnett, W. A., and F. Jawadi (Eds.). Nonlinear Modeling of Economic and Financial Time-Series, International Symposia in Economic Theory and Econometrics, 20, Emerald Group Publishing, London.

Sundaresan, S., 1989. Intertemporally dependent preferences and the volatility of consumption and wealth. Review of Financial Studies, 2, 73-89.

Torous, W., Valkanov, R., and S. Yan, 2005. On predicting stock returns with nearly integrated explanatory variables. Journal of Business, 77, 937-966.

Wei, M., 2005. Human capital, business cycles and asset pricing. Board of Governors of the Federal Reserve, Division of Monetary Affairs, Working Paper.

Whitelaw, R., 1994. Time variations and covariations in the expectation and volatility of stock market returns. Journal of Finance, 49, 515-541.

Xia, Y., 2001. Learning about predictability: the effects of parameter uncertainty on dynamic asset allocation. Journal of Finance, 56, 205-246.

Yogo, M., 2006. A consumption-based explanation of expected stock returns. Journal of Finance, 61(2), 539-580.

Zellner, A., 1986. On assessing prior distributions and Bayesian regression analysis with g-prior distributions. In Goel, P. K., and A. Zellner (Eds.). Bayesian inference and decision techniques: Essays in honour of Bruno de Finetti. North-Holland, Amsterdam, 233-243.

Appendix

A Data Description

A.1 Euro area Data

Euro area aggregates are calculated as weighted-average of euro-11 before 1999 and, thereafter, as break-corrected series covering the real-time composition of the euro area.

GDP

Seasonally adjusted nominal GDP ('stocks') at market prices. From 1999:1 onwards, this series covers nominal GDP of the real-time composition of the euro area, correcting for the breaks caused by the several enlargements, i.e. currently the observations from 2007:4 backwards are extrapolations based on growth rates calculated from the levels series compiled for the euro area 15 in 2008. For period before 1999, the nominal GDP series for the euro area is constructed by aggregating national GDP data for euro 11 using the irrevocable fixed exchange rates of 31 December 1998 for the period 1980:1-1998:4. Again, growth rates from this series are used to backward extend the euro area GDP series.

The euro area seasonally adjusted real GDP series (at 2000 constant prices) has been constructed before 1999 by aggregating national real GDP data using the irrevocable fixed exchange rates. As for the euro area nominal GDP, an artificial euro area real GDP series has also been constructed using the procedure illustrated above. Data are quarterly, seasonally adjusted, expressed in million of euro, and comprise the period 1980:1-2007:4.

Price Deflator

All variables are expressed in real terms by using the GDP deflator. The GDP deflator is calculated as a simple ratio between nominal and real GDP. The year base is 2000 (2000 = 100). Data are quarterly, seasonally adjusted, and comprise the period 1980:1-2007:4.

Monetary Aggregate (M3)

All the data used are denominated in euro. The seasonally adjusted M_3 series for the euro area has been constructed using the index of adjusted stocks for the corresponding real time composition of the currency area. This index corrects for breaks due to enlargement, but as well for reclassifications, exchange rate revaluations and other revaluations. In order to translate the index into outstanding amounts, the M_3 seasonally adjusted index of adjusted stocks for the euro area has been re-based to be equal to the value of the seasonally adjusted stock for the euro area M_3 in January 2008. Before 1999, stocks and flows of the estimated "euro area M3" are derived by by aggregating national stocks and flows at irrevocable fixed exchange rates. Data are seasonally adjusted quarterly averages covering the period 1980:2 to 2007:4.

Short-Run Interest Rate

For short-term interest rates from January 1999 onwards, the euro area three-month Euribor is used. Before 1999, the artificial euro area nominal interest rates used are estimated as weighted-averages of national interest rates calculated with fixed weights based on 1999 GDP at PPP exchange rates. National short-term rates are three-month market rates. Data are quarterly averages, and comprise the period 1980:1-2007:4.

Producer Price Index

World market prices of raw materials. Total index. USD basis, converted into euro. Weighted according to commodity imports of OECD countries, 1989-1991, excluding EU- internal trade. Share in total index: 100%. Data are quarterly, seasonally adjusted, and comprise the period 1980:1-2007:4.

Consumption

Total final private consumption. Data are quarterly, seasonally adjusted, expressed in million of euro, and comprise the period 1980:1-2007:4. The construction principle is similar to that described for disposable income.

Disposable Income

Total compensation of employees. From 1999:1 onwards, this series covers nominal disposable income of the real-time composition of the euro area, correcting for the breaks caused by the several enlargements, i.e. currently the observations from 2007:4 backwards are extrapolations based on growth rates calculated from the levels series compiled for the euro area 15 in 2008. For period before 1999, the nominal disposable income series for the euro area is constructed by aggregating national disposable income data for euro 11 using the irrevocable fixed exchange rates of 31 December 1998 for the period 1980:1-1998:4. Again, growth rates from this series are used to backward extend the euro area disposable income series.

The euro area seasonally adjusted real disposable income series (at 2005 constant prices) has been constructed before 1999 by aggregating national real disposable income data using the irrevocable fixed exchange rates. As for the euro area nominal disposable income, an artificial euro area real disposable income series has also been constructed using the procedure illustrated above. Data are quarterly, seasonally adjusted, expressed in million of euro, and comprise the period 1980:1-2007:4.

Aggregate Wealth

Aggregate wealth is defined as the net worth of households and nonprofit organizations, this is, the sum of financial wealth and housing wealth. Original series are provided at quarterly frequency from the euro area quarterly sectoral accounts for the period 1999:1-2007:4 and at annual frequency from the monetary union financial accounts for the period 1995-1998 and from national sources for the period 1980-1994. Quarterly data before 1999 are back-casted and interpolated using quadratic smoothing and corrected for breaks. Data are quarterly, seasonally adjusted, expressed in million of Euro, and comprise the period 1980:1-2007:4.

Financial Wealth

Net financial wealth is the difference between financial assets (currency and deposits, debt securities, shares and mutual fund shares, insurance reserves, net others) and financial liabilities (excluding mortgage loans) held by households and non-profit institutions serving households. Original series are provided at quarterly frequency from the euro area quarterly sectoral accounts for the period 1999:1-2007:4 and at annual frequency from the monetary union financial accounts for the period 1995-1998 and from national sources for the period 1980-1994. Quarterly data before 1999 are back-casted and interpolated using quadratic smoothing and corrected for breaks. Data are quarterly, seasonally adjusted, expressed in million of Euro, and comprise the period 1980:1-2007:4.

Housing Wealth

Net housing wealth is the difference between gross housing wealth and mortgage loans held by households and non-profit institutions serving households. Original series are provided at annual frequency and quarterly data are back-casted and interpolated using quadratic smoothing. Housing wealth data are at current replacement costs net of capital depreciation based on ECB estimates. Data are quarterly, seasonally adjusted, expressed in million of Euro, and comprise the period 1980:1-2007:4.

Stock Market Index

The source is the Financial Market Data Bank Project (FMDB) for the EMU-DS market. Data are quarterly and comprise the period 1980:1-2007:4.

Housing Price Index

The data on euro area house price index comes from the ECB. Data are quarterly and comprise the period 1980:1-2007:4.

Exchange rate

Exchange rate corresponds to real effective exchange rate. Data are quarterly. The series comprises the period 1980:1-2007:4 and the source is the Bank for International Settlements (BIS).

Credit

Credit is proxied by loans for house purchase, which is a component of the loans to Households by Monetary and Financial Institutions (MFI). Original series are provided at annual frequency and quarterly data are back-casted and interpolated using quadratic smoothing. Data are quarterly, seasonally adjusted, expressed in million of Euro, and comprise the period 1980:1-2007:4.

A.2 US Data

GDP

The source is Bureau of Economic Analysis, NIPA Table 1.1.5, line 1. Data for GDP are quarterly, seasonally adjusted , and comprise the period 1947:1-2008:4.

Price Deflator

All variables were deflated by the CPI, All items less food, shelter, and energy (US city average, 1982-1984=100) ("CUSR0000SA0L12E"). Data are quarterly (computed from monthly series by using end-of-period values), seasonally adjusted , and comprise the period 1967:1-2008:4. The source is the Bureau of Labor Statistics.

Monetary Aggregate

Monetary Aggregate corresponds to M2. Data are quarterly, seasonally adjusted , and comprise the period 1960:1-2008:4. The sources are the OECD, Main Economic Indicators (series "USA.MABMM201.STSA") and the Board of Governors of the Federal Reserve System, Release H6.

Short-Run Interest Rate

Short-Run Interest Rate is defined as the Federal Funds effective rate. Data are quarterly (computed from monthly series by using the compounded rate), and comprise, respectively, the periods 1957:2-2008:4. The source is the Board of Governors of the Federal Reserve System, Release H15 (series "RIFSPFF_N.M" and "RIFSGFSM03_N.M").

Producer Price Indexes

Producer Price Indexes include: (*a*) the producers' price index, Materials and components for construction (1982=100) (series "WPUSOP2200"); (*b*) the producers' price index, All commodities (1982=100) (series "WPU00000000"); (*c*) the producers' price index, Crude materials (stage of processing), (1982=100) (series "WPUSOP1000"); (*d*) the producers' price index, Intermediate materials, supplies and components (1982=100) (series "WPUSOP2000"). Data are quarterly (computed from monthly series by using end-of-period values), and comprise the period 1947:1-2008:4. All series are seasonally adjusted using Census X12 ARIMA. The source is the Bureau of Labor Statistics.

Unemployment Rate

Unemployment rate is defined as the civilian unemployment rate (16 and over) (series "LNS14000000"). Data are quarterly (computed from monthly series by using end-of-period values), seasonally adjusted and comprise the period 1948:1-2008:4. The source is the Bureau of Labor Statistics, Current Population Survey.

Consumption

Consumption is defined as the expenditure in non-durable consumption goods and services. Data are quarterly, seasonally adjusted at an annual rate, measured in billions of dollars (2000 prices), in per capita terms and expressed in the logarithmic form. Series comprises the period 1947:1-2008:4. The source is U.S. Department of Commerce, Bureau of Economic Analysis, NIPA Table 2.3.5.

Disposable Income

After-tax labor income is defined as the sum of wage and salary disbursements (line 3), personal current transfer receipts (line 16) and employer contributions for employee pension and insurance funds (line 7) minus personal contributions for government social insurance (line 24), employer contributions for government social insurance (line 8) and taxes. Taxes are defined as: [(wage and salary disbursements (line 3)) / (wage and salary disbursements (line 3)+ proprietor' income with inventory valuation and capital consumption adjustments (line 9) + rental income of persons with capital consumption adjustment (line 12) + personal dividend income (line 15) + personal interest income (line 14))] * (personal current taxes (line 25)). Data are quarterly, seasonally adjusted at annual rates, measured in billions of dollars (2000 prices), in per capita terms and expressed in the logarithmic form. Series comprises the period 1947:1-2008:4. The source of information is U.S. Department of Commerce, Bureau of Economic Analysis, NIPA Table 2.1..

Aggregate wealth

Aggregate wealth is defined as the net worth of households and nonprofit organizations. Data are quarterly, seasonally adjusted at an annual rate, measured in billions of dollars (2000 prices), in per capita terms and expressed in the logarithmic form. Series comprises the period 1952:2-2008:4. The source of information is Board of Governors of Federal Reserve System, Flow of Funds Accounts, Table B.100, line 41 (series FL152090005.Q).

Financial wealth

Financial wealth is defined as the sum of financial assets (deposits, credit market instruments, corporate equities, mutual fund shares, security credit, life insurance reserves, pension fund reserves, equity in noncorporate business, and miscellaneous assets - line 8 of Table B.100 - series FL154090005.Q) minus financial liabilities (credit market instruments excluding home mortgages, security credit, trade payables, and deferred and unpaid life insurance premiums - line 30 of Table B.100 - series FL154190005.Q). Data are quarterly, seasonally adjusted at an annual rate, measured in billions of dollars (2000 prices), in per capita terms and expressed in the logarithmic form. Series comprises the period 1952:2-2008:4. The source of information is Board of Governors of Federal Reserve System, Flow of Funds Accounts, Table B.100.

Housing wealth

Housing wealth (or home equity) is defined as the value of real estate held by households (line 4 of Table B.100 - series FL155035015.Q) minus home mortgages (line 32 of Table B.100 - series FL153165105.Q). Data are quarterly, seasonally adjusted at an annual rate, measured in billions of dollars (2000 prices), in per capita terms and expressed in the logarithmic form. Series comprises the period 1952:2-2008:4. The source of information is Board of Governors of Federal Reserve System, Flow of Funds Accounts, Table B.100.

Stock Market Index

Stock Market Index corresponds to S&P 500 Composite Price Index (close price adjusted for dividends and splits). Data are quarterly (computed from monthly series by using end-of-period values), and comprise the period 1950:1-2008:4.

Housing Price Index

Housing prices are measured using two sources: (*a*) the Price Index of New One-Family Houses sold including the Value of Lot provided by the US Census, an index based on houses sold in 1996, available for the period 1963:1-2008:4; and (*b*) the House Price Index computed by the Office of Federal Housing Enterprise Oversight (OFHEO), available for the period 1975:1-2008:4. Data are quarterly, seasonally adjusted.

Other Housing Market Indicators are provided by the US Census. We use the Median Sales Price of New Homes Sold including land and the New Privately Owned Housing Units Started. The data for the Median Sales Price of New Homes Sold including land are quarterly, seasonally adjusted using Census X12 ARIMA, and comprise the period 1963:1-2008:4. The data for the New Privately Owned Housing Units Started are quarterly (computed by the sum of corresponding monthly values), seasonally adjusted and comprise the period 1959:1-2008:4.

Exchange Rate

Exchange rate corresponds to real effective exchange rate (series "RNUS"). Data are quarterly (computed from monthly series by using end-of-period values). The series comprises the period 1964:1-2008:4 and the source is the Bank for International Settlements (BIS).

Asset Returns

Asset returns were computed using the MSCI-US Total Return Index, which measure the market performance, including price performance and income from dividend payments. I use the index which includes gross dividends, this is, approximating the maximum possible dividend reinvestment. The amount reinvested is the dividend distributed to individuals resident in the country of the company, but does not include tax credits. Series comprises the period 1970:1-2008:4. The source of information is Morgan Stanley Capital International (MSCI).

Credit

Credit corresponds to consumer credit. Data are quarterly, seasonally adjusted at an annual rate, measured in billions of dollars (2000 prices), in per capita terms and expressed in the logarithmic form. Series comprises the period 1952:2-2008:4. The source of information is Board of Governors of Federal Reserve System, Flow of Funds Accounts, Table B.100, line 34 (series FL153166000.Q).

Brokers and Dealers' Leverage Ratio

Brokers and dealers' leverage ratio is defined as assets divided by equity where equity is the difference between assets and liabilities. Data are quarterly, seasonally adjusted at an annual rate, measured in billions of dollars (2000 prices), in per capita terms and expressed in the logarithmic form. Series comprises the period 1952:2-2008:4. The source of information is Board of Governors of Federal Reserve System, Flow of Funds Accounts, Table L.129, lines 1 and 13 (series FL664090005.Q and FL664190005.Q).

A.3 UK Data

GDP

The source is Office for National Statistics (ONS), series "YBHA". Data for GDP are quarterly, seasonally adjusted , and comprise the period 1955:1-2008:4.

Price Deflator

All variables were deflated by the GDP deflator (series "YBGB"). Data are quarterly, seasonally adjusted, and comprise the period 1955:1-2008:4. The source is the Office for National Statistics.

Monetary Aggregate

Monetary Aggregate corresponds to M4. Data are quarterly, seasonally adjusted , and comprise the period 1963:2-2008:4. The source is the Office for National Statistics, series "AUYN".

Short-Run Interest Rate

Short-Run Interest Rate is defined as the 3-month Treasury Bill rate. Data are quarterly (computed from monthly series by using the compounded rate), and comprise the period 1963:2-2008:4. The source is the Datastream, series "UK3MTHINE".

Producer Price Index

Producer Price Indexes include the producers' price index, Input prices (materials and fuel) (series "RNNK"). Data are quarterly (computed from monthly series by using end-of-period values), and comprise the period 1974:1-2008:4. All series are seasonally adjusted using Census X12 ARIMA. The source is the Office for National Statistics.

Unemployment Rate

Unemployment rate is defined as the civilian unemployment rate (16 and over) (series "MGSX"). Data are quarterly (computed from monthly series by using end-of-period values), seasonally adjusted and comprise the period 1971:1-2008:4. The source is the Office for National Statistics.

Consumption

Consumption is defined as total consumption (ZAKV) less consumption of durable (UTIB) and semi-durable goods (UTIR). Data are quarterly, seasonally adjusted at an annual rate, measured in millions of pounds (2001 prices), in per capita and expressed in the logarithmic form. Series comprises the period 1963:1-2008:4. The source is Office for National Statistics (ONS).

Disposable Income

After-tax labor income is defined as the sum of wages and salaries (ROYJ), social benefits (GZVX), self employment (ROYH), other benefits (RPQK + RPHS + RPHT - ROYS - GZVX + AIIV), employers social contributions (ROYK) less social contributions (AIIV) and taxes. Taxes are defined as (taxes on income (RPHS) and other taxes (RPHT)) x ((wages and salaries (ROYJ) + self employment (ROYH)) / (wages and salaries (ROYJ) + self employment (ROYH) + other income (ROYL - ROYT + NRJN - ROYH)). Data are quarterly, measured in millions of pounds (2001 prices), in per capita terms and expressed in the logarithmic form. Series comprises the period 1974:3-2008:4. The sources of information are: Fernandez-Corugedo et al. (2007) - provided by the Office for National Statistics (ONS) -, for the period 1974:3-1986:4; and the Office for National Statistics (ONS), for the period 1987:1-2008:4.

Aggregate wealth

Aggregate wealth is defined as the net worth of households and nonprofit organizations, this is, the sum of financial wealth and housing wealth. Data are quarterly, seasonally adjusted at an annual rate, measured in millions of pounds (2001 prices), in per capita terms and expressed in the logarithmic form. Series comprises the period 1975:1-2008:4. The sources of information are: Fernandez-Corugedo et al.

(2007) - provided by the Office for National Statistics (ONS) -, for the period 1975:1-1986:4; and the Office for National Statistics (ONS), for the period 1987:1-2008:4.

Financial wealth

Financial wealth is defined as the net financial wealth of households and nonprofit organizations (NZEA). Data are quarterly, seasonally adjusted at an annual rate, measured in millions of pounds (2001 prices), in per capita terms and expressed in the logarithmic form. Series comprises the period 1970:1-2008:4. The sources of information are: Fernandez-Corugedo et al. (2007) - provided by the Office for National Statistics (ONS) -, for the period 1970:1-1986:4; and the Office for National Statistics (ONS), for the period 1987:1-2008:4.

Housing wealth

Housing wealth is defined as the housing wealth of households and nonprofit organizations and is computed as the sum of tangible assets in the form of residential buildings adjusted by changes in house prices (CGRI), the dwellings (of private sector) of gross fixed capital formation (GGAG) and Council house sales (CTCS). Data are quarterly, seasonally adjusted at an annual rate, measured in millions of pounds (2001 prices), in per capita terms and expressed in the logarithmic form. Series comprises the period 1975:1-2008:4. The sources of information are: Fernandez-Corugedo et al. (2007) - provided by the Office for National Statistics (ONS) -, for the period 1975:1-1986:4; and the Office for National Statistics (ONS), for the period 1987:1-2008:4. For data on house prices, the sources of information are: Office of the Deputy Prime Minister (ODPM), Halifax Plc and the Nationwide Building Society.

Stock Market Index

Stock Market Index corresponds to FTSE-All shares Index. Data are quarterly (computed from monthly series by using end-of-period values), and comprise the period 1975:1-2008:4.

Housing Price Index

Housing Price Index corresponds to Nationwide: All Houses Price Index. Data are quarterly, seasonally adjusted using Census X12 ARIMA, and comprise the period 1955:1-2008:4.

Exchange rate

Exchange rate corresponds to real effective exchange rate (series "RNGB"). Data are quarterly (computed from monthly series by using end-of-period values). The series comprises the period 1964:1-2008:4 and the source is the Bank for International Settlements (BIS).

Asset Returns

Asset returns were computed using the MSCI-UK Total Return Index, which measure the market performance, including price performance and income from dividend payments. I use the index which includes gross dividends, this is, approximating the maximum possible dividend reinvestment. The amount reinvested is the dividend distributed to individuals resident in the country of the company, but does not include tax credits. Series comprises the period 1970:1-2008:4. The source of information is Morgan Stanley Capital International (MSCI).

Credit

Credit corresponds to mortgage loans. Data are quarterly, seasonally adjusted at an annual rate, measured in billions of dollars (2000 prices), in per capita terms and expressed in the logarithmic form. Series comprises the period 1983:1-2007:4. The source of information is the Halifax mortgage affordability index from Halifax Plc.

List of Tables

Table 1: Macroeconomic, financial and macro-financial predictors of risk premium.

Author(s)	Variable
	Macroeconomic activity
Sharpe (1964), Lintner (1965), Lucas (1978) and Breeden (1979)	Consumption growth
Parker and Julliard (2005)	Ultimate consumption risk
Cooper and Priestley (2009)	Output gap
	Financial indicators
Campbell and Shiller (1988)	Dividend yield
Fama and Schwert (1977), Hodrick (1992) and Ang and Bekaert (2007)	Short-term interest rates
Campbell (1987) and Fama and French (1989)	Default and term spreads
Adrian et al. (2010)	Leverage ratio of brokers and dealers
	Macro-financial indicators
Lettau and Ludvigson (2001)	Consumption-wealth ratio
Bansal and Yaron (2004)	Long-run risk
Julliard (2004)	Labour income risk
Lustig and van Nieuwerburgh (2005)	Housing collateral ratio
Rangvid (2006)	Price-to-output ratio
Santos and Veronesi (2006)	Labour income-to-consumption ratio
Yogo (2006) and Piazzesi et al. (2007)	Consumption composition risk
Julliard and Sousa (2007)	Change in consumption-wealth ratio
Sousa (2010a)	Wealth composition risk
Sousa (2010b)	Wealth-to-income ratio

Table 2: Bayesian Model Averaging using the Occam's Window method: EA evidence for the 10 "best" models.

	p!=0	EV	SD	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Constant	100.0	-1.389	0.859	-2.039	-0.974	-0.808	-1.621	-2.254	-1.870	-1.984	-2.255	0.017	-2.115
ΔC_{t-1}	9.4	0.159	0.704										
ΔC_{t-12}	2.1	-0.001	0.047										
r_{t-1}	100.0	0.349	0.091	0.334	0.381	0.373	0.328	0.345	0.378	0.311	0.334	0.304	0.332
$bond_{t-1}$	38.7	-0.506	0.735		-1.478	-1.173	-1.371				-0.716		
$\Delta bond_{t-1}$	9.5	-0.195	0.772							-2.074			
og_{t-1}	10.5	-0.022	0.091									-0.333	
π_{t-1}	54.2	-4.968	5.229	-10.135				-10.313	-11.124	-10.031	-8.399		-10.326
$\Delta\pi_{t-1}$	46.7	2.881	3.556	6.507				6.478	6.855	6.331	5.682		6.550
Δi_{t-1}	1.9	0.000	0.003										
Δm_{t-1}	0.6	-0.008	0.153										
Δhp_{t-1}	1.4	-0.005	0.104										
Δe_{t-1}	1.4	-0.001	0.019										
Δcp_{t-1}	97.5	-0.252	0.092	-0.279	-0.219	-0.249	-0.244	-0.300	-0.286	-0.287	-0.277	-0.228	-0.300
$cday_{t-1}$	56.3	2.571	2.602	4.739			3.047	5.094	4.932	4.674	4.731		5.214
$\Delta cday_{t-1}$	1.4	0.060	0.560										
lc_{t-1}	73.5	-2.389	1.850	-3.913		-1.451	-2.607	-4.173	-3.679	-4.000	-3.566		-4.405
rwy_{t-1}	5.6	0.003	0.031										0.101
wy_{t-1}	6.1	0.011	0.054					0.176					
$divyld_{t-1}$	12.9	0.004	0.011						0.028				
$spgdp_{t-1}$	89.3	-0.135	0.067	-0.178	-0.120	-0.124	-0.155	-0.196	-0.142	-0.175	-0.200		-0.182
$\Delta cred_{t-1}$	1.6	-0.009	0.112										
nVar				7	4	5	6	8	8	8	8	3	8
$\bar{R}^2$				0.392	0.294	0.320	0.349	0.405	0.405	0.404	0.404	0.248	0.402
post prob				0.125	0.067	0.044	0.038	0.036	0.035	0.034	0.032	0.028	0.028

Note: Number of selected models: 64. Cumulative posterior probability: 46.7%. p!=0 denotes the posterior inclusion probability, EV corresponds to the mean of the posterior distribution of the parameter and SD stands for standard deviation of the posterior distribution of the parameter. All results are based on the Occam's Window method.

Table 3: Bayesian Model Averaging using the Markov Chain Monte Carlo Model Composite (MC3) method: EA evidence for the 10 "best" models.

	prob	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Constant											
ΔC_{t-1}	0.043										
ΔC_{t-12}	0.019										
r_{t-1}	0.990	X	X	X	X	X	X	X	X	X	X
$bond_{t-1}$	0.358	X		X				X			
$\Delta bond_{t-1}$	0.031										
og_{t-1}	0.249		X								X
π_{t-1}	0.080										
$\Delta \pi_{t-1}$	0.063										
Δi_{t-1}	0.026										
Δm_{t-1}	0.014										
Δhp_{t-1}	0.024										
Δe_{t-1}	0.019										
Δcp_{t-1}	0.625	X	X		X			X	X		
$cday_{t-1}$	0.063										
$\Delta cday_{t-1}$	0.032										
lc_{t-1}	0.129							X			
rwy_{t-1}	0.032										
wy_{t-1}	0.017										
$divyld_{t-1}$	0.196					X			X		
$spgdp_{t-1}$	0.487	X		X				X		X	
$\Delta cred_{t-1}$	0.053										
nVar		4	3	3	2	2	1	5	3	2	2
post prob		0.105	0.067	0.051	0.047	0.041	0.041	0.032	0.029	0.026	0.025

Note: Number of selected models: 614. Cumulative posterior probability: 46.5%. prob denotes the posterior inclusion probability. All results are based on the Markov Chain Monte Carlo Model Composite (MC3) method.

Table 4: Bayesian Model Averaging using the Occam's Window method: US evidence for the 10 "best" models.

	p!=0	EV	SD	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Constant	100.0	2.170	1.247	2.407	2.553	2.566	2.708	0.020	0.021	2.380	2.626	2.261	2.429
ΔC_{t-1}	0.0	0.000	0.000										
ΔC_{t-12}	0.0	0.000	0.000										
r_{t-1}	4.8	-0.004	0.026										
$bond_{t-1}$	0.0	0.000	0.000										
$\Delta bond_{t-1}$	59.7	-1.659	1.651	-2.652		-2.798		-2.882				-2.538	-3.195
og_{t-1}	2.1	-0.000	0.007										
π_{t-1}	0.0	0.000	0.000										
$\Delta \pi_{t-1}$	35.1	1.445	2.323			4.260	3.966						
Δi_{t-1}	0.0	0.000	0.000										
Δm_{t-1}	2.9	-0.019	0.177										
Δhp_{t-1}	0.0	0.000	0.000										
Δe_{t-1}	0.0	0.000	0.000										
Δcp_{t-1}	1.0	-0.002	0.050										
$cday_{t-1}$	85.5	0.991	0.567	1.100	1.168	1.174	1.239			1.088	1.203	1.033	1.111
$\Delta cday_{t-1}$	0.0	0.000	0.000										
lc_{t-1}	1.6	-0.008	0.069										
rwy_{t-1}	0.0	0.000	0.000										
wy_{t-1}	10.3	-0.020	0.075							-0.198		-0.173	
$divyld_{t-1}$	4.4	-0.000	0.001										
$spgdp_{t-1}$	7.1	-0.002	0.012										
Δu_{t-1}	3.7	-0.000	0.006										-0.026
φ_{t-1}	3.0	-0.038	0.347										
$\Delta cred_{t-1}$	5.8	0.019	0.109								0.407		
$aSBRDLR_{t-1}$	1.2	-0.000	0.000										
nVar				2	1	3	2	1	0	2	2	3	3
$\bar{R}^2$				0.094	0.058	0.124	0.084	0.043	0.000	0.071	0.071	0.104	0.104
post prob				0.135	0.125	0.104	0.068	0.045	0.031	0.029	0.027	0.025	0.025

Note: Number of selected models: 37. Cumulative posterior probability: 61.4%. p!=0 denotes the posterior inclusion probability, EV corresponds to the mean of the posterior distribution of the parameter and SD stands for standard deviation of the posterior distribution of the parameter. All results are based on the Occam's Window method.

Table 5: Bayesian Model Averaging using the Markov Chain Monte Carlo Model Composite (MC3) method: US evidence for the 10 "best" models.

	prob	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Constant											
ΔC_{t-1}	0.016										
$\Delta C_{t-1,t-12}$	0.023										
r_{t-1}	0.016										
$bond_{t-1}$	0.016										
$\Delta bond_{t-1}$	0.240			X	X						
og_{t-1}	0.021										X
π_{t-1}	0.019					X				X	
$\Delta \pi_{t-1}$	0.068										
Δi_{t-1}	0.021										
Δm_{t-1}	0.026										
Δhp_{t-1}	0.019										
Δe_{t-1}	0.013										
Δcp_{t-1}	0.017										
$cday_{t-1}$	0.542	X		X		X		X	X		
$\Delta cday_{t-1}$	0.029										
lc_{t-1}	0.055										
rwy_{t-1}	0.026										
wy_{t-1}	0.072						X	X			
$divyld_{t-1}$	0.047										
$spgdp_{t-1}$	0.068										
Δu_{t-1}	0.020										
φ_{t-1}	0.016										
$\Delta cred_{t-1}$	0.026								X		
$aSBRDLR_{t-1}$	0.032										
nVar		1	0	2	1	2	1	2	2	1	1
post prob		0.196	0.130	0.077	0.069	0.038	0.018	0.016	0.015	0.014	0.010

Note: Number of selected models: 652. Cumulative posterior probability: 58.3%. prob denotes the posterior inclusion probability. All results are based on the Markov Chain Monte Carlo Model Composite (MC3) method.

Table 6: Bayesian Model Averaging using the Occam's Window method: UK evidence for the 10 "best" models.

	p!=0	EV	SD	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Constant	100.0	-0.117	0.191	-0.062	-0.041	-0.090	-0.053	-0.060	-0.064	-0.083	-0.070	-0.051	-0.061
ΔC_{t-1}	22.4	-0.4551	1.032		-1.992								
$\Delta C_{t-1,t-12}$	3.7	0.033	0.209										
r_{t-1}	3.7	-0.002	0.022										
$bond_{t-1}$	81.4	-3.271	1.919	-4.243	-4.202	-2.956	-4.198	-4.274	-3.980		-4.022	-4.258	-4.239
$\Delta bond_{t-1}$	78.9	2.342	1.473	2.983	2.966	2.490	3.045	3.259	2.977		2.922	2.975	3.030
og_{t-1}	10.2	-0.017	0.068				-0.128						
π_{t-1}	0.0	0.000	0.000										
$\Delta\pi_{t-1}$	2.4	0.012	0.177										
Δi_{t-1}	6.0	-0.000	0.004					-0.013					
Δm_{t-1}	1.5	0.000	0.001										
Δhp_{t-1}	0.0	0.000	0.000										
Δe_{t-1}	5.1	0.014	0.087						0.215				
Δcp_{t-1}	4.2	-0.011	0.092										-0.242
$cday_{t-1}$	5.6	0.091	0.416										
$\Delta cday_{t-1}$	1.3	-0.002	0.219										
lc_{t-1}	2.5	0.008	0.1000										
rwy_{t-1}	20.8	-0.009	0.022			-0.060							
wy_{t-1}	71.1	-0.115	0.086	-0.162	-0.179		-0.167	-0.166	-0.154		-0.118	-0.170	-0.162
$divyld_{t-1}$	95.2	0.083	0.037	0.100	0.097	0.085	0.097	0.100	0.096	0.028	0.099	0.098	0.100
$spgdp_{t-1}$	3.8	-0.004	0.021										
Δu_{t-1}	3.8	0.000	0.009									0.027	
φ_{t-1}	0.0	0.000	0.000										
$\Delta cred_{t-1}$	0.0	0.000	0.000										
nVar				4	5	4	5	5	5	1	5	5	5
$\bar{R}^2$				0.239	0.263	0.222	0.256	0.251	0.246	0.075	0.244	0.243	0.243
post prob				0.201	0.087	0.076	0.058	0.043	0.031	0.030	0.028	0.028	0.027

Note: Number of selected models: 37. Cumulative posterior probability: 60.9%. p!=0 denotes the posterior inclusion probability, EV corresponds to the mean of the posterior distribution of the parameter and SD stands for standard deviation of the posterior distribution of the parameter. All results are based on the Occam's Window method.

Table 7: Bayesian Model Averaging using the Markov Chain Monte Carlo Model Composite (MC3) method: UK evidence for the 10 "best" models.

	prob	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Constant											X
ΔC_{t-1}	0.044										
$\Delta C_{t-1,t-12}$	0.019										
r_{t-1}	0.021			X							X
$bond_{t-1}$	0.227			X							X
$\Delta bond_{t-1}$	0.190										
og_{t-1}	0.030										
π_{t-1}	0.030										
$\Delta \pi_{t-1}$	0.025										
Δi_{t-1}	0.013										
Δm_{t-1}	0.022										
Δhp_{t-1}	0.028										
Δe_{t-1}	0.078										
Δcp_{t-1}	0.014										
$cday_{t-1}$	0.115					X					
$\Delta cday_{t-1}$	0.027										
lc_{t-1}	0.041										
rwy_{t-1}	0.165						X	X	X		
wy_{t-1}	0.242			X						X	X
$divyld_{t-1}$	0.554	X		X		X	X				X
$spgdp_{t-1}$	0.174				X				X		
Δu_{t-1}	0.026										
φ_{t-1}	0.017										
$\Delta cred_{t-1}$	0.020										
$aSBRDLR_{t-1}$											
nVar		1	0	4	1	2	2	1	2	1	5
post prob		0.106	0.087	0.083	0.050	0.031	0.028	0.022	0.020	0.018	0.014

Note: Number of selected models: 653. Cumulative posterior probability: 45.9%. prob denotes the posterior inclusion probability. All results are based on the Markov Chain Monte Carlo Model Composite (MC3) method.

Table 8: Predictive variables included in the model from the empirical finance literature.

	CRR_1986	C_1987; F_1990	H_1989	FH_1991	FH_1993	W_1994; PS_1998; FH_1999	PT_1995	JS_2007a	JS_2007b	BH_1999
ΔC_{t-1}										
$\Delta C_{t-1,t-12}$										
r_{t-1}	X	X	X	X	X					X
$bond_{t-1}$	X	X	X	X	X	X	X			X
$\Delta bond_{t-1}$	X	X					X			X
og_{t-1}	X				X		X			
π_{t-1}	X				X		X			
$\Delta\pi_{t-1}$	X			X	X					
$cday_{t-1}$									X	
$\Delta cday_{t-1}$									X	
cay_{t-1}								X		
Δcay_{t-1}								X		
lc_{t-1}										
rwy_{t-1}										
wy_{t-1}										
$divyld_{t-1}$			X	X		X	X			X
φ_{t-1}										
$aSBRDLR_{t-1}$										
nVar	6	3	3	4	5	2	5	2	2	4

Table 9: Predictive variables included in the model from the empirical finance literature (cont.).

	C-CAPM	S_2010b	LL_2001	S_2010a	PJ_2005	LvN_2005	SV_2006	Y_2006; PST_2007	AMS_2010
ΔC_{t-1}	X								
$\Delta C_{t-1,t-12}$					X				
r_{t-1}									
$bond_{t-1}$									
$\Delta bond_{t-1}$									
og_{t-1}									
π_{t-1}									
$\Delta\pi_{t-1}$									
$cday_{t-1}$				X					
$\Delta cday_{t-1}$									
cay_{t-1}			X						
Δcay_{t-1}									
lc_{t-1}							X		
rwy_{t-1}						X			
wy_{t-1}		X							
$divyld_{t-1}$									
φ_{t-1}								X	
$aSBRDLR_{t-1}$									X
nVar	1	1	1	1	1	1	1	1	1

Table 10: Bayesian Model Averaging using the Markov Chain Monte Carlo Model Composite method: EA evidence for a selection of 19 models taken from the empirical finance literature.

Model	CRR_1986	C_1987; F_1990	H_1989	FH_1991	FH_1993
RMSE	0.792	0.847	0.802	0.793	0.800
Model vs. Constant	0.890	0.920	0.895	0.890	0.894
Model vs. AR1	0.932	0.964	0.938	0.932	0.936
$\bar{R}^2$	0.158	0.127	0.174	0.174	0.158
post prob	0.353	0.202	0.036	0.052	0.288
Model	W_1994; PS_1998; FH_1999	PT_1995	JS_2007a	JS_2007b	BH_1999
RMSE	0.975	0.874	0.994	0.994	0.796
Model vs. Constant	0.987	0.935	0.997	0.997	0.892
Model vs. AR1	1.034	0.979	1.044	1.044	0.934
$\bar{R}^2$	0.005	0.080	0.000	0.000	0.171
post prob	0.000	0.001	0.003	0.004	0.039
Model	C-CAPM	S_2010b	LL_2001	S_2010a	PJ_2005
RMSE	0.995	0.982	0.994	0.998	0.962
Model vs. Constant	0.997	0.991	0.997	0.999	0.981
Model vs. AR1	1.044	1.037	1.044	1.046	1.027
$\bar{R}^2$	0.000	0.009	0.000	0.000	0.028
post prob	0.001	0.001	0.003	0.004	0.007
Model	LvN_2005	SV_2006			
RMSE	0.999	0.998			
Model vs. Constant	0.999	0.999			
Model vs. AR1	1.047	1.046			
$\bar{R}^2$	0.000	0.000			
post prob	0.000	0.003			

Note: "post prob" denotes the posterior probability and RMSE corresponds to the root mean-squared error. "Model vs. Constant" is the ratio of the RMSE of the model and the RMSE of the constant expected returns benchmark model. "Model vs AR1" is the ratio of the RMSE of the model and the RMSE of the autoregressive benchmark model. $\bar{R}^2$ stands for the adjusted-R^2 statistic. All results are based on the Markov Chain Monte Carlo Model Composite (MC^3) method.

Table 11: In-sample performance of the averaged model vis-a-vis benchmark models: EA evidence.

weighted-average model	Occam Window + Equal	Occam Window + Bayes
RMSE	0.620	0.626
Model vs. Constant	0.787	0.791
Model vs. AR1	0.824	0.829
Model	MCMC + Equal	MCMC + Bayes
RMSE	0.863	0.789
Model vs. Constant	0.929	0.888
Model vs. AR1	0.973	0.930

Note: RMSE corresponds to the root mean-squared error. "Model vs. Constant" is the ratio of the RMSE of the model and the RMSE of the constant expected returns benchmark model. "Model vs AR1" is the ratio of the RMSE of the model and the RMSE of the autoregressive benchmark model. "Equal" and "Bayes" stand for the equally weighted-average model and average model based on the posterior probabilities, respectively. The results are based on the Occam Window and the Markov Chain Monte Carlo Model Composite (MC^3) methods.

Table 12: Bayesian Model Averaging using the Markov Chain Monte Carlo Model Composite method: US evidence for a selection of 19 models taken from the empirical finance literature.

Model	CRR_1986	C_1987; F_1990	H_1989	FH_1991	FH_1993
RMSE	0.916	0.951	0.993	0.973	0.953
Model vs. Constant	0.957	0.975	0.997	0.987	0.976
Model vs. AR1	0.927	0.944	0.965	0.955	0.945
$\bar{R}^2$	0.039	0.026	0.000	0.000	0.009
post prob	0.002	0.010	0.000	0.000	0.000
Model	W_1994; PS_1998; FH_1999	PT_1995	JS_2007a	JS_2007b	BH_1999
RMSE	0.994	0.943	0.982	0.941	0.949
Model vs. Constant	0.997	0.971	0.991	0.991	0.974
Model vs. AR1	0.965	0.940	0.959	0.959	0.943
$\bar{R}^2$	0.000	0.019	0.003	0.044	0.021
post prob	0.000	0.000	0.026	0.262	0.000
Model	C-CAPM	S_2010b	LL_2001	S_2010a	PJ_2005
RMSE	1.000	0.978	0.982	0.942	0.993
Model vs. Constant	0.991	0.989	0.991	0.971	0.996
Model vs. AR1	0.959	0.957	0.959	0.940	0.964
$\bar{R}^2$	0.000	0.015	0.011	0.050	0.000
post prob	0.038	0.021	0.045	0.517	0.017
Model	LvN_2005	SV_2006	Y_2006; PST_2007	AMS_2010	
RMSE	0.998	1.000	0.993	0.994	
Model vs. Constant	0.999	1.000	0.996	0.997	
Model vs. AR1	0.967	0.968	0.964	0.965	
$\bar{R}^2$	0.000	0.000	0.000	0.000	
post prob	0.003	0.004	0.054	0.000	

Note: "post prob" denotes the posterior probability and RMSE corresponds to the root mean-squared error. "Model vs. Constant" is the ratio of the RMSE of the model and the RMSE of the constant expected returns benchmark model. "Model vs AR1" is the ratio of the RMSE of the model and the RMSE of the autoregressive benchmark model. $\bar{R}^2$ stands for the adjusted-R^2 statistic. All results are based on the Markov Chain Monte Carlo Model Composite (MC3) method.

Table 13: In-sample performance of the averaged model vis-a-vis benchmark models: US evidence.

weighted-average model	Occam Window + Equal	Occam Window + Bayes
RMSE	0.885	0.880
Model vs. Constant	0.941	0.938
Model vs. AR1	0.911	0.908
Model	MCMC + Equal	MCMC + Bayes
RMSE	0.951	0.943
Model vs. Constant	0.975	0.971
Model vs. AR1	0.944	0.940

Note: RMSE corresponds to the root mean-squared error. "Model vs. Constant" is the ratio of the RMSE of the model and the RMSE of the constant expected returns benchmark model. "Model vs AR1" is the ratio of the RMSE of the model and the RMSE of the autoregressive benchmark model. "Equal" and "Bayes" stand for the equally weighted-average model and average model based on the posterior probabilities, respectively. The results are based on the Occam Window and the Markov Chain Monte Carlo Model Composite (MC3) methods.

Table 14: Bayesian Model Averaging using the Markov Chain Monte Carlo Model Composite method: UK evidence for a selection of 19 models taken from the empirical finance literature.

Model	CRR_1986	C_1987; F_1990	H_1989	FH_1991	FH_1993
RMSE	0.936	0.962	0.913	0.912	0.943
Model vs. Constant	0.967	0.981	0.955	0.955	0.971
Model vs. AR1	0.946	0.959	0.934	0.933	0.949
$\bar{R}^2$	0.000	0.003	0.054	0.044	0.000
post prob	0.001	0.007	0.001	0.001	0.001
Model	W_1994; PS_1998; FH_1999	PT_1995	JS_2007a	JS_2007b	BH_1999
RMSE	0.916	0.869	0.969	0.949	0.880
Model vs. Constant	0.957	0.932	0.985	0.985	0.938
Model vs. AR1	0.935	0.911	0.962	0.962	0.917
$\bar{R}^2$	0.063	0.078	0.010	0.030	0.077
post prob	0.017	0.004	0.100	0.258	0.004
Model	C-CAPM	S_2010b	LL_2001	S_2010a	PJ_2005
RMSE	0.983	0.962	0.985	0.970	0.999
Model vs. Constant	0.985	0.981	0.992	0.985	1.000
Model vs. AR1	0.962	0.959	0.970	0.963	0.977
$\bar{R}^2$	0.006	0.027	0.004	0.019	0.000
post prob	0.145	0.014	0.087	0.176	0.024
Model	LvN_2005	SV_2006	Y_2006; PST_2007		
RMSE	0.959	0.988	0.987		
Model vs. Constant	0.979	0.994	0.993		
Model vs. AR1	0.957	0.972	0.971		
$\bar{R}^2$	0.030	0.000	0.002		
post prob	0.008	0.052	0.102		

Note: "post prob" denotes the posterior probability and RMSE corresponds to the root mean-squared error. "Model vs. Constant" is the ratio of the RMSE of the model and the RMSE of the constant expected returns benchmark model. "Model vs AR1" is the ratio of the RMSE of the model and the RMSE of the autoregressive benchmark model. $\bar{R}^2$ stands for the adjusted-R^2 statistic. All results are based on the Markov Chain Monte Carlo Model Composite (MC3) method.

Table 15: In-sample performance of the averaged model vis-a-vis benchmark models: UK evidence.

weighted-average model	Occam Window + Equal	Occam Window + Bayes
RMSE	0.753	0.762
Model vs. Constant	0.868	0.873
Model vs. AR1	0.848	0.853
Model	MCMC + Equal	MCMC + Bayes
RMSE	0.912	0.950
Model vs. Constant	0.955	0.974
Model vs. AR1	0.934	0.953

Note: RMSE corresponds to the root mean-squared error. "Model vs. Constant" is the ratio of the RMSE of the model and the RMSE of the constant expected returns benchmark model. "Model vs AR1" is the ratio of the RMSE of the model and the RMSE of the autoregressive benchmark model. "Equal" and "Bayes" stand for the equally weighted-average model and average model based on the posterior probabilities, respectively. The results are based on the Occam Window and the Markov Chain Monte Carlo Model Composite (MC3) methods.

Table 16: Bayesian Model Averaging: EA evidence for different model priors.

	Fixed	Random	Uniform	Custom
	PIP	PIP	PIP	PIP
ΔC_{t-1}	0.195	0.034	0.198	0.198
ΔC_{t-12}	0.118	0.040	0.124	0.119
r_{t-1}	0.984	0.882	0.984	0.987
$bond_{t-1}$	0.403	0.084	0.379	0.404
$\Delta bond_{t-1}$	0.163	0.037	0.175	0.165
og_{t-1}	0.242	0.211	0.248	0.240
π_{t-1}	0.413	0.043	0.389	0.411
$\Delta\pi_{t-1}$	0.322	0.035	0.302	0.325
Δi_{t-1}	0.124	0.028	0.124	0.134
Δm_{t-1}	0.117	0.018	0.097	0.108
Δhp_{t-1}	0.111	0.028	0.122	0.128
Δe_{t-1}	0.113	0.026	0.115	0.108
Δcp_{t-1}	0.853	0.336	0.854	0.864
$cday_{t-1}$	0.390	0.034	0.350	0.390
$\Delta cday_{t-1}$	0.136	0.024	0.145	0.137
lc_{t-1}	0.551	0.076	0.539	0.553
rwy_{t-1}	0.173	0.029	0.169	0.158
wy_{t-1}	0.157	0.027	0.160	0.158
$divyld_{t-1}$	0.268	0.171	0.300	0.283
$spgdp_{t-1}$	0.735	0.171	0.696	0.709
$\Delta cred_{t-1}$	0.133	0.059	0.157	0.142
AvnVar	6.701	2.393	6.629	6.722
# Models visited	29525	13675	29941	29607
Corr PMP	0.986	0.999	0.974	0.985

Note: AvnVar denotes average number of regressors, PIP corresponds to the posterior inclusion probability, PMP refers to posterior model probability.

Table 17: Bayesian Model Averaging: US evidence for different model priors.

	Fixed	Random	Uniform	PIP
	PIP	PIP	PIP	PIP
ΔC_{t-1}	0.087	0.005	0.089	0.077
ΔC_{t-12}	0.096	0.007	0.099	0.095
r_{t-1}	0.100	0.006	0.089	0.097
$bond_{t-1}$	0.080	0.004	0.090	0.102
$\Delta bond_{t-1}$	0.554	0.073	0.549	0.538
og_{t-1}	0.106	0.012	0.107	0.107
π_{t-1}	0.111	0.007	0.123	0.109
$\Delta\pi_{t-1}$	0.359	0.027	0.363	0.373
Δi_{t-1}	0.104	0.008	0.097	0.110
Δm_{t-1}	0.113	0.003	0.108	0.111
Δhp_{t-1}	0.078	0.005	0.099	0.086
Δe_{t-1}	0.093	0.005	0.081	0.085
Δcp_{t-1}	0.090	0.006	0.089	0.090
$cday_{t-1}$	0.681	0.150	0.663	0.675
$\Delta cday_{t-1}$	0.101	0.011	0.095	0.101
lc_{t-1}	0.163	0.007	0.187	0.169
rwy_{t-1}	0.101	0.004	0.110	0.106
wy_{t-1}	0.178	0.013	0.161	0.158
$divyld_{t-1}$	0.141	0.009	0.154	0.153
$spgdp_{t-1}$	0.177	0.009	0.189	0.191
Δu_{t-1}	0.118	0.006	0.115	0.114
φ_{t-1}	0.104	0.007	0.099	0.094
$\Delta cred_{t-1}$	0.116	0.009	0.135	0.123
$aSBRDLR_{t-1}$	0.105	0.009	0.117	0.128
AvnVar	3.956	0.402	4.008	3.992
# Models visited	27663	3555	28224	28199
Corr PMP	0.986	1.000	0.987	0.992

Note: AvnVar denotes average number of regressors, PIP corresponds to the posterior inclusion probability, PMP refers to posterior model probability.

Table 18: Bayesian Model Averaging: UK evidence for different model priors.

	Fixed	Random	Uniform	PIP
	PIP	PIP	PIP	PIP
ΔC_{t-1}	0.267	0.017	0.281	0.264
ΔC_{t-12}	0.161	0.006	0.153	0.150
r_{t-1}	0.138	0.008	0.158	0.132
$bond_{t-1}$	0.556	0.024	0.560	0.529
$\Delta bond_{t-1}$	0.561	0.019	0.554	0.536
og_{t-1}	0.217	0.025	0.215	0.214
π_{t-1}	0.106	0.007	0.102	0.118
$\Delta\pi_{t-1}$	0.105	0.007	0.099	0.115
Δi_{t-1}	0.138	0.010	0.146	0.137
Δm_{t-1}	0.135	0.010	0.119	0.136
Δhp_{t-1}	0.131	0.011	0.130	0.134
Δe_{t-1}	0.194	0.024	0.209	0.208
Δcp_{t-1}	0.114	0.009	0.117	0.133
$cday_{t-1}$	0.262	0.034	0.245	0.242
$\Delta cday_{t-1}$	0.126	0.005	0.127	0.135
lc_{t-1}	0.143	0.017	0.151	0.144
rwy_{t-1}	0.318	0.037	0.318	0.321
wy_{t-1}	0.442	0.036	0.447	0.438
$divyld_{t-1}$	0.769	0.138	0.760	0.744
$spgdp_{t-1}$	0.206	0.063	0.220	0.229
Δu_{t-1}	0.103	0.012	0.106	0.111
φ_{t-1}	0.116	0.011	0.121	0.125
$\Delta cred_{t-1}$	0.140	0.013	0.144	0.141
AvnVar	5.444	0.542	5.480	5.436
# Models visited	30246	4557	30440	30883
Corr PMP	0.979	1.000	0.981	0.976

Note: AvnVar denotes average number of regressors, PIP corresponds to the posterior inclusion probability, PMP refers to posterior model probability.

Table 19: Bayesian Model Averaging: EA evidence for different Zellner g-priors.

	UIP	BRIC	RIC	HQ	EBL	hyper
	PIP	PIP	PIP	PIP	PIP	PIP
ΔC_{t-1}	0.037	0.007	0.007	0.027	0.745	0.714
ΔC_{t-12}	0.039	0.019	0.017	0.047	0.601	0.575
r_{t-1}	0.899	0.788	0.784	0.891	0.969	0.967
$bond_{t-1}$	0.102	0.020	0.025	0.085	0.682	0.644
$\Delta bond_{t-1}$	0.041	0.016	0.014	0.035	0.682	0.658
og_{t-1}	0.192	0.115	0.127	0.211	0.630	0.605
π_{t-1}	0.046	0.011	0.007	0.036	0.862	0.850
$\Delta\pi_{t-1}$	0.035	0.010	0.010	0.042	0.774	0.761
Δi_{t-1}	0.025	0.006	0.007	0.022	0.612	0.580
Δm_{t-1}	0.019	0.003	0.006	0.015	0.613	0.591
Δhp_{t-1}	0.032	0.021	0.014	0.038	0.666	0.634
Δe_{t-1}	0.024	0.007	0.008	0.024	0.585	0.560
Δcp_{t-1}	0.338	0.139	0.140	0.328	0.950	0.935
$cday_{t-1}$	0.032	0.008	0.007	0.032	0.842	0.818
$\Delta cday_{t-1}$	0.026	0.007	0.009	0.019	0.648	0.613
lc_{t-1}	0.087	0.011	0.015	0.070	0.867	0.848
rwy_{t-1}	0.026	0.009	0.005	0.030	0.645	0.617
wy_{t-1}	0.032	0.007	0.010	0.026	0.660	0.626
$divyld_{t-1}$	0.185	0.081	0.086	0.167	0.673	0.639
$spgdp_{t-1}$	0.211	0.068	0.071	0.187	0.886	0.869
$\Delta cred_{t-1}$	0.063	0.032	0.026	0.059	0.646	0.596
AvnVar	2.493	1.385	1.395	2.389	15.239	14.701
# Models visited	13739	7615	7906	13650	43078	44814
Corr PMP	0.999	1.000	0.999	0.999	0.998	0.997

Note: AvnVar denotes average number of regressors, PIP corresponds to the posterior inclusion probability, PMP refers to posterior model probability.

Table 20: Bayesian Model Averaging: US evidence for different Zellner g-priors.

	UIP	BRIC	RIC	HQ	EBL	hyper
	PIP	PIP	PIP	PIP	PIP	PIP
ΔC_{t-1}	0.007	0.003	0.002	0.006	0.437	0.102
ΔC_{t-12}	0.006	0.003	0.002	0.005	0.446	0.105
r_{t-1}	0.004	0.002	0.003	0.005	0.430	0.099
$bond_{t-1}$	0.005	0.001	0.002	0.006	0.437	0.091
$\Delta bond_{t-1}$	0.068	0.035	0.028	0.074	0.531	0.145
og_{t-1}	0.006	0.004	0.008	0.010	0.438	0.100
π_{t-1}	0.005	0.001	0.003	0.006	0.437	0.100
$\Delta\pi_{t-1}$	0.024	0.006	0.005	0.020	0.507	0.132
Δi_{t-1}	0.007	0.002	0.002	0.009	0.430	0.101
Δm_{t-1}	0.005	0.003	0.003	0.011	0.449	0.100
Δhp_{t-1}	0.006	0.002	0.004	0.007	0.425	0.095
Δe_{t-1}	0.007	0.001	0.004	0.003	0.440	0.095
Δcp_{t-1}	0.004	0.002	0.003	0.006	0.431	0.098
$cday_{t-1}$	0.158	0.079	0.075	0.184	0.553	0.150
$\Delta cday_{t-1}$	0.010	0.002	0.002	0.007	0.427	0.094
lc_{t-1}	0.008	0.002	0.003	0.005	0.485	0.124
rwy_{t-1}	0.003	0.002	0.004	0.004	0.456	0.104
wy_{t-1}	0.020	0.006	0.006	0.015	0.443	0.107
$divyld_{t-1}$	0.010	0.003	0.003	0.005	0.462	0.111
$spgdp_{t-1}$	0.014	0.005	0.004	0.012	0.457	0.108
Δu_{t-1}	0.005	0.002	0.002	0.007	0.451	0.099
φ_{t-1}	0.007	0.004	0.002	0.009	0.435	0.095
$\Delta cred_{t-1}$	0.009	0.003	0.005	0.010	0.438	0.100
$aSBRDLR_{t-1}$	0.012	0.005	0.005	0.008	0.445	0.095
AvnVar	0.410	0.175	0.179	0.432	10.889	2.549
# Models visited	3448	1539	1463	3645	51483	13683
Corr PMP	1.000	1.000	1.000	1.000	0.906	1.000

Note: AvnVar denotes average number of regressors, PIP corresponds to the posterior inclusion probability, PMP refers to posterior model probability.

Table 21: Bayesian Model Averaging: UK evidence for different Zellner g-priors.

	UIP	BRIC	RIC	HQ	EBL	hyper
	PIP	PIP	PIP	PIP	PIP	PIP
ΔC_{t-1}	0.021	0.004	0.005	0.020	0.620	0.205
ΔC_{t-12}	0.009	0.002	0.004	0.013	0.552	0.189
r_{t-1}	0.009	0.002	0.002	0.007	0.552	0.181
$bond_{t-1}$	0.030	0.007	0.011	0.024	0.664	0.241
$\Delta bond_{t-1}$	0.020	0.009	0.008	0.022	0.698	0.236
og_{t-1}	0.019	0.005	0.005	0.017	0.597	0.200
π_{t-1}	0.010	0.003	0.003	0.007	0.537	0.170
$\Delta\pi_{t-1}$	0.010	0.002	0.003	0.006	0.551	0.179
Δi_{t-1}	0.008	0.002	0.002	0.012	0.561	0.185
Δm_{t-1}	0.007	0.004	0.003	0.008	0.572	0.185
Δhp_{t-1}	0.010	0.003	0.005	0.007	0.547	0.184
Δe_{t-1}	0.020	0.006	0.004	0.022	0.560	0.182
Δcp_{t-1}	0.008	0.004	0.004	0.007	0.523	0.170
$cday_{t-1}$	0.034	0.011	0.009	0.031	0.568	0.191
$\Delta cday_{t-1}$	0.008	0.003	0.004	0.008	0.537	0.170
lc_{t-1}	0.019	0.003	0.004	0.012	0.536	0.179
rwy_{t-1}	0.047	0.018	0.012	0.043	0.587	0.194
wy_{t-1}	0.037	0.012	0.013	0.031	0.581	0.204
$divyld_{t-1}$	0.140	0.058	0.055	0.131	0.749	0.278
$spgdp_{t-1}$	0.067	0.028	0.027	0.073	0.558	0.189
Δu_{t-1}	0.014	0.002	0.003	0.014	0.528	0.163
φ_{t-1}	0.012	0.004	0.003	0.013	0.525	0.172
$\Delta cred_{t-1}$	0.009	0.002	0.003	0.009	0.578	0.194
AvnVar	0.569	0.195	0.193	0.535	13.282	4.440
# Models visited	4708	1665	1629	4597	51905	19566
Corr PMP	1.000	1.000	1.000	1.000	0.997	1.000

Note: AvnVar denotes average number of regressors, PIP corresponds to the posterior inclusion probability, PMP refers to posterior model probability.

Table 22: In-sample performance of the averaged model at different horizons: EA evidence.

	$H = 1$		$H = 2$	
Model	MCMC + Equal	MCMC + Bayes	MCMC + Equal	MCMC + Bayes
RMSE	0.863	0.789	0.857	0.777
Model vs. Constant	0.929	0.888	0.926	0.882
Model vs. AR1	0.973	0.930	0.911	0.867
	$H = 3$		$H = 4$	
Model	MCMC + Equal	MCMC + Bayes	MCMC + Equal	MCMC + Bayes
RMSE	0.815	0.709	0.797	0.659
Model vs. Constant	0.903	0.842	0.893	0.812
Model vs. AR1	0.877	0.818	0.853	0.776
		$H = 8$		
Model		MCMC + Equal	MCMC + Bayes	
RMSE		0.718	0.528	
Model vs. Constant		0.847	0.727	
Model vs. AR1		0.738	0.633	

Note: RMSE corresponds to the root mean-squared error. "Model vs. Constant" is the ratio of the RMSE of the model and the RMSE of the constant expected returns benchmark model. "Model vs AR1" is the ratio of the RMSE of the model and the RMSE of the autoregressive benchmark model. "Equal" and "Bayes" stand for the equally weighted-average model and average model based on the posterior probabilities, respectively. The results are based on the Markov Chain Monte Carlo Model Composite (MC3) method.

Table 23: In-sample performance of the averaged model at different horizons: US evidence.

	$H = 1$		$H = 2$	
Model	MCMC + Equal	MCMC + Bayes	MCMC + Equal	MCMC + Bayes
RMSE	0.951	0.943	0.917	0.902
Model vs. Constant	0.975	0.971	0.958	0.950
Model vs. AR1	0.944	0.940	0.905	0.897
	$H = 3$		$H = 4$	
Model	MCMC + Equal	MCMC + Bayes	MCMC + Equal	MCMC + Bayes
RMSE	0.893	0.872	0.858	0.825
Model vs. Constant	0.945	0.934	0.927	0.908
Model vs. AR1	0.876	0.865	0.847	0.830
		$H = 8$		
Model		MCMC + Equal	MCMC + Bayes	
RMSE		0.753	0.621	
Model vs. Constant		0.868	0.788	
Model vs. AR1		0.736	0.668	

Note: RMSE corresponds to the root mean-squared error. "Model vs. Constant" is the ratio of the RMSE of the model and the RMSE of the constant expected returns benchmark model. "Model vs AR1" is the ratio of the RMSE of the model and the RMSE of the autoregressive benchmark model. "Equal" and "Bayes" stand for the equally weighted-average model and average model based on the posterior probabilities, respectively. The results are based on the Markov Chain Monte Carlo Model Composite (MC3) method.

Table 24: In-sample performance of the averaged model at different horizons: UK evidence.

	$H = 1$		$H = 2$	
Model	MCMC + Equal	MCMC + Bayes	MCMC + Equal	MCMC + Bayes
RMSE	0.912	0.950	0.842	0.813
Model vs. Constant	0.955	0.974	0.918	0.902
Model vs. AR1	0.934	0.953	0.875	0.859
	$H = 3$		$H = 4$	
Model	MCMC + Equal	MCMC + Bayes	MCMC + Equal	MCMC + Bayes
RMSE	0.787	0.709	0.753	0.651
Model vs. Constant	0.887	0.842	0.868	0.807
Model vs. AR1	0.820	0.778	0.779	0.725
		$H = 8$		
Model		MCMC + Equal	MCMC + Bayes	
RMSE		0.679	0.527	
Model vs. Constant		0.824	0.726	
Model vs. AR1		0.685	0.604	

Note: RMSE corresponds to the root mean-squared error. "Model vs. Constant" is the ratio of the RMSE of the model and the RMSE of the constant expected returns benchmark model. "Model vs AR1" is the ratio of the RMSE of the model and the RMSE of the autoregressive benchmark model. "Equal" and "Bayes" stand for the equally weighted-average model and average model based on the posterior probabilities, respectively. The results are based on the Markov Chain Monte Carlo Model Composite (MC3) method.

Table 25: Out-of-sample performance of the averaged model at different horizons: EA evidence.

	$H = 1$		$H = 2$	
Model	MCMC + Equal	MCMC + Bayes	MCMC + Equal	MCMC + Bayes
RMSE	0.931	0.935	0.932	0.962
Model vs. Constant	0.946	0.947	0.954	0.969
Model vs. AR1	0.990	0.992	0.938	0.953
	$H = 3$		$H = 4$	
Model	MCMC + Equal	MCMC + Bayes	MCMC + Equal	MCMC + Bayes
RMSE	0.901	0.885	0.909	0.865
Model vs. Constant	0.959	0.951	0.973	0.950
Model vs. AR1	0.932	0.924	0.930	0.907
		$H = 8$		
Model		MCMC + Equal	MCMC + Bayes	
RMSE		1.015	0.768	
Model vs. Constant		1.087	0.945	
Model vs. AR1		0.947	0.823	

Note: RMSE corresponds to the root mean-squared error. "Model vs. Constant" is the ratio of the RMSE of the model and the RMSE of the constant expected returns benchmark model. "Model vs AR1" is the ratio of the RMSE of the model and the RMSE of the autoregressive benchmark model. "Equal" and "Bayes" stand for the equally weighted-average model and average model based on the posterior probabilities, respectively. The results are based on the Markov Chain Monte Carlo Model Composite (MC3) method. The out-of-sample forecast period corresponds to the last 10 years of available data.

Table 26: Out-of-sample performance of the averaged model at different horizons: US evidence.

	$H = 1$		$H = 2$	
Model	MCMC + Equal	MCMC + Bayes	MCMC + Equal	MCMC + Bayes
RMSE	0.960	0.975	0.954	1.062
Model vs. Constant	0.974	0.982	0.944	0.997
Model vs. AR1	0.943	0.951	0.892	0.941
	$H = 3$		$H = 4$	
Model	MCMC + Equal	MCMC + Bayes	MCMC + Equal	MCMC + Bayes
RMSE	0.954	1.083	0.944	1.310
Model vs. Constant	0.971	1.034	1.017	1.198
Model vs. AR1	0.900	0.959	0.929	1.095
		$H = 8$		
Model		MCMC + Equal	MCMC + Bayes	
RMSE		0.922	1.190	
Model vs. Constant		1.134	1.288	
Model vs. AR1		0.961	1.092	

Note: RMSE corresponds to the root mean-squared error. "Model vs. Constant" is the ratio of the RMSE of the model and the RMSE of the constant expected returns benchmark model. "Model vs AR1" is the ratio of the RMSE of the model and the RMSE of the autoregressive benchmark model. "Equal" and "Bayes" stand for the equally weighted-average model and average model based on the posterior probabilities, respectively. The results are based on the Markov Chain Monte Carlo Model Composite (MC3) method. The out-of-sample forecast period corresponds to the last 10 years of available data.

Table 27: Out-of-sample performance of the averaged model at different horizons: UK evidence.

	$H = 1$		$H = 2$	
Model	MCMC + Equal	MCMC + Bayes	MCMC + Equal	MCMC + Bayes
RMSE	0.917	1.001	0.840	0.767
Model vs. Constant	0.888	0.928	0.846	0.809
Model vs. AR1	0.868	0.907	0.806	0.771
	$H = 3$		$H = 4$	
Model	MCMC + Equal	MCMC + Bayes	MCMC + Equal	MCMC + Bayes
RMSE	0.780	1.228	0.781	1.245
Model vs. Constant	0.860	1.079	0.919	1.160
Model vs. AR1	0.795	0.997	0.825	1.042
		$H = 8$		
Model		MCMC + Equal	MCMC + Bayes	
RMSE		0.811	0.760	
Model vs. Constant		1.057	1.023	
Model vs. AR1		0.879	0.851	

Note: RMSE corresponds to the root mean-squared error. "Model vs. Constant" is the ratio of the RMSE of the model and the RMSE of the constant expected returns benchmark model. "Model vs AR1" is the ratio of the RMSE of the model and the RMSE of the autoregressive benchmark model. "Equal" and "Bayes" stand for the equally weighted-average model and average model based on the posterior probabilities, respectively. The results are based on the Markov Chain Monte Carlo Model Composite (MC3) method. The out-of-sample forecast period corresponds to the last 10 years of available data.

List of Figures

Figure 1: Bayesian Model Averaging: Model inclusion - EA evidence.

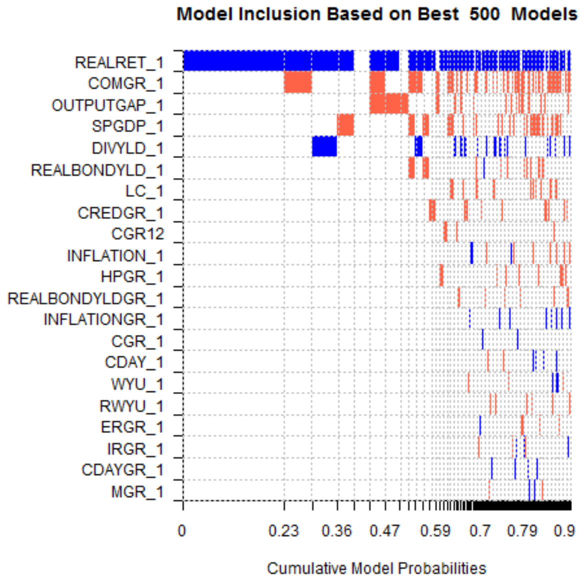

Figure 2: Bayesian Model Averaging: Model inclusion - US evidence.

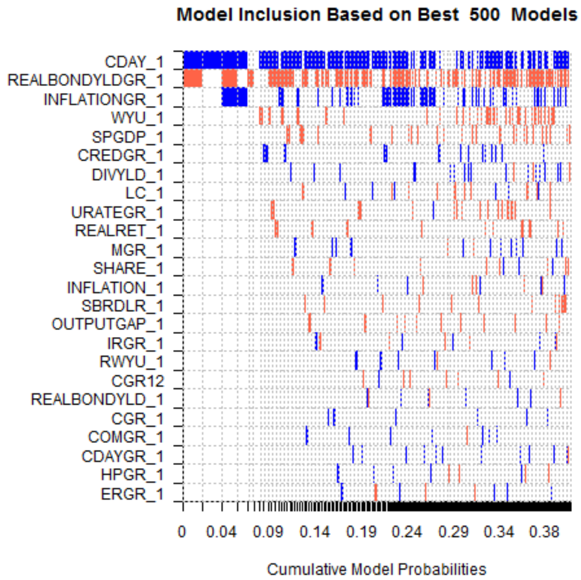

Figure 3: Bayesian Model Averaging: Model inclusion - UK evidence.

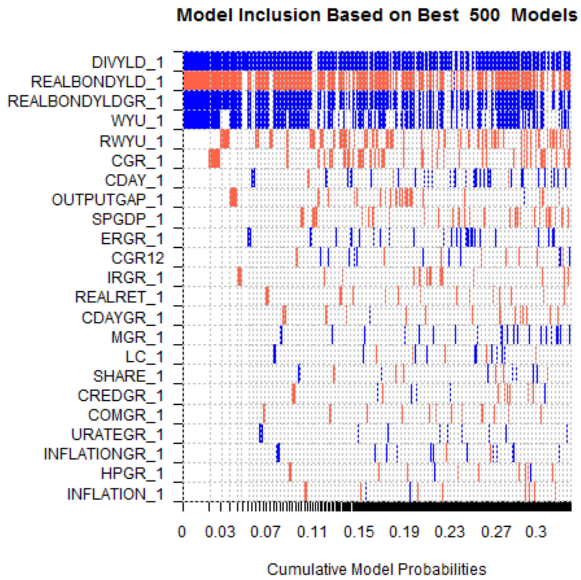

Figure 4: Bayesian Model Averaging: Recursive posterior probabilities - EA evidence.

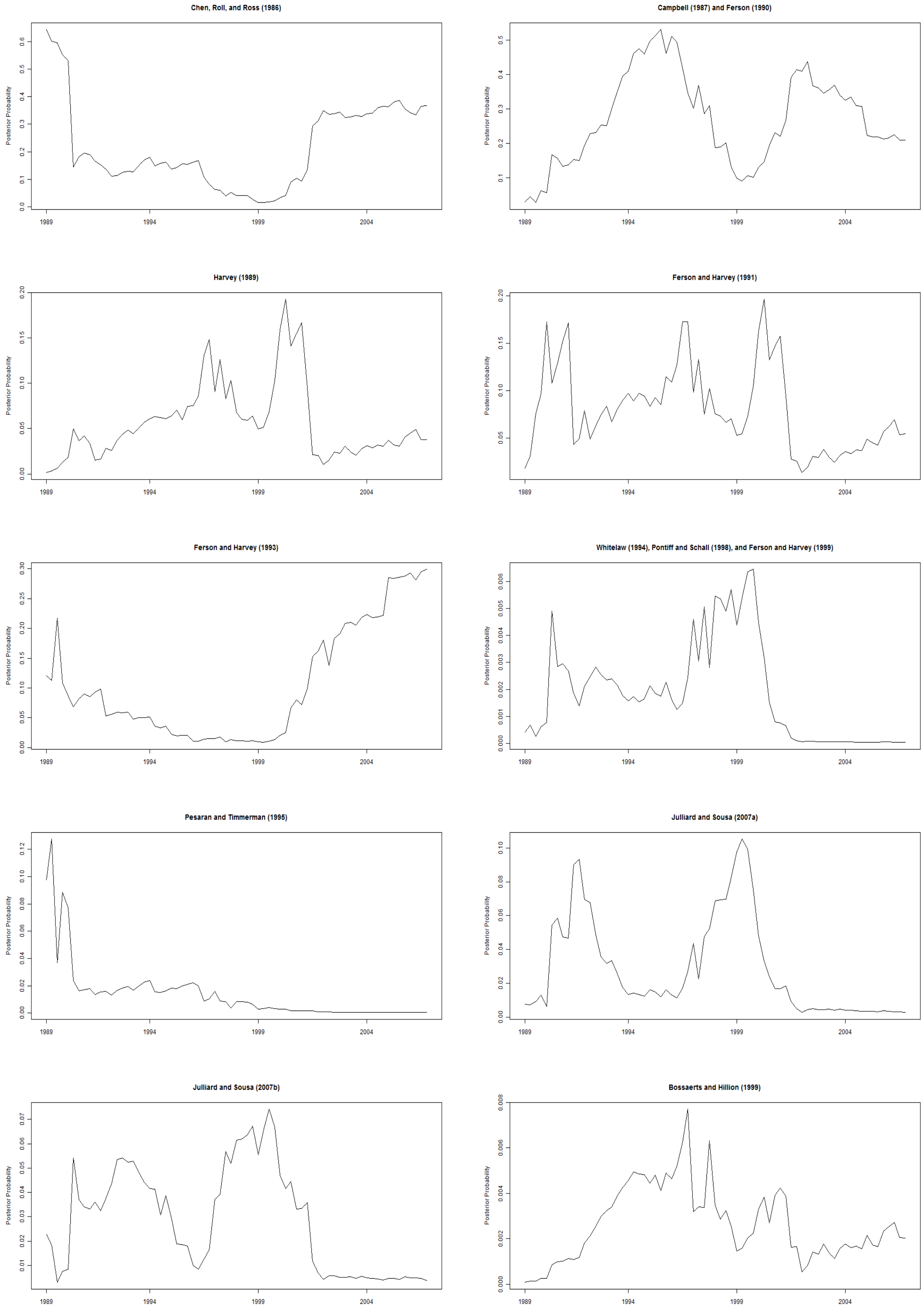

Figure **5**: Bayesian Model Averaging: Recursive posterior probabilities - EA evidence (cont.).

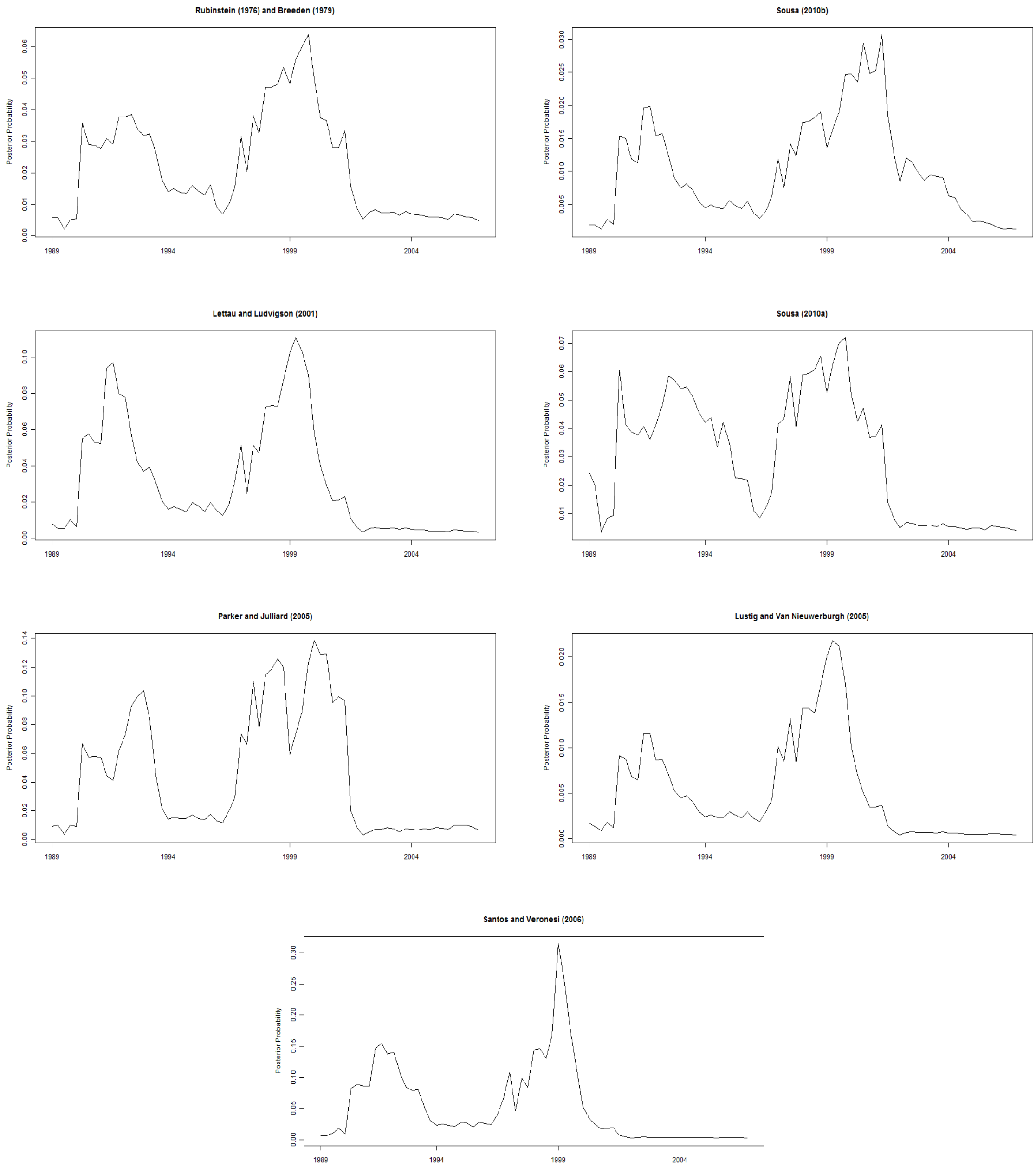

Figure **6**: Bayesian Model Averaging: Recursive posterior probabilities - US evidence.

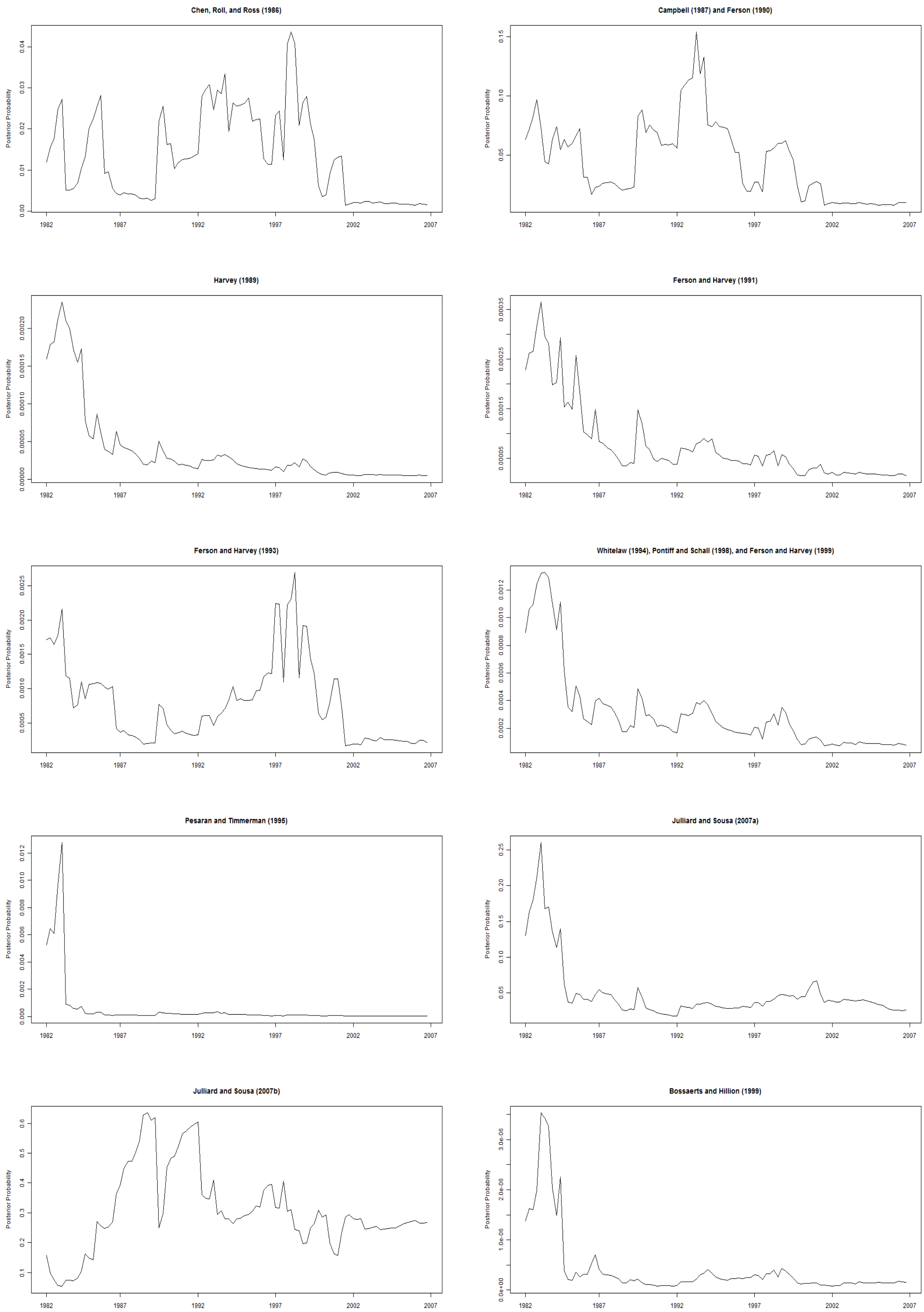

Figure **7**: Bayesian Model Averaging: Recursive posterior probabilities - US evidence (cont.).

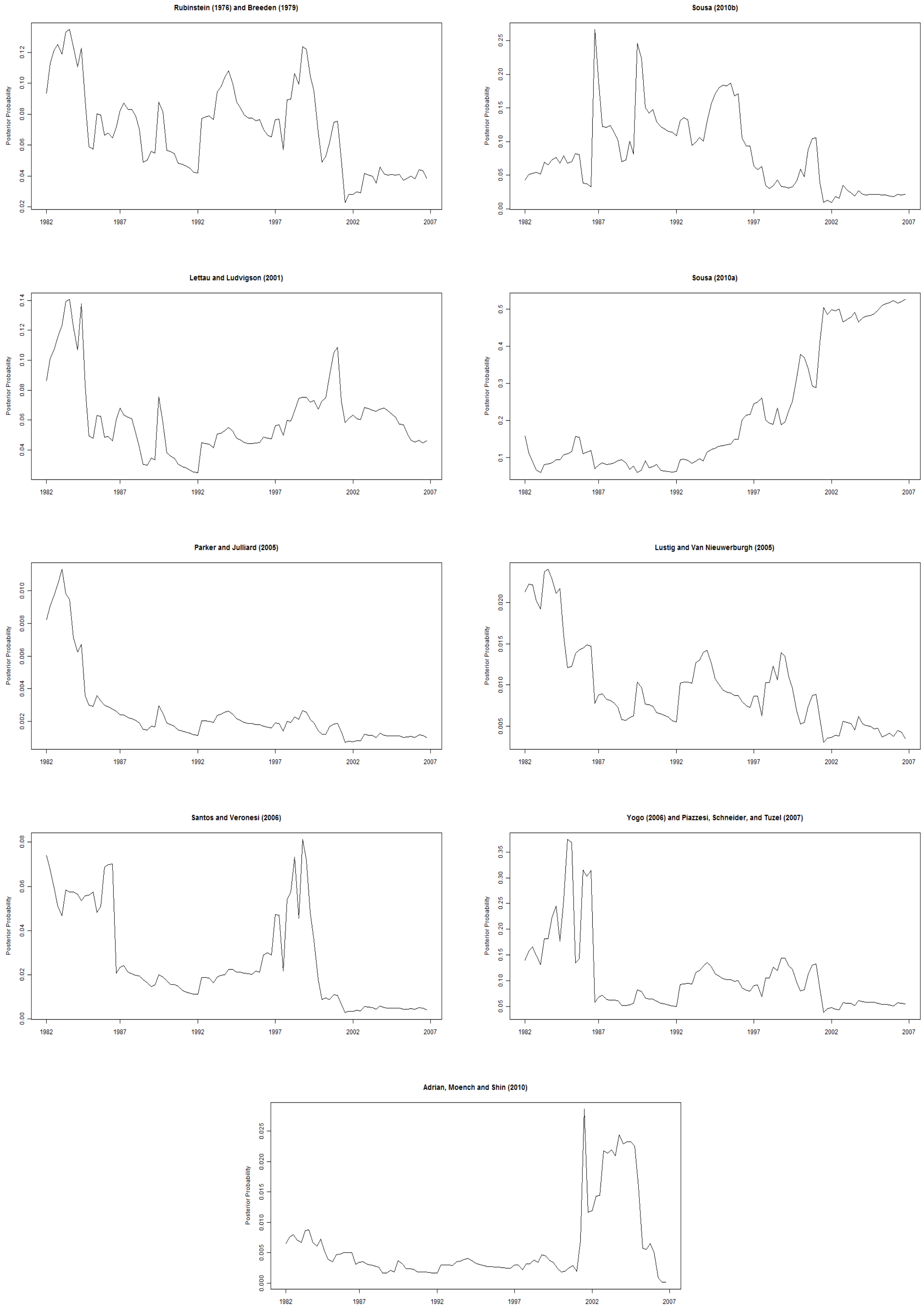

Figure **8**: Bayesian Model Averaging: Recursive posterior probabilities - UK evidence.

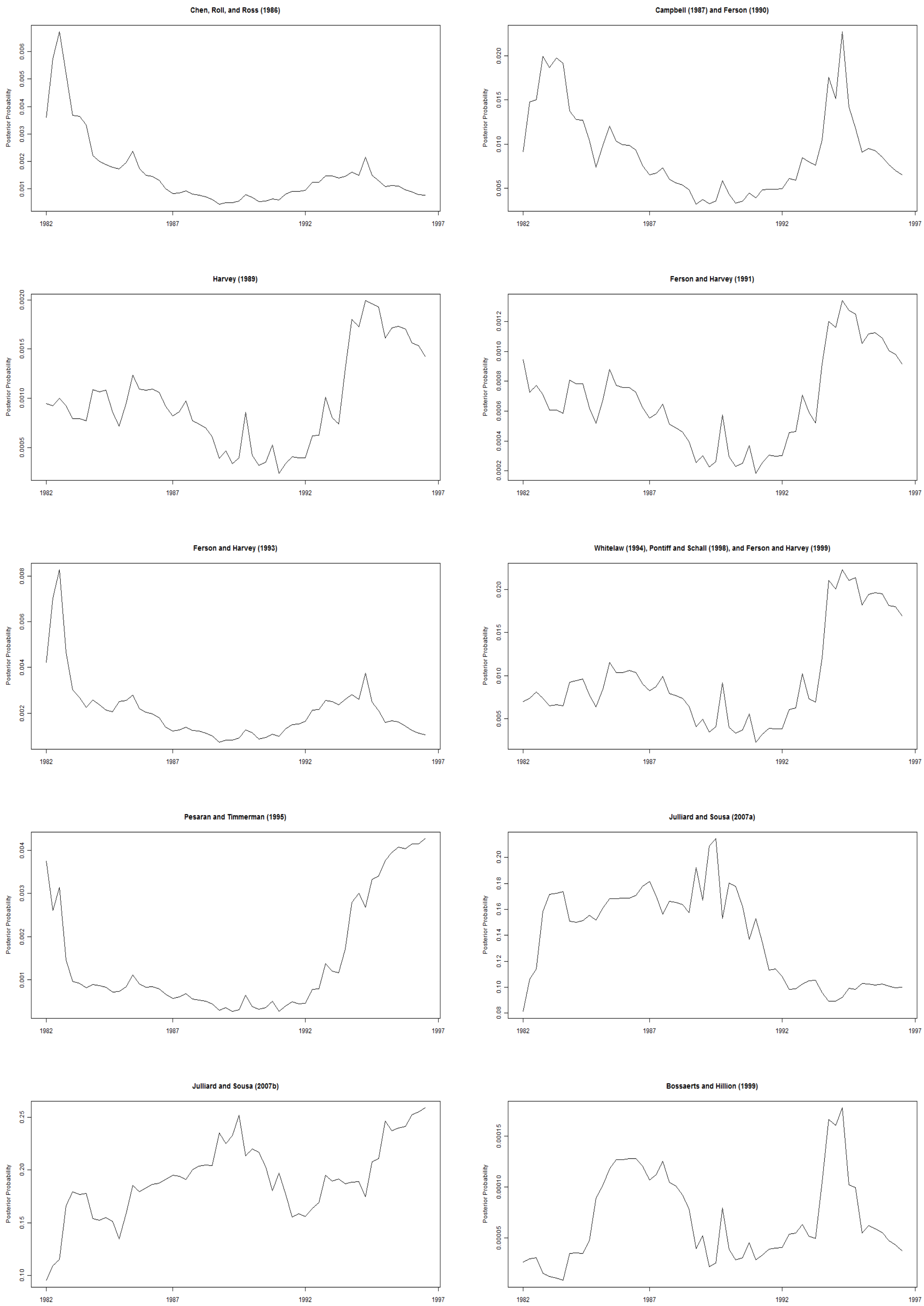

Figure **9**: Bayesian Model Averaging: Recursive posterior probabilities - UK evidence (cont.).