scottish institute for research in economics

SIRE DISCUSSION PAPER

SIRE-DP-2011-39

UK MACROECONOMIC FORECASTING WITH MANY PREDICTORS: WHICH MODELS FORECAST BEST AND WHEN DO THEY DO SO?

GARY KOOP AND DIMITRIS KOROBILIS
University of Strathclyde

www.sire.ac.uk

UK Macroeconomic Forecasting with Many Predictors: Which Models Forecast Best and When Do They Do So?*

Gary Koop
University of Strathclyde

Dimitris Korobilis
University of Strathclyde

July 29, 2009

ABSTRACT

Block factor methods offer an attractive approach to forecasting with many predictors. These extract the information in these predictors into factors reflecting different blocks of variables (e.g. a price block, a housing block, a financial block, etc.). However, a forecasting model which simply includes all blocks as predictors risks being over-parameterized. Thus, it is desirable to use a methodology which allows for different parsimonious forecasting models to hold at different points in time. In this paper, we use dynamic model averaging and dynamic model selection to achieve this goal. These methods automatically alter the weights attached to different forecasting model as evidence comes in about which has forecast well in the recent past. In an empirical study involving forecasting output and inflation using 139 UK monthly time series variables, we find that the set of predictors changes substantially over time. Furthermore, our results show that dynamic model averaging and model selection can greatly improve forecast performance relative to traditional forecasting methods. *Keywords:* Bayesian, state space model, factor model, dynamic model averaging *JEL Classification*: E31, E37, C11, C53

*Both authors are Fellows at the Rimini Centre for Economic Analysis. We would like to thank George Kapetanios and the Bank of England Econometrics group for their help and the provision of the data.

1 Introduction

Macroeconomists interested in forecasting variables such as output growth and inflation often have many potential predictors. For instance, Stock and Watson (2002) forecast various US macroeconomic variables using up to 215 predictors. Since the pioneering work of Geweke (1977), dynamic factor models have typically been used to deal with the problems caused by this proliferation of predictors. These models extract the common fluctuations in the large number of predictors into a much smaller number of factors. Stock and Watson (2006) is a recent survey paper on forecasting using dynamic factor models. In the UK, Kapetanios, Labbard and Price (2009) discuss the Bank of England's suite of forecasting models, including a discussion of dynamic factor methods.[1] Dynamic factor methods involve a forecasting model of the form (see, e.g., Stock and Watson, 2002):

$$y_t = \gamma(L)\, y_t + \beta(L)\, f_t + \varepsilon_t, \tag{1}$$

where f_t is an $q-$vector of factors, $\gamma(L) = \gamma_1 L + .. + \gamma_p L^p$ and $\beta(L) = \beta_1 L + .. + \beta_p L^p$ are polynomials in the lag operator.[2]

Dynamic factor models have been used successfully in a wide variety of forecasting exercises. However, they suffer from the drawback that coefficients cannot easily be interpreted in terms of the underlying macroeconomic variables. For this reason, dynamic factor models with a block structure have recently been proposed (e.g. Ng, Moench and Potter, 2008). These are motivated by the observation that the large panels of macroeconomic time series available for forecasting often fall naturally into different groups or blocks. For instance, the researcher might have several different measures of prices (the price block), several different measures of demand (the demand block), different interest rates and stock prices (the financial block), etc. In such cases, factor methods can be used on the variables in each block separately, resulting in a set of factors each of which has a macroeconomic interpretation (e.g. the price factor, the demand factor, the financial factor, etc.). In this paper, we adopt such a blocking of variables using the 139 monthly variables (1990:1 through 2008:11) from the Bank of England's forecasting suite. When (1) is extended to allow for block factors, then the forecasting equation takes the form:

$$y_t = \gamma(L)\, y_t + \sum_{b=1}^{B} \beta^{(b)}(L)\, f_t^{(b)} + \varepsilon_t, \tag{2}$$

where $f_t^{(b)}$ for $b = 1, .., B$ denotes the factor extracted from the b^{th} block of predictors.

A drawback in all the models discussed so far is that they assume that parameters are constant. To some extent, the use of recursive or rolling forecasting methods can account for time variation in coefficients. But recent research (e.g. Groen, Paap and Ravazzolo, 2008) argues strongly that it is better to build a formal model of time variation in parameters rather to than rely of recursive or rolling methods to accurately pick up parameter shifts. Such time-varying parameter (TVP) models are commonly estimated using state space methods (see, e.g., Cogley and Sargent, 2005, or Cogley, Morozov and Sargent, 2005). In our context, this would involves extending (2) as:

$$y_t = \gamma_t(L)\, y_t + \sum_{b=1}^{B} \beta_t^{(b)}(L)\, f_t^{(b)} + \varepsilon_t, \tag{3}$$

and adding a state equation to model the evolution of coefficients:

$$\theta_{t+1} = \theta_t + \eta_t \tag{4}$$

where θ_t is a vector containing all the coefficients in $\alpha_t(L)$ and $\beta_t^{(b)}$. As an example, a commonly expressed view of inflation is that its persistence has changed markedly since the 1970s. Since persistence depends on $\alpha_t(L)$, allowing for its time variation is potentially of great importance in developing a good model for forecasting inflation.

[1] See also Artis, Bannerjee and Macellino (2005) for an empirical forecasting exercise using UK data and Marcellino, Stock and Watson (2003) for a Euro area exercise.

[2] This model is used for one-step ahead forecasting. When forecasting h periods ahead the direct method can be used and $\gamma(L)$ and $\beta(L)$ re-defined so that the right-hand side includes variables lagged h periods.

However, even forecasting models such as that given by (3) and (4) suffer from the criticism that the same set of predictors is used at all points in time. Papers such as Pesaran and Timmermann (1995) present forecasting models where the set of predictors can change over time and show this to be of importance in helping to improve forecast performance. For instance, when forecasting inflation, it is possible that the predictors in the 1970s were different than in the 1990s. Or it is possible that the predictors in recessions are different than the predictors in expansion. We want a forecasting framework which allows for this.

This paper addresses all these issues. That is, we develop a forecasting model which begins with a block dynamic factor model such as (2), extends it to have time variation in coefficients through a model such as (3) and (4) but also allows for the forecasting model to change over time through a technique known as dynamic model averaging (DMA). This framework is applied in an exercise involving monthly UK data where we forecast output growth and inflation.

In terms of the econometrics, the major innovation is the use of DMA and we explain it in detail below. But the basic issues that must be addressed when doing DMA can be explained briefly as follows: In macroeconomic forecasting exercises such as the present one, the number of predictors can be very large. Even when reducing the set of predictors by extracting common factors and using constant coefficient models such as (2), we can still end up with a large number of potential models. For instance, even if we have a constant parameter model and the only predictors are the first factor from each of B blocks of factors (in our case we have $B = 8$) we will end up with 2^B models when we define each model by whether each factor is included or excluded from the forecasting model. If the researcher wants to select a single model from this huge set through sequential hypothesis testing procedures, she can potentially run into serious pre-test and data mining problems. If the researcher wishes to do model averaging, substantial computational problems arise (see Koop and Potter, 2004, for a discussion of these issues). When we want to allow for the forecasting model to change over time, then such computational problems increase hugely. With B potential predictors, the number of combinations of models which must be estimated in order to forecast through time τ is of the order $2^{B\tau}$. Estimating this many models when models are of the form given by (3) and (4) will typically be computationally infeasible. Accordingly, in this paper, we consider a strategy developed in the engineering literature by Raftery, Karny, Andrysek and Ettler (2007) which they refer to as dynamic model averaging or DMA (although, as discussed below, it can also be used for dynamic model selection or DMS). The method they propose seems ideally suited for our forecasting exercise since it satisfies all the desirable features outlined above. That is, it allows for the forecasting model to change over time (i.e. different predictors can be relevant at different times) while, at the same time, allowing for coefficients in each model to evolve over time. It involves only standard econometric methods for state space models such as the Kalman filter and simulation smoother but (via some empirically-sensible approximations) achieves vast gains in computational efficiency so as to allow DMA or DMS to be done in real time.

Using DMA and DMS in our block factor model with time varying coefficients, we find substantial improvements in forecast performance relative to alternative forecasting strategies. Furthemore, especially for inflation, the set of predictors in the best forecasting model varies substantially over time in a manner which conventional modelling strategies would miss.

2 Dynamic Model Averaging

To explain how DMA and DMS work, we begin by writing the block dynamic factor model with time varying coefficients given by (3) and (4) in standard state space form notation:

$$\begin{aligned} y_t &= Z_t\theta_t + \varepsilon_t \\ \theta_{t+1} &= \theta_t + \eta_t, \end{aligned} \tag{5}$$

where y_t is the dependent variable being forecast, Z_t is a $1 \times m$ vector of observations on explanatory variables that are available to forecast y_t. The discussion below relates to the case where we are forecasting one period in the future ($h = 1$) and, thus, Z_t will contain an intercept, $y_{t-1},..,y_{t-p_y+1}, f_{t-1}^{(b)},.., f_{t-p_f}^{(b)}$ (for $b = 1,..,B$) where p_y is the lag length for the dependent variable and p_f the lag length for the factors.

When forecasting $h > 1$ periods ahead, we use the direct method and variables are lagged appropriately (i.e. Z_t contains information lagged h periods).

We construct $f_t^{(b)}$ using principal components methods[3] involving all the variables in block b.[4] Thus, all elements of Z_t can be interpreted as exogenous or lagged dependent variables. θ_t is an $m \times 1$ vector of regression coefficients, ε_t is $N(0, H_t)$ and η_t is $N(0, Q_t)$. This is a state space model of the sort commonly used in empirical macroeconomics (see, e.g., among many others, Cogley and Sargent, 2005, Cogley, Morozov and Sargent, 2005, Primiceri, 2005). Standard methods (e.g. involving the Kalman filter and smoother) for estimation and prediction exist with such models (and are given in our Technical Appendix).

However, (5) assumes that the same explanatory variables can be used for forecasting at all points in time. In our introduction, we have explained why this might not be a good idea in a forecasting exercise. Furthermore, our empirical work will show that models such as (5) which simply maintain the same set of predictors in all time periods forecast very poorly due to over-parameterization problems. Accordingly, we allow for K models which are characterized by having different subsets of the explanatory variables, $Z_t^{(k)} \subseteq Z_t$ for $k = 1, .., K$,

$$\begin{aligned} y_t &= Z_t^{(k)} \theta_t^{(k)} + \varepsilon_t^{(k)} \\ \theta_{t+1}^{(k)} &= \theta_t^{(k)} + \eta_t^{(k)}, \end{aligned} \tag{6}$$

$\varepsilon_t^{(k)}$ is $N\left(0, H_t^{(k)}\right)$ and $\eta_t^{(k)}$ is $N\left(0, Q_t^{(k)}\right)$. The fact that we are letting different models hold at each point in time and will do model averaging justifies the terminology DMA. Alternatively, we can select the single best model at each point in time and do DMS.

We have now defined the set of models we will work with and, for each model, standard econometric methods can be used to forecast. However, to complete our algorithm, we need to specify how models evolve over time (i.e. we need a way of specifying how explanatory variables enter/leave the model in real time). To explain the issues relating to this let $L_t \in \{1, 2, .., K\}$ denote which model applies at time t. When the number of models is small, a natural specification would involve a transition matrix, P, with elements $p_{ij} = \Pr(L_t = i | L_{t-1} = j)$ for $i, j = 1, .., K$. That is, if model j holds at time $t-1$, then p_{ij} specifies the probability that the forecasting model will switch to model i at time t. Such Markov switching specifications are widely used in many contexts in econometrics and inference in such models is well-understood. The problem with our using such a specification in our case is that, when the number of models gets large, it becomes computationally infeasible. In our empirical work, we will have $K = 2^B$ with $B = 8$ and accordingly P will be a 256×256 matrix. Clearly, this will lead to imprecise inferences and very slow computation.[5] These are the reasons why DMA has not been done previously in substantive empirical problems in macroeconomics. Recently, Raftery, Karny, Andrysek and Ettler (2007) have proposed an approximate method in an industrial application. Their approximations have the huge advantage that Kalman filtering and smoothing methods can be used, allowing for fast real time forecasting. It is this approximate method we use in this paper.

Complete details of the Raftery et al (2007) algorithm are given in the Technical Appendix. Here we explain only the main ideas. In general, Bayesian estimation of state space models such as (5) involve Markov Chain Monte Carlo (MCMC) methods which take draws of the states conditional on the other model parameters (such as H_t and Q_t) and then draw the other parameters conditional on the states. With the large number of models we are working with, it is computationally impossible to use such MCMC methods. Accordingly, one aspect of the Raftery et al (2007) algorithm is to avoid MCMC. They do this by obtaining a plug-in estimate of H_t and assuming $Q_t = \left(1 - \lambda^{-1}\right) \Sigma_{t-1}$ where $0 < \lambda \leq 1$ and Σ_t is the covariance matrix of the estimation error in the Kalman filter (i.e. the estimation error is $\left(\theta_t - \widehat{\theta}_t\right)$ where

[3] In our recursive forecasting exercise, we extract factors recursively so that the factors at time τ are constructed using information through time τ.

[4] The alternative would be to treat $f_t^{(b)}$ as a latent variable and simulate it in the context of a Markov Chain Monte Carlo algorithm as in Ng, Moench and Potter (2008). This is theoretically straightforward, but computationally infeasible when doing DMA with a reasonable number of blocks.

[5] In related models, Chen and Liu (2000) show how computation time up to t typically involves mixing over K^t terms.

$\widehat{\theta}_t$ is the filtered estimate, see Technical Appendix). Such approximations have been used frequently in the state space literature. Raftery et al (2007) provide a detailed justification of this approximation and relate the resulting approach to statistical methods such as age-weighting and windowing. λ is known as a "forgetting factor", which is motivated by the fact that this specification implies that observations j periods in the past have weight λ^j. An alternative way of interpreting λ is to note that it implies an effective window size of $\frac{1}{1-\lambda}$. It is common to choose a value of λ near one, suggesting a gradual evolution of coefficients. $\lambda = 1$ implies θ_t is constant over time. As λ decreases, a greater and greater degree of coefficient change is allowed for. As $\lambda \rightarrow 0$ we end up in a case where only the most recent observation is of use for forecasting (or equivalently, large structural breaks are occurring in every time period).

The second aspect of the Raftery et al (2007) algorithm involves an approximation that allows for the fast and efficient calculation of posterior model probabilities at each point in time. As notation, let $\pi_{t|r,k} = \Pr\left(L_t = k|y^r\right)$ denote the probability that model k applies at time t using information through time r where this information is denoted by $y^r = (y_1, .., y_r)'$.When forecasting at time t, we can use $\pi_{t|t-1,k}$ to either do forecast averaging or forecasting using a single best model. That is, DMS can be done by simply selecting the model with highest $\pi_{t|t-1,k}$ to be the forecasting model at time t. Alternatively, DMA involves forecasting from all K models and using $\pi_{t|t-1,k}$ for $k = 1, .., K$ as weights when constructing an average forecast.

In general, we can obtain $\pi_{t|t-1,k}$ in an iterative manner. That is, we have the relationship

$$\pi_{t-1|t-1,k} = \frac{\pi_{t-1|t-2,k} p_k\left(y_{t-1}|y^{t-2}\right)}{\sum_{s=1}^{K} \pi_{t-1|t-2,l} p_s\left(y_{t-1}|y^{t-2}\right)}, \tag{7}$$

where $p_s\left(y_{t-1}|y^{t-2}\right)$ is the predictive density for model s (as shown in the Technical Appendix this is simply a Normal density) evaluated at y_{t-1}. If we were to use a conventional Markov switching process for the models (as described above) with transition probabilities such as P with elements p_{ks} then an iterative algorithm combining (7) and

$$\pi_{t|t-1,k} = \sum_{s=1}^{K} \pi_{t-1|t-1,l} p_{ks} \tag{8}$$

can be immediately seen. However, as discussed previously, such a strategy is impossible in our case since P is of too large a dimension. Raftery et al (2007) surmount this problem by replacing (8) by

$$\pi_{t|t-1,k} = \frac{\pi_{t-1|t-1,k}^{\alpha}}{\sum_{l=1}^{K} \pi_{t-1|t-1,l}^{\alpha}}, \tag{9}$$

where $0 < \alpha \leq 1$ is another forgetting factor similar to λ. Its interpretation is similar to λ, but in terms of the evolution of models (not the evolution of states). Raftery et al (2007) and Smith and Miller (1986) provide a detailed discussion for why this approximation is empirically sensible. Insight into the interpretation of α is obtained by noting that (9) can be written as:

$$\begin{aligned} \pi_{t|t-1,k} &\propto \left[\pi_{t-1|t-2,k} p_k\left(y_{t-1}|y^{t-2}\right)\right]^{\alpha} \\ &= \prod_{i=1}^{t-1} \left[p_k\left(y_{t-i}|y^{t-i-1}\right)\right]^{\alpha^i}. \end{aligned}$$

It can be seen that $\pi_{t|t-1,k}$ will be larger and, thus, DMS will be more likely to select model k at time t if it has forecast well in the recent past (where forecast performance is measured by the predictive density, $p_k\left(y_{t-i}|y^{t-i-1}\right)$). The interpretation of "recent past" is controlled by the forgetting factor, α and we have the same exponential decay at the rate α^i for observations i periods ago as we had associated with λ. Thus, if $\alpha = 0.99$ (our benchmark value and also the value used by Raftery et al, 2007) and we are using monthly data, forecast performance five years ago receives about 50% as much weight as forecast performance last period. Forecast performance one year ago receives about 90% as much weight as last month's performance. If $\alpha = 0.95$, then forecast performance five years ago receives only about 5% as much weight as last period's performance. These considerations suggest that we focus on the interval

$\alpha \in (0.95, 0.99)$ with our benchmark choice being $\alpha = 0.99$. Note also that $\alpha = 1$ implies that the same model is used in every time period. For similar reasons, we will also focus on $\lambda \in (0.95, 0.99)$ with our benchmark choice being $\lambda = 0.99$.

Note that one could choose values for λ and α based on forecast performance in some way, but this is would bias our results in favour of DMA and is not a valid procedure for out-of-sample forecasting. Alternatively when forecasting at time τ we could consider a grid of values for λ and α and select the value which yielded the highest value for an information criterion or the marginal likelihood. In essence, this amounts to treating α and λ as unknown parameters. However, this would greatly add to the computational burden, perhaps so much as to make it impossible to do forecasting in real time. Hence, we follow Raftery et al (2007) and simply select values for the forgetting factors, but our Empirical Appendix carries out a sensitivity analysis.

Complete details on the Raftery et al (2007) approach to DMA are provided in the Technical Appendix. The purpose of this section was to explain the basic ideas of DMA and DMS. In particular, we have shown why it is necessary to use such approximations; have stressed the fact that with these approximations no MCMC algorithm is required (only Kalman filtering and smoothing or similar iterative updating algorithms); and provided explanation of the forgetting factors λ and α which are important in DMA and DMS.

Finally, our treatment of H_t is similar to that used by Raftery et al (2007) who use a plug-in method where we simply replace H_t by an estimate $\widehat{H}_t$. Details are provided in the Technical Appendix. We use a rolling version of their estimator to allow for changes in volatility. The reader interested in further discussion of DMA, including its relationship with conventional Bayesian model averaging (BMA), is referred to Raftery et al (2007).

3 Empirical Work

3.1 Data and Modelling Issues

The Bank of England maintains a data set of many variables that it uses in its suite of UK forecasting models (see Kapetanios, Labbard and Price, 2009). From this we have taken the 139 monthly variables for which complete data is available from 1990M1 through 2008M11. The variables fall naturally into eight categories that we use as our blocks from which we extract factors. Thus, we have an international block which contains various US and European output, unemployment and price variables. The output block contains various indices of production as well as surveys of manufacturers on their present output plans. The price block contains various measures of inflation and inflation expectations as well as variables relating to wages and the price of oil. The demand block uses various measures of consumer confidence and sales. The financial block has stock prices, dividend yields, the exchange rates for several important currencies relative to the £ and various interest rates and spreads. The housing block contains variables reflecting house prices and sales. The money block contains various measures of the money stock and bank deposits and the labour block contains various measures of employment and unemployment. When doing factor analysis it is common to transform all variables all variables to stationarity. We have done this, making the same choices as the Bank of England for the transformations. The complete list of variables and transformations used in given in the Data Appendix.

Remember that our models are all based on (3) and involve lags of the dependent variable, factors from each block as well as lags of these factors. We extract the first factor for each block and include it and one lag of it as potential predictors as well as two lags of the dependent variable (a choice which is adequate to clean up any autocorrelation in the errors) and an intercept. Note that this strategy leads to 18 potential predictors. All of our models in (6) will involve subsets of these predictors. If (as done in many implementations of DMA and BMA) we define our models according to whether each individual predictor is included or excluded from the model we would end up with 2^{18} different models. Even though DMA leads to great computational simplifications, it still can be computationally demanding and handling this many models is infeasible. For this reason we limit our set of models by assuming: i) all models contain the intercept and lags of the dependent variable and ii) models are defined by whether each factor and its lag are jointly included or excluded. So, for example, we have models where the price factor and its lag are included and models where neither the price factor nor its lag are included. But

we rule out models which contain only the price factor (but not its lag) or only the lagged price factor. With these assumptions we have 2^8 models.

We use principal components to extract factors using all the variables in each block. We forecast inflation (the annual percentage change in the all-items CPI) and output growth (the percentage change in the index of production for all production industries). When forecasting inflation, the all-items CPI variable is deleted from the price block. When forecasting output growth, the index of production for all production industries is deleted from the output block.

Our main results are for $\alpha = 0.99$ and $\lambda = 0.99$, but we do discuss sensitivity to these choices in an Empirical Appendix. The Kalman filter and iterative algorithm for drawing model probabilities are both initialized diffusely as described in the Technical Appendix.

3.2 Forecasting Performance

We compare our forecasts using DMA and DMS to several alternative forecasting methods. These can be interpreted as special cases of DMA or DMS, but either with particular choices for single models or particular choices for α and λ. The first two of our alternative methods allow for time variation of parameters, but no time variation in the models. These are an AR(2) model with time varying parameters (labelled TVP-AR(2)) and the TVP block factor model in (3) including all the predictors (labelled TVP-Factor). The TVP part of the model is specified using $\lambda = 0.99$. Our third alternative uses DMA, but does not allow the coefficients to vary over time in each model (i.e. this is a special case of DMA where $\lambda = 1$, but $\alpha = 0.99$). Finally, we present results using Bayesian model averaging (BMA) which is a special case of DMA where $\lambda = \alpha = 1$. We evaluate forecast performance over the period 1992M3 through 2008M11.

The standard metric of Bayesian forecast comparison is the sum of log predictive likelihoods (see, e.g., Geweke and Amisano, 2007). This has the advantage that it involves entire predictive distribution and not simply point forecasts. The predictive likelihood is the predictive density for y_t given data through time $t-1$ evaluated at the actual outcome (i.e. in model k the predictive density is $p_k\left(y_t|y^{t-1}\right)$). The formula for the predictive density is given in the Technical Appendix. In addition to the sum of log predictive likelihoods, we also present results for two standard measures of the performance of the point forecasts, mean squared forecast error and mean absolute forecast error defined as:

$$MSFE = \frac{\sum_{\tau=\tau_0}^{T}\left[y_\tau - E\left(y_\tau|Data_{\tau-h}\right)\right]^2}{T-\tau_0+1}$$

and

$$MAFE = \frac{\sum_{\tau=\tau_0+1}^{T}\left|y_\tau - E\left(y_\tau|Data_{\tau-h}\right)\right|}{T-\tau_0+1}.$$

where $Data_{\tau-h}$ denotes the information available through period $\tau-h$ where h is the forecast horizon and $E\left(y_\tau|Data_{\tau-h}\right)$ is the point forecast of y_τ.

Tables 1 and 2 present these measures of forecast comparison for three forecast horizons, $h = 1, 6$ and 12, for output growth and inflation, respectively. The main story coming out of these tables is a strong one: DMA and DMS almost always forecast better than the other approaches and never forecast much worse than the best alternative. Particularly for output growth, the forecasting gains of DMA and DMS relative to the alternative approaches are quite substantial. This story comes through strongest in the log predictive likelihoods where there is only one case where DMA and DMS are not the best two forecasting methods. The exception is for inflation forecasting for $h = 1$. In this case, the parsimonious TVP-AR(2) has (by a small amount) the highest sum of log predictive likelihoods. However, this result does not carry over to MAFE and MSFE where DMA and DMS show a better forecast performance.

A second story is that time-varying parameter models which simply include all potential predictors consistently perform extremely poorly. This clearly shows the benefits of DMA. A naive researcher might think that simply working with a single flexible TVP model such as (3) and (4) would be adequate since it might be able to approximate what is done in DMA. That is, TVP models allow for the marginal effects of predictors to evolve which, in theory, could allow for predictors to (approximately) drop in or

out of the model over time (i.e. their marginal effects could be near zero in some time periods but evolve to having more substantive effects in others). This is clearly not happening in this data set. DMA is adding something of great benefit for forecasting. The particularly poor performance of the TVP-Factor model suggests one reason why this is so: the TVP-Factor model is making some big forecast errors at some points in time. This is probably due to a changes in forecasting model and/or structural breaks which the TVP-Factor model cannot adjust to quickly enough. This is a point we will return to later in our discussion of how the forecasting model changes over time.

A third story relates to forecast shrinkage. The fact that the parsimonious TVP-AR(2) often forecasts well, typically much better than the TVP-Factor approach which includes all the predictors, indicates the benefit of shrinkage. The gains in forecast performance one would expect the extra information in the factors to provide is clearly outweighed by the fact that models such as (3) and (4), or even (2), introduce a large number of new parameters. It is only through use of DMA or DMS that we can realize the benefits of this extra information, since most of these new parameters are shrunk to zero. And it is worth emphasizing that the way this shrinkage occurs is changing over time. It is worth stressing, however, the extremely poor forecast performance of the TVP-AR(2) model for inflation when $h = 12$. In this case, it is clear that the predictors do contain important information that the TVP-AR(2) is missing. This suggests that, while DMA and DMS are safe forecasting options (they typically forecast best, but even when not the best, they never forecast poorly relative to alternative approaches), extremely parsimonious models like the TVP-AR(2) are riskier. Even though they often forecast well, sometimes their forecasts are way off.

DMA and DMS allow for variation in models and variation in parameters. Our fourth story relates to the relative roles of these two aspects in improving forecast performance. Tables 1 and 2 (with some exceptions) indicate that DMA with constant parameters performs fairly well, usually beating BMA (which has no variation in either parameters or models) by a substantial amount.[6] This suggests that variation in models is more important than variation in parameters in the good forecast performance of DMA and DMS.

Finally, what do the tables say about the relative forecast performance of DMA and DMS? In this regard, there is no clear pattern. Sometimes DMA forecasts better and sometimes DMS forecasts better (and sometimes the story of log-predictive likelihoods is different than the story told by MAFEs and MFSEs).

[6] We should stress that all models are estimated recursively, so statements like "no variation in either parameters or models" apply for a given forecast time, τ.

Table 1: Measures of Forecast Performance (Output growth)			
Forecast Method	Sum of log pred. like.	MSFE	MAFE
	$h = 1$		
DMA ($\alpha = \lambda = 0.99$)	-200.720	97.501	98.550
DMS ($\alpha = \lambda = 0.99$)	-200.390	99.106	99.346
TVP-Block	-228.514	119.151	112.204
TVP-AR(2)	-207.214	104.559	102.439
DMA ($\lambda = 1, \alpha = 0.99$)	-201.970	98.427	98.795
BMA (DMA with $\alpha = \lambda = 1$)	-206.330	105.020	101.240
	$h = 6$		
DMA ($\alpha = \lambda = 0.99$)	-217.430	115.700	110.910
DMS ($\alpha = \lambda = 0.99$)	-217.480	116.020	110.910
TVP-Block	-258.235	182.993	138.614
TVP-AR(2)	-221.453	116.018	110.909
DMA ($\lambda = 1, \alpha = 0.99$)	-217.640	115.660	110.950
BMA (DMA with $\alpha = \lambda = 1$)	-217.510	115.790	111.040
	$h = 12$		
DMA ($\alpha = \lambda = 0.99$)	-226.920	117.540	110.250
DMS ($\alpha = \lambda = 0.99$)	-227.520	119.560	112.160
TVP-Block	-265.8394	189.564	138.4474
TVP-AR(2)	-232.1462	120.54	112.3358
DMA ($\lambda = 1, \alpha = 0.99$)	-227.370	118.290	110.900
BMA (DMA with $\alpha = \lambda = 1$)	-228.550	120.480	111.940

Table 2: Measures of Forecast Performance (Inflation)			
Forecast Method	Sum of log pred. like.	MSFE	MAFE
	$h = 1$		
DMA ($\alpha = \lambda = 0.99$)	-13.118	9.637	32.722
DMS ($\alpha = \lambda = 0.99$)	-12.615	9.445	32.408
TVP-Block	-22.716	13.396	38.727
TVP-AR(2)	-10.537	9.838	32.785
DMA ($\lambda = 1, \alpha = 0.99$)	-14.631	9.639	32.628
BMA (DMA with $\alpha = \lambda = 1$)	-21.437	9.973	33.065
	$h = 6$		
DMA ($\alpha = \lambda = 0.99$)	-196.780	75.870	89.099
DMS ($\alpha = \lambda = 0.99$)	-196.380	81.795	95.183
TVP-Block	-216.125	106.112	103.547
TVP-AR(2)	-202.954	83.351	97.057
DMA ($\lambda = 1, \alpha = 0.99$)	-207.520	88.206	97.246
BMA (DMA with $\alpha = \lambda = 1$)	-215.750	103.150	105.200
	$h = 12$		
DMA ($\alpha = \lambda = 0.99$)	-270.870	190.020	134.390
DMS ($\alpha = \lambda = 0.99$)	-271.840	194.780	135.550
TVP-Block	-280.243	202.460	147.685
TVP-AR(2)	-567.132	961.581	272.127
DMA ($\lambda = 1, \alpha = 0.99$)	-287.460	206.980	139.180
BMA (DMA with $\alpha = \lambda = 1$)	-301.540	234.100	148.760

3.3 Which are the Most Important Predictors and Does this Change over Time?

Of the different forecasting approaches given in the preceding section, only DMA and DMS allow for different forecasting models at different times. Accordingly, in this section we present results only for these two approaches. Given the huge number of models we cannot possibly present empirical results for every model. Instead we summarize results in two different ways. We begin with figures which illustrate that, although we have 8 factor blocks which could potentially be selected, most of the time the best model used by DMS is much more parsimonious with only a few of these blocks. Formally, if we let $Size_k$ be the number of blocks of factors in model k, then

$$E\left(Size_t\right) = \sum_{k=1}^{K} \pi_{t|t-1,k} Size_k$$

can be interpreted as the expected or average number of blocks of factors used by DMA at time t. Figures 1 and 2 plots $E\left(Size_t\right)$ for our six empirical exercises (i.e. two forecast variables and three forecast horizons).

The patterns in Figures 1 and 2 both tend to indicate that, as time goes by and more data is available for estimation, more factors are chosen (although, for inflation, this tendency stops after about 2000). This is as expected. Both Figures 1 and 2 are indicating a high degree of parsimony (with the exception of inflation forecasting at long horizons, we always have $E\left(Size_t\right) < 3.5$), but are somewhat different from one another. When forecasting output growth, DMA is placing most weight on forecasting models with only 1 or 2 factor blocks (and, at the beginning of the sample, the TVP-AR(2) with no factor blocks at all is receiving most of the posterior weight). When inflation is the dependent variable, DMA is choosing somewhat less parsimonious models, but for $h = 1$ and $h = 6$ DMA is placing most weight on models with two or three factors. The interesting exception is inflation forecasting for $h = 12$ where DMA is choosing less parsimonious models with up to five factor blocks. This explains the extremely poor forecast performance of the TVP-AR(2) for this case noted above (see Table 2).

Figures 1 and 2 present evidence that the best forecasting model is changing over time and that DMA and DMS tend to be quite parsimonious. However, they shed little light on which factor blocks are the most important at various points in time. To shed light on this issue, remember that the Raftery et al (2007) DMA algorithm provides us with time-varying probabilities associated with every model (i.e. $\pi_{t|t-1,k}$ for $k = 1, .., K$). For any factor block at any point in time, we can use these to calculate the total probability associated with models containing a particular block factor (and its lag). That is, for $b = 1, .., B$, we can calculate $\sum_{k \subset b} \pi_{t|t-1,k}$ where the notation implies the summation is taken over models which include the b^{th} factor block. Figures 3 through 8 plot these probabilities for our 8 factor blocks for output growth and inflation, respectively.

The patterns for output growth and inflation are quite different, reflecting the higher degree of parsimony which DMA finds when forecasting the former. For output growth (see Figures 3 through 5), we are only occasionally finding lines which go near one. But it is also the case that it is rare for the lines in Figures 3 through 5 to be horizontal lines at zero. This indicates that there is no factor block which is always or never an important predictor. For instance, for $h = 1$, after 2000, Figure 3 indicates that DMA is averaging over many forecasting models with appreciable weight (e.g. more than 10%) attached to models containing each of the 8 factor blocks. However, there is no one factor block that is always an important predictor. That is, even the highest line (corresponding to the output block) in Figure 3, never goes much above 0.50. Thus, the other 50% of the weight used by DMA when averaging forecasts is coming from models which do not contain this most important of predictors.

For inflation, the patterns are more complicated. In contrast to output growth, for much of our sample period, there are variables which are clearly important predictors in the sense that some lines in Figure 6 through 8 are near one. This is particularly true for the price and money blocks. Remember that the price block includes lagged information on wages, oil prices and the various disaggregated components of the CPI. DMA is finding information in this block of variables that is often helpful for predicting current inflation. However, the patterns in Figures 6 through 8 vary over time and across forecast horizons. For instance, for $h = 6$ the price block is most important near the beginning of the sample, but for $h = 1$

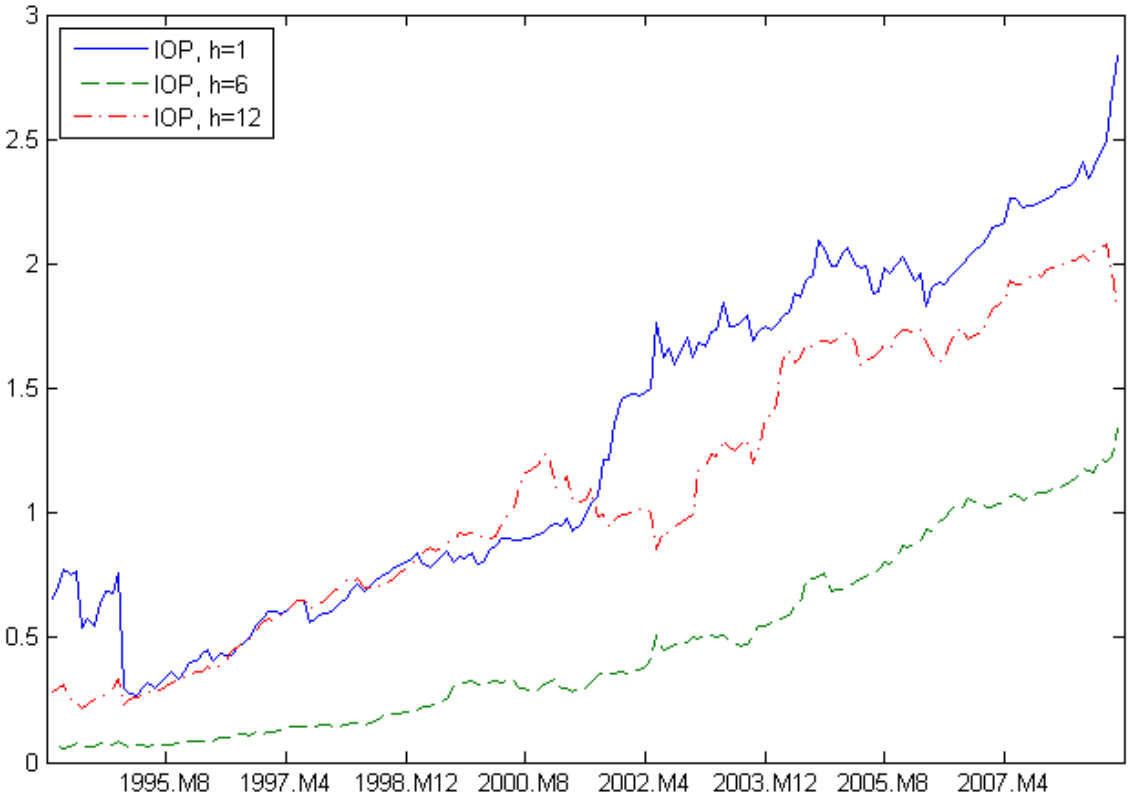

Figure 1: Figure 1: Expected Number of Predictors in Each Forecasting Exercise (Output growth)

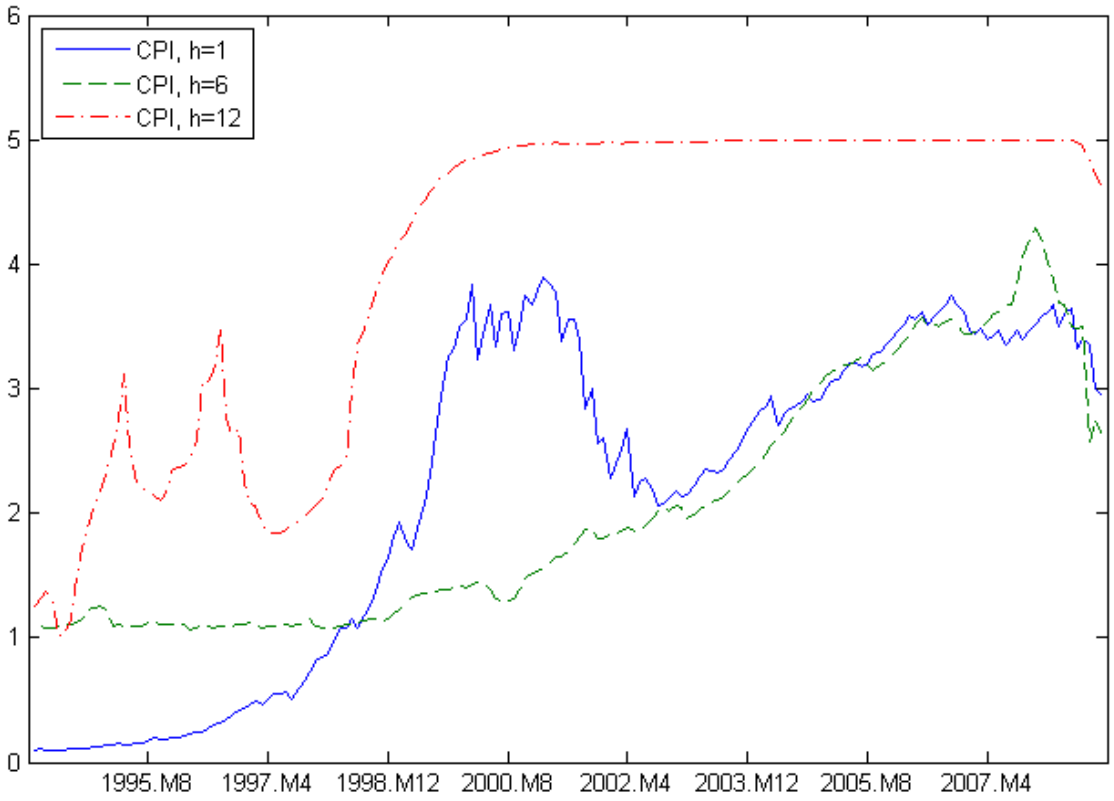

Figure 2: Figure 2: Expected Number of Predictors in Each Forecasting Exercise (Inflation)

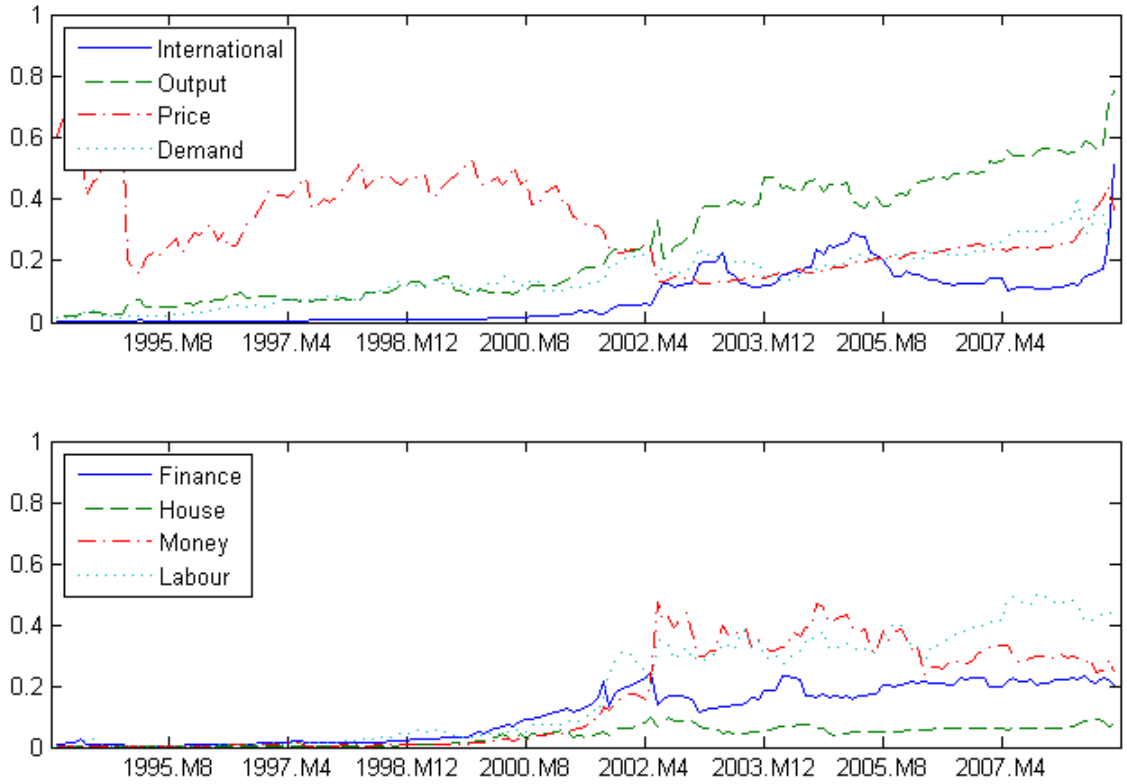

Figure 3: Figure 3: Total Probability of Models which Contain Each Factor Block (Output growth, $h = 1$)

and $h = 12$ it is most important in the latter half of the sample. For $h = 12$ we find that the price block is almost always a very important predictor, except for a brief spell around 1998 where is becomes completely unimportant. The money block is an important predictor at all forecast horizons after the year 2000, but patterns before 2000 are more variable.

Clearly we are finding a great deal of variation over time in terms of what the best predictors. Such variation could not be modelled using conventional forecasting approaches or conventional theoretical models. For instance, standard Phillips' curve arguments would imply that the unemployment rate is always a good predictor for inflation. If this were true, we would find the labour factor (which includes various measures of the unemployment rate) to always be a good predictor for inflation and the total probability of models which contain the labour factor to be near one in Figures 6 through 8. We are not finding patterns such as this for the labour block (or for any other factor block) . Sometimes (i.e. after 1998 for $h = 12$) we are finding the labour factor to be important in forecasting inflation, but not at other times. But at most times, models which contain the labour factor will be allocated appreciable (e.g. more than 10%) of the weight when doing DMA.

When forecasting inflation, almost all of the factor blocks play an appreciable role at some time or forecast horizon or another. Somewhat suprisingly, it is only the housing block which never receives appreciable weight. The demand and financial blocks often come through as being important predictors (but often are unimportant). The picture we are finding is one where DMA is averaging over many parsimonious models (as opposed to selecting just a few parsimonious models) and this set of models is changing substantially over time. These characteristics would be hard to mimic in conventional approaches.

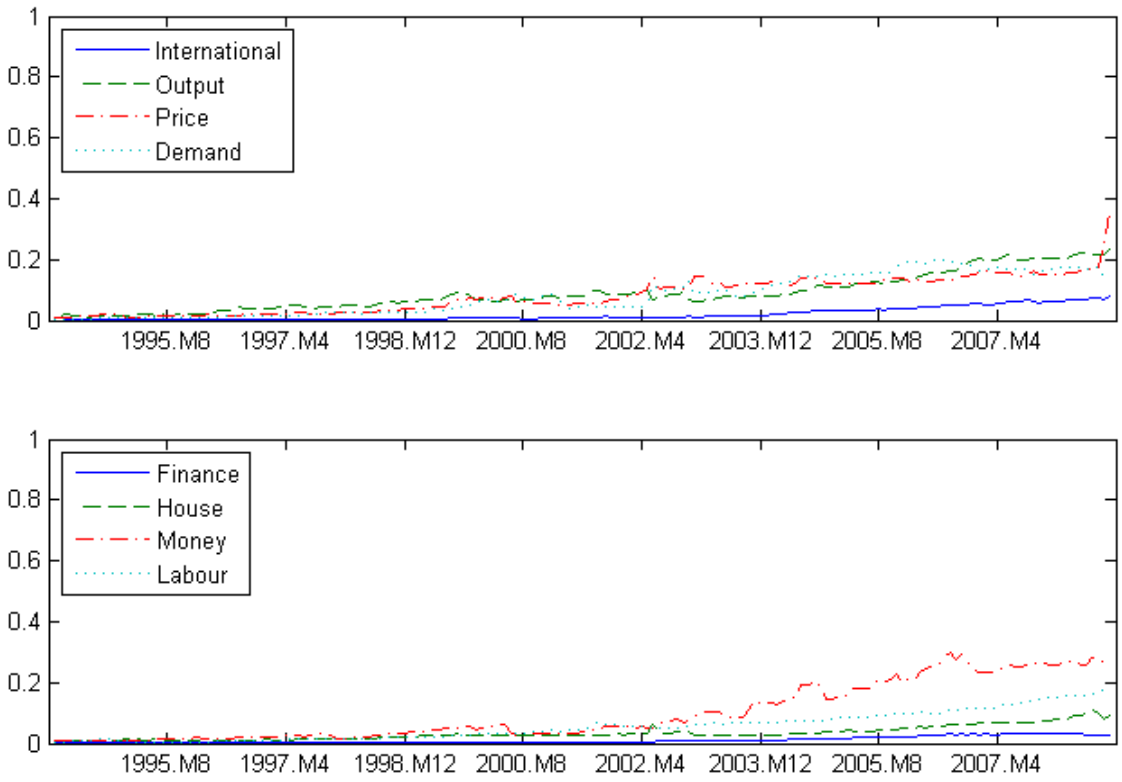

Figure 4: Figure 4: Total Probability of Models which Contain Each Factor Block (Output growth, $h = 6$)

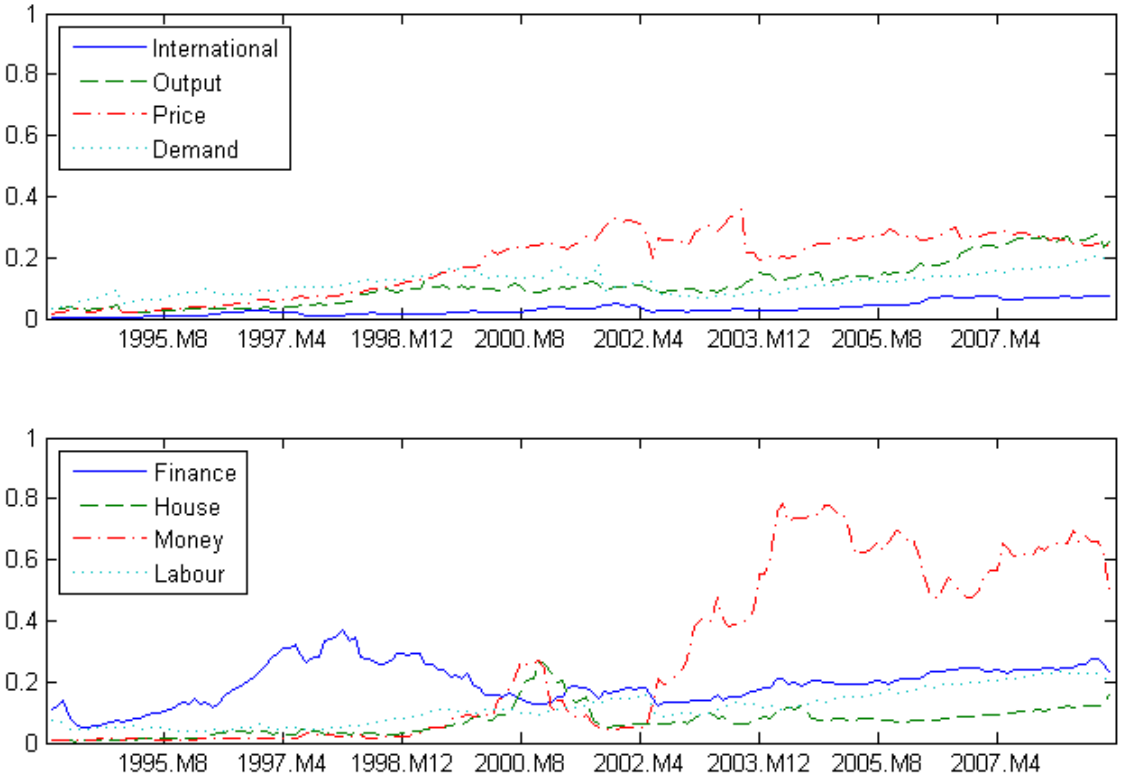

Figure 5: Figure 5: Total Probability of Models which Contain Each Factor Block (Output growth, $h = 12$)

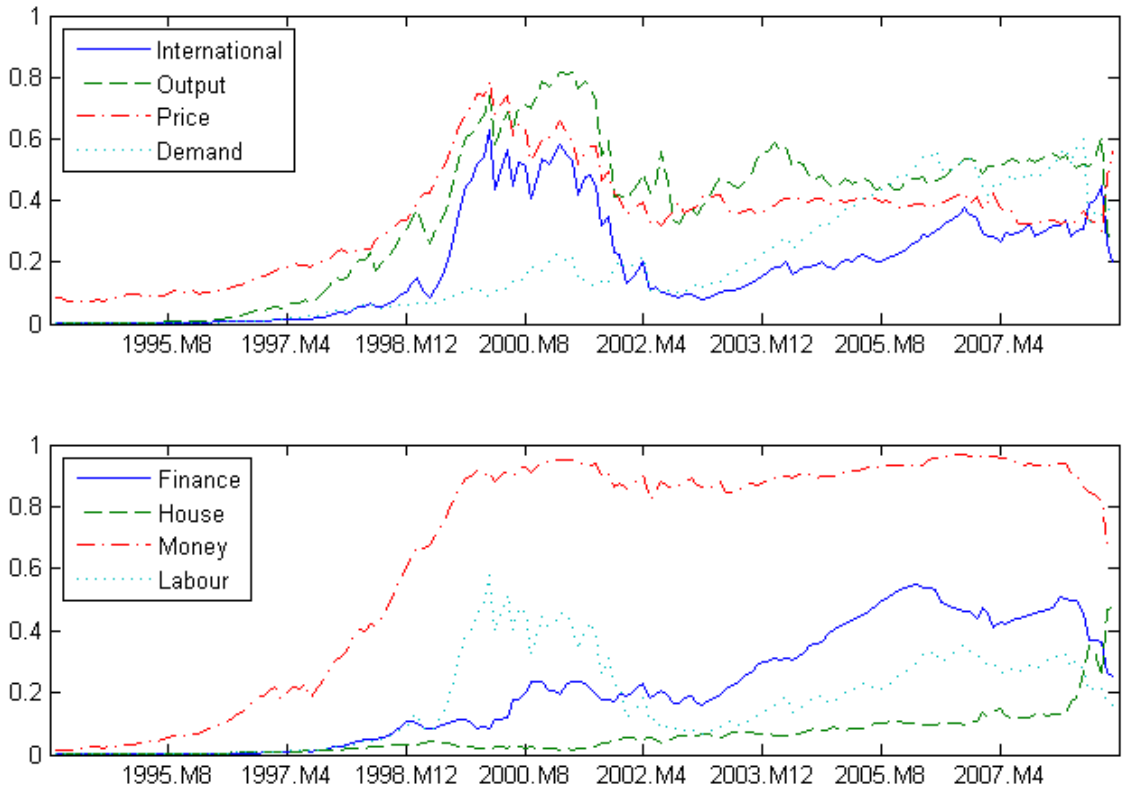

Figure 6: Figure 6: Total Probability of Models which Contain Each Factor Block (Inflation, $h = 1$)

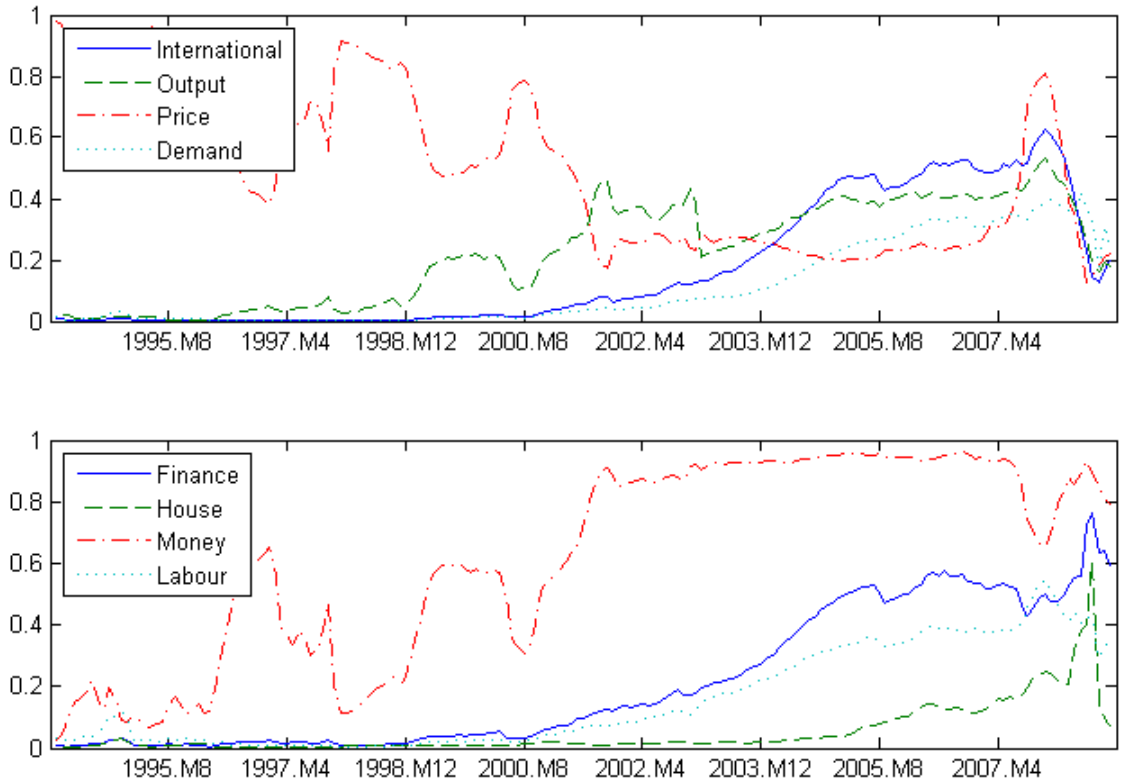

Figure 7: Figure 7: Total Probability of Models which Contain Each Factor Block (Inflation, $h = 6$)

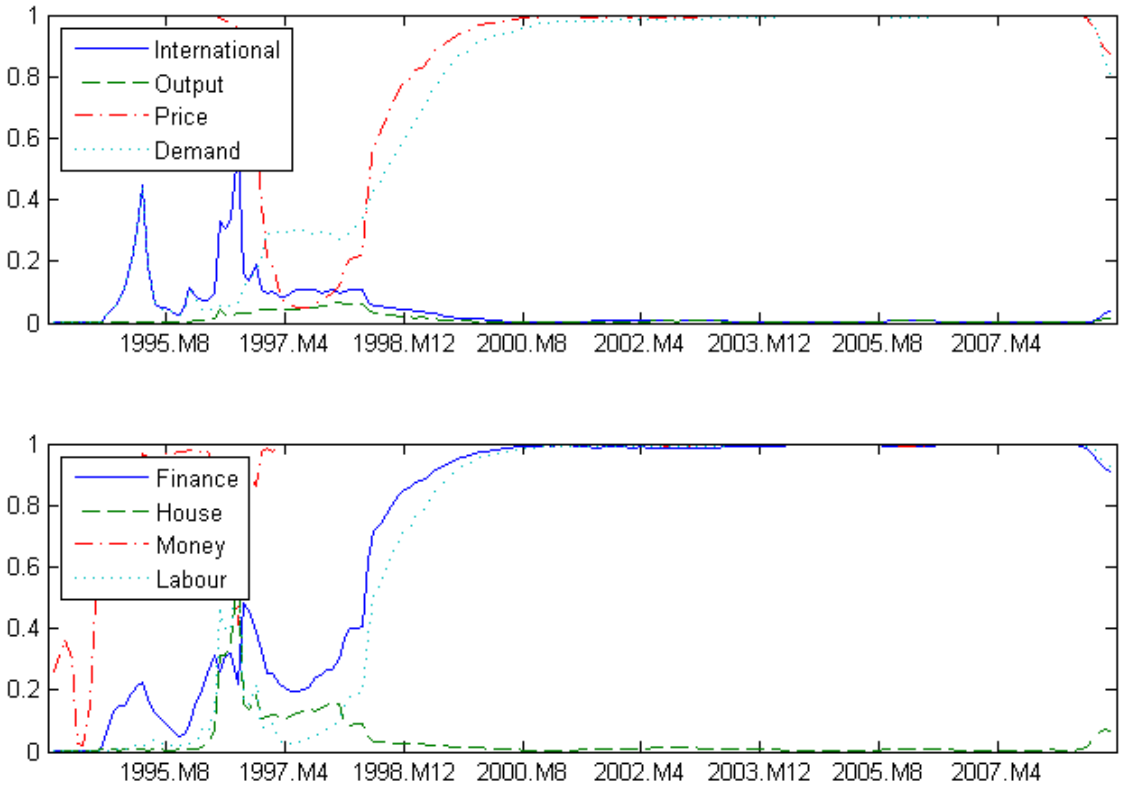

Figure 8: Figure 8: Total Probability of Models which Contain Each Factor Block (Inflation, $h = 12$)

4 Conclusions

In this paper, we have argued that DMA and DMS hold many attractions for the macroeconomic forecaster. They allow for the coefficients in a model to evolve over time, but also allow for model used for forecasting to change over time. In practice, this means that DMS switches between various parsimonious models over time and DMA tends to place a great deal of weight on such parsimonious models. The alternative of working with one general model, including all potential predictors, is unattractive due to over-parameterization worries. The alternative of choosing one single parsimonious model is also unattractive since a good parsimonious forecasting model at some times could be a bad model at other times. But, to our knowledge, other than the regression-based US application of Koop and Korobilis (2009), DMA and DMS have not been used by macroeconomic forecasters. Relative to our previous work, the present paper extends the use of DMA and DMS to block factor models with the monthly data used in the Bank of England's forecasting suite. We find improvements in forecast performance relative to several popular alternatives. Furthermore, the use of DMA and DMS provides insight on which factors predict inflation and output growth and whether they change over time. We find that DMA is averaging over many different parsimonious models and the set of parsimonious models is changing substantially over time. For output growth, there are no factor block which stand out as being a consistently important predictor. But every factor block is playing a role in the DMA average forecast at some time or other. For inflation, the money and price blocks stand out as being fairly consistently important predictors (although there are some times when they are not). But all the other factor blocks are important at some times or for some forecast horizons. The general pattern, though, is one where the best forecasting model is changing over time. This feature is automatically picked up by DMA or DMS, but not with conventional forecast procedures.

References

Artis M, Bannerjee A, Macellino, M. Factor forecasts for the UK. Journal of Forecasting 2005; 24; 279-298.

Chen R, Liu J. Mixture Kalman filters. Journal of the Royal Statistical Society, Series B 2000; 62; 493-508.

Cogley T, Sargent T. Drifts and volatilities: Monetary policies and outcomes in the post WWII U.S. Review of Economic Dynamics 2005; 8; 262-302.

Cogley T, Morozov S, Sargent T. Bayesian fan charts for U.K. inflation: Forecasting and sources of uncertainty in an evolving monetary system. Journal of Economic Dynamics and Control 2005; 29; 1893-1925.

Geweke J 1977. The Dynamic Factor Analysis of Economic Time Series. In: Aigner D, Goldberger A (Eds), Latent variables in socio economic models. North Holland: Amsterdam, 1977.

Geweke J, Amisano J. Hierarchical Markov normal mixture models with applications to financial asset returns, 2007 manuscript available at http://www.biz.uiowa.edu/faculty/jgeweke/papers/paperA/paper.pdf.

Groen J, Paap R, Ravazzolo F. Real-time inflation forecasting in a changing world, 2008 manuscript.

Kapetanios G, Labbard V., Price S. Forecast combination and the Bank of England's suite of statistical forecasting models. Economic Modelling 2009; forthcoming.

Koop G, Korobilis, D. Forecasting inflation using dynamic model averaging, 2009 manuscript available at

http://personal.strath.ac.uk/gary.koop/koop_korobilis_forecasting_inflation_using_DMA.pdf.

Koop G, Potter S. Forecasting in dynamic factor models using Bayesian model averaging. The Econometrics Journal 2004; 7; 550-565.

Marcellino M, Stock J, Watson M. Macroeconomics forecasting in the Euro area: Country specific versus area wide information. European Economic Review 2003; 47; 1-18.

Ng S, Moench E, Potter S. (2008). Dynamic hierarchical factor models, manuscript available at http://www.columbia.edu/~sn2294/papers/dhfm-short.pdf.

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

Primiceri G. Time varying structural vector autoregressions and monetary policy. Review of Economic Studies 2005; 72; 821-852.

Raftery A, Karny M, Andrysek J, Ettler P. (2007). Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill, Technical report 525, Department of Statistics, University of Washington.

Smith J, and Miller J. A non-Gaussian state space model and application to prediction records. Journal of the Royal Statistical Society, Series B 1986; 48; 79-88.

Stock J, Watson M 2006. Macroeconomic Forecasting Using Many Predictors. In: Elliott G, Granger C, Timmerman A (Eds), Handbook of economic forecasting. North Holland: Amsterdam, 2006.

Stock J, Watson M. Macroeconomic forecasting using diffusion indexes. Journal of Business and Economic Statistics 2002; 20; 147-162.

Stock J, Watson M. Forecasting inflation. Journal of Monetary Economics 1999; 44; 293-335.

Stock J, Watson M. Evidence on structural instability in macroeconomic time series relations. Journal of Business and Economic Statistics 1996; 14; 11-30.

Technical Appendix

In this appendix, we describe the DMA algorithm of Raftery et al (2007). With one minor exception (the treatment of H_t), our methods are identical to theirs and the reader is referred to their paper for further details, explanation and motivation for their approach.

We begin with the standard state space model in (5). For given values of H_t and Q_t, the Kalman filter and smoother can be used to carry out recursive estimation or forecasting. That is, if we let $y^t = (y_1, .., y_t)'$ then Kalman filtering begins with the result that

$$\theta_{t-1}|y^{t-1} \sim N\left(\widehat{\theta}_{t-1}, \Sigma_{t-1}\right) \tag{10}$$

where formulae for $\widehat{\theta}_{t-1}$ and Σ_{t-1} are provided below and proceeds using:

$$\theta_t|y^{t-1} \sim N\left(\widehat{\theta}_{t-1}, R_t\right), \tag{11}$$

where

$$R_t = \Sigma_{t-1} + Q_t.$$

Raftery et al (2007) replaces this latter equation by:

$$R_t = \frac{1}{\lambda}\Sigma_{t-1} \tag{12}$$

or, equivalently, $Q_t = \left(1 - \lambda^{-1}\right)\Sigma_{t-1}$ where $0 < \lambda \leq 1$.

The next step towards forecasting in this one model case involves:

$$\theta_t|y^t \sim N\left(\widehat{\theta}_t, \Sigma_t\right), \tag{13}$$

where

$$\widehat{\theta}_t = \widehat{\theta}_{t-1} + R_t Z_t \left(H_t + Z_t R_t Z_t'\right)^{-1}\left(y_t - Z_t\widehat{\theta}_{t-1}\right) \tag{14}$$

and

$$\Sigma_t = R_t - R_t Z_t \left(H_t + Z_t R_t Z_t'\right)^{-1} Z_t R_t. \tag{15}$$

Recursive forecasting is done using the predictive distribution

$$y_t|y^{t-1} \sim N\left(Z_t\widehat{\theta}_{t-1}, H_t + Z_t R_t Z_t'\right). \tag{16}$$

We now switch to the notation for the multi-model case in (6) and let Θ_t denote the vector of all the coefficients in all the models. In the standard single model case, Kalman filtering is based on (10), (11) and (13). In the multi-model case, for model k, these three equations become:

$$\Theta_{t-1}|L_{t-1} = k, y^{t-1} \sim N\left(\widehat{\theta}_{t-1}^{(k)}, \Sigma_{t-1}^{(k)}\right), \tag{17}$$

$$\Theta_t|L_t = k, y^{t-1} \sim N\left(\widehat{\theta}_{t-1}^{(k)}, R_t^{(k)}\right) \tag{18}$$

and

$$\Theta_t|L_t = k, y^t \sim N\left(\widehat{\theta}_t^{(k)}, \Sigma_t^{(k)}\right), \tag{19}$$

where $\widehat{\theta}_t^{(k)}, \Sigma_t^{(k)}$ and $R_t^{(k)}$ are obtained via Kalman filtering in the usual way using (**??**), (14) and (15) except with (k) superscripts added to denote model k. To make clear the notation in these equations, note that, conditional on $L_t = k$, the prediction and updating equations will only provide information

on $\theta_t^{(k)}$ and not the full vector Θ_t. Hence, we have only written (17), (18) and (19) in terms of the distributions which hold for $\theta_t^{(k)}$.

The previous results were all conditional on $L_t = k$, and we need a method for unconditional prediction (i.e. not conditional on a particular model). This is done as described in the body of the text. That is, $\pi_{t|t-1,k}$ for $k = 1.,,.K$ can be used for averaging across models when forecasting at time t. It can be calculated iteratively using (9) and (7). Recursive forecasting can be done by averaging over predictive results for every model (obtained using (16) for each model) using $\pi_{t|t-1,k}$. So, for instance, DMA point predictions are given by:

$$E\left(y_t|Data_{t-1}\right) = \sum_{k=1}^{K} \pi_{t|t-1,k} Z_t^{(k)} \widehat{\theta}_{t-1}^{(k)}$$

where $Data_{t-1}$ denotes all data information available at time $t-1$. Predictive standard deviations can be calculated using this and predictive second moments:

$$E\left(y_t^2|Data_{t-1}\right) = \sum_{k=1}^{K} \pi_{t|t-1,k} \left(Z_t^{(k)} \widehat{\theta}_{t-1}^{(k)}\right)^2.$$

DMS proceeds by selecting the model with the highest value for $\pi_{t|t-1,k}$ at each point in time and using it for forecasting.

All the recursions above are started by choosing initial values for $\pi_{0|0,l}$, $\theta_0^{(s)}$ and $\Sigma_0^{(s)}$ for $s = 1, .., K$. In our empirical work we draw these from relatively diffuse priors for the initial conditions. In particular, we set $\pi_{0|0,l} = \frac{1}{K}$ (so that, initially, all models are equally likely), $\theta_0^{(l)} = 0$ and $\Sigma_0^{(l)} = 10I$.

The preceding discussion is all conditional on H_t. Raftery et al (2007) recommend a simple plug in method where $H_t^{(k)} = H^{(k)}$ and is replaced with a consistent estimate. When forecasting macroeconomic variables, however, it is likely that the error variance is changing over time. In theory, we could use a stochastic volatility or ARCH specification for $H_t^{(k)}$. However, to do this would greatly add to the computational burden. Thus, we prefer a simple plug-in approach which is a rolling version of the recursive method of Raftery et al (2007). To be precise, let

$$\widetilde{H}_t^{(k)} = \frac{1}{t-t^*} \sum_{j=t^*+1}^{t} \left[\left(y_t - Z_t^{(k)} \widehat{\theta}_{t-1}^{(k)}\right)^2 - Z_t^{(k)} R_t^{(k)} Z_t^{(k)\prime}\right].$$

Raftery el al (2007) uses this formula with $t^* = 0$, but we set $t^* = t - 20$. To avoid the rare possibility that $\widetilde{H}_t^{(k)} < 0$, Raftery et al (2007) use $\widehat{H}_t^{(k)}$ as an estimate of $H_t^{(k)}$ where:

$$\widehat{H}_t^{(k)} = \begin{cases} \widetilde{H}_t^{(k)} \text{ if } \widetilde{H}_t^{(k)} > 0 \\ \widehat{H}_{t-1}^{(k)} \text{ otherwise} \end{cases}$$

and we follow this practice.

Empirical Appendix

The results in the body of the paper have $\lambda = \alpha = 0.99$. We also present special cases of DMA where there is no time variation in parameters ($\lambda = 1, \alpha = 0.99$) and no time variation in either parameters or models ($\lambda = \alpha = 1$). In this appendix, we present a wider range of results for both forgetting factors. As discussed in the text, the interval $\lambda, \alpha \in [0.95, 1]$ is of most empirical interest.

****rewrite below

With regards to output growth, results are qualitatively the same as in Tables 1 and 2. If anything, setting $\lambda = \alpha = 0.90$ slightly improves forecasting performance of DMA and DMS (although this choice leads to much worse forecast performance for TVP-Factor). With regards to inflation, results are quite similar, although there is one exception. If we set the forgetting factors very high (0.99), then for $h = 1$, we are finding some evidence that our versions of AR(2) models forecast roughly as well as DMA or DMS. But this does not occur for other forecast horizons or even slightly lower values for the forgetting factors.

Table A1: Sensitivity Analysis: Output Growth

Forecast Method	Sum of log pred. like.	MSFE	MAFE
	$h = 1$		
DMA, $\alpha = \lambda = 0.95$	-197.200	87.155	95.129
DMS, $\alpha = \lambda = 0.95$	-176.450	81.181	89.393
DMA, $\alpha = 0.99, \lambda = 0.95$	-198.500	93.621	98.424
DMS, $\alpha = 0.99, \lambda = 0.95$	-194.400	92.568	99.535
DMA, $\alpha = 0.95, \lambda = 0.99$	-197.980	92.529	95.545
DMS, $\alpha = 0.95, \lambda = 0.99$	-178.750	79.670	87.277
	$h = 6$		
DMA, $\alpha = \lambda = 0.95$	-220.540	120.650	111.840
DMS, $\alpha = \lambda = 0.95$	-206.620	113.000	106.200
DMA, $\alpha = 0.99, \lambda = 0.95$	-218.360	115.300	109.920
DMS, $\alpha = 0.99, \lambda = 0.95$	-219.000	116.810	110.510
DMA, $\alpha = 0.95, \lambda = 0.99$	-218.790	118.440	110.380
DMS, $\alpha = 0.95, \lambda = 0.99$	-206.010	110.580	107.480
	$h = 12$		
DMA, $\alpha = \lambda = 0.95$	-225.310	111.110	107.840
DMS, $\alpha = \lambda = 0.95$	-208.850	104.940	104.790
DMA, $\alpha = 0.99, \lambda = 0.95$	-227.820	115.420	109.140
DMS, $\alpha = 0.99, \lambda = 0.95$	-224.910	115.480	109.640
DMA, $\alpha = 0.95, \lambda = 0.99$	-224.570	114.810	110.060
DMS, $\alpha = 0.95, \lambda = 0.99$	-208.250	104.740	105.730

Table A2: Sensitivity Analysis: Inflation			
Forecast Method	Sum of log pred. like.	MSFE	MAFE
	$h = 1$		
DMA, $\alpha = \lambda = 0.95$	0.061	9.572	31.964
DMS, $\alpha = \lambda = 0.95$	21.985	8.370	30.113
DMA, $\alpha = 0.99, \lambda = 0.95$	-9.986	9.531	31.843
DMS, $\alpha = 0.99, \lambda = 0.95$	-6.191	9.467	31.340
DMA, $\alpha = 0.95, \lambda = 0.99$	-3.552	9.593	32.691
DMS, $\alpha = 0.95, \lambda = 0.99$	15.689	8.916	32.311
	$h = 6$		
DMA, $\alpha = \lambda = 0.95$	-150.270	53.067	74.600
DMS, $\alpha = \lambda = 0.95$	-133.940	54.993	74.495
DMA, $\alpha = 0.99, \lambda = 0.95$	-163.430	54.974	75.781
DMS, $\alpha = 0.99, \lambda = 0.95$	-160.610	55.288	76.619
DMA, $\alpha = 0.95, \lambda = 0.99$	-189.360	70.320	85.178
DMS, $\alpha = 0.95, \lambda = 0.99$	-172.060	65.516	81.878
	$h = 12$		
DMA, $\alpha = \lambda = 0.95$	-170.730	65.627	84.929
DMS, $\alpha = \lambda = 0.95$	-158.820	70.852	86.663
DMA, $\alpha = 0.99, \lambda = 0.95$	-195.010	95.909	96.682
DMS, $\alpha = 0.99, \lambda = 0.95$	-196.180	102.770	100.560
DMA, $\alpha = 0.95, \lambda = 0.99$	-252.640	141.700	118.790
DMS, $\alpha = 0.95, \lambda = 0.99$	-237.050	144.090	119.210

Data Appendix

The following is the list of variables from the Bank of England's monthly forecasting suite. These are the variables for which complete data exists since 1990M1. The data runs from 1990M1 through 2008M11. The variables are listed in the blocks used to construct the factors. The acronym for each variable is the same as that used by the Bank of England. All variables are transformed to stationarity and the necessary transformation for each variable is also listed.

International Block

EAESINPRG: Europe (EA) industrial production, VOLA (source: DST). Transformation: percentage change.

EMESPISDH: Europe (EA) industrial production of durable consumer goods, WDA (source: DST). Transformation: difference.

EMESPISNH: Europe (EA) industrial production of non-durable consumer goods, WDA (source: DST). Transformation: percentage change.

EMOCIPEOG: Europe (EM) industrial production: manufacturing (source: DST). Transformation: percentage change.

EMOCIPIGG: Europe (EM) industrial production: investment goods (source: DST). Transformation: percentage change.

EMOCIPING: Europe (EM) industrial production: intermediate goods (source: DST). Transformation: percentage change.

EAESPISDG: Euro area (EUR13): Industrial production - Durable consumer goods, Index (2000=100), SA. (source: DST). Transformation: percentage change.

EAESPISNG: Euro area (EUR13): Industrial production - Non-durable consumer goods, Index (2000=100), SA. (source: DST). Transformation: percentage change.

USIP336VG: US industrial production - automobile and light duty motor vehicle, VOLA (source: DST). Transformation: percentage change.

USIPMAUPG: US industrial production - automotive products (consumer goods), VOLA (source: DST). Transformation: percentage change.

USIPMDUCG: US industrial production - durable consumer goods, VOLA (source: DST). Transformation: percentage change.

USIPMNOCG: US industrial production - non-durable consumer goods, VOLA (source: DST). Transformation: percentage change.

USOCIPMNG: US industrial production - manufacturing, VOLA (source: DST). Transformation: percentage change.

USUMCONSH: US consumer sentiment, volume index, not seasonally adjusted (source: DST). Transformation: percentage change.

USEMPNAGE: US Employed - nonfarm industries, total (payroll survey), VOLA (source: DST). Transformation: percentage change.

USESUNEMO: US unemployment rate, VOLA (source: DST). Transformation: difference.

USUNEMPP: US unemployed, total (16 years and over), VOLA (source: DST). Transformation: percentage change.

USCPF: US CPI - all urban sample, all items price index, not seasonally adjusted (source: DST). Transformation: percentage change.

Output Block

CBIEXP: CBI monthly trends enquiry: Excluding seasonal variations, do you consider that in volume terms, your present export order book is above normal? (source: CBI). Transformation: difference.

CBIFG: CBI monthly trends enquiry: Adequacy of stocks of finished goods (source: CBI). Transformation: difference.

CBIORD: CBI monthly trends enquiry: Excluding seasonal variations, do you consider that in volume terms, your present total order book is above normal? (source: CBI). Transformation: difference.

CBIOUT: CBI monthly trends enquiry: What, excluding seasonal variations, is the expected trend over the next 4 months with regards to your volume of output? (source: CBI). Transformation: difference.

CKYW: IOP: Industry C,D,E: All production industries: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKYX: IOP: Industry C: Mining & quarrying: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKYY: IOP: Industry D: Manufacturing: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKYZ: IOP: Industry E: Electricity, gas and water supply: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKZA: IOP: Industry DA: Manuf of food, drink & tobacco: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKZB: IOP: Industry DB: Manuf of textile & textile products: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKZC: IOP: Industry DC: Manuf of leather & leather products: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKZD: IOP: Industry DD: Manuf of wood & wood products: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKZE: IOP: Industry DE: Pulp/paper/printing/publishing industries: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKZF: IOP: Industry DF: Manuf coke/petroleum prod/nuclear fuels: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKZG: IOP: Industry DG: Manuf of chemicals & man-made fibres: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKZH: IOP: Industry DH: Manuf of rubber & plastic products: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKZI: IOP: Industry DI: Manuf of non-metallic mineral products: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKZJ: IOP: Industry DJ: Manuf of basic metals & fabricated prod: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKZK: IOP: Industry DK: Manuf of machinery & equipment: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKZL: IOP: Industry DL: Manuf of electrical & optical equipment: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

CKZM: IOP: Industry DM: Manuf of transport equipmnt: CVMSA NAYear=100 (source: ONS). Transformation: percentage change.

Price Block

CBIPR: CBI monthly trends enquiry: What, excluding seasonal variations, is the expected trend over the next 4 months with regards to average price for domestic orders? (source: CBI). Transformation: difference.

D7G7: CPI INDEX 00: ALL ITEMS- estimated pre-97 2005=100 (source: ONS). Transformation: percentage change (annual percentage change used for forecasting this variable).

D7G8: CPI INDEX 01: FOOD AND NON-ALCOHOLIC BEVERAGES 2005=100 (source: ONS). Transformation: percentage change.

D7G9: CPI INDEX 02: ALCOHOLIC BEVERAGES,TOBACCO & NARCOTICS- estimated pre-97 2005=100 (source: ONS). Transformation: percentage change.

D7GA: CPI INDEX 03: CLOTHING AND FOOTWEAR- estimated pre-97 2005=100 (source: ONS). Transformation: percentage change.

D7GB: CPI INDEX 04: HOUSING, WATER AND FUELS- estimated pre-97 2005=100 (source: ONS). Transformation: percentage change.

D7GC: CPI INDEX 05: FURN, HH EQUIP & ROUTINE REPAIR OF HOUSE- est. pre-97 2005=100 (source: ONS). Transformation: percentage change.

D7GD: CPI INDEX 06: HEALTH- estimated pre-97 2005=100 (source: ONS). Transformation: percentage change.

D7GE: CPI INDEX 07: TRANSPORT- estimated pre-97 2005=100 (source: ONS). Transformation: percentage change.

D7GF: CPI INDEX 08: COMMUNICATION- estimated pre-97 2005=100 (source: ONS). Transformation: percentage change.

D7GG: CPI INDEX 09: RECREATION & CULTURE- estimated pre-97 2005=100 (source: ONS). Transformation: percentage change.

D7GH: CPI INDEX 10: EDUCATION- estimated pre-97 2005=100 (source: ONS). Transformation: percentage change.

D7GI: CPI INDEX 11: HOTELS, CAFES AND RESTAURANTS- estimated pre-97 2005=100 (source: ONS). Transformation: percentage change.

D7GJ: CPI INDEX 12: MISCELLANEOUS GOODS AND SERVICES- estimated pre-97 2005=100 (source: ONS). Transformation: percentage change.

PLLA: PPI: 6292990000: NSI M & F purchased by Man: Excl FBPT Excl CCL NSA (source: ONS). Transformation: second log difference.

PLLU: PPI: 7209200000: Output of manufactured products (source: ONS). Transformation: second log difference.

PLLW: PPI: 7209299890: Prod of man ind excl.f,b, p & t SA (source: ONS). Transformation: second log difference.

PVNQ: PPI: 7209200890: NSO: All manufacturing (excluding duty), SA (source: ONS). Transformation: second log difference.

OILBRNI_P: Price of Brent Crude - 1 month fwd,, fob US$/BBL (source: DST). Transformation: percentage change.

OILBRNP_P: Price of Brent Crude - physical delivery, fob US$/BBL (source: DST). Transformation: percentage change.

Demand Block

MREVQ1: CBI distributive trades reported motor traders sales (source: CBI). Transformation: difference.

MREVQ2: CBI distributive trades reported motor traders orders (source: CBI). Transformation: difference.

MREVQ3: CBI distributive trades reported motor traders sales for time of year (source: CBI). Transformation: difference.

MREVQ4: CBI distributive trades reported motor traders stocks (source: CBI). Transformation: difference.

RETREQ1: CBI distributive trades reported retailing sales (source: CBI). Transformation: difference.

RETREQ2: CBI distributive trades reported retailing orders (source: CBI). Transformation: difference.

RETREQ3: CBI distributive trades reported retailing sales for time of year (source: CBI). Transformation: difference.

RETREQ4: CBI distributive trades reported retailing stocks (source: CBI). Transformation: difference.

WHREQ1: CBI distributive trades reported wholesaling sales (source: CBI). Transformation: difference.

WHREQ2: CBI distributive trades reported wholesaling orders (source: CBI). Transformation: difference.

WHREQ3: CBI distributive trades reported wholesaling sales for time of year (source: CBI). Transformation: difference.

WHREQ4: CBI distributive trades reported wholesaling stocks (source: CBI). Transformation: difference.

GFKBALSA: GFK consumer confidence aggregate balance. Transformation: difference.

GFKQ1: GFK consumer confidence: How does the financial situation of your household compare to what it was 12 months ago? Transformation: difference.

GFKQ2: GFK consumer confidence: How do you think the financial position of your household will change over the next 12 months? Transformation: difference.

GFKQ3: GFK consumer confidence: How do you think the general economic situation of this country has changed over the last 12 months? Transformation: difference.

GFKQ4: GFK consumer confidence: How do you think the general economic situation of this country will develop over the next 12 months? Transformation: difference.

GFKQ7: GFK consumer confidence: How do you think the level of unemployment will change over the next 12 months? Transformation: difference.

GFKQ8: GFK consumer confidence: Do you think that there is an advantage for people to make major purchases at this time? Transformation: difference.

GFKQ9: GFK consumer confidence: Over the next 12 months how do you think the amount of money you will spend on major purchases will compare with what you spent over the last 12 months? Transformation: difference.

EAPS: Retail sales index (RSI):Volume seasonally adjusted: All Retailers: All Business Index (source: ONS). Transformation: percentage change.

EAPT: RSI:Predominantly food stores (vol sa):All Business Index (source: ONS). Transformation: percentage change.

EAPU: RSI:Non-specialised stores (vol sa):All Business Index (source: ONS). Transformation: percentage change.

EAPW: RSI:Other non-food stores (vol sa):All Business Index (source: ONS). Transformation: percentage change.

EAPX: RSI:textiles:clothing:footwear (vol sa):All Business Index (source: ONS). Transformation: percentage change.

EAPY: RSI:Household goods stores (vol sa):All Business Index (source: ONS). Transformation: percentage change.

EAPZ: RSI:Non-store retailing & repair (vol sa):All Business Index (source: ONS). Transformation: percentage change.

EAQV: RSI:Value seasonally Adjusted:All Retailers:All Business Index (source: ONS). Transformation: percentage change.

EAQW: RSI:Predominantly food stores (val sa):All Business Index (source: ONS). Transformation: percentage change.

GMAZ: OS visits to UK :Earnings:#M-(Cur.Price-SA) (source: ONS). Transformation: percentage change.

GMBB: UK visits abroad:Expenditure abroad:#M-(Cur.Price-SA) (source: ONS). Transformation: percentage change.

BQKO: BOP:IM:CVM:SA:Total Trade in Goods (source: ONS). Transformation: percentage change.

BQKQ: BOP:EX:CVM:SA:Total Trade in Goods (source: ONS). Transformation: percentage change.

ELAR: BOP:IM:price index:NSA:Finished manufactures: SITC 7+8 (source: ONS). Transformation: percentage change.

Financial Block

FTALLSH_DY: FTSE All share dividend yield (source: DST). Transformation: percentage change.

FTALLSH_PI: FTSE All share price index (source: DST). Transformation: percentage change.

FTSE100_PI: FTSE 100 price index (source: DST). Transformation: percentage change.

A_JYS: Exchange rate: Japanese Yen/£, monthly average (source: FST). Transformation: percentage change.

A_SFS: Exchange rate: Swiss franc/£, monthly average (source: FST). Transformation: difference.

A_USS: Exchange rate: US$/£, monthly average (source: FST). Transformation: percentage change.

SERI: Exchange rate: Euro/£, monthly average, prior to Euro, use DM with conversion rate 1.9583 (source: FST). Transformation: difference.

A_AMIJ: 3 month £ inter-bank rate (mean LIBID/LIBOR), 10.30AM, monthly average (source: FST). Transformation: difference.

A_BEDR: Bank of England repo rate, monthly average (source: FST). Transformation: difference.

A_VOMA: Overnight £ inter-bank rate (mean LIBID/LIBOR), 8:30 AM, monthly average (source: FST). Transformation: difference.

A_VSMA: 6 month £ inter-bank rate (mean LIBID/LIBOR), 8:30 AM, monthly average (source: FST). Transformation: difference.

AJNB: Treasury bills: average discount rate (source: ONS). Transformation: difference.

VRPSPOTI10Y: 10 year VRP (variable roughness penaly model) spot rate (inflation), (source: BoE). Transformation: difference.

VRPSPOTI5Y: 5 year VRP spot rate (inflation), (source: BoE). Transformation: difference.

VRPSPOTN5Y: 5 year VRP spot rate (nominal), (source: BoE).Transformation: difference.

VRPSPOTR10: 10 year VRP spot rate (real). Transformation: difference.

VRPSPOTR5Y: 5 year VRP spot rate (real). Transformation: difference.

Housing Block

HFAXPA: Halifax house price index, all houses (all buyers), 1983 = 100, SA (source: HAC). Transformation: percentage change.

RICSPR: RICS housing market survey, prices, England and Wales, net balance, SA (source: HAC). Transformation: difference.

RICSSASTK: Ratio of RICS sales series to RICS stock series (source: HAC). Transformation: difference.

Money Block

AUZJ_AUYN: Amounts outstanding of monetary financial institutions' sterling M4 liabilities to private sector, SA (source: MST).

AVAG: Monthly average amount outstanding of total sterling notes and coin in circulation outside the Bank of England (in sterling millions), SA (source: MST). Transformation: percentage change.

VQJL: Monthly amounts outstanding of UK resident banks' (inc. Central Bank) sterling net lending to private sector (in sterling millions), SA (source: MST). Transformation: percentage change.

VQJM: Monthly amounts outstanding of monetary financial institutions' sterling net lending to private sector (in sterling millions), SA (source: MST). Transformation: percentage change.

VQXK: Money Stock: Retail Deposits and Cash in M4, SA (source: MST). Transformation: percentage change.

VQXL: Monthly amounts outstanding of UK resident banks' (inc. Central Bank) sterling retail deposits from private sector (in sterling millions), SA (source: MST). Transformation: percentage change.

VQZA: Money Stock: Retail Deposits and Cash in M4, NSA (source: MST). Transformation: percentage change.

Labour Block

BCJD: Total Claimant count SA (UK) - thousands (source: ONS). Transformation: difference.

BCJE: Claimant count rate - all - SA (UK) % (source: ONS). Transformation: difference.

LOKA: UK Employee jobs (SA): DA (15-16) Food products, beverages & tobacco (source: ONS). Transformation: percentage change.

LOKB: UK Employee jobs (SA): DB/DC (17-19) Manu. of clothing, textiles, leather (source: ONS). Transformation: percentage change.

LOKC: UK Employee jobs (SA): DD (20) Wood & wood products (source: ONS). Transformation: percentage change.

LOKD: UK Employee jobs (SA): DE (21-22) Paper, pulp, publishing & recording media (source: ONS). Transformation: percentage change.

LOKE: UK Employee jobs (SA): DG (24) Chemicals, chemical products & man-made fibres (source: ONS). Transformation: percentage change.

LOKF: UK Employee jobs (SA): DH (25) Rubber & plastic products (LMT B.12) (source: ONS). Transformation: percentage change.

LOKG: UK Employee jobs (SA): DI/DJ (26-28) Non-metallic mineral products, metals (source: ONS). Transformation: percentage change.

LOKH: UK Employee jobs (SA): DK (29) Machinery & eqpt. (LMT B.12) (source: ONS). Transformation: percentage change.

LOKI: UK Employee jobs (SA): DL (30-33) Electrical & optical eqpt. (LMT B.12) (source: ONS). Transformation: percentage change.

LOKJ: UK Employee jobs (SA): DM (34-35) Transport equipment (LMT B.12) (source: ONS). Transformation: percentage change.

LOKK: UK Employee jobs (SA): DF,DN (23,36-37) Coke, nuclear fuel & other manu. (source: ONS). Transformation: percentage change.

MGRZ: LFS: In employment: UK: All: Aged 16+: 000s:SA: Annual = Spring qtr (Mar to May) (source: ONS). Transformation: percentage change.

MGSC: LFS: Unemployed: UK: All: Aged 16+: 000s: SA: Annual = Spring qtr (Mar-May) (source: ONS). Transformation: percentage change.

YBUS: LFS: Total actual weekly hours worked (millions): UK: All: SA (source: ONS). Transformation: percentage change.

YEJF: Employee jobs: All jobs: Production Inds. (C-E): (000s): (SA): UK (source: ONS). Transformation: percentage change.

LNKY: AEI (including bonuses), private sector (source: ONS). Transformation: percentage change.

LNMQ: AEI (including bonuses), whole economy (source: ONS). Transformation: percentage change.