

# Dynamic Bayesian Predictive Synthesis in Time Series Forecasting

Kenichiro McAlinn\* & Mike West†

September 21, 2017

## Abstract

We discuss model and forecast combination in time series forecasting. A foundational Bayesian perspective based on agent opinion analysis theory defines a new framework for density forecast combination, and encompasses several existing forecast pooling methods. We develop a novel class of dynamic latent factor models for time series forecast synthesis; simulation-based computation enables implementation. These models can dynamically adapt to time-varying biases, miscalibration and inter-dependencies among multiple models or forecasters. A macroeconomic forecasting study highlights the dynamic relationships among synthesized forecast densities, as well as the potential for improved forecast accuracy at multiple horizons.

*JEL Classification:* C11; C15; C53; E37

*Keywords:* Agent opinion analysis, Bayesian forecasting, Density forecast combination, Dynamic latent factors models, Macroeconomic forecasting

---

\*Booth School of Business, University of Chicago, Chicago, IL 60637. Email: kenichiro.mcalinn@chicagobooth.edu

†Department of Statistical Science, Duke University, Durham, NC 27708-0251.

Email: Mike.West@duke.edu. Tel: +1 919 684 4210. Fax: +1 919 684 8594

# 1 Introduction

Recent research at the interfaces of applied/empirical macroeconomics and Bayesian methodology development reflects renewed interest in questions of model and forecast comparison, calibration, and combination. A number of creative ideas for forecast density pooling define new empirical models fitted by Bayesian methods (e.g. [Terui and van Dijk 2002](#); [Hall and Mitchell 2007](#); [Amisano and Giacomini 2007](#); [Hoogerheide et al. 2010](#); [Kascha and Ravazzolo 2010](#); [Geweke and Amisano 2011, 2012](#); [Billio et al. 2012, 2013](#); [Aastveit et al. 2014](#); [Kapetanios et al. 2015](#); [Aastveit et al. 2015](#); [Pettenuzzo and Ravazzolo 2016](#); [Negro et al. 2016](#)), with examples of improved forecast performance in studies in macroeconomics and finance. This growing body of research responds in part to the need to improve information flows to policy and decision makers at national and international levels. From a methodological viewpoint, the field is also stimulated by the increased availability of formal forecasting models that yield full density forecasts; whether these arise from multiple professional forecasters and models, and sets of competing econometric models, the result is a need to define formal, reliable methodology for integrating such predictive information. Importantly, such methods should allow for learning and integrating information about the forecasters or models, characterizing their anticipated biases, relationships, and dependencies among them, and how these characteristics change with time.

We contribute to this field with a class of models that implement *Bayesian predictive synthesis* (BPS) in a sequential, dynamic setting for time series forecasting. Our work is partly motivated by a concern for foundational/theoretical underpinnings of methodology. A forecast combination rule/algorithm may demonstrate success in a specific case study, but understanding potential conceptual and theoretical foundations is of interest in order to advance broader understanding—through transparency of implicit underlying assumptions—and hence open paths to possible generalizations of practical importance. Our study of this links to past literature on subjective Bayesian “agent/expert opinion analysis” (e.g. [Lindley et al. 1979](#); [West 1984](#); [Genest and Schervish 1985](#); [West 1988](#); [West and Crosse 1992](#); [West 1992](#); [Dawid et al. 1995](#); [French 2011](#)). In that context, a decision maker regards multiple models or forecasters as providers of “forecast data” to be used in her prior-posterior updating (see also [West and Harrison 1997](#), Sect 16.3.2). It turns out that this perspective—referred to as Bayesian predictive synthesis (BPS)—theoretically identifies the existence of a subclass of Bayesian updating rules in the form of latent factor models; explicitly, *the set of forecast densities are those of inherent latent factors linked to the outcome of interest*. This opens a path to exploring BPS approaches based on various assumptions about latent factor regression model forms. We explore this in some illuminating examples, showing special cases linked to various existing density pooling methods.

The direct time series extension involves *dynamic latent factor models* in which sequences of forecast densities define time-varying priors for inherent latent factor processes linked to the time series of interest. We discuss and develop this new class of models for dynamic BPS. A study of

forecasting U.S. inflation demonstrates application, showing how dynamic BPS is able to adaptively learn over time about time-varying biases and miscalibration of multiple models or forecasters, and generate useful insights into time-varying patterns of dependencies among them while also improving forecast accuracy. The study includes illuminating comparison with other standard and recent methods of forecast density pooling. Importantly, we argue for defining BPS models specific to forecasting goals, and demonstrate this in multi-step ahead forecasting in the case study.

Section 2 summarizes the foundations of BPS in agent opinion analysis theory free from the time series context, with key examples. Full details of the emerging class of dynamic latent factor models are then developed for time series in Section 3. In Section 4, we develop a detailed case study in forecasting quarterly U.S. inflation, comparisons to other approaches, and aspects of analysis of synthetic data. Additional comments in Section 5 conclude the paper. Three appendices provide full details of the MCMC methodology for Bayesian analysis of dynamic latent factor BPS models underlying the case study, as well as additional graphical summaries from that analysis and of a separate synthetic data set created to validate the results in the case study.

**Some notation:** We use lower case bold font for vectors and upper case bold font for matrices. Vectors are columns by default. Distributional notation  $y \sim N(f, v)$ ,  $\mathbf{x} \sim N(\mathbf{a}, \mathbf{A})$  and  $k \sim G(a, b)$  are for the univariate normal, multivariate normal and gamma distributions, respectively. The delta Dirac function is  $\delta_x(y)$ , the probability mass function of  $y$  degenerate at a point  $x$ . We use, for example,  $N(y|f, v)$  to denote the actual density function of  $y$  when  $y \sim N(f, v)$ . Index sets  $s:t$  stand for  $s, s+1, \dots, t$  when  $s < t$ , such as in  $y_{1:t} = \{y_1, \dots, y_t\}$ .

## 2 Background and Foundations

### 2.1 Subjective Bayesian Forecasting and Information Processing

We begin with background and examples free from the time series context, for clarity. Section 3 then moves to the dynamic setting with direct translation of the theory summarized here.

In agent opinion analysis (West 1992) a Bayesian decision maker  $\mathcal{D}$  aims to predict an outcome  $y$  and to use information from  $J$  individual agents (models, forecasters, or forecasting agencies, etc.) labelled  $\mathcal{A}_j$ , ( $j = 1:J$ ). To begin,  $\mathcal{D}$  has prior  $p(y)$ ; the agents provide her with forecast information in terms of their pdfs  $h_j(y)$ . These forecast densities represent the individual inferences from the agents; for  $\mathcal{D}$ , they are data and define the information set  $\mathcal{H} = \{h_1(\cdot), \dots, h_J(\cdot)\}$ . Thus  $\mathcal{D}$  will predict  $y$  using her implied posterior  $p(y|\mathcal{H})$ . Given the complex nature of  $\mathcal{H}$ —a set of  $J$  density functions, in a setting where there will be varying dependencies among agents as well as individual biases—a fully specified Bayesian model  $p(y, \mathcal{H}) = p(y)p(\mathcal{H}|y)$  is not easily conceptualized.

The relevant agent opinion analysis theory (Genest and Schervish 1985; West and Crosse 1992; West 1992) is summarized as follows. Before observing  $\mathcal{H}$ , suppose that  $\mathcal{D}$  specifies her prior  $p(y)$

and her *prior expectation* of the product of agent densities, denoted by  $m(\mathbf{x}) = E[\prod_{j=1:J} h_j(x_j)]$ . Here  $\mathbf{x} = x_{1:J} = (x_1, \dots, x_J)'$  is a vector of  $J$  dummy variables and, critically, the expectation is over the *a priori* uncertain functions  $h_j(\cdot)$ . That is,  $\mathcal{D}$  specifies only these two summary aspects of her full, but otherwise incompletely specified joint prior  $p(y, \mathcal{H})$ . Then, it follows that there exists a subset of Bayesian models  $p(y, \mathcal{H})$  under which the required posterior has the form

$$p(y|\mathcal{H}) = \int_{\mathbf{x}} \alpha(y|\mathbf{x}) \prod_{j=1:J} h_j(x_j) d\mathbf{x} \quad (1)$$

where  $\mathbf{x} = x_{1:J} = (x_1, \dots, x_J)'$  is a  $J$ -dimensional latent vector and  $\alpha(y|\mathbf{x})$  a conditional pdf for  $y$  given  $\mathbf{x}$ . Critically, this shows that there exist latent variables  $\mathbf{x}$  with the interpretations that: (i) the forecast distribution of agent  $j$  is that of latent variable  $x_j$ , (ii) the latent variables are generated independently from the  $h_j(\cdot)$ , and (iii) the latent variables relate as a group to the outcome  $y$  through a conditional (regression) distribution  $\alpha(y|\mathbf{x})$ .

Interpret the latent factor vector  $\mathbf{x}$  as follows. Suppose the (purely hypothetical) agent information  $h_j(y) = \delta_{x_j}(y)$  – delta Dirac functions – for  $j = 1:J$ . That is,  $\mathcal{A}_j$  makes a perfect prediction  $y = x_j$  for some specified value  $x_j$ .  $\mathcal{D}$ 's posterior is then  $\alpha(y|\mathbf{x})$  which reflects her views of  $y$  based on supposedly exact predicted values (or “oracle” values) from the agents. We refer to the  $x_j$  as the *latent agent states* and to  $\alpha(y|\mathbf{x})$  as  $\mathcal{D}$ 's *calibration function*.

Critically, the representation of eqn. (1) does not require a full specification of  $p(y, \mathcal{H})$ , and does not indicate what the functional form of  $\alpha(y|\mathbf{x})$  is; this opens the path to developing models based on different specifications of  $\alpha(y|\mathbf{x})$ . This posterior  $p(y|\mathcal{H})$  must be consistent with the specified  $p(y)$  and  $m(\mathbf{x})$ , seen by integrating eqn. (1) with respect to  $\mathcal{H}$  to yield  $p(y) = \int \alpha(y|\mathbf{x}) m(\mathbf{x}) d\mathbf{x}$ . Note the implied role and interpretation of  $m(\mathbf{x})$  here: prior to observing  $\mathcal{H}$ , the pdf  $m(\mathbf{x})$  is  $\mathcal{D}$ 's prior for the latent agent states; on observing  $\mathcal{H}$ , her posterior for  $\mathbf{x}$  is the product of the  $h_j(x_j)$ . This shows how  $\mathcal{D}$  can incorporate views, and historical information, about agent-specific biases, patterns of miscalibration, inter-dependencies among agents and their relative expertise and expected forecast accuracy implicitly through  $m(\mathbf{x})$  and/or more directly through specification of  $\alpha(y|\mathbf{x})$ . This, and the nature of the overall framework, is best seen through examples.

## 2.2 Examples: Mixture BPS and Connections to Other Methods

A first general point is that BPS defines a theoretical foundation for *any* pooling method that implicitly relies on the idea of latent factors generated from the  $h_j(\cdot)$ . Aastveit et al. (2015) is a key example, as their pooling method/model is explicitly constructed based on simulation from the  $h_j(\cdot)$  (using sequential Monte Carlo methods). These authors use a specific class of models that can now be seen as implicitly defining – in a time series extension – BPS models with particular (rather complicated and time-varying) calibration functions. The BPS theory indicates that this is now justified from a foundational Bayesian perspective.

An example class takes  $\alpha(y|\mathbf{x}) = a_0(\mathbf{x})\pi_0(y) + \sum_{1:J} a_j(\mathbf{x})\delta_{x_j}(y)$ , a mixture of point masses with a base prediction  $\pi_0(y)$  using *agent state-dependent probabilities*  $a_{0:J}(\mathbf{x})$ . With this specific form of calibration function, it follows from eqn. (1) that

$$p(y|\mathcal{H}) = q_0(y|\mathcal{H})\pi_0(y) + \sum_{1:J} q_j(y|\mathcal{H})h_j(y) \quad (2)$$

where the probability weights  $q_j(\mathcal{H}) = \int a_j(\mathbf{x})\delta_{x_j}(y) \prod_j h_j(x_j)d\mathbf{x}$  (for  $j = 1:J$ ) depend on  $\mathcal{H}$  and naturally act to recalibrate the contributions from each  $h_j(\cdot)$  modulo the specified forms of the calibration weights  $a_j(\mathbf{x})$ . Some specific examples connect with existing methods as well as to new and practically relevant features.

**(A) Constant weights:** If the calibration function mixture weights are constant,  $a_j(\mathbf{x}) = a_j$  for  $j = 1:J$ , the implied posterior in eqn. (2) has agent density  $j$  weights  $q_j(y|\mathcal{H}) = a_j$ . Thus, the BPS framework has several special cases: (i) the traditional Bayesian model averaging (BMA) approach, in which the  $a_j$  are assigned as model probabilities based on past data and experience; (ii) other mixture approaches in which  $\mathcal{D}$  is free to choose the weights  $a_j$  to optimize any objective function as in stylized optimal prediction pools (e.g. [Geweke and Amisano 2011](#)), or to specify them based on past performance in forecasting or decision problems (e.g. [Pettenuzzo and Ravazzolo 2016](#)).

**(B) Agent-specific outcome dependent weights:** Suppose the calibration function mixture weight for agent  $j$  depends only on latent agent factor  $j$ , i.e.,  $a_j(\mathbf{x}) = a_j(x_j)$  for  $j = 1:J$ . Then the implied posterior in eqn. (2) has agent density  $j$  weights  $q_j(y|\mathcal{H}) = a_j(y)$  which acts to modify and “recalibrate” the contribution of forecast  $h_j(y)$ . This defines an “outcome dependent” mixture posterior, the weights on agent forecast densities depending on  $y$ , i.e., defines “generalized density combination” methods ([Kapetanios et al. 2015](#)) as special cases.

**(C) General dependence weights:** Generally,  $a_j(\mathbf{x})$  can be chosen to incorporate  $\mathcal{D}$ ’s views of agent-specific biases, miscalibration and—critically and not represented in other approaches—dependencies among agents.  $\mathcal{D}$  might choose weight functions  $a_j(\mathbf{x})$  as “kernels” to correct for expected location biases, to decay at values of  $\mathbf{x}$  viewed as extreme under  $m(\mathbf{x})$ , to take smaller values in cases of expected high positive dependency among agents when the  $x_j$  values are clustered together, to reflect increased uncertainty in cases of high dispersion of the  $x_j$ , and so forth. In the posterior of eqn. (2), the weight function  $q_j(y|\mathcal{H})$  on  $h_j(y)$  is outcome dependent as in example (B) and [Kapetanios et al. \(2015\)](#), but now—depending on choices of  $a_j(\mathbf{x})$ —can capture aspects of cross-agent dependencies as well as agent-specific calibration features.

These examples serve as relevant background and connect to some prior approaches. Parallel research reported in [Johnson and West \(2016\)](#) develops Bayesian methodology for general classes of such mixture models, based on various assumed forms of calibration weights  $a_j(\mathbf{x})$ . For the current paper, we move to quite different and core examples based on normal theory (rather than mixtures) and that underlie our main methodological extensions to dynamic latent factor models.

### 2.3 Examples: Normal Latent Factor BPS

Directly relevant to the dynamic BPS models below, a class of examples arises when the implied joint prior  $\alpha(y|\mathbf{x})m(\mathbf{x})$  is multivariate normal or T, which easily and intuitively allow  $\mathcal{D}$  to incorporate views of agent biases, miscalibration, and dependencies. Suppose a case in which  $(y, \mathbf{x})$  are jointly normal with margins  $p(y) = N(y|f, q)$  and  $m(\mathbf{x}) = N(\mathbf{x}|\mathbf{m}, \mathbf{M})$  where: (i)  $\mathbf{m} = f\mathbf{1} + \mathbf{b}$  for some  $J$ -vector  $\mathbf{b} = (b_1, \dots, b_J)'$  and where  $\mathbf{1}$  is the  $J$ -vector of ones; (ii)  $\mathbf{M}$  is a  $J \times J$  variance matrix  $\mathbf{M}$  with diagonal elements  $qs_j$  for some scales  $s_j > 0$ . Here  $\mathcal{D}$  expects agent forecast distributions to be normal (although of course her full, but incompletely specific prior recognizes that is just an expectation), and includes explicit, agent-specific location bias  $b_j$  and scale bias  $s_j$  relative to  $\mathcal{D}$ 's prior. The covariances in  $\mathbf{M}$  explicitly recognize cross-agent dependencies; if  $\mathcal{D}$  regards agents  $i, j$  as very highly dependent, then the  $i, j$  correlation should be large and positive. The full joint normal for  $(y, \mathbf{x})$  is completed with a covariance vector  $\mathbf{r} = C(\mathbf{x}, y) = (r_1, \dots, r_J)'$ . The implied conditional  $\alpha(y|\mathbf{x})$  is then normal with  $E(y|\mathbf{x}) = f + \sum_{j=1:J} \theta_j(x_j - m_j)$  and  $V(y|\mathbf{x}) = v$  given by  $v = q - \sum_{j=1:J} r_j \theta_j$  where the  $\theta_j$  are the elements of the  $J$ -vector  $\mathbf{M}^{-1}\mathbf{r}$ .

Suppose now the agents provide normal forecasts,  $h_j(y) = N(y|h_j, H_j)$  for  $j = 1:J$ . Then  $\mathcal{D}$ 's posterior  $p(y|\mathcal{H})$  is, from eqn. (2), normal with  $E(y|\mathcal{H}) = f + \sum_{j=1:J} \theta_j(x_j - m_j)$  and variance  $V(y|\mathcal{H}) = v + \sum_{j=1:J} \theta_j^2 H_j$ . Thus, the posterior effectively corrects for quantified biases in location and scale for each agent, while also now incorporating the quantified views of cross-agent dependencies implicitly through the implied regression coefficients  $\theta_j$ . In special cases, this justifies approaches using simple linear pooling of point forecasts, but now with explicit bias corrections and uncertainty quantification via the full posterior. Also, the posterior mean and variance of  $p(y|\mathcal{H})$  above arise with non-normal agent densities when  $(h_j, H_j)$  are agent  $j$ 's mean and variance.

From a methodological viewpoint, it is the role of the bias and dependence parameters in defining the regression model  $\alpha(y|\mathbf{x})$  itself that is key. This calibration density can be rewritten as

$$\alpha(y|\mathbf{x}) = N(y|\mathbf{F}'\boldsymbol{\theta}, v) \quad \text{with} \quad \mathbf{F} = (1, \mathbf{x}')' \quad \text{and} \quad \boldsymbol{\theta} = (\theta_0, \theta_1, \dots, \theta_J)' \quad (3)$$

where  $\theta_0 = f - \sum_{j=1:J} \theta_j m_j$ . Thus, the practically relevant *effective* calibration parameters are  $(\boldsymbol{\theta}, v)$ , and historical data will inform  $\mathcal{D}$  on these. This defines the starting point for analysis and extension to time series contexts, as now follows.

## 3 Dynamic BPS using Latent Factor Regression Models

The new methodological developments forming the core of this paper adapt and extend the basic BPS framework to sequential forecasting in time series. In particular, we develop dynamic extensions of Section 2.3 involving time-varying parameters to characterize and formally allow for agent-specific biases, patterns of miscalibration, inter-dependencies, and relative expertise/forecast accuracy as time evolves and data is processed. We do this in the context of a scalar time series.

Among the connections in recent literature we have already noted that [Hoogerheide et al. \(2010\)](#) and [Aastveit et al. \(2015\)](#) relate directly in key aspects of technical structure. In addition to opportunities for time-varying parameter models– a special case of the broader DLM setting developed in the following sections– these authors develop empirical methods using forecasts *simulated* from the sets of agent models. Based on the BPS theory, this is now understood to be foundationally justified since agent forecast distributions are those of implied latent factors that relate to the outcome of interest. As we see below, practical Bayesian analysis of dynamic BPS models naturally involves simulation of these latent states from the agent distributions in forecasting computations; however, they must be simulated from different distributions– the appropriate conditional posteriors– for model fitting and analysis.

### 3.1 Dynamic Sequential Setting

The decision maker  $\mathcal{D}$  is sequentially predicting a time series  $y_t, t = 1, 2, \dots$ , and at each time point receives forecast densities from each agent. At each time  $t - 1$ ,  $\mathcal{D}$  aims to forecast  $y_t$  and receives current forecast densities  $\mathcal{H}_t = \{h_{t1}(y_t), \dots, h_{tJ}(y_t)\}$  from the set of agents. The full information set used by  $\mathcal{D}$  is thus  $\{y_{1:t-1}, \mathcal{H}_{1:t}\}$ . As data accrues,  $\mathcal{D}$  learns about relationships among agents, their forecast and dependency characteristics, so that a Bayesian model will involve parameters that define the BPS framework and for which  $\mathcal{D}$  updates information over time. The implication for the temporal/dynamic extension of the BPS model of Section 2 is that  $\mathcal{D}$  has a time  $t - 1$  distribution for  $y_t$  of the form

$$p(y_t | \Phi_t, y_{1:t-1}, \mathcal{H}_{1:t}) \equiv p(y_t | \Phi_t, \mathcal{H}_t) = \int \alpha_t(y_t | \mathbf{x}_t, \Phi_t) \prod_{j=1:J} h_{tj}(x_{tj}) dx_{tj} \quad (4)$$

where  $\mathbf{x}_t = x_{t,1:J}$  is a  $J$ -dimensional latent agent state vector at time  $t$ ,  $\alpha_t(y_t | \mathbf{x}_t, \Phi_t)$  is  $\mathcal{D}$ 's conditional calibration pdf for  $y_t$  given  $\mathbf{x}_t$ , and  $\Phi_t$  represents time-varying parameters defining the calibration pdf– parameters for which  $\mathcal{D}$  has current beliefs represented in terms of a current (time  $t - 1$ ) posterior  $p(\Phi_t | y_{1:t-1}, \mathcal{H}_{1:t-1})$ . The methodological focus can now rest on evaluation of models based on various assumptions about the form of  $\alpha_t(y_t | \mathbf{x}_t, \Phi_t)$  and its defining dynamic state parameters  $\Phi_t$ . Naturally, we look to tractable dynamic linear regression models, a subset of the broader class of dynamic linear models, or DLMs ([West and Harrison 1997](#); [Prado and West 2010](#)), as a first approach to defining a computationally accessible yet flexible framework for dynamic BPS.

### 3.2 Latent Factor Dynamic Linear Models

Consider a dynamic regression for BPS calibration that extends the basic example of eqn. (3) to the time series setting. That is, eqn. (3) becomes the dynamic version

$$\alpha_t(y_t | \mathbf{x}_t, \Phi_t) = N(y_t | \mathbf{F}'_t \boldsymbol{\theta}_t, v_t) \quad \text{with} \quad \mathbf{F}_t = (1, \mathbf{x}'_t)' \quad \text{and} \quad \boldsymbol{\theta}_t = (\theta_{t0}, \theta_{t1}, \dots, \theta_{tJ})', \quad (5)$$

the latter being the  $(1 + J)$ -vector of time-varying bias/calibration coefficients and the conditional variance  $v_t$  defining residual variation beyond that explained by the regression on latent agent factors. Explicitly, the functional model parameters are now  $\Phi_t = (\theta_t, v_t)$  at each time  $t$ . This BPS specification defines the first component of the standard conjugate form dynamic linear model (DLM) ([West and Harrison 1997](#), Section 4)

$$y_t = \mathbf{F}'_t \boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim N(0, v_t), \quad (6a)$$

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N(0, v_t \mathbf{W}_t) \quad (6b)$$

where  $\boldsymbol{\theta}_t$  evolves in time according to a linear/normal random walk with innovations variance matrix  $v_t \mathbf{W}_t$  at time  $t$ , and  $v_t$  is the residual variance in predicting  $y_t$  based on past information and the set of agent forecast distributions. The residuals  $\nu_t$  and evolution *innovations*  $\boldsymbol{\omega}_s$  are independent over time and mutually independent for all  $t, s$ .

The DLM specification is completed using standard discount factor based methods, long used in the core Bayesian forecasting literature (e.g. [West and Harrison 1997](#); [Prado and West 2010](#)) and of increasing use in econometric and financial applications in more recent times, with discount factors sometimes now referred to as “forgetting factors” (e.g. [Dangl and Halling 2012](#); [Koop and Korobilis 2013](#); [Gruber and West 2016, 2017](#); [Zhao et al. 2016](#)). First, the time-varying intercept and agent coefficients  $\boldsymbol{\theta}_t$  follow the random walk evolution of eqn. (6b) where  $\mathbf{W}_t$  is defined via a standard, single discount factor specification ([West and Harrison 1997](#), Sect 6.3; [Prado and West 2010](#), Sect 4.3), using a state evolution discount factor  $\beta \in (0, 1]$ . Second, the residual variance  $v_t$  follows a standard beta-gamma random walk volatility model ([West and Harrison 1997](#), Sect 10.8; [Prado and West 2010](#), Sect 4.3), with  $v_t = v_{t-1}\delta/\gamma_t$  for some discount factor  $\delta \in (0, 1]$  and where  $\gamma_t$  are beta distributed innovations, independent over time and independent of  $\nu_s, \boldsymbol{\omega}_r$  for all  $t, s, r$ . Given choices of discount factors underlying these two components, and a (conjugate normal/inverse-gamma) initial prior for  $(\boldsymbol{\theta}_0, v_0)$  at  $t = 0$ , the model is fully specified.

Eqns. (6) define a dynamic latent factor model: the  $\mathbf{x}_t$  vectors in each  $\mathbf{F}_t$  are latent variables. At time  $t - 1$ , the set of agent densities becomes available for forecasting  $y_t$ ; then, from the BPS foundation, each  $x_{tj}$  is a latent draw from  $h_{tj}(\cdot)$ . Note that the latent factor generating process has the  $x_{tj}$  drawn *independently* from their  $h_{tj}(\cdot)$  and externally to the BPS model. That is, coupled with eqns. (6a,6b), we have

$$p(\mathbf{x}_t | \Phi_t, y_{1:t-1}, \mathcal{H}_{1:t}) \equiv p(\mathbf{x}_t | \mathcal{H}_t) = \prod_{j=1:J} h_{tj}(x_{tj}) \quad (7)$$

for all time  $t$  and with  $\mathbf{x}_t, \mathbf{x}_s$  conditionally independent for all  $t \neq s$ . Critically, the independence of the  $x_{tj}$  *conditional on the*  $h_{tj}(\cdot)$  must not be confused with the question of  $\mathcal{D}$ 's modeling and estimation of the *dependencies among agents*; this is simply central and integral, and reflected through the effective DLM parameters  $\Phi_t = (\boldsymbol{\theta}_t, v_t)$ .

We use standard DLMs to define synthesis functions  $\alpha_t(y_t | \mathbf{x}_t, \Phi_t)$  in view of the central role of such models in Bayesian forecasting, and the flexibility they have in adapting to time-varying structure and relationships. That said, the BPS theory does not imply any specific structure for the synthesis functions; the decision maker  $\mathcal{D}$  is free to make alternative model specifications.

Some methods of point forecast combination restrict weights on forecasts to the unit simplex and to sum to one, i.e., as probabilities. In the dynamic BPS setting, restricting the dynamic regression coefficients would be technically challenging, but more importantly we do not see it as relevant. We can expect such constraints to lead to BPS models that underperform compared to the unrestricted case. For example, consider the case where all agents overestimate the quantity of interest by some positive value. Under the restrictive case, there is no combination of weights that can achieve that quantity, while the unrestricted analysis will appropriately adapt. If of interest, the posteriors for  $\theta_t$  over time can of course be explored to investigate whether or not the model:data combination provides support for any kind of such restriction.

### 3.3 Bayesian Analysis and Computation

At any current time  $t$ ,  $\mathcal{D}$  has available the history of the BPS analysis to that point, including the now historical information  $\{y_{1:t}, \mathcal{H}_{1:t}\}$ . Over times  $1:t$ , the BPS analysis will have involved inferences on both the latent agent states  $\mathbf{x}_{1:T}$  as well as the dynamic BPS model parameters  $\Phi_{1:T}$ . Importantly, inferences on the former provide insights into the nature of dependencies among the agents, as well as individual agent forecast characteristics. The former addresses key and topical issues of overlap and redundancies among groups of forecasting models or individuals, as well as information sharing and potential herding behavior within groups of forecasters. The “output” of full posterior summaries for the  $\mathbf{x}_t$  series is thus a key and important feature of BPS.

For posterior analysis, the holistic view is that  $\mathcal{D}$  is interested in computing the posterior for the full set of past latent agent states (latent factors) and dynamic parameters  $\{\mathbf{x}_{1:t}, \Phi_{1:t}\}$ , rather than restricting attention to forward filtering to update posteriors for current values  $\{\mathbf{x}_t, \Phi_t\}$ ; the latter is of course implied by the former. This analysis is enabled by Markov chain Monte Carlo (MCMC) methods, and then forecasting from time  $t$  onward follows by theoretical and simulation-based extrapolation of the model; both aspects involve novelties in the BPS framework but are otherwise straightforward extensions of traditional methods in Bayesian time series ([West and Harrison 1997](#), Chap 15; [Prado and West 2010](#)).

**Posterior Computations via MCMC.** The dynamic latent factor model of eqns. (6a,6b,7) leads to a two-component block Gibbs sampler for sets of the latent agent states  $\mathbf{x}_t$  and DLM dynamic parameters  $\Phi_t$ . These are iteratively resimulated from two conditional posteriors noted below, with obvious initialization based on agent states drawn independently from priors  $h_*(*)$ .

First, conditional on values of agent states, the next MCMC step draws new parameters from

$p(\Phi_{1:t} | \mathbf{x}_{1:t}, y_{1:t})$ . By design, this is a discount-based dynamic linear regression model, and sampling uses the standard forward filtering, backward sampling (FFBS) algorithm (e.g. Frühwirth-Schnatter 1994; West and Harrison 1997, Sect 15.2; Prado and West 2010, Sect 4.5).

Second, conditional on values of dynamic parameters, the MCMC draws new agent states from  $p(\mathbf{x}_{1:t} | \Phi_{1:t}, y_{1:t}, \mathcal{H}_{1:t})$ . It is immediate that the  $\mathbf{x}_t$  are conditionally independent over time  $t$  in this conditional distribution, with time  $t$  conditionals

$$p(\mathbf{x}_t | \Phi_t, y_t, \mathcal{H}_t) \propto N(y_t | \mathbf{F}'_t \boldsymbol{\theta}_t, v_t) \prod_{j=1:J} h_{tj}(x_{tj}) \quad \text{where } \mathbf{F}_t = (1, x_{t1}, x_{t2}, \dots, x_{tJ})'.$$

In cases when each of the agent forecast densities is normal, this yields a multivariate normal for  $\mathbf{x}_t$  that is trivially sampled. In other cases, this will involve either a Metropolis-Hastings simulator or an augmentation method. A central, practically relevant case is when agent forecasts are T distributions; each  $h_{tj}(\cdot)$  can then be represented as a scale mixture of normals, and augmenting the posterior MCMC to include the implicit underlying latent scale factors generates conditional normals for each  $x_t$  coupled with conditional inverse gammas for those scales. This is again a standard MCMC approach and much used in Bayesian time series, in particular (e.g. Frühwirth-Schnatter 1994; West and Harrison 1997, Chap 15).

Full technical details of the MCMC computations, and additional discussion, is given in the supplementary Appendix A.

**Forecasting 1-Step Ahead.** At time  $t$  we forecast 1-step ahead by generating “synthetic futures” from the BPS model, as follows: (i) For each sampled  $\Phi_t$  from the posterior MCMC above, draw  $v_{t+1}$  from its discount volatility evolution model, and then  $\boldsymbol{\theta}_{t+1}$  conditional on  $\boldsymbol{\theta}_t, v_{t+1}$  from the evolution model eqn. (6b)– this gives a draw  $\Phi_{t+1} = \{\boldsymbol{\theta}_{t+1}, v_{t+1}\}$  from  $p(\Phi_{t+1} | y_{1:t}, \mathcal{H}_{1:t})$ ; (ii) Draw  $\mathbf{x}_{t+1}$  via independent sampling of the  $h_{t+1,j}(x_{t+1,j})$ , ( $j = 1:J$ ); (iii) Draw  $y_{t+1}$  from the conditional normal of eqn. (6a) given these sampled parameters and agent states. Repeating this generates a random sample from the 1-step ahead forecast distribution for time  $t + 1$ .

### 3.4 Multi-Step Ahead Forecasting

Forecasting over multiple horizons is often of greater importance than 1-step ahead forecasting. Economic policy makers, for example, forecast/assess macroeconomic variables over a year or multiple years, drawing from their own forecast models, judgemental inputs, other economists and forecasters, in order to advise policy decisions. However, forecasting over longer horizons is typically more difficult than over shorter horizons, and models calibrated on the short-term basis can often be quite poor in the longer-term. As noted in Section 1, fitting of time series models is inherently based on 1-step ahead, as DLM (and other) model equations make explicit. So, when multi-step ahead forecasting is primary, new ideas for forecast calibration and combination are needed.

BPS provides a natural and flexible framework to synthesize forecasts over multiple horizons, with potential to improve forecasting at multiple horizons simultaneously, as we now discuss.

**Direct projection for multi-step forecasting.** At time  $t$ , the agents provide  $k$ -step ahead forecast densities  $h_{t,1:J}(x_{t+k})$ . The direct approach follows traditional DLM updating and forecasting via simulation as for 1-step ahead. That is: (i) project the BPS model forward from time  $t$  to  $t+k$  by simulating the dynamic model parameters  $\Phi_{t+1}, \Phi_{t+2}, \dots, \Phi_{t+k}$  using sequential, step ahead extension of the 1-step case; (ii) draw independently from each of the  $h_{t,1:J}(x_{t+k})$  to give a sampled vector  $x_{t+k}$ ; then (iii) draw  $y_{t+k}$  from the conditional normal given these sampled parameters and states. While this is theoretically correct, it fails to update/calibrate based on the horizon of interest, relying wholly on the model as fitted— with its essential basis in 1-step forecasting accuracy— even though  $\mathcal{D}$  may be mainly interested in forecasting several steps ahead.

**BPS( $k$ ) for customized multi-step forecasting.** BPS can be customized to forecasting goals, allowing  $\mathcal{D}$  to focus on the horizon of interest. This responds in part to the reality that a model that forecasts well in the short-term may be useless for multiple steps ahead, while another model may be ideal for several steps ahead but poorer in the short-term. As noted by others (e.g. [Aastveit et al. 2014](#), and references therein) this calls for combination methods specific to the forecast horizon. We therefore consider multiple BPS models in parallel for different forecast horizons.

This involves a modification of Section 3 in which the model at time  $t-1$  for predicting  $y_t$  changes as follows. With a *specific forecast horizon*  $k > 1$ , modify the BPS so that the agents'  $k$ -step ahead forecast densities made at time  $t-k$ , i.e.,  $h_{t-k,j}(x_{tj})$  replace  $h_{tj}(x_{tj})$  in the resulting model analysis. This changes the interpretation of the dynamic model parameters  $\{\theta_t, v_t\}$  to be explicitly geared to the  $k$ -step horizon. Bayesian model fitting then naturally “tunes” the model to the horizon  $k$  of interest. Forecasting the chosen  $k$ -steps ahead now simply involves extrapolating the model via simulation, as above, but now in this modified and horizon-specific BPS model.

We denote this customized model strategy by BPS( $k$ ) to distinguish it from the direct extrapolation of BPS. Note that this is fundamentally different from the traditional method of model extrapolation as it directly updates, calibrates, and learns using the horizon of interest. The applied study in Section 4 below bears out the view that this can be expected to improve forecasting accuracy over multiple horizons. One cost, of course, is that a bank of BPS models is now required for any set of horizons of interest; that is, different models will be built for different horizons  $k$ , so increasing the computational effort required. We do not link analyses across horizons, i.e., do not relate dynamic BPS model parameters in BPS(4) to those of BPS(1), for example. A model that is good at short-term forecasting may be hopeless at longer horizons, and vice-versa, so there is no notion of a formal relationship between BPS models, and their parameters, at different horizons.

We further note contextual relevance of this perspective in macroeconomics when  $\mathcal{D}$  is a consumer of forecasts from groups, agencies or model developers. Such agents may use different

models, data, advisors, and approaches for different horizons. When the forecast generating models/methods are known,  $\mathcal{D}$  may redefine the BPS model accordingly; however, generally in practice these underlying models, strategies, data, and advisors will not be wholly known and understood.

## 4 US Macroeconomic Time Series

### 4.1 Data, Forecasting Models and Implementation

**Time Series Data.** We analyze quarterly U.S. macroeconomic data, focusing on forecasting inflation rates with both 1-quarter and 4-quarter ahead interests. The study involves three quarterly macro series: annual inflation rate ( $p$ ), short-term nominal interest rate ( $r$ ), and unemployment rate ( $u$ ) in the U.S. economy from 1961/Q1 to 2014/Q4, a context of topical interest (Cogley and Sargent 2005; Primiceri 2005; Koop et al. 2009; Nakajima and West 2013a). The inflation rate is the annual percentage change in a chain-weighted GDP price index, the interest rate is the yield on three-month Treasury bills, and the unemployment rate is seasonally adjusted and includes all workers over 16 years of age. Data are recorded as the latest last vintage when the data was collected (2015/1Q). Prior studies (e.g. Nakajima and West 2013a) use data over the period of 1963/Q1-2011/Q4; we extend this to more recent times, 1961/Q1 to 2014/Q4 (see data display in supplementary Appendix B). We focus on forecasting inflation using past values of the three indices as candidate predictors underlying a set of four time series models– the  $J = 4$  agents– to be evaluated, calibrated, and synthesized. The time frame includes key periods that warrant special attention: the early 1990s recession, the Asian and Russian financial crises in the late 1990s, the dot-com bubble in the early 2000s, and the sub-prime mortgage crisis and great recession of the late 2000s. These periods exhibit sharp shocks to the U.S. economy generally, and test the predictive ability of any models and strategies under stress. For any forecast calibration and aggregation strategy to be effective and useful, its predictive performance must be robust under these conditions; most traditional macroeconomic models suffer significant deficiencies in such times.

**Agent Models and BPS Specification.** The  $J = 4$  agents represent the two major structures of time series forecast models: factor and lag. Labeling them  $M^*$ , the agent models for inflation  $y_t \equiv p_t$  use predictors: M1-  $p_{t-1}$ ; M2-  $p_{t-1:3}, r_{t-1:3}, u_{t-1:3}$ ; M3-  $p_{t-1:3}$ ; M4-  $p_{t-1}, r_{t-1}, u_{t-1}$ . Thus, each has a time-varying autoregressive term in inflation rate  $p$ , while two also have dynamic regressions on lagged interest rate  $r$  and unemployment rate  $u$ , the differences being in lags chosen and model complexity. In each, residual volatility follows a standard beta-gamma random walk. Each  $M^*$  is a standard DLM so that model fitting and generation of forecasts is routine. While these models are simple compared to more sophisticated models used to forecast inflation (such as Stock and Watson 1999, 2007; Stella and Stock 2012; Belmonte et al. 2014 as well as Bayesian TVP-VARs seen in Nakajima and West 2013a), part of the utility and appeal of predictive synthesis, and in

forecast combinations in general, is in gaining improvements using relatively simple models and not resorting to complicated models that have technical/computational difficulties. Additionally, we can expect that adding these sophisticated models to the set of agents will only improve the synthesis. That said, the chosen models are, in fact, not so trivial in terms of forecasting performance *per se*, and most more elaborate inflation models are limited in forecasting ability, especially at longer horizons. So, while the selected simpler univariate models are not by any means “proposed” as state-of -the-art, they are of non-trivial interest as well as providing inputs to the main subject of the paper and, as we detail below, serve the main purposes in demonstrating application of dynamic latent factor BPS models in this now foundationally defined Bayesian approach. Using these easily implemented models, we see the BPS analysis exhibiting the ability to adapt to time-varying biases in individual models, automatically adapt/reweight models differently at differing time periods based on “local” performance, identify dependencies among forecast models (and adapt the synthesized forecasts to them), and define improved forecasting at multiple horizons with BPS models customized to horizon.

Prior specifications for the DLM state vector and discount volatility model in each of the  $J$  agent models is based on  $\theta_0|v_0 \sim N(\mathbf{0}, (v_0/s_0)\mathbf{I})$  and  $1/v_0 \sim G(n_0/2, n_0s_0/2)$  with  $n_0 = 2, s_0 = 0.01$ , using the usual  $(\theta, v)$  DLM notation ([West and Harrison 1997](#), Chap 4). Each agent model uses standard discount factor ( $\beta$ ) specification for state evolution variances and discount factor ( $\delta$ ) for residual volatility; we use  $(\beta, \delta) = (0.99, 0.95)$  in each of these agent models. All DLM-based agent forecast densities  $h_*(\cdot)$  are then those of predictive T distributions.

In the dynamic BPS models for forecast horizons  $k = 1$  and  $k = 4$ , we take initial priors as  $\theta_0|v_0 \sim N(\mathbf{m}_0, (v_0/s_0)\mathbf{I})$  with  $\mathbf{m}_0 = (0, 1'/J)'$  and  $1/v_0 \sim G(n_0/2, n_0s_0/2)$  with  $n_0 = 10, s_0 = 0.002$ . BPS for 1-step ahead forecasting is based on  $(\beta, \delta) = (0.95, 0.99)$ , while BPS(4), customized to 4-quarter ahead forecasting as discussed in Section 3.4, uses  $\theta_0|v_0 \sim N(\mathbf{m}_0, 10^{-4}(v_0/s_0)\mathbf{I})$  and  $(\beta, \delta) = (0.99, 0.99)$ . Differences by forecast horizon echo earlier discussion about different model choices being relevant to different forecast goals.

In general, discount factors should be set between  $0.9 - 0.99$ , expressing views on the variability of  $\theta_t$ , through  $v_t W_t$  (via  $\beta$ ), and  $v_t$  (via  $\delta$ ). If the decision maker  $\mathcal{D}$  believes that the synthesis will benefit from weights that are extremely flexible (e.g. where good agents change over multiple periods of time), then she can specify a lower discount factor  $\beta$ . In responding to dynamics, taking discount factors too small will lead to overly adaptive models that may show some small improvements in short-term forecasting accuracy but be quite poor at longer horizons, whereas discount values too close to 1 lead to under-adaptive and “over-smoothing” models. Relevant choices are, of course, always context dependent. In this example, we can assume that the simple models used will significantly underperform in the long term forecasts, as they overfit to short term dynamics. For this reason, we set a higher discount factor  $\beta$  to smooth out the parameters because we can expect that there is very little signal to extract from the agents.

The prior mean  $\mathbf{m}_0$  of  $\theta_0$  equally weights the latent agent states as a default neutral position,

while the prior conditional variances are large, and the degrees of freedom low, so that initial data will have a substantial impact in modifying the implied, relatively vague priors. We have explored analyses across ranges of priors and discount factors, and chosen these values as they lead to good agent-specific and BPS forecasting accuracy; our central conclusions with respect to BPS do not change materially with different values close to those chosen for the summary examples. It is always relevant to use rather diffuse initial priors and to experiment with those and discount factors on an initial period of training data, and then monitor forecasting performance over time based on a chosen model and its defining discount factors. The sequential framework allows for intervention to change model posteriors, and fixed parameters including the key discount factors, and practitioners should always be open to intervention opportunities. Looking at the behavior of the parameters during an initial learning period with training data gives insight into— among other things— choice of discount factors and prior specifications. For example, if the coefficients look jumpy/static during the learning period, it might be wise to increase/decrease the discount factor  $\beta$  to achieve a good level of adaptability.

**Data Analysis and Forecasting.** The 4 agent models are analyzed and synthesized as follows. First, the agent models are analyzed in parallel over 1961/Q1-1977/Q1 as a training period, simply running the DLM forward filtering to the end of that period to calibrate the agent forecasts. This continues over 1977/Q2-1989/Q4 now accompanied by the calibration of the other forecast combination methods. Also, at each quarter  $t$  during this period, the MCMC-based BPS analysis is run using data from 1977/Q2 up to time  $t$ ; that is, we repeat the analysis with an increasing “expanding window” of past data as we move forward in time. We do this for the traditional 1-step ahead focused BPS model, and— separately and in parallel— for a 4-step ahead focused BPS(4) model, as discussed in Section 3.4. This continues over the third period to the end of the series, 1990/Q1-2014/Q4; now we also record and compare forecasts as they are sequentially generated. This testing period spans over a quarter century, and we are able to explore predictive performance over periods of drastically varying economic circumstances, check robustness, and compare benefits and characteristics of each strategy. Out-of-sample forecasting is thus conducted and evaluated in a way that mirrors the realities facing decision and policy makers.

**Forecast Accuracy and Comparisons.** We compare forecasts from BPS with standard Bayesian model uncertainty analysis (or Bayesian model averaging— BMA) in which the agent densities are mixed with respect to sequentially updated model probabilities (e.g. [Harrison and Stevens 1976](#); [West and Harrison 1997](#), Sect 12.2), and with the state-of-the-art density combination method of [Billio et al. \(2013\)](#) using the DeCo package ([Casarin et al. 2015](#)). In addition, we compare with simpler, equally-weighted averages of agent forecast densities using both linear pools (equally-weighted arithmetic means of forecast densities) and logarithmic pools (equally-weighted harmonic means of forecast densities), with some theoretical underpinnings (e.g. [West 1984](#)). While these

strategies might seem overly simplistic, they have been shown to dominate some more complex aggregation strategies in some contexts, at least in terms of direct point forecasts in empirical studies (e.g. [Genre et al. 2013](#)). For point forecasts from all methods, we compute and compare mean squared forecast error (MSFE) over the forecast horizons of interest. In comparing density forecasts with BPS, we also evaluate log predictive density ratios (LPDR); at horizon  $k$  and across time indices  $t$ , this is

$$\text{LPDR}_{1:t}(k) = \sum_{i=1:t} \log\{p_s(y_{t+k}|y_{1:t})/p_{\text{BPS}}(y_{t+k}|y_{1:t})\}$$

where  $p_s(y_{t+k}|y_{1:t})$  is the predictive density under model or model combination aggregation strategy indexed by  $s$ , compared against the corresponding BPS forecasts at this horizon. As used by several authors recently (e.g. [Nakajima and West 2013a](#); [Aastveit et al. 2015](#)), LPDR measures provide a direct statistical assessment of relative accuracy at multiple horizons that extend traditional 1-step focused Bayes' factors. They weigh and compare dispersion of forecast densities along with location, to elaborate on raw MSFE measures; comparing both measurements, i.e., point and density forecasts, gives a broader understanding of the predictive abilities of the different strategies.

## 4.2 Dynamic BPS and Forecasting

Comparing predictive summaries over the out-of-sample period, BPS improves forecasting accuracy relative to the 4 agent models, and dominates BMA, DeCo, and the pooling strategies; see numerical summaries in Table 1. Looking at point forecast accuracy, BPS exhibits improvements of no less than 10% over all models and strategies for 1- and 4-step ahead forecasts ( $\text{BPS}(k)$  at  $k = 4$  for the latter). As might be expected, BPS substantially improves characterization of forecast uncertainties as well as adaptation in forecast locations, reflected in the LPDR measures. Further, our expectations of improved multi-step forecasting using horizon-specific BPS are borne out: direct projection of the standard BPS model to 4-step ahead forecasts perform poorly, mainly as a result of under-dispersed forecast densities from each agent. In contrast,  $\text{BPS}(4)$  model performs substantially better, being customized to the 4-quarter horizon.

We further our analysis by reviewing summary graphs showing aspects of analyses evolving over time during the testing period, a period that includes challenging economic times that impede good predictive performance. We take the 1-step and 4-step contexts in sequence. Additional summaries and figures appear in supplementary Appendix B.

**1-Step Ahead Forecasting.** Several figures summarize sequential analysis for 1-step forecasting. Fig. 1 confirms that BPS performs uniformly better than, or on par with, the agent models and other methods based on LPDR measures that reflect relative dispersion of forecast densities as well as location. Major shocks and times of increased volatility have substantial impact on the relative

performance. Four notable “shock” periods are: 1992/Q3-Q4 (early 90s recession), 1997/Q4-1998/Q1 (Asian and Russian financial crisis), 2001/Q2-2003/Q1 (dot-com bubble), and 2009/Q2-2010/Q1 (sub-prime mortgage crisis). Even under the influence of these shocks, BPS is able to perform well with most of its improvements over other models and strategies coming from swift adaptation.

Fig. 2 compares on-line 1-step ahead forecast standard deviations. Economic (and other) decision makers are often faced with forecasts that have large forecast uncertainties; while honest in reflecting uncertainties, resulting optimal decisions may then be so unreliable as to be useless. Large economic models that require complex estimation methods, but have useful properties for policy makers, often produce large forecast standard deviations that might come from the complexity of the model, data, estimation method, or all of the above without necessarily knowing the source of uncertainty. BPS, on the other hand, synthesizes the forecasts and by doing so, has the ability to *decrease* forecast uncertainties relative to the agents, without overly underestimating real risks; this is evident in the example here, where BPS leads the agents (and other strategies) in terms of LPDR performance. Fig. 2 shows that some part of this comes from generally reduced forecast uncertainties—coupled with more accurate point forecasts—at this 1-step horizon. We caution that reduced uncertainties are not always expected or achieved, as exemplified below.

Given the pictures for LPDR scores and 1-step forecast uncertainties above, it is no surprise that BPS almost uniformly dominates in terms of raw point forecast accuracy as well. Apart from an initially unstable period at beginning of the time frame, MSFE measures bear this out (see supplementary Appendix B). Point forecast accuracy is particularly improved at crisis periods, with MSFE staying relatively level under BPS while significantly increasing for other models and methods. In summary, BPS is able to adapt to maintain improved forecasting performance both in terms of point forecasts and risk characterization, a key positive feature for decision makers who are tasked with forecasting risk and quantiles, especially under critical situations such as economic crises.

We also note that, over the prior period 1977/Q2-1989/Q4, BMA—characteristically—effectively degenerated, with posterior probabilities increasingly favoring agent M3; thus, at the start of the test period, BMA-based forecast densities are very close to those from M3 alone. BPS, on the other hand, allows for continual adaptation as agent models change in their relative forecasting abilities; over the test period, agent M2 tends to play a dominant role in BPS, notable in terms of the on-line estimates of BPS agent coefficients in  $\theta_{tj}$ ; see Fig. 3. An interesting point to note is how BPS successfully adapts its coefficients during the sub-prime mortgage crisis by significantly down-weighting M3. As a dynamic model, BPS will not degenerate, continually allowing for “surprises” in changes in relative forecast performance across the agents.

For comparison, the DeCo analysis generates weights that heavily favor M3 from the training period; see Fig. 4. While dynamic in theory, the inferred weights end up being quite stable over time. Contrasting this to BPS, it is clear that the benefits of BPS stem largely from dynamic adaptability. As the posterior mean trajectories of Fig. 3 indicate, BPS is constantly updating and

calibrating without over-learning, partly due to discount learning (forgetting) and the latent biases/dependencies being effectively transferred to the coefficients. Through understanding what aspects of BPS (intercept, discounting, latent biases/dependencies, etc...) contribute to the gains seen in this example is of interest, it is clear that they allow the coefficients to quickly adapt over time and improve forecasts.

**4-Step Ahead Forecasting.** Several figures summarize sequential analysis for 4-step forecasting, using both the direct extrapolation to 4-quarters ahead under the BPS model and the customized BPS(4) model. Each BPS strategy performs consistently better than agents and other strategies in point forecasting, while BPS(4) makes significant improvements in terms of both point and distribution forecasts compared to direct BPS extrapolation. This is clearly seen in the time trajectories of LPDR under BPS(4) in Fig. 5, and in point forecast accuracy measured by mean square forecast errors (see supplementary Appendix B). Direct BPS extrapolation performs relatively poorly in terms of both point forecast accuracy and LPDR (see figures in supplementary Appendix B) as it is inherently calibrated to 1-step model fit. In particular, it fails to adequately represent the increased uncertainty associated with longer term forecasts. Looking at the forecast standard deviations in Fig. 6, it is clear that BPS(4) is able to improve by adjusting to the increased forecast uncertainties. Then, even though forecast uncertainties increase substantially, they are clearly more than balanced by improved location forecasts. This again bears out the recommendation to directly synthesize forecasts on the horizon of interest.

On-line trajectories of estimates of the BPS(4) coefficients  $\theta_t$  as they are sequentially updated and adapt (see figure in supplementary Appendix B) show notable reduction in adaptability over time relative to the 1-step BPS coefficients of Fig. 3. This is to be expected as the models' forecasts are less reliable at longer horizons, so the data-based information advising the changes in posteriors over time is limited. The dynamic intercept term serves as a comparison base as it moves away from zero, playing a more active role in BPS(4) than in the 1-step case. Additionally, the 4-step ahead coefficient values (indicated here by just the on-line means, of course) are quite different from 1-step coefficients, reasonably reflecting the differing forecasting abilities of the agents at differing horizons. BPS(4) is able to adapt to the 4-step ahead forecast, differently from the 1-step BPS, and dominates in performance compared to all other methods as a result.

### 4.3 Retrospective Analysis

Based on the full MCMC analysis of all data in 1990/Q1-2014/Q4, we review aspects of retrospective posterior inference.

**BPS Coefficients.** Figs. 7 shows trajectories with uncertainties under BPS (1-step); these can be compared with on-line point summaries in Fig. 3 earlier discussed. We see the expected smoothing

of estimated trajectories of coefficients. Similar figures are given for BPS(4) model coefficients in supplementary Appendix B. To the extent that the role of the intercept terms can be regarded as reflecting (lack of) effectiveness of the synthesized models, these figures confirm that the agents' predictions are much more questionable at 4-steps ahead than at 1-step ahead. Intercepts increase up to and during the sub-prime mortgage crisis due to the increased inability of the models to forecast well during this time.

**Latent Agent States and Forecast Dependencies.** BPS naturally allows for– and adapts to– dependencies among agents as they evolve over time. In many cases, models and data used by agents are typically unknown to the decision maker  $\mathcal{D}$  and therefore posterior inference on dependencies among agents is of special interest; even when agents are chosen statistical models– as in this example– the questions of inter-dependence and potential redundancy in forecast value are hard and open questions in all approaches to aggregation.

As noted early, the conceptual and theoretical basis of BPS allows direct investigation of agent dependencies, as the inherent latent agent states  $x_{tj}$ – when inferred based on the observed data– carry the relevant information. From the full MCMC analysis to the end of the test data period, we have full posterior samples for the states  $x_t$ – in both the direct BPS and customized BPS(4). For illustration, we focus on the 1-step BPS analysis. Fig. 8 displays posterior trajectories for the errors in latent states, namely  $y_t - x_{tj}$  over time  $t$ , in terms of posterior means and intervals (similar trajectories of the  $x_{tj}$  themselves, together with the inflation outcomes  $y_t$ , are in supplementary Appendix B). The patterns over time in each of these reflect the strong, positive dependencies among agents that are to be expected given the nature of the agent models.

To explore dependencies, we simply investigate the posterior for  $x_{1:T}$ . This is not of standard form and is represented in terms of the MCMC-based posterior sample. One simple set of summaries is based on just computing empirical  $R^2$  measures: from the MCMC sample, compute the approximate posterior variance matrix of  $x_t$  at each  $t$ , and from that extract implied sets of conditionals variances of any  $x_{tj}$  given any subset of the other  $x_{ti}, i \neq j$ . We do this for  $i = 1:J \setminus j$ , defining the MC-empirical  $R^2$  for agent  $j$  based on all other agents, i.e., measuring the redundancy of agent  $j$  in the context of all  $J$  agents– the *complete conditional dependencies*. We do this also using each single agent  $i \neq j$ , defining paired MC-empirical  $R^2$  measures of how dependent agents  $i, j$  are– the *bivariate dependencies*. Fig. 9 displays trajectories over time for these two measures from the BPS 1-step analysis; a corresponding figure from BPS(4) is in supplementary Appendix B.

Overall, we see high complete conditional dependencies at both forecast horizons, as expected due to the nature of the 4 models and their evaluation on the same data. Dependencies are substantial and much higher for 1-step forecasts than for 4-step ahead forecasts, reflecting decreasing concordance with increasing horizon, and all decrease over the test period. The predictability of M2 based on the others drops at a greater rate after about the start of 2002, in part due to poorer and less reliable performance during the dot-com crisis. The paired measures are all very low

compared to the complete conditionals, and again naturally lower overall in 4-step forecasting. Concordance of M2 and M3 decreases for 1-step but increases slightly for 4-step ahead forecasts, reflecting dynamics in relationships that differ with forecast horizon; from earlier discussion of forecast accuracy, this can be explained by how, in 1-step ahead forecasts, M2 improves while M3 deteriorates during the sub-prime mortgage crisis. In contrast, for 4-step ahead forecasts we see forecast errors converging between the two, explaining the increase in concordance as all models performed equally poorly.

#### 4.4 Simulation Studies

We have explored a range of synthetic data sets to fully evaluate the above analysis using the same four models but with known parameters, and with simulated data generated with random switching between models. The results echo and amplify those of the macroeconomic study. In particular, at 1-step ahead, BPS outperforms the best model and best traditional strategy by nearly 20% in terms of point forecasts, as well as significantly improving in terms of LPDR measures of density forecasts. At 4-steps ahead, BPS(4) very substantially improves on all models and on BPS, the latter being partly due to improved characterization of forecast uncertainties under BPS(4), coupled with somewhat improved point forecasts.

Detailed graphical summaries of analysis of one such synthetic data set appear in the supplementary Appendix C. These graphs include on-line filtering and forecast summaries as well as aspects of the retrospective posterior analysis from the dynamic BPS model, to compare with those from the real data analysis.

### 5 Summary Comments

Drawing on theory of Bayesian agent opinion analysis, BPS provides a theoretically and conceptually sound framework to compare and synthesize density forecasts that has been developed here for dynamic contexts of sequential time series forecasting. With this new framework and extension, decision makers are able to dynamically calibrate, learn, and update weights for ranges of forecasts from dynamic models, with multiple lags and predictors as exemplified here, as well as from more subjective sources such as individual forecasters or agencies.

The U.S. macroeconomic data study illustrates how effective and practical BPS is under settings that are increasingly important and topical in macroeconomics and econometrics. By dynamically synthesizing the forecasts, BPS improves forecast performance and dominates other standard strategies, such as BMA and pooling, over short and long horizons and for both point and distribution forecasts. Further analysis shows evidence that BPS is also robust in its forecast abilities under economic distress, which is critically important for practical applications. Additionally, posterior inference of the full time series provides the decision maker with information on how agents are

related, and how that relationship dynamically evolves through time; this has potential to inform BPS modeling for continued forecast synthesis into the future.

In addition to applications to U.S. macroeconomic data, BPS has the potential to be applied to other fields where multiple forecasts, whether from forecasters or models, are available. This includes areas such as finance (e.g., stocks, indexes, and bonds), business (e.g., product demand and earnings), meteorology, and risk (e.g., seismic and environmental risk). Methodological extensions are needed for multivariate synthesis, non-normal forecasts and discrete data, and missing or incomplete/partial forecasts. It will certainly be of interest and practical importance to develop studies in which agents are represented by sets of more elaborate macroeconomic models, such as VAR, dynamic threshold models, dynamic stochastic general equilibrium (DSGE) models, and to integrate forecasts coming from professional forecasters and economists. In the latter contexts, developing BPS models that integrate partial forecast information—such as quantiles of agent forecast distributions rather than full density forecasts—is of practical importance. We believe that further such studies will define increasing empirical support for the utility of the approach and attract applied researchers. Coupled with this, some of the foundational theory of Bayesian agent opinion analysis in the single agent case ([West 1992](#)) explicitly addressed partial forecast information, but there has been no extension, to date, to multiple agents and the dynamic/time series setting.

Computational questions are also relevant; as developed and exemplified, analysis in the sequential time series context relies on repeat reanalysis using MCMC, with a new simulation analysis required as each new time period arises. This is in common with the application of Bayesian dynamic latent factor models of other forms in the sequential forecasting context, including, in particular, dynamic latent threshold models (e.g. [Nakajima and West 2013a,b, 2015](#); [Zhou et al. 2014](#)) whose use in defining sets of candidate agents for BPS is of some applied interest. One view is that a substantial computational burden is nowadays a minor issue and, in fact, a small price to pay for the potential improvements in forecasting accuracy and insights that our example illustrates. That said, some methods of sequential model analysis based on sequential Monte Carlo (SMC, e.g. [Lopes and Tsay 2011](#)) may provide for more efficient computations, at least in terms of CPU cycles, in some stylized versions of the overall BPS model framework.

## Acknowledgements

We are grateful for useful discussions with colleagues that helped with perspective in developing this work, including Matt Johnson (Duke University) and Jouchi Nakajima (Bank for International Settlements), and the organizers and participants of the 5th European Seminar on Bayesian Econometrics (ESOBE) held at the Study Center Gerzensee, Switzerland, October 29-30th 2015.

We also thank the editors and four anonymous referees for very detailed, positive and constructive comments on the original version of this paper.

## References

- Aastveit, K. A., Gerdrup, K. R., Jore, A. S., Thorsrud, L. A., 2014. Nowcasting GDP in real time: A density combination approach. *Journal of Business & Economic Statistics* 32, 48–68.
- Aastveit, K. A., Ravazzolo, F., van Dijk, H. K., 2015. Combined density Nowcasting in an uncertain economic environment. *Journal of Business & Economic Statistics* -, -, to appear.
- Amisano, G. G., Giacomini, R., 2007. Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics* 25, 177–190.
- Belmonte, M. A., Koop, G., Korobilis, D., 2014. Hierarchical shrinkage in time-varying parameter models. *Journal of Forecasting* 33(1), 480–94.
- Billio, M., Casarin, R., Ravazzolo, F., van Dijk, H. K., 2012. Combination schemes for turning point predictions. *Quarterly Review of Finance and Economics* 52, 402–412.
- Billio, M., Casarin, R., Ravazzolo, F., van Dijk, H. K., 2013. Time-varying combinations of predictive densities using nonlinear filtering. *Journal of Econometrics* 177, 213–232.
- Casarin, R., Grassi, S., Ravazzolo, F., van Dijk, H. K., 2015. Parallel sequential monte carlo for efficient density combination: The deco matlab toolbox. *Journal of Statistical Software* 68.
- Cogley, T., Sargent, T. J., 2005. Drifts and volatilities: Monetary policies and outcomes in the post WWII U.S. *Review of Economic Dynamics* 8, 262–302.
- Dangl, T., Halling, M., 2012. Predictive regressions with time-varying coefficients. *Journal of Financial Economics* 106, 157–181.
- Dawid, A. P., DeGroot, M. H., Mortera, J., Cooke, R., French, S., Genest, C., Schervish, M. J., Lindley, D. V., McConway, K. J., Winkler, R. L., 1995. Coherent combination of experts' opinions. *Test* 4, 263–313.
- French, S., 2011. Aggregating expert judgement. *Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales (Serie A: Matematicas)* 105, 181–206.
- Frühwirth-Schnatter, S., 1994. Data augmentation and dynamic linear models. *Journal of Time Series Analysis* 15, 183–202.
- Genest, C., Schervish, M. J., 1985. Modelling expert judgements for Bayesian updating. *Annals of Statistics* 13, 1198–1212.
- Genre, V., Kenny, G., Meyler, A., Timmermann, A., 2013. Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting* 29, 108–121.

- Geweke, J., Amisano, G. G., 2012. Prediction with misspecified models. *The American Economic Review* 102, 482–486.
- Geweke, J. F., Amisano, G. G., 2011. Optimal prediction pools. *Journal of Econometrics* 164, 130–141.
- Gruber, L. F., West, M., 2016. GPU-accelerated Bayesian learning in simultaneous graphical dynamic linear models. *Bayesian Analysis* 11, 125–149.
- Gruber, L. F., West, M., 2017. Bayesian forecasting and scalable multivariate volatility analysis using simultaneous graphical dynamic linear models. *Econometrics and Statistics* (published online March 12), arXiv:1606.08291.
- Hall, S. G., Mitchell, J., 2007. Combining density forecasts. *International Journal of Forecasting* 23, 1–13.
- Harrison, P. J., Stevens, C. F., 1976. Bayesian forecasting. *Journal of the Royal Statistical Society (Series B: Methodological)* 38, 205–247.
- Hoogerheide, L., Kleijn, R., Ravazzolo, F., Van Dijk, H. K., Verbeek, M., 2010. Forecast accuracy and economic gains from Bayesian model averaging using time-varying weights. *Journal of Forecasting* 29, 251–269.
- Johnson, M. C., West, M., 2016. Bayesian predictive synthesis for probabilistic forecast calibration and combination. Tech. rep., Department of Statistical Science, Duke University.
- Kapetanios, G., Mitchell, J., Price, S., Fawcett, N., 2015. Generalised density forecast combinations. *Journal of Econometrics* 188, 150–165.
- Kascha, C., Ravazzolo, F., 2010. Combining inflation density forecasts. *Journal of Forecasting* 29, 231–250.
- Koop, G., Korobilis, D., 2013. Large time-varying parameter VARs. *Journal of Econometrics* 177 (2), 185–198.
- Koop, G., Leon-Gonzalez, R., Strachan, R. W., 2009. On the evolution of the monetary policy transmission mechanism. *Journal of Economic Dynamics and Control* 33, 997–1017.
- Lindley, D. V., Tversky, A., Brown, R. V., 1979. On the reconciliation of probability assessments. *Journal of the Royal Statistical Society (Series A: General)* 142, 146–180.
- Lopes, H. F., Tsay, R. S., 2011. Particle filters and Bayesian inference in financial econometrics. *Journal of Forecasting* 30, 168–209.

- Nakajima, J., West, M., 2013a. Bayesian analysis of latent threshold dynamic models. *Journal of Business & Economic Statistics* 31, 151–164.
- Nakajima, J., West, M., 2013b. Bayesian dynamic factor models: Latent threshold approach. *Journal of Financial Econometrics* 11, 116–153.
- Nakajima, J., West, M., 2015. Dynamic network signal processing using latent threshold models. *Digital Signal Processing* 47, 6–15.
- Negro, M. D., Hasegawa, R. B., Schorfheide, F., 2016. Dynamic prediction pools: an investigation of financial frictions and forecasting performance. *Journal of Econometrics* 192(2), 391–405.
- Pettenuzzo, D., Ravazzolo, F., 2016. Optimal portfolio choice under decision-based model combinations. *Journal of Applied Econometrics* 31 (7), 1312–1332.
- Prado, R., West, M., 2010. *Time Series: Modelling, Computation & Inference*. Chapman & Hall/CRC Press.
- Primiceri, G. E., 2005. Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies* 72, 821–852.
- Stella, A., Stock, J. H., 2012. A state-dependent model for inflation forecasting. FRB International Finance Discussion Paper (1062).
- Stock, J. H., Watson, M. W., 1999. Forecasting inflation. *Journal of Monetary Economics* 44(2), 293–335.
- Stock, J. H., Watson, M. W., 2007. Why has us inflation become harder to forecast? *Journal of Money, Credit and banking* 39(s1), 3–33.
- Terui, N., van Dijk, H. K., 2002. Combined forecasts from linear and non- linear time series models. *International Journal of Forecasting* 18(3), 421–438.
- West, M., 1984. Bayesian aggregation. *Journal of the Royal Statistical Society (Series A: General)* 147, 600–607.
- West, M., 1988. Modelling expert opinion (with discussion). In: Bernardo, J. M., DeGroot, M. H., Lindley, D. V., Smith, A. F. M. (Eds.), *Bayesian Statistics 3*. Oxford University Press, pp. 493–508.
- West, M., 1992. Modelling agent forecast distributions. *Journal of the Royal Statistical Society (Series B: Methodological)* 54, 553–567.
- West, M., Crosse, J., 1992. Modelling of probabilistic agent opinion. *Journal of the Royal Statistical Society (Series B: Methodological)* 54, 285–299.

West, M., Harrison, P. J., 1997. Bayesian Forecasting & Dynamic Models, 2nd Edition. Springer Verlag.

Zhao, Z. Y., Xie, M., West, M., 2016. Dynamic dependence networks: Financial time series forecasting & portfolio decisions (with discussion). *Applied Stochastic Models in Business and Industry* 32, 311–339, arXiv:1606.08339.

Zhou, X., Nakajima, J., West, M., 2014. Bayesian forecasting and portfolio decisions using dynamic dependent sparse factor models. *International Journal of Forecasting* 30, 963–980.

# Dynamic Bayesian Predictive Synthesis in Time Series Forecasting

Kenichiro McAlinn & Mike West

## Tables and Figures

	MSFE <sub>1:T</sub>				LPDR <sub>1:T</sub>		
	1-step	%	4-step	%	1-step	4-step	BPS(4)
M1	0.0634	-23.83	0.4227	-14.68	-13.84	71.43	-94.56
M2	0.0598	-16.80	0.4156	-12.75	-8.55	68.16	-97.82
M3	0.0616	-20.31	0.4208	-14.16	-9.06	60.08	-105.90
M4	0.0811	-58.40	0.4880	-32.39	-22.71	67.46	-98.53
BMA	0.0617	-20.51	0.4882	-32.45	-9.00	65.65	-100.33
LinP	0.0575	-12.30	0.4275	-15.98	-8.84	85.50	-80.48
LogP	0.0579	-13.09	0.4275	-15.98	-7.86	68.23	-97.75
DeCo	0.0571	-11.52	0.4156	-12.75	-	-	-
BPS	0.0512	-	0.4001	-8.55	-	-	-
BPS(4)	-	-	0.3686	-	-	-	-

Table 1: US inflation rate forecasting 1990/Q1-2014/Q4: Forecast evaluations for quarterly U.S. inflation over the 25 years 1990/Q1-2014/Q4, comparing mean squared forecast errors and log predictive density ratios for this  $T = 100$  quarters. The column % denotes improvements over BPS and BPS( $k$ ) for 1- and 4-step ahead forecasts, respectively. Note: LPDR<sub>1:T</sub> is respect to BPS and BPS(4) and nonexistent for DeCo due to lack of analytic predictive distributions.

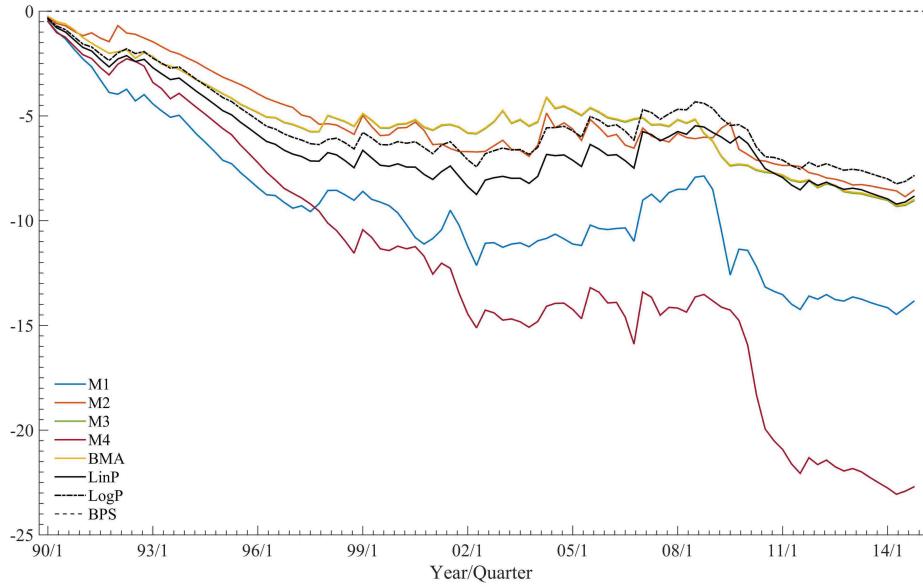


Figure 1: US inflation rate forecasting 1990/Q1-2014/Q4: 1-step ahead log predictive density ratios  $\text{LPDR}_{1:t}(1)$  sequentially revised at each of the  $t = 1:100$  quarters. The baseline at 0 over all  $t$  corresponds to the standard BPS model.

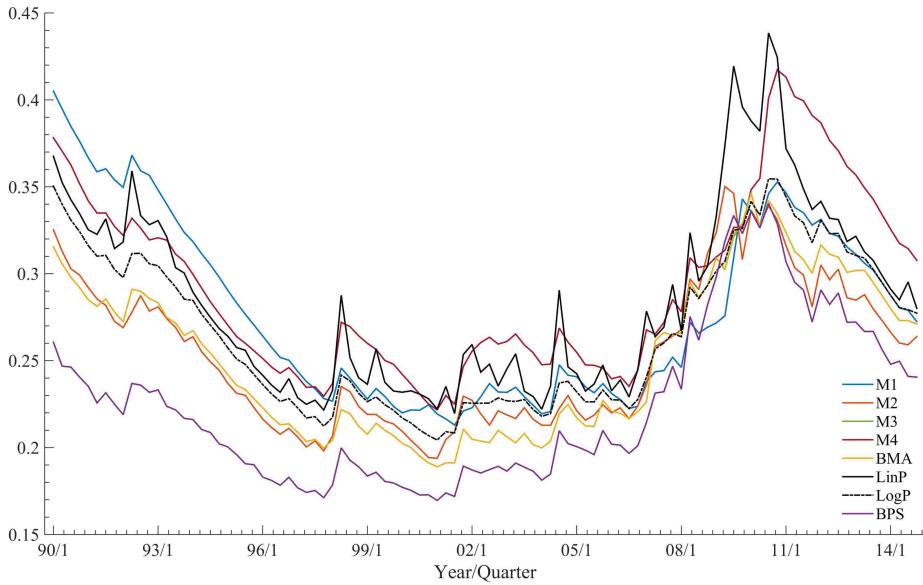


Figure 2: US inflation rate forecasting 1990/Q1-2014/Q4: 1-step ahead forecast standard deviations sequentially computed at each of the  $t = 1:100$  quarters.

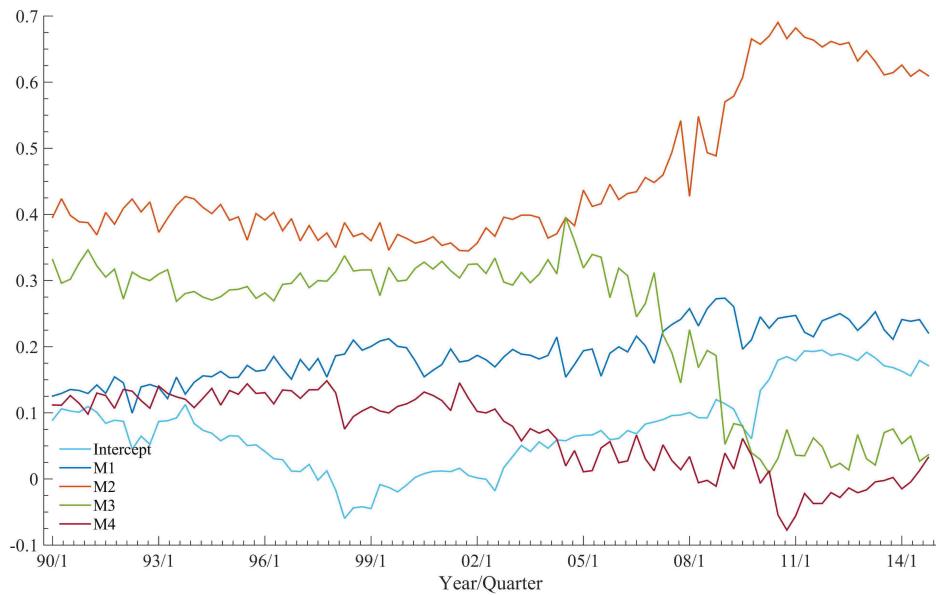


Figure 3: US inflation rate forecasting 1990/Q1-2014/Q4: On-line posterior means of BPS model coefficients sequentially computed at each of the  $t = 1:100$  quarters.

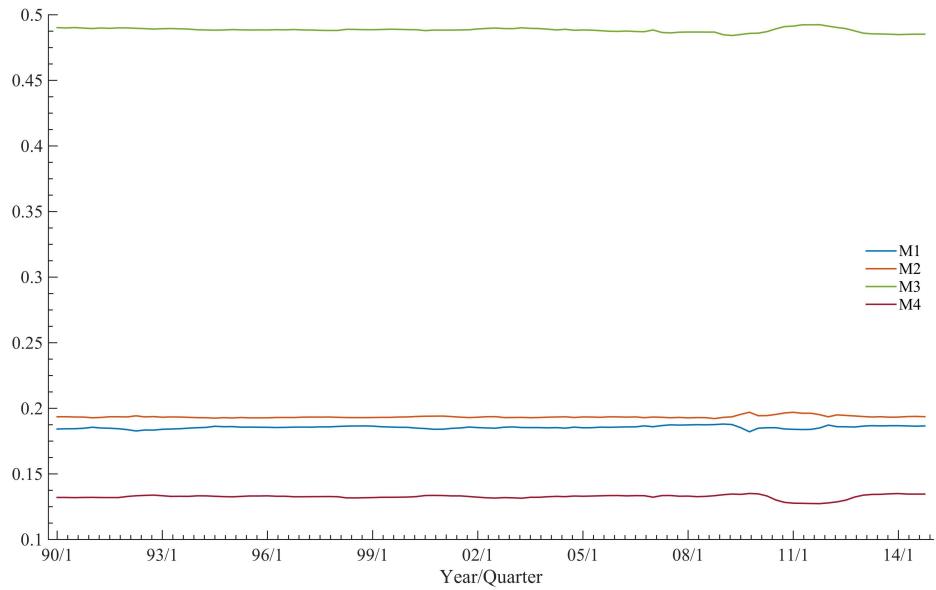


Figure 4: US inflation rate forecasting 1990/Q1-2014/Q4: On-line means of DeCo model weights sequentially computed at each of the  $t = 1:100$  quarters.

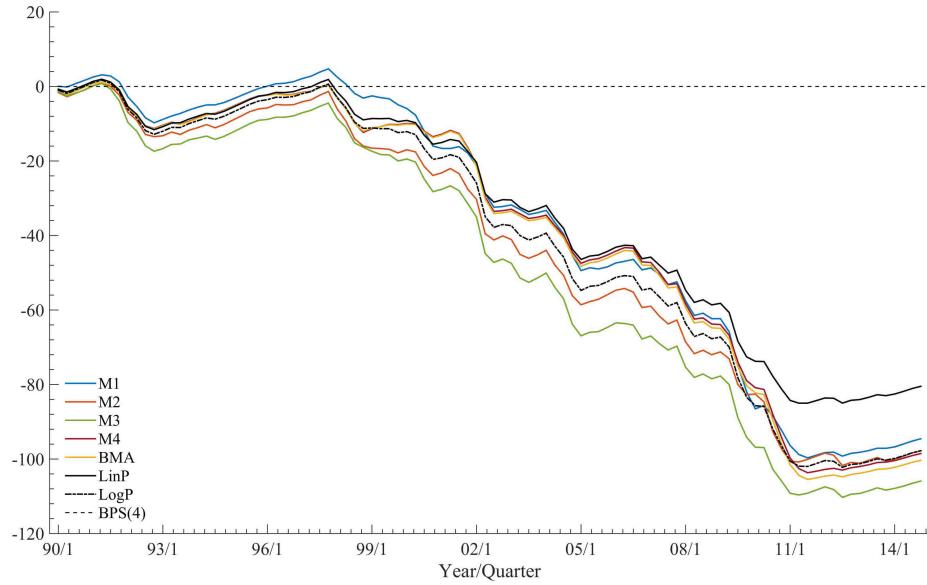


Figure 5: US inflation rate forecasting 1990/Q1-2014/Q4: 4-step ahead log predictive density ratios,  $LPDR_{1:t}(4)$  sequentially revised at each of the  $t = 1:100$  quarters using the 4-step ahead customized BPS(4) model (baseline at 0 over time).

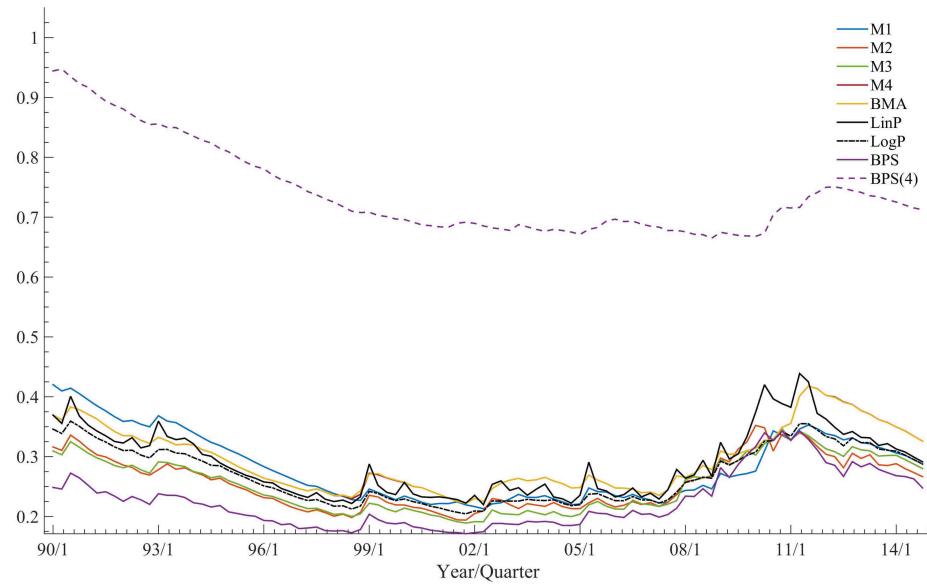


Figure 6: US inflation rate forecasting 1990/Q1-2014/Q4: 4-step ahead forecast standard deviations sequentially computed at each of the  $t = 1:100$  quarters.

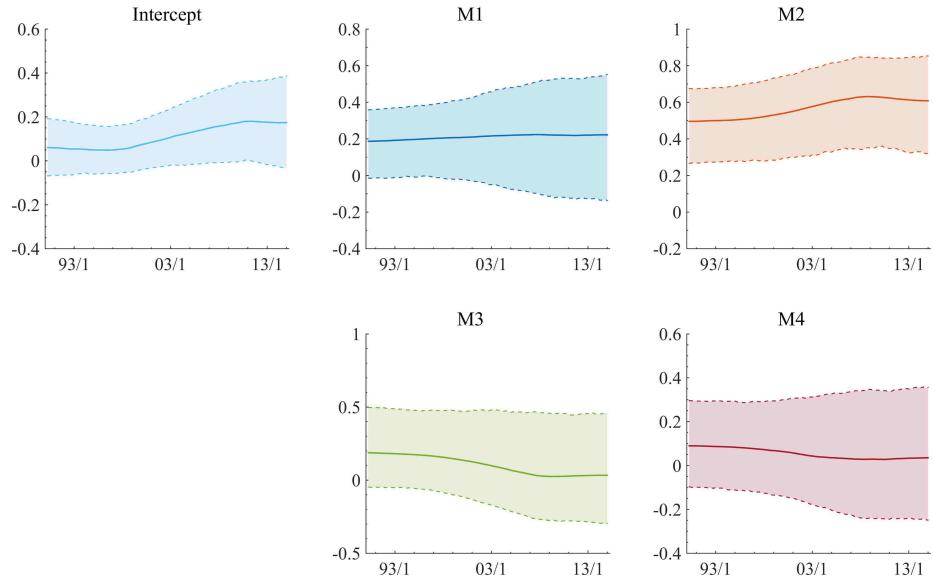


Figure 7: US inflation rate forecasting 1990/Q1-2014/Q4: Retrospective posterior trajectories of the BPS model coefficients based on data from the full  $t = 1:100$  quarters. Posterior means (solid) and 95% credible intervals (shaded).

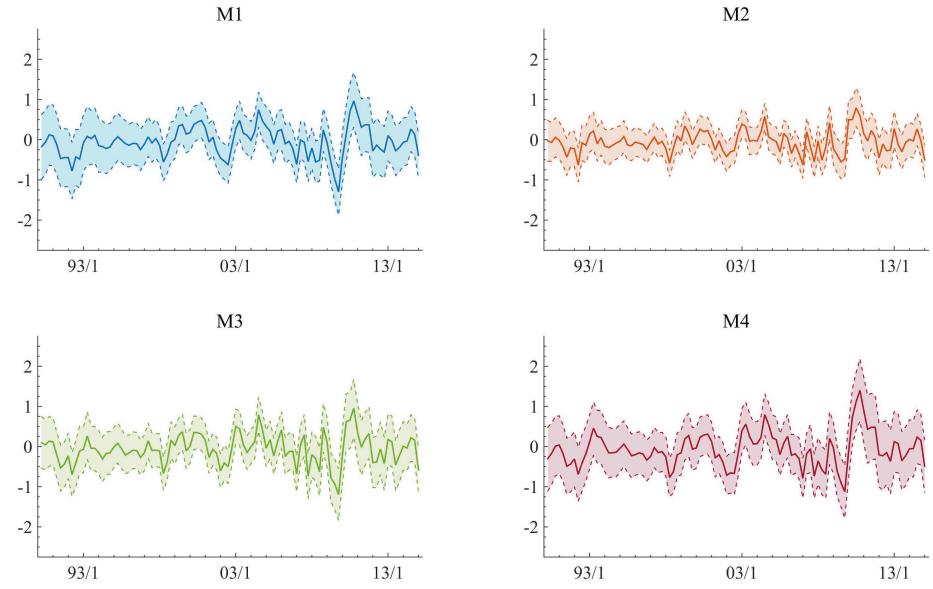


Figure 8: US inflation rate forecasting 1990/Q1-2014/Q4: BPS model-based posterior trajectories of the error in latent agents states  $y_t - x_{tj}$  for  $j = 1:4$  over the  $t = 1:100$  quarters. Posterior means (solid) and 95% credible intervals (shaded) from the MCMC analysis.

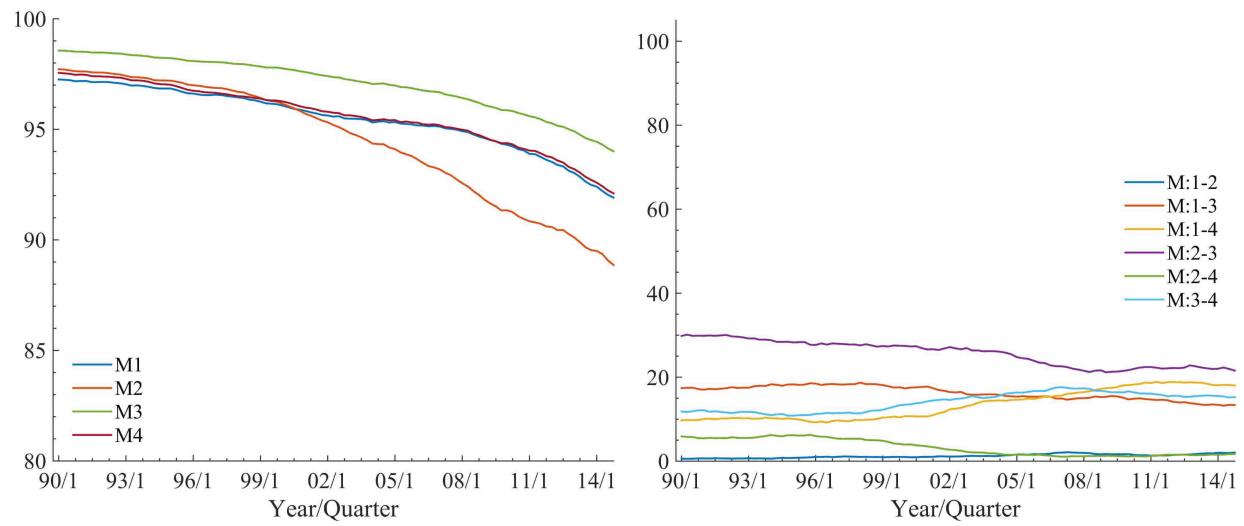


Figure 9: US inflation rate forecasting 1990/Q1-2014/Q4: BPS model-based trajectories of 1-step ahead MC-empirical  $R^2$  (left) and paired MC-empirical  $R^2$  (right) in the posterior for the latent agent states  $x_{jt}$  for  $j = 1:4$  over the  $t = 1:100$  quarters.

# Dynamic Bayesian Predictive Synthesis in Time Series Forecasting

Kenichiro McAlinn & Mike West

## Supplementary Material

## A Appendix: Summary of MCMC for Dynamic BPS

### A.1 Overview and Initialization

This appendix summarizes algorithmic details of implementation of the MCMC computations for dynamic BPS model fitting of Section 3.3. This involves a standard set of steps in a customized two-component block Gibbs sampler: one component samples the latent agent states, and the second samples the dynamic BPS model states/parameters. The latter involves the FFBS algorithm central to MCMC in all conditionally normal DLMs ( Frühwirth-Schnatter 1994; West and Harrison 1997, Sect 15.2; Prado and West 2010, Sect 4.5).

In our sequential learning and forecasting context, the full MCMC analysis is performed anew at each time point as time evolves and new data are observed. We detail MCMC steps for a specific time  $t$  here, based on all data up until that time point.

Standing at time  $T$ , the decision maker has historical information  $\{y_{1:T}, \mathcal{H}_{1:T}\}$ , initial prior  $\theta_0 \sim N(\mathbf{m}_0, \mathbf{C}_0 v_0 / s_0)$  and  $1/v_0 \sim G(n_0/2, n_0 s_0/2)$ , and discount factors  $(\beta, \delta)$ . The MCMC is run iteratively as follows.

**Initialization:** First, initialize by setting  $\mathbf{F}_t = (1, x_{t1}, \dots, x_{tJ})'$  for each  $t = 1:T$  at some chosen initial values of the latent agent states. Initial values can be chosen arbitrarily. One obvious and appropriate choice—our recommended default choice—is to simply generate agent states from their priors, i.e., from the agent forecast distributions,  $x_{tj} \sim h_{tj}(x_{tj})$  independently for all  $t = 1:T$  and  $j = 1:J$ . This is easily implemented in cases when the agent forecasts are T or normal distributions, or can be otherwise directly sampled; we use this in our analyses reported in the paper, and recommend it as standard. An obvious alternative initialization is to simply set  $x_{tj} = y_t$  for each  $j, t$ , though we prefer to initialize with some inherent dispersion in starting values. Ultimately, since the MCMC is rapidly convergent, choice of initial values is not critical.

### A.2 Two Sampling Steps in Each MCMC Iterate

Following initialization, the MCMC iterates repeatedly to resample two coupled sets of conditional posteriors to generate the MCMC samples from the target posterior  $p(\mathbf{x}_{1:T}, \Phi_{1:T}|y_{1:T}, \mathcal{H}_{1:T})$ . These two conditional posteriors and algorithmic details of their simulation are as follows.

### A.2.1 Per MCMC Iterate Step 1: Sampling BPS DLM parameters $\Phi_{1:T}$

Conditional on any values of the latent agent states, we are in the setting of a conditionally normal DLM with the agent states as known predictors based on their specific values. The BPS conjugate DLM form,

$$\begin{aligned} y_t &= \mathbf{F}'_t \boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim N(0, v_t), \\ \boldsymbol{\theta}_t &= \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N(0, v_t \mathbf{W}_t), \end{aligned}$$

has known elements  $\mathbf{F}_t, \mathbf{W}_t$  and specified initial prior at  $t = 0$ . The implied conditional posterior for  $\Phi_{1:T}$  then does not depend on  $\mathcal{H}_{1:T}$ , reducing to  $p(\Phi_{1:T} | \mathbf{x}_{1:T}, y_{1:T})$ . This is simulated using the efficient and standard FFBS algorithm, modified to incorporate the discount stochastic volatility components for  $v_t$  (e.g. [Frühwirth-Schnatter 1994](#); [West and Harrison 1997](#), Sect 15.2; [Prado and West 2010](#), Sect 4.5). In detail, this proceeds as follows.

**Forward filtering:** For each  $t = 1:T$  in sequence, perform the standard one-step filtering updates to compute and save the sequence of sufficient statistics for the on-line posteriors  $p(\boldsymbol{\theta}_t, v_t | y_{1:t}, \mathbf{x}_{1:t})$  at each  $t$ . The summary technical details are as follows:

1. *Time  $t - 1$  posterior:*

$$\begin{aligned} \boldsymbol{\theta}_{t-1} | v_{t-1}, \mathbf{x}_{1:t-1}, y_{1:t-1} &\sim N(\mathbf{m}_{t-1}, \mathbf{C}_{t-1} v_{t-1} / s_{t-1}), \\ v_{t-1}^{-1} | \mathbf{x}_{1:t-1}, y_{1:t-1} &\sim G(n_{t-1}/2, n_{t-1} s_{t-1} / 2), \end{aligned}$$

with point estimates  $\mathbf{m}_{t-1}$  of  $\boldsymbol{\theta}_{t-1}$  and  $s_{t-1}$  of  $v_{t-1}$ .

2. *Update to time  $t$  prior:*

$$\begin{aligned} \boldsymbol{\theta}_t | v_t, \mathbf{x}_{1:t-1}, y_{1:t-1} &\sim N(\mathbf{m}_{t-1}, \mathbf{R}_t v_t / s_{t-1}) \quad \text{with} \quad \mathbf{R}_t = \mathbf{C}_{t-1} / \delta, \\ v_t^{-1} | \mathbf{x}_{1:t-1}, y_{1:t-1} &\sim G(\beta n_{t-1} / 2, \beta n_{t-1} s_{t-1} / 2), \end{aligned}$$

with (unchanged) point estimates  $\mathbf{m}_{t-1}$  of  $\boldsymbol{\theta}_t$  and  $s_{t-1}$  of  $v_t$ , but with increased uncertainty relative to the time  $t - 1$  posteriors, the level of increased uncertainty being defined by the discount factors.

3. *1-step predictive distribution:*  $y_t | \mathbf{x}_{1:t}, y_{1:t-1} \sim T_{\beta n_{t-1}}(f_t, q_t)$  where

$$f_t = \mathbf{F}'_t \mathbf{m}_{t-1} \quad \text{and} \quad q_t = \mathbf{F}'_t \mathbf{R}_t \mathbf{F}_t + s_{t-1}.$$

4. Filtering update to time  $t$  posterior:

$$\begin{aligned}\boldsymbol{\theta}_t | v_t, \mathbf{x}_{1:t}, y_{1:t} &\sim N(\mathbf{m}_t, \mathbf{C}_t v_t / s_t), \\ v_t^{-1} | \mathbf{x}_{1:t}, y_{1:t} &\sim G(n_t/2, n_t s_t/2),\end{aligned}$$

with defining parameters as follows:

- i. For  $\boldsymbol{\theta}_t | v_t : \mathbf{m}_t = \mathbf{m}_{t-1} + \mathbf{A}_t e_t$  and  $\mathbf{C}_t = r_t (\mathbf{R}_t - q_t \mathbf{A}_t \mathbf{A}'_t)$ ,
- ii. For  $v_t : n_t = \beta n_{t-1} + 1$  and  $s_t = r_t s_{t-1}$ ,

based on 1-step forecast error  $e_t = y_t - f_t$ , the state adaptive coefficient vector (a.k.a. “Kalman gain”)  $\mathbf{A}_t = \mathbf{R}_t \mathbf{F}_t / q_t$ , and volatility estimate ratio  $r_t = (\beta n_{t-1} + e_t^2 / q_t) / n_t$ .

**Backward sampling:** Having run the forward filtering analysis up to time  $T$ , the backward sampling proceeds as follows.

- a. At time  $T$ : Simulate  $\Phi_T = (\boldsymbol{\theta}_T, v_T)$  from the final normal/inverse gamma posterior  $p(\Phi_T | \mathbf{x}_{1:T}, y_{1:T})$  as follows. First, draw  $v_T^{-1}$  from  $G(n_T/2, n_T s_T/2)$ , and then draw  $\boldsymbol{\theta}_T$  from  $N(\mathbf{m}_T, \mathbf{C}_T v_T / s_T)$ .
- b. Recurse back over times  $t = T-1, T-2, \dots, 0$ : At time  $t$ , sample  $\Phi_t = (\boldsymbol{\theta}_t, v_t)$  as follows:
  - i. Simulate the volatility  $v_t$  via  $v_t^{-1} = \beta v_{t+1}^{-1} + \gamma_t$  where  $\gamma_t$  is an independent draw from  $\gamma_t \sim G((1-\beta)n_t/2, n_t s_t/2)$ ,
  - ii. Simulate the state  $\boldsymbol{\theta}_t$  from the conditional normal posterior  $p(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t+1}, v_t, \mathbf{x}_{1:T}, y_{1:T})$  with mean vector  $\mathbf{m}_t + \delta(\boldsymbol{\theta}_{t+1} - \mathbf{m}_t)$  and variance matrix  $\mathbf{C}_t(1-\delta)(v_t/s_t)$ .

### A.2.2 Per MCMC Iterate Step 2: Sampling the latent agent states $\mathbf{x}_{1:T}$

Conditional on most recently sampled values of the BPS DLM parameters  $\Phi_{1:T}$ , the MCMC iterate completes with resampling of the latent agent states from their full conditional posterior  $p(\mathbf{x}_{1:t} | \Phi_{1:t}, y_{1:t}, \mathcal{H}_{1:t})$ . It is immediate that the  $\mathbf{x}_t$  are conditionally independent over time  $t$  in this conditional distribution, with time  $t$  conditionals

$$p(\mathbf{x}_t | \Phi_t, y_t, \mathcal{H}_t) \propto N(y_t | \mathbf{F}'_t \boldsymbol{\theta}_t, v_t) \prod_{j=1:J} h_{tj}(x_{tj}) \quad \text{where } \mathbf{F}_t = (1, x_{t1}, x_{t2}, \dots, x_{tJ})'. \quad (8)$$

Several comments are relevant to studies with different forms of the agent forecast densities.

1. *Normal agent forecast densities:* In cases when each of the agent forecast densities is normal, the posterior in eqn. (8) yields a multivariate normal distribution for  $\mathbf{x}_t$ . Computation of its defining parameters and then drawing a new sample vector  $\mathbf{x}_t$  are trivial. Specifically, suppose that  $h_{tj}(x_{tj})$  is density of the normal  $N(h_{tj}, H_{tj})$ , and write  $\mathbf{h}_t = (h_{t1}, h_{t2}, \dots, h_{tJ})'$

and  $\mathbf{H}_t = \text{diag}(H_{t1}, H_{t2}, \dots, H_{tJ})$ . Then the posterior distribution for each  $x_t$  is

$$p(\mathbf{x}_t | \Phi_t, y_t, \mathcal{H}_t) = N(\mathbf{h}_t + \mathbf{b}_t c_t, \mathbf{H}_t - \mathbf{b}_t \mathbf{b}'_t g_t) \quad (9)$$

where  $c_t = y_t - \theta_{t0} - \mathbf{h}'_t \boldsymbol{\theta}_{t,1:J}$ ,  $g_t = v_t + \boldsymbol{\theta}'_{t,1:J} \mathbf{H}_t \boldsymbol{\theta}_{t,1:J}$  and  $\mathbf{b}_t = \mathbf{H}_t \boldsymbol{\theta}_{t,1:J} / g_t$ . This is easily computed and then sampled independently for each  $1:T$  to provide resimulated agent states over  $1:T$ .

2. In some cases, as in our study in this paper, the agent forecast densities will be those of Student T distributions. In our case study the four agents represent conjugate dynamic linear models in which all forecast densities are T, with parameters varying over time and with step-ahead forecast horizon. In such cases, standard Bayesian augmentation methods apply to enable simulation. Each T distribution is expressed as a scale mixture of normals, with the mixing scale parameters introduced as inherent latent variables with inverse gamma distributions. This expansion of the parameter space makes the T distributions conditional normals, and the mixing scales are resampled (from implied conditional posterior inverse gamma distributions) for each MCMC iterate along with the agent states. This is again a standard MCMC approach and much used in Bayesian time series, as in other areas (e.g. [Frühwirth-Schnatter 1994](#); [West and Harrison 1997](#), Chap 15). Then, conditional on the current values of these latent scales, sampling the  $x_t$  reduces technically to that conditional normals above.

Specifically, suppose that  $h_{tj}(x_{tj})$  is density of the normal  $T_{n_{tj}}(h_{tj}, H_{tj})$ ; the notation means that  $(x_{tj} - h_{tj})/\sqrt{H_{tj}}$  has a standard Student T distribution with  $n_{tj}$  degrees of freedom. Then latent scale factors  $\phi_{tj}$  exist such that: (i) conditional on  $\phi_{tj}$ , latent agent factor  $x_{tj}$  has a conditional normal density  $x_{tj}|\phi_{tj} \sim N(h_{tj}, H_{tj}/\phi_{tj})$  independently over  $t, j$ ; (ii) the  $\phi_{tj}$  are independent over  $t, j$  with gamma distributions,  $\phi_{tj} \sim G(n_{tj}/2, n_{tj}/2)$ . Then, at each MCMC step, the above normal update for latent agent states is replaced by the following two steps:

- i. Based on current values of all  $\phi_{tj}$  simulated at the last MCMC iterate, the normal update above applies to resample each  $x_t$  vector; the only modification in the multivariate normal conditional of eqn. (9) is that the diagonal matrix  $\mathbf{H}_t$  is now given by  $\mathbf{H}_t = \text{diag}(H_{t1}/\phi_{t1}, H_{t2}/\phi_{t2}, \dots, H_{tJ}/\phi_{tJ})$ .
- ii. Conditional on these new samples of the  $x_t$ , updated samples of the latent scales are drawn– independently for all  $t, j$ – from the implied set of conditional gamma posteriors  $\phi_{tj}|x_{tj} \sim G((n_{tj} + 1)/2, (n_{tj} + d_{tj})/2)$  where  $d_{tj} = (x_{tj} - h_{tj})^2/H_{tj}$ .
3. In some cases, agent densities may be more elaborate mixtures of normals, such as (discrete or continuous) location and/or scale mixtures that represent asymmetric distributions. The same augmentation strategy can be applied in such cases, with augmented parameters including location shifts in place of, or in addition to, scale shifts.

4. In other cases, we may be able to directly simulate the agent forecast distributions and evaluate forecast density functions at any point, but do not have access to analytic forms. One class of examples is when the agents are simulation models, e.g., DSGE models. Another involves forecasts in terms of histograms. In such cases, MCMC will proceed using some form of Metropolis-Hastings algorithm, or accept/reject methods, or importance sampling for the latent agent states.

For example, suppose we only have access to simulations from the agent forecast distributions, in terms of  $I$  independent draws from each collated in the simulated vectors  $\mathbf{x}_t^{(i)}$  for  $i = 1:I$ . We can apply importance sampling as follows: (a) compute the marginal likelihood values  $p(y_t|\Phi_t, \mathbf{x}_t^{(i)}, \mathcal{H}_t)$  for each  $i = 1:I$ ; (b) compute and normalize the implied importance sampling weights  $w_{ti} \propto N(y_t|\Phi_t, \mathbf{x}_t^{(i)}, \mathcal{H}_t)$ , and then (c) resample latent agent states for this MCMC stage according to the probabilities these weights define.

## B Appendix: Additional Graphical Summaries from Inflation Forecasting Analysis

This appendix lays out additional graphical summaries of results from the inflation forecasting analysis in the paper, providing material supplementary to that discussed in Section 4.

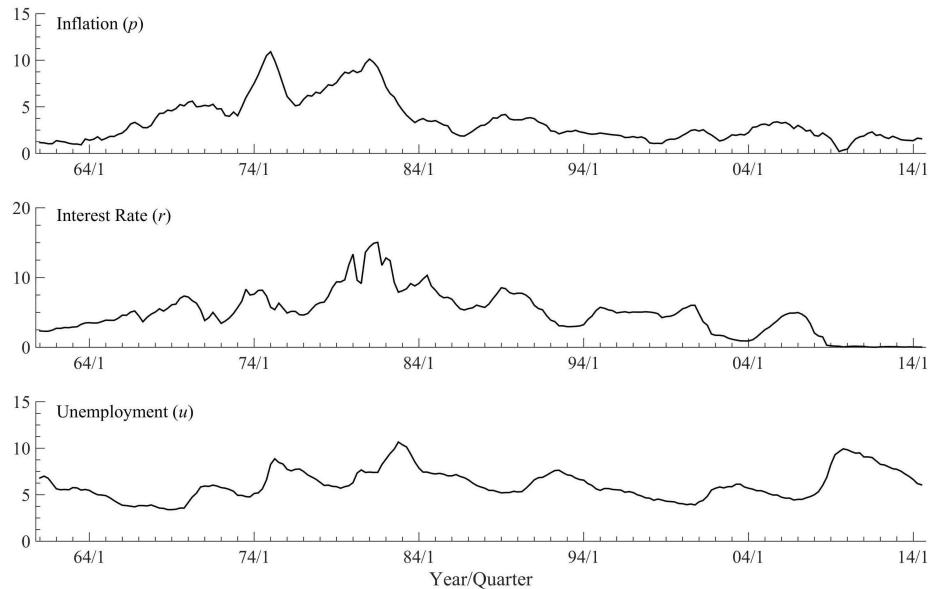


Figure B1: US inflation rate forecasting 1990/Q1-2014/Q4: U.S. macroeconomic time series (indices  $\times 100$  for % basis): annual inflation rate ( $p$ ), short-term nominal interest rate ( $r$ ), and unemployment rate ( $u$ ).

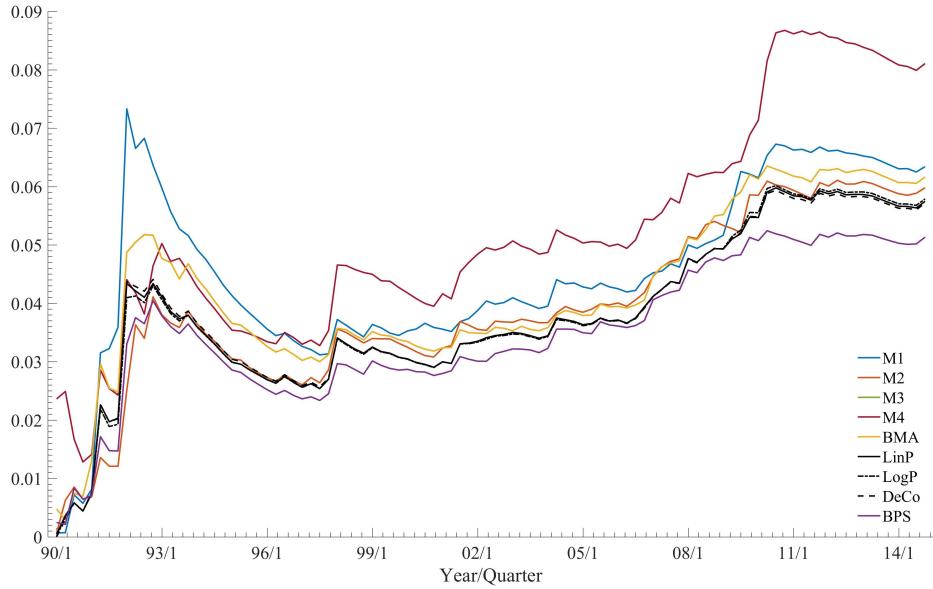


Figure B2: US inflation rate forecasting 1990/Q1-2014/Q4: Mean squared 1-step ahead forecast errors  $\text{MSFE}_{1:t}(1)$  sequentially revised at each of the  $t = 1:100$  quarters. As with forecast uncertainties, point forecast accuracy is particularly improved under BPS at crisis periods, with MSFE staying relatively level while significantly increasing for other models and methods.

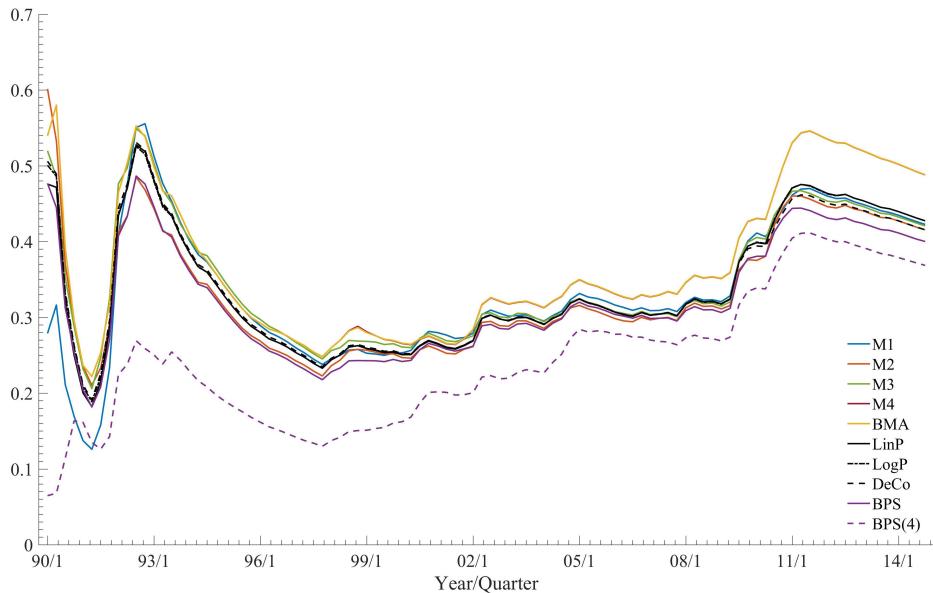


Figure B3: US inflation rate forecasting 1990/Q1-2014/Q4: Mean squared 4-step ahead forecast errors  $\text{MSFE}_{1:t}(4)$  sequentially revised at each of the  $t = 1:100$  quarters. Customized to the 4-step ahead horizon, BPS(4) dominates in point forecast accuracy over all models and methods, including the direct BPS extrapolation.

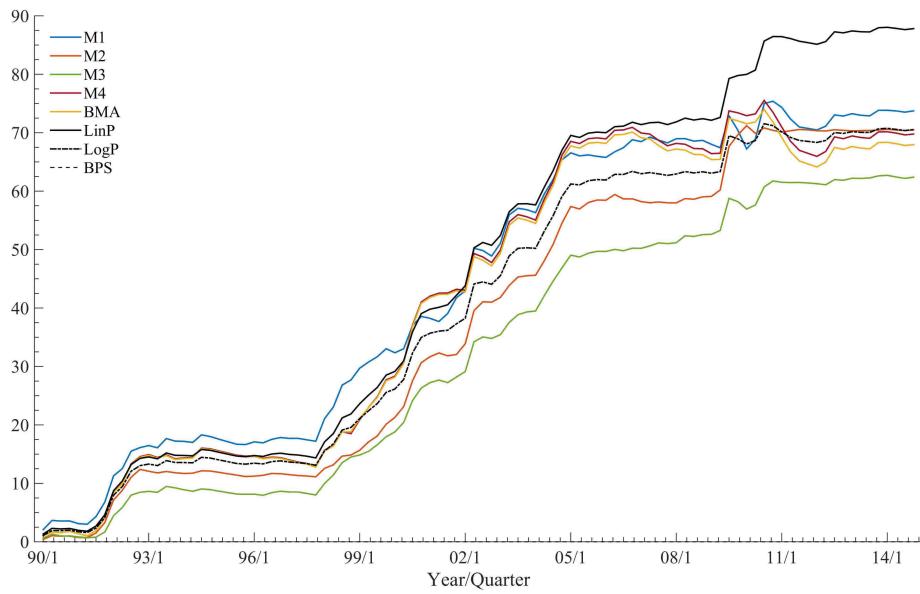


Figure B4: US inflation rate forecasting 1990/Q1-2014/Q4: 4-step ahead log predictive density ratios  $LPDR_{1:t}(4)$  sequentially revised at each of the  $t = 1:100$  quarters using direct projection from the 1-step ahead BPS model (baseline at 0 over time). Direct BPS extrapolation performs relatively poorly as– being inherently calibrated to 1-step model fit– it fails to adequately represent the increased uncertainty associated with long term forecasts in addition to relying on less accurate model forecasts. In contrast, BPS(4) is able to improve by adjusting to the increased forecast uncertainties.

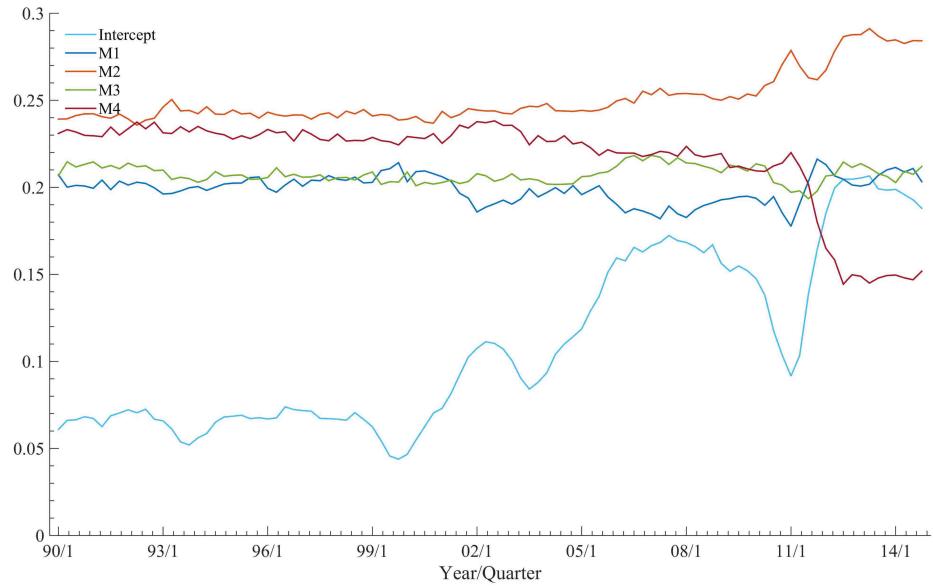


Figure B5: US inflation rate forecasting 1990/Q1-2014/Q4: On-line posterior means of BPS(4) model coefficients  $\theta_t$  sequentially computed at each of the  $t = 1:100$  quarters. There is a notable reduction in adaptability over time relative to the 1-step BPS coefficients; this is expected as the agents' forecasts are less reliable at longer horizons, so the data-based information advising the changes in posteriors over time is relatively limited. Further, the 4-step ahead trajectories are quite different from the 1-step trajectories, reasonably reflecting the differing forecasting abilities of the models at differing horizons.

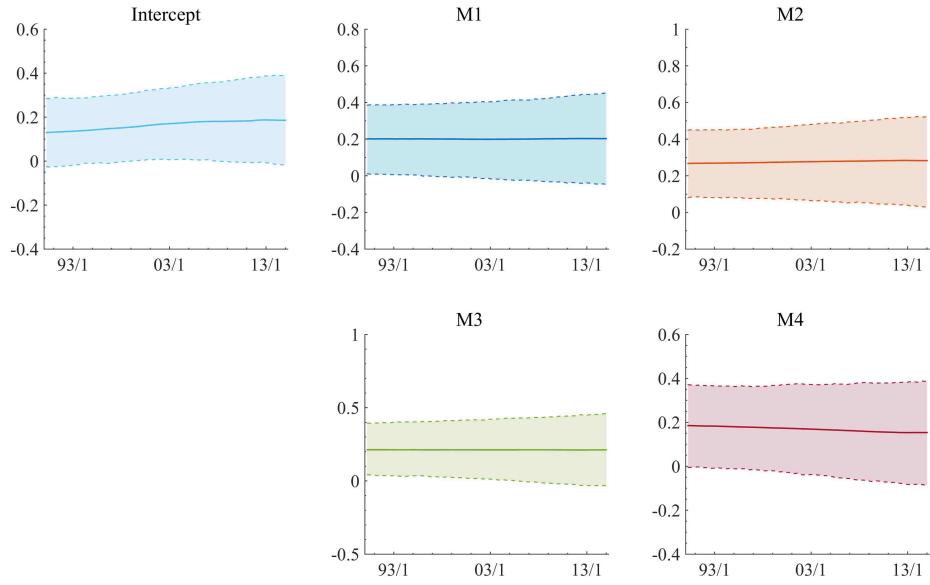


Figure B6: US inflation rate forecasting 1990/Q1-2014/Q4: Retrospective posterior trajectories of the BPS(4) model coefficients based on data from the full  $t = 1:100$  quarters. Posterior means (solid) and 95% credible intervals (shaded).

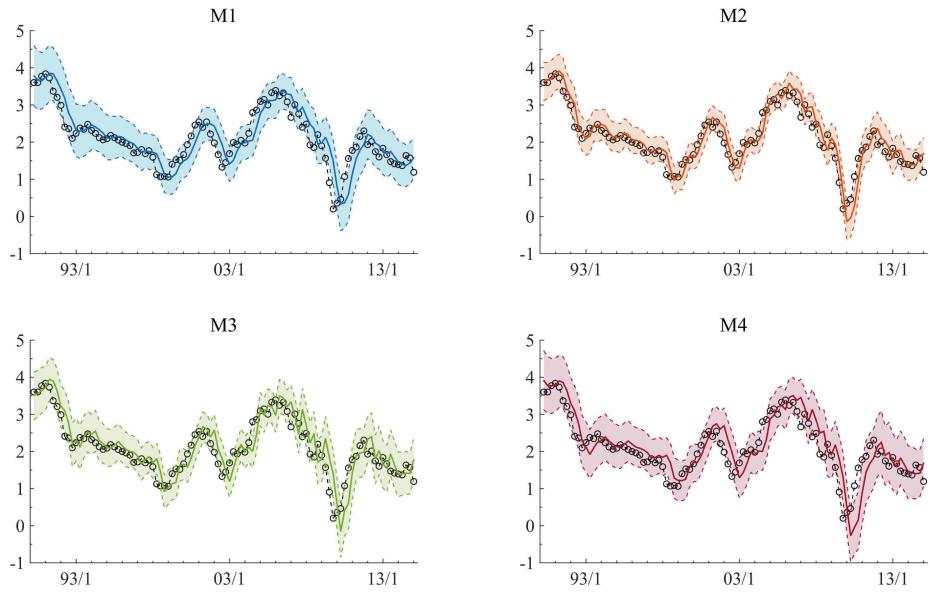


Figure B7: US inflation rate forecasting 1990/Q1-2014/Q4: BPS model-based posterior trajectories of the latent agent states  $x_{tj}$  for  $j = 1:4$  over the  $t = 1:100$  quarters. Posterior means (solid) and 95% credible intervals (shaded) from the MCMC analysis, with data  $y_t \equiv p_t$  (circles).

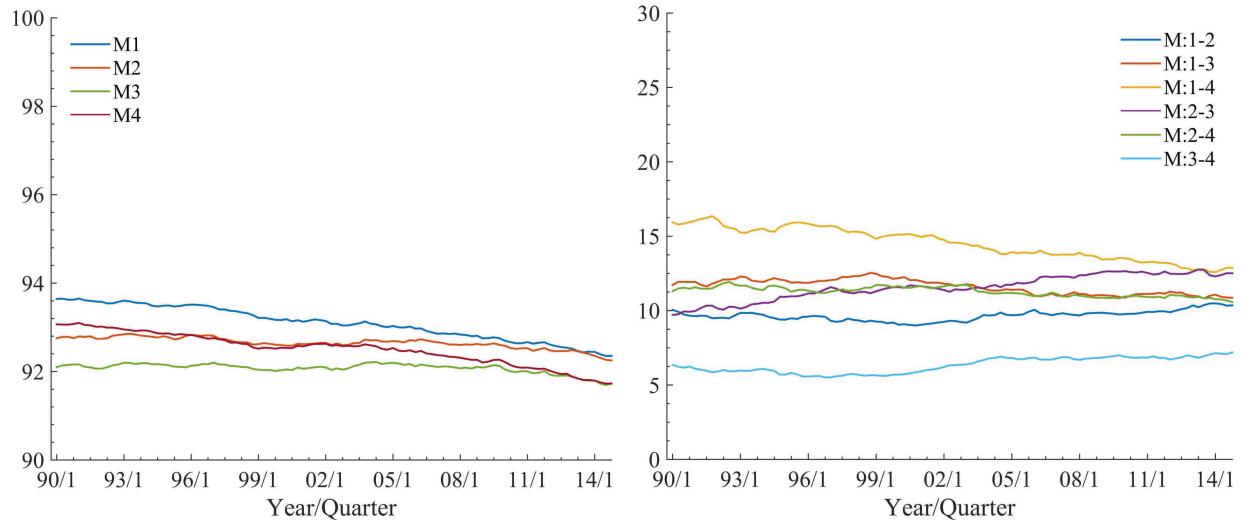


Figure B8: US inflation rate forecasting 1990/Q1-2014/Q4: BPS(4) model-based trajectories of 4-step ahead MC-empirical  $R^2$  (left) and paired MC-empirical  $R^2$  (right) in the posterior for the latent agent states  $x_{jt}$  for  $j = 1:4$  over the  $t = 1:100$  quarters.

## C Appendix: Additional Assessment Summaries for Simulation Study

This appendix lays out graphical summaries of results of analysis on one simulated data as noted in Section 4.4.

### C.1 Sequential Forecast Performance and On-line Inference

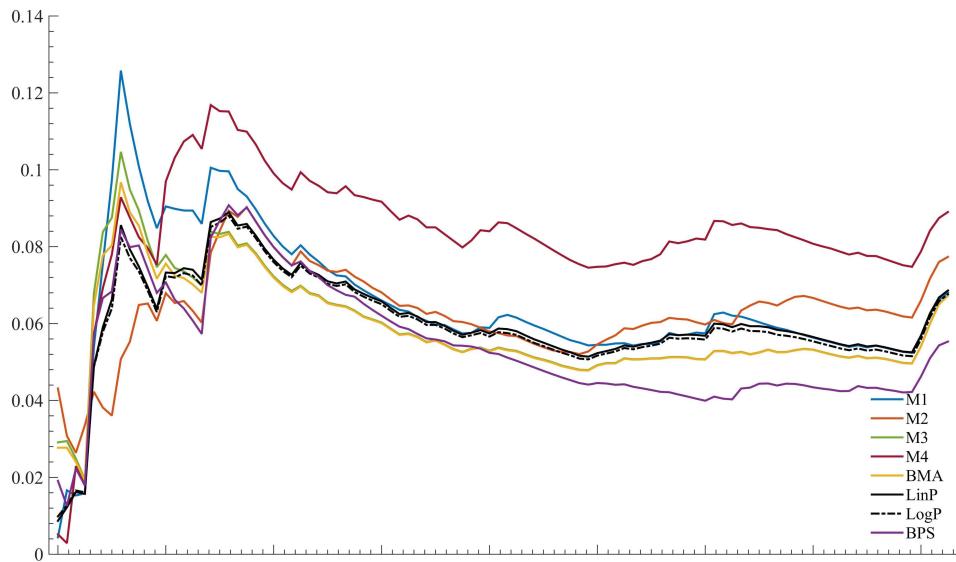


Figure C1: Simulated data forecasting: Mean squared 1-step ahead forecast errors  $\text{MSFE}_{1:t}$  in the analysis of the synthetic data. Note that BPS outperforms for the latter half of the analysis.

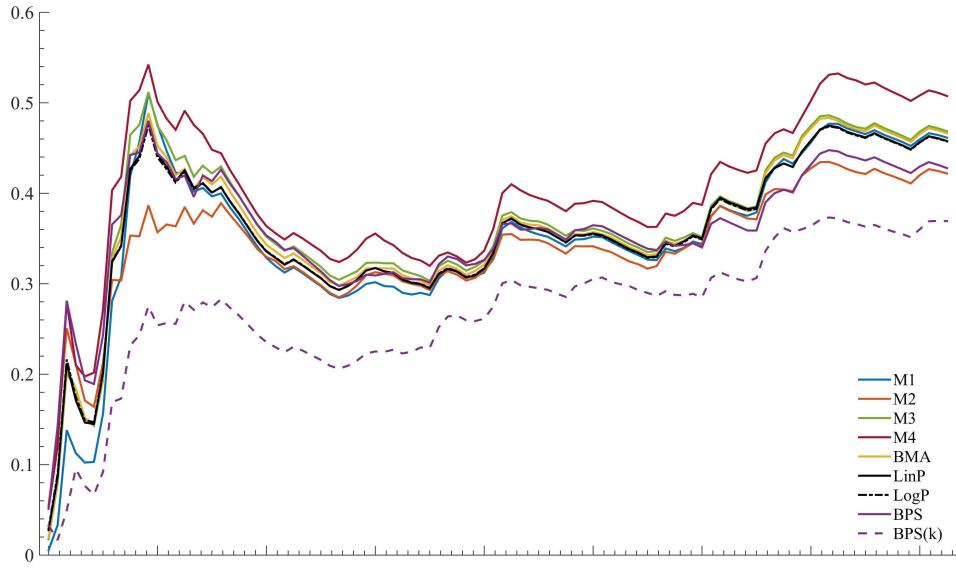


Figure C2: Simulated data forecasting: Mean squared 4-step ahead forecast errors  $\text{MSFE}_{1:t}$  in the analysis of the synthetic data. Note that, while BPS underperforms compared to the best model and BMA, BPS( $k$ ) dominates all models and strategies through  $1:T$ .

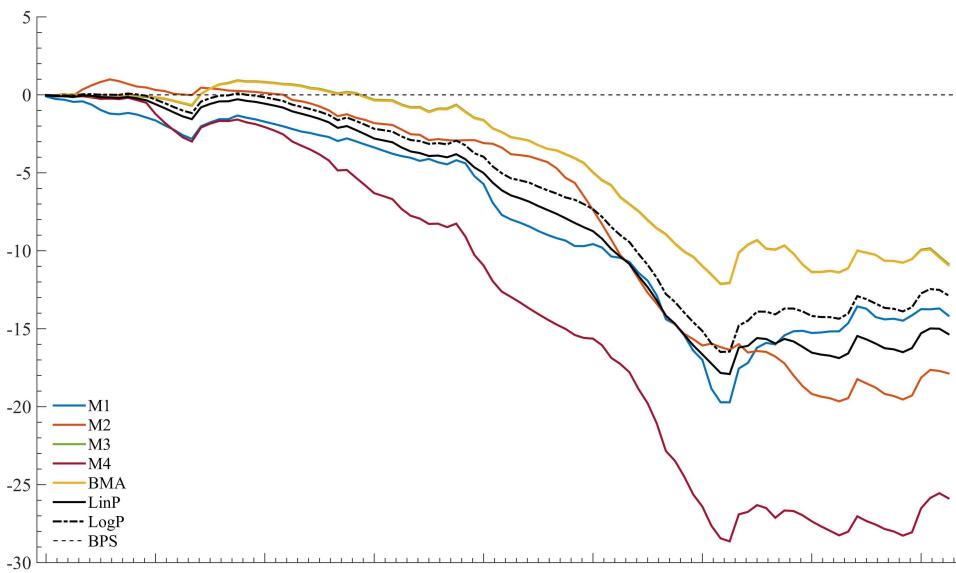


Figure C3: Simulated data forecasting: Cumulative sum of 1-step ahead log predictive density ratios  $\text{LPDR}_{1:t}$  in the analysis of the synthetic data. Though BPS underperforms slightly in the beginning, it dominates the others in the later half of the analysis.

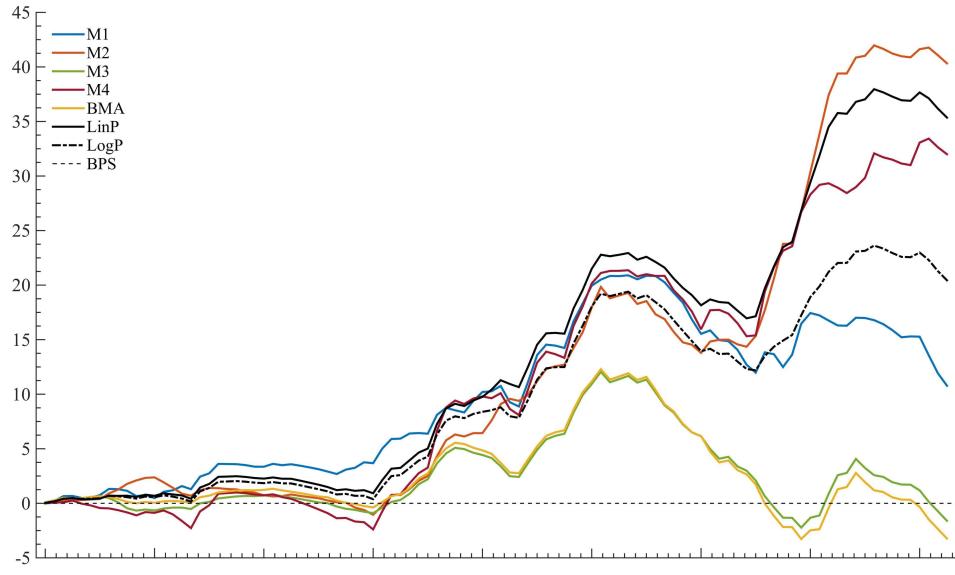


Figure C4: Simulated data forecasting: Cumulative sum of 4-step ahead log predictive density ratios  $\text{LPDR}_{1:t}$  in the analysis of the synthetic data. BPS overfits to the forecasts and fails to correctly estimate the forecast standard deviation, leading to poor density forecasts.

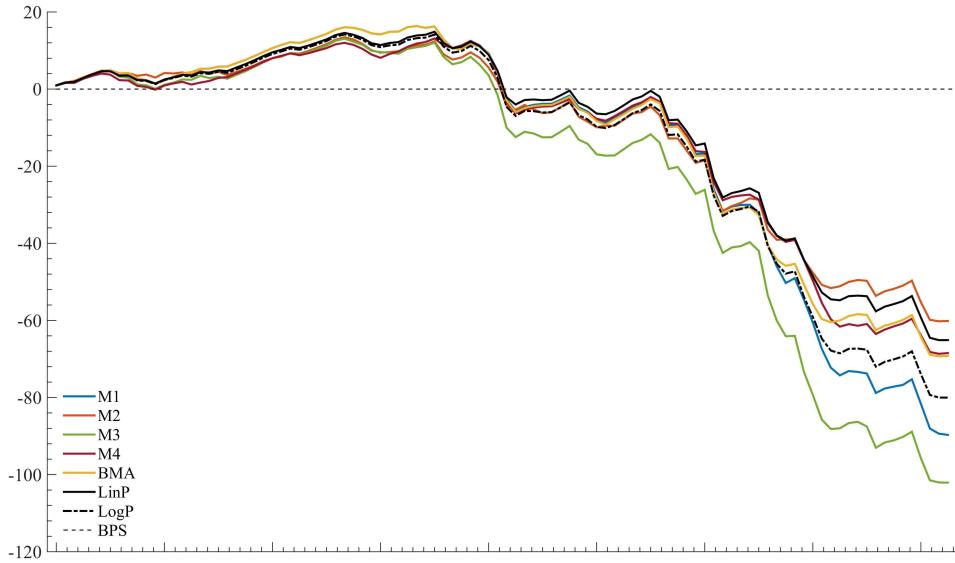


Figure C5: Simulated data forecasting: Cumulative sum of 4-step ahead log predictive density ratios  $\text{LPDR}_{1:t}$  in the analysis of the synthetic data. Contrary to Fig. C4, BPS( $k$ ) succeeds in forecasting uncertainty and is able to achieve superior density forecasts. Note that, BPS( $k$ ) does better when the model switching is less frequent (Fig. C6).

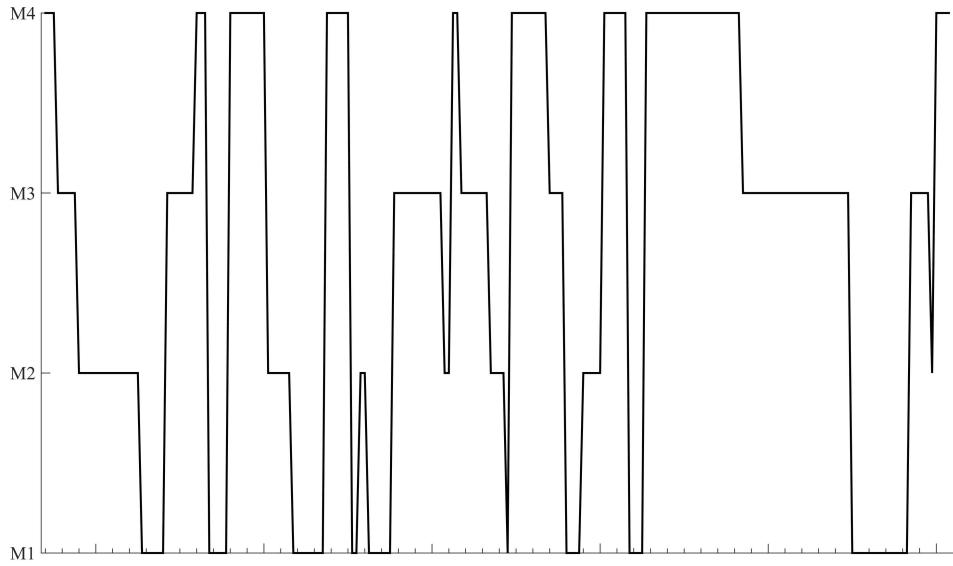


Figure C6: Simulated data forecasting: Models chosen in simulating the synthetic data.

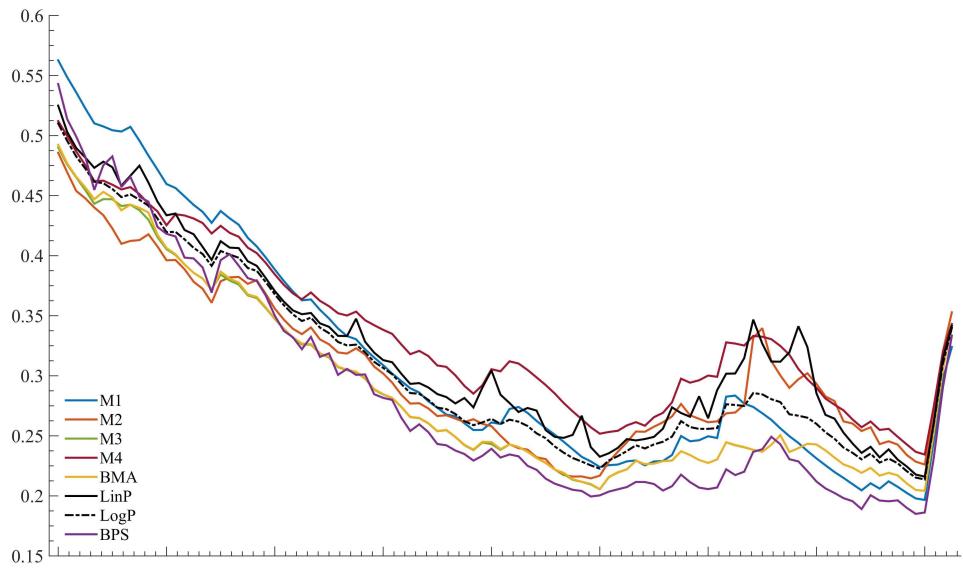


Figure C7: Simulated data forecasting: 1-step ahead forecast standard deviations in the analysis of the synthetic data.

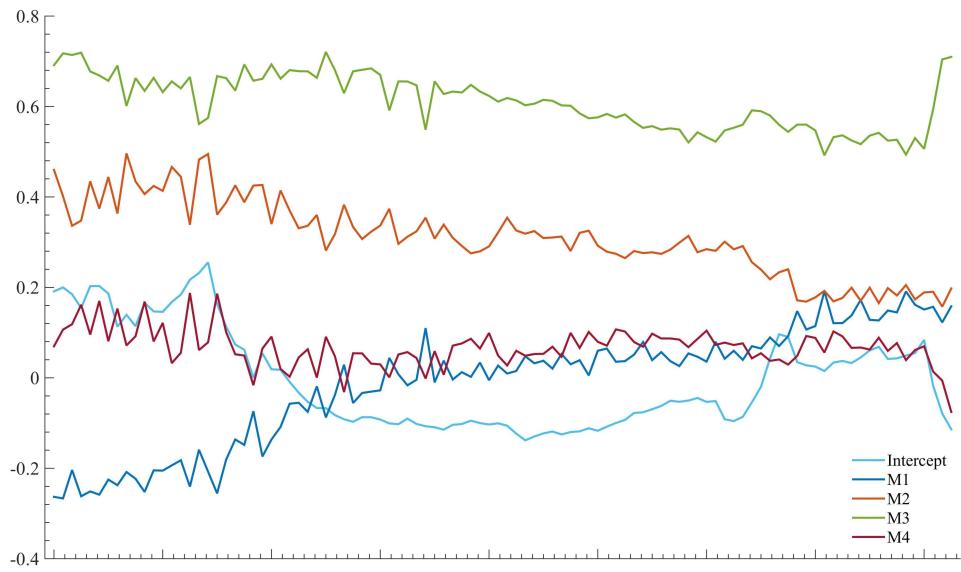


Figure C8: Simulated data forecasting: 1-step ahead forecast coefficients of BPS in the analysis of the synthetic data.

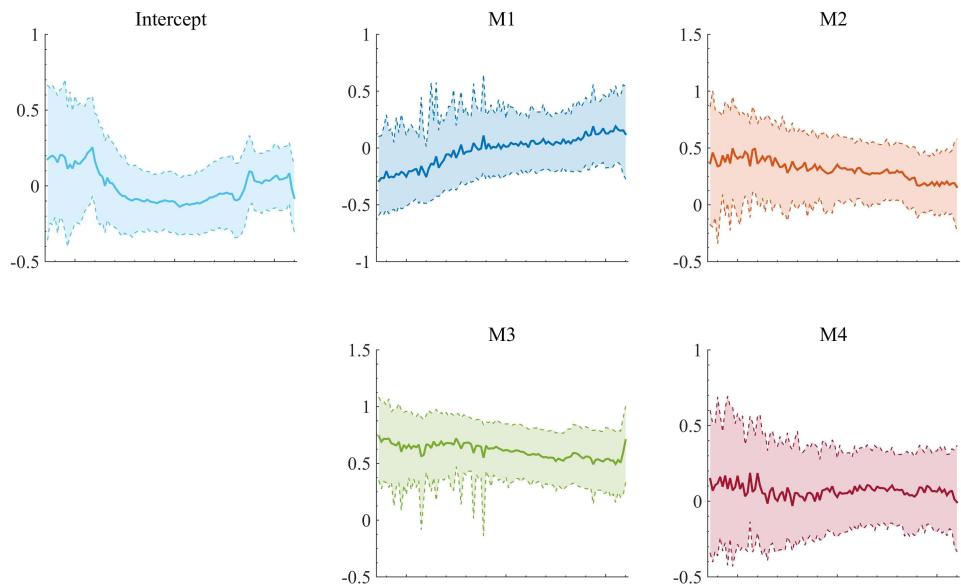


Figure C9: Simulated data forecasting: 1-step ahead forecast coefficients of BPS in the analysis of the synthetic data with credible intervals: posterior means (solid) and 95% credible intervals (dotted).

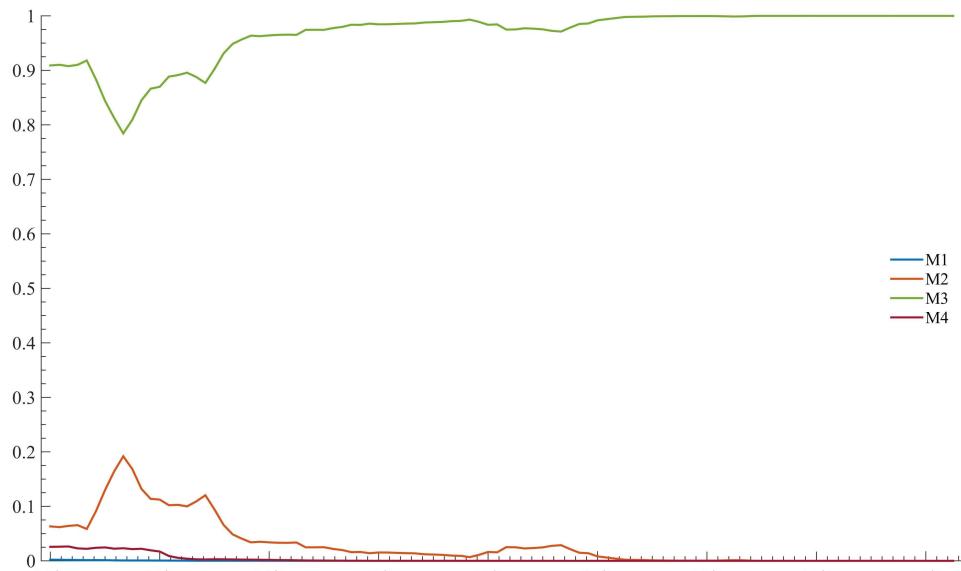


Figure C10: Simulated data forecasting: 1-step ahead forecast weights of BMA in the analysis of the synthetic data.

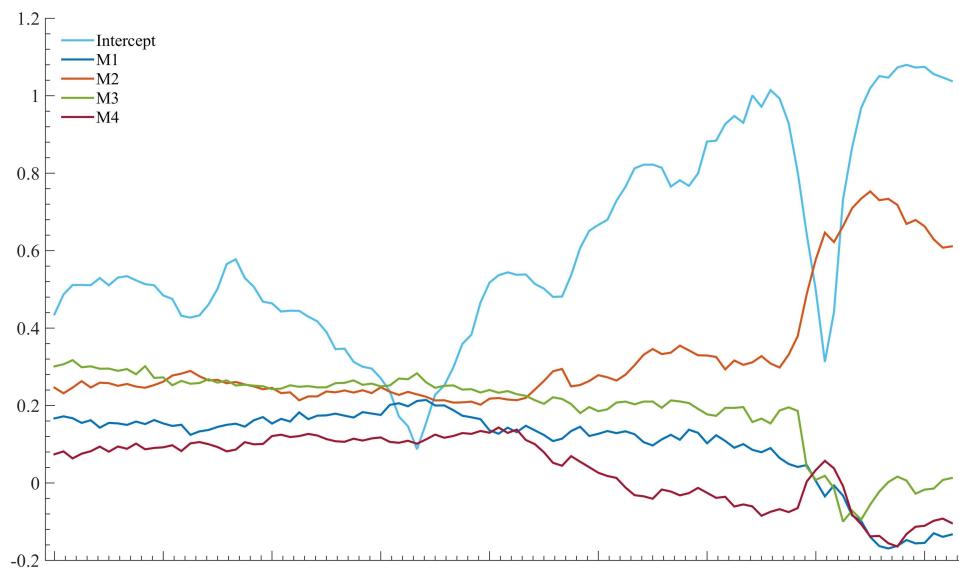


Figure C11: Simulated data forecasting: 4-step ahead forecast coefficients of BPS in the analysis of the synthetic data.

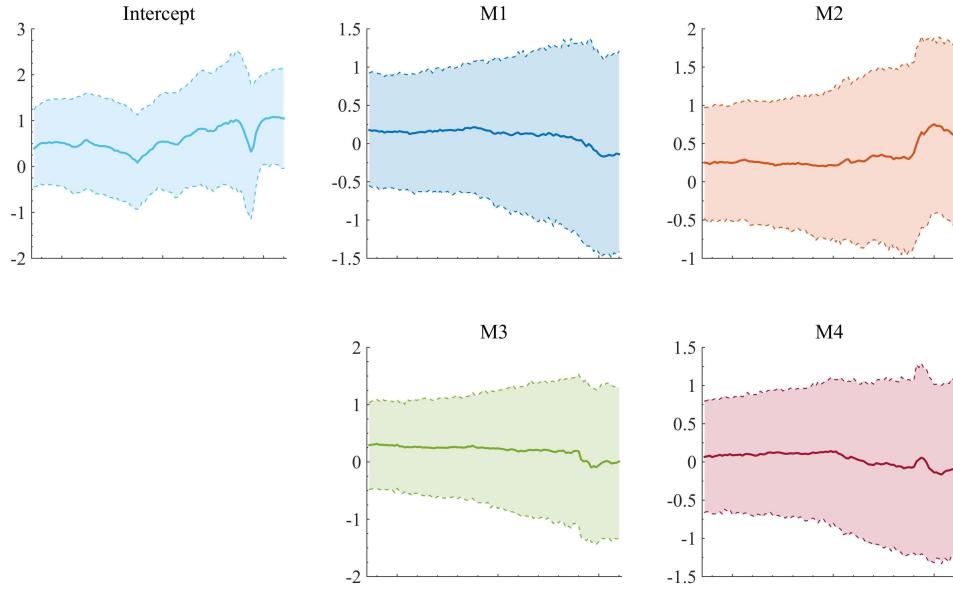


Figure C12: Simulated data forecasting: 4-step ahead forecast coefficients of BPS in the analysis of the synthetic data with credible intervals: posterior means (solid) and 95% credible intervals (dotted)..

## C.2 Retrospective Posterior Analysis

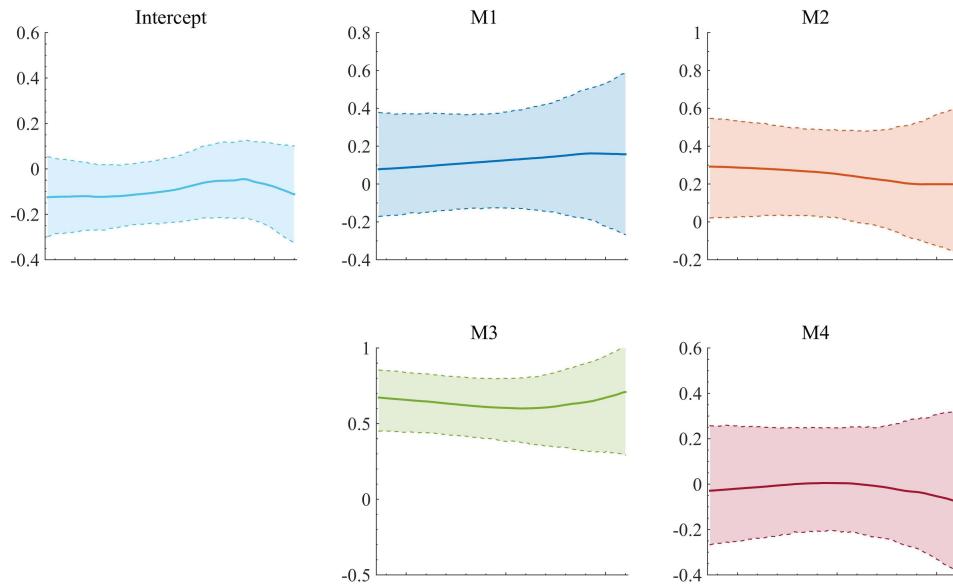


Figure C13: Simulated data forecasting: Posterior trajectories of the BPS coefficients in the analysis of the synthetic data for 1-step ahead forecasts.

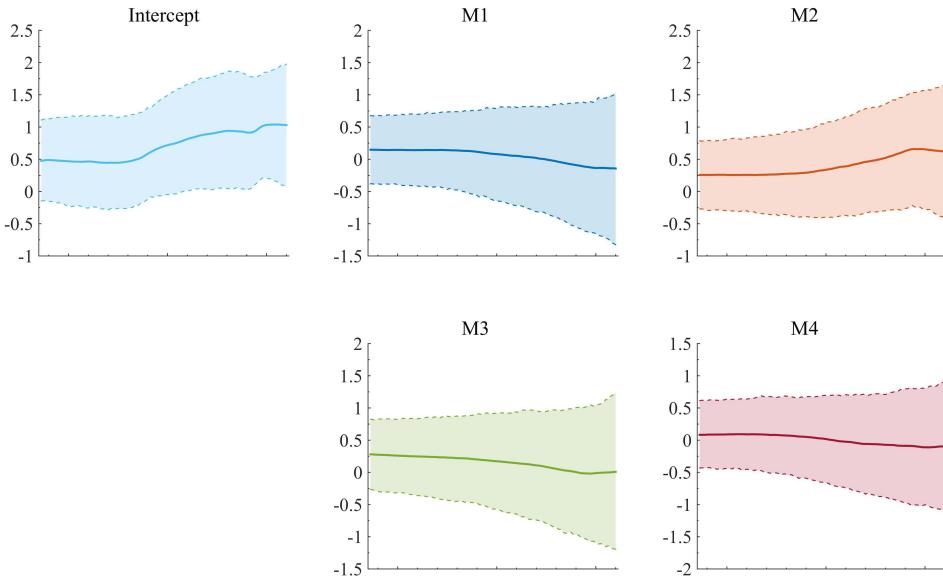


Figure C14: Simulated data forecasting: Posterior trajectories of the BPS coefficients in the analysis of the synthetic data for 4-step ahead forecasts.

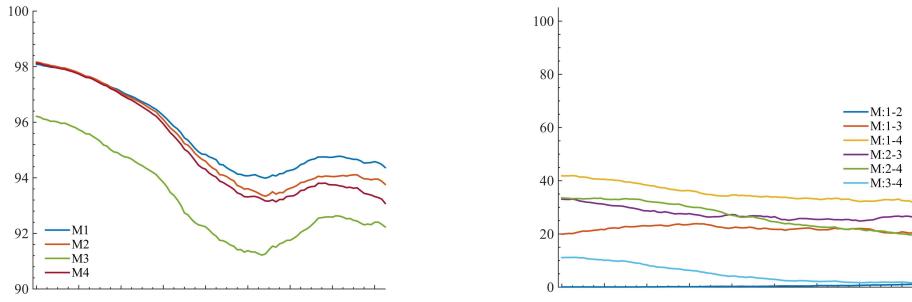


Figure C15: Simulated data forecasting: MC-empirical  $R^2$  (Left) and Paired MC-empirical  $R^2$  (Right) of the forecasts in the analysis of synthetic data using 1-step ahead forecasts.

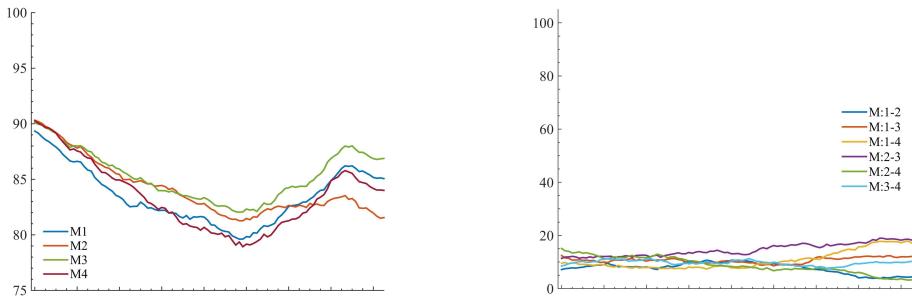


Figure C16: Simulated data forecasting: MC-empirical  $R^2$  (Left) and Paired MC-empirical  $R^2$  (Right) of the forecasts in the analysis of synthetic data using 4-step ahead forecasts.