Dynamic Bayesian Predictive Synthesis in Time Series Forecasting

Kenichiro McAlinn* & Mike West

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

January 29, 2016

Abstract

We discuss model and forecast comparison, calibration, and combination from a foundational perspective. Bayesian predictive synthesis (BPS) defines a coherent theoretical basis for combining multiple forecast densities, whether from models, individuals, or other sources, and extends existing forecast pooling and Bayesian model mixing methods. Time series extensions are implicit dynamic latent factor models, allowing adaptation to time-varying biases, mis-calibration, and dependencies among models or forecasters. Bayesian simulation-based computation enables implementation. A macroeconomic time series study highlights insights into 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

*Corresponding author. Tel: +1 919 684 4210. Fax: +1 919 684 8594

E-mail: kenichiro.mcalinn@duke.edu (Ken McAlinn), mw@stat.duke.edu (Mike West)

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 issues promote this interest. The increased adoption of formal forecasting models that yield full density forecasts has generated an interest in adaptive methods of density forecast combination, leading to new "density pooling" algorithms that aim to correct forecast biases and extend traditional "ad hoc" pooling rules. From a Bayesian perspective, formal model uncertainty and mixing analysis is "optimal" when a closed set of models generate forecast densities; in practice, it suffers from several limitations. First, Bayesian model combination ignores forecast and decision goals, scoring models on purely statistical grounds; in policy or financial applications, for example, the "best" decision models may not be highest posterior probability, or Bayesian averaged, models. Second, time series forecasting accuracy typically differs with forecast horizon; a model that forecasts well one quarter ahead may be useless for several quarters ahead, so calling for combination methods specific to the forecast horizon. Related to this, formal Bayesian model probabilities inherently score only 1-step ahead forecasting accuracy. Third, Bayesian model probabilities converge/degenerate– typically fast in time/with sample size– and always to the "wrong" model; this leads to lack of adaptability in sequential forecasting in time series. Fourth, the traditional statistical framework does not easily– if at all– apply to contexts where multiple forecasters, or groups of forecasters, generate forecast densities from their own models and perspectives; there is a need for integration of information about the forecasters, their anticipated biases, and– critically– relationships and dependencies among them. Note that these comments criticize the mechanical application of formal Bayesian model probabilities, and *not* Bayesian thinking and methodology *per se*.

Recent literature includes creative ideas for forecast density pooling, defining new empirical models fitted by Bayesian methods (e.g. 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). Some of these direct/empirical models for forecast density calibration and combination demonstrate improved forecast performance in studies in macroeconomics and finance. These methods advance the broader field that, since the seminal paper by Bates and Granger (1969), has drawn on interests and expertise from business, economics, technology, meteorology, management science, military intelligence, seismic risk, and environmental risk, among other areas (e.g. Clemen 1989; Clemen and Winkler 1999; Timmermann 2004; Clemen and Winkler 2007). Recent interest in macroeconomics, driven partly by a need to improve information flows to policy and decision makers at national and international levels, represents development with potential to impact more broadly. The challenge is to define formal, reliable methodology for integrating predictive information from multiple professional forecasters and models, and sets of competing econometric models in a more traditional statistical context.

The developments reported here respond to this movement, motivated by an interest in the question of foundational underpinnings of some of the specific algorithmic/empirical models for forecast density combination recently introduced. While a new combination rule/algorithm may demonstrate success in a specific case study, 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 methodological generalizations of practical import. We address this by linking to the historical literature on subjective Bayesian thinking about the broad field of assessing and combining subjective probabilities. In particular, we revisit 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) that requires a formal, subjective Bayesian context for treating multiple models or forecasters as providers of "forecast data" to be used in prior-posterior updating by a coherent decision maker (See also; West and Harrison 1997, Sect 16.3.2).

The resulting ideas of *Bayesian Predictive Synthesis* (BPS) build on foundational theory in West (1992) that defines a rather broad framework with specific functional forms of posterior predictive distributions that "synthesize" sets of forecast densities. The framework provides interpretation of traditional and recently introduced pooling methods as special cases. More importantly from a practical time series forecasting perspective, development of BPS for sequential forecasting of time series enables the use of flexible, adaptive Bayesian dynamic models that are able to respond to changes in characteristics of sets of models and forecasters over time. We discuss this, and develop dynamic latent factor regression models (West and Harrison 1997; Prado and West 2010; West 2013) to both exemplify the framework and to define practically relevant and useful special cases. As we demonstrate in a topical macroeconomic time series study, BPS has the potential to define fully Bayesian, interpretable models that can adapt to time-varying biases and mis-calibration of multiple models or forecasters, and generate useful insights into patterns of relationships and dependencies among them while also improving forecast accuracy.

Section 2 briefly summarizes the foundations of BPS, underlying our novel theory and methods for dynamic problems in Section 3. In Section 4, analyses of U.S macroeconomic time series with a focus on forecasting quarterly inflation illustrates and highlights the benefits of the new framework. Additional comments in Section 5 conclude the paper.

Some notation: We use lower case bold font for vectors and upper case bold font for matrices. Vectors are columns by default. Distributional notation $y \sim N(f, v)$, $\boldsymbol{x} \sim N(\boldsymbol{a}, \boldsymbol{A})$ 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 $s, s+1, \ldots, t$ when $s < t$, such as in $y_{1:t} = \{y_1, \ldots, y_t\}$.

2 Background and Foundations

We develop dynamic models for time series that build on theoretical foundation in West (1992) concerned with predicting a single outcome y. In time series generalizations, y will be the outcome of a univariate time series at one time point, typically real-valued, although the foundational theory is general. We first summarize the basic ideas and key historical result in this single outcome setting.

A Bayesian decision maker $\mathcal{D}$ is interested in predicting the outcome y and aims to incorporate information from J individual agents (models, forecasters, or forecasting agencies, etc.) labelled $\mathcal{A}_j$, $(j = 1{:}J)$. To begin, $\mathcal{D}$ has prior $p(y)$; then each $\mathcal{A}_j$ provides $\mathcal{D}$ with forecast information in terms of a p.d.f. $h_j(y)$. These forecast densities represent the individual inferences from the agents, and define the information set $\mathcal{H} = \{h_1(\cdot), \dots, h_J(\cdot)\}$ now available to $\mathcal{D}$. Formal subjective Bayesian analysis indicates that $\mathcal{D}$ will predict y using the implied posterior $p(y|\mathcal{H})$ from a full Bayesian prior-to-posterior analysis. 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.

West (1992) extended prior theory (Genest and Schervish 1985; West and Crosse 1992) to show that there exists a restricted class of Bayesian models $p(y, \mathcal{H})$ under which the required posterior has the form

$$p(y|\mathcal{H}) = \int \alpha(y|\boldsymbol{x}) \prod_{j=1:J} h_j(x_j) dx_j \tag{1}$$

where $\boldsymbol{x} = x_{1:J} = (x_1, \dots, x_J)'$ is a $J-$dimensional latent vector and $\alpha(y|\boldsymbol{x})$ a conditional p.d.f. for y given $\boldsymbol{x}$. This posterior form relies on, and must be consistent with, $\mathcal{D}$'s prior

$$p(y) = \int \alpha(y|\boldsymbol{x}) m(\boldsymbol{x}) d\boldsymbol{x} \qquad \text{where} \qquad m(\boldsymbol{x}) = E[\prod_{j=1:J} h_j(x_j)], \tag{2}$$

the expectation in the last formula being over $\mathcal{D}$'s distribution $p(\mathcal{H})$. Critically, the representation of eqn. (1) does not require a full specification of $p(y, \mathcal{H})$ (and hence $p(\mathcal{H})$), but only the prior $p(y)$ and the marginal expectation function $m(\boldsymbol{x})$ of eqn. (2). These specifications alone do not, of course, indicate what the functional form of $\alpha(y|\boldsymbol{x})$ is, which opens the path to developing models based on different specifications. Key to considering this is the interpretation of the latent vector $\boldsymbol{x}$. From eqn. (1), note two implications/interpretations:

- Suppose each agent $\mathcal{A}_j$ *simulates* a single draw from $h_j(y)$; label these draws $\boldsymbol{x} = x_{1:J}$. Then $\mathcal{D}$ can immediately simulate from the implied posterior $p(y|\mathcal{H})$ by sampling $y \sim \alpha(y|\boldsymbol{x})$.
- Suppose the hypothetical agent information $h_j(y) = \delta_{x_j}(y)$ 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 $\alpha(y|\boldsymbol{x})$.

This aids understanding of the role of $\alpha(y|\boldsymbol{x})$ as $\mathcal{D}$'s model for converting sets of simulated, or supposedly exact predicted values (or "oracle" values) from agents into his/her revised predictions of y. We refer to the x_j as the *latent agent states*.

From eqn. (2), note that $(y, \boldsymbol{x})$ have an implicit joint distribution with margins $p(y)$ and $m(\boldsymbol{x})$, so we can consider this is understanding the ways in which the framework allows $\mathcal{D}$ to incorporate views, and historical information, about agent-specific biases, patterns of mis-calibration, inter-dependencies among agents and their relative expertise/accuracy. The margin for latent agent states $m(\boldsymbol{x})$ is $\mathcal{D}$'s prior expectation of the product of agent densities; an example with $m(\boldsymbol{x})$ having positive dependencies among a subset of the x_j indicates that $\mathcal{D}$ anticipates positive concordance among the corresponding predictive densities $h_j(\cdot)$ of that subset of agents.

Examples 1. Johnson and West (2016) explore BPS examples in which $\alpha(y|\boldsymbol{x})$ is a mixture of point masses with a base prediction, $\alpha(y|\boldsymbol{x}) = a_0(\boldsymbol{x})\pi_0(y) + \sum_{1:J} a_j(\boldsymbol{x})\delta_{x_j}(y)$ for some agent state-dependent probabilities $a_{0:J}(\boldsymbol{x})$. That leads to a new interpretation of standard Bayesian model averaging, and justifies constant or data-dependent weighting in linear pooling of densities. In particular, it provides a formal theoretical basis for outcome-dependent density pooling as in recently successful empirical methods (e.g. Kapetanios et al. 2015; Aastveit et al. 2015; Pettenuzzo and Ravazzolo 2016, among others), with practical import in that BPS allows for the integration of information about agent inter-dependencies that are neglected or ignored by other approaches. The general BPS approach also explicitly allows and encourages customization of mixture weights with respect to forecast goals; this underpins, for example, the use of sets of BPS models in parallel for different forecast horizon (e.g. Aastveit et al. 2014, and references therein).

Examples 2. A second class of examples arises when the implied joint prior $\alpha(y|\boldsymbol{x})m(\boldsymbol{x})$ is multivariate normal or T, which easily and intuitively allow for: (i) ranges of agent biases and mis-calibration, viewed through shifts in means and/or variances of implied conditional distributions of individual conditional distributions $(x_j|y)$; and (ii) inter-dependencies, reflected in patterns of correlations and other aspects of conditional dependence among the x_j (West and Crosse 1992; West and Harrison 1997, Sect 16.3.2).

As a specific, conditionally normal example, suppose that the joint prior for $(y, \boldsymbol{x})$ is consistent with the margin $p(y)$ and a conditional for $(\boldsymbol{x}|y)$ that is normal with mean $\boldsymbol{\mu} + \boldsymbol{\beta}y$ and variance matrix $\boldsymbol{V}$; here $\boldsymbol{\mu}$ and $\boldsymbol{\beta}$ account for agent-specific biases while the diagonal elements of $\boldsymbol{V}$ reflect aspects of $\mathcal{D}$'s views about agent precisions and consistency, given the bias "corrections" in $\boldsymbol{\mu}$ and $\boldsymbol{\beta}$. The conditional has the form of a factor model with y as a single common factor underlying $\boldsymbol{x}$, and the variation among the entries in the "factor loading" vector $\boldsymbol{\beta}$ reflect key aspects of dependencies among agents. The residual variance matrix $\boldsymbol{V}$ may be diagonal, or there may be some additional agent-agent dependencies via non-zero correlations. Together with a normal prior $p(y)$, this yields the agent calibration density

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

and residual variance v. The agent bias and dependence parameters are theoretically mapped into the *effective* calibration parameters $(\boldsymbol{\theta}, v)$. In terms of the calibration and combination of agent forecast densities when available, the implied prior on these effective parameters is all that is needed, even though that may come indirectly through priors on $(\boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{V})$. Variants involving mixing over the scale parameters yield related conditional T distributions.

3 Dynamic BPS

The new methodological developments forming the core of this paper adapt and extend the basic BPS framework summarized above to sequential forecasting in time series. In particular, we develop dynamic extensions of conditionally normal BPS models (Section 2/Examples 2) involving time-varying parameters to characterize and formally allow for agent-specific biases, patterns of mis-calibration, 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, for clarity and examples, although the ideas and approach are immediately adaptable to multivariate cases.

Among the connections in recent literature mentioned in Section 1, 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 sets of models. This relates directly, as BPS provides a complete theoretical framework with implied underlying latent agent states arising from the agent distributions. 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

$$p(y_t|\boldsymbol{\Phi}_t, \boldsymbol{y}_{1:t-1}, \mathcal{H}_{1:t}) \equiv p(y_t|\boldsymbol{\Phi}_t, \mathcal{H}_t) = \int \alpha_t(y_t|\boldsymbol{x}_t, \boldsymbol{\Phi}_t) \prod_{j=1:J} h_{tj}(x_{tj}) dx_{tj} \tag{4}$$

where $\boldsymbol{x}_t = x_{t,1:J}$ is a $J-$dimensional latent agent state vector at time t, $\alpha_t(y_t|\boldsymbol{x}_t, \boldsymbol{\Phi}_t)$ is $\mathcal{D}$'s conditional calibration p.d.f. for y_t given $\boldsymbol{x}_t$, and $\boldsymbol{\Phi}_t$ represents time-varying parameters defining the calibration p.d.f.– parameters for which $\mathcal{D}$ has current beliefs represented in terms of a current (time $t-1$) posterior $p(\boldsymbol{\Phi}_t|y_{1:t-1}, \mathcal{H}_{1:t-1})$. The methodological focus can now rest on evaluation of models based on various assumptions about the form of $\alpha_t(y_t|\boldsymbol{x}_t, \boldsymbol{\Phi}_t)$ and its defining dynamic state parameters $\boldsymbol{\Phi}_t$. 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

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

the latter being the $(1+J)-$vector of time-varying bias/calibration coefficients. This defines the first component of the standard conjugate form DLM (West and Harrison 1997, Section 4)

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

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

where $\boldsymbol{\theta}_t$ evolves in time according to a linear/normal random walk with innovations variance matrix $v_t\boldsymbol{W}_t$ 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 ν_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 methods: (i) The time-varying intercept and agent coefficients $\boldsymbol{\theta}_t$ follow the random walk evolution of eqn. (6b) where $\boldsymbol{W}_t$ is defined via a standard, single discount factor specification (West and Harrison 1997, Sect 6.3; Prado and West 2010, Sect 4.3); (ii) 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 γ_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 $(\boldsymbol{\theta}_1, v_1)$ at $t=0$, the model is fully specified.

Eqns. (6) define a dynamic latent factor model: the $\boldsymbol{x}_t$ vectors in each $\boldsymbol{F}_t$ are latent variables. At each time, they are conceived as arising as single draws from the set of agent densities $h_{tj}(\cdot)$, the latter becoming available at time $t-1$ for forecasting y_t. Note that the latent factor generating process has the x_{tj} drawn *independently* from their $h_{tj}(\cdot)$– based on the BPS foundational theory of eqn. (4)– and externally to the BPS model. That is, coupled with eqns. (6a,6b), we have

$$p(\boldsymbol{x}_t|\boldsymbol{\Phi}_t, \boldsymbol{y}_{1:t-1}, \mathcal{H}_{1:t}) \equiv p(\boldsymbol{x}_t|\mathcal{H}_t) = \prod_{j=1:J} h_{tj}(x_{tj}) \tag{7}$$

for all time t and with $\boldsymbol{x}_t, \boldsymbol{x}_s$ conditionally independent for all $t \neq s$. Importantly, the independence of the x_{tj} 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 $\boldsymbol{\Phi}_t$.

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 x_* as well as the dynamic BPS model parameters $\boldsymbol{\Phi}_*$. 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 behaviors within groups of forecasters. The "output" of full posterior summaries for the $\boldsymbol{x}_t$ 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 $\{\boldsymbol{x}_{1:t}, \boldsymbol{\Phi}_{1:t}\}$, rather than restricting attention to forward filtering to update posteriors for current values $\{\boldsymbol{x}_t, \boldsymbol{\Phi}_t\}$; 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 structure of eqns. (6a,6b,7) leads easily to a two-component block Gibbs sampler for the latent agent states and dynamic parameters. These are iteratively resimulated from the 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 $p(\boldsymbol{\Phi}_{1:t}|\boldsymbol{x}_{1:t}, y_{1:t})$. 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 $p(\boldsymbol{x}_{1:t}|\boldsymbol{\Phi}_{1:t}, y_{1:t}, \mathcal{H}_{1:t})$. It is immediate that the $\boldsymbol{x}_t$ are conditionally independent over time t in this conditional distribution, with time t conditionals

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

In cases when each of the agent forecast densities is normal, this yields a multivariate normal for $\boldsymbol{x}_t$ 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 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 $\boldsymbol{x}_t$ 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).

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 $\boldsymbol{\Phi}_t$ 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 $\boldsymbol{\Phi}_{t+1} = \{\boldsymbol{\theta}_{t+1}, v_{t+1}\}$ from $p(\boldsymbol{\Phi}_{t+1}|y_{1:t}, \mathcal{H}_{1:t})$; (ii) Draw $\boldsymbol{x}_{t+1}$ 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 equal or 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, judgmental 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 $\boldsymbol{\Phi}_{t+1}, \boldsymbol{\Phi}_{t+2}, \ldots, \boldsymbol{\Phi}_{t+k}$ 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 $\boldsymbol{x}_{t+k}$; 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 is open to models customized to forecasting goals, and so provides $\mathcal{D}$ a strategy to focus modeling on the horizon of interest. This involves a trivial 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 $\{\boldsymbol{\theta}_t, v_t\}$ 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 Sections 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 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 US 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 US 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. 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, shown in Fig. 1. 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 US 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- $p_{t-3}, r_{t-3}, u_{t-3}$; M3- p_{t-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. Prior specifications for the DLM state vector and discount volatility model in each is based on $\boldsymbol{\theta}_0|v_0 \sim N(\mathbf{0}, v_0\boldsymbol{I})$ and $1/v_0 \sim G(1, 0.01)$, using the usual $(\boldsymbol{\theta}, v)$ DLM notation (West and Harrison 1997, Chap 4). Each agent model uses standard discount factor (β) specification for state evolution variances and discount factor (δ) for residual volatility; we use $(\beta, \delta) = (0.99, 0.95)$ in each of these agent models. The DLM-based forecast densities $h_{t-k,j}(x_{tj})$ are then those of predictive T distributions. MCMC-based model fitting adapts to introduce latent scale factors as noted in Section 3.3.

In the dynamic BPS models for forecast horizons $k = 1$ and $k = 4$, we take initial priors as $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}, 0.25\boldsymbol{I})$ with $\boldsymbol{m} = (0, \mathbf{1}'/p)'$ and $1/v_0 \sim G(5, 0.01)$. 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 $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}, 10^{-4}\boldsymbol{I})$ and $(\beta, \delta) = (0.99, 0.99)$. Differences by forecast horizon echo earlier discussion about different model choices being relevant to different forecast goals. 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; conclusions with respect to BPS do not change materially with different values close to those chosen for the summary examples.

Data Analysis and Forecasting. The 4 agent models are analyzed and synthesized as follows. First, the 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. This continues over 1977/Q2-1989/Q4, now accompanied by the standard, sequentially updated BMA analysis. Also, at each quarter t during this period, the MCMC-based BPS analysis is run using from 1977/Q2 data up to time t; that is, we repeat the analysis with an increasing "moving window" of past data as we move forward in time. We do this for the traditional 1-step 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 (that more recently has been referred to as 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). 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 (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

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

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, so 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 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 (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, 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.

1-Step Ahead Forecasting. Figs. 2-5 summarize sequential analysis for 1-step forecasting. Fig. 2 shows the 1-step ahead measures MSFE$_{1:t}(1)$ for each time t. BPS almost uniformly dominates, except at the beginning of the time period where the MSFE is somewhat unstable. Four "shock" periods are notable and increase forecast errors: 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. The sub-prime mortgage crisis period highlights this, with MSFE staying relatively level under BPS while the others significantly increase.

Fig. 3 confirms that BPS performs uniformly better than, or on par with, the other models and BMA based on LPDR measures that measure relative distributional form and dispersion of forecast densities as well as location. Major shocks and times of increased volatility have substantial impact on the relative performance, again most notable at the beginning of the sub-prime mortgage crisis. 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.

Fig. 4 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. 4 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.

We finally 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, BPS tends more highly weight agent M2, notable in terms of the on-line estimates of BPS agent coefficients in $\boldsymbol{\theta}_{tj}$; see Fig. 5. 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.

4-Step Ahead Forecasting. Figs. 6-10 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; see Figs. 6, 7, and 8. BPS performs relatively poorly in terms of LPDR as– being inherently calibrated to 1-step model fit– it fails to adequately represent the increased uncertainty associated with long term forecasts. Looking at the forecast standard deviations in Fig. 9, 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 as illustrated in Figs. 6 and 8. This again bears out the recommendation to directly synthesize forecasts on the horizon of interest.

Fig. 10 shows on-line estimates of the BPS(4) coefficients $\boldsymbol{\theta}_t$ as they are sequentially updated and adapt over time during the test period. There is a notable reduction in adaptability over time relative to the 1-step BPS coefficients (Fig. 5); 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 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 dominate in performance compared to all other methods as a result.

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 improve 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.

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. Retrospective posteriors for BPS (1-step) and BPS(4) model coefficients are summarized in Figs. 11 and 12, respectively, to compare with on-line point summaries in Figs. 5 and 10 earlier discussed. We see the expected smoothing of estimated trajectories of coefficients. 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 for 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 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 $\boldsymbol{x}_t$– in both the direct BPS and customized BPS(4). For illustration, we focus on the 1-step BPS analysis; Fig. 13 displays posterior trajectories for the x_{tj}, together with the inflation outcomes y_t; Fig. 14 is similar, but vertically centers the display by plotting trajectories for the forecast deviations $y_t - x_{tj}$. 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 $\boldsymbol{x}_{1:T}$. This is not a 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 $\boldsymbol{x}_t$ 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 $i = 1{:}J \backslash j$, 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. 15 and 16 displays trajectories over time for these two measures, based on each of the BPS 1-step and BPS(4) analyses.

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. On the contrary, for 4-step ahead forecasts (Fig. 16), we see forecast errors converging between the two, explaining the increase in concordance as all models performed equally poorly.

5 Additional 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. It will be of interest 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.

The US 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 US macroeconomic data, BPS has the potential to be applied to other fields and data where multiple forecasts, whether from forecasters or models, are available. This includes financial data, such as stocks, indexes, and bonds, business data, such as product demand and earnings, meteorological data, and risk forecasts, including seismic and environmental risk. There are a number of methodological extensions that are needed for further investigation and among these, we note, are multivariate synthesis, non-normal forecasts and discrete data, and missing or incomplete/partial forecasts. 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 of Japan), 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.

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.

Bates, J. M., Granger, C. W. J., 1969. The combination of forecasts. Operational Research Quarterly 20, 451–468.

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.

Clemen, R. T., 1989. Combining forecasts: A review and annotated bibliography. International Journal of Forecasting 5, 559–583.

Clemen, R. T., Winkler, R. L., 1999. Combining probability distributions from experts in risk analysis. Risk Analysis 19, 187–203.

Clemen, R. T., Winkler, R. L., 2007. Aggregating probability distributions. In: W. Edwards, R. M., von Winterfeldt, D. (Eds.), Advances in Decision Analysis: From Foundations to Applications. Cambridge University Press, Ch. 9, pp. 154–176.

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.

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.

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., 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.

Pettenuzzo, D., Ravazzolo, F., 2016. Optimal portfolio choice under decision-based model combinations. Journal of Applied Econometrics -, –, forthcoming.

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.

Timmermann, A., 2004. Forecast combinations. In: Elliott, G., Granger, C. W. J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Vol. 1. North Holland, Ch. 4, pp. 135–196.

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., 2013. Bayesian dynamic modelling. In: Damien, P., Dellaportes, P., Polson, N. G., Stephens, D. A. (Eds.), Bayesian Theory and Applications. Clarendon: Oxford University Press, Ch. 8, pp. 145–166.

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.

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.

Bayesian Predictive Synthesis

Kenichiro McAlinn & Mike West

Tables and Figures

	$\text{MSFE}_{1:T}$		$\text{LPDR}_{1:T}$		
	1-step	4-step	1-step	4-step	BPS(4)
M1	0.0634	0.4227	-13.84	71.43	-94.56
M2	0.0598	0.4156	-8.55	68.16	-97.82
M3	0.0616	0.4208	-9.06	60.08	-105.90
M4	0.0811	0.4880	-22.71	67.46	-98.53
BMA	0.0617	0.4882	-9.00	65.65	-100.33
LinP	0.0575	0.4275	-8.84	85.50	-80.48
LogP	0.0579	0.4275	-7.86	68.23	-97.75
BPS	0.0512	0.4001			
BPS(4)	-	0.3686			

Table 1: US inflation rate forecasting 1990/Q1-2014/Q4: Forecast evaluations for quarterly US inflation over the 25 years 1990/Q1-2014/Q4, comparing mean squared forecast errors and log predictive density ratios for this $T = 100$ quarters.

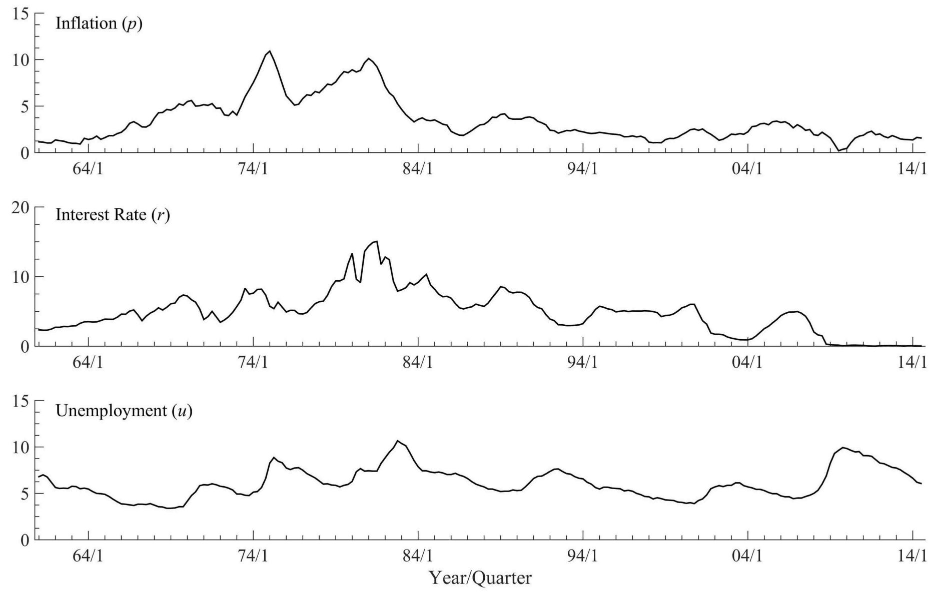

Figure 1: US inflation rate forecasting 1990/Q1-2014/Q4: US macroeconomic time series (indices $\times 100$ for $\%$ basis): annual inflation rate (p), short-term nominal interest rate (r), and unemployment rate (u).

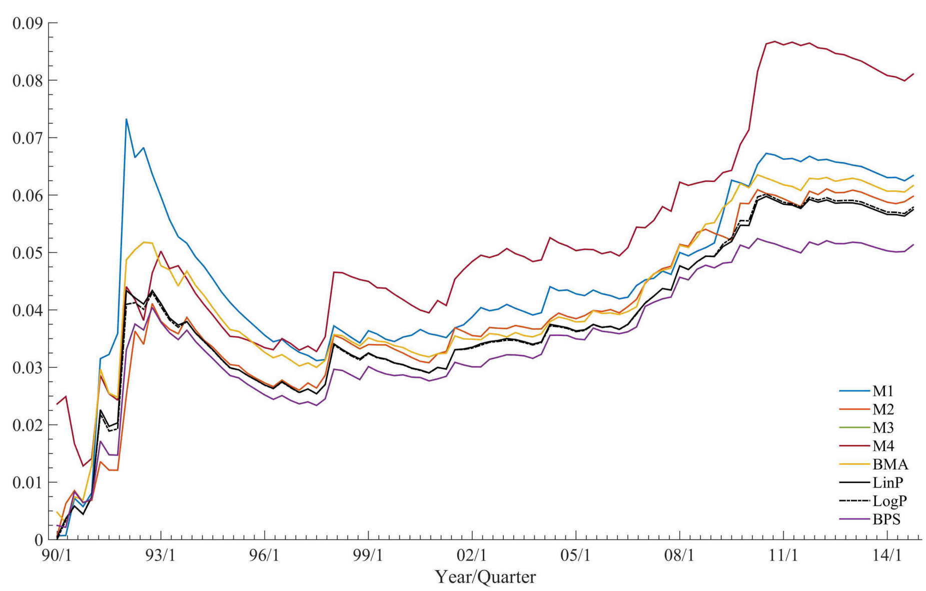

Figure 2: US inflation rate forecasting 1990/Q1-2014/Q4: Mean squared 1-step ahead forecast errors $\text{MSFE}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}100$ quarters.

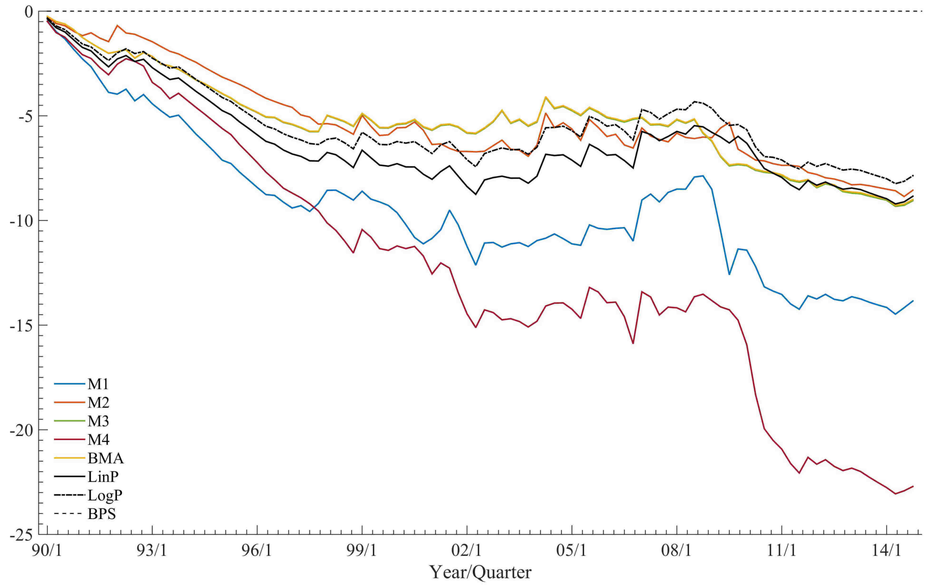

Figure 3: US inflation rate forecasting 1990/Q1-2014/Q4: 1-step ahead log predictive density ratios $\text{LPDR}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}100$ quarters. The baseline at 0 over all t corresponds to the standard BPS model.

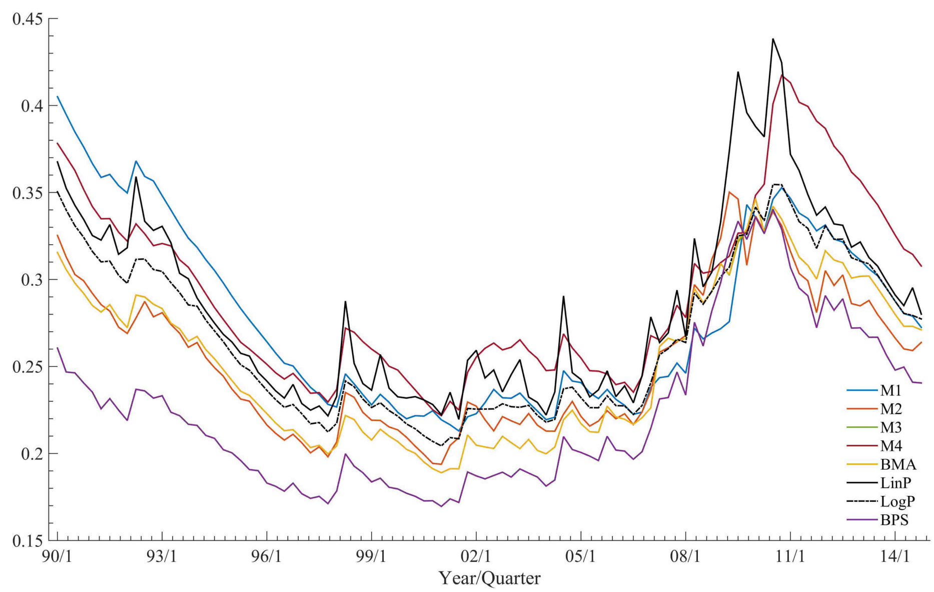

Figure 4: US inflation rate forecasting 1990/Q1-2014/Q4: 1-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}100$ quarters.

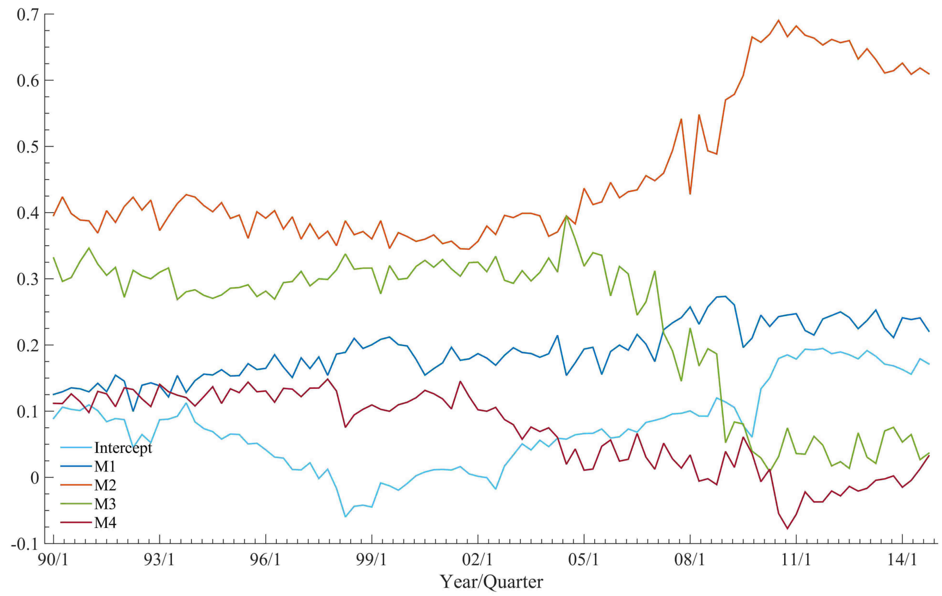

Figure 5: US inflation rate forecasting 1990/Q1-2014/Q4: On-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}100$ quarters.

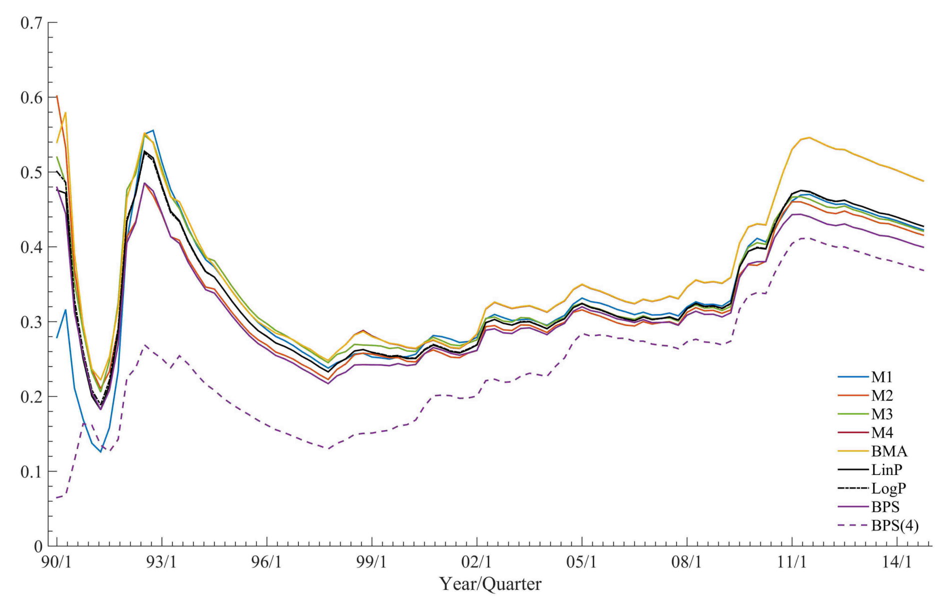

Figure 6: US inflation rate forecasting 1990/Q1-2014/Q4: Mean squared 4-step ahead forecast errors $\mathrm{MSFE}_{1:t}(4)$ sequentially revised at each of the $t = 1{:}100$ quarters.

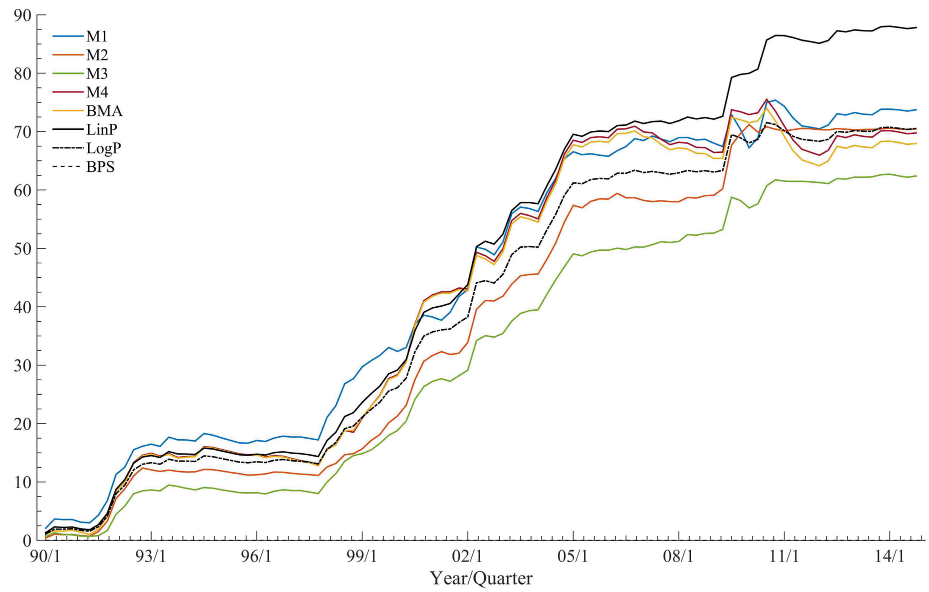

Figure 7: US inflation rate forecasting 1990/Q1-2014/Q4: 4-step ahead log predictive density ratios $\mathrm{LPDR}_{1:t}(4)$ sequentially revised at each of the $t = 1{:}100$ quarters using direct projection from the 1-step ahead BPS model (baseline at 0 over time).

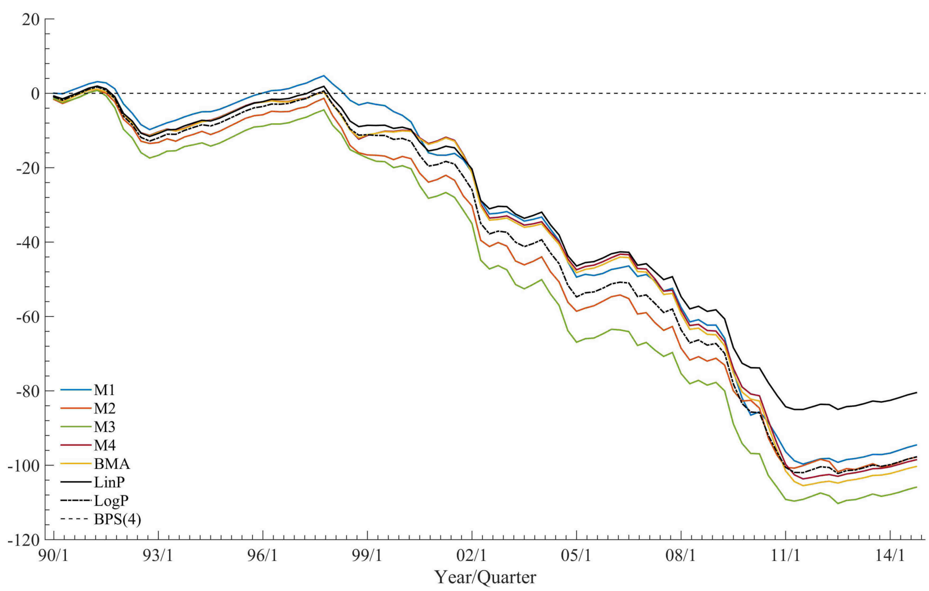

Figure 8: US inflation rate forecasting 1990/Q1-2014/Q4: 4-step ahead log predictive density ratios, $\mathrm{LPDR}_{1:t}(4)$ sequentially revised at each of the $t = 1{:}100$ quarters using the 4-step ahead customized BPS(4) model (baseline at 0 over time).

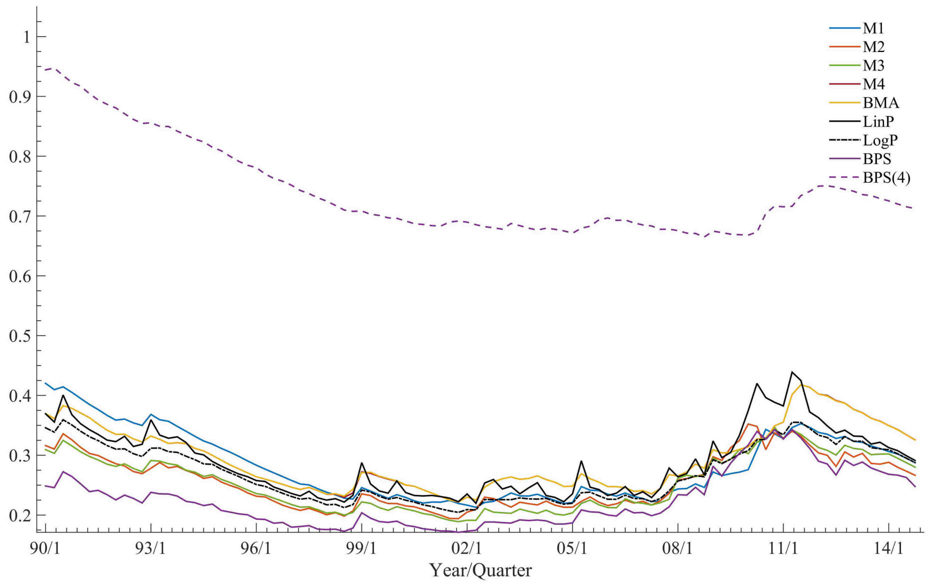

Figure 9: US inflation rate forecasting 1990/Q1-2014/Q4: 4-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}100$ quarters.

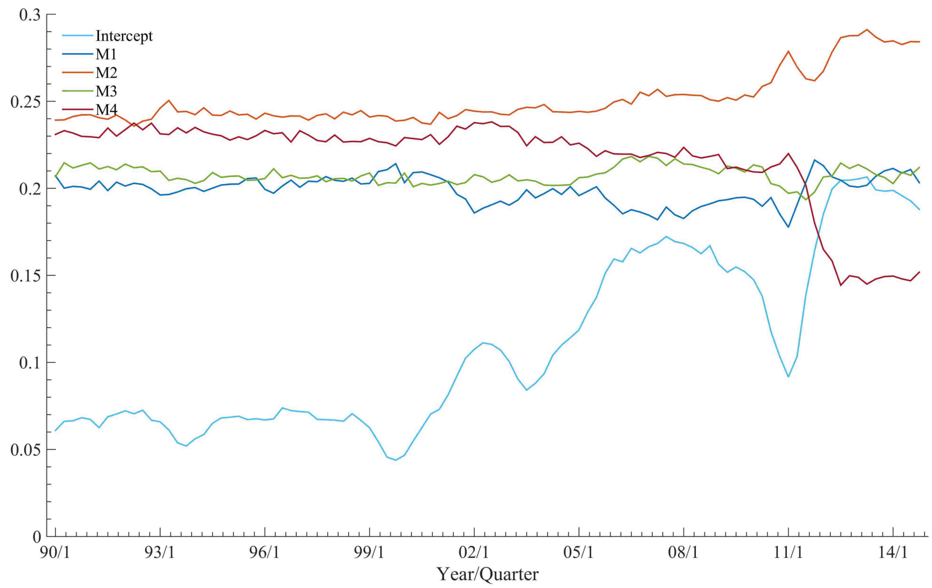

Figure 10: US inflation rate forecasting 1990/Q1-2014/Q4: On-line posterior means of BPS(4) model coefficients sequentially computed at each of the $t = 1{:}100$ quarters.

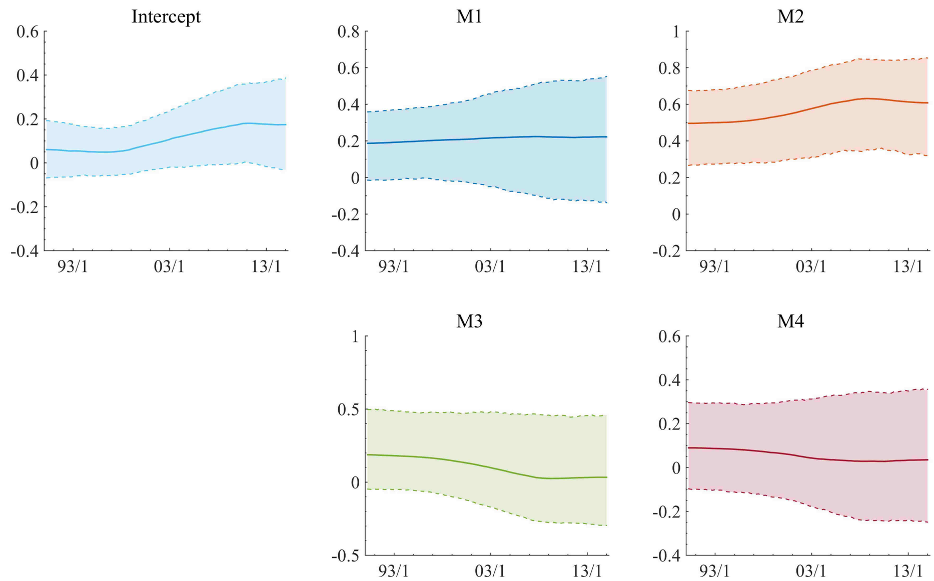

Figure 11: US inflation rate forecasting 1990/Q1-2014/Q4: Retrospective posterior trajectories of the BPS model coefficients based on data from the full $t = 1{:}100$ quarters. Posterior means (solid) and 95% credible intervals (shaded).

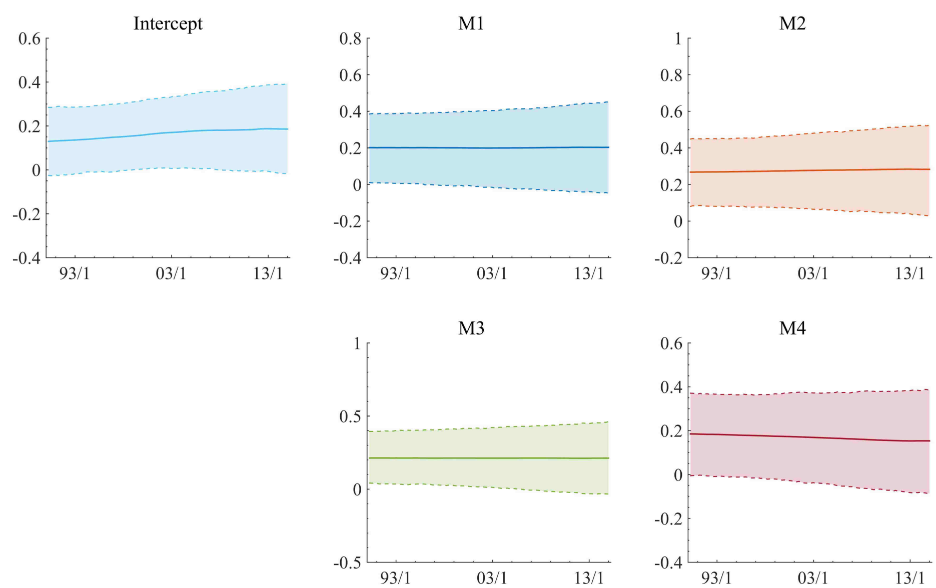

Figure 12: US inflation rate forecasting 1990/Q1-2014/Q4: Retrospective posterior trajectories of the BPS(4) model coefficients based on data from the full $t = 1{:}100$ quarters. Posterior means (solid) and 95% credible intervals (shaded).

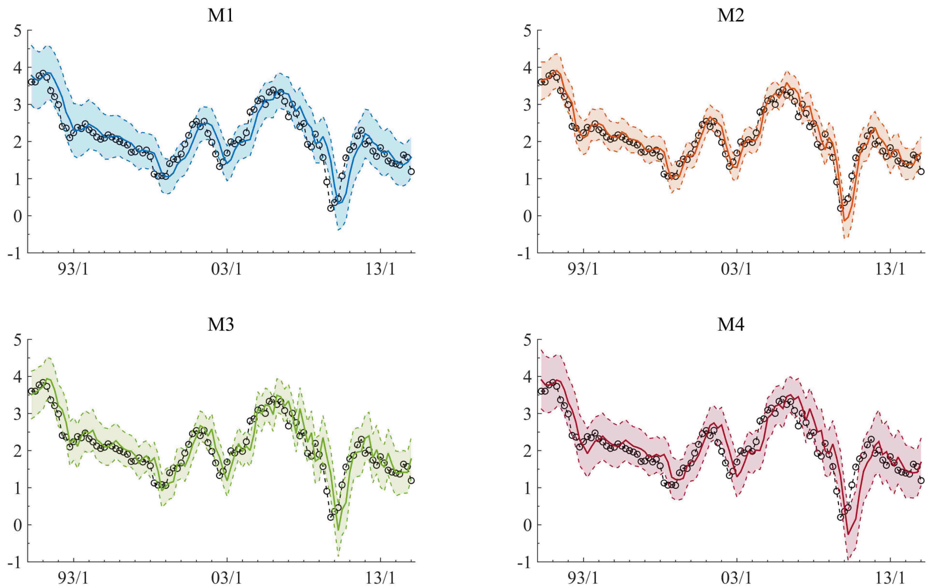

Figure 13: US inflation rate forecasting 1990/Q1-2014/Q4: BPS model-based posterior trajectories of the latent agent states x_{tj} for $j = 1{:}4$ over the $t = 1{:}100$ quarters. Posterior means (solid) and 95% credible intervals (shaded) from the MCMC analysis, with data $y_t \equiv p_t$ (circles).

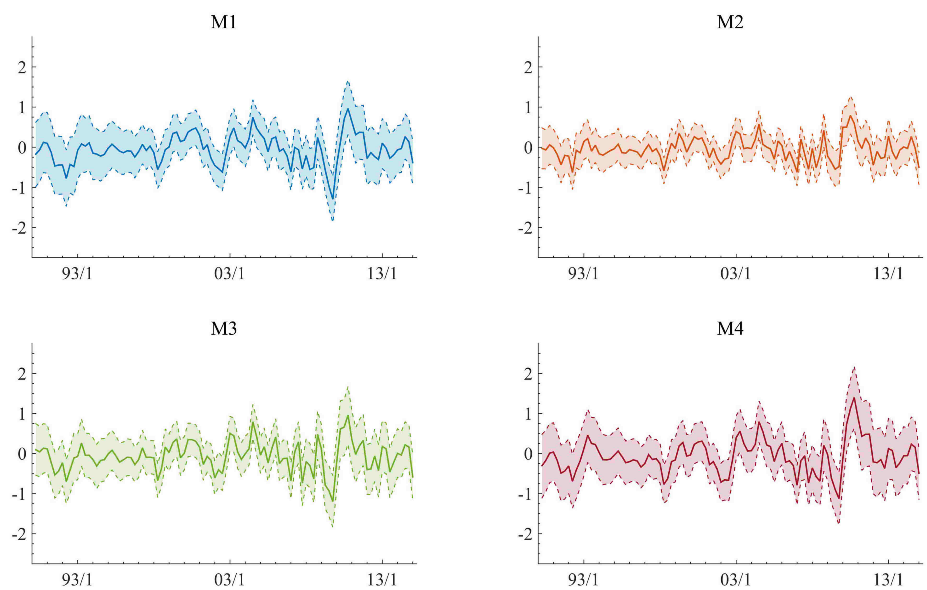

Figure 14: US inflation rate forecasting 1990/Q1-2014/Q4: BPS model-based posterior trajectories of the error in latent agents states $y_t - x_{tj}$ for $j = 1{:}4$ over the $t = 1{:}100$ quarters. Posterior means (solid) and 95% credible intervals (shaded) from the MCMC analysis.

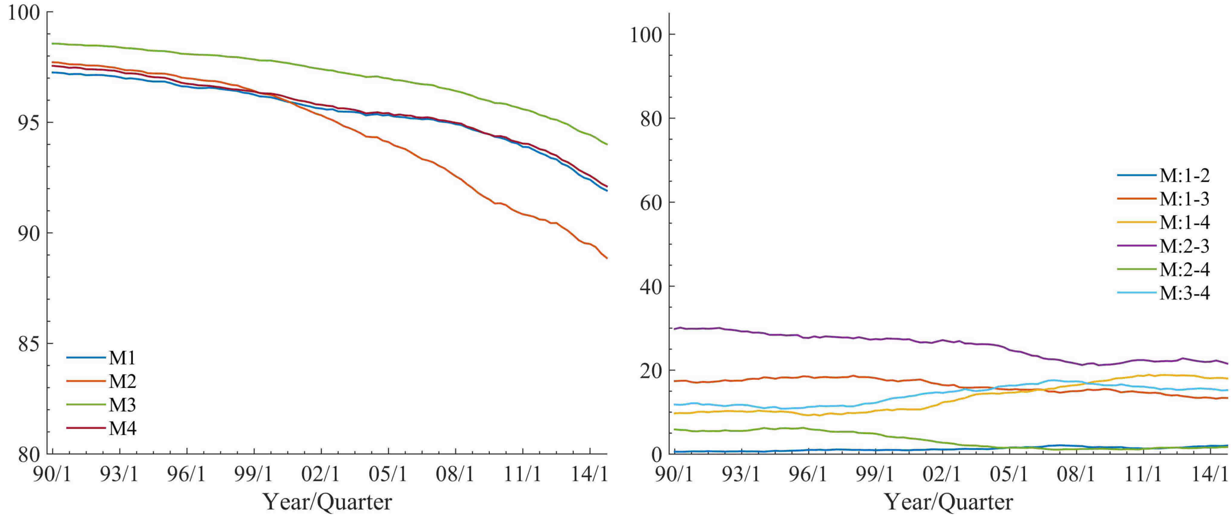

Figure 15: US inflation rate forecasting 1990/Q1-2014/Q4: BPS model-based trajectories of 1-step ahead MC-empirical R^2 (left) and paired MC-empirical R^2 (right) in the posterior for the latent agent states x_{jt} for $j = 1{:}4$ over the $t = 1{:}100$ quarters.

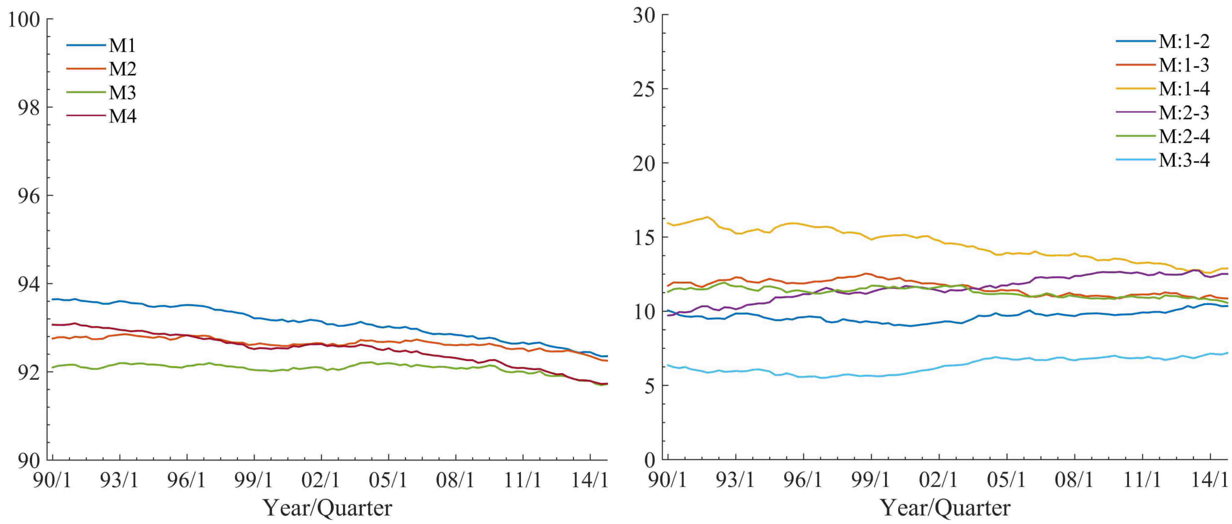

Figure 16: US inflation rate forecasting 1990/Q1-2014/Q4: BPS(4) model-based trajectories of 4-step ahead MC-empirical R^2 (left) and paired MC-empirical R^2 (right) in the posterior for the latent agent states x_{jt} for $j = 1{:}4$ over the $t = 1{:}100$ quarters.