CAMP Working Paper Series
No 1/2019

# Multivariate Bayesian Predictive Synthesis in Macroeconomic Forecasting

Kenichiro McAlinn, Knut Are Aastveit, Jouchi Nakajima & Mike West

**BI** NORWEGIAN
BUSINESS SCHOOL

# Multivariate Bayesian Predictive Synthesis
# in Macroeconomic Forecasting[*]

Kenichiro McAlinn[1,2†]  Knut Are Aastveit[3,4], Jouchi Nakajima[5], & Mike West[2]

[1]Booth School of Business, University of Chicago

[2]Department of Statistical Science, Duke University

[3]Norges Bank

[4]BI Norwegian Business School

[5]Bank for International Settlements

January 16, 2019

## Abstract

We present new methodology and a case study in use of a class of Bayesian predictive synthesis (BPS) models for multivariate time series forecasting. This extends the foundational BPS framework to the multivariate setting, with detailed application in the topical and challenging context of multi-step macroeconomic forecasting in a monetary policy setting. BPS evaluates– sequentially and adaptively over time– varying forecast biases and facets of miscalibration of individual forecast densities for multiple time series, and– critically– their time-varying interdependencies. We define BPS methodology for a new class of dynamic multivariate latent factor models implied by BPS theory. Structured dynamic latent factor BPS is here motivated by the application context– sequential forecasting of multiple US macroeconomic time series with forecasts generated from several traditional econometric time series models. The case study highlights the potential of BPS to improve of forecasts of multiple series at multiple forecast horizons, and its use in learning dynamic relationships among forecasting models or agents.

[†]Corresponding author.

*E-mail*: kenichiro.mcalinn@chicagobooth.edu, knut-are.aastveit@norges-bank.no, jouchi.nakajima@bis.org and mike.west@duke.edu

# 1 Introduction

In dynamic socio-economic systems generating multiple, inter-related time series, the dependencies across series can critically impact on decisions, policies, and their outcomes. In particular, studies of inter-dependencies among macroeconomic time series can provide fundamental insights into the state of the economy to advise economic policy making. The interest is to use such relationships to improve forecasts over multiple horizons, to help to guide the policy decisions and understand their impact. Central banks set national target interest rates based on (implicit or explicit) utility/loss considerations that weigh future outcomes of inflation and measures of the real economy and labor market. Understanding the (time-varying) dependencies of these measures– especially the dynamics over multiple horizons– is simply critical. Driven by this, multivariate models, ranging from vector autoregressive models (VAR) to dynamic stochastic generalized equilibrium models (DSGE), have been developed and used by researchers and policy makers. A huge literature reflects the critical nature of the field, from the early works of, for example, Sims (1993), Stock and Watson (1996), and Sims and Zha (1998), to more recent advances in dynamic Bayesian models in Cogley and Sargent (2005), Primiceri (2005), Benati and Surico (2008), Koop et al. (2009), Koop et al. (2010), Nakajima (2011), Nakajima and West (2013a,b), and Zhou et al. (2014), among others.

Concerned with accurate and useful forecasts, policy makers routinely rely on multiple sources, employing multiple models, forecasters, and economists, to produce forecasts. To ensure appropriate normative decision making as well as reflecting increased uncertainty into the future, it has become popular, particularly for central banks, to provide probabilistic (density) forecasts. For example, forecasts reported in the monetary policy reports of the Bank of England, Norges Bank, Swedish Riksbank, and recently also for the Federal Reserve Bank, have reflected this change. To respond to this increased usage of density forecasts, there has been a recent resurgence in interest in forecast comparison, calibration, and combination of density forecasts in macroeconomics, econometrics, and statistics. These new developments range from combining predictive densities using weighted linear combinations of prediction models, evaluated using various scoring rules (e.g. Hall and Mitchell 2007; Amisano and Giacomini 2007; Jore et al. 2010; Hoogerheide et al. 2010; Kascha and Ravazzolo 2010; Geweke and Amisano 2011, 2012; Gneiting and Ranjan 2013; Aastveit et al. 2014), to more complex combination approaches that allows for time-varying weights with possibly both learning and model set incompleteness (e.g. Billio et al. 2013; Casarin et al. 2015; Aastveit et al. 2018; Pettenuzzo and Ravazzolo 2016; Del Negro et al. 2016).

The extensive literature on forecast combination has, for the most part, focused on forecasting a single series. This is true from the seminal paper by Bates and Granger (1969) to applications in 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), as it is to the recent developments reported above. In contrast, the literature on multivariate forecasting is dominated by traditional statistical model comparison and variable selection (e.g. Chan et al. 2012; Korobilis 2013; Nakajima and West 2013a). Little attention has yet been given to forecast comparison, calibration, and combination in the context of forecasting multiple series. It is relevant to note that Satopää et al. (2014) consider

probability predictions aggregated over multiple series, though forecasts are probabilities and not distributions, and the probabilities are continuously updated until the outcome is realized– a very different context from ours here. A few authors (e.g. Andersson and Karlsson 2008; Amendola and Storti 2015; Amisano and Geweke 2017) recognize the lack of multivariate approaches but restrict attention to direct extensions of univariate methods and with models combined linearly using one metric for overall performance. This is limiting in several ways: in ignoring inter-dependencies among series that can be detrimental in informing decisions, in ignoring the reality that some models might be good at forecasting one series but poor in another, and for the fact that some or all models maybe be poor overall. Partly reflecting the lack of formal statistical frameworks for holistic multivariate forecast model assessment and combination, economic policy makers use "ad hoc" strategies, which either rely on the policy maker's "favorite" model, or ignore inter-dependencies all together. The need for coherent methodology that gives policy makers flexibility in incorporating multivariate density forecasts from multiple sources cannot be understated.

The developments of the current paper address the above issues, challenges, and needs. The new models, methodology, and motivating case study presented builds on the conceptual basis and underlying theory of dynamic *Bayesian Predictive Synthesis* (BPS) introduced in a univariate forecasting setting (McAlinn and West 2017). BPS is a coherent Bayesian framework for evaluation, calibration, comparison, and context- and data-informed combination of multiple forecast densities. The approach applies whether the forecast densities arise from sets of models, forecasters, agencies, or institutions. As detailed in McAlinn and West (2017), the framework includes, as special cases, a range of existing univariate density forecasts combination methods. Our multivariate framework here explicitly allows modeling and estimation of time-varying forecast biases and facets of miscalibration of individual forecast densities, time-varying inter-dependencies among models or forecasters over multiple series, and addresses the above noted challenges in multivariate settings.

Methodological details are developed for one specific subclass of BPS models– a new, flexible class of dynamic latent factor models with seemingly-unrelated regression structure (DFSUR models). While motivated by the case study context and representing just one specific class of multivariate structures that might be used in BPS, these DFSUR models are also of broader interest in defining new methodology for multivariate time series. In our setting, each individual model generating multivariate forecast densities is linked to a set of multivariate dynamic latent factor processes– the relationships across each set of latent factors are then a key focus in understanding and leading to forecast combination that address inter-dependencies.

The case study uses multivariate BPS methodology for forecasting and evaluation of several macroeconomic time series relevant in economic policy decision making contexts. First, forecasting six U.S macroeconomic time series, we show that forecasts from our multivariate BPS framework outperform forecasts from other commonly used aggregation methods, including standard Bayesian model averaging and optimal prediction pools (Geweke and Amisano 2011; Amisano and Geweke 2017). BPS also– naturally and inherent in the approach– outperforms forecasts from each of the individual models (agents) synthesized. Forecast improvements arise from several sources: the exploitation of dynamics in agent-specific biases; modeling of inter-dependencies across agents; the fact that BPS defines an encompassing model that is able to define baseline forecasts even when/if

agent forecast are all poor; and the use of the new class of dynamic BPS factor models that allow for and enable estimation of time-varying parameters related to these various features.

Second, multivariate BPS allows us to both account for and to analyze time-varying inter-dependencies among different agents and different series. We demonstrate that this is a crucial advance in aggregation methodology. For example, we show that relative forecast improvements from BPS appear to increase during periods where the dependencies among agents and series change, exemplified in periods of crisis and structural change in the macoeconomic setting. We further discuss and exemplify who how these dynamic patterns of inter-dependencies can provide economic insights relevant to policy decision makers.

Third, to illustrate how BPS can be directly relevant for economic policy and decision making, we conduct a scenario analysis using conditional forecasts. From a policy/decision making stand-point, a critical interest is that of predicting future outcomes of key economic series conditional on specific "current" policy actions. Our example emulates the actual decisions made by the Federal Reserve Bank in the late 2008 stages of the global financial crisis; we find that the individual agent forecasts do not produce reasonable conditional forecasts, while the BPS conditional forecasts significantly improve forecast accuracy and relevance for advising decision making.

Section 2 summarizes the BPS framework and implications in terms of the broad class of implied theoretical models for dynamic multivariate problems, introduces the novel class of DFSUR models customized to the multivariate BPS context, and summarizes key aspects of Bayesian model fitting, sequential analysis and forecasting. The case study concerning U.S macroeconomic time series analysis and forecasting, illustrating and highlighting the role and benefits of the new framework using the class of DFSUR models in the BPS context, is developed in Section 3. Additional comments on aspects of the work as well as broader potential impact and open questions appear in the concluding Section 4. An appendix of detailed supplementary material summarizes Bayesian computational methods (MCMC) for fitting and using DFSUR models in the BPS context, and contains more extensive graphical and tabular summaries of results of the multiple BPS forecasting analyses from the case study.

## 2    Dynamic Multivariate BPS

The new developments forming the methodological core of this paper adapt and extend the basic BPS framework of McAlinn and West (2017) to multivariate density forecast synthesis with practical decision goals in mind. McAlinn and West (2017) defined formal, coherent methodology for integrating density forecasts from multiple, potentially competing statistical model– or forecasters, or institutions– in a univariate time series setting. The dynamic BPS approach there has a foundation in coherent Bayesian reasoning with predictive and decision analytic goals, based on historical developments in assessing and combining subjective probabilities (Lindley et al. 1979; West 1984; Genest and Schervish 1985; West 1988; West and Crosse 1992; West 1992; Dawid et al. 1995; French 2011). Drawing on key theoretical results from that Bayesian "agent/expert opinion analysis" literature, McAlinn and West (2017) define a class of time-varying parameter, latent factor models in which each of the univariate latent factors relates to one of the set of models or forecast-

ers generating predictions. The models are developed methodologically and shown to have promise in understanding relationships among forecasting models, their biases and inter-dependencies over time, and can improve short and medium term forecasting for univariate time series.

We now develop the new, multivariate extension of dynamic BPS, beginning with a brief summary of the key background theory free from the time series context.

**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. 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 $y_{1:t} = \{y_1, \ldots, y_t\}$.

## 2.1 BPS Background

Consider forecasting a $q{\times}1-$ dimensional vector of outcomes $\boldsymbol{y}$. Outcomes are typically real-valued, as is the case in our applications below, though the foundational theory is general. A Bayesian decision maker $\mathcal{D}$ is to receive forecast distributions for $\boldsymbol{y}$ from each of $J$ *agents*; in our application, the agents are different Bayesian time series models, while in other contexts they may include professional forecasters, or forecasting agencies, etc., labelled $\mathcal{A}_j$, $(j = 1{:}J)$. Then $\mathcal{D}$ aims to incorporate the information provided by the set of agent forecast distributions in her thinking and forecasting $\boldsymbol{y}$ and any resulting decisions. Agent $\mathcal{A}_j$ provides a probability density function $h_j(\boldsymbol{y})$. These forecast densities represent the individual inferences from the agents, and define the information set $\mathcal{H} = \{h_1(\cdot), \ldots, h_J(\cdot)\}$ now available to $\mathcal{D}$. Formal subjective Bayesian analysis dictates that, $\mathcal{D}$ will then use the information set $\mathcal{H}$ to predict $\boldsymbol{y}$ using the implied posterior $p(\boldsymbol{y}|\mathcal{H})$ from a full Bayesian prior-to-posterior analysis.

To obtain a full Bayesian prior-to-posterior analysis, West (1992) extended prior theory (Genest and Schervish 1985; West and Crosse 1992) to show that there is a subset of all Bayesian models in which $\mathcal{D}$'s posterior has the mathematical form

$$p(\boldsymbol{y}|\mathcal{H}) = \int_{\boldsymbol{X}} \alpha(\boldsymbol{y}|\boldsymbol{X}) \prod_{j=1:J} h_j(\boldsymbol{x}_j) d\boldsymbol{x}_j \tag{1}$$

where each $\boldsymbol{x}_j$ is a latent $q{\times}1-$dimensional vector, $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_J]'$ collects these latent vectors in a $J{\times}q-$dimensional matrix, and $\alpha(\boldsymbol{y}|\boldsymbol{X})$ is a conditional p.d.f. for $\boldsymbol{y}$ given $\boldsymbol{X}$. The interpretation is as follows. First, in the subjective view of $\mathcal{D}$ there *must* exist latent factors $\boldsymbol{x}_j$ potentially related to $\boldsymbol{y}$ and such that agent $\mathcal{A}_j$'s forecast density is that of $\boldsymbol{x}_j$. Second, *conditional on learning $\mathcal{H}$*, the $\mathcal{D}$ regards the latent factors as conditionally independent with $\boldsymbol{x}_j \sim h_j(\boldsymbol{x}_j)$. Note that this does not imply that $\mathcal{D}$ regards the forecasts as independent, since under her prior the $\boldsymbol{h}_j(\cdot)$ are uncertain and likely highly inter-dependent, and the key element $\alpha(\boldsymbol{y}|\boldsymbol{X})$ is how $\mathcal{D}$ expresses her views of dependencies. Third, this key element $\alpha(\boldsymbol{y}|\boldsymbol{X})$ is $\mathcal{D}$'s regression model relating the $\boldsymbol{x}_j$ as a collective to the outcome $\boldsymbol{y}$. Refer to $\alpha(\boldsymbol{y}|\boldsymbol{X})$ as the BPS *synthesis function*, and to the $\boldsymbol{x}_j$ as *latent agent states*.

The key eqn. (1) does define the functional form of $\alpha(\boldsymbol{y}|\boldsymbol{X})$. McAlinn and West (2017) show

that, for scalar outcomes $\boldsymbol{y} = y$, many forecast and model combination methods (e.g. Bayesian model averaging; Geweke and Amisano 2011; Kapetanios et al. 2015; Pettenuzzo and Ravazzolo 2016; Aastveit et al. 2018, among others) can be considered as special cases of eqn. (1), realized via different choices of the form of the BPS synthesis function $\alpha(\cdot|\cdot)$. For vector outcomes $\boldsymbol{y}$, eqn. (1) similarly allows flexibility for $\mathcal{D}$ to specify $\alpha(\boldsymbol{y}|\boldsymbol{X})$ to reflect decision goals and incorporate views and historical information about, for example, agent-specific biases, patterns of miscalibration, inter-dependencies among agents, and their relative expertise and expected forecast accuracy. Any specific BPS model will be created by assuming a specific model form for the synthesis p.d.f.

## 2.2 Dynamic Sequential Setting

For a $q-$vector time series $\boldsymbol{y}_t, t = 1, 2, \ldots$, decision maker $\mathcal{D}$ receives forecast densities from each agent sequentially over time. At time $t - 1$, $\mathcal{D}$ receives current forecast densities $\mathcal{H}_t = \{h_{t1}(\boldsymbol{y}_t), \ldots, h_{tJ}(\boldsymbol{y}_t)\}$ from the set of agents and aims to forecast $\boldsymbol{y}_t$. The full information set used by $\mathcal{D}$ at time $t$ is thus $\{\boldsymbol{y}_{1:t-1}, \mathcal{H}_{1:t}\}$. As $\mathcal{D}$ observes more information, her views of the agent biases and calibration characteristics, as well as of inter-dependencies among agents are repeatedly updated. A formal, parametrized Bayesian dynamic model is the vehicle for structuring this sequential learning in a general state-space context. This defines the dynamic BPS framework.

The time series extension of eqn. (1) implies that $\mathcal{D}$ has a time $t - 1$ distribution for $\boldsymbol{y}_t$ as

$$p(\boldsymbol{y}_t|\boldsymbol{\Phi}_t, \boldsymbol{y}_{1:t-1}, \mathcal{H}_{1:t}) \equiv p(\boldsymbol{y}_t|\boldsymbol{\Phi}_t, \mathcal{H}_t) = \int \alpha_t(\boldsymbol{y}_t|\boldsymbol{X}_t, \boldsymbol{\Phi}_t) \prod_{j=1:J} h_{tj}(\boldsymbol{x}_{tj}) d\boldsymbol{x}_{tj} \tag{2}$$

where $\boldsymbol{X}_t = [\boldsymbol{x}_{t1}, \ldots, \boldsymbol{x}_{tJ}]'$ is a $J \times q-$dimensional matrix of latent agent states at time $t$, the conditional p.d.f. $\alpha_t(\boldsymbol{y}_t|\boldsymbol{X}_t, \boldsymbol{\Phi}_t)$ is $\mathcal{D}$'s synthesis p.d.f. for $\boldsymbol{y}_t$ given $\boldsymbol{X}_t$, and involves time-varying parameters $\boldsymbol{\Phi}_t$ for which $\mathcal{D}$ has current beliefs represented in terms of her (time $t - 1$) posterior $p(\boldsymbol{\Phi}_t|\boldsymbol{y}_{1:t-1}, \mathcal{H}_{1:t-1})$.

This general framework defines the $\boldsymbol{x}_{tj}$ as realizations of inherent dynamic latent factors– the *latent agent states* at time $t$– and synthesis is achieved by relating these latent factor processes to the time series $\boldsymbol{y}_t$ via models of the time-varying synthesis function $\alpha_t(\boldsymbol{y}_t|\boldsymbol{X}_t, \boldsymbol{\Phi}_t)$. The foundational theory does not specify this p.d.f., and methodology is based on specific chosen forms. For the multivariate extension of McAlinn and West (2017), we look to a specific class of models that extends the traditional seemingly unrelated regression model (SUR; Zellner 1962) to a dynamic Bayesian framework, as a first approach to defining a computationally accessible yet flexible framework for dynamic multivariate BPS.

## 2.3 Multivariate Latent Factor Dynamic Models

Consider a dynamic multivariate BPS synthesis function

$$\alpha_t(\boldsymbol{y}_t|\boldsymbol{X}_t, \boldsymbol{\Phi}_t) = N(\boldsymbol{y}_t|\boldsymbol{F}_t\boldsymbol{\theta}_t, \boldsymbol{V}_t) \tag{3}$$

with

$$\boldsymbol{F}_t = \begin{pmatrix} 1 & \boldsymbol{f}'_{t1} & 0 & \boldsymbol{0}' & \cdots & \cdots & 0 & \boldsymbol{0}' \\ 0 & \boldsymbol{0}' & 1 & \boldsymbol{f}'_{t2} & & & & \vdots \\ \vdots & & & & \ddots & & & \vdots \\ 0 & \boldsymbol{0}' & \cdots & \cdots & \cdots & \cdots & 1 & \boldsymbol{f}'_{tq} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\theta}_t = \begin{pmatrix} \boldsymbol{\theta}_{t1} \\ \boldsymbol{\theta}_{t2} \\ \vdots \\ \boldsymbol{\theta}_{tq} \end{pmatrix} \tag{4}$$

where for each series $r = 1{:}q$, the $J{\times}1-$vector $\boldsymbol{f}_{tr} = (x_{tr1}, x_{tr2}, ..., x_{trJ})'$ is a realization of the set of $J$ latent agents states for series $r$, and the $(J+1){\times}1-$vector $\boldsymbol{\theta}_{tr} = (1, \theta_{tr1}, \theta_{tr2}, ..., , \theta_{trJ})'$ contains an intercept and coefficients representing time-varying bias/calibration weights of the $J$ latent agent states for series $r$. Note that $\boldsymbol{F}_t$ has $q$ rows and $(J+1)q$ columns, and $\boldsymbol{\theta}_t$ is a column $(J+1)q-$vector. Observation noise is reflected in the– likely volatile– residual $q{\times}q$ variance matrix $\boldsymbol{V}_t$, and the general time-varying parameter of eqn. (2) is set as $\boldsymbol{\Phi}_t = \{\boldsymbol{\theta}_t, \boldsymbol{V}_t\}$.

This defines the first component of a conditionally linear, conditionally normal model– a subclass of multivariate dynamic linear models but with latent factors as predictors. Modeling time evolution of the parameter processes $\boldsymbol{\Phi}_t = (\boldsymbol{\theta}_t, \boldsymbol{V}_t)$ is needed to complete model specification. We do this using the first step into dynamic models, with traditional random walk models to allow for– but not anticipate direction in– stochastic changes over time in both regressions $\boldsymbol{\theta}_t$ and matrix volatilities $\boldsymbol{V}_t$, as is traditional in Bayesian time series literatures; see, for example, West and Harrison (1997) (Chap. 16) and Prado and West (2010) (Chap. 10). Thus we take

$$\boldsymbol{y}_t = \boldsymbol{F}_t\boldsymbol{\theta}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim N(\boldsymbol{0}, \boldsymbol{V}_t), \tag{5a}$$

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N(\boldsymbol{0}, \boldsymbol{W}_t) \tag{5b}$$

where $\boldsymbol{\theta}_t$ evolves in time according to a linear/normal random walk with innovations variance matrix $\boldsymbol{W}_t$ at time $t$, and $\boldsymbol{V}_t$ is the residual variance in predicting $\boldsymbol{y}_t$ based on past information and the set of agent forecast distributions.

Model specification is completed using standard discount methods. As with the univariate DLM, the time-varying intercept and agent coefficients $\boldsymbol{\theta}_t$ follow the random walk evolution of eqn. (5b) where $\boldsymbol{W}_t$ is defined via a standard, single discount factor specification (Prado and West 2010, Chap. 10). The residual variance matrix $\boldsymbol{V}_t$ follows a standard inverse Wishart random walk volatility model, also based on discounting with a second discount factor.

We now have a class of dynamic, multivariate latent factor models in which latent factors are realized as draws from the set of agent densities $h_{tj}(\cdot)$, becoming available to $\mathcal{D}$ at $t-1$ for forecasting $\boldsymbol{y}_t$. Thus, coupled with eqns. (5a,5b), we have the time $t$ *prior* for the latent states– conditional on $\mathcal{H}_{1:t}$, as

$$p(\boldsymbol{X}_t | \boldsymbol{\Phi}_t, \boldsymbol{Y}_{1:t-1}, \mathcal{H}_{1:t}) \equiv p(\boldsymbol{X}_t | \mathcal{H}_t) = \prod_{j=1:J} h_{tj}(\boldsymbol{x}_{tj}) \tag{6}$$

with $\boldsymbol{X}_t, \boldsymbol{X}_s$ conditionally independent for all $t \neq s$. Again it is important to stress that the *conditional* independence of the $\boldsymbol{x}_{tj}$ given the $h_{tj}(\cdot)$ must not be confused with the $\mathcal{D}$'s modeling and estimation of the dependencies among agents. This dependence is central and integral, and is reflected through the effective dynamic parameters $\boldsymbol{\Phi}_t = (\boldsymbol{\theta}_t, \boldsymbol{V}_t)$.

6

## 2.4 Bayesian Analysis and Computation

At any current time $t$, $\mathcal{D}$ has historical information $\{\boldsymbol{y}_{1:t}, \mathcal{H}_{1:t}\}$ and the history of the BPS analysis up until that point. This will have defined inferences on past latent agent states $\boldsymbol{X}_*$ and the dynamic BPS model parameters $\boldsymbol{\Phi}_* = (\boldsymbol{\theta}_*, \boldsymbol{V}_*)$. The former, importantly, provides insight into the dependencies, biases, and other characteristics pertaining to $\boldsymbol{y}_{1:t}$, among agents and individual agents. Posterior summaries for $\boldsymbol{X}_t$ over time inform on this– a key feature of BPS. This inference is topical, as issues of herding (overlap and redundancies) among groups of agents (either models or individuals) is of practical importance, and understanding how these characteristics change over time and across series is key.

Posterior analysis is enabled by Markov chain Monte Carlo (MCMC) methods, followed by forecasting from time $t$ onward utilizing theoretical and simulation-based extrapolation of the model. $\mathcal{D}$ is interested in the inference on the full set of past latent agent states and dynamic parameters $\{\boldsymbol{X}_{1:t}, \boldsymbol{\Phi}_{1:t}\}$, as well as forward filtering to update posteriors for current values $\{\boldsymbol{X}_t, \boldsymbol{\Phi}_t\}$. Posterior MCMC-based computation uses nowadays traditional methods, and extends the MCMC method used in McAlinn and West (2017) for the univariate case with several modifications.

**Posterior Computations via MCMC.** At a given current time $t$, the multivariate dynamic latent factor model structure of eqns. (5a,5b,6) leads easily to a three-component block Gibbs sampler for the latent agent states, dynamic coefficient parameters, and dynamic volatility parameters. The components are iteratively resampled from the three conditional posteriors noted below, initialized given agent states drawn independently from priors $h_*(*)$.

First, conditional on the agent states and residual volatility, the MCMC step draws new dynamic coefficient parameters from $p(\boldsymbol{\theta}_{1:t}|\boldsymbol{X}_{1:t}, \boldsymbol{V}_{1:t}, \boldsymbol{y}_{1:t})$. This is the full (normal) posterior for the sequence of states in the implied conditional multivariate DLM, and is efficiently sampled using an extension of the traditional forward filtering, backward sampling (FFBS) algorithm (e.g. Prado and West 2010, chap 10).

Second, the MCMC step draws new dynamic volatility matrices $\boldsymbol{V}_t$ from the full joint conditional posterior $p(\boldsymbol{V}_{1:t}|\boldsymbol{X}_{1:t}, \boldsymbol{\theta}_{1:t}, \boldsymbol{y}_{1:t})$– conditional on the agent states and dynamic coefficient parameters. This employs the standard FFBS algorithm for inverse Wishart discount volatility models (Prado and West 2010, chap. 10)

Third, conditional on values of dynamic parameters $\boldsymbol{\Phi}_{1:t} = (\boldsymbol{\theta}_{1:t}, \boldsymbol{V}_{1:t})$, the MCMC draws new agent states from $p(\boldsymbol{X}_{1:t}|\boldsymbol{\Phi}_{1:t}, \boldsymbol{y}_{1:t}, \mathcal{H}_{1:t})$. As with the univariate case, the $\boldsymbol{X}_t$ are conditionally independent over time $t$ in this conditional distribution, with time $t$ conditionals $p(\boldsymbol{X}_t|\boldsymbol{\Phi}_t, \boldsymbol{y}_t, \mathcal{H}_t) \propto N(\boldsymbol{y}_t|\boldsymbol{F}_t\boldsymbol{\theta}_t, \boldsymbol{V}_t) \prod_{j=1:J} h_{tj}(\boldsymbol{x}_{tj})$. In cases when all of the agents' forecasts are multivariate normal, the posterior is a multivariate normal that is trivially sampled using the properties of conditional normal. For a more central and practically important case of forecasts being multivariate T distribution, 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 multivariate normals for each $\boldsymbol{X}_t$ coupled with conditional inverse gammas for those scales. In other cases, augmenting the MCMC utilizing Metropolis-Hastings simulator or an augmentation can be used. More discussion of these algorithmic details is given in Appendix A.

**Forecasting 1-Step Ahead.** At time $t$ we forecast 1-step ahead by generating "synthetic futures" from the BPS model, as follows. First, for each sampled $\boldsymbol{\Phi}_t$ from the posterior MCMC above, draw $\boldsymbol{V}_{t+1}$ from its discount volatility evolution model, and then $\boldsymbol{\theta}_{t+1}$ conditional on $\boldsymbol{\theta}_t, \boldsymbol{V}_{t+1}$ from the evolution model eqn. (5b)– this gives a draw $\boldsymbol{\Phi}_{t+1} = \{\boldsymbol{\theta}_{t+1}, \boldsymbol{V}_{t+1}\}$ from $p(\boldsymbol{\Phi}_{t+1}|\boldsymbol{y}_{1:t}, \mathcal{H}_{1:t})$. Second, draw $\boldsymbol{X}_{t+1}$ via independent sampling of the $h_{t+1,j}(\boldsymbol{x}_{t+1,j})$, $(j = 1{:}J)$. Third, bring these samples together and draw a synthetic 1-step outcome $\boldsymbol{y}_{t+1}$ from the conditional normal of eqn. (5a) given these sampled parameters and agent states. Repeating this generates a random Monte Carlo sample from the 1-step ahead forecast distribution for time $t + 1$.

## 2.5 Multi-Step Ahead Forecasting

In many applications involving multivariate analysis, long term forecasting and analysis is often of equal or greater importance than the basic 1-step ahead horizon. For example, in terms of economic policy and macroeconomic time series– based on traditional monthly or quarterly data– the most important horizons of interest are 1-3 years ahead. This is especially true when dealing with monthly data, as knowing a month ahead has very little utility compared to understanding the long term dynamics and structure over multiple years. Thus, economic policy makers advise policy decisions based on inputs from their own forecast models, judgemental inputs, views of other economists, and forecast over the next year or 2-3 years. However, forecast difficulty increases as the horizon increases, especially when models are calibrated on the short-term basis. Traditional statistical evaluation of time series models is inherently based on 1-step ahead forecasts, raising interest in developing BPS to address longer-term forecasting goals.

BPS provides two methods for multi-step ahead forecasting, as laid out in McAlinn and West (2017). The first method is direct sequential projection of $\boldsymbol{\Phi}_{t+1}, \boldsymbol{\Phi}_{t+2}, \dots, \boldsymbol{\Phi}_{t+k}$, updating the parameters over time until it reaches $t + k$ ($k$ being the horizon of interest) and plugging in the relevant forecasts, $\boldsymbol{X}_{t+k}$, sampled from $h_{t,1:J}(\boldsymbol{x}_{t+k})$. The second is denoted by BPS($k$), referring to applying the BPS model to synthesise $k$-step ahead forecasts directly, as expanded upon below. BPS($k$) often outperforms direct projection in terms of forecast accuracy. This is sensible, as we can expect some agents to perform differently (relative to other agents) for different forecast horizons. In the context of macroeconomic forecasts, we might observe that economists, who rely on qualitative information and policy experience, to outperform purely quantitative models on a long term basis and thus calibrating the forecasts on their short term predictive ability– for which quantitative models often are superior– can be problematic.

**BPS($k$) for customized multi-step forecasting.** BPS provides $\mathcal{D}$ a flexible strategy to focus on the horizon $k$ of interest as it is customizable to the forecasting goals. This involves a trivial modification of methodology in Section 2 in which the model at time $t - 1$ for predicting $\boldsymbol{y}_t$ changes as follows. For a *specific forecast horizon* $k > 1$, replace $h_{t-k,j}(\boldsymbol{x}_{tj})$ with $h_{tj}(\boldsymbol{x}_{tj})$ so BPS is calibrated using the forecasts from $t - k$. Doing this results in dynamic model parameters $\{\boldsymbol{\theta}_t, \boldsymbol{V}_t\}$ to be explicitly geared to the $k$-step horizon, calibrating and "tuning" to the horizon $k$ of interest. Forecasting, then, simply follows the model extrapolation via simulation as in Section 2.4.

8

# 3 Case Study in US Macroeconomic Forecasting

## 3.1 Data, Forecasting Models and Implementation

**Time Series Data.** We analyze monthly US macroeconomic data, focusing on forecasting six macroeconomic time series with 1-, 12-, and 24-month ahead interests. The study involves the following monthly macro series: annual inflation rate ($p$), wage ($w$), unemployment rate ($u$), consumption ($c$), investment ($i$), and short-term nominal interest rate ($r$) in the US economy from 1986/1 to 2015/12, a context of topical interest (Cogley and Sargent 2005; Primiceri 2005; Koop et al. 2009; Nakajima and West 2013a). The inflation rate is the consumer price index for all urban consumers: all items less food and energy, not seasonally adjusted, wage is the average hourly earnings of production and nonsupervisory employees: total private, not seasonally adjusted, the unemployment rate is the civilian unemployment rate, seasonally adjusted, consumption is the personal consumption expenditures, seasonally adjusted annual rate, investment is the ISM manufacturing: new orders index, and the interest rate is the effective federal funds rate, not seasonally adjusted; the first four being annual changes, investment being monthly changes, and monthly interest rates. Fig. 1 shows the data for the six series over the time span considered. We focus on forecasting the six series, with an emphasis on inflation, using past values of the six series as candidate predictors underlying a set of five time series models– the $J = 5$ agents– to be evaluated, calibrated, and synthesized.

During the period of analysis, the sub-prime mortgage crisis and great recession of the late 2000s warrant special attention. This period involved a series of significant, unique shocks to the US economy, so any analysis is challenged in terms of predictive ability in short and longer terms. For any combination strategy to be effective and useful, its predictive performance must be robust under these conditions. Additionally, due to the structural changes in the overall economy (e.g. Aastveit et al. 2017), there is also interest in understanding changes in the inter-dependencies among series over the crisis periods. On this goal, multivariate analysis offers opportunity for improved understanding that simple univariate analyses just cannot.

In our BPS($k$) analyses, for $k = 12$ we take investment as the cumulative value of the previous year of monthly differences, since investment measures monthly difference and forecasting the change from 11th to 12th month is of little relevance to the policy maker. Similarly, for $k = 24$, investment is the cumulative value of the 24 months of monthly differences. Additionally, inflation, wage, unemployment, and consumption are defined as changes from the current time. Agents will produce forecasts according to the target value by either summing the forecasts to the target horizon, or summing certain periods within the horizon. In this way, the target forecast is directly in line with what policy makers are interested in and focus on for decision making.

**Agent Models and BPS Specification.** For the $J = 5$ agents we use time varying parameter vector autoregressive (TVP-VAR) models that cover multiple dynamic structures utilized in the literature (Cogley and Sargent 2005; Primiceri 2005; Koop et al. 2009; Nakajima and West 2013a) and in practice. Labelled as M*, the agent models are: M1- VAR(1); M2- VAR(12); M3- VAR(3); M4- VAR(1:3:9); M5- VAR(1:6:12). The numbers in parentheses are the lags and the number be-

tween colons represent intervals (e.g. 1:3:9 uses lags of 1, 3, 6, and 9). Each M* is a standard TVP-VAR model (or exchangeable time series; Prado and West 2010, Chap 10) with the residual volatility following a matrix-beta/Wishart random walk so that model fitting and generation of forecasts is routine. Though more recent variants of these core models– such as, for example, Bayesian latent threshold TVP-VARs with stochastic volatility as in Nakajima and West (2013a)– might be considered, one benefit and appeal of forecast synthesis is making improvements over a set of relatively simple models. That is, we explore BPS applied to rather standard, and currently accepted variants of models that are traditional and whose basic model forms (up to assumptions about lags and variables to include) are accepted in the applied macroeconomic forecasting community. We note that, while using more complex models do not necessarily produce more accurate forecasts, it is important to discuss the potential synthesis accuracy under different contexts. For example, since the agent models considered here are relatively simple, we might expect the biases and inter-dependencies to be relatively simple as well, a context where BPS benefits from learning them. We can expect this to be also true for forecasts coming from forecasters, as it is observed that many forecasts exhibit patterns of bias and herding under different economic situations. More complex models, on the other hand, will have low bias and high variance, which implies little signal to be learned in the synthesis. In these settings, BPS might not benefit from learning its biases and inter-dependencies, though it will benefit and improve accuracy by averaging out its variance.

In the dynamic BPS models for forecast horizons $k = 1, 12, 24$, we take initial priors using independence across series $r$ and with $\boldsymbol{\theta}_{tr} \sim N(\boldsymbol{a}_0, \boldsymbol{R}_0)$ with $\boldsymbol{a}_0 = (0, \mathbf{1}'/J)'$ and where $\boldsymbol{R}_0$ is diagonal with diagonal elements of 1 apart from: (a) elements of 0.001 for the intercept coefficients; and (b) elements of 0.1 for coefficients on investment. Coupled with this, we take $\boldsymbol{V}_0 \sim IW(7, 7 * 0.01\boldsymbol{I})$. Discount factors in the BPS(1) model are set at $(\beta, \delta) = (0.99, 0.99)$, and variants for BPS($k$) for $k = 12, 24$-month ahead forecasting are as discussed in Section 2.5, though now we increase the prior variance on the intercept from 0.001 to 0.01 and 0.1, for $k = 12, 24$, respectively, to reflect the increased uncertainty about the relevance of the agent forecasts at longer horizons.

We have explored analyses across ranges of choices of initial priors and discount factors, and chosen these values as they lead to good agent-specific and BPS forecasting accuracy; conclusions about the main questions– how BPS can improve forecasts while generating insights into agent characteristics and dependencies over time– to not change materially with different values close to those chosen for the summary examples.

**Data Analysis and Forecasting.** The 5 agent models are analyzed and synthesized as follows. First, the agent models are analyzed in parallel over 1986/1-1993/6 as a training period to calibrate the VARs. This continues over 1993/7-2001/12 while at each month $t$ during this period, the MCMC-based BPS analysis is run in parallel using data from 1993/7 up to time $t$ in an "expanding window" fashion, adding data as we move forward in time. We do this for the traditional 1-step focused BPS model, and– separately and in parallel– for the $k = 12, 24$-step ahead focused BPS($k$) model as discussed in Section 2.5. This continues over the third period to the end of the series, 2001/1-2015/12, generating forecasts (for the agents and BPS) for each $t$ until the end of the testing period. This testing period spans over a decade and a half and includes 180 data points,

providing a good measure on how the agents and BPS perform under different economic situations; most notably before, during, and after the sub-prime mortgage crisis. Out-of-sample forecasting is thus conducted and evaluated in a way that mirrors the realities facing decision and policy makers.

**Forecast Accuracy Metrics.** Following the recent literature on macroeconomic forecasting, we compare both point and density forecasts to give a broader understanding of the predictive abilities of the agents and BPS. For the point forecasts, we compute and compare mean squared forecast error (MSFE) over the forecast horizons of interest and for each series. For density forecasts with BPS, we evaluate log predictive density ratios (LPDR); at horizon $k$ and across time indices $t$ for the joint set of series, this is

$$\text{LPDR}_{1:t}(k) = \sum_{s=1:t} \log\{p_j(\boldsymbol{y}_{s+k}|\boldsymbol{y}_{1:s})/p_{\text{BPS}}(\boldsymbol{y}_{s+k}|\boldsymbol{y}_{1:s})\}$$

where $p_j(\boldsymbol{y}_{s+k}|\boldsymbol{y}_{1:s})$ is the predictive density under each agent indexed by $j$, at each time $s$ over the next $k$ time points. The LPDR measures are, at each $t$, baselined against the corresponding BPS forecasts over each horizon $k$. LPDR provides a direct statistical assessment of the distributional accuracy of a forecast relative to, in this case, BPS for multiple horizons, extending the 1-step focused Bayes' factors. They compare the location and dispersion of the forecasts, giving an assessment of risk, elaborating on MSFE measure, and have been increasingly used in broader model comparison and forecast accuracy studies (e.g. Nakajima and West 2013a; Aastveit et al. 2018) as well as mirroring the assessment done in central banks.

**Comparison Methods.** We compare BPS with standard model combination using Bayesian model averaging (BMA) and with optimal linear pooling (OLP: Geweke and Amisano 2011; Amisano and Geweke 2017). Both strategies are linear pools generating 1-step ahead forecast forms

$$p(\boldsymbol{y}_{t+1}|\boldsymbol{w}_t, \mathcal{H}_{t+1}) = \sum_{j=1}^{J} w_{tj}h_{t+1,j}(\boldsymbol{x}_{t+1,j}), \tag{7}$$

where $\sum_{1=j}^{J} w_{tj} = 1$ and $w_{tj} \geq 0$. As shown in McAlinn and West (2017), all such form can be interpreted as special cases of BPS (eqn. 2) based on different choices of synthesis model. In BMA, these time-varying weights in the pool are simply posterior model probabilities $w_{tj} = p(\mathcal{A}_j|\boldsymbol{y}_{1:t}) \propto p(\boldsymbol{y}_{1:t}|\mathcal{A}_j)p(\mathcal{A}_j)$. With uniform prior probabilities the weights are proportional to the marginal likelihoods generated under each model and are simply updated over time. BMA is appealing in an $\mathcal{M}$-closed setting (where the true data generating model is nested), as the posterior model probability of the true model will increase towards one with accruing data. In contrast, OLP is a decision-guided strategy with weights chosen adaptively over time; at time $t$, they are chosen to maximize the log score $\sum_{s=1}^{t} \log p(\boldsymbol{y}_s|\boldsymbol{w}_t, \mathcal{H}_s)$. A key strength of OLP (Geweke and Amisano 2011; Amisano and Geweke 2017) is that weights do not necessarily degenerate asymptotically, in contrast to BMA. We sequentially revise weights over time, generating optimal OLP weights at each $t$ that are then used in eqn. (7) to produce implied one-step ahead aggregated predictive distributions.

11

## 3.2 Dynamic BPS and Forecasting

**1-step ahead forecasting.** Table 1 summarizes predictive measures for 1-step ahead forecasts. Looking at point forecasts, BPS exhibits clear improvements over agent forecasts on 5 out of the 6 series, namely inflation, wage, consumption, investment, and interest rate. The gains are at least 2%, except compared to one model for interest rate, and OLP for consumption and investment. Even for the series for which BPS does not show substantial improvement over the models, the difference between BPS and the best model is within 1%. It is also notable that the best model differs for each series. VAR(1:3:9) is best for inflation while it is the worst for wage, for example. Under traditional model combination strategies, such as BMA or OLP in which each model is assessed only on 1-step ahead density forecast accuracy and for the full multivariate forecast, accuracy is inherently aggregated over series. BPS, due to its flexible synthesis function, is able to synthesize forecasts on each series, while retaining the inter-series dependencies. This leads to BPS improving on multiple series without trading-off one over another. In fact, the results in Table 1 shows that BPS exhibits improvements for all 6 series relative to BMA and OLP (while OLP does outperform BMA for most series).

BPS demonstrates an ability to substantially improve characterization of forecast uncertainties as well as adaptation in forecast locations, reflected in the LPDR measures. The best model in terms of LPDR is only best for the wage series in terms of MSFE and performs only modestly for the other series. This indicates how LPDR measures for multiple series favor overall performance and do not distinguish models that are good for some series but poor for others. Model combination schemes that are dependent on likelihood measures– including BMA and OLP– favor the average performing model (VAR(3) in the case of 1-step ahead forecasts). The fact that OLP has slightly improved LPDR relative to BMA is due to the weights not degenerating. In contrast, BPS dynamically synthesizes forecasts for each series, while improving uncertainty assessment per series and inferring the interdependence among series. This leads to improved overall distribution forecasts as well as point forecasts. This feature of BPS is critical, as $\mathcal{D}$ typically has priorities among the series being forecast.

We now review summary graphs showing aspects of analysis evolving over time during the testing period, a period that includes challenging economic times that impede good predictive performance. Figs. 2–8 summarize sequential analysis for 1-step forecasting.

Fig. 2 shows the 1-step ahead measures $\text{MSFE}_{1:t}(1)$ over time $t$ in forecasting inflation. The other series are omitted for the sake of brevity, but the patterns in inflation are consistent (see supplementary appendix material). Additionally, forecasting inflation is one of the most important tasks for an economic policy maker, and therefore focusing on inflation is appropriate for this example. While BPS does not outperform the other models over the whole testing period, we see that it is on par with the best models considered. BPS ends up improving on the other models based on its performance during and after the sub-prime mortgage crisis, demonstrating how BPS dynamically adapts over time to produce robust forecasts over crisis periods and changing regimes.

Fig. 3 confirms that BPS performs uniformly better than the other models based on LPDR measures. The gradual decline in LPDR and a more drastic decline after the crisis is indicative of how BPS dynamically adapts its location and uncertainty to improve its distribution forecasts.

One crucial aspect of the BPS model is that it can adapt coefficients specific to each series.

Fig. 4 shows the on-line posterior means of BPS model coefficients for 1-step ahead forecasts for each series. We first note how the coefficients for each series evolve and are different across series, reflecting how different models are better at forecasting different series and how the relative accuracy differs in time. Additionally, we note that the dynamic coefficients can sometimes be somewhat volatile, which is reflective of the level of uncertainty in the latent agent states.

The very adaptive coefficient trajectories for the 1-step ahead forecast analysis–particularly compared to those from the 12- and 24-step analyses (Figs. C16-C27)– arise due to a number of factors. One feature is the use of a fairly constraining initial prior on the intercept, heavily favoring small values, so the on-line posteriors for the BPS coefficients adapt more dramatically than were the intercept to be "freer" to move around. This is coupled with generally strong and positive inter-dependencies among agent forecasts that lead to high collinearity. Coefficients trajectories are most volatile during the stable pre-crisis period, where agent forecasts are particularly similar, and less erratic when agent forecasts diverge more substantially post-crisis. For 12- and 24-step ahead forecasts, while the agent forecasts are generally poorer, their inter-dependencies are much weaker yielding more stable coefficient trajectories based on sustained differences in relative forecasting accuracy across agents even in the context of high uncertainties.

Coefficient trajectories for inflation in Fig. 4(a) exhibit structural change after the sub-prime mortgage crisis. VAR(1) and VAR(3), which are relatively simple models with short lags, have the highest coefficients up until the crisis, but these quickly drop off with more weight given to VAR(1:3:9) and VAR(1:6:12) that are more complex models with longer lags. This can be viewed as a structural change where simpler dynamics are being replaced by longer, more complex, dynamics after the crisis.

In contrast, Fig. 4(b) displays relatively stable trajectories in coefficients for wage. The VAR(3) model has the highest, stable coefficient. In comparison to inflation, the simplest VAR(1) model has limited impact, while the most complex VAR(12) model has a persistent negative coefficient which, to some degree, balances the impact of the simpler, short-lag models.

In Fig. 4(c), the estimated trajectory of coefficients for unemployment exhibits an increase for the VAR(1) model and a gradual decrease for the more complex VAR(3) model after the sub-prime mortgage crisis, as well as overall small effects from the other more complex models. Due to unemployment being heavily impacted by the crisis, this characteristic is understandable. Long term unemployment trends become irrelevant in light of the recent shock to the economy, and the coefficients reflect that shift.

For consumption coefficients in Fig. 4(d), we see a grouping of agents: VAR(1)-VAR(1:6:12) and VAR(3)-VAR(12). Before the crisis, we see that these two groups converging with concordant coefficients (except for VAR(1:3:9), which is almost always negative), then quickly re-separating again after the crisis. This suggests that consumption is mainly driven by biannual lags.

On coefficients for investment in Fig. 4(e), we see the effect of our use of a rather restrictive initial prior due to experience in the training data period of extremely high uncertainty in the agent forecasts linked to the highly volatile nature of this series. This results in relatively stable coefficients trajectories, with all being above zero and around equal weight. However, there are still clear patterns that emerge; notably, we see an upward spike in VAR(1:6:12) at around 2003

13

and some fluctuations during the subprime mortgage crisis.

For the interest rate series, Fig. 4(f) shows coefficient trajectories favoring more complex models with longer lags. Interestingly, we see a gradual decrease in coefficients on the VAR(1) model up until the sub-prime mortgage crisis, at which point it stays level. We can infer that long term dynamics were taking over short term dynamics leading up to the crisis, bringing up interesting questions about lending and credit characteristics pre-crisis. We also note that the introduction of zero interest rates after the crisis does not seem to effect the coefficients at all.

Figs. 5-6 exhibit selected aspects of inferences on the trends in uncertainty and dependence between and within agents over time. The posterior latent agent variances (the diagonal elements of the covariance matrix) reflect how forecast uncertainties change over time; see Fig. 5. Complex models for multiple series– that require estimation methods that are also complex– often produce large forecast standard deviations coming from the model, data, estimation method, or all of the above. Large VAR models are popular in practice, due to modeling flexibility and interpretability, but naturally lead to inflated uncertainty measures due to large numbers of parameters, collinearities and resulting estimation uncertainties. BPS, on the other hand, has smaller uncertainty in synthesized forecasts, resulting in decreased forecast uncertainty relative to each of the agents. This is a critical benefit of BPS. Though underestimating real risk is as dangerous as overestimating it, the LPDR results indicate that the BPS uncertainty estimates are valid– point forecasts are generally improved and lower predictive uncertainties couple with that to lead to a win-win analysis.

We now consider aspects of posterior distributions on the latent agent factor processes, investigating inter-dependencies among agent forecasts and their temporal evolutions. One set of numerical summaries is the full retrospective posterior correlations among elements of each $X_t$. High positive dependencies among agents will generate high negative correlations among the corresponding dynamic regression coefficients on the latent factors, and vice-versa, discovering the underlying temporal dependencies. We look at these measures of dependence at 3 specific time points that represent different regimes in the testing period; pre-sub-prime mortgage crisis (2003/12), immediately after the crisis (2008/12), and post-crisis (2013/12). Correlations are arranged in order agents. See Fig. 6.

At the beginning of the period where the series are relatively stable (Fig. 6(a)), we note strong (negative and positive dependencies) within agents, with some notable positive dependence between 12-month (long term) investment and interest rate and strong negative dependence for unemployment and consumption across all agents. The negative correlation within series is expected, as these are modeled by the agents and updated through BPS. Across agents, we note some diagonal patterns that are present, albeit weak. For example, looking at the off-diagonal block between VAR(1) and VAR(12), we can see negative dependencies between all series except for interest rates. This indicates some long terms dependencies, although we do expect them to exist since VAR(1) is nested within VAR(12).

Immediately after the crisis (Fig. 6(b)), we note a drastic change in correlations across and within agents appearing, as seen in the positive correlations radiating from and within series and negative correlations across agents. Focusing on VAR(1:3:9), there are strong negative correlations with VAR(3), but comparatively very little between the other agents. Other patterns, notably block

14

patterns appear between VAR(1), VAR(12), and VAR(3) that were not present pre-crisis. These emergent patterns suggest a large shift in dependence during the crisis that would have been overlooked without BPS. Post-crisis (Fig. 6(c)), we see a new dependence structure appearing, with almost no dependence across– and even within– series, except for mild dependencies within series and some diagonal structure seen in Fig. 6(a) reappearing. In particular, the correlations between VAR(1:3:9) and each of VAR(12) and VAR(3) vanishes.

The patterns of changes in dependencies among agents and across series over time can provide insights into how the economy changes with respect to economic shocks defining different regimes. Figs. 6(a) and 6(c) are both snapshots of relatively stable periods, yet the characteristics exhibited through the estimated dependencies are different. These differences in economic structure are not unexpected, though graphically visualizing the differences through the lens of agents and BPS-defined inferences on inter-dependencies provides new insight and perceptions into the overall change in economy. We further explore this analysis by highlighting how each component of the correlation changes over time. Fig. 7 represents three distinct patterns of correlation observed.

First, Fig. 7(a) relates to the correlation between VAR(1:6:12): inflation ($p$) and VAR(1:3:9): interest rate ($r$) over 2001/1-2015/12, and presents a clear pattern of crisis correlation where the dependence is practically flat except during the recession. Note that the correlation is between two separate agents, something that is unattainable in a typical modeling/combination framework. As these two models capture different dynamics (VAR(1:6:12) is bi-annual and VAR(1:3:9) is quarterly), we can glean insight into the complex dependence between series over time. The figure clearly shows three major regimes: a stable period (pre-mid 2004 and post-2011), where the correlation is near zero, a pre-crisis period (mid 2004 to 2008), where there is a small but persistent increase in correlation, and a crisis period (2008 to 2011), where there is a large negative correlation.

Second, Fig. 7(b) presents a business-cycle pattern, measuring the dynamic correlation between unemployment ($u$) and investment ($i$) for VAR(3). Similar to Fig. 7(a), the correlation pattern emerges strongly during the crisis, though contrary to Fig. 7(a), the correlation is persistent and fluctuating pre-crisis. Interestingly, the shock in the correlation from the crisis slowly diminishes over time displaying a business-cycle pattern. BPS analysis thus suggests a need to consider this correlation in the recovery phase after the recession.

Finally, Fig. 7(c)– the correlation between inflation ($p$) and unemployment ($u$) for VAR(1)– displays strong evidence of structural change between these economic variables. Pre-crisis, the correlation is positive with some cyclical fluctuations. Post-crisis, there is a steady decrease in correlation, where it becomes negative at around 2012. This result suggests a weakening of the Phillips curve after the crisis, a phenomenon widely referenced in the literature.

Finally, Fig. 8 measures the Kullback-Leibler divergence of the posterior on agent latent states $\boldsymbol{X}_t$ from their prior at each time $t$, namely $D_{KL}(t) = \mathbb{E}[\log(p(\boldsymbol{X}_t|\boldsymbol{y}_{1:t}, \mathcal{H}_{1:t})/h(\boldsymbol{X}_t))]$ at each $t$ where $h(\boldsymbol{X}_t)$ is the product of individual agent forecast densities (the time $t$ prior for latent states), $p(\boldsymbol{X}_t) = p(\boldsymbol{X}_t|\boldsymbol{y}_{1:t}, \mathcal{H}_{1:t})$ is the posterior, and the expectation is with respect to $p(\boldsymbol{X}_t)$. This can be numerically approximated using the posterior Monte Carlo samples $\{\boldsymbol{X}_{1:t,i}, \boldsymbol{\Phi}_{1:t,i} : i = 1{:}n\}$; specifically, $D_{KL}(t) \approx \sum_{i=1:n} \log(p(\boldsymbol{X}_{ti}|\boldsymbol{\Phi}_{1:t,i}, \boldsymbol{y}_{1:t}, \mathcal{H}_{1:t})/h(\boldsymbol{X}_{ti}))/n$ where the terms in each summand are analytically available. The divergence measures change in information so is a direct

measurement/summary of the extent to which biases and inter-dependencies are learned through BPS. If an agent produces significantly biased forecasts during economic distress, BPS will learn and calibrate that bias, and the difference will be captured in the divergence. Similarly, we expect and observe inter-dependencies across agents to increase during crises and shocks– see Fig. 6(c)– and in such contexts the divergence will naturally increase. In Fig. 8 we see three major spikes– during the dot-com bubble, then before and after the great financial crisis. The first and third spikes are understandable; Fig. 6(c) indicates that strong agent inter-dependencies appear during a crisis. The second spike is perhaps more interesting, as early 2008 was when– in hindsight– early signs of the crisis were emerging. Because the subtle shifts in agent forecasts (increased biases and inter-dependencies not reported in the prior agent forecasts) are captured using BPS, the divergence metric is able to pick up the early signals of a recession. Although the results presented here are retrospective (i.e. the posterior is in regard to $t = 1{:}T$), it is indicative of the strengths of BPS in capturing crises and shocks measured through prior-posterior divergences on the fundamental latent factors that BPS defines.

**k-step ahead forecasting.**  Long term forecasting for economic policy is far more important than 1-step ahead forecasting. Analysis forecasts 12- (one year) and 24- (two years) steps ahead to demonstrate BPS over the set of agents at practically important horizons.

Tables 2 and 3 summarize predictive measures for the two forecast horizons. For point forecasts, BPS($k$) outperforms all other models for all series with the exception of wage growth. The improvements hold for all $k$-step ahead forecasts considered and the improvements of BPS($k$) significantly increase with $k$. The improvements come from BPS($k$) directly synthesizing the $k$-step ahead forecasts from the agents, calibrating, adapting, and learning the latent dependencies and biases over the $k$-step ahead quantity of interest. For 24-step ahead forecasts of inflation, one of the most important series for a central bank when setting their key policy rate, BPS($k$) greatly improves on the agent models, with massive gains over the best agent model. As with 1-step ahead forecasting, it is also notable that agent model performances vary significantly across series. In contrast to BPS, traditional model combination schemes, such as BMA, fail to improve over all series by sacrificing improved accuracy for one series over others; in fact, for both 12- and 24-step ahead forecasts, BMA-based analysis degenerates to the VAR(1) model. In addition, BPS($k$) significantly improves in quantifying uncertainty in forecasts, as evident in the comparison of LPDR. Creating long-term forecasts for multiple time series is a very difficult problem due to the nature of these models being built and trained on 1-step ahead forecasting metrics (likelihood) and failing to propagate forward accurately. In contrast, BPS($k$) synthesizes the $k$-step ahead forecasts directly, adjusting and calibrating uncertainty according to the actual quantity of interest. Thus, no matter how the agent uncertainty forecasts are over- or under-estimating, BPS($k$) can re-adjust accordingly by learning how the agents over- or under-estimate. The consistency of the LPDR improvements over multiple $k$-steps demonstrate this key feature of BPS($k$).

As with the sequential MSFE results for 1-step ahead forecasts, we focus now on multi-step MSFE results for inflation. The characteristics of the results for other series are similar but omitted for brevity. Figs. 9 and 11 exhibit MSFE comparisons for inflation over the testing period for $k =$

12, 24-step ahead forecasts. Although the scale is different for each $k$, there are notable common characteristics that characterize the BPS($k$) results. For example, agent models experience several large shocks in precision over the testing period; this occurs, in particular, around the time of the advent of the sub-prime mortgage crisis in the late 2000s. These shocks particularly effect the precision of the agent forecasts, especially the 24-step ahead forecasts. In comparison, BPS($k$) stays relatively robust throughout multiple shocks and structural breaks.

In LPDR evolutions over time (Figs. 10 and 12), BPS($k$) improves over the agent models over all of the time period considered, except for slight increases in favoring the simpler VAR(1) model immediately post crisis. BPS($k$) is able to adapt to maintain improved forecasting performance both in terms of location and uncertainty assessment, a key positive feature for decision makers tasked with forecasting risk and quantiles for long horizons under possible shocks and regime change.

Figs. C22-C27 exhibit the on-line posterior means of BPS model coefficients for the 24-step ahead forecasts. The coefficients for 6- and 12-step ahead forecasts are omitted– similar conclusions arise in those analyses. Overall, the BPS($k$) coefficients are relatively stable compared to the 1-step ahead results, due somewhat to the lack of signal from the agent forecasts. The agent forecasts' ability for 24-steps are considerably worse than from their 1-step ahead counterparts, leading to less useful information to be synthesized by BPS($k$). The lack of signal from all of the agent models leads to less movement in the coefficients, and in turn, an increase in adaptability in the intercept.

**Scenario analysis.** We conclude our empirical study with a scenario analysis using conditional forecasts. Scenario analysis is a growing, critical component in economic policy/decision making, with interest in evaluating the potential effects of certain policy decisions on future outcomes in terms of other key economic variables (see, e.g., Bańbura et al. 2015). We conduct a scenario analysis using conditional forecasts– emulating real policy making decisions– to explore BPS for economic policy and decision making under contexts of interest. We focus on the "zero" interest rate policy after the financial crisis and reflect the real-world policy of the Federal Reserve Board at 2008/10, in the beginning months of the financial crisis. At that time point, we examine the hypothetical impact on inflation and unemployment of a zero interest rate policy that will be maintained for the next two years. We fix interest rate at 0.1% for each month of the two years beginning at 2008/10 and forecast inflation and unemployment at 2010/10. This uses the relevant 24-step ahead conditional forecasts from the five agent models integrated into the BPS(24) analysis now conditional on the constrained future path of interest rates.

We compare the distributional forecasts of the five agents and BPS against the actual realization of inflation and unemployment at 2010/10 (Fig. 13). For inflation, BPS not only has a lower mean error than each agent, but has a significantly tighter uncertainty band. This is simply critical for policy making, as large uncertainty assessments– seen in the agents– reflect practically unusable forecasts, not to mention that all agents forecast negative inflation. The results for unemployment echo those for inflation, with BPS displaying superior mean and uncertainty estimates. Although VAR(1:3:9) does slightly improve on BPS for the mean, the forecast from VAR(1:3:9) is practically unusable for policy making with such a large uncertainty band. On the other hand, BPS produces similarly accurate mean forecasts while greatly shrinking its uncertainty.

The scenario analysis using conditional forecasts highlights the usefulness and practicality of BPS for advising economic policy and decision making. Individual forecasts from agents did not produce pragmatic conditional forecasts, in terms of mean, uncertainty, or both, for use in policy assessment. In contrast, we see that the BPS synthesized conditional forecasts significantly improve forecast accuracy and usability for decision making, even though the underlying agent forecasts are imperfect. The results from the scenario analysis, combined with its performance from the forecasting exercise, indicate the potential utility of BPS in terms of forecasting, inference, and real-world decision making.

# 4   Summary and Additional Comments

Multivariate BPS defines a theoretically and conceptually sound framework to compare and synthesize multivariate density forecasts in a dynamic context. The approach will enable decision makers to dynamically calibrate, learn, and update predictions based on ranges of forecasts from sets of models, as well as from more subjective sources such as individual forecasters or agencies. While it will be of interest to develop future studies in which agents are represented by sets of more elaborate macroeconomic models– such as dynamic threshold models and dynamic stochastic general equilibrium (DSGE) models– and to integrate forecasts coming from professional forecasters and economists, the current case study already demonstrates the real practical potential. In our sequential BPS analysis of multiple US macroeconomic series, we have highlighted questions of forecast synthesis methodology with respect to forecasting goals: interest in 12 or 24 month-ahead forecasting demands– from a formal Bayesian perspective– analysis customized to the horizon, and the results bear out the practical relevance of that perspective. The studies show that the flexible and interpretable DFSUR models can (i) adapt to time-varying biases and miscalibration of multiple models or forecasters, (ii) adaptively and practically account for– while generating useful insights into– patterns of time-varying relationships and dependencies among sets of models or forecasters, and (iii) improving forecast accuracy– in some cases, most substantially– for each of several multiple macroeconomic series together, at multiple horizons. The predictive performance of BPS is robust in times of severe economic distress, which is important for practical applications.

Additionally, and most importantly, inference on the inter-dependencies among forecasting models– linked to the BPS foundational latent factor structure and aspects of inference on time-varying parameters characterizing that structure– provides both illumination of the inter-dependencies, and how they may vary across subsets of the multivariate series. This also provides the decision maker with the opportunity to respond and change or intervene in the BPS modeling for continued forecast synthesis into the future. These features are combined with the very practically focused use of scenario analysis within which, in our case study, the BPS approach defines relevant and potentially useful inputs to policy decision making processes even though each of the individual forecasters is alone of limited value.

Multivariate BPS has further potential in applications to other fields and data where inter-dependencies between series have impact on the decision making, and where multiple forecasts, whether from forecasters or models, are available. Such applications include financial data, such

as stocks, indexes, and bonds with portfolio decisions in mind. Further methodological extensions that warrant investigation include non-normal forecasts and discrete data, and missing or incomplete/partial forecasts. A second, major area for extension arises from the observation that our BPS setting is not, in fact, as general as it could be in terms of building on the historical Bayesian agent analysis literature. Referring back to Section 2.1, each agent $\mathcal{A}_j$ could provide a forecast density $h_j(z_j)$ for a set of outcomes $z_j$ that is not, in fact, exactly the outcome $y$ of interest to $\mathcal{D}$. For example, $z_j$ may be a subset of the macroeconomic variables in $y$, but not all of them; and/or it may include additional, relevant variables not included in $y$. The same foundational theory applies, and this provides potential to explore BPS when different models or forecasters have differing areas of expertise as well as different strategies in forecasting collections of related series.

There are several aspects of multivariate analysis that can be further explored as well. In our main analysis, we had agents who produced multivariate forecasts. However, for forecast surveys, forecasts may be conditional forecasts, received individually for each series. Under this context, using BPS can be a powerful tool to uncover the dependence structure for each agent. This also raises an interesting question of whether synthesizing multivariate forecasts improves over synthesizing univariate, conditional forecasts, especially when interests lie in forecasting only one or few series within a large series (similar to the scenario analysis in the empirical application).

Interesting and practically relevant questions concern the scalability of multivariate BPS, with both modeling and computational concerns. As a theoretically well-founded framework, BPS does not have implicit limitations in terms of scalability, though practical relevance of synthesis models and computational implementation naturally raise scalability issues. For example, the dynamic SUR model will struggle to handle a large number of series due to the inefficiency associated with estimating the time-varying covariance matrix. Scalability questions are, of course, not specific to BPS; they are questions faced by all multivariate time series models, and are exacerbated in BPS as the multivariate models are imposed on inherently latent factor processes. The same general considerations arise in considering larger numbers of models/agents. Any serious consideration of substantial scaling will need to adopt model developments relevant more broadly– such as building on some sparse graphical modeling or latent factor ideas that appear in other areas of recent research in time series, scalable extensions of SUR-style structures, or direct reduction ideas such as clustering series (or agents, or both) into groups and modeling cross-group relationships using a modified BPS approach. Some potential lies in exploring dynamic latent factor models (e.g. Aguilar and West 2000) at the level of the latent agent factors, i.e., new kinds of hierarchical latent factor models, to begin to address scalability. Other ideas with significant potential might explore coupled systems of dynamic models for univariate series, such as developing more highly structured simultaneous graphical dynamic linear models (SGDLMs: Gruber and West 2016, 2017) for multivariate latent agent factor processes. In other contexts requiring fast data processing in sequential analyses, it may be that some form of sequential Monte Carlo (SMC, e.g. Lopes and Tsay 2011) may prove useful, especially if enabled by decoupling of multivariate series into small or univariate but linked subsets, within each of which customized SMC methods might be more effective. Again, this might be approached by exploiting and extending to a BPS framework concepts of decouple/recouple in modeling increasingly high-dimensional time series elsewhere (e.g. Gruber and West 2016; Chen

et al. 2017; Gruber and West 2017). Beyond our development of DFSUR models in the current study, these are general questions at the forefront of research in multivariate dynamic modeling generally, and worth pointing out for potential future research.

# References

Aastveit, K. A., Carriero, A., Clark, T. E., and Marcellino, M. (2017), "Have standard VARs remained stable since the crisis?" *Journal of Applied Econometrics*, 32, 931–951.

Aastveit, K. A., Gerdrup, K. R., Jore, A. S., and 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., and van Dijk, H. K. (2018), "Combined density nowcasting in an uncertain economic environment," *Journal of Business & Economic Statistics*, 36, 131–145.

Aguilar, O. and West, M. (2000), "Bayesian dynamic factor models and portfolio allocation," *Journal of Business & Economic Statistics*, 18, 338–357.

Amendola, A. and Storti, G. (2015), "Model uncertainty and forecast combination in high-dimensional multivariate volatility prediction," *Journal of Forecasting*, 34, 83–91.

Amisano, G. and Geweke, J. (2017), "Prediction using several macroeconomic models," *Review of Economics and Statistics*, 99, 912–925.

Amisano, G. G. and Giacomini, R. (2007), "Comparing density forecasts via weighted likelihood ratio tests," *Journal of Business & Economic Statistics*, 25, 177–190.

Andersson, M. K. and Karlsson, S. (2008), "Bayesian forecast combination for VAR models," in *Bayesian Econometrics (Advances in Econometrics, Volume 23)*, eds. S. Chib, W., Griffiths, Koop, G., and Terrell, D., Emerald Group Publishing Limited, pp. 501–524.

Bańbura, M., Giannone, D., and Lenza, M. (2015), "Conditional forecasts and scenario analysis with vector autoregressions for large cross-sections," *International Journal of Forecasting*, 31, 739–756.

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

Benati, L. and Surico, P. (2008), "Evolving US monetary policy and the decline of inflation predictability," *Journal of the European Economic Association*, 6, 634–646.

Billio, M., Casarin, R., Ravazzolo, F., and 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., and van Dijk, H. K. (2015), "Parallel sequential Monte Carlo for efficient density combination: The DeCo MATLAB Toolbox," *Journal of Statistical Software, Articles*, 68, 1–30.

Chan, J. C., Koop, G., Leon-Gonzalez, R., and Strachan, R. W. (2012), "Time varying dimension models," *Journal of Business & Economic Statistics*, 30, 358–367.

Chen, X., Irie, K., Banks, D., Haslinger, R., Thomas, J., and West, M. (2017), "Scalable Bayesian modeling, monitoring and analysis of dynamic network flow data," *Journal of the American Statistical Association*, in press- published online July 10, 2016, arXiv:1607.02655.

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

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

— (2007), "Aggregating probability distributions," in *Advances in Decision Analysis: From Foundations to Applications*, eds. W. Edwards, R. M. and von Winterfeldt, D., Cambridge University Press, chap. 9, pp. 154–176.

Cogley, T. and 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., and Winkler, R. L. (1995), "Coherent combination of experts' opinions," *Test*, 4, 263–313.

Del Negro, M., Hasegawa, R. B., and Schorfheide, F. (2016), "Dynamic prediction pools: An investigation of financial frictions and forecasting performance," *Journal of Econometrics*, 192, 391–405.

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. and Schervish, M. J. (1985), "Modelling expert judgements for Bayesian updating," *Annals of Statistics*, 13, 1198–1212.

Geweke, J. and Amisano, G. G. (2012), "Prediction with misspecified models," *The American Economic Review*, 102, 482–486.

Geweke, J. F. and Amisano, G. G. (2011), "Optimal prediction pools," *Journal of Econometrics*, 164, 130–141.

Gneiting, T. and Ranjan, R. (2013), "Combining predictive distributions," *Electronic Journal of Statistics*, 7, 1747–1782.

Gruber, L. F. and West, M. (2016), "GPU-accelerated Bayesian learning in simultaneous graphical dynamic linear models," *Bayesian Analysis*, 11, 125–149.

— (2017), "Bayesian forecasting and scalable multivariate volatility analysis using simultaneous graphical dynamic linear models," *Econometrics and Statistics*, 3, 3–22, arXiv:1606.08291.

Hall, S. G. and Mitchell, J. (2007), "Combining density forecasts," *International Journal of Forecasting*, 23, 1–13.

Hoogerheide, L., Kleijn, R., Ravazzolo, F., Van Dijk, H. K., and Verbeek, M. (2010), "Forecast accuracy and economic gains from Bayesian model averaging using time-varying weights," *Journal of Forecasting*, 29, 251–269.

Jore, A. S., Mitchell, J., and Vahey, S. P. (2010), "Combining forecast densities from VARs with uncertain instabilities," *Journal of Applied Econometrics*, 25, 621–634.

Kapetanios, G., Mitchell, J., Price, S., and Fawcett, N. (2015), "Generalised density forecast combinations," *Journal of Econometrics*, 188, 150–165.

Kascha, C. and Ravazzolo, F. (2010), "Combining inflation density forecasts," *Journal of Forecasting*, 29, 231–250.

Koop, G., Korobilis, D., et al. (2010), "Bayesian multivariate time series methods for empirical macroeconomics," *Foundations and Trends in Econometrics*, 3, 267–358.

Koop, G., Leon-Gonzalez, R., and Strachan, R. W. (2009), "On the evolution of the monetary policy transmission mechanism," *Journal of Economic Dynamics and Control*, 33, 997–1017.

Korobilis, D. (2013), "VAR forecasting using Bayesian variable selection," *Journal of Applied Econometrics*, 28, 204–230.

Lindley, D. V., Tversky, A., and 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. and Tsay, R. S. (2011), "Particle filters and Bayesian inference in financial econometrics," *Journal of Forecasting*, 30, 168–209.

McAlinn, K. and West, M. (2017), "Dynamic Bayesian predictive synthesis in time series forecasting," *Journal of Econometrics*, forthcoming, arXiv:1601.07463.

Nakajima, J. (2011), "Time-varying parameter VAR model with stochastic volatility: An overview of methodology and empirical applications," *Monetary and Economic Studies*, 29, 107–142.

Nakajima, J. and West, M. (2013a), "Bayesian analysis of latent threshold dynamic models," *Journal of Business & Economic Statistics*, 31, 151–164.

— (2013b), "Bayesian dynamic factor models: Latent threshold approach," *Journal of Financial Econometrics*, 11, 116–153.

Pettenuzzo, D. and Ravazzolo, F. (2016), "Optimal portfolio choice under decision-based model combinations," *Journal of Applied Econometrics*, 31, 1312–1332.

Prado, R. and 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.

Satopää, V. A., Jensen, S. T., Mellers, B. A., Tetlock, P. E., and Ungar, L. H. (2014), "Probability aggregation in time-series: Dynamic hierarchical modeling of sparse expert beliefs," *The Annals of Applied Statistics*, 8, 1256–1280.

Sims, C. A. (1993), "A nine-variable probabilistic macroeconomic forecasting model," in *Business Cycles, Indicators and Forecasting*, eds. Stock, J. H. and Watson, M. W., University of Chicago Press, pp. 179–212.

Sims, C. A. and Zha, T. (1998), "Bayesian methods for dynamic multivariate models," *International Economic Review*, 96, 949–968.

Stock, J. H. and Watson, M. W. (1996), "Evidence on structural instability in macroeconomic time series relations," *Journal of Business & Economic Statistics*, 14, 11–30.

Timmermann, A. (2004), "Forecast combinations," in *Handbook of Economic Forecasting*, eds. Elliott, G., Granger, C. W. J., and Timmermann, A., North Holland, vol. 1, chap. 4, pp. 135–196.

West, M. (1984), "Bayesian aggregation," *Journal of the Royal Statistical Society (Series A: General)*, 147, 600–607.

— (1988), "Modelling expert opinion (with discussion)," in *Bayesian Statistics 3*, eds. Bernardo, J. M., DeGroot, M. H., Lindley, D. V., and Smith, A. F. M., Oxford University Press, pp. 493–508.

— (1992), "Modelling agent forecast distributions," *Journal of the Royal Statistical Society (Series B: Methodological)*, 54, 553–567.

West, M. and Crosse, J. (1992), "Modelling of probabilistic agent opinion," *Journal of the Royal Statistical Society (Series B: Methodological)*, 54, 285–299.

West, M. and Harrison, P. J. (1997), *Bayesian Forecasting & Dynamic Models*, Springer Verlag, 2nd ed.

Zellner, A. (1962), "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias," *Journal of the American Statistical Association*, 57, 348–368.

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

# Tables and Figures

Table 1: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead forecast evaluations for monthly US macroeconomic series over the 15 years 2001/1-2015/12, comparing mean squared forecast errors and log predictive density ratios for this $T = 180$ months. The column $\%$ denotes improvements over the standard BPS model. Note: $\text{LPDR}_{1:T}$ is relative to the standard BPS model.

$\text{MSFE}_{1:T}$

| 1-step | Infl | % | Wage | % | Unemp | % |
|---|---|---|---|---|---|---|
| VAR(1) | 0.0141 | −8.22 | 0.1444 | −35.91 | 0.0206 | 0.74 |
| VAR(12) | 0.0160 | −22.93 | 0.1110 | −4.44 | 0.0230 | −10.73 |
| VAR(3) | 0.0147 | −13.24 | 0.1105 | −3.96 | 0.0219 | −5.67 |
| VAR(1:3:9) | 0.0135 | −3.76 | 0.1198 | −12.77 | 0.0222 | −7.18 |
| VAR(1:6:12) | 0.0137 | −5.14 | 0.1449 | −36.40 | 0.0215 | −3.70 |
| BMA | 0.0146 | −12.20 | 0.1111 | −4.53 | 0.0218 | −5.26 |
| OLP | 0.0137 | −5.09 | 0.1170 | −10.06 | 0.0212 | −2.65 |
| BPS | 0.0130 | − | 0.1063 | − | 0.0207 | − |

$\text{MSFE}_{1:T}$

| 1-step | Cons | % | Invest | % | Interest | % |
|---|---|---|---|---|---|---|
| VAR(1) | 0.3908 | −3.50 | 13.2183 | −2.99 | 0.0275 | −35.07 |
| VAR(12) | 0.4697 | −24.41 | 15.3571 | −19.65 | 0.0246 | −20.92 |
| VAR(3) | 0.3982 | −5.48 | 13.3210 | −3.79 | 0.0211 | −3.74 |
| VAR(1:3:9) | 0.4049 | −7.25 | 13.8918 | −8.24 | 0.0204 | −0.55 |
| VAR(1:6:12) | 0.3889 | −3.02 | 13.4301 | −4.64 | 0.0228 | −12.02 |
| BMA | 0.3971 | −5.18 | 13.2145 | −2.96 | 0.0215 | −5.80 |
| OLP | 0.3776 | −0.04 | 12.8692 | −0.27 | 0.0214 | −5.49 |
| BPS | 0.3775 | − | 12.8346 | − | 0.0203 | − |

| 1-step | $\text{LPDR}_{1:T}$ |
|---|---|
| VAR(1) | −92.05 |
| VAR(12) | −118.63 |
| VAR(3) | −45.81 |
| VAR(1:3:9) | −49.02 |
| VAR(1:6:12) | −67.49 |
| BMA | −47.28 |
| OLP | −43.54 |
| BPS | − |

Table 2: US macroeconomic forecasting 2001/1-2015/12: 12-step ahead forecast evaluations for monthly US macroeconomic series over the 15 years 2001/1-2015/12, comparing mean squared forecast errors and log predictive density ratios for this $T = 180$ months. The column % denotes improvements over BPS(12). Note: LPDR$_{1:T}$ is relative to BPS(12).

|  | MSFE$_{1:T}$ | | | | | |
| 12-step | Infl | % | Wage | % | Unemp | % |
| --- | --- | --- | --- | --- | --- | --- |
| VAR(1) | 0.5317 | $-143.15$ | 0.4453 | 19.50 | 1.2028 | $-10.66$ |
| VAR(12) | 0.4272 | $-95.35$ | 0.7750 | $-40.12$ | 1.6918 | $-55.65$ |
| VAR(3) | 0.5789 | $-164.74$ | 0.5215 | 5.72 | 1.1788 | $-8.45$ |
| VAR(1:3:9) | 0.4541 | $-107.69$ | 1.1207 | $-102.62$ | 1.6353 | $-50.46$ |
| VAR(1:6:12) | 0.5342 | $-144.30$ | 0.8934 | $-61.52$ | 1.3585 | $-24.99$ |
| OLP | 0.4794 | $-119.23$ | 0.4636 | 16.18 | 1.1634 | $-7.04$ |
| BPS(12) | 0.2187 | $-$ | 0.5531 | $-$ | 1.0869 | $-$ |

|  | MSFE$_{1:T}$ | | | | | |
| 12-step | Cons | % | Invest | % | Interest | % |
| --- | --- | --- | --- | --- | --- | --- |
| VAR(1) | 7.2471 | $-23.21$ | 7067.67 | $-65.55$ | 5.5916 | $-68.74$ |
| VAR(12) | 18.4145 | $-213.07$ | 8824.02 | $-106.68$ | 6.1707 | $-86.22$ |
| VAR(3) | 7.3142 | $-24.35$ | 6378.42 | $-49.40$ | 4.8222 | $-45.52$ |
| VAR(1:3:9) | 10.3823 | $-76.51$ | 9111.99 | $-113.43$ | 4.6622 | $-40.69$ |
| VAR(1:6:12) | 10.1116 | $-71.91$ | 10013.45 | $-134.54$ | 7.4612 | $-125.16$ |
| OLP | 6.9217 | $-17.68$ | 6758.88 | $-58.31$ | 5.3488 | $-61.41$ |
| BPS(12) | 5.8818 | $-$ | 4269.33 | $-$ | 3.3137 | $-$ |

| 12-step | LPDR$_{1:T}$ |
| --- | --- |
| VAR(1) | $-119.05$ |
| VAR(12) | $-535.09$ |
| VAR(3) | $-366.85$ |
| VAR(1:3:9) | $-463.46$ |
| VAR(1:6:12) | $-361.20$ |
| OLP | $-338.48$ |
| BPS(12) | $-$ |

Table 3: US macroeconomic forecasting 2001/1-2015/12: 24-step ahead forecast evaluations for monthly US macroeconomic series over the 15 years 2001/1-2015/12, comparing mean squared forecast errors and log predictive density ratios for this $T = 180$ months. The column % denotes improvements over BPS(24). Note: LPDR$_{1:T}$ is relative to BPS(24).

| | MSFE$_{1:T}$ | | | | | |
|---|---|---|---|---|---|---|
| 24-step | Infl | % | Wage | % | Unemp | % |
| VAR(1) | 3.9536 | −331.10 | 2.4117 | 7.71 | 16.46 | −55.68 |
| VAR(12) | 2.7373 | −198.47 | 4.5054 | −72.41 | 18.32 | −73.28 |
| VAR(3) | 3.8504 | −319.85 | 3.1877 | −21.98 | 13.78 | −30.35 |
| VAR(1:3:9) | 4.8627 | −430.23 | 8.8723 | −239.52 | 21.06 | −99.17 |
| VAR(1:6:12) | 4.4141 | −381.32 | 8.4162 | −222.06 | 16.99 | −60.65 |
| OLP | 2.7301 | −197.69 | 2.6218 | −0.33 | 14.95 | −41.35 |
| BPS(24) | 0.9171 | − | 2.6132 | − | 10.58 | − |

| | MSFE$_{1:T}$ | | | | | |
|---|---|---|---|---|---|---|
| 24-step | Cons | % | Invest | % | Interest | % |
| VAR(1) | 56.27 | −104.54 | 51937 | −776.23 | 31.68 | −480.56 |
| VAR(12) | 118.09 | −329.23 | 38151 | −543.65 | 25.89 | −374.58 |
| VAR(3) | 46.80 | −70.09 | 39671 | −569.30 | 21.84 | −300.31 |
| VAR(1:3:9) | 78.73 | −186.15 | 80278 | −1254.37 | 25.41 | −365.71 |
| VAR(1:6:12) | 72.54 | −163.67 | 86671 | −1362.23 | 62.16 | −1039.32 |
| OLP | 50.65 | −84.10 | 41648 | −602.65 | 25.41 | −365.72 |
| BPS(24) | 27.51 | − | 5927 | − | 5.46 | − |

| 24-step | LPDR$_{1:T}$ |
|---|---|
| VAR(1) | −445.81 |
| VAR(12) | −489.98 |
| VAR(3) | −462.48 |
| VAR(1:3:9) | −808.31 |
| VAR(1:6:12) | −804.49 |
| OLP | −339.92 |
| BPS(24) | − |
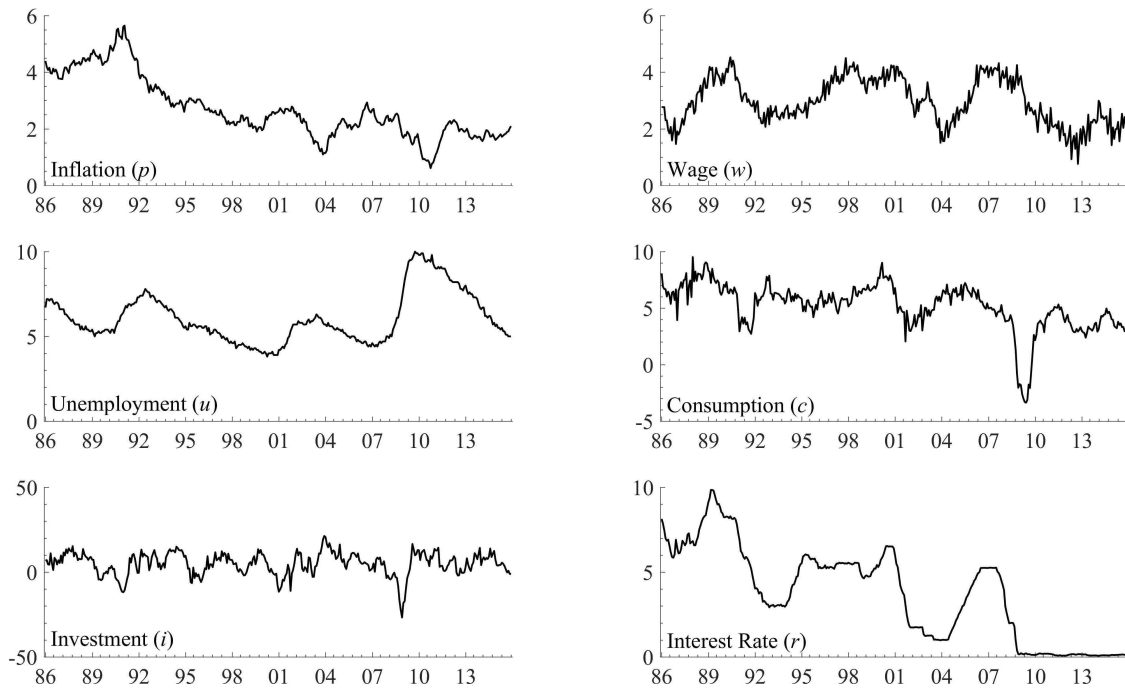
Figure 1: US macroeconomic data 1986/1-2015/12: US macroeconomic time series (indices $\times 100$ for % basis): annual inflation rate ($p$), wage ($w$), unemployment rate ($u$), consumption ($c$), investment ($i$), and short-term nominal interest rate ($r$).
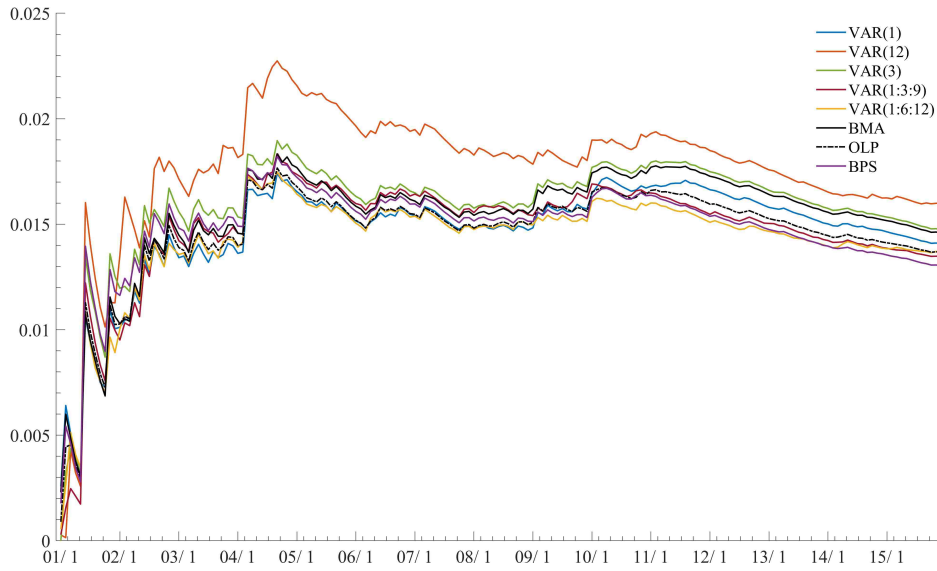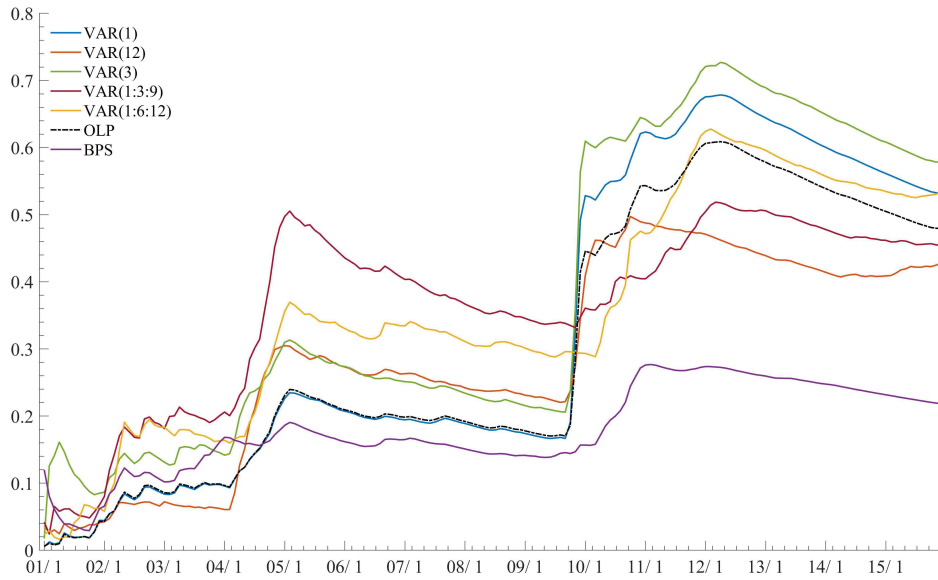
Figure 2: US macroeconomic forecasting 2001/1-2015/12: Mean squared 1-step ahead forecast errors MSFE$_{1:t}(1)$ of inflation $(p)$ sequentially revised at each of the $t = 1:180$ months.
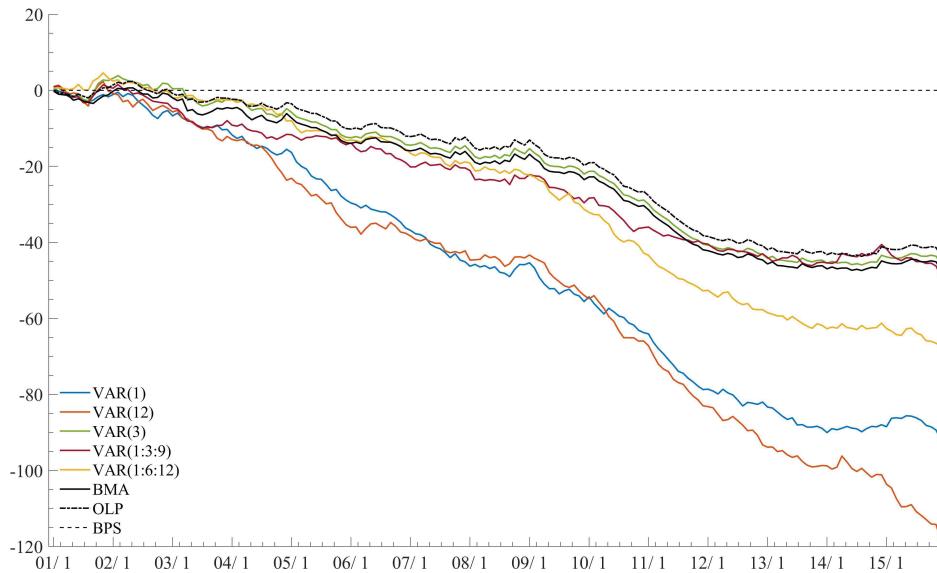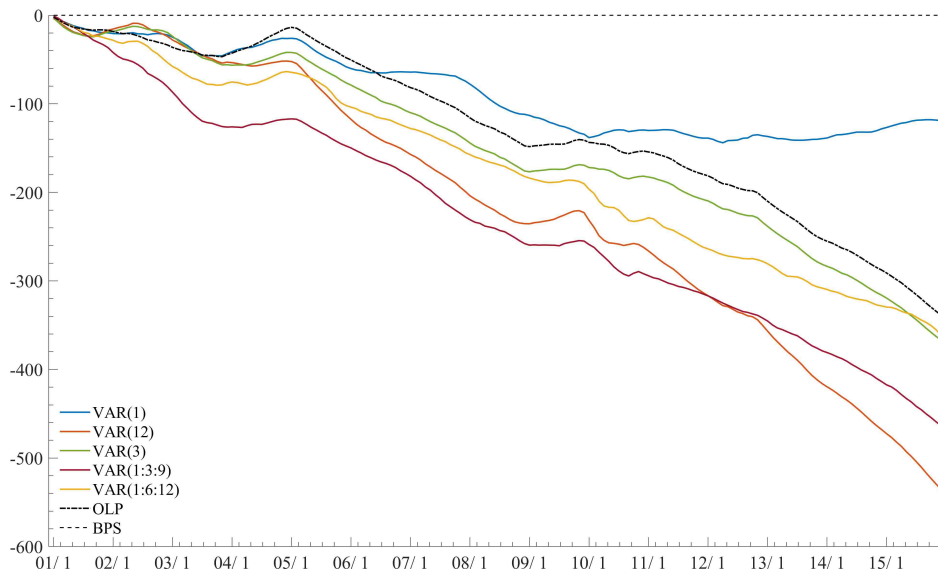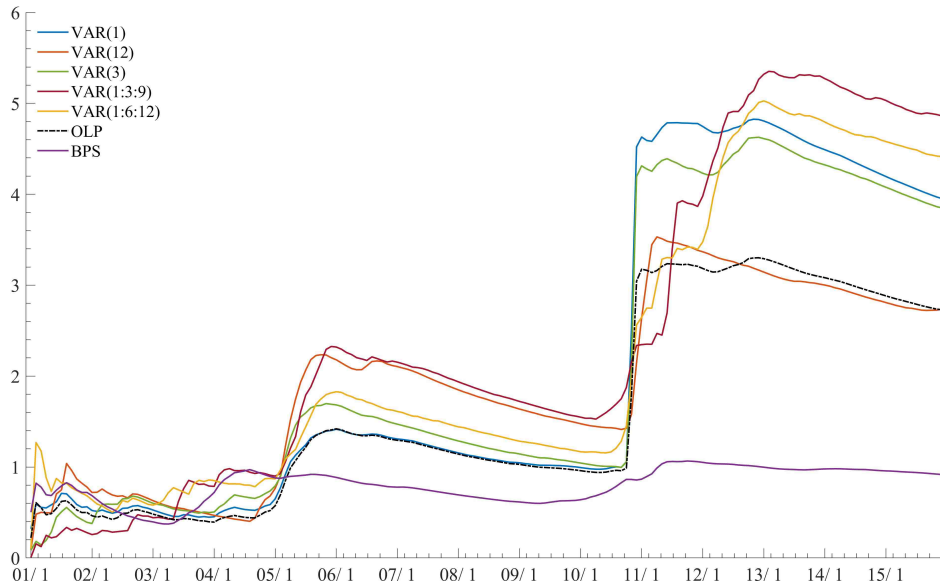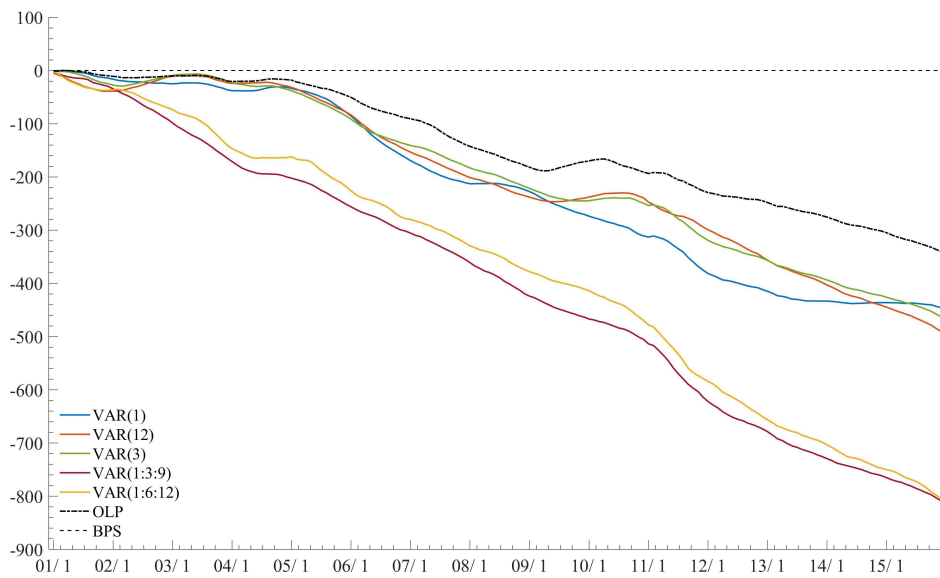


Figure 3: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead log predictive density ratios LPDR$_{1:t}(1)$ sequentially revised at each of the $t = 1:180$ months. The baseline at 0 over all $t$ corresponds to the standard BPS model.

(a) Inflation ($p$)

(b) Wage ($w$)

(c) Unemployment ($u$)

(d) Consumption ($c$)

(e) Investment ($i$)

(f) Interest rate ($r$)

Figure 4: US macroeconomic forecasting 2001/1-2015/12: On-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1:180$ months.
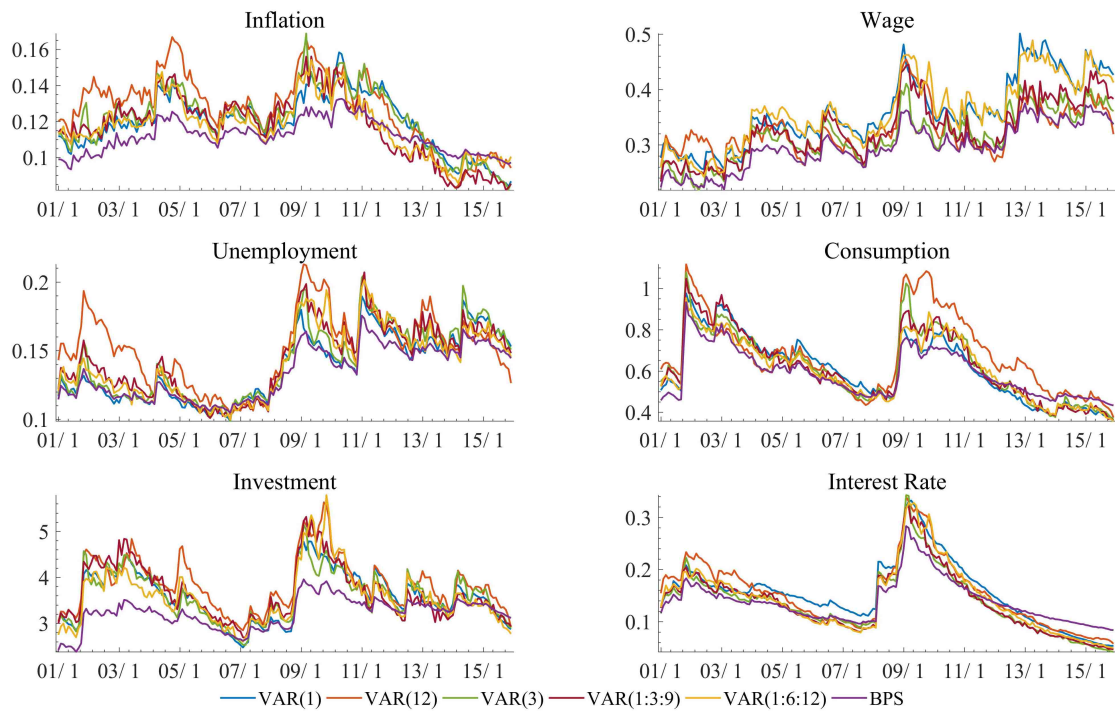
Figure 5: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}180$ months.

(a) 2003/12
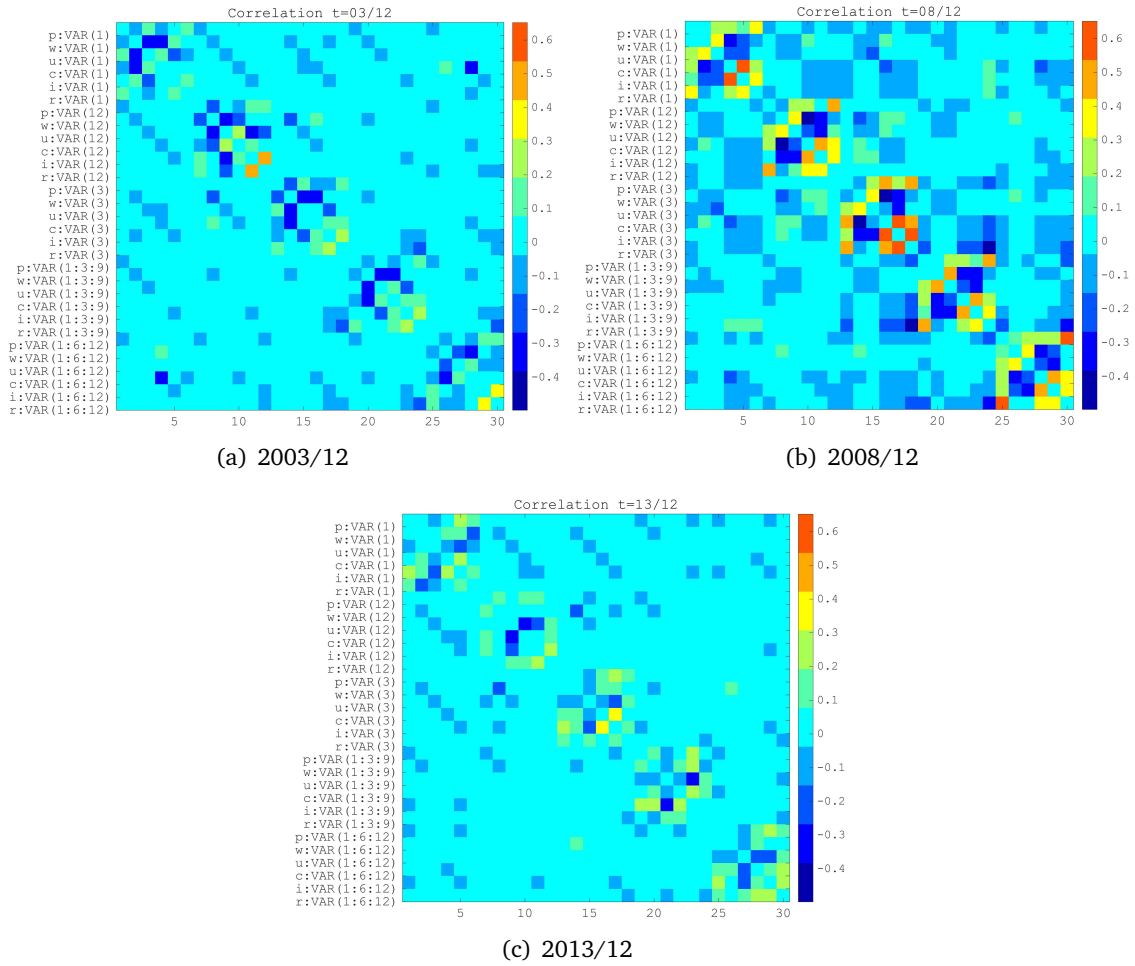
(b) 2008/12

(c) 2013/12

Figure 6: US macroeconomic forecasting 2001/1-2015/12: Retrospective posterior correlations of latent agent factors at 2003/12, 2008/12, and 2013/12.
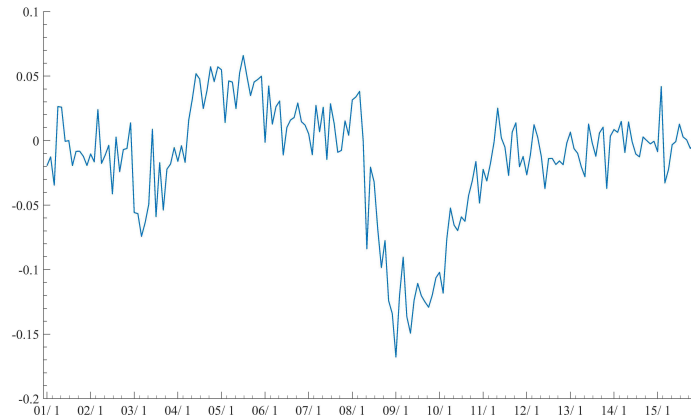
(a) VAR(1:6:12): inflation ($p$) and VAR(1:3:9): interest rate ($r$)



(b) VAR(3): unemployment ($u$) and VAR(3): investment ($i$)



(c) VAR(1): inflation ($p$) and VAR(1): unemployment ($u$)

Figure 7: US macroeconomic forecasting 2001/1-2015/12: Dynamic retrospective posterior correlations of latent agent factors over 2001/1-2015/12.
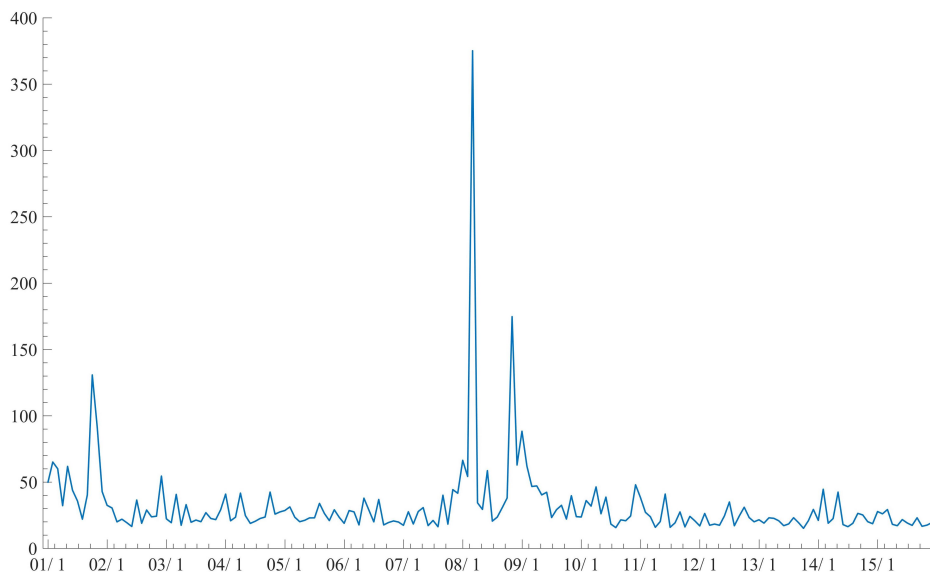
Figure 8: US macroeconomic forecasting 2001/1-2015/12: Retrospective Kullback-Leibler divergence between prior agent forecasts and posterior agent forecasts computed at each of the $t = 1{:}180$ months.

Figure 9: US macroeconomic forecasting 2001/1-2015/12: Mean squared 12-step ahead forecast errors MSFE$_{1:t}(12)$ of inflation ($p$) sequentially revised at each of the $t = 1{:}180$ months.



Figure 10: US macroeconomic forecasting 2001/1-2015/12: 12-step ahead log predictive density ratios LPDR$_{1:t}(12)$ sequentially revised at each of the $t = 1{:}180$ months. The baseline at 0 over all $t$ corresponds to BPS(12).

34

Figure 11: US macroeconomic forecasting 2001/1-2015/12: Mean squared 24-step ahead forecast errors MSFE$_{1:t}(24)$ of inflation ($p$) sequentially revised at each of the $t = 1{:}180$ months.



Figure 12: US macroeconomic forecasting 2001/1-2015/12: 24-step ahead log predictive density ratios LPDR$_{1:t}(24)$ sequentially revised at each of the $t = 1{:}180$ months. The baseline at 0 over all $t$ corresponds to BPS(24).
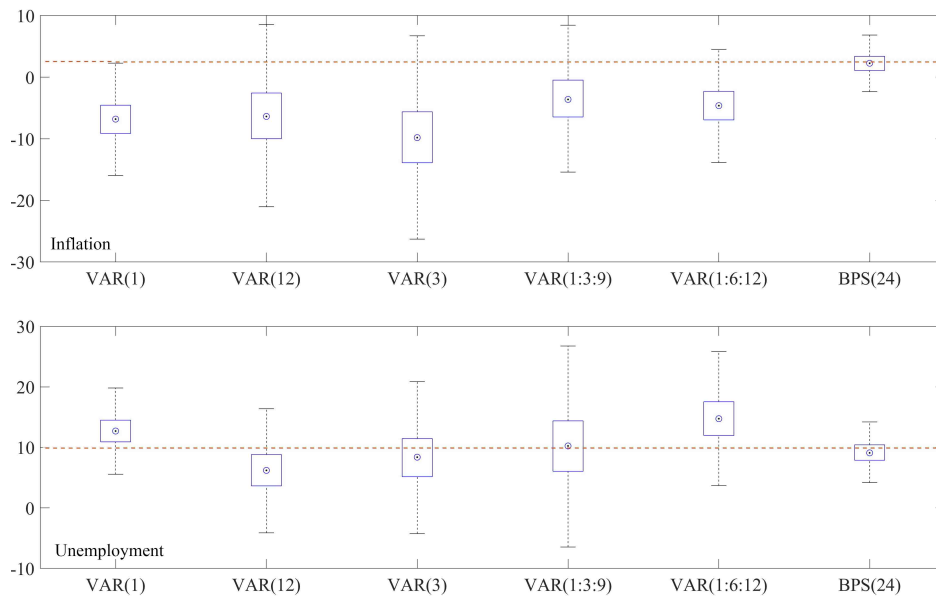
Figure 13: US macroeconomic forecasting 2001/1-2015/12: Comparison of distributional forecasts of inflation $(p)$ and unemployment $(u)$ using scenario analysis at 2008/10 with conditional 24-step ahead forecasts under zero interest rate policy.

<div align="center">

# Multivariate Bayesian Predictive Synthesis
# in Macroeconomic Forecasting

Kenichiro McAlinn, Knut Are Aastveit, Jouchi Nakajima, & Mike West

## Supplementary Material

</div>

## 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 2.4. This involves a standard set of steps in a customized three-component block Gibbs sampler: the first component samples the latent agent states, the second, the second samples the dynamic BPS model states/parameters, and the third component samples the observation variance. The latter two involves a modified 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 analysis based on data over times $t = 1{:}T$ for any chosen $T$.

Standing at time $t = 0$, the decision maker has initial information summarized via in terms of $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}_0, \boldsymbol{C}_0)$ and $\boldsymbol{V}_0 \sim IW(n_0, \boldsymbol{D}_0)$, independently. Here the $q \times q$ variance matrix $\boldsymbol{V}_0$ has the inverse Wishart distribution with $n > 0$ degrees of freedom and prior "sum-of-squares" matrix $\boldsymbol{D}_0$. Equivalently, the precision matrix $\boldsymbol{V}_0^{-1} \sim W(h_0, \boldsymbol{D}_0^{-1})$, the Wishart distribution with $h_0 = n_0 + q - 1$ and mean $h_0 \boldsymbol{D}_0^{-1}$ so that the initial estimate $\boldsymbol{D}_0/h_0$ is the prior harmonic mean of $\boldsymbol{V}_0$. Model specification is completed with two chosen discount factors: $\beta$, defining the extent of time variation in the evolution of the states $\boldsymbol{\theta}_t$, and $\delta$ defining levels of variation in the evolution of the volatility matrices $\boldsymbol{V}_t$.

At time $T$, the decision maker has accrued information $\{y_{1:T}, \mathcal{H}_{1:T}\}$. The MCMC analysis is then run iteratively as follows.

**Initialization:**  First, initialize by setting

$$
\boldsymbol{F}_t = \begin{pmatrix}
1 & \boldsymbol{f}'_{t1} & 0 & \boldsymbol{0}' & \cdots & \cdots & 0 & \boldsymbol{0}' \\
0 & \boldsymbol{0}' & 1 & \boldsymbol{f}'_{t2} & & & & \vdots \\
\vdots & & & & \ddots & & & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & 1 & \boldsymbol{f}'_{tq}
\end{pmatrix},
\tag{8}
$$

for each $t = 1{:}T$, with elements set at some chosen initial values of the latent agent states. Initial values can be chosen arbitrarily. One obvious and appropriate choice– our recommended default

<div align="center">

1

</div>

choice– is to simply generate agent states from their priors, i.e., from the agent forecast distributions, $\boldsymbol{x}_{tj} \sim h_{tj}(\boldsymbol{x}_{tj})$ 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 $\boldsymbol{x}_{tj} = \boldsymbol{y}_t$ for each $t, j$, 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. Given initial values of agent factor vector $\boldsymbol{x}_{tj} = (x_{t1j}, x_{t2j}, ..., x_{tqj})'$ for each agent $j = 1{:}J$ and each time $t$, the $\boldsymbol{F}_t$ matrices are initialized with series-specific row entries from $\boldsymbol{f}_{tr} = (x_{tr1}, x_{tr2}, ..., x_{trJ})'$ for each $r = 1{:}q$.

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

Following initialization, the MCMC iterates repeatedly to resample three sets of conditional posteriors to generate the MCMC samples from the target posterior $p(\boldsymbol{X}_{1:T}, \boldsymbol{\theta}_{1:T}, \boldsymbol{V}_{1:T} | \boldsymbol{y}_{1:T}, \mathcal{H}_{1:T})$. These conditional posteriors and algorithmic details of their simulation are as follows.

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

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

$$\boldsymbol{y}_t = \boldsymbol{F}_t \boldsymbol{\theta}_t + \boldsymbol{\nu}_t, \quad \nu_t \sim N(0, \boldsymbol{V}_t),$$
$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \omega_t \sim N(0, \boldsymbol{W}_t),$$

has known elements $\boldsymbol{F}_t, \boldsymbol{W}_t$ and specified initial prior at $t = 0$. The implied conditional posterior for $\boldsymbol{\theta}_{1:T}$ then does not depend on $\mathcal{H}_{1:T}$, reducing to $p(\boldsymbol{\theta}_{1:T} | \boldsymbol{X}_{1:T}, \boldsymbol{V}_{1:T}, \boldsymbol{y}_{1:T})$. This is simulated using the efficient and standard FFBS algorithm (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 $p(\boldsymbol{\theta}_t | \boldsymbol{X}_{1:t}, \boldsymbol{V}_{1:t}, \boldsymbol{y}_{1:t})$ at each $t$. The summary technical details are as follows:
>
> 1. *Time $t-1$ posterior:*
>
> $$\boldsymbol{\theta}_{t-1} | \boldsymbol{X}_{1:t-1}, \boldsymbol{V}_{1:t-1}, \boldsymbol{y}_{1:t-1} \sim N(\boldsymbol{m}_{t-1}, \boldsymbol{C}_{t-1}),$$
>
> with point estimate $\boldsymbol{m}_{t-1}$ of $\boldsymbol{\theta}_{t-1}$.
>
> 2. *Update to time $t$ prior:*
>
> $$\boldsymbol{\theta}_t | \boldsymbol{X}_{1:t-1}, \boldsymbol{V}_{1:t-1}, \boldsymbol{y}_{1:t-1} \sim N(\boldsymbol{m}_{t-1}, \boldsymbol{R}_t) \quad \text{with} \quad \boldsymbol{R}_t = \boldsymbol{C}_{t-1}/\delta,$$

2

with (unchanged) point estimates $\boldsymbol{m}_{t-1}$ of $\boldsymbol{\theta}_t$, 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:* $\boldsymbol{y}_t | \boldsymbol{X}_{1:t}, \boldsymbol{V}_{1:t}, \boldsymbol{y}_{1:t-1} \sim T_{\beta n_{t-1}}(\boldsymbol{f}_t, \boldsymbol{Q}_t)$ where

$$\boldsymbol{f}_t = \boldsymbol{F}_t \boldsymbol{m}_{t-1} \quad \text{and} \quad \boldsymbol{Q}_t = \boldsymbol{F}_t \boldsymbol{R}_t \boldsymbol{F}'_t + \boldsymbol{V}_t.$$

4. *Filtering update to time $t$ posterior:*

$$\boldsymbol{\theta}_t | \boldsymbol{V}_{1:t}, \boldsymbol{X}_{1:t}, \boldsymbol{y}_{1:t} \sim N(\boldsymbol{m}_t, \boldsymbol{C}_t),$$

with defining parameters $\boldsymbol{m}_t = \boldsymbol{m}_{t-1} + \boldsymbol{A}_t \boldsymbol{e}_t$ and $\boldsymbol{C}_t = \boldsymbol{R}_t - \boldsymbol{A}_t \boldsymbol{Q}_t \boldsymbol{A}'_t$, based on 1-step forecast error $\boldsymbol{e}_t = \boldsymbol{y}_t - \boldsymbol{f}_t$ and the state adaptive coefficient vector (a.k.a. "Kalman gain") $\boldsymbol{A}_t = \boldsymbol{R}_t \boldsymbol{F}'_t \boldsymbol{Q}_t^{-1}$.

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

a. *At time $T$:* Simulate $\boldsymbol{\theta}_T$ from the final multivariate normal posterior

$$p(\boldsymbol{\theta}_T | \boldsymbol{X}_{1:T}, \boldsymbol{V}_{1:T}, \boldsymbol{y}_{1:T}) = N(\boldsymbol{m}_T, \boldsymbol{C}_T).$$

b. *Recurse back over times $t = T - 1, T - 2, \ldots, 0$:* At each time $t$, simulate the state $\boldsymbol{\theta}_t$ from the conditional posterior $p(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t+1}, \boldsymbol{X}_{1:t}, \boldsymbol{V}_{1:t}, \boldsymbol{y}_{1:t})$; this is multivariate normal with mean vector $\boldsymbol{m}_t + \delta(\boldsymbol{\theta}_{t+1} - \boldsymbol{m}_t)$ and variance matrix $\boldsymbol{C}_t(1 - \delta)$.

### A.2.2 Per MCMC Iterate Step 2: Sampling BPS DLM parameters $\boldsymbol{V}_{1:T}$

Conditional on the sampled values of the BPS DLM parameters $\boldsymbol{\theta}_{1:T}$ and latent agent states $\boldsymbol{X}_{1:T}$, the next step in the MCMC iterate samples the full conditional posterior of the sequence of volatility matrices, generating a draw from $p(\boldsymbol{V}_{1:T} | \boldsymbol{X}_{1:T}, \boldsymbol{\theta}_{1:T}, \boldsymbol{y}_{1:T})$.

**Forward filtering:** For each $t = 1{:}T$ in sequence, update and save the forward filtering summaries $(n_t, \boldsymbol{D}_t)$ of on-line posteriors

$$\boldsymbol{V}_t | \boldsymbol{X}_{1:t}, \boldsymbol{\theta}_{1:t}, \boldsymbol{y}_{1:t} \sim IW(n_t, \boldsymbol{D}_t),$$

given by $n_t = h_t - q + 1$ where $h_t = \beta h_{t-1} + 1$, and $\boldsymbol{D}_t = \beta \boldsymbol{D}_{t-1} + (\boldsymbol{y}_t - \boldsymbol{F}'_t \boldsymbol{\theta}_t)(\boldsymbol{y}_t - \boldsymbol{F}'_t \boldsymbol{\theta}_t)'$.

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

a. *At time $T$:* Simulate $\boldsymbol{V}_T$ from the final inverse Wishart posterior $IW(n_T, \boldsymbol{D}_T)$.

3

b. *Recurse back over times* $t = T-1, T-2, \ldots, 0$ : At time $t$, sample $\boldsymbol{V}_t$ from the conditional posterior $p(\boldsymbol{V}_t | \boldsymbol{V}_{t+1}, \boldsymbol{X}_{1:t}, \boldsymbol{\theta}_{1:t}, \boldsymbol{y}_{1:t})$. Algorithmically, this is achieved via

$$\boldsymbol{V}_t^{-1} = \beta \boldsymbol{V}_{t+1}^{-1} + \boldsymbol{\Upsilon}_t \quad \text{where} \quad \boldsymbol{\Upsilon}_t \sim W((1-\beta)h_t, \boldsymbol{D}_t^{-1}),$$

and where the $\Upsilon_t$ are independent over $t$.

### A.2.3 Per MCMC Iterate Step 3: Sampling the latent agent states $\boldsymbol{X}_{1:T}$

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

$$p(\boldsymbol{X}_t | \boldsymbol{\Phi}_t, \boldsymbol{y}_t, \mathcal{H}_t) \propto N(\boldsymbol{y}_t | \boldsymbol{F}_t \boldsymbol{\theta}_t, \boldsymbol{V}_t) \prod_{j=1:J} h_{tj}(\boldsymbol{x}_{tj}). \tag{9}$$

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

1. *Multivariate normal agent forecast densities:* In cases when each of the agent forecast densities is normal, the posterior in eqn. (9) yields a multivariate normal distribution for vectorized $\boldsymbol{X}_t$. Computation of its defining parameters and then drawing a new sample vector $\boldsymbol{X}_t$ are trivial.

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 five agents represent conjugate exchangeable dynamic linear models in which all forecast densities are multivariate T, with parameters varying over time and with step-ahead forecast horizon. In such cases, standard Bayesian augmentation methods apply to enable simulation. Each multivariate T distribution is expressed as a scale mixture of multivariate normals, with the mixing scale parameters introduced as inherent latent variables with inverse gamma distributions. This expansion of the parameter space makes the multivariate T distributions conditional multivariate normals, and the mixing scales are resampled (from implied conditional posterior inverse gamma distributions) 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 $\boldsymbol{X}_t$ reduces technically to that conditional normals above.

   Specifically, suppose that $h_{tj}(\boldsymbol{x}_{tj})$ is density of the normal $T_{n_{tj}}(\boldsymbol{h}_{tj}, \boldsymbol{H}_{tj})$; the notation means that $(\boldsymbol{x}_{tj} - \boldsymbol{h}_{tj})/\sqrt{\boldsymbol{H}_{tj}}$ has a standard multivariate 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 $\boldsymbol{x}_{tj}$ has a conditional multivariate normal density $\boldsymbol{x}_{tj} | \phi_{tj} \sim N(\boldsymbol{h}_{tj}, \boldsymbol{H}_{tj}/\phi_{tj})$ 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 normal simulations conditional on the $\phi_{tj}$. Following this, we resample values of the $\phi_{tj}$ from their trivially implied conditional gamma posteriors.

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 matrix $\boldsymbol{X}_t^{(i)}$ for $i = 1{:}I$. We can apply importance sampling as follows: *(a)* compute the marginal likelihood values $p(\boldsymbol{y}_t|\boldsymbol{\Phi}_t, \boldsymbol{X}_t^{(i)}, \mathcal{H}_t)$ for each $i = 1{:}I$; *(b)* compute and normalize the implied importance sampling weights $w_{ti} \propto N(\boldsymbol{y}_t|\boldsymbol{\Phi}_t, \boldsymbol{X}_t^{(i)}, \mathcal{H}_t)$, and then *(c)* resample latent agent states for this MCMC stage according to the probabilities these weights define.

# B  Appendix: Additional Graphical Summaries from Macroeconomic Analysis

This appendix lays out additional graphical summaries of results from the macroeconomic forecasting analysis in the paper, providing material supplementary to that discussed in Section 3.
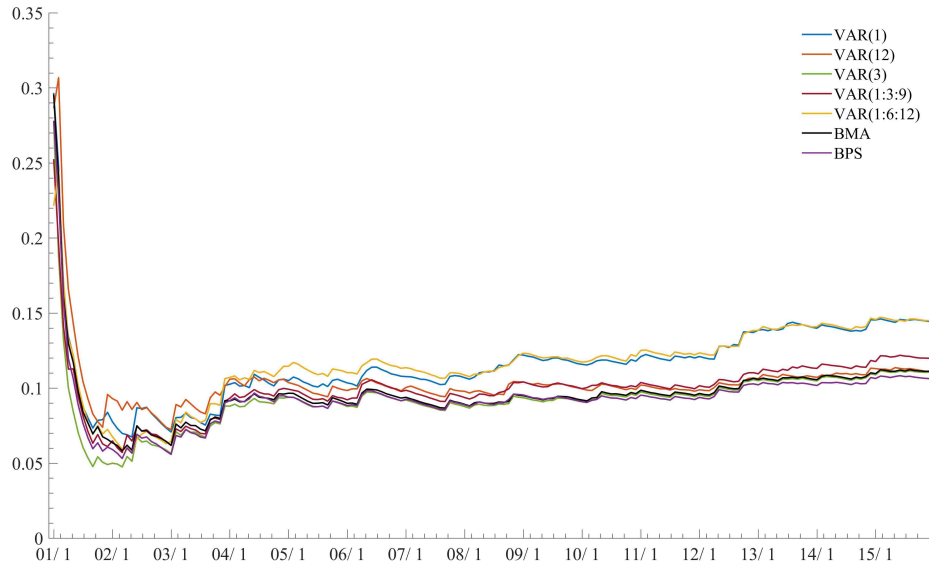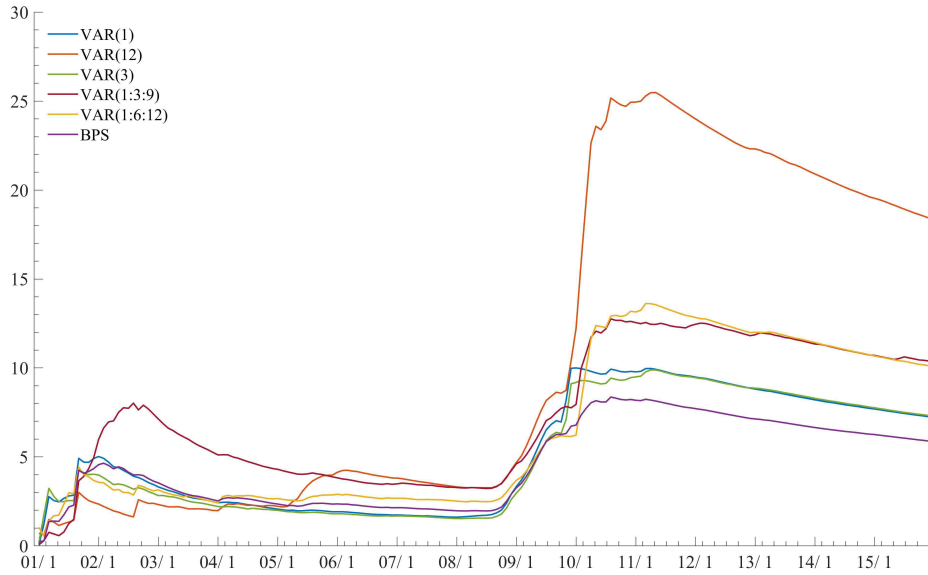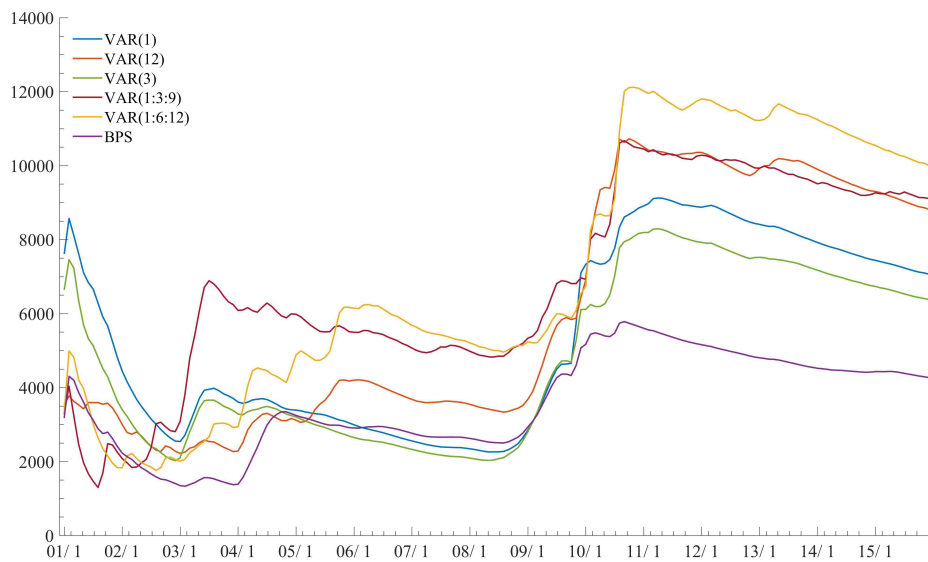


Figure C1: US macroeconomic forecasting 2001/1-2015/12: Mean squared 1-step ahead forecast errors $\text{MSFE}_{1:t}(1)$ of wage $(w)$ sequentially revised at each of the $t = 1:180$ months.
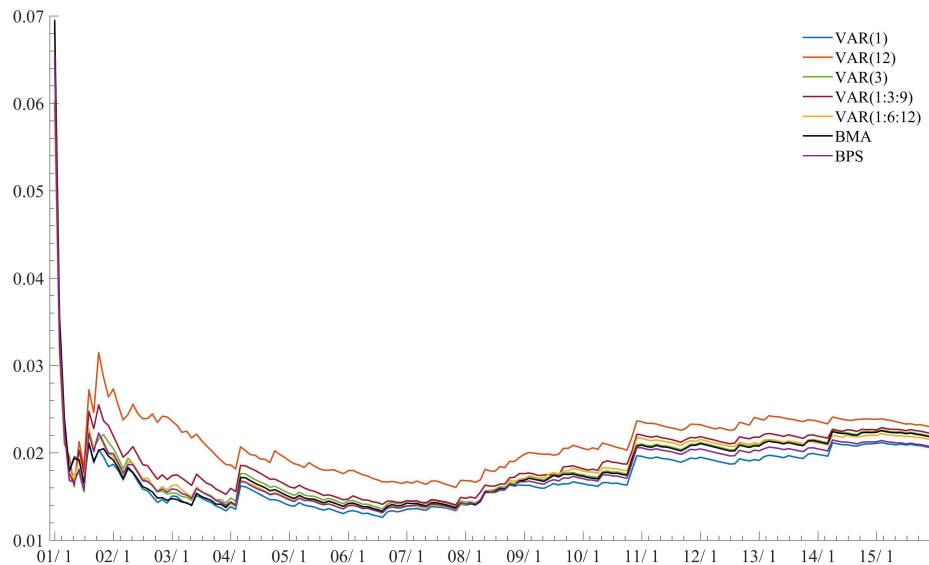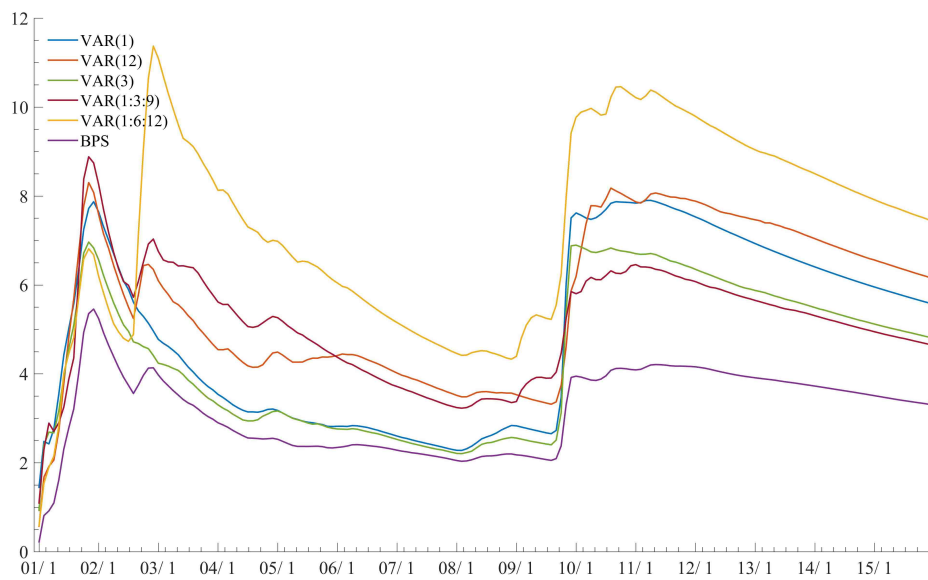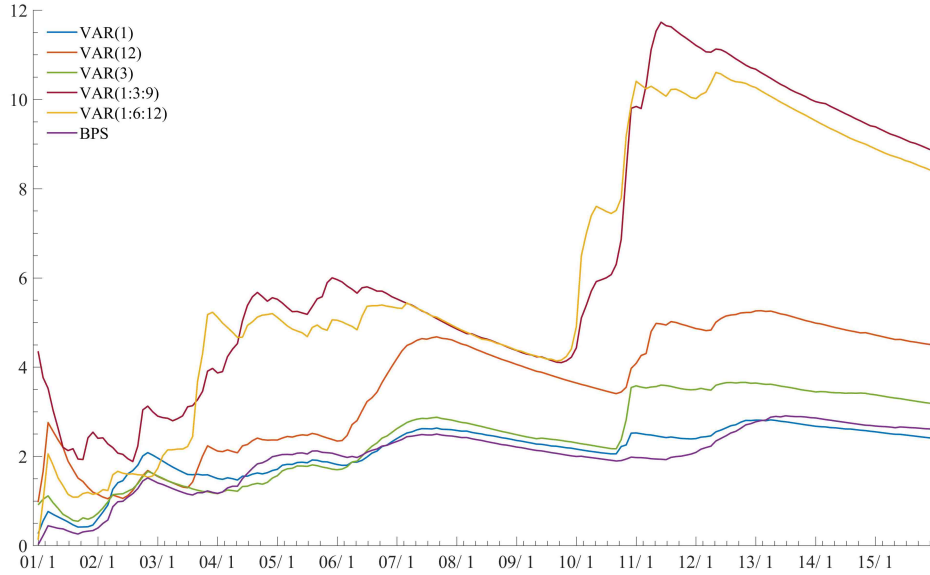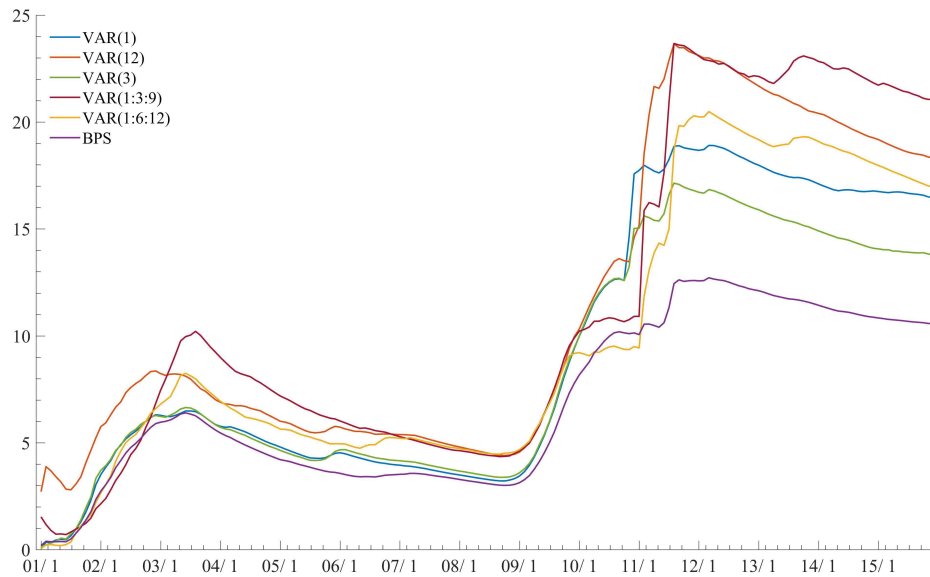


Figure C2: US macroeconomic forecasting 2001/1-2015/12: Mean squared 1-step ahead forecast errors $\text{MSFE}_{1:t}(1)$ of unemployment rate $(u)$ sequentially revised at each of the $t = 1:180$ months.
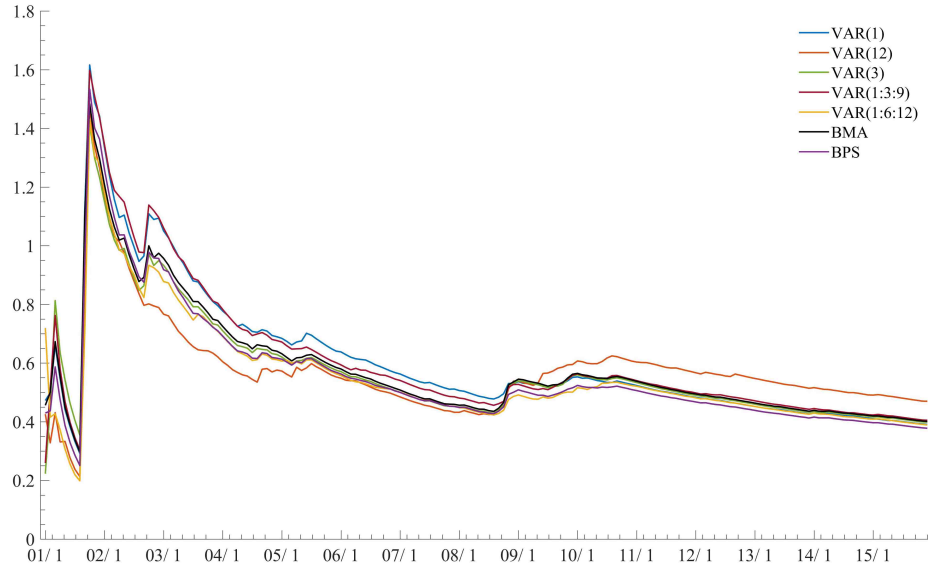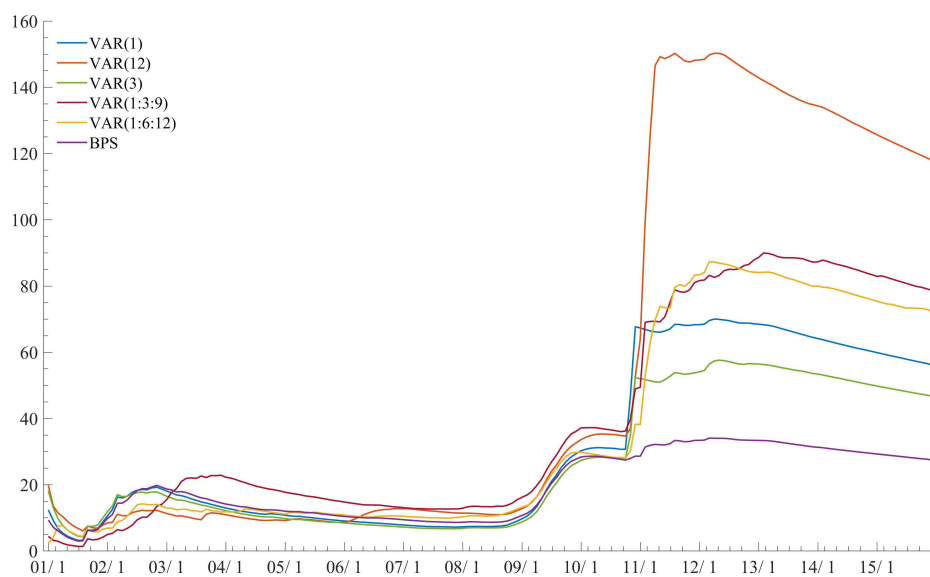
Figure C3: US macroeconomic forecasting 2001/1-2015/12: Mean squared 1-step ahead forecast errors $\text{MSFE}_{1:t}(1)$ of consumption $(c)$ sequentially revised at each of the $t = 1{:}180$ months.
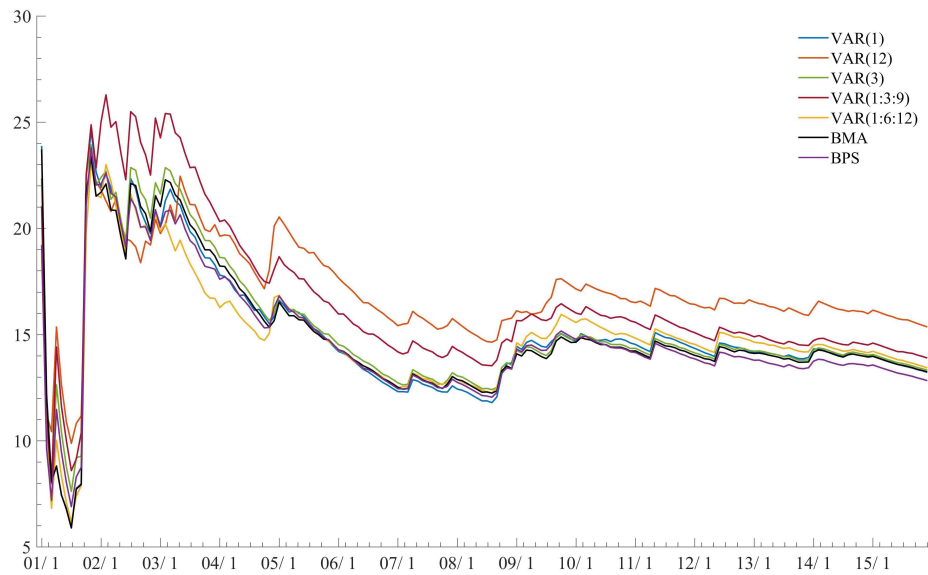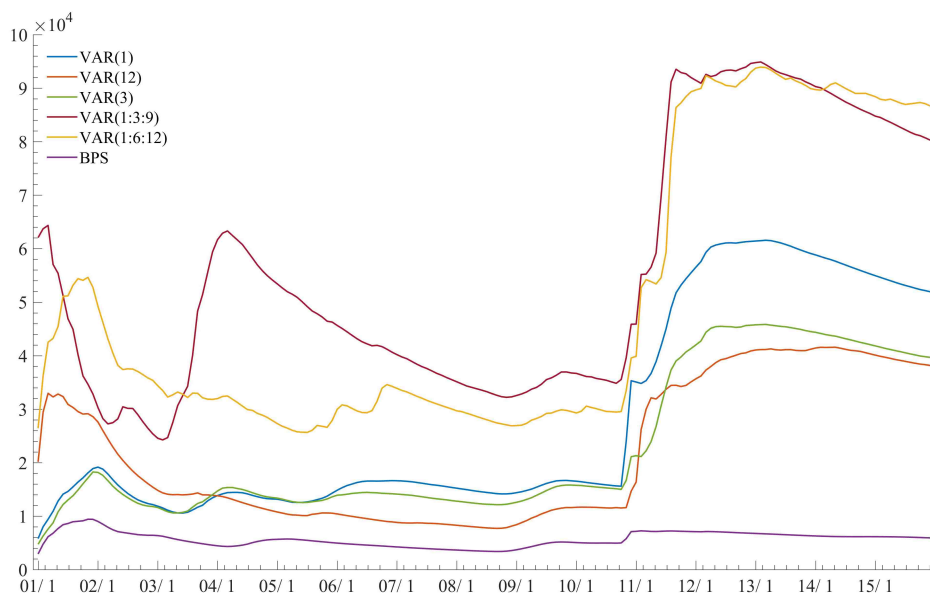


Figure C4: US macroeconomic forecasting 2001/1-2015/12: Mean squared 1-step ahead forecast errors $\text{MSFE}_{1:t}(1)$ of investment $(i)$ sequentially revised at each of the $t = 1{:}180$ months.
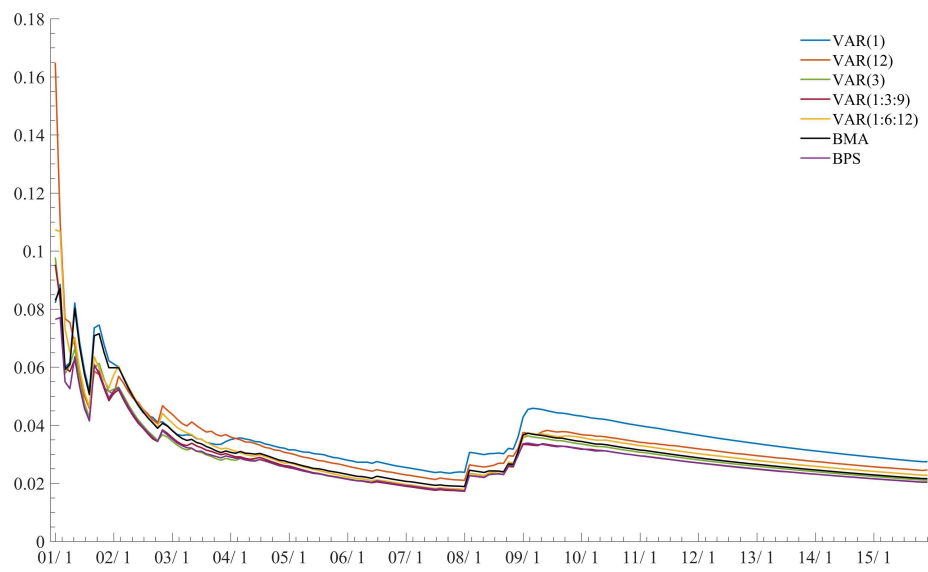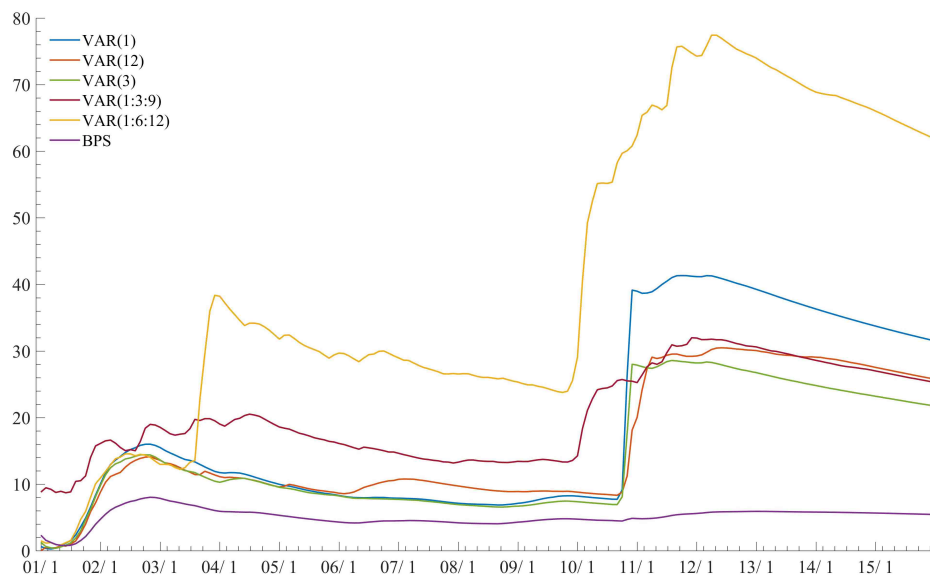
Figure C5: US macroeconomic forecasting 2001/1-2015/12: Mean squared 1-step ahead forecast errors $\text{MSFE}_{1:t}(1)$ of interest rate $(r)$ sequentially revised at each of the $t = 1{:}180$ months.
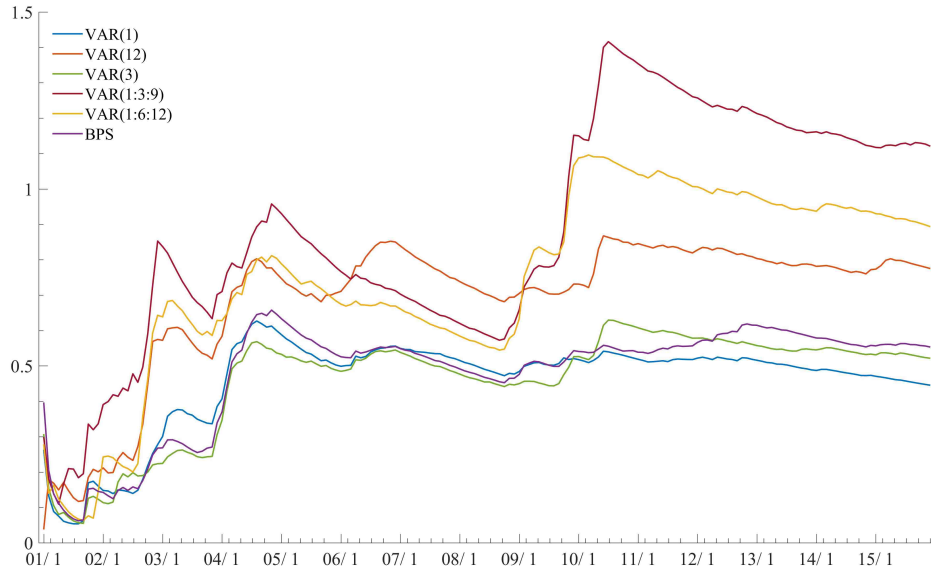
Figure C6: US macroeconomic forecasting 2001/1-2015/12: Mean squared 12-step ahead forecast errors $\text{MSFE}_{1:t}(12)$ of wage $(w)$ sequentially revised at each of the $t = 1:180$ months.
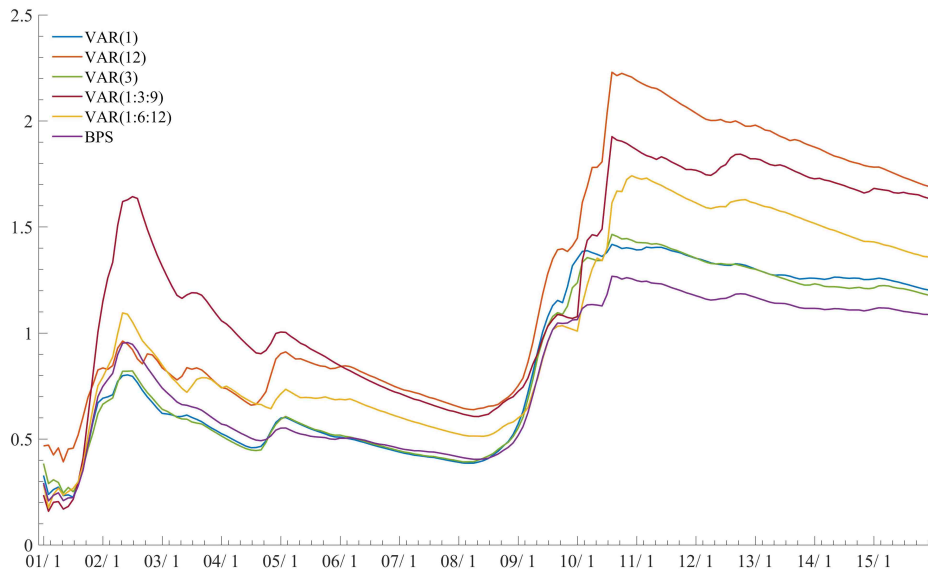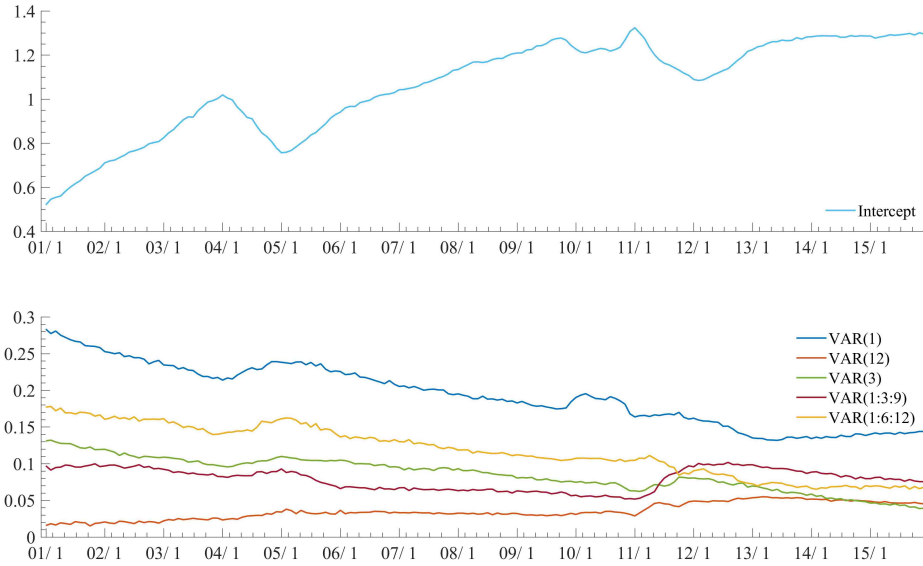


Figure C7: US macroeconomic forecasting 2001/1-2015/12: Mean squared 12-step ahead forecast errors $\text{MSFE}_{1:t}(12)$ of unemployment rate $(u)$ sequentially revised at each of the $t = 1:180$ months.

Figure C8: US macroeconomic forecasting 2001/1-2015/12: Mean squared 12-step ahead forecast errors $\text{MSFE}_{1:t}(12)$ of consumption $(c)$ sequentially revised at each of the $t = 1{:}180$ months.



Figure C9: US macroeconomic forecasting 2001/1-2015/12: Mean squared 12-step ahead forecast errors $\text{MSFE}_{1:t}(12)$ of investment $(i)$ sequentially revised at each of the $t = 1{:}180$ months.

Figure C10: US macroeconomic forecasting 2001/1-2015/12: Mean squared 12-step ahead forecast errors $\text{MSFE}_{1:t}(12)$ of interest rate $(r)$ sequentially revised at each of the $t = 1{:}180$ months.

Figure C11: US macroeconomic forecasting 2001/1-2015/12: Mean squared 24-step ahead forecast errors $\text{MSFE}_{1:t}(24)$ of wage $(w)$ sequentially revised at each of the $t = 1{:}180$ months.



Figure C12: US macroeconomic forecasting 2001/1-2015/12: Mean squared 24-step ahead forecast errors $\text{MSFE}_{1:t}(24)$ of unemployment rate $(u)$ sequentially revised at each of the $t = 1{:}180$ months.

Figure C13: US macroeconomic forecasting 2001/1-2015/12: Mean squared 24-step ahead forecast errors $\text{MSFE}_{1:t}(24)$ of consumption ($c$) sequentially revised at each of the $t = 1:180$ months.



Figure C14: US macroeconomic forecasting 2001/1-2015/12: Mean squared 24-step ahead forecast errors $\text{MSFE}_{1:t}(24)$ of investment ($i$) sequentially revised at each of the $t = 1:180$ months.

13

Figure C15: US macroeconomic forecasting 2001/1-2015/12: Mean squared 24-step ahead fore-cast errors $\mathrm{MSFE}_{1:t}(24)$ of interest rate $(r)$ sequentially revised at each of the $t = 1:180$ months.

14

Figure C16: US macroeconomic forecasting 2001/1-2015/12: On-line posterior means of BPS(12) model coefficients for inflation ($p$) sequentially computed at each of the $t = 1{:}180$ months.
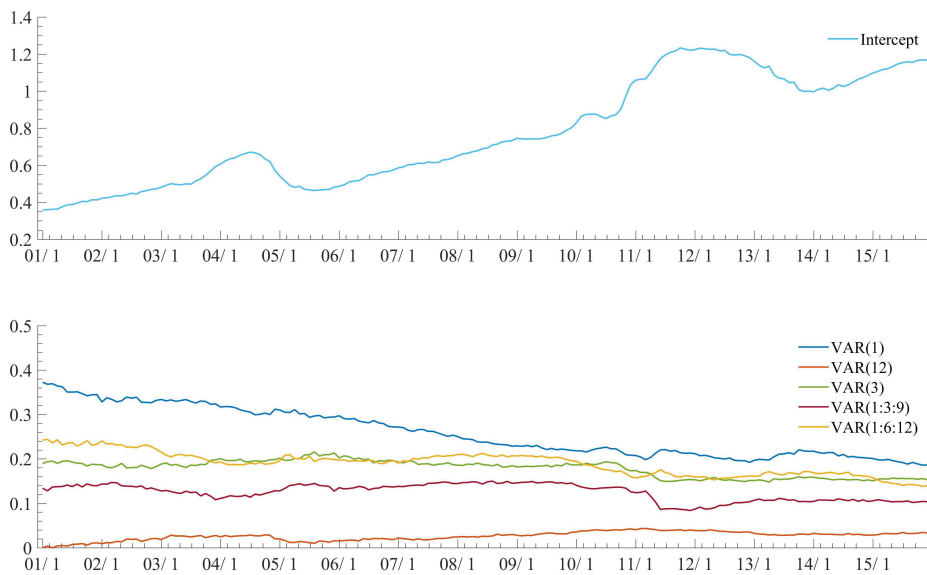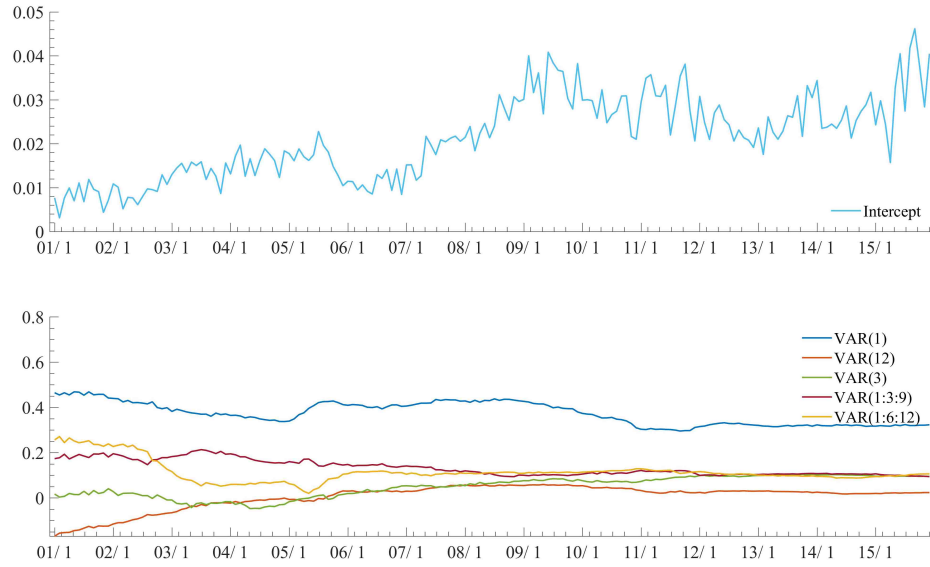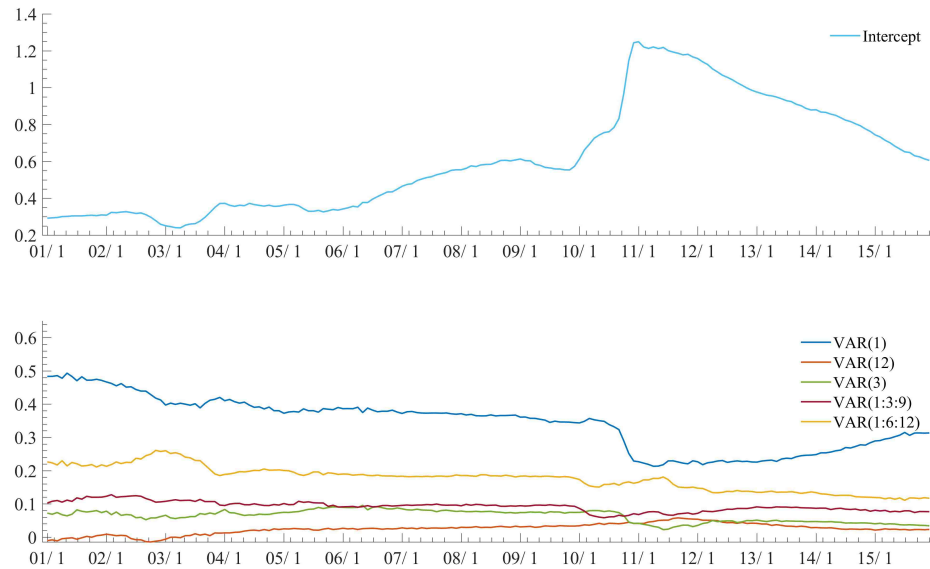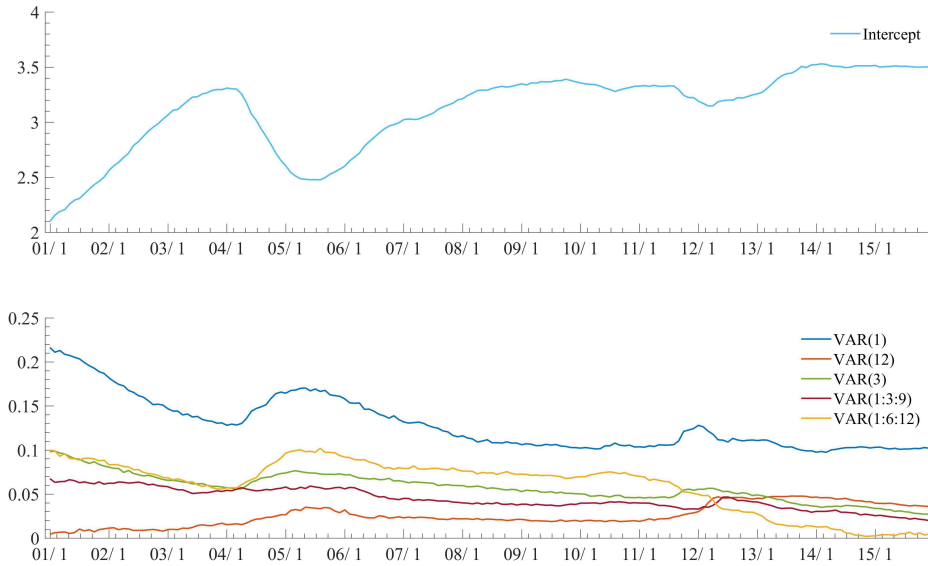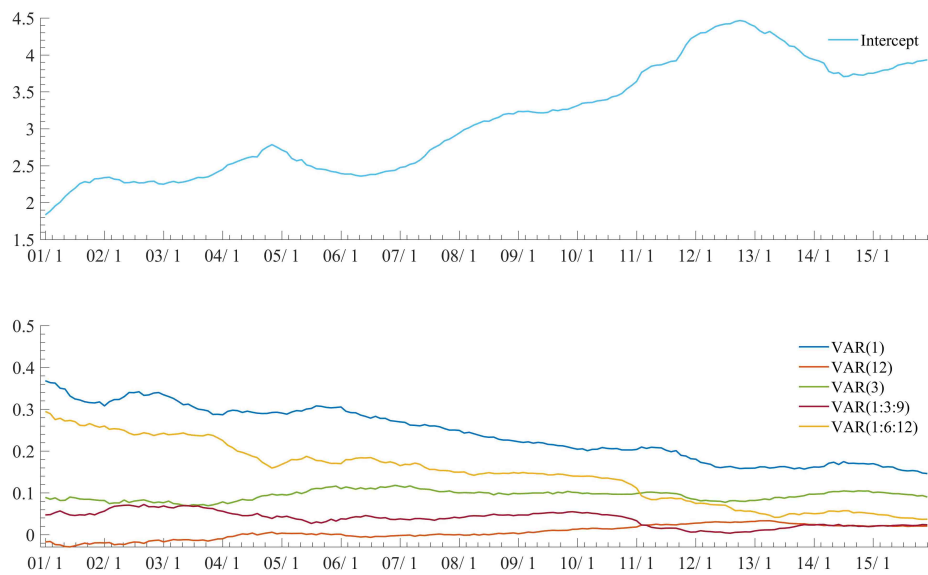


Figure C17: US macroeconomic forecasting 2001/1-2015/12: On-line posterior means of BPS(12) model coefficients for wage ($w$)sequentially computed at each of the $t = 1{:}180$ months.
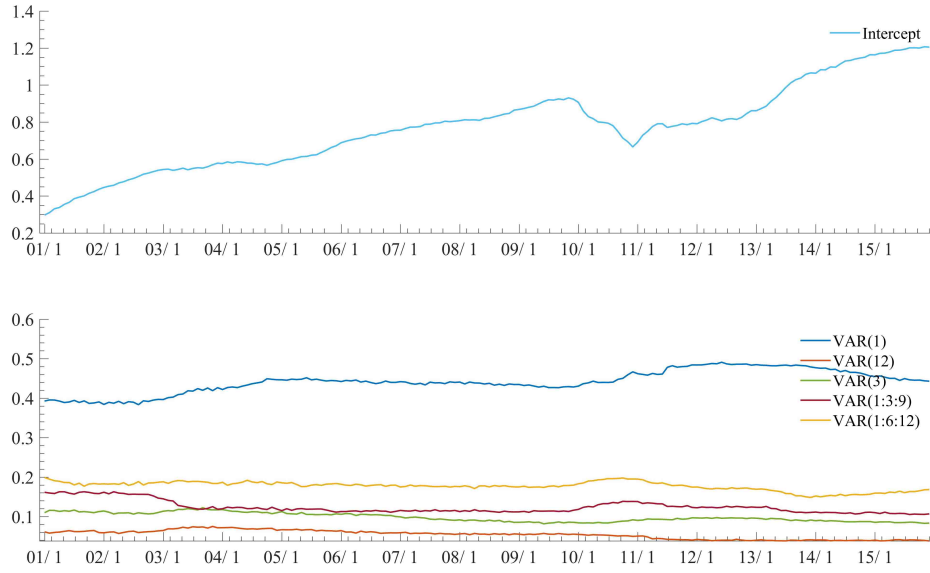
Figure C18: US macroeconomic forecasting 2001/1-2015/12: On-line posterior means of BPS(12) model coefficients for unemployment ($u$) sequentially computed at each of the $t = 1:180$ months.
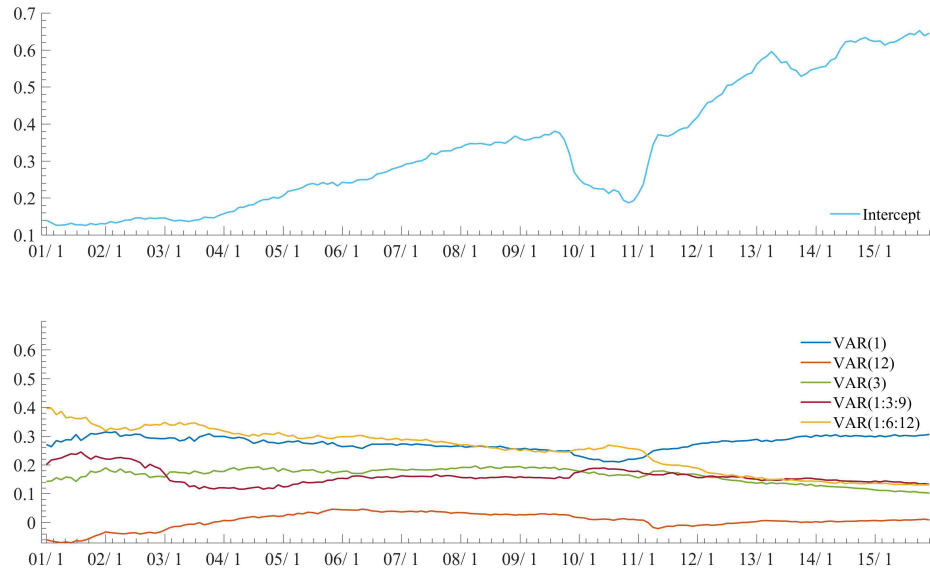


Figure C19: US macroeconomic forecasting 2001/1-2015/12: On-line posterior means of BPS(12) model coefficients for consumption ($c$) sequentially computed at each of the $t = 1:180$ months.

Figure C20: US macroeconomic forecasting 2001/1-2015/12: On-line posterior means of BPS(12) model coefficients for investment ($i$) sequentially computed at each of the $t = 1{:}180$ months.



Figure C21: US macroeconomic forecasting 2001/1-2015/12: On-line posterior means of BPS(12) model coefficients for interest rate ($r$) sequentially computed at each of the $t = 1{:}180$ months.

17

Figure C22: US macroeconomic forecasting 2001/1-2015/12: On-line posterior means of BPS(24) model coefficients for inflation $(p)$ sequentially computed at each of the $t = 1{:}180$ months.



Figure C23: US macroeconomic forecasting 2001/1-2015/12: On-line posterior means of BPS(24) model coefficients for wage $(w)$ sequentially computed at each of the $t = 1{:}180$ months.
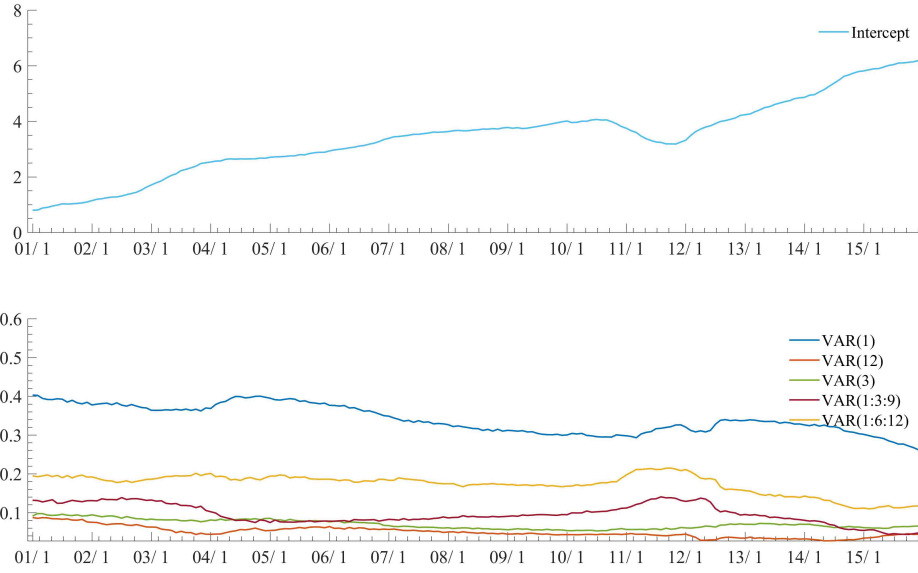
18

Figure C24: US macroeconomic forecasting 2001/1-2015/12: On-line posterior means of BPS(24) model coefficients for unemployment ($u$) sequentially computed at each of the $t = 1{:}180$ months.
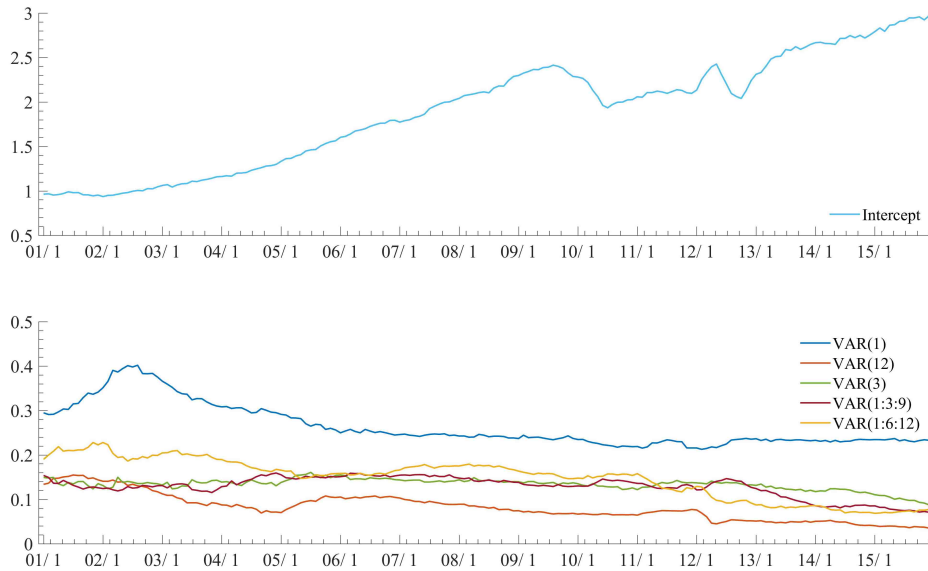


Figure C25: US macroeconomic forecasting 2001/1-2015/12: On-line posterior means of BPS(24) model coefficients for consumption ($c$) sequentially computed at each of the $t = 1{:}180$ months.
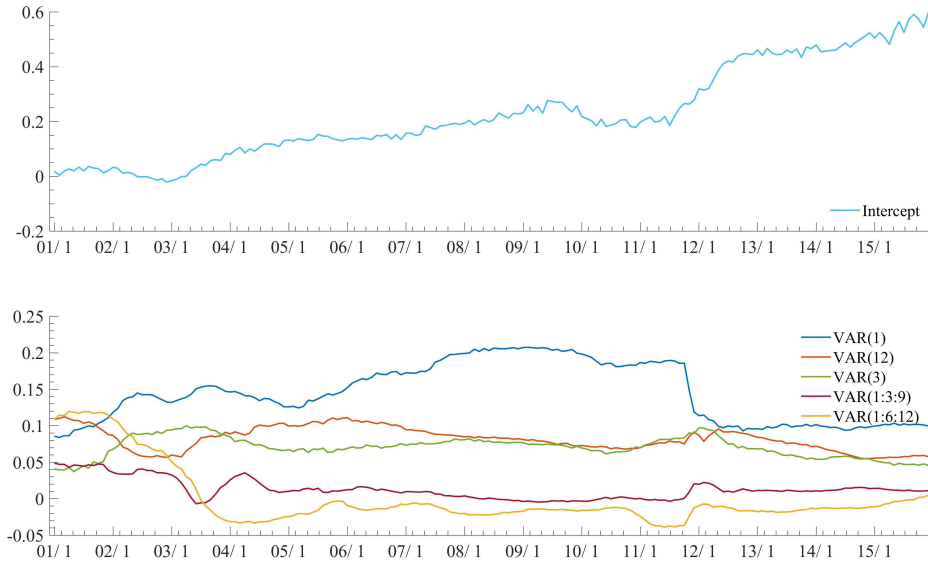
Figure C26: US macroeconomic forecasting 2001/1-2015/12: On-line posterior means of BPS(24) model coefficients for investment ($i$) sequentially computed at each of the $t = 1{:}180$ months.
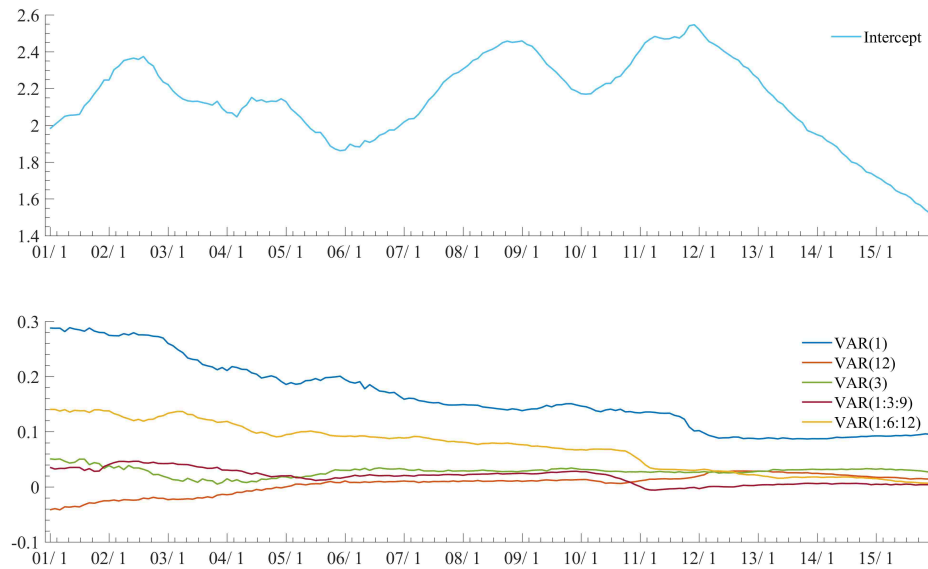


Figure C27: US macroeconomic forecasting 2001/1-2015/12: On-line posterior means of BPS(24) model coefficients for interest rate ($r$) sequentially computed at each of the $t = 1{:}180$ months.

**Centre for Applied Macroeconomics and Commodity Prices (CAMP)**
will bring together economists working on applied macroeconomic issues, with special
emphasis on petroleum economics.

BI Norwegian Business School
Centre for Applied Macro - Petroleum economics (CAMP)
N-0442 Oslo

http://www.bi.no/camp