Dynamic Modeling and Bayesian Predictive Synthesis

by

Kenichiro McAlinn

Department of Statistical Science
Duke University

Date: ______________________

Approved:

____
Mike West, Supervisor

____
James O. Berger

____
Andrew Cron

____
Carl F. Mela

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Statistical Science
in the Graduate School of Duke University
2017

ABSTRACT

Dynamic Modeling and Bayesian Predictive Synthesis

by

Kenichiro McAlinn

Department of Statistical Science
Duke University

Date: ______________________
Approved:

____
Mike West, Supervisor

____
James O. Berger

____
Andrew Cron

____
Carl F. Mela

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistical Science in the Graduate School of Duke University
2017

Abstract

This dissertation discusses model and forecast comparison, calibration, and combination from a foundational perspective. For nearly five decades, the field of forecast combination has grown exponentially. Its practicality and effectiveness in important real world problems concerning forecasting, uncertainty, and decisions propels this. Ample research– theoretical and empirical– into new methods and justifications have been produced. However, its foundations– the philosophical/theoretical underpinnings up0n which methods and strategies are built– have been unexplored in recent literature. *Bayesian predictive synthesis* (BPS) defines a coherent theoretical basis for combining multiple forecast densities, whether from models, individuals, or other sources, and generalizes existing forecast pooling and Bayesian model mixing methods. By understanding the underlying foundation that defines the combination of forecasts, multiple extensions are revealed, resulting in significant advances in the understanding and efficacy of the methods for decision making in multiple fields.

The extensions discussed in this dissertation are into the temporal domain. Many important decision problems involve time series, including policy decisions in macroeconomics and investment decisions in finance, where decisions are sequentially updated over time. Time series extensions of BPS are implicit dynamic latent factor models, allowing adaptation to time-varying biases, miscalibration, and dependencies among models or forecasters. Multiple studies using different

data and different decision problems are presented, demonstrating the effectiveness of dynamic BPS, in terms of forecast accuracy and improved decision making, and highlighting the unique insight it provides.

Chapter 1 introduces the idea of Bayesian predictive synthesis and outlines this dissertation. Chapter 2 formulates the theoretical foundations of BPS and extends it to dynamic decisions, as well as Bayesian computation that enables BPS. Chapter 3 provides a study of forecasting US inflation using univariate BPS. Chapter 4 provides a study of multiple US macroeconomic time series using multivariate BPS. Chapter 5 applies BPS to studies of return predictability in finance. Chapter 6 extends the BPS framework to solve problems of temporal misspecification, using applications in finance. Chapter 7 extends the BPS framework for mixed frequency forecasting and provides a study of nowcasting quarterly GDP using monthly macroeconomic data. Finally, Chapter 8 concludes the dissertation with open questions and discussion of potential future research.

To Rika

Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

Symbols

$\mathcal{D}$	A decision maker.
$\mathcal{A}$	An agent.
$\mathcal{H}$	An information set.
$\boldsymbol{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Normal distribution.
$v \sim G(a, b)$	Gamma distribution.
$\boldsymbol{V} \sim W^{-1}(c, \boldsymbol{D})$	Inverse Wishart distribution.
$diag(\cdot)$	Diagonal matrix.
$s : t$	Indices $s, s+1, ..., t-1, t$.

Abbreviations

(V)AR	(Vector) Autoregressive (model).
BITS	Bayesian inter-temporal synthesis.
BMA	Bayesian Model Averaging.
BPS	Bayesian predictive synthesis.
DeCo	Density combination.
DLM	Dynamic linear model.
FASST	Frequency auto-splicing synthesis.
LPDR	Log predictive density ratios.
MCMC	Markov chain Monte Carlo.

MFS	Mixed frequency synthesis.
(U-)MIDAS	(Unrestricted-) Mixed data sampling.
MSFE	Mean squared forecast errors.
OLS	Ordinary least squares.
TV(P)	Time-varying (parameter).

Acknowledgements

First and foremost, I would like to thank my advisor, Mike West, whose invaluable insight and guidance has helped me become the researcher, academic, and (subjective) Bayesian that I am. Without his support and trust, I would not have been able to realize my research and this dissertation.

I would like to thank my dissertation committee members, Jim Berger, Andrew Cron, and Carl Mela, as well as Merlise Clyde and Robert Winkler, who were on my preliminary committee, for their time and support. Additional thanks to Christian Macaro and everyone at SAS, who supported me through their econometric fellowship. Further financial support was provided through the BEST Award from the BEST Foundation and the Data Expeditions program from the Information Initiative at Duke.

I would also like to thank my undergraduate and graduate advisors Teruo Nakatsuma, who taught me the world of Bayesian statistics, and Michael Ainge, who taught me how to become a true academic and thinker. Without these two providing me with good priors during my early graduate career, I would not have been able to be where I am today.

Thanks go out to my cohort (namely my dungeon mates, Joe, Lorin, Matt, and Victor), for (better or worse) letting me be who I am. Without them, my graduate years would surely have been bland (or worse, mediocre). I also extend my thanks to my friends, in and out of Duke, whose discussions and support helped me through the years. Additional thanks go out to the artists, thinkers, and philosophers who inhabit my head. The inspiration they have provided me formed the way I think and helped define who I am.

Special thanks to my parents, Jerry and Sachiko McAlinn, who instilled me with the foundations to be a good student. My father, through our conversations and discussions, and my mother, who kept pushing me to be curious, is what brought me to this point and all the points before and after. I could not have done this without their unconditional love and support.

Finally, my eternal gratitude goes out to my wife, Rika, who has supported me through all of my graduate years with endless patience. Without her pushing me to be better, I would not have made it this far.

1

Introduction

1.1 Bayesian analysis of predictive synthesis

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

Contemporary literature includes creative ideas for forecast density pooling, defining new empirical models fitted by Bayesian methods (e.g. Terui and van Dijk, 2002; Hall and Mitchell, 2007; Amisano and Giacomini, 2007; Hoogerheide et al., 2010; Kascha and Ravazzolo, 2010; Geweke and Amisano, 2011, 2012; Billio et al., 2012, 2013; Aastveit et al., 2014; Kapetanios et al., 2015; Aastveit et al., 2016; Pettenuzzo and Ravazzolo, 2016; Del Negro et al., 2016). Some of these direct/empirical models for forecast density calibration and combination demonstrate improved forecast performance in studies in macroeconomics and finance. These methods advance the broader field that, since the seminal paper by Bates and Granger (1969), has drawn on interests and expertise from business, economics, technology, meteorology, management science, military intelligence, seismic risk, and environmental risk, among other areas (e.g. Clemen, 1989; Clemen and Winkler, 1999; Timmermann, 2004; Clemen and Winkler, 2007). Recent interest in macroeconomics, driven partly by a need to improve information flows to policy and decision makers at national and international levels, represents development with potential to impact more broadly. The challenge is to define formal, reliable methodology for integrating predictive information from multiple professional forecasters and models, and sets of competing econometric models in a more traditional statistical context.

This dissertation responds to this movement, motivated by an interest in the question of foundational underpinnings of some of the specific algorithmic or empirical models for forecast density combination recently introduced. While a new combination rule/algorithm may demonstrate success in a specific case study, understanding potential conceptual and theoretical foundations is of interest in order to advance broader understanding– through transparency of implicit underlying assumptions– and hence open paths to possible methodological generalizations of practical import. This dissertation addresses this by linking to the historical literature on subjective Bayesian thinking about the broad field of assessing and combining subjective probabilities. In particular, we revisit Bayesian "agent/expert opinion analysis" (e.g. Lindley et al., 1979; West, 1984; Genest and Schervish, 1985; West, 1988; West and Crosse, 1992; West, 1992; Dawid et al., 1995; French, 2011), which requires a formal, subjective Bayesian context for treating multiple models or forecasters as providers of "forecast data" to be used in prior-posterior updating by a coherent decision maker (See also; West and Harrison, 1997, Sect 16.3.2).

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

Formalizing a coherent framework for predictive distributions, and in a broader sense "information," leads to broader extensions where the framework for BPS can be utilized to build new methodologies and models that solve topical problems in economics and finance. These problems include: dealing with mixed frequency data, temporal misspecification, and dimension reduction. These new methodologies are discussed and practical decision problems are explored through applications.

1.2 Outline

This dissertation is organized as follows. Chapter 2 introduces the background, foundation, and theory of BPS and how it is extended to dynamic settings for univariate and multivariate time series data. The Bayesian computation that enables dynamic BPS and a simulation study using US macroeconomic data are presented to demonstrate the key characteristics of BPS under a controlled setting. Chapter 3 applies the univariate BPS framework to a topical US macroeconomic data study. This study lays out the basic features and characteristics of BPS for short and long term forecasting. Chapter 4 extends the application in Chapter 3 to a study of multiple US macroeconomic time series. Chapters 3 and 4 together lay down the basic framework of how BPS is applied and utilized, on which the applications in later chapters are built. Chapters 5-7 apply BPS to different data and problems that are topical in economics and finance. Chapter 5 presents two studies involving return predictability in finance. The first study explores the return predictability of US sector indices and how BPS can improve on the literature. The second study utilizes the concept of information projection to synthesize sector information to improve the return predictability of the US market index. Chapter 6 presents two studies in the problem of temporal misspecification. The two studies explore temporal misspecification in rolling window size and data frequency for long term forecasting and how the BPS framework mitigates the misspecification problem and improves forecasts. Chapter 7 proposes a new method for forecasting using mixed frequency time series using the BPS framework. A study involving nowcasting quarterly GDP using monthly macroeconomic data illustrates the features of the proposed method. Chapter 8 concludes the dissertation with open questions and discussion of potential future research. Finally, the Appendices provide details of MCMC algorithms for the models and methods proposed.

2

Bayesian Predictive Synthesis

2.1 Background and foundations

This dissertation develops dynamic models for time series that build on theoretical foundations in West (1992) concerned with predicting a single outcome y, or multiple outcomes $\boldsymbol{y}$. In time series generalizations, y, or $\boldsymbol{y}$, will be the outcome of a univariate/multivariate time series at one time point, typically real-valued, although the foundational theory is general. Without loss of generality, we first summarize the basic ideas and key historical result in the single outcome setting.

A Bayesian decision maker $\mathcal{D}$ is interested in predicting the outcome y and aims to incorporate information from J individual agents (models, forecasters, or forecasting agencies, etc.) labelled $\mathcal{A}_j$, $(j = 1{:}J)$. To begin, $\mathcal{D}$ has prior $p(y)$, $\mathcal{D}$'s prior belief on the outcome; then each $\mathcal{A}_j$ provides $\mathcal{D}$ with forecast information in terms of a probability density function $h_j(x)$, where x is a quantity related to y, the special and pertinent case being $x = y$. 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 indicates that $\mathcal{D}$ will predict y using the implied posterior $p(y|\mathcal{H})$ from a full Bayesian prior-to-posterior analysis, that is ,

$$p(y|\mathcal{H}) \propto p(y, \mathcal{H}) = p(\mathcal{H}|y)p(y).$$

Given the complex nature of $\mathcal{H}$– a set of J density functions, in a setting where there will be varying dependencies among agents as well as individual biases– a fully specified Bayesian model $p(y, \mathcal{H}) = p(y)p(\mathcal{H}|y)$ is not easily conceptualized.

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

$$p(y|\mathcal{H}) = \int \alpha(y|\boldsymbol{x}) \prod_{j=1:J} h_j(x_j)dx_j \tag{2.1}$$

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

$$p(y) = \int \alpha(y|\boldsymbol{x})m(\boldsymbol{x})d\boldsymbol{x} \qquad \text{where} \qquad m(\boldsymbol{x}) = E[\prod_{j=1:J} h_j(x_j)], \tag{2.2}$$

the expectation in the last formula being over $\mathcal{D}$'s distribution $p(\mathcal{H})$. Critically, the representation of eqn. (2.1) does not require a full specification of $p(y, \mathcal{H})$ (and hence $p(\mathcal{H})$), but only the prior $p(y)$ and the marginal expectation function $m(\boldsymbol{x})$ of eqn. (2.2). The above theorem forms a partially specified nonparametric representation of Bayes' theorem, where $\mathcal{D}$'s specifications for $p(y)$ and $m(\boldsymbol{x})$ are consistent with a complete joint prior $p(y, \mathcal{H})$ (consistency condition).

These specifications alone do not, of course, indicate what the functional form of $\alpha(y|\boldsymbol{x})$ is, which opens the path to developing models based on different specifications. Key to considering this is the interpretation of the latent vector $\boldsymbol{x}$. From eqn. (2.1), note two implications/interpretations:

- Suppose each agent $\mathcal{A}_j$ *simulates* a single draw from $h_j(x_j)$; label these draws $\boldsymbol{x} = x_{1:J}$. Then $\mathcal{D}$ can immediately simulate from the implied posterior $p(y|\mathcal{H})$ by sampling $y \sim \alpha(y|\boldsymbol{x})$.
- Suppose the hypothetical agent information $h_j(x_j) = \delta_{x_j}(x_j)$ for $j = 1{:}J$. That is, $\mathcal{A}_j$ makes a perfect prediction $y = x_j$ for some specified value x_j. $\mathcal{D}$'s posterior is then $\alpha(y|\boldsymbol{x})$.

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

From eqn. (2.2), note that $(y, \boldsymbol{x})$ have an implicit joint distribution with margins $p(y)$ and $m(\boldsymbol{x})$, so we can consider this is understanding the ways in which the framework allows $\mathcal{D}$ to incorporate views, and historical information, about agent-specific biases, patterns of miscalibration, inter-dependencies among agents and their relative expertise/accuracy. The margin for latent agent states $m(\boldsymbol{x})$ is $\mathcal{D}$'s prior expectation of the product of agent densities; an example with $m(\boldsymbol{x})$ having positive dependencies among a subset of the x_j indicates that $\mathcal{D}$ anticipates positive concordance among the corresponding predictive densities $h_j(\cdot)$ of that subset of agents.

Example 1. A class of examples arises when the implied joint prior $\alpha(y|\boldsymbol{x})m(\boldsymbol{x})$ is multivariate normal or T, which easily and intuitively allows for: (i) ranges of agent biases and mis-calibration, viewed through shifts in means and/or variances of implied conditional distributions of individual conditional distributions $(x_j|y)$; and (ii) inter-dependencies, reflected in patterns of correlations and other aspects of conditional dependence among the x_j (West and Crosse, 1992; West and Harrison, 1997, Sect 16.3.2).

As a specific, conditionally normal example, suppose that the joint prior for

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

$$\alpha(y|\boldsymbol{x}) = N(y|\boldsymbol{F}'\boldsymbol{\theta}, v) \quad \text{with} \quad \boldsymbol{F} = (1, \boldsymbol{x}')' \quad \text{and} \quad \boldsymbol{\theta} = (\theta_0, \theta_1, ..., \theta_J)' \qquad (2.3)$$

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

2.2 Dynamic synthesis of forecasts

The new methodological developments forming the core of this dissertation adapt and extend the basic BPS framework summarized above to sequential forecasting in time series. In particular, we develop dynamic extensions of conditionally normal BPS models involving time-varying parameters to characterize and formally allow for agent-specific biases, patterns of mis-calibration, inter-dependencies, and relative expertise/forecast accuracy as time evolves and data is processed. We do this in the context of a scalar time series, for clarity and examples, although the ideas and approach are immediately extendable to multivariate cases, explored later.

Among the connections in recent literature mentioned in Section 1, Hoogerheide et al. (2010) and Aastveit et al. (2016) relate directly in key aspects of technical structure. In addition to opportunities for time-varying parameter models– a special case of the broader DLM setting developed in the following sections– these authors develop empirical methods using forecasts *simulated* from sets of models. This relates directly, as BPS provides a complete theoretical framework with implied underlying latent agent states arising from the agent distributions. As we see below, practical Bayesian analysis of dynamic BPS models naturally involves simulation of these latent states from the agent distributions in forecasting computations; however, they must be simulated from different distributions– the appropriate conditional posteriors– for model fitting and analysis.

2.2.1 Dynamic sequential setting: Univariate

The decision maker $\mathcal{D}$ is sequentially predicting a time series $y_t, t = 1, 2, \ldots,$ and at each time point receives forecast densities from each $\mathcal{A}_j$. At each time $t-1$, $\mathcal{D}$ aims to forecast y_t and receives current forecast densities $\mathcal{H}_t = \{h_{t1}(x_{t1}), \ldots, h_{tJ}(x_{tJ})\}$ from the set of agents. The full information set used by $\mathcal{D}$ is thus $\{\boldsymbol{y}_{1:t-1},\ \mathcal{H}_{1:t}\}$. As data accrues, $\mathcal{D}$ learns about the latent relationships among agents, their forecast and dependency characteristics, so that a Bayesian model will involve parameters that define the BPS framework and for which $\mathcal{D}$ updates information over time. The implication for the temporal/dynamic extension of the BPS model of Section 2.1 is that $\mathcal{D}$ has a time $t-1$ distribution for y_t of the form

$$p(y_t|\boldsymbol{\Phi}_t, \boldsymbol{y}_{1:t-1}, \mathcal{H}_{1:t}) \equiv p(y_t|\boldsymbol{\Phi}_t, \mathcal{H}_t) = \int \alpha_t(y_t|\boldsymbol{x}_t, \boldsymbol{\Phi}_t) \prod_{j=1:J} h_{tj}(x_{tj}) dx_{tj} \qquad (2.4)$$

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

Latent Factor Dynamic Linear Models

Consider a dynamic regression for BPS calibration that extends the basic example of eqn. (2.3) to the time series setting. That is, eqn. (2.3) becomes the dynamic version

$$\alpha_t(y_t|\boldsymbol{x}_t, \boldsymbol{\Phi}_t) = N(y_t|\boldsymbol{F}_t'\boldsymbol{\theta}_t, v_t) \quad \text{with} \quad \boldsymbol{F}_t = (1, \boldsymbol{x}_t')' \quad \text{and} \quad \boldsymbol{\theta}_t = (\theta_{t0}, \theta_{t1}, ..., \theta_{tJ})', \tag{2.5}$$

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

$$y_t = \boldsymbol{F}_t'\boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim N(0, v_t), \tag{2.6a}$$

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N(0, v_t\boldsymbol{W}_t) \tag{2.6b}$$

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

ν_t and evolution *innovations* $\boldsymbol{\omega}_s$ are independent over time and mutually independent for all t, s.

Note that under the above dynamic setting, the fundamental specifications of $m(\boldsymbol{x}_t)$ and $p(y_t)$ effectively disappear as a result of the prior-posterior updating of DLMs. That is, $\mathcal{D}$'s prior for $t+1$, $p(y_{t+1})$, is the posterior from t, $p(y_t|\boldsymbol{\Phi}_t, \mathcal{H}_t)$, and $m(\boldsymbol{x}_{t+1}) = E(\prod_{j=1:J} h_{tj}(x_{tj})|D_{t-1})$, so $\mathcal{D}$'s expectations are consistent with the latest set of forecasts at hand. Substituting the above into eqn. (2.2), we see that it satisfies the consistency condition and hence the theorem. This is a unique characteristic of dynamic models, and care must be made in the specifications of $m(\boldsymbol{x}_t)$ and $p(y_t)$ for stationary models that do not share the sequential prior-posterior updating scheme of DLMs.

The DLM specification is completed using standard discount methods: (i) The time-varying intercept and agent coefficients $\boldsymbol{\theta}_t$ follow the random walk evolution of eqn. (2.6b) where $\boldsymbol{W}_t$ is defined via a standard, single discount factor specification (West and Harrison 1997, Sect 6.3; Prado and West 2010, Sect 4.3); (ii) The residual variance v_t follows a standard beta-gamma random walk volatility model (West and Harrison 1997, Sect 10.8; Prado and West 2010, Sect 4.3), with $v_t = v_{t-1}\delta/\gamma_t$ for some discount factor $\delta \in (0, 1]$ and where γ_t are beta distributed innovations, independent over time and independent of $\nu_s, \boldsymbol{\omega}_r$ for all t, s, r. Given choices of discount factors underlying these two components, and a (conjugate normal/inverse-gamma) initial prior for $(\boldsymbol{\theta}_1, v_1)$ at $t = 0$, the model is fully specified.

Eqns. (2.6) define a dynamic latent factor model: the $\boldsymbol{x}_t$ vectors in each $\boldsymbol{F}_t$ are latent variables. At each time, they are conceived as arising as single draws from the set of agent densities $h_{tj}(\cdot)$, the latter becoming available at time $t-1$ for forecasting y_t. Note that the latent factor generating process has the x_{tj} drawn *independently* from their $h_{tj}(\cdot)$– based on the BPS foundational theory of eqn. (2.4)–

and externally to the BPS model. That is, coupled with eqns. (2.6a,2.6b), we have

$$p(\boldsymbol{x}_t|\boldsymbol{\Phi}_t, \boldsymbol{y}_{1:t-1}, \mathcal{H}_{1:t}) \equiv p(\boldsymbol{x}_t|\mathcal{H}_t) = \prod_{j=1:J} h_{tj}(x_{tj}) \tag{2.7}$$

for all time t and with $\boldsymbol{x}_t, \boldsymbol{x}_s$ conditionally independent for all $t \neq s$. Importantly, the independence of the x_{tj} must not be confused with the question of $\mathcal{D}$'s modeling and estimation of the dependencies among agents: this is simply central and integral, and reflected through the effective DLM parameters $\boldsymbol{\Phi}_t$.

While we choose a simple form of a DLM for the synthesis function $\alpha_t(y_t|\boldsymbol{x}_t, \boldsymbol{\Phi}_t)$ in view of its computational simplicity and flexibility, the theory does not imply any specific structure for the synthesis function and the decision maker is free to define model specifications and/or informative priors based on prior information/belief. For example, if the models behind the agents are known to be similar, the decision maker might want to set the prior correlations to be high amongst agents; or if the decision maker believes there are clear regime changes that favor certain agents at certain periods of time, a regime switching approach or an indicator in the state equation might be suitable.

We also note that most methods in the forecast combination literature focus on forecasts averaging with point weights (in θ_t) that are restricted to the unit simplex, as well as the weights summing to one (e.g. Clemen, 1989; Timmermann, 2004; Hall and Mitchell, 2007; Hoogerheide et al., 2010; Geweke and Amisano, 2012). To modify eqn. (2.5) so the coefficients sum to one, we can apply the technique used in Irie and West (2016), where the sum of coefficients is always restricted to the same value. For coefficients restricted to the unit simplex but not summing to one, it is significantly more complicated, as we now have a non-linear state space model in eqn. (2.6b). The benefit of having coefficients restricted to the unit simplex is interpretability. However, beyond that, there is no real benefit in the restriction. In the dynamic BPS setting, not only is restricting the coefficients computationally and technically more intensive, we can expect it to underperform compared to the unrestricted DLM case presented in this dissertation. Consider, for example, the case where all agents overestimate the quantity of interest by some positive value. Under the restrictive case, there is no combination of coefficients that can achieve that quantity, while the unrestrictive case can with negative coefficients. For these reasons, we utilize the unrestricted DLM instead of the conventional restricted versions.

2.2.2 *Dynamic sequential setting: Multivariate*

For the multivariate time series $\boldsymbol{y}_t, t = 1, 2, \dots,$ with dimension q for the number of series, the decision maker $\mathcal{D}$ sequentially receives forecast densities from each agent to forecast all series at each time point. At each time $t-1$, $\mathcal{D}$ receives current forecast densities $\mathcal{H}_t = \{h_{t1}(\boldsymbol{x}_{t1}), \dots, h_{tJ}(\boldsymbol{x}_{tJ})\}$ from the set of agents and aims to forecast $\boldsymbol{y}_t$, forming the full information set used by $\mathcal{D}$; $\{\boldsymbol{Y}_{1:t-1},\ \mathcal{H}_{1:t}\}$. As $\mathcal{D}$ observes more information, the dependencies/biases/characteristics among agents and for each agent are learnt, updating information over time with a Bayesian model, with its associated parameters, that define the BPS framework.

In the temporal/dynamic domain, the BPS model of Section 2.1 implies that $\mathcal{D}$ has a time $t-1$ multivariate distribution for $\boldsymbol{y}_t$ of the form

$$p(\boldsymbol{y}_t|\boldsymbol{\Phi}_t, \boldsymbol{Y}_{1:t-1}, \mathcal{H}_{1:t}) \equiv p(\boldsymbol{y}_t|\boldsymbol{\Phi}_t, \mathcal{H}_t) = \int \alpha_t(\boldsymbol{y}_t|\boldsymbol{X}_t, \boldsymbol{\Phi}_t) \prod_{j=1:J} h_{tj}(\boldsymbol{x}_{tj}) d\boldsymbol{x}_{tj} \qquad (2.8)$$

where $\boldsymbol{X}_t = \boldsymbol{X}_{t,1:J}$ is a $J \times q-$dimensional latent agent state matrix at time t, $\alpha_t(\boldsymbol{y}_t|\boldsymbol{X}_t, \boldsymbol{\Phi}_t)$ is $\mathcal{D}$'s conditional calibration p.d.f. for $\boldsymbol{y}_t$ given $\boldsymbol{X}_t$, and $\boldsymbol{\Phi}_t$ represents time-varying parameters defining the calibration p.d.f.– parameters for which $\mathcal{D}$ has current beliefs represented in terms of a current (time $t-1$) posterior $p(\boldsymbol{\Phi}_t|\boldsymbol{Y}_{1:t-1}, \mathcal{H}_{1:t-1})$.

At this point, the methodological focus rests on the evaluation of agents based on various assumptions about $\alpha_t(\boldsymbol{y}_t|\boldsymbol{X}_t, \boldsymbol{\Phi}_t)$ along with its defining dynamic state parameters $\boldsymbol{\Phi}_t$. For the multivariate extension of the univariate case in Section 2.2.1, we look to extend the seemingly unrelated regression (SUR; Zellner, 1962) to a dynamic Bayesian framework, as a first approach to defining a computationally accessible yet flexible framework for dynamic multivariate BPS.

Multivariate Latent Factor Dynamic Linear Models

Consider a dynamic multivariate regression for the BPS synthesis function that extends the univariate DLM in Section 2.2.1 to the multivariate domain. We thus specify

$$\alpha_t(\boldsymbol{y}_t|\boldsymbol{X}_t, \boldsymbol{\Phi}_t) = N(\boldsymbol{y}_t|\boldsymbol{F}_t'\boldsymbol{\theta}_t, \boldsymbol{V}_t) \tag{2.9}$$

with

$$\boldsymbol{F}_t = \begin{pmatrix} 1 & \boldsymbol{x}_{t1}' & 0 & \mathbf{0} & \cdots & \cdots & 0 & \mathbf{0} \\ 0 & \mathbf{0} & 1 & \boldsymbol{x}_{t2}' & & & & \vdots \\ \vdots & & & & \ddots & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 1 & \boldsymbol{x}_{tq}' \end{pmatrix} \quad \text{and} \quad \boldsymbol{\theta}_t = (\boldsymbol{\theta}_{t1}, \boldsymbol{\theta}_{t2}, ..., \boldsymbol{\theta}_{tq})', \tag{2.10}$$

where each $\boldsymbol{x}_{tq} = (x_{tq1}, x_{tq2}, ..., x_{tqJ})'$ in the former is a $1 \times J-$vector of realizations of agents states for series q and the latter is a $(1+J)q \times 1-$vector of time-varying bias/calibration coefficients, each $\boldsymbol{\theta}_{tq} = (1, \theta_{tq1}, \theta_{tq2}, ..., , \theta_{tqJ})$ containing the coefficients related to the agents' forecasts of series q. This defines the first component of the dynamic multivariate DLM

$$\boldsymbol{y}_t = \boldsymbol{F}_t'\boldsymbol{\theta}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim N(\mathbf{0}, \boldsymbol{V}_t), \tag{2.11a}$$

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N(\mathbf{0}, \boldsymbol{W}_t) \tag{2.11b}$$

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.

Similar to the univariate DLM with discount stochastic volatility used in Section 2.2.1, the specification for the above multivariate DLM 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. (2.11b) where $\boldsymbol{W}_t$ is defined via a standard, single discount factor specification (Prado and West 2010, Chap 10). The residual variance $\boldsymbol{\nu}_t$ follows a standard inverse Wishart random walk volatility model with some discount factor $\delta \in (0, 1]$. Given choices of discount factors underlying these two components, and a initial prior for $(\boldsymbol{\theta}_1, \boldsymbol{V}_1)$ at $t = 0$, the model is fully specified.

The above formulation in eqns. (2.11) define a dynamic multivariate latent factor model, where each agent forecast vector, $\boldsymbol{x}_{tq}$, in $\boldsymbol{F}_t$ are latent variables. At each time, the latent variables are realized as single draws from the set of agent densities $h_{tj}(\cdot)$, becoming available to $\mathcal{D}$ at $t-1$ for forecasting $\boldsymbol{y}_t$. As with the univariate case, $\boldsymbol{x}_t$ is drawn from $h_{tj}(\cdot)$ *independently* between agents, based on the BPS foundational theory of eqn. (2.8). This leads to, coupled with eqns. (2.11a,2.11b), the conditional posterior for the latent states,

$$p(\boldsymbol{X}_t|\boldsymbol{\Phi}_t, \boldsymbol{Y}_{1:t-1}, \mathcal{H}_{1:t}) \equiv p(\boldsymbol{X}_t|\mathcal{H}_t) = \prod_{j=1:J} h_{tj}(\boldsymbol{x}_{tj}) \tag{2.12}$$

for all time t and with $\boldsymbol{X}_t, \boldsymbol{X}_s$ conditionally independent for all $t \neq s$. Though the theory in eqn. (2.4) implies independence of $\boldsymbol{x}_{tj}$, this 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 DLM parameters $\boldsymbol{\Phi}_t$. In particular, the multivariate DLM specification above formulates the dependence among agents within and across series.

One aspect to note about the specification for eqns. (2.11) is that this does not provide a conjugate prior-posterior updating as in the case of univariate DLMs. For this to be conjugate, the above needs to be in the form of an exchangeable time series (Prado and West 2010, Chap 10). That is,

$$\boldsymbol{F}_t = (1, \boldsymbol{x}'_{t1}, \boldsymbol{x}'_{t2}, \cdots, \boldsymbol{x}'_{tJ}) \quad \text{and} \quad \boldsymbol{\Theta}_t = (\boldsymbol{\theta}'_{t1}, \boldsymbol{\theta}'_{t2}, ..., \boldsymbol{\theta}'_{tq})', \tag{2.13}$$

where all of the agents' forecasts for all series feeds into $\mathcal{D}$'s forecast of each series (i.e. the latent factors are the same across all series). While this is computationally superior to eqns. (2.11), it lacks interpretability and practicality. In the exchangeable time series framework, all forecasts across agents and across series feed into the forecasts of each series (e.g. forecasts of inflation feeding into the forecast of unemployment and vice versa). While we can assume, and expect, the latent states across agents and series to be dependent, we cannot assume that forecasts for other series to positively improve the forecast of another series directly. Eqns. (2.11), on the other hand, take into account the inter-agent dependencies across and between series that we expect (in the variance matrix $\boldsymbol{W}_t$), while filtering out non-direct forecasts from each series, focusing on the forecasts that matter. Additionally, $\boldsymbol{F}_t$ is a $1 + Jq$-dimension vector, which is often quite large and causes practical problems with the estimation. In comparison, eqn. (2.11) has a $1 + J$-dimensional vector per series. For these reasons, we have selected the specification in eqns. (2.11) for the synthesis function. That being said, we do not conclude that an exchangeable time series or more complicated Bayesian VARs should not be used. Theorem 2.8 does not imply the form of the synthesis function, and exploration of different specifications should be encouraged.

2.2.3 Multi-step ahead synthesis

Forecasting over multiple horizons is often of equal or greater importance than 1-step ahead forecasting. Economic policy makers, for example, forecast/assess macroeconomic variables over a year or multiple years, drawing from their own forecast models, judgmental inputs, other economists and forecasters, in order to advise policy decisions. However, forecasting over longer horizons is typically more difficult than over shorter horizons, and models calibrated on the short-term basis can often be quite poor in the longer-term. As noted in Chapter 1, fitting of time series models is inherently based on 1-step ahead, as DLM (and other) model equations make explicit. When multi-step ahead forecasting is primary, new ideas for forecast calibration and combination are needed. BPS provides a natural and flexible framework to synthesize forecasts over multiple horizons, with potential to improve forecasting at multiple horizons simultaneously, as we now discuss (for univariate synthesis, though multivariate synthesis directly follows from the results).

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

BPS(k) **for customized multi-step forecasting.** BPS is open to models customized to forecasting goals, and so provides $\mathcal{D}$ a strategy to focus modeling on the horizon of interest. This involves a trivial modification of Section 2.2.1 in which the model at time $t-1$ for predicting y_t changes as follows. With a *specific forecast horizon* $k>1$, modify the BPS so that the agents' k-step ahead forecast densities made at time $t-k$, i.e., $h_{t-k,j}(x_{tj})$ replace $h_{tj}(x_{tj})$ in the resulting model analysis. This changes the interpretation of the dynamic model parameters $\{\boldsymbol{\theta}_t, v_t\}$ to be explicitly geared to the k-step horizon. Bayesian model fitting then naturally "tunes" the model to the horizon k of interest. Forecasting the chosen k-steps ahead now simply involves extrapolating the model via simulation, as above, but now in this modified and horizon-specific BPS model.

We denote this customized model strategy by BPS(k) to distinguish it from the direct extrapolation of BPS. Note that this is fundamentally different from the traditional method of model extrapolation as it directly updates, calibrates, and learns using the horizon of interest. The applied studies in subsequent chapters bear out the view that this can be expected to improve forecasting accuracy over multiple horizons. One cost, of course, is that a bank of BPS models is now required for any set of horizons of interest; that is, different models will be built for different horizons $k,$ so increasing the computational effort required. If the model specification of the synthesis function is correct, we can expect that the direct extrapolation of BPS to equal BPS(k), i.e. the estimated parameters $\boldsymbol{\Phi}$ are equivalent. However, this is rarely– if ever– the case in real data applications.

We further note contextual relevance of this perspective in applications when $\mathcal{D}$ is a consumer of forecasts from groups, agencies or model developers. Such agents may use different models, data, advisors, and approaches for different horizons. When the forecast generating models/methods are known, $\mathcal{D}$ may redefine the BPS model accordingly; however, generally in practice these underlying models, strategies, data, and advisors will not be wholly known and understood.

2.3 Bayesian computation

2.3.1 Univariate synthesis

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

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

Posterior Computations via MCMC. The dynamic latent factor model structure of eqns. (2.6a,2.6b,2.7) leads easily to a two-component block Gibbs sampler for the latent agent states and dynamic parameters. These are iteratively resimulated from the two conditional posteriors noted below, with obvious initialization based on agent states drawn independently from priors $h_*(*)$.

First, conditional on values of agent states, the next MCMC step draws new parameters from $p(\boldsymbol{\Phi}_{1:t}|\boldsymbol{X}_{1:t}, \boldsymbol{y}_{1:t})$. By design, this is a discount-based dynamic linear regression model, and sampling uses the standard forward filtering, backward sampling (FFBS) algorithm (e.g. Frühwirth-Schnatter 1994; West and Harrison 1997, Sect 15.2; Prado and West 2010, Sect 4.5).

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

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

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

Forecasting 1-Step Ahead. At time t we forecast 1-step ahead by generating "synthetic futures" from the BPS model, as follows: (i) For each sampled $\boldsymbol{\Phi}_t$ from the posterior MCMC above, draw v_{t+1} from its discount volatility evolution model, and then $\boldsymbol{\theta}_{t+1}$ conditional on $\boldsymbol{\theta}_t, v_{t+1}$ from the evolution model eqn. (2.6b)– this gives a draw $\boldsymbol{\Phi}_{t+1} = \{\boldsymbol{\theta}_{t+1}, v_{t+1}\}$ from $p(\boldsymbol{\Phi}_{t+1}|\boldsymbol{y}_{1:t}, \mathcal{H}_{1:t})$; (ii) Draw $\boldsymbol{x}_{t+1}$ via independent sampling of the $h_{t+1,j}(x_{t+1,j})$, $(j = 1{:}J)$; (iii) Draw y_{t+1} from the conditional normal of eqn. (2.6a) given these sampled parameters and agent states. Repeating this generates a random sample from the 1-step ahead forecast distribution for time $t+1$.

2.3.2 *Multivariate synthesis*

At any current time t, $\mathcal{D}$ has the now historical information $\{\boldsymbol{Y}_{1:t}, \mathcal{H}_{1:t}\}$, the past observed data and agents' forecast of that data, with the history of the BPS analysis up until that point. Up until that time, the BPS analysis will have inferred on both the latent agent states $\boldsymbol{X}_*$ and the dynamic BPS model parameters $\boldsymbol{\Phi}_*$. The former, importantly, provides insight into the dependencies, biases, and other characteristics pertaining to $\boldsymbol{Y}_{1:t}$, among agents and individual agents. Thus, the full posterior summaries for the multivariate $\boldsymbol{X}_t$ contains these characteristics among and within the multiple series; 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.

The posterior analysis is enabled by Markov chain Monte Carlo (MCMC) methods, followed by forecasting from time t onward utilizing the theoretical and simulation based extrapolation of the model. That is to say, $\mathcal{D}$ is interested in inference on the full set of past latent agent states (latent factors) and dynamic parameters $\{\boldsymbol{X}_{1:t}, \boldsymbol{\Phi}_{1:t}\}$, as well as forward filtering to update posteriors for current values $\{\boldsymbol{X}_t, \boldsymbol{\Phi}_t\}$, and posterior computation enables this. These aspects of posterior computation extend the MCMC method used for the univariate case, though there are specific modifications in terms of fitting the model in eqns. (2.11).

Posterior Computations via MCMC. The multivariate dynamic latent factor model structure of eqns. (2.11a,2.11b,2.12) 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 $p(\mathbf{\Theta}_{1:t}|\boldsymbol{X}_{1:t}, \boldsymbol{Y}_{1:t}, \boldsymbol{V}_{1:t})$. Given the residual volatility, $\boldsymbol{V}_{1:t}$, this mirrors the standard discount-dynamic multivariate dynamic linear regression model, and is thus sampled using an extension of the forward filtering, backward sampling (FFBS) algorithm (e.g. Prado and West 2010, Chap 10).

Second, due to the model specification, the MCMC step draws new dynamic volatility parameters, $p(\boldsymbol{V}_{1:t}|\boldsymbol{X}_{1:t}, \boldsymbol{Y}_{1:t}, \mathbf{\Theta}_{1:t})$, given the agent states and dynamic coefficient parameters. This is done using standard inverse Wishart analysis with discount stochastic volatility. Note that the univariate dynamic latent agent model (and the exchangeable time series framework) enjoys full conditionally conjugate analysis and the joint posterior $p(\mathbf{\Phi}_{1:t}|\boldsymbol{X}_{1:t}, \boldsymbol{Y}_{1:t})$ can be obtained directly, merging the above two steps into one.

Third, conditional on values of dynamic parameters, the MCMC draws new agent states from $p(\boldsymbol{X}_{1:t}|\mathbf{\Phi}_{1:t}, \boldsymbol{Y}_{1:t}, \mathcal{H}_{1:t})$. As with the univariate case, $\boldsymbol{X}_t$ are conditionally independent over time t in this conditional distribution, with time t conditionals

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

where

$$\boldsymbol{F}_t = \begin{pmatrix} 1 & \boldsymbol{x}_{t1}' & 0 & \mathbf{0} & \cdots & \cdots & 0 & \mathbf{0} \\ 0 & \mathbf{0} & 1 & \boldsymbol{x}_{t2}' & & & & \vdots \\ \vdots & & & & \ddots & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 1 & \boldsymbol{x}_{tq}' \end{pmatrix}.$$

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 normals. For a more central and practically important case of forecasts being multivariate T distributions, each $h_{tj}(\cdot)$ can be represented as a scale mixture of normals, and augmenting the posterior MCMC to include the implicit underlying latent scale factors generates conditional multivariate normals for each $\boldsymbol{X}_t$ coupled with conditional inverse gammas for those scales. For non-normal forecasts, where $p(\boldsymbol{X}_t|\boldsymbol{\Phi}_t, \boldsymbol{y}_t, \mathcal{H}_t) \propto N(\boldsymbol{y}_t|\boldsymbol{F}_t'\boldsymbol{\theta}_t, \boldsymbol{V}_t)$, is not analytically tractable, we can either use a Metropolis-Hastings algorithm or importance sampling scheme to sample $p(\boldsymbol{X}_t|\boldsymbol{\Phi}_t, \boldsymbol{y}_t, \mathcal{H}_t)$. Further discussion can be found in Appendix A.

Forecasting 1-Step Ahead. Forecasting 1-step ahead is done in the same fashion as the univariate case. At time t we forecast 1-step ahead by generating "synthetic futures" from the BPS model, as follows: (i) For each sampled $\boldsymbol{\Phi}_t$ from the posterior MCMC above, draw $\boldsymbol{V}_{t+1}$ 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. (2.6b)– this gives a draw $\boldsymbol{\Phi}_{t+1} = \{\boldsymbol{\theta}_{t+1}, \boldsymbol{V}_{t+1}\}$ from $p(\boldsymbol{\Phi}_{t+1}|\boldsymbol{Y}_{1:t}, \mathcal{H}_{1:t})$; (ii) Draw $\boldsymbol{X}_{t+1}$ via independent sampling of the $h_{t+1,j}(\boldsymbol{x}_{t+1,j}), (j = 1{:}J)$; (iii) Draw $\boldsymbol{y}_{t+1}$ from the conditional normal of eqn. (2.6a) given these sampled parameters and agent states. Repeating this generates a random sample from the 1-step ahead forecast distribution for time $t + 1$.

2.4 Simulation study

In this simulated example, we sample synthetic data using the macroeconomic data used in Chapter 3 (three quarterly macro series: annual inflation rate (p), short-term nominal interest rate (r), and unemployment rate (u) in the US economy from 1961/Q1 to 2014/Q4). From the data we form four models, M1- p_{t-1};

M2- $p_{t-3}, r_{t-3}, u_{t-3}$; M3- p_{t-3}; M4- $p_{t-1}, r_{t-1}, u_{t-1}$, where each subscript indicates the lag(s) of data used, that generate four unique time series with the following constant coefficients on the intercept and factor(s) in the models:

M1- $(0.05, 0.95)$,

M2- $(0.1, 1, -0.2, -0.1, 0.2, -0.1, 0.05, -0.3, 0.3, 0.05)$,

M3- $(0.2, 1.3, -0.3, -0.2)$,

M4- $(0.1, 0.9, 0.1, -0.1)$,

with a constant standard deviation of 0.1 for the observation error. The data, y_t is generated by switching between the generated four unique time series with a stationary probability of 85$\%$ and transition probability of 5$\%$ each.

Agents

Each M* is a standard DLM so that model fitting and generation of forecasts is routine. Prior specifications for the DLM state vector and discount volatility model in each is based on $\boldsymbol{\theta}_0|v_0 \sim N(\mathbf{0}, v_0\boldsymbol{I})$ and $1/v_0 \sim G(1, 0.01)$, using the usual $(\boldsymbol{\theta}, v)$ DLM notation (West and Harrison, 1997, Chap 4). Each agent model uses standard discount factor (β) specification for state evolution variances and discount factor (δ) for residual volatility; we use $(\beta, \delta) = (0.99, 0.95)$ in each of these agent models. The DLM-based forecast densities $h_{t-k,j}(x_{tj})$ are then predictive T distributions.

BPS specifications

For the dynamic BPS models for forecast horizons $k = 1$ and $k = 4$, we take initial priors as $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}, \boldsymbol{I})$ with $\boldsymbol{m} = (0, \mathbf{1}'/p)'$ and $1/v_0 \sim G(5, 0.01)$. BPS for 1-step ahead forecasting is based on $(\beta, \delta) = (0.95, 0.99)$, while BPS$(4)$, customized to

4-quarter ahead forecasting, uses $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}, 10^{-4}\boldsymbol{I})$ and $(\beta, \delta) = (0.99, 0.99)$. Differences by forecast horizon echo earlier discussion about different model choices being relevant to different forecast goals.

Comparison of out-of-sample forecast performance

We compare forecasts from BPS with standard Bayesian model uncertainty analysis (that more recently has been referred to as Bayesian model averaging– BMA) in which the agent densities are mixed with respect to sequentially updated model probabilities (e.g. Harrison and Stevens, 1976; West and Harrison, 1997, Sect 12.2). In addition, we compare with simpler, equally-weighted averages of agent forecast densities: using both linear pools (equally-weighted arithmetic means of forecast densities) and logarithmic pools (equally-weighted harmonic means of forecast densities), with some theoretical underpinnings (e.g. West 1984).

For point forecasts from all methods, we compute and compare mean squared forecast errors (MSFE) over the forecast horizons of interest. In comparing density forecasts with BPS, we also evaluate log predictive density ratios (LPDR); at horizon k and across time indices t, this is

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

where $p_s(y_{t+k}|y_{1:t})$ is the predictive density under model or model combination aggregation strategy indexed by s, compared against the corresponding BPS forecasts at this horizon.

Table 2.1 summarizes the forecast capabilities for each model and strategy over the whole period. For 1-step ahead forecasts, in terms of $\mathrm{MSFE}_{1:T}$, BPS performs 17.77$\%$ better than the best individual model (M3) and 17.69$\%$ better than the best aggregation strategy (BMA) (Fig. 2.1). Both models and strategies work similarly well, with aggregation strategies being on par or slightly better than the best forecast models. Likewise, the results of $\text{LPDR}_{1:T}$ shows clearly that BPS outperforms the other methods (Fig. 2.2). Further investigation of the coefficients (see Figure 2.3), shows that BPS is successful in correctly choosing the best model (M3) but not degenerative like BMA (see Figure 2.4), showing that BPS is superior in adaptability and flexibility, which leads to its improved predictive performance.

The results for the 4-step ahead forecasts are indicative of three things: (1) BPS performs slightly worse than the best individual model (M2: -2.48$\%$) and the best aggregation strategy (BMA: -2.06$\%$), (2) BPS performs worse than most in terms of distribution forecasts, and (3) BPS(k) is superior in both point and distribution forecasts compared to BPS (9.66$\%$ improvement in terms of $\text{MSFE}_{1:T}$). One explanation for the poor performance in $\text{LPDR}_{1:T}$ is that BPS overly optimizes its parameters to the horizon it models, thus giving a tighter estimate (smaller standard deviation; see Figure 2.5) that does not account for the poor long term predictive abilities of the models. Since BPS(k) can directly model the forecasts at the horizon of interest, it is able to take into account the forecast uncertainties that come with forecasting long horizons evidenced by the drastically improved $\text{LPDR}_{1:T}$.

Table 2.1: Simulated data forecasting 1990/Q1-2014/Q4: 1- and 4-step ahead forecast evaluations comparing mean squared forecast errors and log predictive density ratios for the $T = 100$ quarters. Note: $\text{LPDR}_{1:T}$ is relative to BPS.

	$\text{MSFE}_{1:T}$		$\text{LPDR}_{1:T}$		
	1-step	4-step	1-step	4-step	BPS(4)
M1	0.06807	0.34387	-14.16	12.53	-55.63
M2	0.07734	0.29409	-17.87	44.48	-23.68
M3	0.06728	0.32987	-10.85	11.84	-56.33
M4	0.08902	0.38116	-25.88	27.25	-40.92
BMA	0.06721	0.29534	-10.93	43.56	-24.61
LinP	0.06860	0.33018	-15.34	36.92	-31.25
LogP	0.06765	0.32942	-12.87	25.43	-42.74
BPS	0.05532	0.30156	-	-	-
BPS(k)	-	0.27242	-	-	-

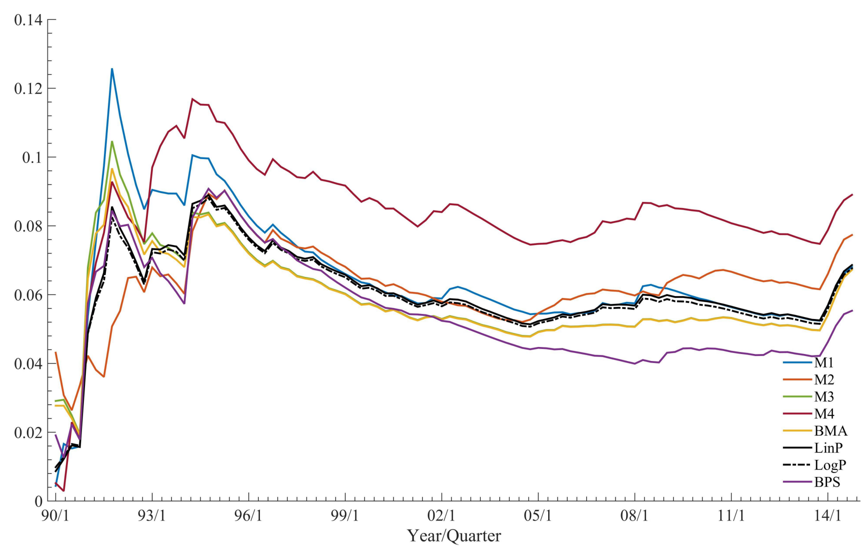

FIGURE 2.1: Simulated data forecasting 1990/Q1-2014/Q4: Mean squared 1-step ahead forecast errors $\text{MSFE}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}100$ quarters.

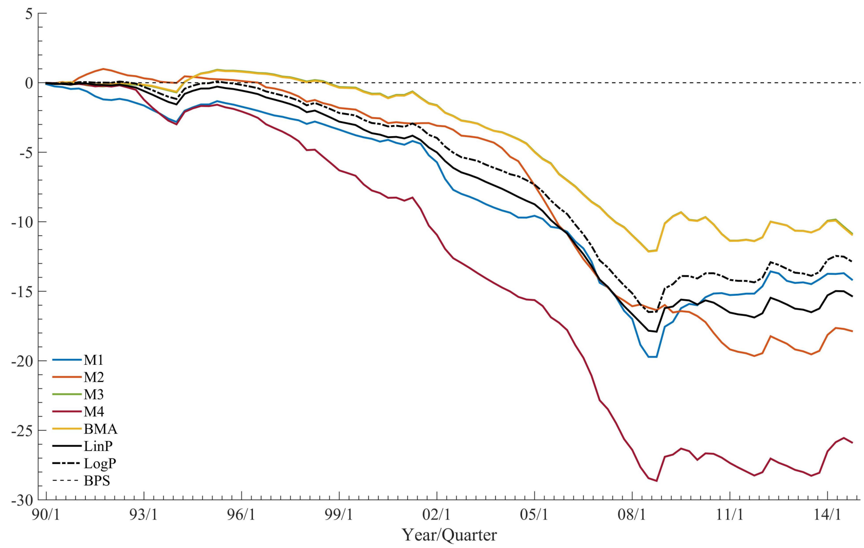

FIGURE 2.2: Simulated data forecasting 1990/Q1-2014/Q4: 1-step ahead log predictive density ratios $\text{LPDR}_{1:t}$ sequentially revised at each of the $t = 1{:}100$ quarters. The baseline at 0 over all t corresponds to the standard BPS model.

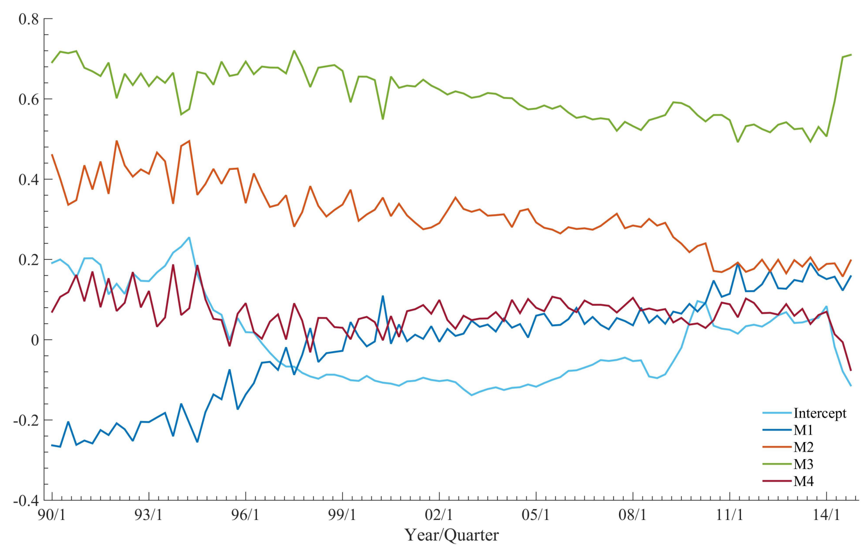

FIGURE 2.3: Simulated data forecasting 1990/Q1-2014/Q4: On-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}100$ quarters.

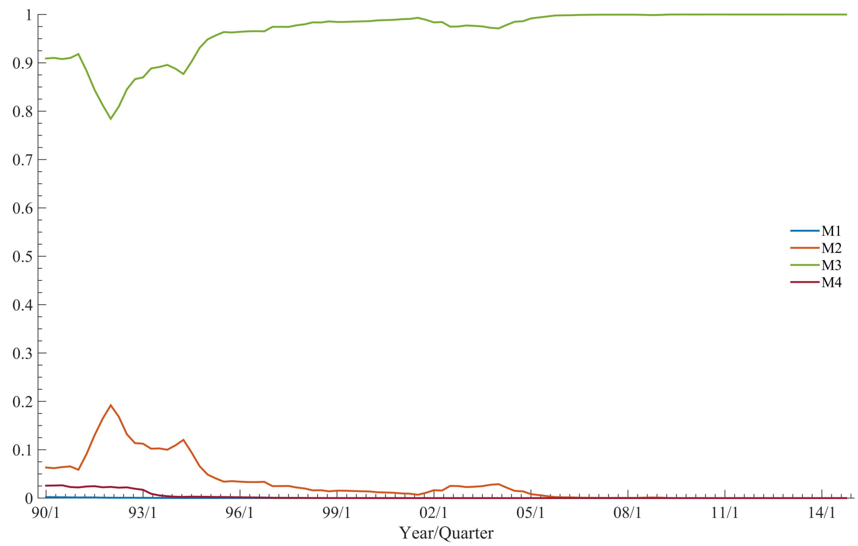

FIGURE 2.4: Simulated data forecasting 1990/Q1-2014/Q4: On-line model probabilities for BMA sequentially computed at each of the $t = 1{:}100$ quarters.

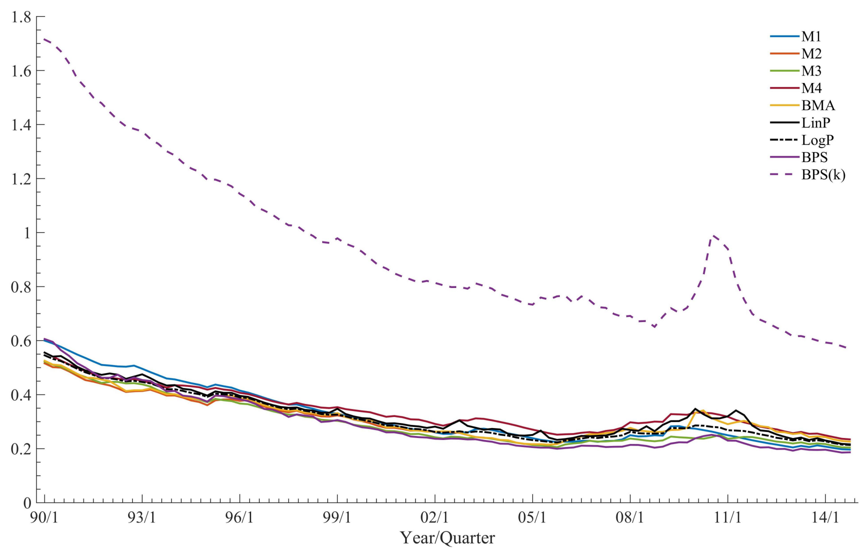

FIGURE 2.5: Simulated data forecasting 1990/Q1-2014/Q4: 4-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}100$ quarters.

3

Univariate BPS: US Inflation

3.1 US inflation forecasting

Inflation is one of the key metrics for the state of economy. Decision makers, from economic policy makers to financial portfolio managers to business managers, heavily rely on inflation projections to make decisions on monetary policy, portfolio weights, and product pricing. In order to make informed evidence-based decisions for the future, decision makers consult forecasters, models, etc. for inflation forecasts, to be better informed on the state of the economy, market, and demand. Due to its importance, there has been a long history of research, from both researchers and practitioners, on the subject of forecasting inflation (Stock and Watson, 1999, 2007; Stella and Stock, 2012; Belmonte et al., 2014). For our application of univariate BPS, we forecast inflation using a topical macroeconomic data set and models, echoing the decision process of economic policy makers.

3.1.1 Data

We analyze quarterly US macroeconomic data, focusing on forecasting inflation rates with both 1-quarter and 4-quarter ahead interests. The study involves three quarterly macro series: annual inflation rate (p), short-term nominal interest rate (r), and unemployment rate (u) in the US economy from 1961/Q1 to 2014/Q4, a context of topical interest (Cogley and Sargent, 2005; Primiceri, 2005; Koop et al., 2009; Nakajima and West, 2013a). The inflation rate is the annual percentage change in a chain-weighted GDP price index, the interest rate is the yield on three-month Treasury bills, and the unemployment rate is seasonally adjusted and includes all workers over 16 years of age. Prior studies (e.g. Nakajima and West, 2013a) use data over the period of 1963/Q1-2011/Q4; we extend this to more recent times, 1961/Q1 to 2014/Q4, shown in Fig. 3.1. We focus on forecasting inflation using past values of the three indices as candidate predictors underlying a set of four time series models– the $J = 4$ agents– to be evaluated, calibrated, and synthesized. The time frame includes key periods that warrant special attention: the early 1990s recession, the Asian and Russian financial crises in the late 1990s, the dot-com bubble in the early 2000s, and the sub-prime mortgage crisis and great recession of the late 2000s. These periods exhibit sharp shocks to the US economy generally, and test the predictive ability of any models and strategies under stress. For any forecast calibration and aggregation strategy to be effective and useful, its predictive performance must be robust under these conditions; most traditional macroeconomic models suffer significant deficiencies in such times.

3.1.2 Agents and BPS specification

The $J = 4$ agents represent the two major structures of time series forecast models: factor and lag. Labeling them M*, the agent models for inflation $y_t \equiv p_t$ use predictors: M1- p_{t-1}; M2- $p_{t-3}, r_{t-3}, u_{t-3}$; M3- p_{t-3}; M4- $p_{t-1}, r_{t-1}, u_{t-1}$. Thus, each has a time-varying autoregressive term in inflation rate p, while two also have dynamic regressions on lagged interest rate r and unemployment rate u, the differences being in lags chosen and model complexity. In each, residual volatil-

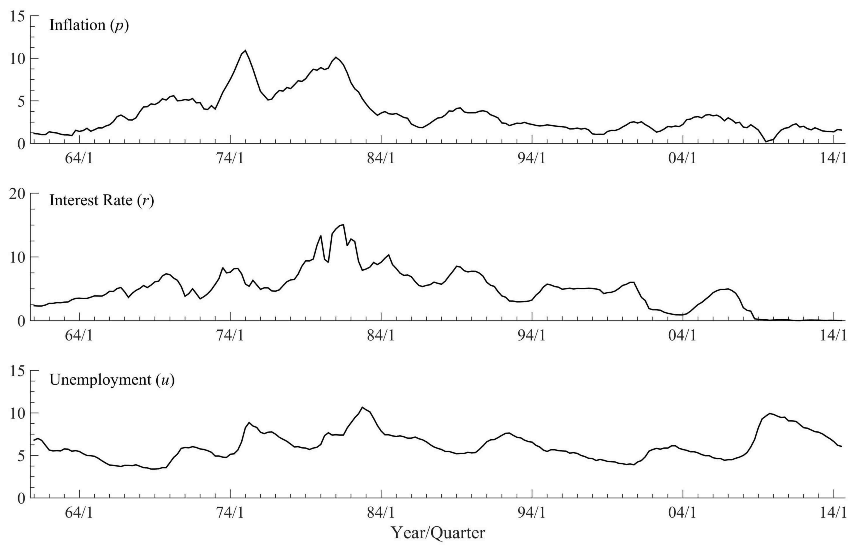

FIGURE 3.1: US inflation rate forecasting 1990/Q1-2014/Q4: US macroeconomic time series (indices $\times$100 for $\%$ basis): annual inflation rate (p), short-term nominal interest rate (r), and unemployment rate (u).

ity follows a standard beta-gamma random walk. Each M* is a standard DLM so that model fitting and generation of forecasts is routine. While these models are simple compared to more sophisticated models used to forecast inflation (such as Stock and Watson (1999, 2007); Stella and Stock (2012); Belmonte et al. (2014) as well as Bayesian VARs seen in Nakajima and West (2013a)), part of the utility and appeal of predictive synthesis, and in forecast combinations in general, is in gaining improvements using relatively simple models and not resorting to complicated models that have technical/computational difficulties. Additionally, we can expect that adding these sophisticated models to the set of agents will only improve the synthesis. Prior specifications for the DLM state vector and discount volatility model in each is based on $\boldsymbol{\theta}_0|v_0 \sim N(\mathbf{0}, v_0\boldsymbol{I})$ and $1/v_0 \sim G(1, 0.01)$, using the usual $(\boldsymbol{\theta}, v)$ DLM notation (West and Harrison, 1997, Chap 4). Each agent model uses standard discount factor (β) specification for state evolution variances and discount factor (δ) for residual volatility; we use $(\beta, \delta) = (0.99, 0.95)$ in each of these agent models. The DLM-based forecast densities $h_{t-k,j}(x_{tj})$ are then those of predictive T distributions. MCMC-based model fitting adapts to introduce latent scale factors as noted in Chapter 2.

In the dynamic BPS models for forecast horizons $k = 1$, we take initial priors as $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}, \boldsymbol{I})$ with $\boldsymbol{m} = (0, \mathbf{1}'/p)'$ and $1/v_0 \sim G(5, 0.01)$. BPS for 1-step ahead forecasting is based on $(\beta, \delta) = (0.95, 0.99)$. BPS$(4)$, customized to 4-quarter ahead forecasting as discussed in Section 2.2.3, uses $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}, 10^{-4}\boldsymbol{I})$ and $(\beta, \delta) = (0.99, 0.99)$. Differences by forecast horizon echo earlier discussion about different model choices being relevant to different forecast goals.

In general, discount factors should be set between $0.9 - 0.99$ for both β and δ. If the decision maker believes that the synthesis will benefit from weights that are extremely flexible (e.g. where good agents change over multiple periods of time), the decision maker can specify a lower discount factor β. In this example, we can assume that the simple models used will significantly underperform in the long term forecasts, as they are designed to fit to short term dynamics. For this reason, we set a higher discount factor β to smooth out the parameters because we can expect that there is very little signal to extract from the agents. Furthermore, these discount factors can be estimated via sampling, though it adds an extra level of computational complexity.

Another key specification is the prior mean of $\boldsymbol{\theta}_0$, $\boldsymbol{m}$. This can be chosen based on the prior belief of the decision maker, for example, when there are strong preferences in agents or there is a lack of data for calibration, though it is sensible to set it so that the coefficients in $\boldsymbol{m}$ on the agents sum to one.

Looking at the behavior of the parameters during the learning period provides insight into what prior specifications $\mathcal{D}$ should use for the analysis. For example, if the coefficients look jumpy/static during the learning period, it might be wise to increase/decrease the discount factor β to achieve a suitable level of adaptability.

We have explored analyses across ranges of priors and discount factors, and chosen these values as they lead to good agent-specific and BPS forecasting accuracy; conclusions with respect to BPS do not change materially with different values close to those chosen for the summary examples.

3.1.3 Data analysis and forecasting

The 4 agent models are analyzed and synthesized as follows. First, the agent models are analyzed in parallel over 1961/Q1-1977/Q1 as a training period, simply running the DLM forward filtering to the end of that period to calibrate the agent forecasts. This continues over 1977/Q2-1989/Q4 now accompanied by the calibration of the other forecast combination methods. Also, at each quarter t during this period, the MCMC-based BPS analysis is run using data from 1977/Q2 up to time t; that is, we repeat the analysis with an increasing "moving window" of past data as we move forward in time. We do this for the traditional 1-step focused BPS model, and– separately and in parallel– for a 4-step ahead focused BPS(4) model, as discussed in Section 2.2.3. This continues over the third period to the end of the series, 1990/Q1-2014/Q4; now we also record and compare forecasts as they are sequentially generated. This testing period spans over a quarter century, and we are able to explore predictive performance over periods of drastically varying economic circumstances, check robustness, and compare benefits and characteristics of each strategy. Out-of-sample forecasting is thus conducted and evaluated in a way that mirrors the realities facing decision and policy makers.

3.1.4 Forecast accuracy and comparisons

We compare forecasts from BPS with BMA, in which the agent densities are mixed with respect to sequentially updated model probabilities (e.g. Harrison and Stevens, 1976; West and Harrison, 1997, Sect 12.2), and a more recent "density combination" method of Billio et al. (2013) using the DeCo package (Casarin et al., 2015). In addition, we compare with simpler, equally-weighted averages of agent forecast densities: using both linear pools (equally-weighted arithmetic means of forecast densities) and logarithmic pools (equally-weighted harmonic means of forecast densities), with some theoretical underpinnings (e.g. West 1984). For point forecasts from all methods, we compute and compare mean squared forecast errors (MSFE) over the forecast horizons of interest. In comparing density forecasts with BPS, we also evaluate log predictive density ratios (LPDR); at horizon k and across time indices t, this is

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

where $p_s(y_{t+k}|y_{1:t})$ is the predictive density under model or model combination aggregation strategy indexed by s, compared against the corresponding BPS forecasts at this horizon. As used by several authors recently (e.g. Nakajima and West, 2013a; Aastveit et al., 2016), LPDR measures provide a direct statistical assessment of relative accuracy at multiple horizons that extend traditional 1-step focused Bayes' factors. They weigh and compare dispersion of forecast densities along with location, so elaborate on raw MSFE measures; comparing both measurements, i.e., point and density forecasts, gives a broader understanding of the predictive abilities of the different strategies.

3.2 Dynamic BPS and forecasting

Comparing predictive summaries over the out-of-sample period, BPS improves forecasting accuracy relative to the 4 agent models, and dominates BMA, DeCo, and the pooling strategies; see numerical summaries in Table 3.1. Looking at point forecast accuracy, BPS exhibits improvements of no less than 10$\%$ over all models and strategies for 1- and 4-step ahead forecasts (BPS(k) at $k = 4$ for the latter). As might be expected, BPS substantially improves characterization of forecast uncertainties as well as adaptation in forecast locations, reflected in the LPDR measures. Further, our expectations of improved multi-step forecasting using horizon-specific BPS are borne out: direct projection of the standard BPS model to 4-step ahead forecasts perform poorly, mainly as a result of under-dispersed forecast densities from each agent. In contrast, BPS(4) model performs substantially better, being customized to the 4-quarter horizon.

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

Table 3.1: US inflation rate forecasting 1990/Q1-2014/Q4: Forecast evaluations for quarterly US inflation over the 25 years 1990/Q1-2014/Q4, comparing mean squared forecast errors and log predictive density ratios for this $T = 100$ quarters. The column $\%$ denotes improvements over BPS and BPS(k) for 1- and 4-step ahead forecasts, respectively. Note: LPDR$_{1:T}$ is relative to BPS and BPS(4), for each column, and nonexistent for DeCo due to lack of analytic predictive distributions.

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

3.2.1 1-step ahead forecasting

Figs. 3.2-3.6 summarize sequential analysis for 1-step forecasting. Fig. 3.2 shows the 1-step ahead measures $\text{MSFE}_{1:t}(1)$ for each time t. BPS almost uniformly dominates, except at the beginning of the time period where the MSFE is somewhat unstable. Four "shock" periods are notable and increase forecast errors: 1992/Q3-Q4 (early 90s recession), 1997/Q4-1998/Q1 (Asian and Russian financial crisis), 2001/Q2-2003/Q1 (dot-com bubble), and 2009/Q2-2010/Q1 (sub-prime mortgage crisis). Even under the influence of these shocks, BPS is able to perform well with most of its improvements over other models and strategies coming from swift adaptation. The sub-prime mortgage crisis period highlights this, with MSFE staying relatively level under BPS while the others significantly increase.

Fig. 3.3 confirms that BPS performs uniformly better than, or on par with, the other models and BMA based on LPDR measures that measure relative distributional form and dispersion of forecast densities as well as location. Major shocks

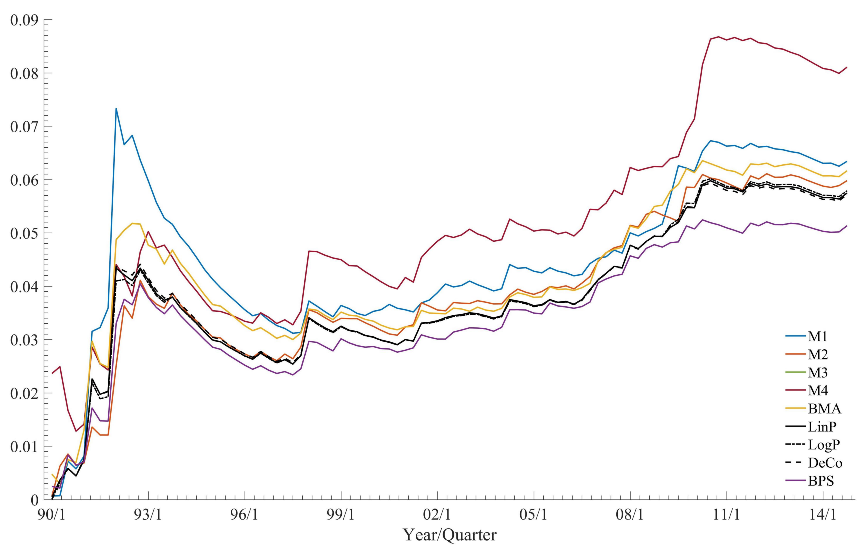

FIGURE 3.2: US inflation rate forecasting 1990/Q1-2014/Q4: Mean squared 1-step ahead forecast errors $\text{MSFE}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}100$ quarters.

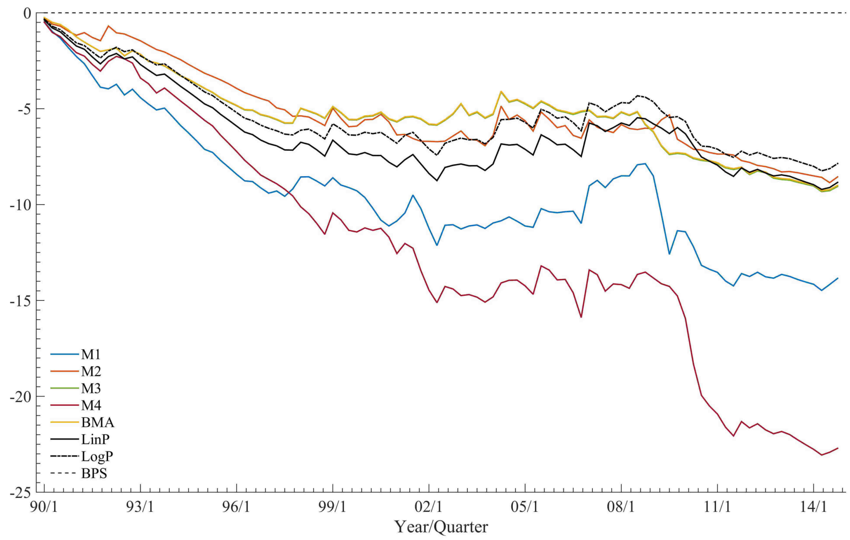

FIGURE 3.3: US inflation rate forecasting 1990/Q1-2014/Q4: 1-step ahead log predictive density ratios $\text{LPDR}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}100$ quarters. The baseline at 0 over all t corresponds to the standard BPS model.

and times of increased volatility have substantial impact on the relative performance, again most notable at the beginning of the sub-prime mortgage crisis. BPS is able to adapt to maintain improved forecasting performance both in terms of point forecasts and risk characterization, a key positive feature for decision makers who are tasked with forecasting risk and quantiles, especially under critical situations such as economic crises.

Fig. 3.4 compares on-line 1-step ahead forecast standard deviations. Economic (and other) decision makers are often faced with forecasts that have large forecast uncertainties; while honest in reflecting uncertainties, resulting optimal decisions may then be so unreliable as to be useless. Large economic models that require complex estimation methods, but have useful properties for policy makers, often produce large forecast standard deviations that might come from the complexity of the model, data, estimation method, or all of the above without necessarily knowing the source of uncertainty. BPS, on the other hand, synthesizes the forecasts

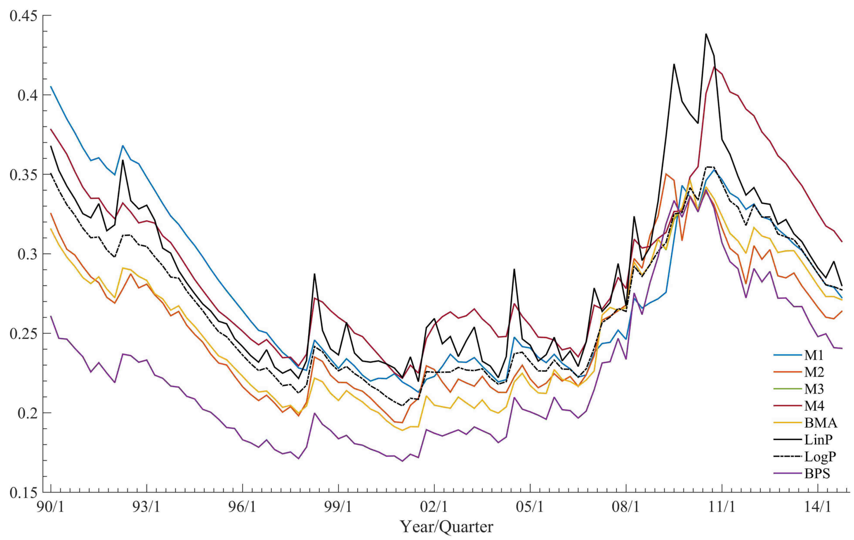

FIGURE 3.4: US inflation rate forecasting 1990/Q1-2014/Q4: 1-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}100$ quarters.

and by doing so, has the ability to *decrease* forecast uncertainties relative to the agents, without overly underestimating real risks; this is evident in the example here, where BPS leads the agents (and other strategies) in terms of LPDR performance. Fig. 3.4 shows that some part of this comes from generally reduced forecast uncertainties– coupled with more accurate point forecasts– at this 1-step horizon. We caution that reduced uncertainties are not always expected or achieved, as exemplified below.

Finally, BMA– characteristically– effectively degenerated, with posterior probabilities increasingly favoring agent M3, over the prior period 1977/Q2-1989/Q4; thus, at the start of the test period, BMA-based forecast densities are very close to those from M3 alone. BPS, on the other hand, allows for continual adaptation as agent models change in their relative forecasting abilities; over the test period, BPS tends more highly weight agent M2, notable in terms of the on-line estimates of BPS agent coefficients in $\boldsymbol{\theta}_{tj}$; see Fig. 3.5. An interesting point to note is how

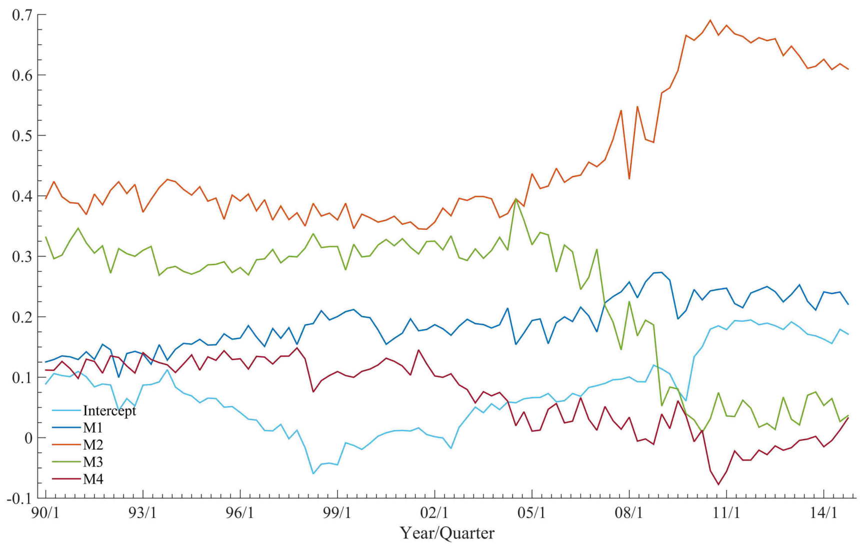

FIGURE 3.5: US inflation rate forecasting 1990/Q1-2014/Q4: On-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}100$ quarters.

BPS successfully adapts its coefficients during the sub-prime mortgage crisis by significantly down-weighting M3. As a dynamic model, BPS will not degenerate, continually allowing for "surprises" in changes in relative forecast performance across the agents.

Similarly, DeCo weights (see Fig. 3.6) heavily favor M3 from the training period and the weights– while dynamic– end up being very static. It is clear that the benefits of BPS is in its dynamic nature of its coefficients in stark contrast to DeCo. As the jagged coefficients indicate, BPS is constantly updating and calibrating without over-learning, partly due to discount learning (forgetting) and the latent biases/dependencies being effectively transferred to the coefficients. Though understanding what aspects of BPS (intercept, discounting, latent biases/dependencies, etc...) contribute to the gains seen in this example is beyond the scope of this dissertation, it is clear that they allow the coefficients to quickly adapt over time and improve forecasts.

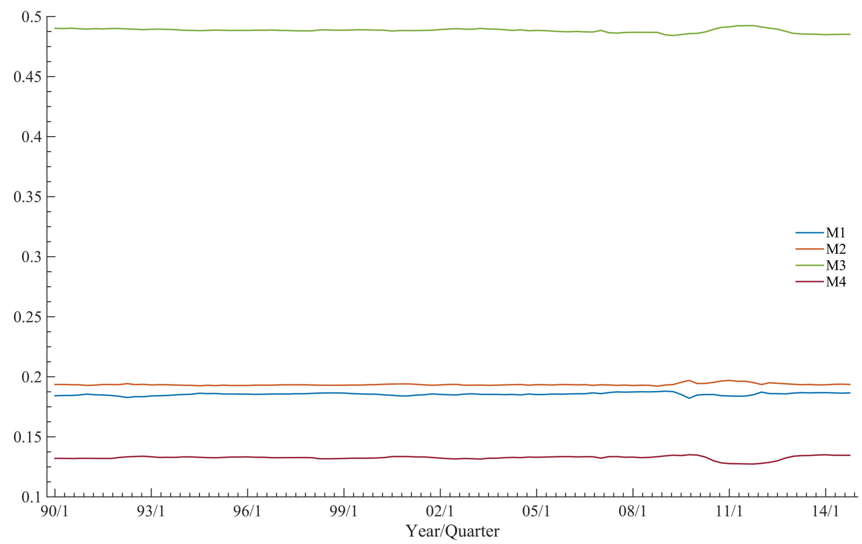

FIGURE 3.6: US inflation rate forecasting 1990/Q1-2014/Q4: On-line means of DeCo model weights sequentially computed at each of the $t = 1{:}100$ quarters.

3.2.2 4-step ahead forecasting

Figs. 3.7-3.11 summarize sequential analysis for 4-step forecasting, using both the direct extrapolation to 4-quarters ahead under the BPS model and the customized BPS(4) model. Each BPS strategy performs consistently better than agents and other strategies in point forecasting, while BPS(4) makes significant improvements in terms of both point and distribution forecasts compared to direct BPS extrapolation; see Figs. 3.7, 3.8, and 3.9. BPS performs relatively poorly in terms of LPDR as– being inherently calibrated to 1-step model fit– it fails to adequately represent the increased uncertainty associated with long term forecasts. Looking at the forecast standard deviations in Fig. 3.10, it is clear that BPS(4) is able to improve by adjusting to the increased forecast uncertainties. Then, even though forecast uncertainties increase substantially, they are clearly more than balanced by improved location forecasts as illustrated in Figs. 3.7 and 3.9. This again bears out the recommendation to directly synthesize forecasts on the horizon of interest rather than direct projection.

Fig. 3.11 shows on-line estimates of the BPS(4) coefficients $\boldsymbol{\theta}_t$ as they are sequentially updated and adapt over time during the test period. There is a notable reduction in adaptability over time relative to the 1-step BPS coefficients (Fig. 3.5); this is expected as the agents' forecasts are less reliable at longer horizons, so the data-based information advising the changes in posteriors over time is limited. The dynamic intercept term serves as a comparison base as it moves away from zero, playing a more active role in BPS(4) than in the 1-step case. Additionally, the 4-step ahead coefficient values (indicated here by just the on-line means) are quite different from 1-step coefficients, reasonably reflecting the differing forecasting abilities of the agents at differing horizons. BPS(4) is able to adapt to the 4-step ahead forecast, unlike the 1-step BPS, and dominate in performance compared to all other methods as a result.

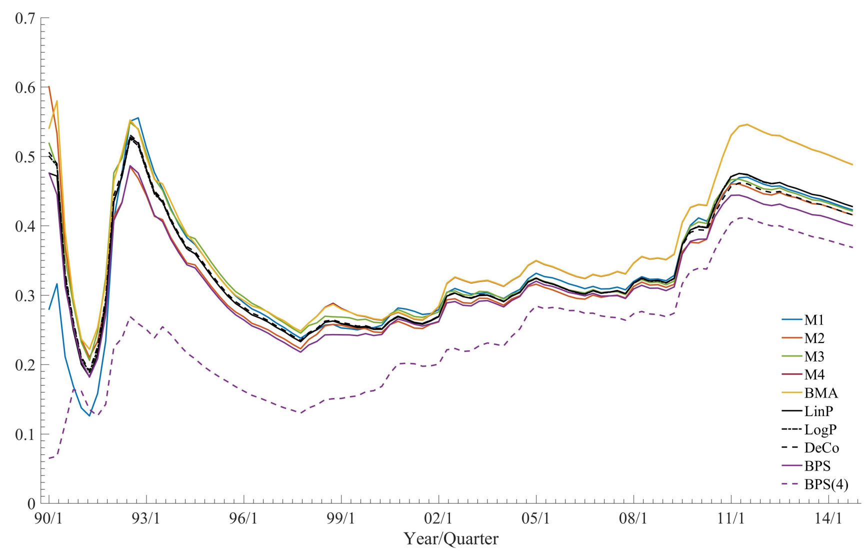

FIGURE 3.7: US inflation rate forecasting 1990/Q1-2014/Q4: Mean squared 4-step ahead forecast errors MSFE$_{1:t}(4)$ sequentially revised at each of the $t = 1{:}100$ quarters.

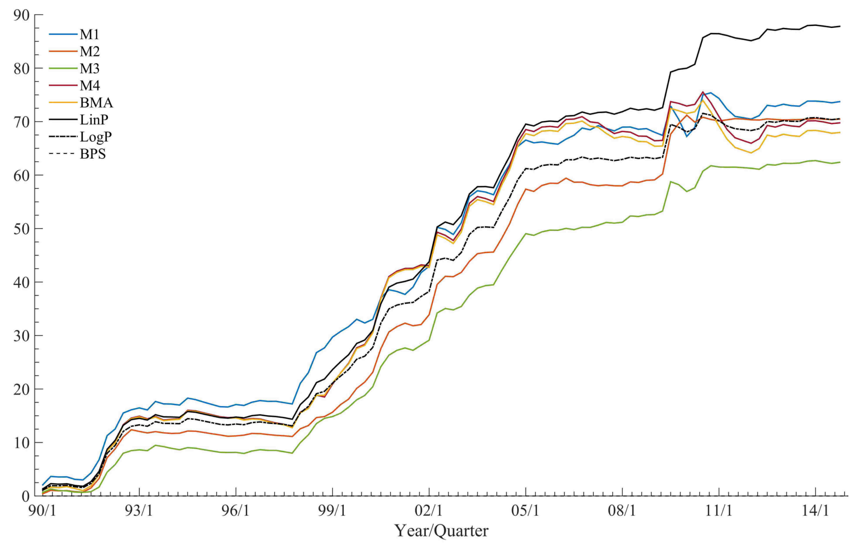

FIGURE 3.8: US inflation rate forecasting 1990/Q1-2014/Q4: 4-step ahead log predictive density ratios $\text{LPDR}_{1:t}(4)$ sequentially revised at each of the $t = 1{:}100$ quarters using direct projection from the 1-step ahead BPS model (baseline at 0 over time).

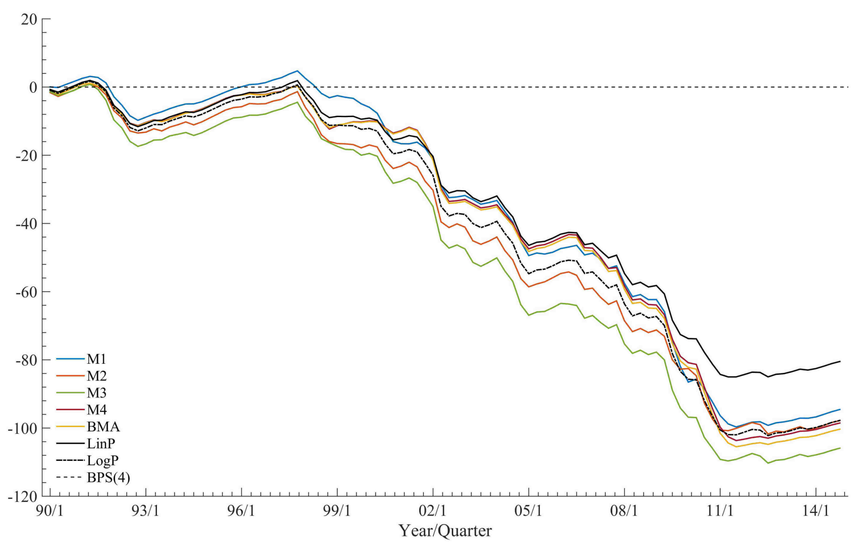

FIGURE 3.9: US inflation rate forecasting 1990/Q1-2014/Q4: 4-step ahead log predictive density ratios, $\text{LPDR}_{1:t}(4)$ sequentially revised at each of the $t = 1{:}100$ quarters using the 4-step ahead customized BPS(4) model (baseline at 0 over time).

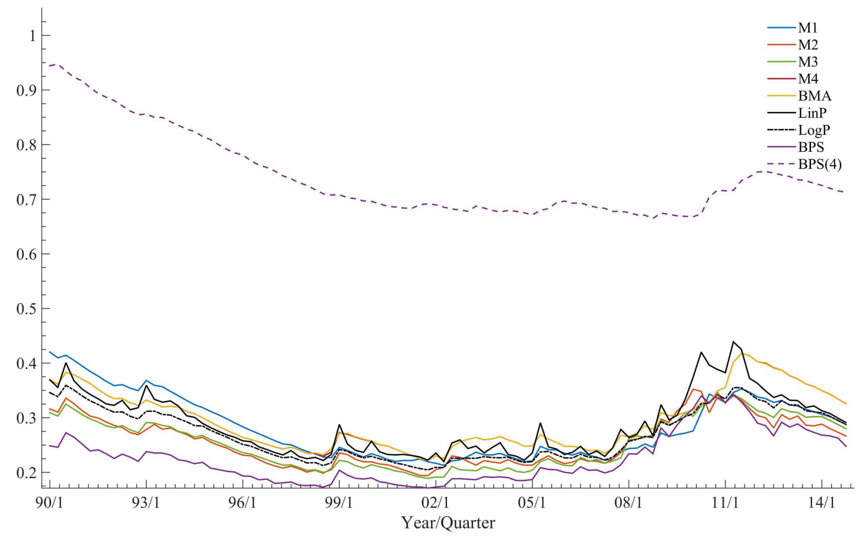

FIGURE 3.10: US inflation rate forecasting 1990/Q1-2014/Q4: 4-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}100$ quarters.

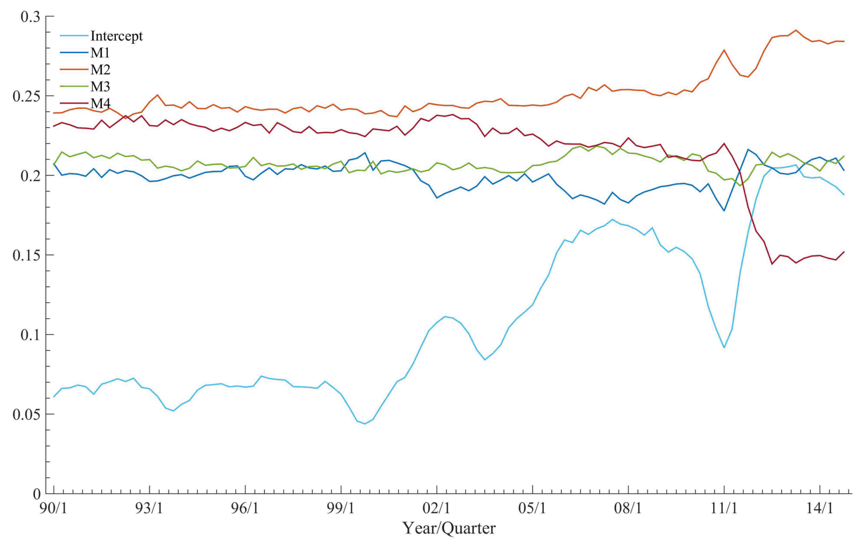

FIGURE 3.11: US inflation rate forecasting 1990/Q1-2014/Q4: On-line posterior means of BPS(4) model coefficients sequentially computed at each of the $t = 1{:}100$ quarters.

3.3 Retrospective analysis

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

BPS Coefficients. Retrospective posteriors for BPS (1-step) and BPS(4) model coefficients are summarized in Figs. 3.12 and 3.13, respectively, to compare with on-line point summaries in Figs. 3.5 and 3.11 earlier discussed. We see the expected smoothing of estimated trajectories of coefficients. To the extent that the role of the intercept terms can be regarded as reflecting (lack of) effectiveness of the synthesized models, these figures confirm that the agents' predictions are much more questionable at 4-steps ahead than at 1-step ahead. Intercepts increase up to and during the sub-prime mortgage crisis due to the increased inability for the models to forecast well during this time.

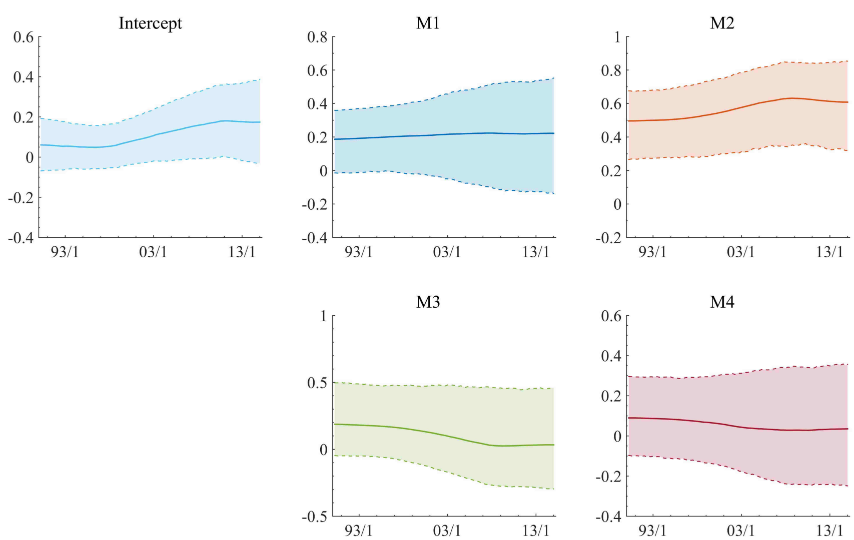

FIGURE 3.12: US inflation rate forecasting 1990/Q1-2014/Q4: Retrospective posterior trajectories of the BPS model coefficients based on data from the full $t = 1{:}100$ quarters. Posterior means (solid) and 95$\%$ credible intervals (shaded).

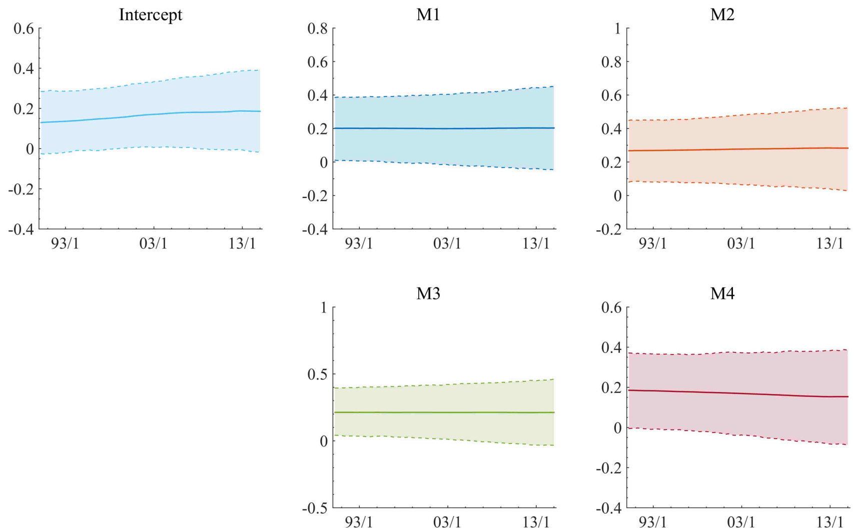

FIGURE 3.13: US inflation rate forecasting 1990/Q1-2014/Q4: Retrospective posterior trajectories of the BPS(4) model coefficients based on data from the full $t = 1{:}100$ quarters. Posterior means (solid) and 95$\%$ credible intervals (shaded).

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

As noted earlier, the conceptual and theoretical basis of BPS allows direct investigation of agent dependencies, as the inherent latent agent states x_{tj}– when inferred based on the observed data– carry the relevant information. From the full MCMC analysis to the end of the test data period, we have full posterior samples for the states $\boldsymbol{x}_t$– in both the direct BPS and customized BPS(4). For illustration, we focus on the 1-step BPS analysis; Fig. 3.14 displays posterior trajectories for the x_{tj}, together with the inflation outcomes y_t; Fig. 3.15 is similar, but vertically centers the display by plotting trajectories for the forecast deviations $y_t - x_{tj}$. The patterns over time in each of these reflect the strong, positive dependencies among agents that are to be expected given the nature of the agent models.

To explore dependencies, we simply investigate the posterior for $\boldsymbol{x}_{1:T}$. This is not a standard form and is represented in terms of the MCMC-based posterior sample. One simple set of summaries is based on just computing empirical R^2 measures: from the MCMC sample, compute the approximate posterior variance matrix of $\boldsymbol{x}_t$ at each t, and from that extract implied sets of conditionals variances of any x_{tj} given any subset of the other $x_{ti}, i \neq j$. We do this for $i = 1{:}J \backslash j$, defining the MC-empirical R^2 for agent j based on all other agents, i.e., measuring the redundancy of agent j in the context of all J agents– the *complete conditional dependencies*. We do this also using each single agent $i \neq j$, defining paired MC-empirical R^2 measures of how dependent agents i, j are– the *bivariate dependencies*. Fig. 3.16 and 3.17 displays trajectories over time for these two measures for each horizon.

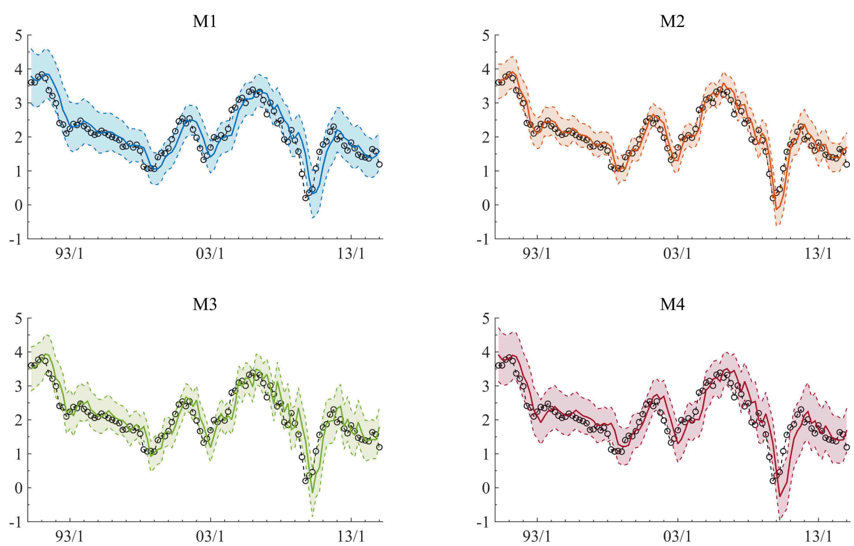

FIGURE 3.14: US inflation rate forecasting 1990/Q1-2014/Q4: BPS model-based posterior trajectories of the latent agent states x_{tj} for $j = 1{:}4$ over the $t = 1{:}100$ quarters. Posterior means (solid) and 95% credible intervals (shaded) from the MCMC analysis, with data $y_t \equiv p_t$ (circles).

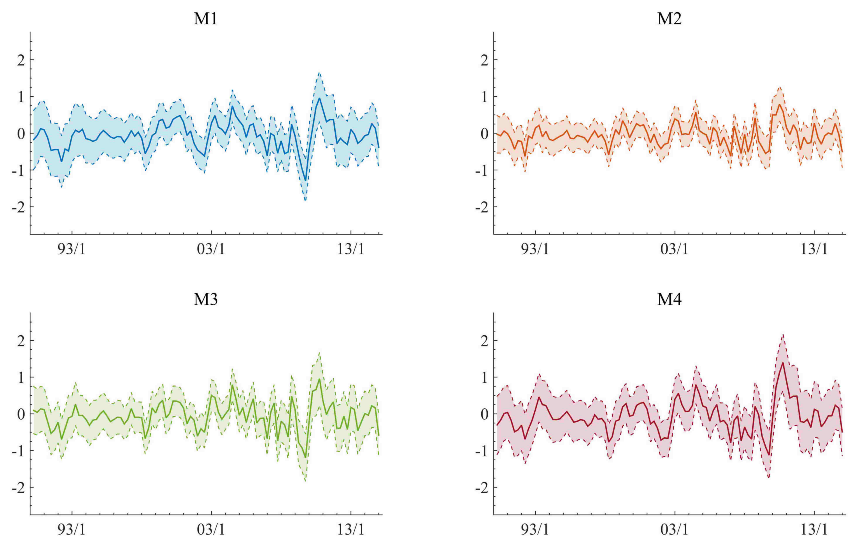

FIGURE 3.15: US inflation rate forecasting 1990/Q1-2014/Q4: BPS model-based posterior trajectories of the error in latent agents states $y_t - x_{tj}$ for $j = 1{:}4$ over the $t = 1{:}100$ quarters. Posterior means (solid) and 95% credible intervals (shaded) from the MCMC analysis.

Overall, we see high complete conditional dependencies at both forecast horizons, as expected due to the nature of the 4 models and their evaluation on the same data. Dependencies are substantial and much higher for 1-step forecasts than for 4-step ahead forecasts, reflecting decreasing concordance with increasing horizon, and all decrease over the test period. The predictability of M2 based on the others drops at a greater rate after about the start of 2002, in part due to poorer and less reliable performance during the dot-com crisis. The paired measures are all very low compared to the complete conditionals. Concordance of M2 and M3 decreases for 1-step but increases slightly for 4-step ahead forecasts, reflecting dynamics in relationships that differ with forecast horizon; from earlier discussion of forecast accuracy, this can be explained by how, in 1-step ahead forecasts, M2 improves while M3 deteriorates during the sub-prime mortgage crisis. For 4-step ahead forecasts (Fig. 3.17), forecast errors converge between the two, explaining the increase in concordance as all models performed equally poorly.

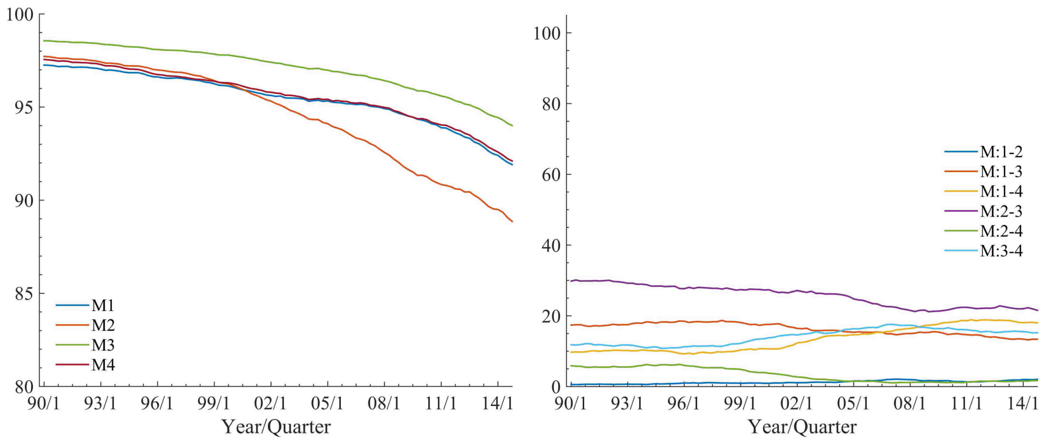

FIGURE 3.16: US inflation rate forecasting 1990/Q1-2014/Q4: BPS model-based trajectories of 1-step ahead MC-empirical R^2 (left) and paired MC-empirical R^2 (right) in the posterior for the latent agent states x_{tj} for $j = 1{:}4$ over the $t = 1{:}100$ quarters.

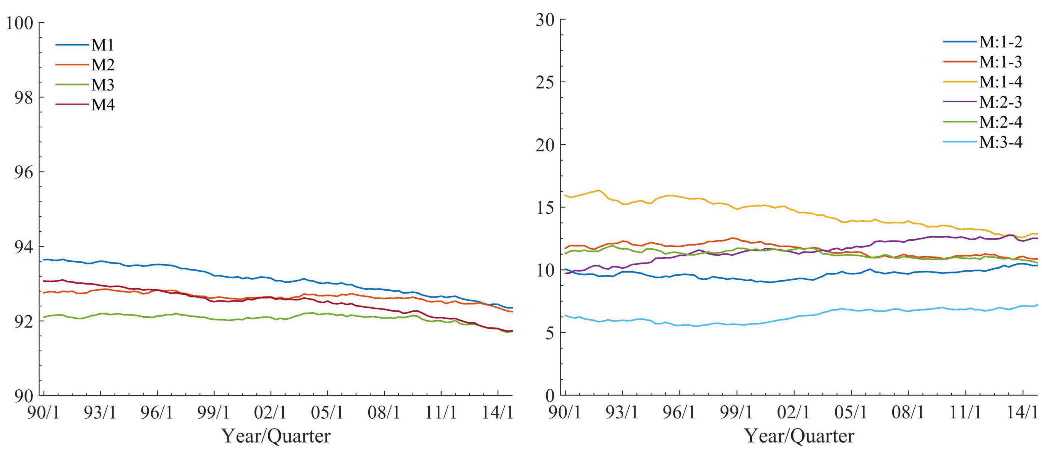

FIGURE 3.17: US inflation rate forecasting 1990/Q1-2014/Q4: BPS(4) model-based trajectories of 4-step ahead MC-empirical R^2 (left) and paired MC-empirical R^2 (right) in the posterior for the latent agent states x_{tj} for $j = 1{:}4$ over the $t = 1{:}100$ quarters.

3.4 Summary

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

4

Multivariate BPS: Macroeconomic Time Series

4.1 US Macroeconomic Time Series

Many critical problems in decision making involve multiple series. For these problems, the cross-series dependence structure plays a crucial role in making informed forecasts, inference, and decisions. In economic policy making, dependencies among macroeconomic series provide fundamental insight into the state of economy, improve forecasts over multiple horizons, and measure/forecast the impact of policy decisions made. Multivariate models, ranging from vector autoregressive models (VAR) to dynamic stochastic generalized equilibrium models (DSGE), have been developed and used in response to this reality. There is a vast literature in macroeconomics and econometrics, from the early works of Sims (1993), Stock and Watson (1996), Sims and Zha (1998) to some recent advances in dynamic Bayesian models in Cogley and Sargent (2005), Primiceri (2005), Benati and Surico (2008), Koop et al. (2009), Nakajima and West (2013a), Koop et al. (2010), and others that address this exact problem for the purpose of making informed economic policy decisions.

In terms of model comparison, calibration, and combination, the literature is dominated by model comparison and variable selection (e.g. Chan et al., 2012; Korobilis, 2013; Nakajima and West, 2013a), while little attention and development has been given to model combination. The existing literature on multivariate model combination (Andersson and Karlsson, 2008; Amendola and Storti, 2015) is limited to direct extensions of univariate methods, where models are combined linearly using one metric for the overall performance of a model. This is limiting in multiple ways, though most critically it lacks the flexibility necessary for multivariate forecasting. For example, some agents (models or individuals) might be good at forecasting one series but poor in another, or do not forecast some series at all. The strategies in the literature are limited in assessing the average performance across series when the full set of multivariate forecasts exist.

Multivariate BPS provides a flexible and coherent structure for multivariate forecasts to be synthesized. For our application, we forecast a topical 6-series macroeconomic data set using models echoing the decision process of economic policy makers.

4.1.1 Data

We analyze monthly US macroeconomic data, focusing on forecasting 6 macroeconomic time series with 1-, 6-, 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. 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. Fig. 4.1 shows the data for the 6 series over the time span considered. We focus on forecasting the 6, with an emphasis on inflation, using past values of the 6 indices as candidate predictors underlying a set of 5 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. As this period exhibited such a sharp shock to the US economy, it tests the predictive ability of any model and strategy under great stress. For a 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, there is also interest in inferring the dependencies between series before and after the crisis; something that is unavailable with a univariate series.

4.1.2 Agents and BPS specifications

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. Labeling them 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 between 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) with the residual volatility following a matrix-beta/Wishart random walk so that model fitting and generation of forecasts is routine. Though more state-of-the-art models (for example, Bayesian TVP-VARs with stochastic volatility seen in Nakajima and West (2013a)) are available, the benefit and appeal of fore-

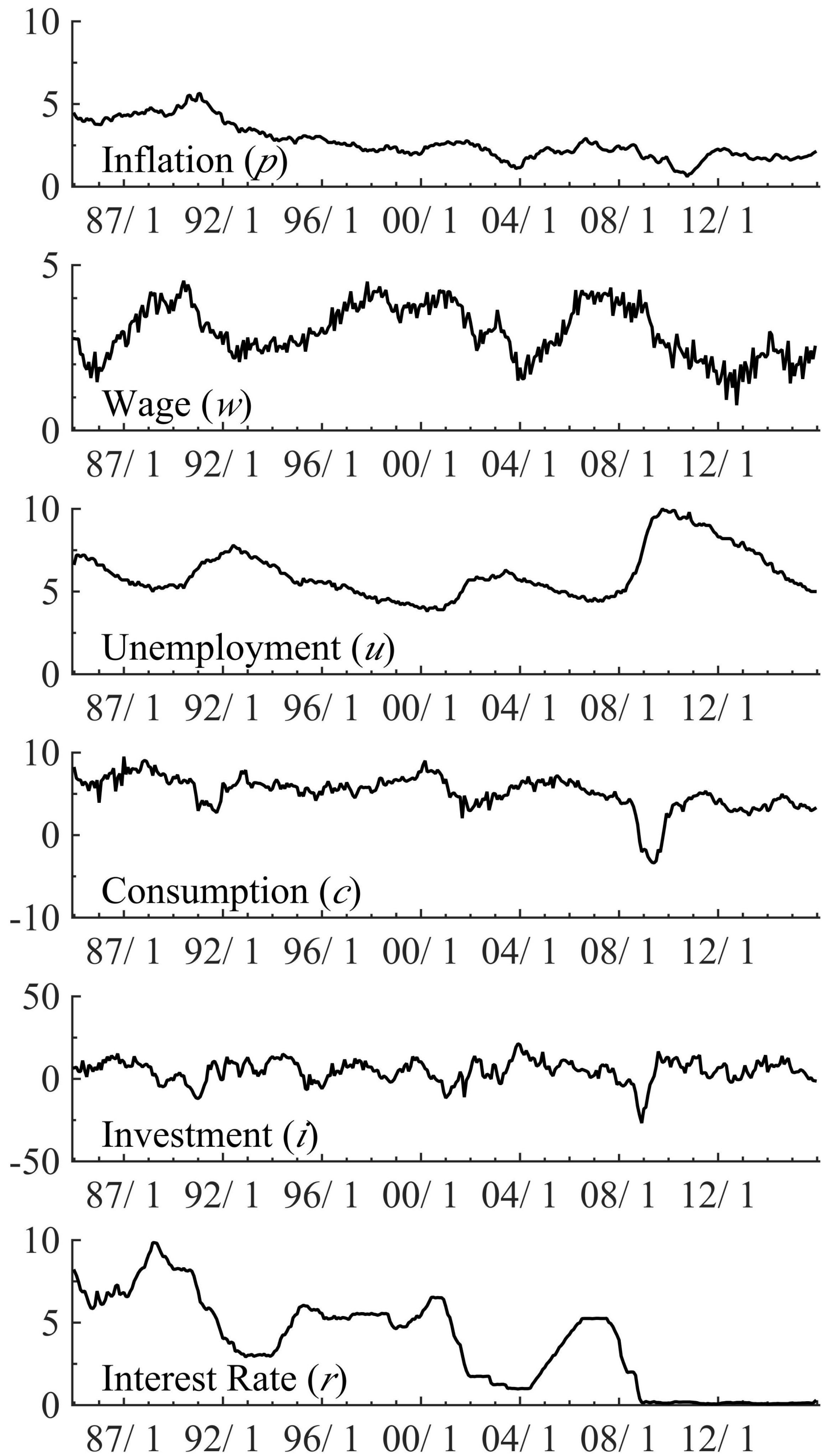

FIGURE 4.1: US macroeconomic data 2001/1-2015/12: US macroeconomic time series (indices $\times 100$ for $\%$ basis): annual inflation rate (p), wage (w), unemployment rate (u), consumption (c), investment (i), and short-term nominal interest rate (r).

cast synthesis is making improvements over a set of relatively simple model and not resorting to complex models that are hard to estimate/calibrate.

Prior specifications for the TVP-VAR state vector and discount volatility model in each is based on $\boldsymbol{\Theta}_0|\boldsymbol{V}_0 \sim N(\boldsymbol{M}_0, \boldsymbol{V}_0)$ and $\boldsymbol{V}_0 \sim \boldsymbol{W}^{-1}(12, 12\boldsymbol{I})$, where $\boldsymbol{M}_0 = [\mathbf{0}; 0.99\boldsymbol{I}]$, using the usual $(\boldsymbol{\Theta}, \boldsymbol{V})$ notation (Prado and West, 2010, Chap 10). Each agent model uses standard discount factor (β) specification for state evolution variances and discount factor (δ) for residual volatility; we use $(\beta, \delta) = (0.99, 0.98)$ in each of these agent models. The forecast densities $h_{t-k,j}(\boldsymbol{x}_{tj})$ are then those of predictive T distributions.

In the dynamic BPS models for forecast horizons $k = 1, 6, 12, 24$, we take initial priors as $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}, \boldsymbol{C})$ with $\boldsymbol{m} = (0, \mathbf{1}'/p)'$ and $\boldsymbol{C} = diag(0.001, \mathbf{1})$ and $\boldsymbol{V}_0 \sim \boldsymbol{W}^{-1}(7, 0.01\boldsymbol{I})$. BPS for 1-step ahead forecasting is based on $(\beta, \delta) = (0.99, 0.95)$, as with BPS$(k)$ for $k = 6, 12, 24$-month ahead forecasting, as discussed in Section 2.2.3.

We have explored analyses across ranges of priors and discount factors, and chosen these values as they lead to good agent-specific and BPS forecasting accuracy; conclusions with respect to BPS do not change materially with different values close to those chosen for the summary examples.

4.1.3 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 agent VAR models. 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 $k = 6, 12, 24$-step ahead focused BPS(k)

model, as discussed in Section 2.2.3. 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 of 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.

4.1.4 Forecast and accuracy and comparisons

As in Section 3.1.3, 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 errors (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

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

where $p_j(\boldsymbol{y}_{t+k}|\boldsymbol{Y}_{1:t})$ is the predictive density under each agent indexed by j, baselined against the corresponding BPS forecasts at this horizon. 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 recently used in similar studies (Nakajima and West, 2013a; Aastveit et al., 2016).

4.2 Dynamic BPS and forecasting

4.2.1 1-step ahead forecasting

Table 4.1 summarizes the predictive measures compared for the 1-step ahead forecasts. Looking at point forecasts, BPS exhibits improvements over 4 out of the 6 series to the agent models; the 4 being inflation, wage, consumption, and investment. Even for the 2 series for which BPS does not show improvement over the models, the difference between the best model is within 1%. On the series BPS makes an improvement, the gains are at least 2%. 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, where a model is assessed on overall performance, it would sacrifice the accuracy for one series for another. 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 series over another.

As with the univariate example in Chapter 3, BPS substantially improves characterization of forecast uncertainties as well as adaptation in forecast locations, reflected in the LPDR measures. Note that the best model, in terms of LPDR, is only best for wage in terms of MSFE and performs average for the other series. This indicates how LPDR measures for multiple series favor overall performance over models that are good for some but bad for others. Model combination schemes that are dependent on likelihood measures (BMA, for example), therefore, would have heavily favored the average performing model. BPS dynamically synthesizes each series, while improving uncertainty assessment (per series and dependence between series), to improve in terms of overall distribution forecasts as well as point forecasts. This feature of BPS is critical, as $\mathcal{D}$ typically has priorities in the series being forecasted.

Table 4.1: 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 BPS. Note: $\text{LPDR}_{1:T}$ is relative to BPS.

	$\text{MSFE}_{1:T}$					
1-step	Infl	%	Wage	%	Unemp	%
VAR(1)	0.1187	-4.4572	0.3800	-16.7632	0.1435	0.4466
VAR(12)	0.1265	-11.3291	0.3331	-2.3560	0.1515	-5.1460
VAR(3)	0.1214	-6.8500	0.3324	-2.1185	0.1480	-2.7157
VAR(1:3:9)	0.1162	-2.2800	0.3461	-6.3567	0.1491	-3.4446
VAR(1:6:12)	0.1170	-2.9578	0.3807	-16.9740	0.1467	-1.7517
BPS	0.1136	-	0.3255	-	0.1441	-

	$\text{MSFE}_{1:T}$					
1-step	Cons	%	Invest	%	Interest	%
VAR(1)	0.6251	-2.3135	3.6357	-2.0875	0.1657	-14.8758
VAR(12)	0.6853	-12.1708	3.9188	-10.0371	0.1568	-8.6896
VAR(3)	0.6311	-3.2879	3.6498	-2.4834	0.1452	-0.6731
VAR(1:3:9)	0.6363	-4.1505	3.7272	-4.6560	0.1430	0.8887
VAR(1:6:12)	0.6237	-2.0753	3.6647	-2.9023	0.1509	-4.6144
BPS	0.6110	-	3.5614	-	0.1442	-

1-step	$\text{LPDR}_{1:T}$
VAR(1)	-86.1728
VAR(12)	-105.3492
VAR(3)	-43.8210
VAR(1:3:9)	-46.4023
VAR(1:6:12)	-63.0610
BPS	-

We further our analysis by reviewing summary graphs showing aspects of analyses evolving over time during the testing period, a period that includes challenging economic times that impede good predictive performance. Figs. 4.2-4.15 summarize sequential analysis for 1-step forecasting.

Fig. 4.2 shows the 1-step ahead measures $\text{MSFE}_{1:t}(1)$ for each time t for inflation. The other series are omitted for the sake of brevity, but the patterns in inflation are consistent. Additionally, forecasting inflation is one of the most im-

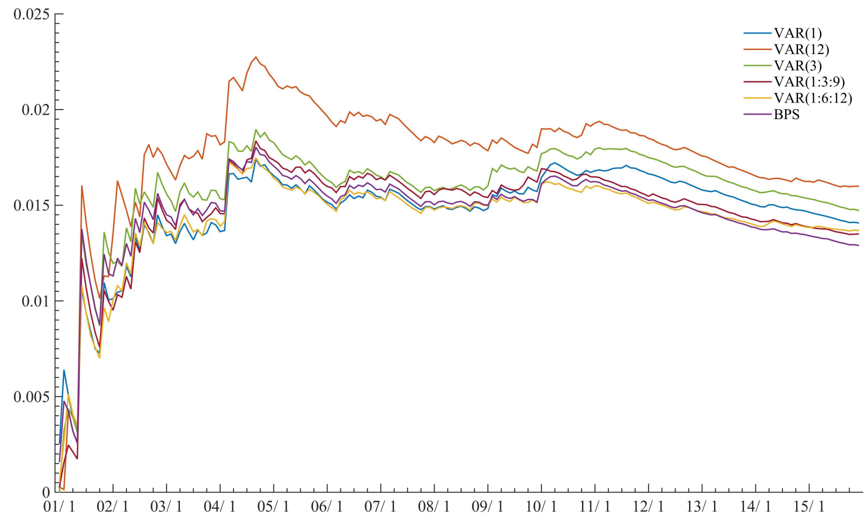

FIGURE 4.2: US macroeconomic forecasting 2001/1-2015/12: Mean squared 1-step ahead forecast errors $\text{MSFE}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}180$ months for inflation.

portant 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. 4.3 confirms that BPS performs uniformly better than the other models based on LPDR measures that measure relative distributional form and dispersion of forecast densities as well as location. Compared to the quarterly data in Chapter 3, the sub-prime mortgage crisis has less of a clear impact on LPDR, though the effects can be seen impacting the measure during and after the shock. 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.

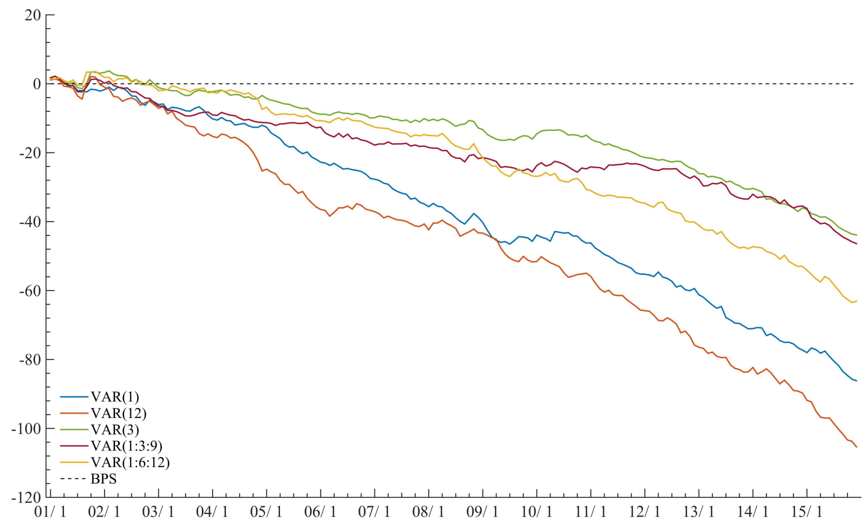

FIGURE 4.3: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead log predictive density ratios $\text{LPDR}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}180$ months. The baseline at 0 over all t corresponds to the standard BPS model.

One of the crucial aspects of the BPS synthesis function used is that it can adapt its coefficients specific to the series. Figs. 4.4-4.9 are 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 are drastically different from each other. For example, the model with the highest coefficient is different from each series, reflecting how certain models are better at forecasting different series.

Focusing on inflation (Fig. 4.4), the coefficients clearly exhibit a 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 quickly drop off, replaced by VAR$(1{:}3{:}9)$ and VAR$(1{:}6{:}12)$, which 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.

Wage (Fig. 4.5), on the other hand, stays relatively stable over time. VAR(3), a quarter worth of lags, has the highest coefficient and stays the highest throughout.

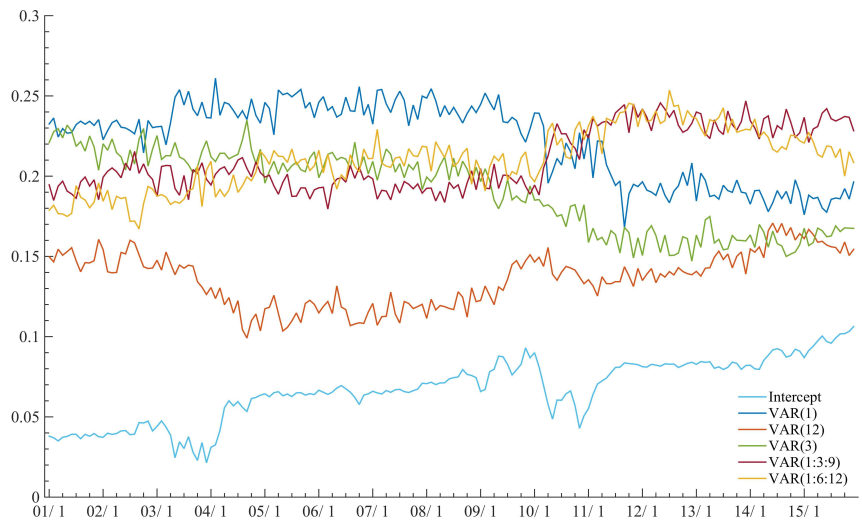

FIGURE 4.4: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead on-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}180$ months for inflation.

In comparison to inflation, VAR(1), the simplest model, stays close to zero, while VAR(12), the most complex model, is in the negative.

Fig. 4.6, coefficients for unemployment, exhibits an increase in VAR(1) and decrease in VAR$(1{:}3{:}9)$ after the sub-prime mortgage crisis. 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.

Coefficients for consumption and investment (Figs. 4.7-4.8) are perplexing, as there is almost no signal to be read off the coefficients. There are multiple reasons why this might occur. The biggest reason is the high variation in the two series. The forecasts from these models, therefore, have large forecast uncertainties that in turn make the BPS coefficients unstable.

Finally, interest rate (Fig. 4.9) coefficients favor more complex models with longer lags. Interestingly, we see a gradual decrease in VAR(1) up until the sub-prime mortgage crisis, at which point it stays level. Long term dynamics, we can

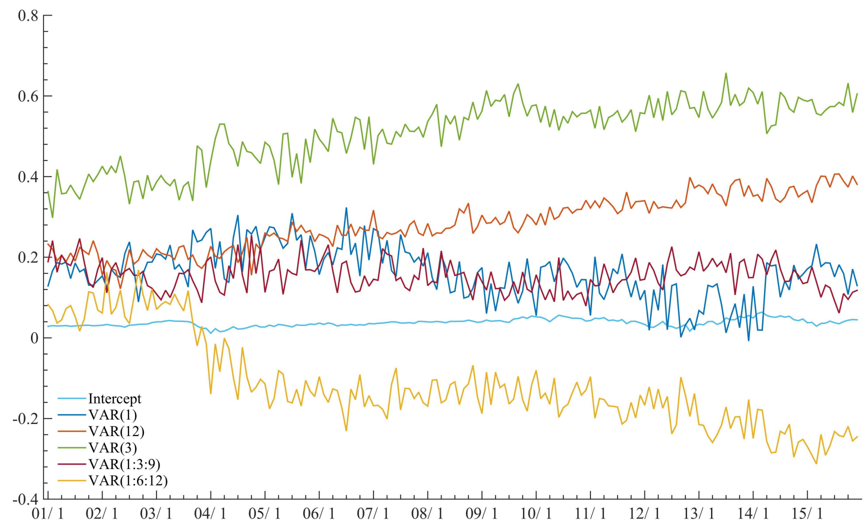

FIGURE 4.5: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead on-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}180$ months for wage.

infer, 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.

Looking at misspecification, i.e., the sum of on-line posterior means of the sum of BPS model coefficients, most of the series are relatively stable around 1 (fully specified), with the exception being inflation and investment. The commonalities between the two is that both have a decreasing trend, indicating greater misspecification over the time period considered. This is expected, as the models used became increasingly unsuitable to the shifting economy over time.

Figs. 4.11-4.15 exhibit selected aspects of the trend in uncertainty and dependence between series over time. The forecast standard deviation (the diagonal elements of the forecast covariance matrix) displays how the uncertainty measures change over time; seen in Fig. 4.11. Complex models for multiple series that require estimation methods that are also complex, often produce large fore-

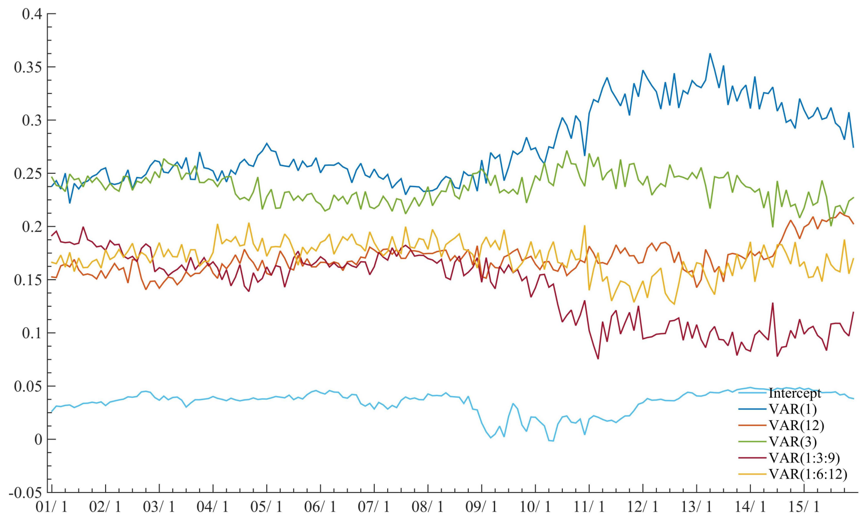

FIGURE 4.6: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead online posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}180$ months for unemployment.

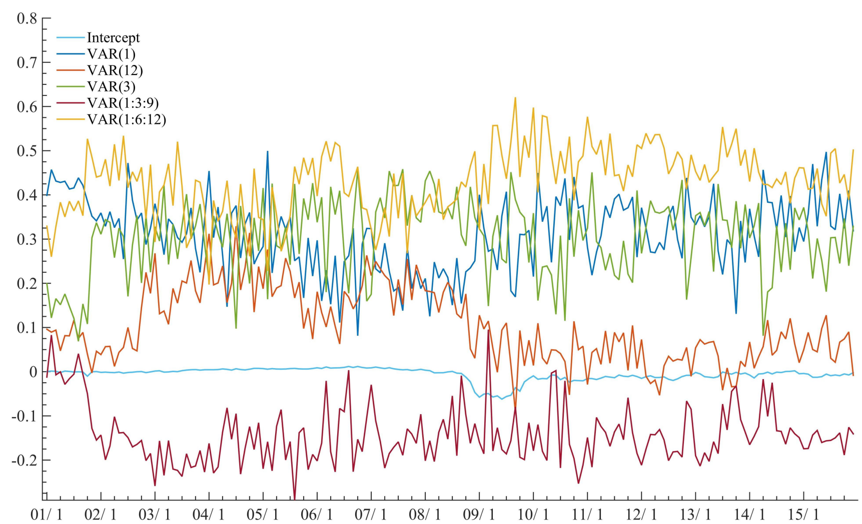

FIGURE 4.7: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead online posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}180$ months for consumption.

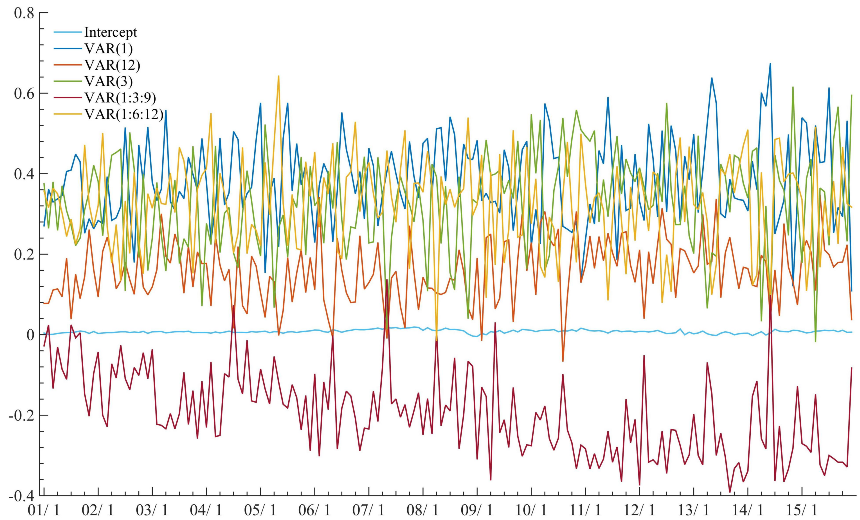

FIGURE 4.8: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead on-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}180$ months for investment.

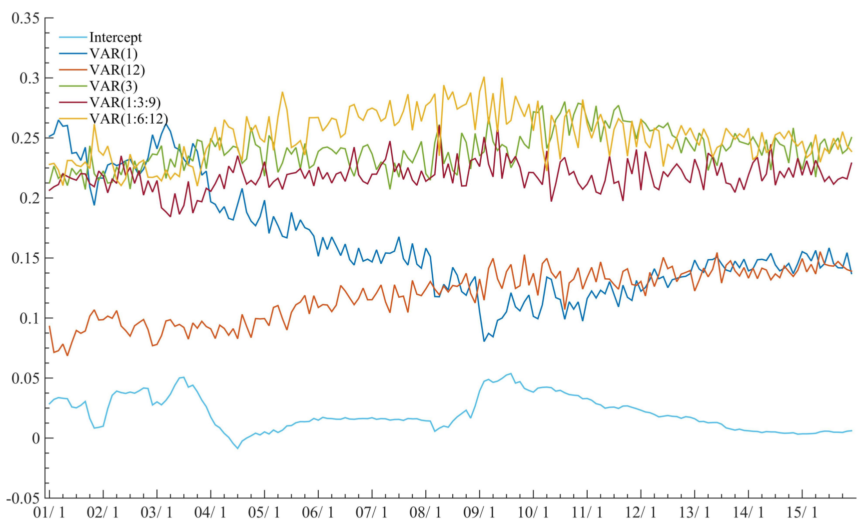

FIGURE 4.9: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead on-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}180$ months for interest rate.

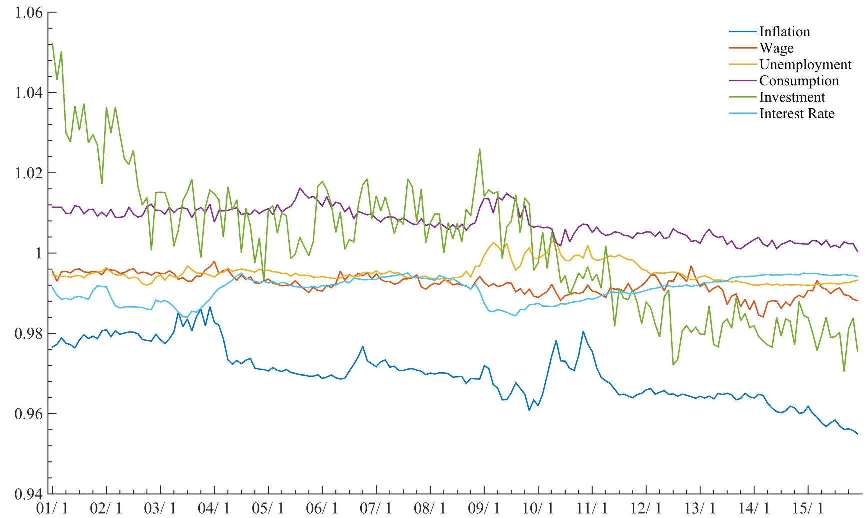

FIGURE 4.10: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead on-line posterior means of the sum of BPS model coefficients (misspecification) sequentially computed at each of the $t = 1{:}180$ months.

cast deviations coming from the model, data, estimation method, or all of the above. Large VAR models are popular in practice, due to its modeling flexibility and interpretability, but naturally lead to inflated uncertainty measures due to its complex nature. BPS, on the other hand, has smaller uncertainty forecasts by synthesizing the forecasts and decreasing the forecast uncertainty relative to the agents. Though underestimating real risk is as dangerous as overestimating it, the LPDR results indicate that the BPS uncertainty estimates are better than the agent forecast uncertainties.

Finally, we move our attention to the posterior BPS forecast dependencies among agent forecasts over time. The forecast dependence among agent forecasts is defined here as the on-line posterior of the covariance of $\boldsymbol{\theta}_t$ in eqn. (2.11b). As mentioned in Chapter 2, the posterior dependencies between agents are effectively transferred onto the dynamic coefficients, thus inspecting the dependence structure of $\boldsymbol{\theta}_t$ is in effect exploring the latent dependencies among agents. Rather than tracking the dependence over time, we look at the dependence at four specific time

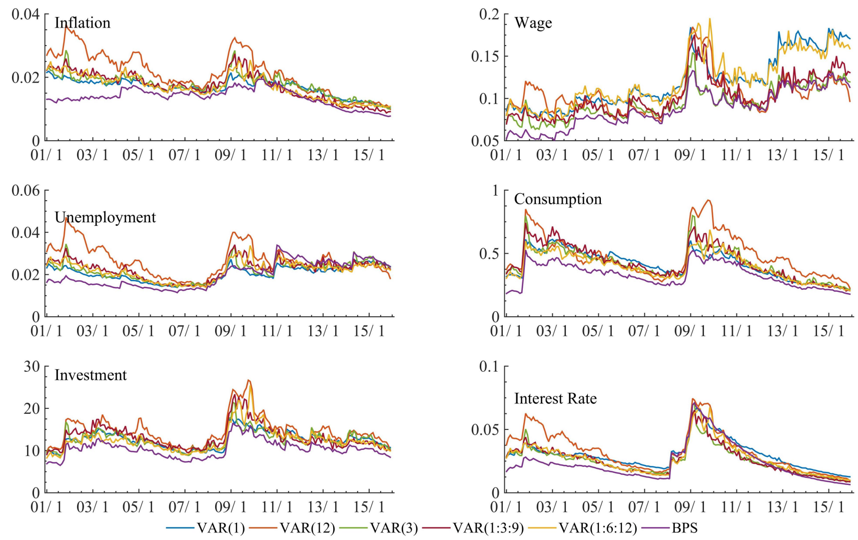

FIGURE 4.11: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}180$ months.

points that represent four different regimes in the testing period; pre-sub-prime mortgage crisis (2003/10), immediately before the crisis (2007/7), immediately after the crisis (2009/3), and post-crisis (2014/6). We also use the correlation instead of the covariance in order to standardize the matrix. The correlation is arranged in order of series and then agents within that series.

Figs. 4.12-4.15 exhibit this dependence among different time points. At the beginning of the testing period (Fig. 4.12), where the series are relatively stable, the dependence is almost zero across series, with some negative correlation within series. The negative correlation within series is expected. If a model coefficient increases, the other model coefficients decrease so the sum of coefficients stays stable from t to $t+1$. Immediately before the crisis (Fig. 4.13), a slight change can be noticed for the change in dependence across series, namely wage, consumption, and investment, and some within, e.g., consumption, with positive correlations. The difference from Fig. 4.12 is notable. The crisis has yet to set in and the change in dependence is already clear, with dependence across series becoming prevalent.

Immediately after the crisis (Fig. 4.14), more non-zero dependence appears across all agent forecasts, especially between unemployment, consumption, investment, and interest rate. Under considerable shock, the dependencies that started to appear in Fig. 4.13 become more rampant. Larger dependencies appear between consumption and investment, and the positive dependence in consumption becomes stronger. Post crisis (Fig. 4.15), we see a new dependence structure appearing, with greater dependence among wage, consumption, and investment, but weaker dependence otherwise compare to Fig. 4.14. Additionally, the positive correlation within consumption also has weakened compared to right after the crisis.

Following the trend of dependence among agents and across series over time provides important insight into how the economy changes over shocks and different regimes. Figs. 4.12 and 4.15 are both snapshots of relatively stable periods, yet the characteristics exhibited through the dependence are starkly different. This

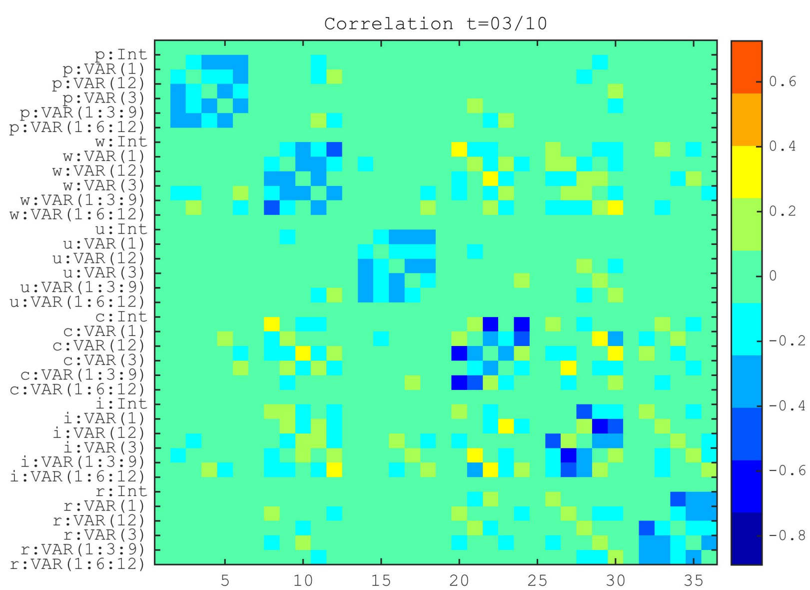

FIGURE 4.12: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead online posterior correlation of BPS model coefficients for 2003/10.

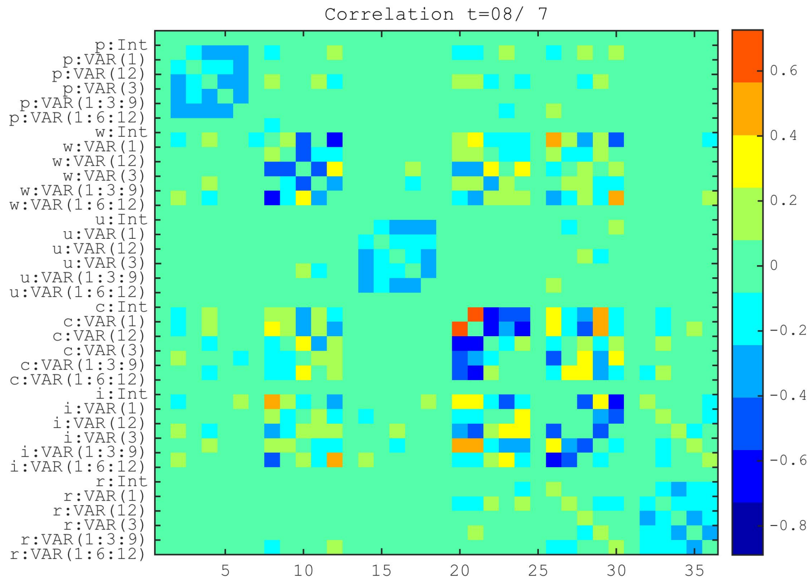

FIGURE 4.13: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead on-line posterior correlation of BPS model coefficients for 2007/07.

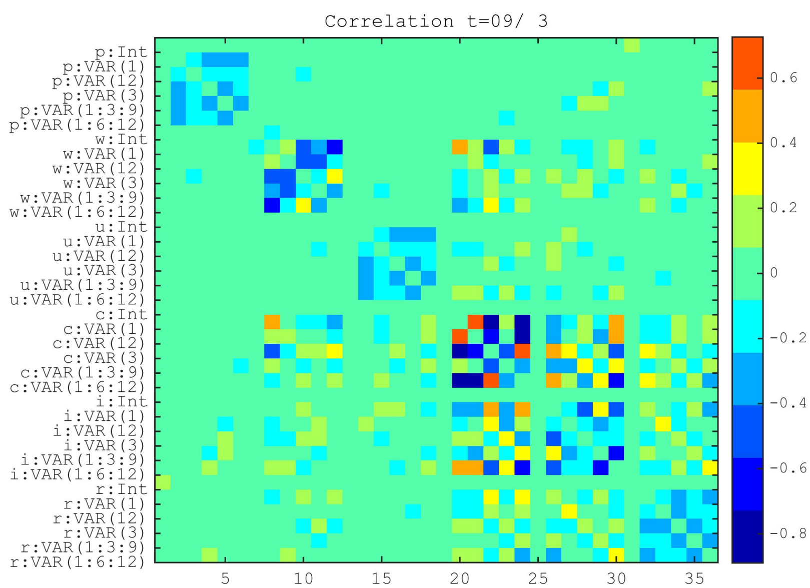

FIGURE 4.14: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead on-line posterior correlation of BPS model coefficients for 2009/03.

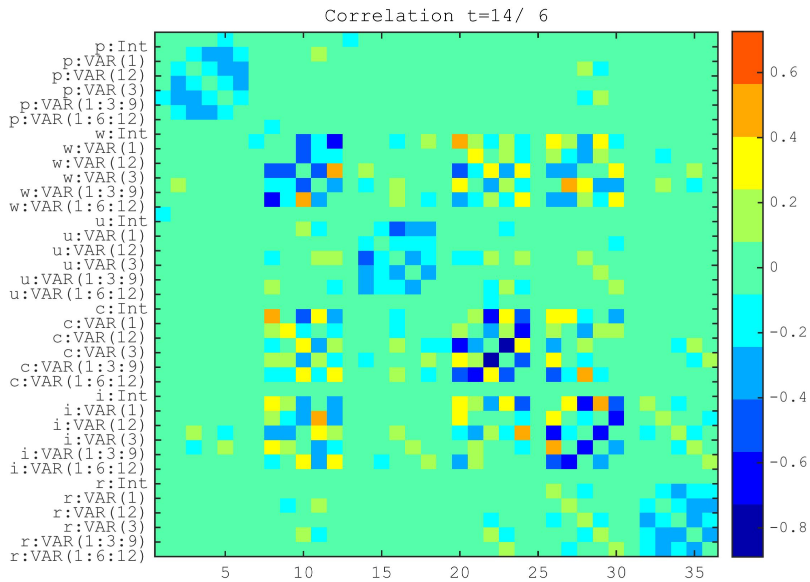

FIGURE 4.15: US macroeconomic forecasting 2001/1-2015/12: 1-step ahead on-line posterior correlation of BPS model coefficients for 2014/06.

difference in economic structure is not unexpected, though graphically visualizing the difference through the lens of agents provides new insight and perception into the overall change in economy.

4.2.2 k-step ahead forecasting

Long term forecasts in economic policy making is equally– if not more– important to 1-step ahead forecasting. For this study, we forecast 6- (half a year), 12- (year), and 24- (two years) step ahead to demonstrate the effectiveness of BPS over the set of agents.

Tables 4.2-4.4 summarizes the predictive measures compared for the three forecast horizons. For point forecasts, BPS(k) outperforms all other models for all series except for coming in second for wage. This improvement holds for all k-step ahead forecasts considered and the improvements of BPS(k) only increase with k. The improvement comes 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, which is one of the most important series for economic policy making, BPS(k) greatly improves on the agent models, improving over 130% on the best agent model. As with 1-step ahead forecasting, it is also notable that agent model performances greatly vary across series. Traditional model combination schemes, again, will fail to improve over all series by sacrificing one series over another, unlike BPS(k)

BPS(k) also significantly improves on the uncertainty forecasts, as evident in the comparison of LPDR. Creating long term uncertainty forecasts for multiple time series is a very difficult problem due to the nature of these models being designed for 1-step ahead forecasts and failing to propagate forward accurately. 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 solely on MSFE results over t for inflation. The characteristics of the results for inflation are similar to those of the other series and are omitted for the sake of brevity. Figs. 4.16-4.18 exhibit MSFE comparisons for inflation over the testing period for $k = 6, 12, 24$-step ahead forecasts. Although the scale is different for each k, there are notable common characteristics that define BPS(k). For example, agent models experience several large shocks in precision over the testing period, in particular the sub-prime mortgage crisis in the late 2000s. These shocks effect the precision of the agent models greatly, especially for 24-step ahead forecasts. In comparison, BPS(k) stays relatively robust throughout multiple shocks and structural breaks.

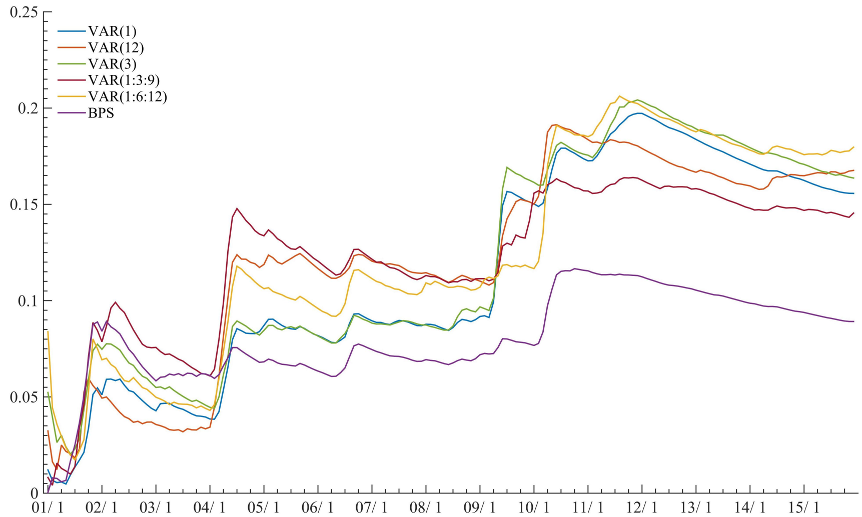

FIGURE 4.16: US macroeconomic forecasting 2001/1-2015/12: Mean squared 6-step ahead forecast errors $\text{MSFE}_{1:t}(6)$ sequentially revised at each of the $t = 1{:}180$ months for inflation.

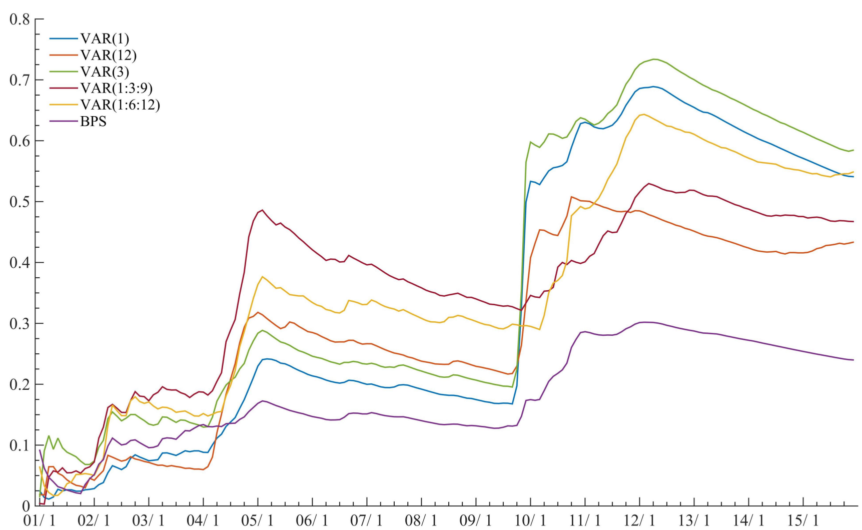

FIGURE 4.17: US macroeconomic forecasting 2001/1-2015/12: Mean squared 12-step ahead forecast errors $\text{MSFE}_{1:t}(12)$ sequentially revised at each of the $t = 1{:}180$ months for inflation.

Table 4.2: US macroeconomic forecasting 2001/1-2015/12: 6-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(6). Note: LPDR$_{1:T}$ is relative to BPS(6).

	MSFE$_{1:T}$					
6-step	Infl	%	Wage	%	Unemp	%
VAR(1)	0.3946	-32.1399	0.5112	-3.5642	0.4451	-0.5160
VAR(12)	0.4094	-37.1158	0.5767	-16.8366	0.5466	-23.4443
VAR(3)	0.4046	-35.4893	0.4908	0.5652	0.4522	-2.1229
VAR$(1{:}3{:}9)$	0.3814	-27.7328	0.5996	-21.4591	0.5316	-20.0687
VAR$(1{:}6{:}12)$	0.4239	-41.9655	0.6143	-24.4435	0.4895	-10.5437
BPS(6)	0.2986	-	0.4936	-	0.4428	-

	MSFE$_{1:T}$					
6-step	Cons	%	Invest	%	Interest	%
VAR(1)	1.3652	-2.0847	8.7751	-22.2396	1.0702	-10.8799
VAR(12)	2.1554	-61.1649	11.0582	-54.0433	1.0866	-12.5850
VAR(3)	1.3499	-0.9354	8.5996	-19.7944	1.0356	-7.2979
VAR$(1{:}3{:}9)$	1.7679	-32.1961	10.1307	-41.1240	1.0063	-4.2685
VAR$(1{:}6{:}12)$	1.8082	-35.2079	10.5685	-47.2220	1.1712	-21.3478
BPS(6)	1.3374	-	7.1786	-	0.9651	-

6-step	LPDR$_{1:T}$
VAR(1)	-171.0656
VAR(12)	-621.8982
VAR(3)	-487.3850
VAR$(1{:}3{:}9)$	-459.5934
VAR$(1{:}6{:}12)$	-413.7586
BPS(6)	-

Looking at LPDR evolutions over time (Figs. 4.19-4.21), BPS(k) improves over the agent models over all of the time period considered, except for slight increases in VAR(1), post crisis. The greatest gains for BPS(k) are made during the sub-prime mortgage crisis, as seen in the drastic drop after 2009. 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.

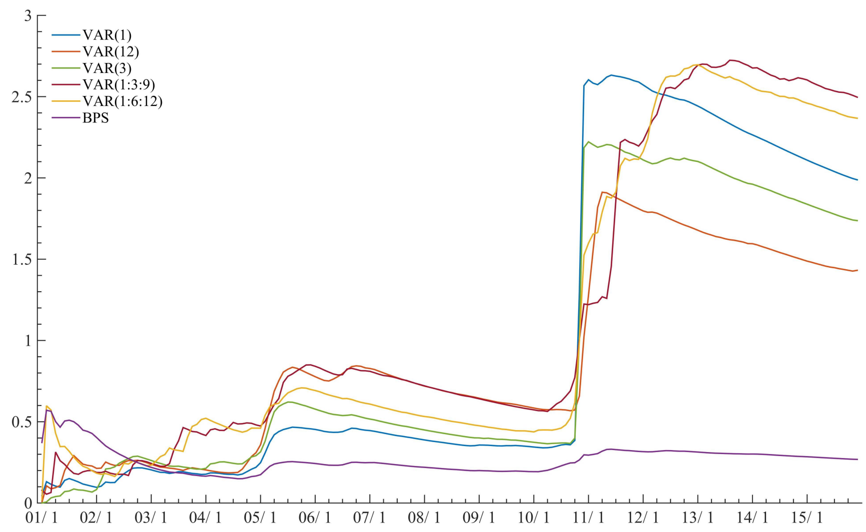

FIGURE 4.18: US macroeconomic forecasting 2001/1-2015/12: Mean squared 24-step ahead forecast errors $\text{MSFE}_{1:t}(24)$ sequentially revised at each of the $t = 1{:}180$ months for inflation.

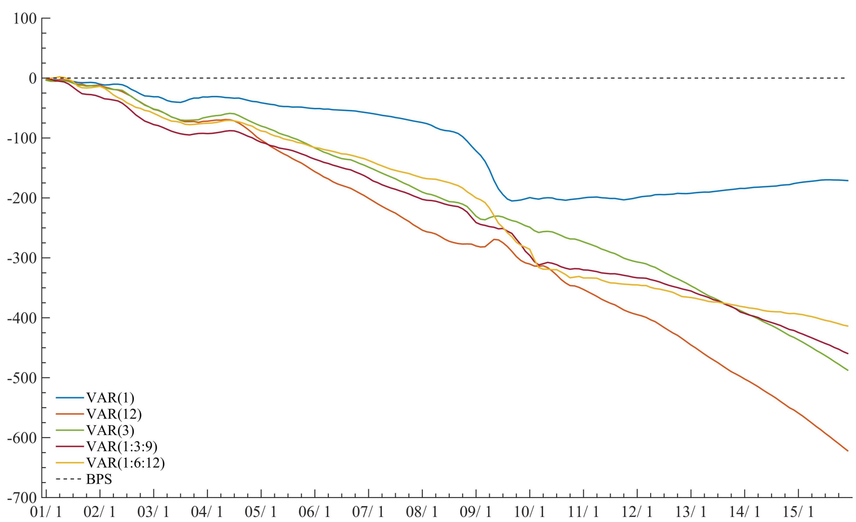

FIGURE 4.19: US macroeconomic forecasting 2001/1-2015/12: 6-step ahead log predictive density ratios $\text{LPDR}_{1:t}(6)$ sequentially revised at each of the $t = 1{:}180$ months. The baseline at 0 over all t corresponds to the standard BPS model.

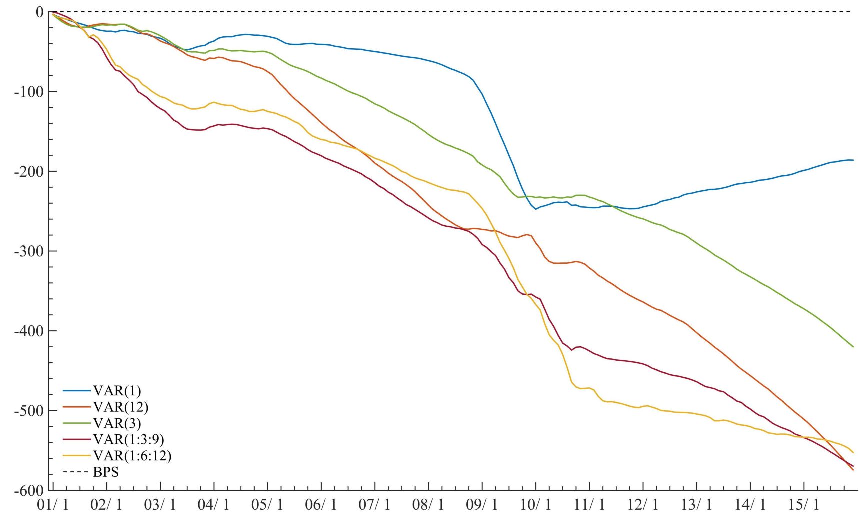

FIGURE 4.20: US macroeconomic forecasting 2001/1-2015/12: 12-step ahead log predictive density ratios $\mathrm{LPDR}_{1:t}(12)$ sequentially revised at each of the $t = 1{:}180$ months. The baseline at 0 over all t corresponds to the standard BPS model.

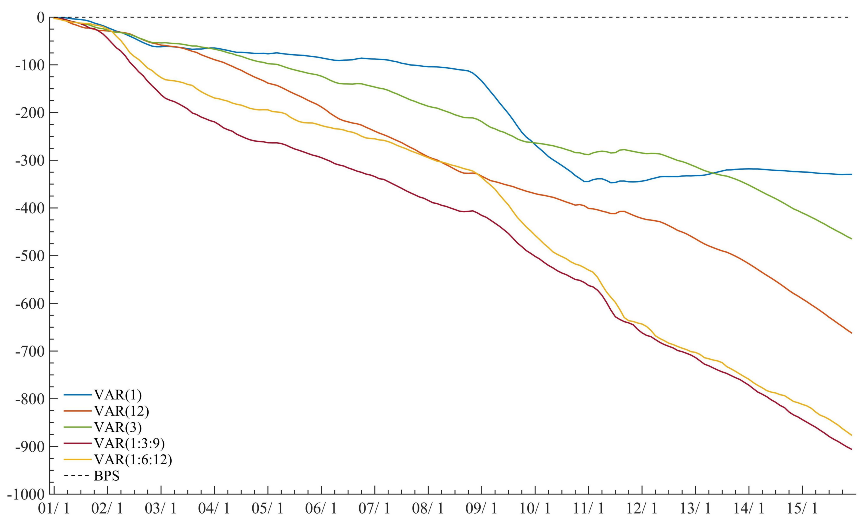

FIGURE 4.21: US macroeconomic forecasting 2001/1-2015/12: 24-step ahead log predictive density ratios $\mathrm{LPDR}_{1:t}(24)$ sequentially revised at each of the $t = 1{:}180$ months. The baseline at 0 over all t corresponds to the standard BPS model.

Table 4.3: 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)	0.7355	-50.0615	0.6675	1.5569	1.0944	-6.4387
VAR(12)	0.6582	-34.2938	0.8761	-29.2033	1.2724	-23.7507
VAR(3)	0.7645	-55.9703	0.7222	-6.5031	1.0787	-4.9131
VAR$(1{:}3{:}9)$	0.6721	-37.1169	1.0602	-56.3535	1.2837	-24.8491
VAR$(1{:}6{:}12)$	0.7366	-50.2885	0.9536	-40.6275	1.1722	-13.9975
BPS(12)	0.4901	-	0.6781	-	1.0282	-

	MSFE$_{1:T}$					
12-step	Cons	%	Invest	%	Interest	%
VAR(1)	2.6814	-21.8409	11.1198	-36.0292	2.3553	-25.4645
VAR(12)	4.2298	-92.1963	11.1509	-36.4094	2.4233	-29.0847
VAR(3)	2.6434	-20.1114	10.2069	-24.8615	2.2074	-17.5840
VAR$(1{:}3{:}9)$	3.2074	-45.7432	12.9534	-58.4592	2.1449	-14.2556
VAR$(1{:}6{:}12)$	3.1846	-44.7034	12.7985	-56.5647	2.7716	-47.6382
BPS(12)	2.2007	-	8.1746	-	1.8773	-

12-step	LPDR$_{1:T}$
VAR(1)	-186.0814
VAR(12)	-574.2636
VAR(3)	-419.9158
VAR$(1{:}3{:}9)$	-569.5800
VAR$(1{:}6{:}12)$	-552.4816
BPS(12)	-

Figs. 4.22-4.27 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 due redundancy. Overall, the BPS(k) coefficients are relatively stable compared to the 1-step ahead results, likely due to the lack of signal from the agent forecasts. The agent forecasts' ability for 24-steps are considerably worse than the 1-step ahead counterpart, 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.

The model misspecification, measured by the sum of on-line posterior means of BPS model coefficients (Fig. 4.28), is significantly larger than that of the 1-step ahead forecasts (Fig. 4.10), confirming how agent model misspecification increases over k. The exception to this is unemployment and consumption, which were stable for 1-step ahead forecasts as well (Fig. 4.10). This persistence is puzzling, as neither series can be specifically characterized as different (e.g. less volatile) from

Table 4.4: 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)	1.4100	-172.1416	1.0076	5.8680	2.7310	-19.2747
VAR(12)	1.1965	-130.9199	1.3878	-29.6519	2.8236	-23.3150
VAR(3)	1.3181	-154.3944	1.1666	-8.9925	2.3746	-3.7070
VAR(1:3:9)	1.5801	-204.9591	2.1644	-102.2126	3.2050	-39.9739
VAR(1:6:12)	1.5385	-196.9375	2.2139	-106.8317	2.8777	-25.6798
BPS(24)	0.5181	-	1.0704	-	2.2897	-

	MSFE$_{1:T}$					
24-step	Cons	%	Invest	%	Interest	%
VAR(1)	4.9316	-38.7673	12.3935	-45.6739	5.6585	-71.3629
VAR(12)	6.5373	-83.9519	12.6490	-48.6776	5.0113	-51.7629
VAR(3)	4.1476	-16.7072	9.1875	-7.9903	4.6347	-40.3577
VAR(1:3:9)	6.3717	-79.2916	25.4858	-199.5617	5.0248	-52.1721
VAR(1:6:12)	5.9665	-67.8905	26.3992	-210.2973	7.8968	-139.1487
BPS(24)	3.5538	-	8.5077	-	3.3021	-

24-step	LPDR$_{1:T}$
VAR(1)	-329.6262
VAR(12)	-661.7998
VAR(3)	-464.2547
VAR(1:3:9)	-905.6553
VAR(1:6:12)	-876.0228
BPS(24)	-

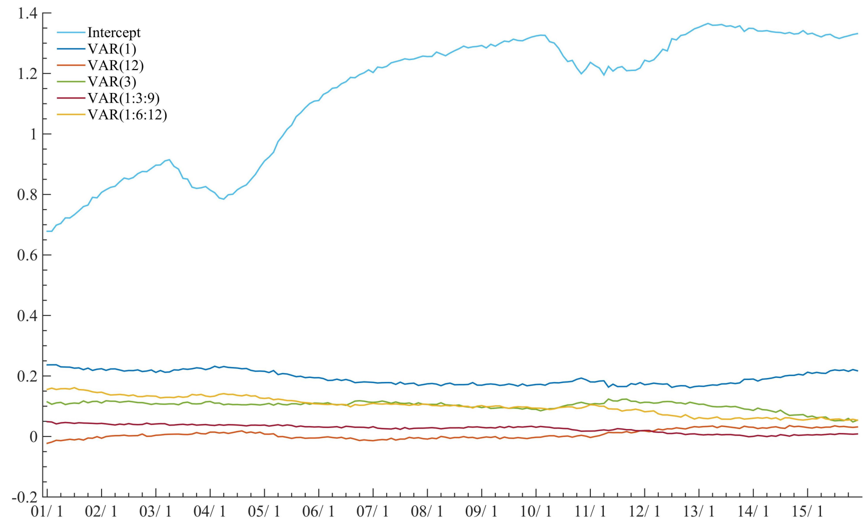

FIGURE 4.22: US macroeconomic forecasting 2001/1-2015/12: 24-step ahead on-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}180$ months for inflation.

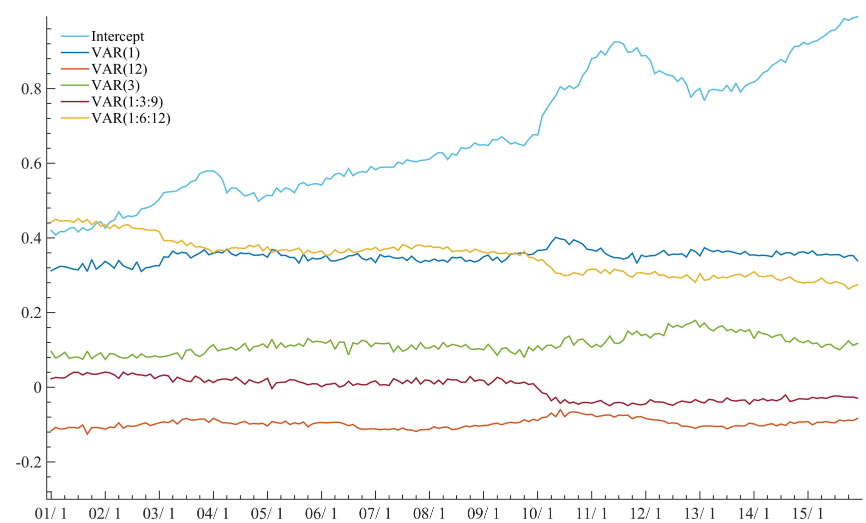

FIGURE 4.23: US macroeconomic forecasting 2001/1-2015/12: 24-step ahead on-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}180$ months for wage.

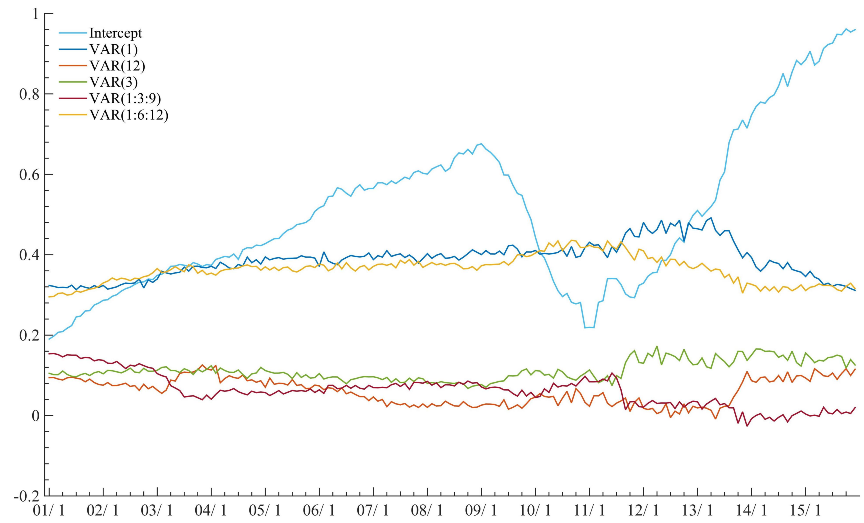

FIGURE 4.24: US macroeconomic forecasting 2001/1-2015/12: 24-step ahead on-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}180$ months for unemployment.

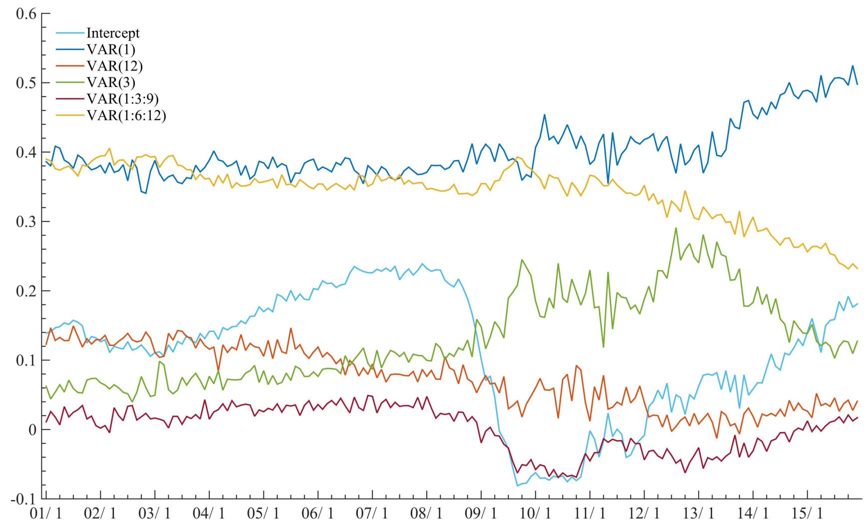

FIGURE 4.25: US macroeconomic forecasting 2001/1-2015/12: 24-step ahead on-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}180$ months for consumption.

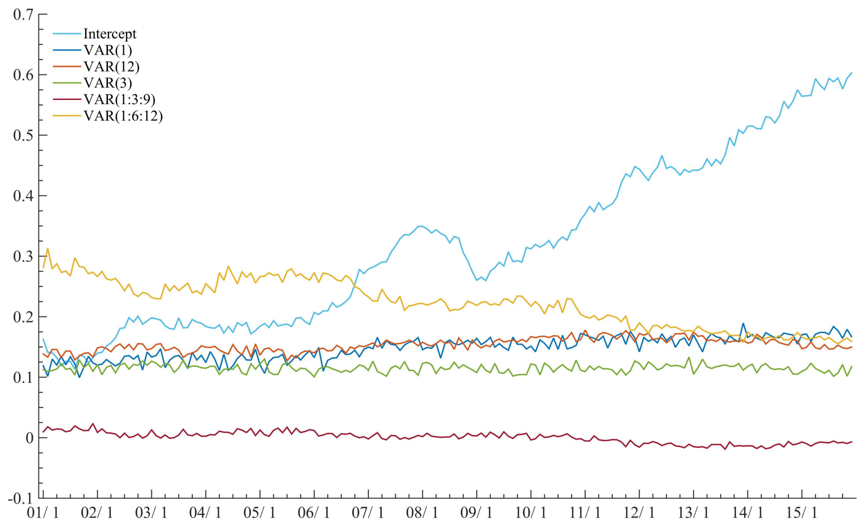

FIGURE 4.26: US macroeconomic forecasting 2001/1-2015/12: 24-step ahead on-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}180$ months for investment.

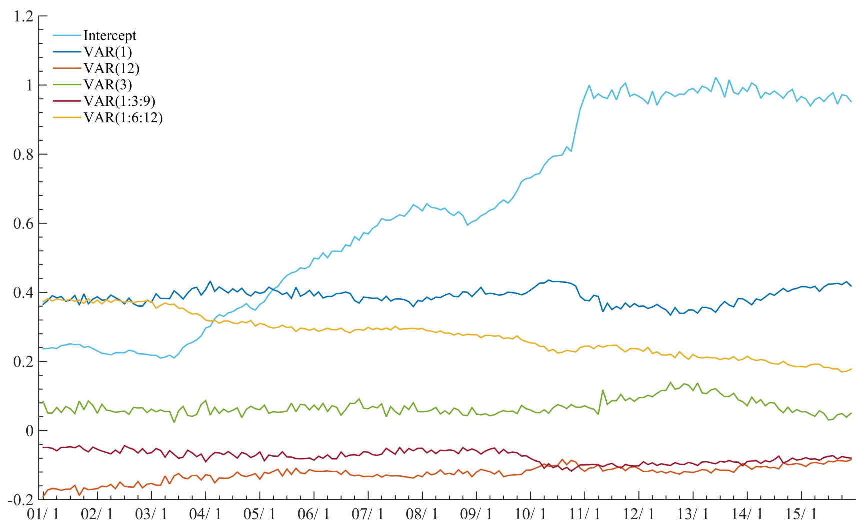

FIGURE 4.27: US macroeconomic forecasting 2001/1-2015/12: 24-step ahead on-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}180$ months for interest rate.

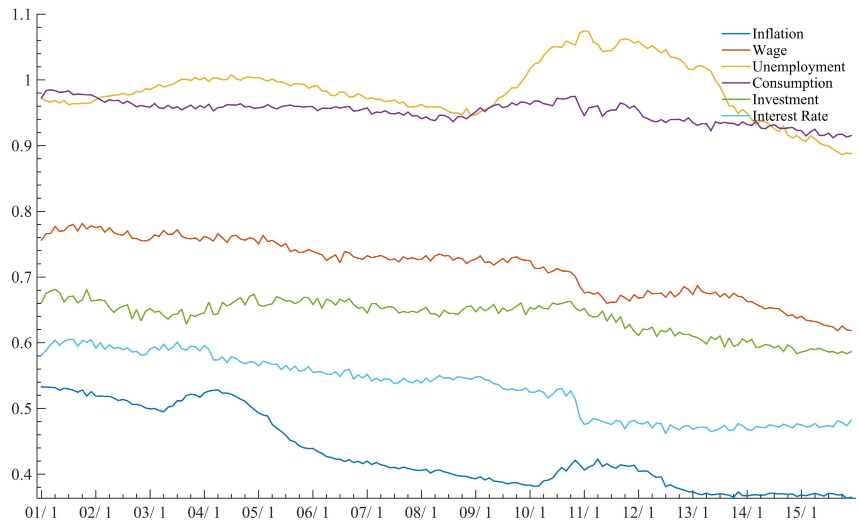

FIGURE 4.28: US macroeconomic forecasting 2001/1-2015/12: 24-step ahead online posterior means of the sum of BPS model coefficients (misspecification) sequentially computed at each of the $t = 1{:}180$ months.

the other series. Additionally, the level of misspecification does not seem to effect the improvements of BPS(k). Further investigation into the connection of BPS(k) misspecification and improvements is warranted.

4.3 Summary

Multivariate forecasting is an increasingly important topic in many fields where decisions concerning uncertainty and dependence between multiple series are as crucial as individual forecast accuracy. The univariate BPS framework is extended to the multivariate setting to synthesize forecasts from multiple agents. A topical multiple US macroeconomic data study demonstrates how multivariate BPS can be effective and practical for macroeconomic policy decisions. Multivariate BPS dynamically synthesizes forecasts to improve forecasts of individual series and cross-series dependence, dominating the agent forecasts over multiple horizons for both point and distribution forecasts. An analysis over the testing period illustrates the robustness of the forecast accuracies of BPS under economic distress, which is critically important for practical applications. Further analysis into the dependence structure of the agents across and within series highlights key insights into the dynamic nature of dependence under different economic situations, providing the decision maker with crucial information into the state of the economy.

5

BPS Applications in Finance

5.1 Equity return predictability

There is a long literature on attempts to predict stock market returns in finance and economics (e.g. Cremers, 2002; Avramov, 2002; Ang and Bekaert, 2007; Lettau and Van Nieuwerburgh, 2008; Cochrane, 2008). However, it is near consensus that no one predictor or set of predictors can outperform simple historical average of returns consistently over time. In their seminal paper, Welch and Goyal (2008) conduct an exhaustive search of predictors to find that predictors considered to be statistically significant in the literature to be unstable in terms of out-of-sample forecasts.

There are several limitations to methods in the literature. First, dynamics are often not considered, as the models in the literature are almost solely static linear regressions. We can expect certain predictors to be "in favor" at certain times, but fall "out of favor" at others. This is particularly true under shocks or structural changes seen in the equity and economic markets in the last couple of decades. Second, research on return predictability often focus on market returns and not on sector indices or individual returns. Predictors having stronger signals for predictability at a finer level (sector, individual, etc.) might be washed out by aggregation. Third, incorporating information is limited to linear combinations of predictors. Under the framework used in the literature, the only way to add information from other sources (e.g. sector information) is by adding predictors to the model. This is limiting in the sense that we quickly run into problems of dimensionality and overfitting. If only a few predictors are expected to be significant at any given time or period of time, most of the information is insignificant or detrimental in terms of out-of-sample forecasting due to overfitting and additional noise.

As an application of BPS to financial data and problems of return predictability, this dissertation considers two distinct problems: model synthesis to improve sector return predictability, and synthesizing projected sector information to improve market return predictability.

The first application approaches the problem of return predictability for sector index returns, pulling information from sector predictors and market predictors. Considering the return predictability of sector index returns, the question is whether predictability comes from sector specific information or is driven by the market as a whole. Compiling predictors for each sector and for the aggregate market, we can build agents that represent different views on drivers of predictability: sector, market, or both. Synthesizing these agents produces a deeper understanding of the predictability of each sector, as well as improving forecast accuracy and economic gains.

The second application approaches the problem of market return predictability utilizing projected sector information. For this the idea of information projection is utilized. Consider an agent analysis framework where each agent represents a sector using specific sector information to produce forecast distributions in regard to market returns. For example, each agent can be a linear regression model (a linear projection of covariates onto the response), with market returns as a response variable and sector specific predictors as covariates, producing a predictive distribution for market returns. This forecast distribution can then be considered as sector specific information filtered in terms of market returns, as they are projections. These agent forecasts (or filtered information) are then synthesized using BPS to learn the dependencies across sectors and biases for each sectors. Additionally, the latent dependencies uncovered by BPS can be considered as a new way of looking at inter-sector dependency, as this dependency is conditional on the market.

5.1.1 Data

We analyze two separate sets of US equity returns data for the two studies. One is the monthly sector returns for each of the 12 sectors on the New York Stock Exchange (Fig. 5.1.), and the second is the monthly market (S&P 500 index) returns (Fig. 5.2.), both from 1985/1-2015/12, a data set of interest in the finance literature (Welch and Goyal, 2008). The 12 sectors are; NoDur: non-durable goods, Durbl: durable goods, Manuf: manufacturing, Enrgy: energy, Chems: chemicals, BusEq: business equipment, Telcm: telecommunications, Utils: utilities, Shops: shops, Hlth: Health, Money: money, and Other: other stocks not included in the above 11. During the period of analysis, the sub-prime mortgage crisis and great recession of the late 2000s warrant special attention. Because this period exhibited such a sharp shock to the US economy, and more so to the equity market, it tests the robustness of the predictive ability of any model and strategy.

Each sector and market index is accompanied by several predictors used by Welch and Goyal (2008), including, **Volatility**: the square root of the sum of daily squared (de-meaned) returns on the value-weighted industry portfolio, **ldp**: the difference between the log of dividends and the log of prices, **ldy**: the difference between the log of dividends and the log of lagged prices, **bmr**: the ratio of book

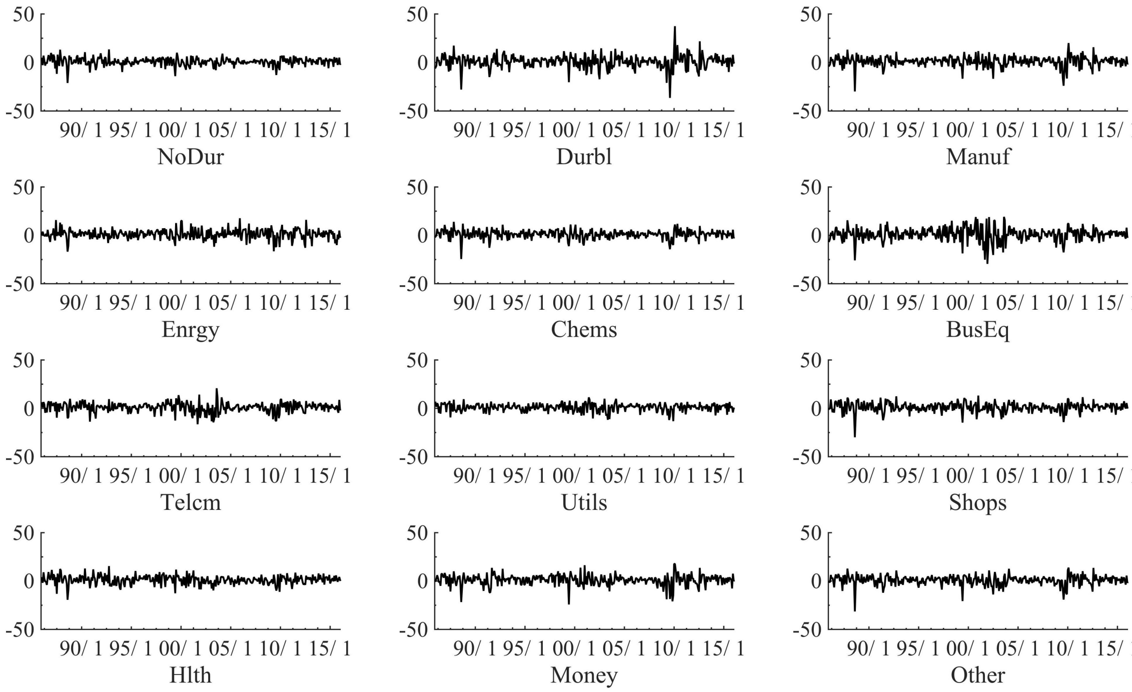

FIGURE 5.1: US sector index forecasting 2000/7-2015/12: US sector index returns (returns $\times 100$ for % basis).

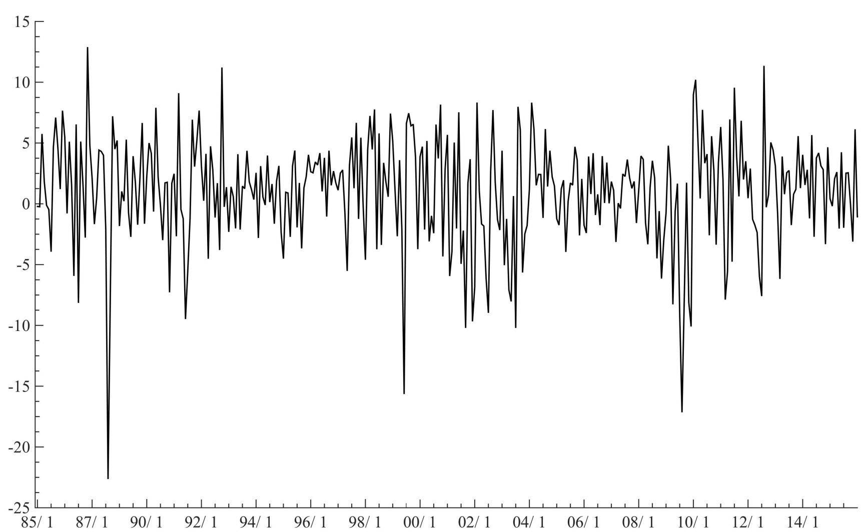

FIGURE 5.2: US market index forecasting 2000/7-2015/12: US market index returns (returns $\times 100$ for % basis).

value to market value for the Dow Jones Industrial Average computed on a quarterly basis and interpolated through a cubic spline to get monthly estimates, and **ntis**: the ratio of 12-month moving sums of net issues divided by the total end-of-year market capitalization. These predictors are considered in the literature to improve the predictability of returns (Welch and Goyal, 2008) and are used to build the agent forecasts used in the BPS analysis.

5.1.2 Agents and BPS specifications

Sector index return predictability

For the analysis for sector index returns, we have $J = 4$ agents. Labeling them M*, the agent models are: M1- historical average; M2- sector specific predictors; M3- aggregate market predictors; M4- sector specific and aggregate market predictors. M1 has a predictive distribution using the historical mean and variance from $1:t$. The other three use the predictors above in a linear regression model, estimated using ordinary least squares (OLS), with an expanding window (sequentially adding data from $t = 1$). This set of models is common in the literature, as the biggest challenge in return predictability is improving on the historical average (M1). M2-M4 represent the views that predictability is driven from sector specific predictors, aggregate market predictors, or a combination of both.

Market index return predictability

For the analysis for market index returns, we have $J = 12$ agents. Each M* is a model that forecasts the market index returns using sector specific predictors for each sector *, generating forecasts using a linear regression model with an expanding window (sequentially adding data from $t = 1$).

The idea here is that each M* is a projection of the sector information onto the quantity of interest: market index returns. By regressing the sector specific predictors onto the market index returns, the agents are essentially filtering the sector information as it pertains to the market index returns. The forecasts coming from each agent, thus, can be seen as filtered/projected sector information in regard to the market returns.

For both studies, the dynamic BPS models take initial priors as $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}, 0.01\boldsymbol{I})$ with $\boldsymbol{m} = (0, \mathbf{1}'/p)'$ and $1/v_0 \sim G(5, 100)$. The discount factor for BPS is based on $(\beta, \delta) = (0.95, 0.99)$.

5.1.3 Data analysis and forecasting

The agent models for both applications are analyzed and synthesized as follows. First, the agent models are analyzed in parallel over 1985/1-1992/9 as a training period, simply running the linear regression model to the end of that period to calibrate the agent forecasts. This continues over 1992/10-2000/6 now accompanied by the MCMC-based BPS analysis running using from 1992/10 data up to time t; that is, we repeat the analysis with an increasing "moving window" of past data as we move forward in time.

This continues over the third period to the end of the series, 2000/7-2015/12; now we also record and compare forecasts as they are sequentially generated. This testing period spans over a decade, and we are able to explore predictive performance over periods of drastically varying economic circumstances, check robustness, and compare benefits and characteristics of each strategy. Out-of-sample forecasting is thus conducted and evaluated in a way that mirrors the realities facing decision makers and portfolio managers.

5.1.4 Forecast accuracy and comparison

As with Section 3.1.4, we compare forecast accuracy using mean squared forecast errors (MSFE) and log predictive density ratios (LPDR).

Additionally, we compare the economic gains from these models. To do this we first construct a two asset portfolio with a risk-free asset (Rf_t) and a risky asset (R_t; sector or market index) for each t. Then, the weights of the optimal portfolio are computed as

$$w_t = (\mathbb{E}(R_t) - Rf_t)/\kappa\sqrt{\mathbb{V}(R_t)},$$

where $\mathbb{E}(R_t)$ and $\mathbb{V}(R_t)$ are the expectation and variance of the forecast of R_t and κ is the risk averseness of $\mathcal{D}$. After obtaining the portfolio gains for each of the agent models and BPS, we compare the results using the Sharpe ratio, which is defined as the average return divided by the standard deviation of returns.

For the market index return application, we do not compare the BPS results with each agent but rather three benchmarks: linear regression model using aggregate market predictors, linear regression model using the aggregate market predictors and all of the sector specific predictors, and the historical average. Linear regression models are static models estimated estimated using OLS in an expanding window.

5.2 US equity sector index returns forecasting

Comparing predictive summaries over the out-of-sample period, we see mixed results in terms of point, distribution, and economic forecasting; see numerical summaries in Table 5.1-5.3. Focusing on the point forecasts, we see that the historical average of returns outperforms other models and BPS except for the health sector, where sector specific predictors improve over the historical average. BPS, while doing fairly well compared to the historical average, does not perform the best in contrast to other studies. In contrast, BPS outperforms everything in terms of distribution forecasts (LPDR) for every sector. This is most likely due to BPS improving uncertainty forecasts through its dynamic stochastic volatility. It is notable

Table 5.1: US sector index forecasting 2000/7-2015/12: Forecast evaluations for monthly US sector indices, comparing mean squared forecast errors.

$\text{RMSE}_{1:T}$	NoDur	Durbl	Manuf	Enrgy	Chems	BusEq
Historical ave.	3.3811	7.6215	5.8097	5.7668	4.0285	7.3199
Sector spec.	3.3833	7.7307	5.8220	5.8343	3.9981	7.6625
Aggregate	3.5680	7.9105	5.9990	5.8525	4.2524	7.5077
Sector + Aggregate	3.5648	8.0263	6.0494	5.9369	4.2247	7.8721
BPS	3.3884	7.7267	5.8673	5.7957	4.0733	7.4918

$\text{RMSE}_{1:T}$	Telcm	Utils	Shops	Hlth	Money	Other
Historical ave.	5.4410	4.1746	4.2240	3.9284	5.6364	5.0819
Sector spec.	5.3413	4.3724	4.2581	3.8982	5.7901	5.1127
Aggregate	5.5965	4.1963	4.4141	4.1080	5.7970	5.2467
Sector + Aggregate	5.4766	4.4452	4.4450	4.1190	5.9035	5.2343
BPS	5.4195	4.2181	4.2680	3.9106	5.7047	5.1127

that there are larger improvements for some sectors compared to others. In particular, the greatest improvements are in BusEq and Durbl. Looking at the economic gains, in terms of the Sharpe ratio, BPS outperforms the rest for all but three out of the 13 sectors examined. This echoes the results of LPDR, with BPS improving on economic gains due to its superior uncertainty estimates.

To explore further, three sectors of interest are examined: Durables, Health, and Money. These three sectors have distinct characteristics that form a wider snapshot of the economy as a whole. Durables represent the consumer side of the economy, more volatile and susceptible to shocks that effect consumer spendings, etc. Health represents the more stable side of the economy, not greatly effected by market shocks. Money represents the credit side of the economy, most effected by market shocks, especially the sub-prime mortgage crisis. With these three sectors, we are able to get a broader sense of how BPS works compared to the other models.

Figs. 5.3-5.5 show the 1-step ahead forecast standard deviations sequentially computed throughout the testing period. Overall, BPS has a significantly different trajectory than that of the other models. The standard deviation of BPS is lower be-

Table 5.2: US sector index forecasting 2000/7-2015/12: Forecast evaluations for monthly US sector indices, comparing log predictive density ratios for this $T = 186$ months. Note: $\text{LPDR}_{1:T}$ is relative to BPS.

$\text{LPDR}_{1:T}$	NoDur	Durbl	Manuf
Historical ave.	-7.9184	-12.2366	-6.1869 0
Sector spec.	-7.3447	-14.5937	-4.7631
Aggregate	-12.2502	-16.3419	-8.3995
Sector + Aggregate	-11.7935	-16.2838	-8.5360
BPS	-	-	-

$\text{LPDR}_{1:T}$	Enrgy	Chems	BusEq
Historical ave.	-1.0418	-3.8302	-20.8570
Sector spec.	-3.2243	-2.6033	-26.3108
Aggregate	-2.9698	-8.5865	-24.3988
Sector + Aggregate	-4.9547	-7.7476	-29.7572
BPS	-	-	-

$\text{LPDR}_{1:T}$	Telcm	Utils	Shops
Historical ave.	-11.4318	-1.0405	-4.7843
Sector spec.	-5.9891	-11.3608	-5.7258
Aggregate	-15.1233	-1.4260	-8.9174
Sector + Aggregate	-8.2104	-13.2259	-10.1732
BPS	-	-	-

$\text{LPDR}_{1:T}$	Hlth	Money	Other
Historical ave.	-6.4069	-8.4800	-4.4928
Sector spec.	-5.7470	-9.6983	-3.8411
Aggregate	-11.0820	-8.8511	-6.7959
Sector + Aggregate	-11.0579	-10.8348	-5.2145
BPS	-	-	-

fore the subprime crisis but spikes during the crisis, quickly adapting to the shock. For Durables and Money, the standard deviation of BPS switches between being larger and smaller than the models compared at different economic periods. For Health, note that the standard deviation of BPS is almost constantly lower than the other models. As BPS dominates the other models in terms of LPDR, the results demonstrate that these models, ubiquitous in the literature, constantly under- or over-estimate uncertainty.

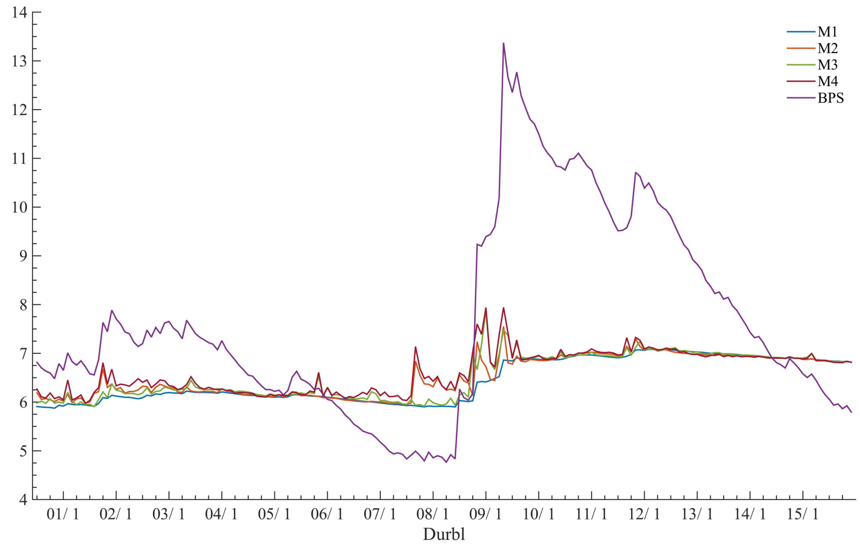

FIGURE 5.3: US sector index forecasting 2000/7-2015/12: 1-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}186$ months for Durables.

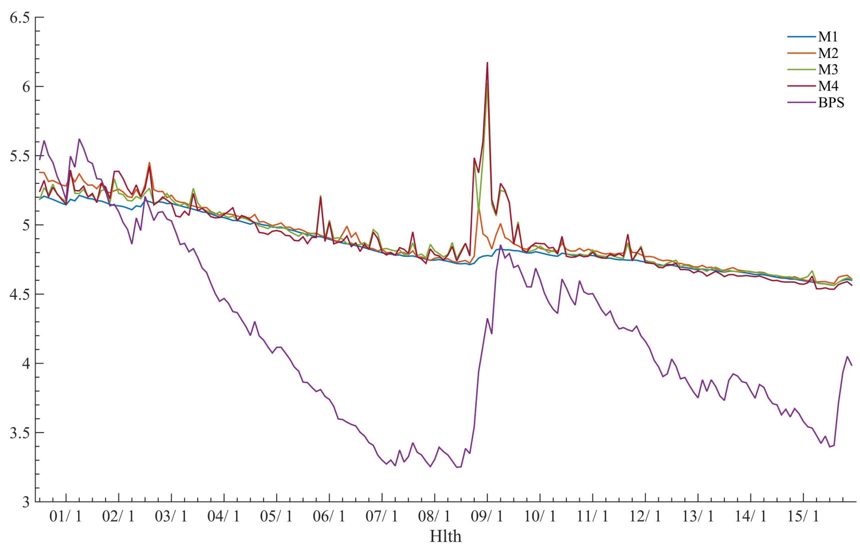

FIGURE 5.4: US sector index forecasting 2000/7-2015/12: 1-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}186$ months for Health.

Table 5.3: US sector index forecasting 2000/7-2015/12: Forecast evaluations for monthly US sector indices, comparing Sharpe ratios for this $T = 186$ months.

Sharpe Ratio$_{1:T}$	NoDur	Durbl	Manuf	Enrgy	Chems	BusEq
Historical ave.	0.9032	0.8259	0.7840	0.7256	0.8297	0.8442
Sector spec.	0.8749	0.6954	0.8256	0.7371	0.9124	0.6958
Aggregate	0.7164	0.3707	0.5354	0.6896	0.6736	0.7457
Sector + Aggregate	0.7447	0.4885	0.5776	0.6527	0.6935	0.6006
BPS	0.8753	0.8332	0.8533	0.8348	0.8660	0.8448

Sharpe Ratio$_{1:T}$	Telcm	Utils	Shops	Hlth	Money	Other
Historical ave.	0.8079	0.7754	0.8898	0.8591	0.7813	0.8391
Sector spec.	0.8417	0.5492	0.8374	0.8654	0.5021	0.7744
Aggregate	0.6615	0.7425	0.7724	0.7929	0.5386	0.5871
Sector + Aggregate	0.6419	0.5242	0.7437	0.7944	0.4874	0.5465
BPS	0.8543	0.8484	0.8790	0.8761	0.8470	0.8526

Moving on to the on-line posterior means of BPS model coefficients (Figs. 5.6-5.8), there is a stark contrast across sectors. First, historical average is always the model with the highest coefficients across sectors. This is expected, as the historical average almost always outperforms models using predictors over time. Second, while Health has stable coefficients throughout, Durables and Money see large shocks in the coefficients around and after the sub-prime mortgage crisis. In particular, Durables has the most difference in coefficients over time. There is a drop in the model with aggregate predictors and sector and aggregate predictors. One interpretation is that during and after the crisis, the consumer economy was hard hit, leading to the Durables sector being effected more than that of the aggregate market as a whole. Models including aggregate information, which no longer provide useful information for forecasting, are then downgraded or discounted. There is, however, an increase of model misspecification as a whole, implied by the increase in the intercept. Health stays stable over the whole period, suggesting that the market shocks had little effect in terms of changes in information provided by the models. For Money, we see all models, except historical average, falling

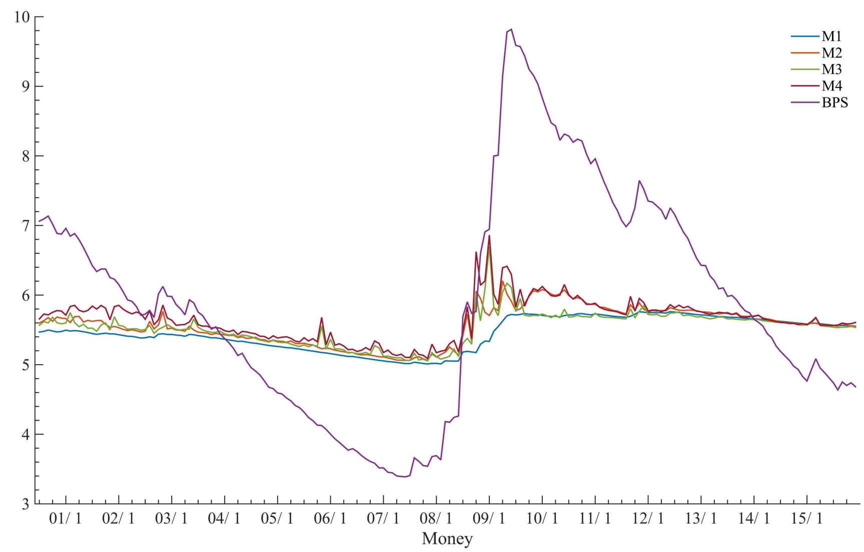

FIGURE 5.5: US sector index forecasting 2000/7-2015/12: 1-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}186$ months for Money.

out of favor after the sub-prime mortgage crisis. This suggests that the shock to the Money sector was not a shock to the fundamentals (realized in the predictors), since the coefficients imply that neither the sector predictors nor the market predictors provide sufficient information.

Finally, we move our focus to the model misspecification (Figs. 5.9-5.11; the sum of on-line posterior means of the BPS model coefficients). Consistent with the results above, we see that Health is relatively stable throughout, while Durables and Money drop around to 0.5. As this result represents the overall level of misspecification of the models considered, the results imply that both sector specific predictors and market predictors offer little information to be synthesized during and after the sub-prime mortgage crisis.

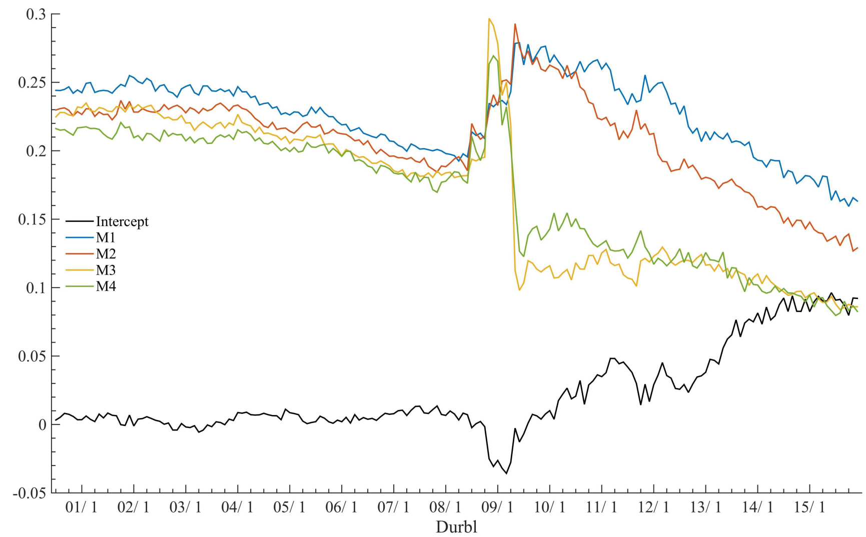

FIGURE 5.6: US sector index forecasting 2000/7-2015/12: On-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}186$ months for Durables.

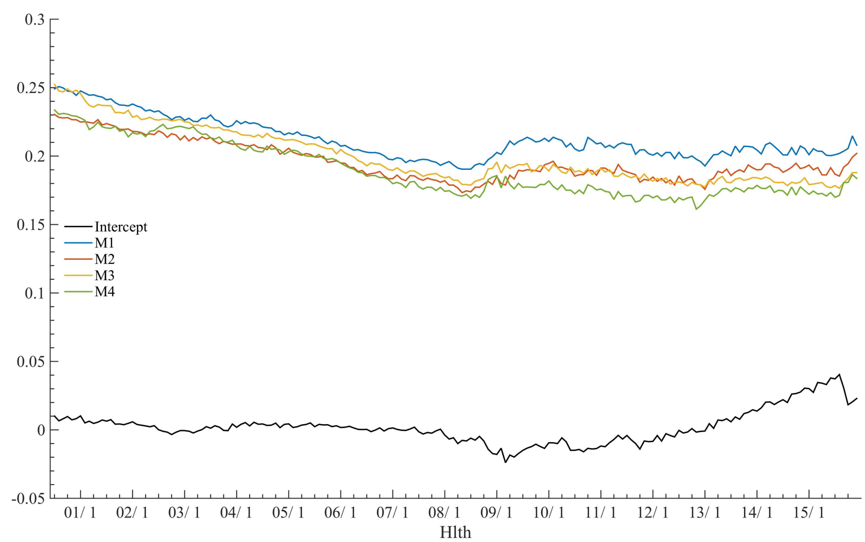

FIGURE 5.7: US sector index forecasting 2000/7-2015/12: On-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}186$ months for Health.

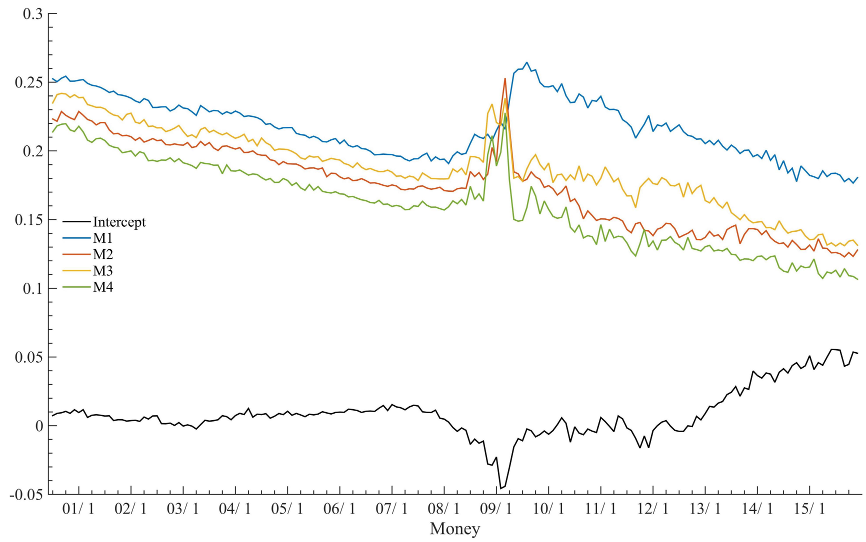

FIGURE 5.8: US sector index forecasting 2000/7-2015/12: On-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}186$ months for Money.

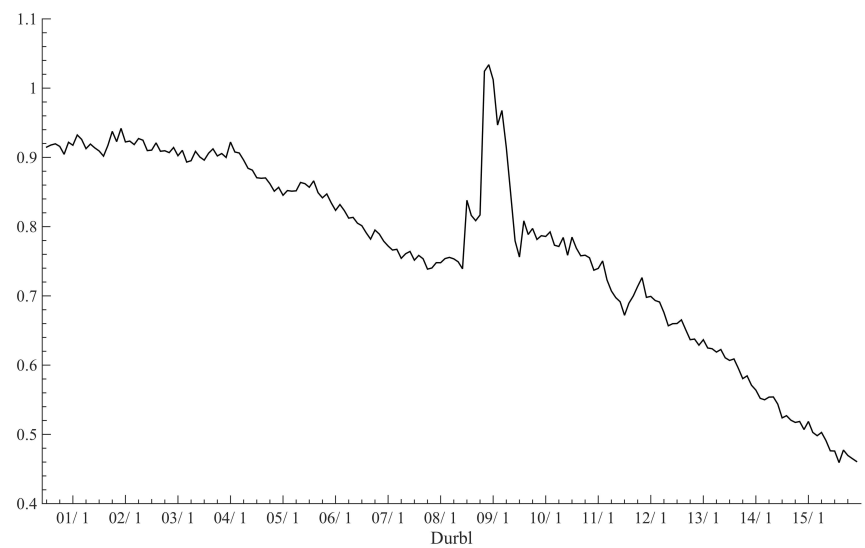

FIGURE 5.9: US sector index forecasting 2000/7-2015/12: On-line posterior means of the sum of BPS model coefficients (misspecification) sequentially computed at each of the $t = 1{:}186$ months for Durables.

FIGURE 5.10: US sector index forecasting 2000/7-2015/12: On-line posterior means of the sum of BPS model coefficients (misspecification) sequentially computed at each of the $t = 1{:}186$ months for Health.

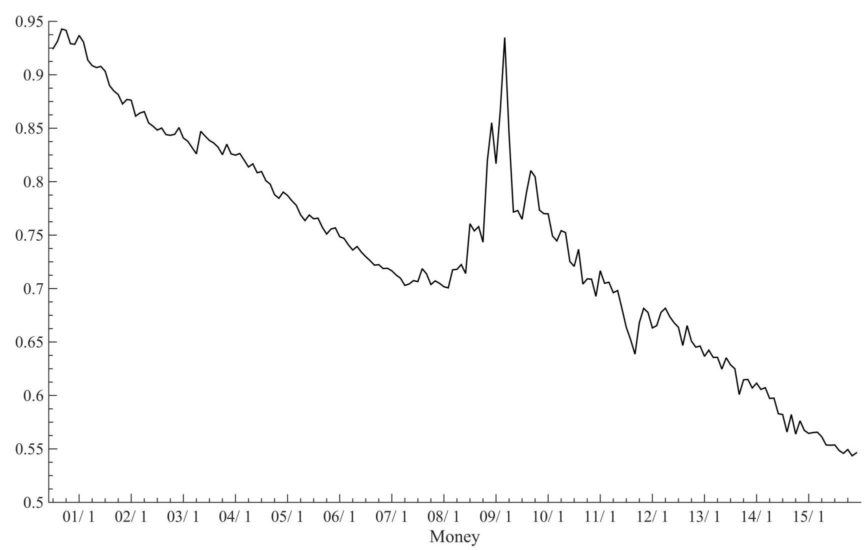

FIGURE 5.11: US sector index forecasting 2000/7-2015/12: On-line posterior means of the sum of BPS model coefficients (misspecification) sequentially computed at each of the $t = 1{:}186$ months for Money.

5.3 US equity market index returns forecasting

For the analysis of US equity market index returns, we compare BPS– synthesizing projected sector information– against conventional linear regression models with market aggregate predictors, aggregate and sector predictors, and historical average. Compared to the mixed results in Section 5.2, BPS exhibits superior performance on all three measures (summarized in Table 5.4), comparing predictive summaries over the out-of-sample period. Noteworthy is BPS improving over the historical average in terms of point forecast, which BPS did not for the results in Section 5.2. We see that OLS (Full), the linear regression with all sector information, performs significantly worse than all models and methods considered. Compare this to BPS, which uses the same set of information as OLS (Full), we see how BPS improves by dynamically learning the dependencies and biases across and within sectors and dynamically adapt to different economic climates.

Figs. 5.12-5.16 summarize sequential analysis for 1-step ahead forecasting. Fig. 5.12 shows the 1-step ahead $\text{MSFE}_{1:t}$ for each time t. Even though the margins are small, BPS almost uniformly dominates, except at the beginning of the time period where the MSFE is somewhat unstable. During the sub-prime mortgage crisis, OLS (Full) significantly increases in forecast error. The biggest reason for this is overfitting, with $5 \times 14 = 70$ predictors used for OLS (Full), it significantly lacks

Table 5.4: US market index forecasting 2000/7-2015/12: Forecast evaluations for monthly US market index, comparing mean squared forecast errors, log predictive density ratios, and Sharpe ratios for this $T = 186$ months. Note: $\text{LPDR}_{1:T}$ is relative to BPS.

	$\text{MSFE}_{1:T}$	$\text{LPDR}_{1:T}$	Sharpe Ratio
OLS (Aggregate)	22.1276	-11.4728	0.5907
OLS (Full)	36.2512	-49.1352	0.3445
Historical Ave.	20.5635	-7.4650	0.8037
BPS	20.1323	-	0.8759

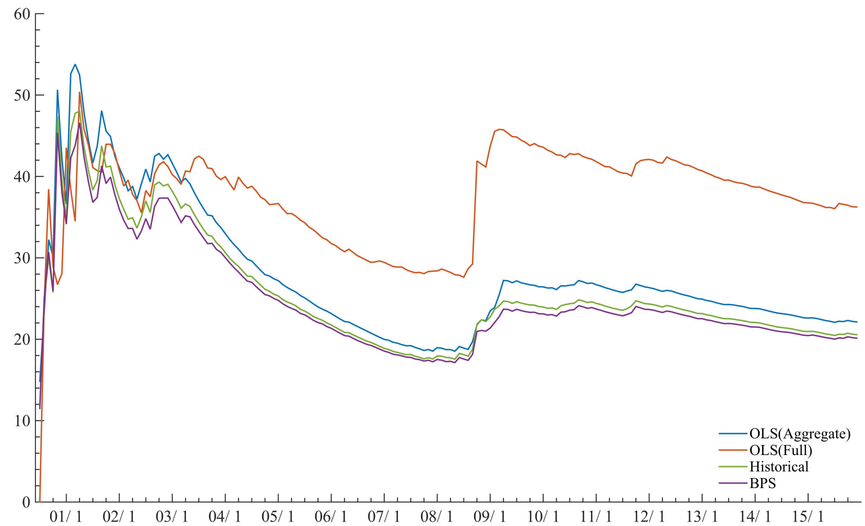

FIGURE 5.12: US market index forecasting 2000/7-2015/12: Mean squared 1-step ahead forecast errors $\text{MSFE}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}186$ months.

adaptability over key structural changes and shocks in the economy. BPS, on the other hand, remains stable over time improving over the rest in a robust fashion due to its dynamic nature.

Fig. 5.13 confirms that BPS performs uniformly better than, or on par with, the other models based on LPDR measures for distribution forecasts. Major shocks and times of increased volatility have substantial impact on the relative performance, again most notable at the beginning of the sub-prime mortgage crisis. BPS is able to adapt to maintain improved forecasting performance both in terms of location and risk characterization, a key positive feature for portfolio managers who are dependent on accurate risk assessment.

The LPDR results are indicative of improved uncertainty assessments by BPS. Looking at the 1-step ahead forecast standard deviations (Fig. 5.14), BPS dynamically adapts over the other methods by shrinking its forecast standard deviation during periods of stability (pre-sub-prime mortgage crisis) and increasing it during periods of shocks and instabilities (post-sub-prime mortgage crisis). Notably, the

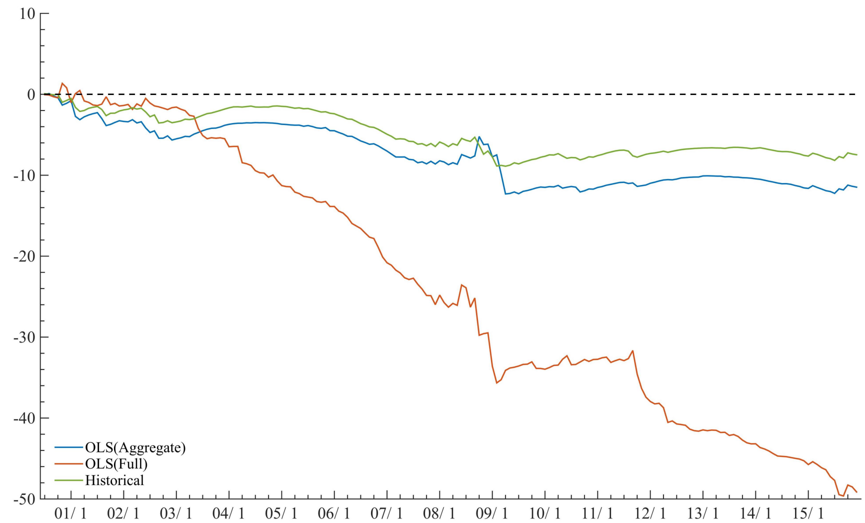

FIGURE 5.13: US inflation rate forecasting 1990/Q1-2014/Q4: 1-step ahead log predictive density ratios $\text{LPDR}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}186$ months. The baseline at 0 over all t corresponds to the standard BPS model.

standard deviation of OLS (Full) has multiple spikes in the forecast standard deviation, even during periods of stability. This, again, is most likely due to OLS (Full) overfitting and as a result fail to assess the correct level of risk.

Fig. 5.15 exhibits the on-line posterior means of BPS model coefficients sequentially computed for each t. Each coefficient represents the information provided by each sector. Thus, when a certain coefficient goes up (down), it can be inferred that the information from that sector is providing more (less) information to the forecast of the market. For example, while coefficients are relatively stable before the sub-prime mortgage crisis, the coefficients change drastically after the crisis. In particular, there is a heavier reliance in information from more stable sectors: Utilities, Energy, and Manufacturing, while relying less on volatile sectors: Money and Durables. This can be seen as a shift in the market, where the financial and consumer side of the market representing proportionally less of the economy.

Finally, the level of misspecification (Fig. 5.16) is stable compared to the results from Section 5.2. There are several clear trends, also seen in the other study,

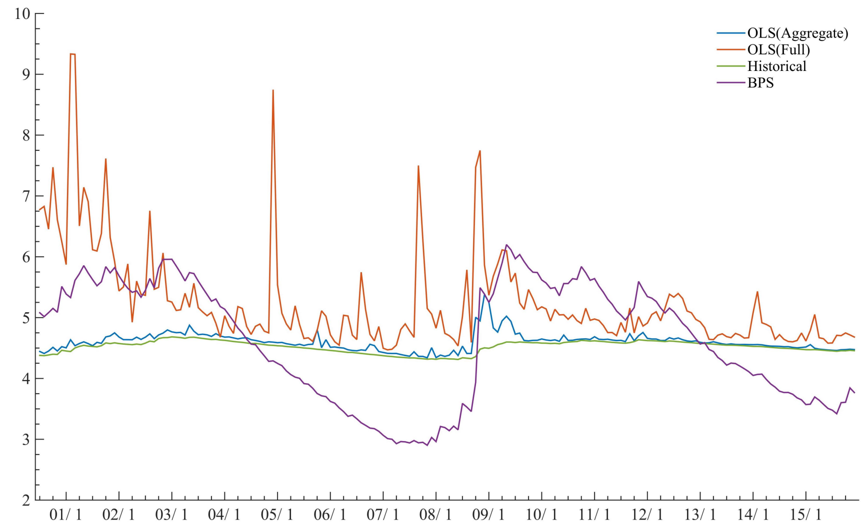

FIGURE 5.14: US market index forecasting 2000/7-2015/12: 1-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}186$ months.

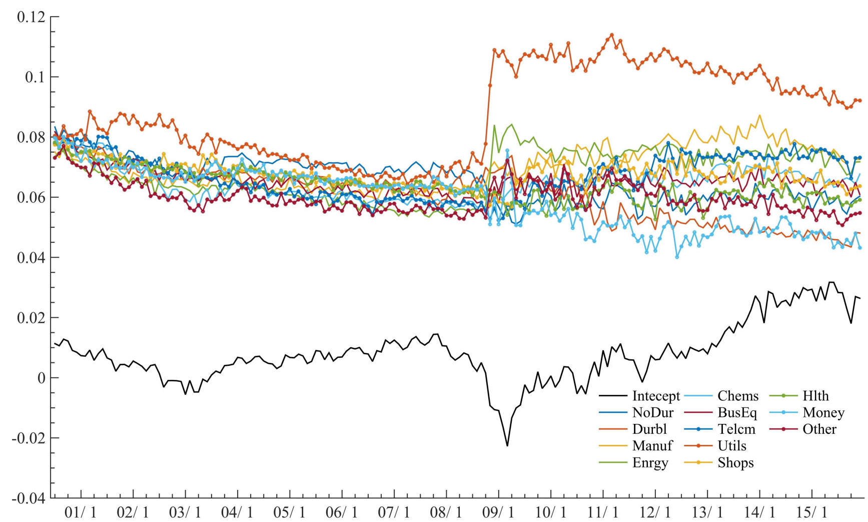

FIGURE 5.15: US market index forecasting 2000/7-2015/12: On-line posterior means of BPS model coefficients sequentially computed at each of the $t = 1{:}186$ months.

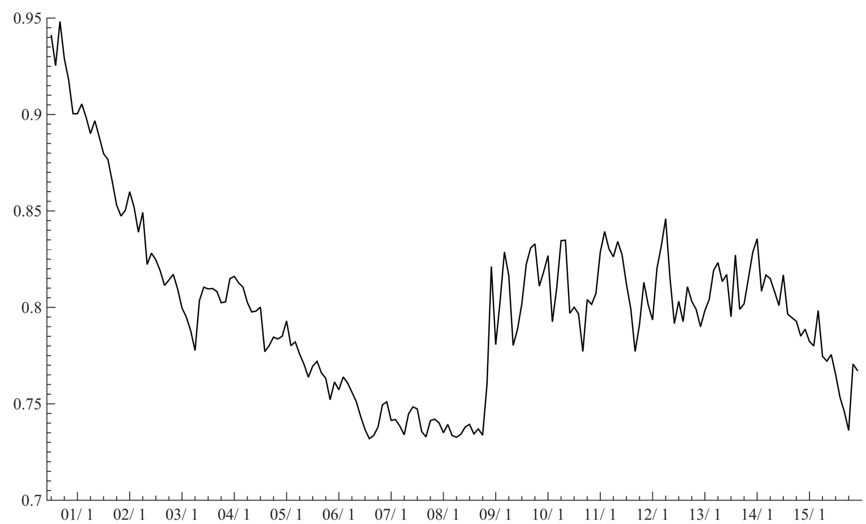

FIGURE 5.16: US market index forecasting 2000/7-2015/12: On-line posterior means of the sum of BPS model coefficients (misspecification) sequentially computed at each of the $t = 1{:}186$ months.

where we see a gradual decrease in misspecification, a shock during the sub-prime mortgage crisis, and volatile periods after the shock. From this, inferences can be made on how sector information fully forecasts market returns. Notably, the volatile period after the crisis is indicative on the uncertainty about the model misspecification as a whole, where signals from sector information are more volatile and less useful compared to pre-crisis.

5.4 Summary

The two finance applications illustrate the effectiveness of BPS under different data and settings, while exploring interesting ways BPS can be used to improve forecasts and economic gains. The first study applies BPS to predictability of sector index returns. By dynamically synthesizing forecasts coming from agents with different sources of information, BPS improves in terms of distribution forecasts and economic gains for most sectors. Further analysis into the BPS estimates and parameters provides insight into characteristics of the returns for each sector.

The second study furthers the application of BPS by defining agents as a projection of sector information onto the market returns, in order to improve return predictability of the market. By creating agents who represent sector information, via filtered projection, BPS is able to synthesize the information from multiple facets of the economy to improve the overall forecast of the market. The BPS synthesized forecasts of market returns dominates other models in terms of point, distribution forecasts, and economic gains. An analysis of on-line BPS coefficients draws interesting parallels and insight into the economic regime changes over the last decade.

6

BPS Extensions for Temporal Misspecification

6.1 Context, literature, and methodology

The literature on model misspecification has been almost solely limited to the parameter space (George and McCulloch, 1993, 1997; George et al., 2008; West, 2003; Clyde and George, 2004; Carvalho et al., 2007, 2008; Yoshida and West, 2010; Wang, 2010; Chen et al., 2011; Korobilis, 2013). However, especially in time series analysis, temporal misspecification plays a significant role as well, though it is almost never discussed or explored. The concept of temporal misspecification expands the discussion of $\mathcal{D}$'s specification of a model to the temporal domain, such as the length of data $\mathcal{D}$ wants to use and which sampling frequency is best suited for the decision problem. There are inherent tradeoffs within the temporal specifications $\mathcal{D}$ can make. This chapter discusses these tradeoffs and demonstrates how the framework of BPS can mitigate these problems and improve in terms of the decision goals.

One problem concerns optimal learning lengths for analyzing complex data with possible structure breaks and shocks. Statistical theory stipulates that a longer learning length is better, as the model accrues more information. However, this is only valid when the model is correctly specified and assuming a true data generating process that is static over time. Neither of these assumptions are considered true under the views of subjective Bayesian thinking. As it is often seen in time series data– economic time series, for example– there are multiple structural changes and shocks that make it impossible to assume one model, with one set of covariates, to hold true over time. This brings up an interesting and challenging tradeoff for $\mathcal{D}$: more data are needed to get better estimates, but shocks and structural changes can "throw off" these estimates or make them obsolete. We call this the learning-adaptability tradeoff, as $\mathcal{D}$ has to sacrifice learning (data) to be adaptable and adaptability for more learning (data). In the finance and economics literature, we see this tradeoff decision being made without clear discussion or acknowledgement. For example, the Dow Jones Industrial Average, one of the standard indices used in financial research, has historical data going back to 1896. However, we see very little research being done with learning periods going back beyond 1960. In more egregious cases, we see researchers using data before or after a shock in order to get stable estimates. Conscious or unconscious, this tradeoff is made constantly by researchers and practitioners.

The second key problem concerns the specification of data frequency. For example, a portfolio manager or policy maker might be tasked to forecast an index a month ahead. For many indices, especially in finance, data are sampled at different frequencies, from annual to tick which is in milliseconds. Again, statistical theory will stipulate that, if the model is correctly specified, more data is better. However, using millisecond tick data to forecast a month ahead would be pointless, as tick data are contaminated with microstructure noise, other dynamics, etc. and include very little information that would be helpful to $\mathcal{D}$'s decision problem. Additionally, the higher the frequency, the longer the forecast horizon will be. Using monthly data to forecast a month ahead would be 1-step ahead, while using daily data to forecast a month ahead would be around 22-step ahead, depending on the length of the month. As models are inherently designed and fit to forecast 1-step ahead quantities, increasing the forecast horizon necessarily makes the forecasts more uncertain. Herein lies the temporal tradeoff: the finer the frequency the more accurate it is in terms of the mean, but less accurate in terms of risk assessment. We call this the accuracy-uncertainty tradeoff.

Both of these problems of temporal misspecification will be addressed in this chapter using topical problems in finance, demonstrating the effectiveness of the BPS based methods proposed.

6.2 Temporal misspecification in rolling window analysis

The learning-adaptability tradeoff is addressed and mitigated by applying the framework of BPS on rolling window size uncertainty. Rolling window methods are common in finance and economics, where static models are still standard. The question of optimal rolling window size is rarely discussed except in some cases (Clark and McCracken, 2009; Rossi and Inoue, 2012; Inoue et al., 2017). Extending the BPS framework to the temporal domain, we can construct agents with different rolling window sizes, or temporal information. For this new method, Bayesian inter-temporal synthesis (BITS), each agent will have more past information than the previous agent (longer windows). The additional (past) information, however, is not guaranteed to be useful information for forecasting, as it may contain shocks, outliers, and structural breaks that skew agent forecasts. Synthesizing these agents under the framework of BPS will dynamically infer biases and dependencies of this added information and adapt accordingly, discounting past data.

6.2.1 Data

For this study, we analyze US equity market returns data from 1985/1-2015/12 as with the application in Chapter 5. This period is of special interest, as the sub-prime mortgage crisis and great recession of the late 2000s occur and cause great temporal uncertainty before, during, and after the crisis.

Accompanied by the market returns are multiple predictors used in Welch and Goyal (2008). These are, **Volatility**: the square root of the sum of daily squared (de-meaned) returns on the value-weighted industry portfolio, **ldp**: the difference between the log of dividends and the log of prices, **ldy**: the difference between the log of dividends and the log of lagged prices, **bmr**: the ratio of book value to market value for the Dow Jones Industrial Average computed on a quarterly basis and interpolated through a cubic spline to get monthly estimates, **ntis**: the ratio of 12-month moving sums of net issues divided by the total end-of-year market capitalization, **eqis**: the ratio of equity issuing activity as a fraction of total issuing activity, **epr**: the difference between the log of earnings and the log of prices, **dpr**: the difference between the log of dividends and the log of earnings, **tbl**: 3- month treasury bill, **lty**: long-term government bond yield, **ltr**: long term rate of returns, **AAA**: long term bond returns for AAA rated bonds, **BAA**: long term bond returns for BAA rated bonds, and **infl**: US inflation.

6.2.2 Agents and BPS specifications

For the analysis of market index returns, we have $J = 4$ agents. Labeling them M*, the agent models vary in rolling window size only using the all of the predictors above: M1- window size=18; M2- window size=24; M3- window size=36; M4- window size=60. All models are linear regressions estimated using OLS. The window sizes are selected to cover a broad range of temporal spaces.

In the dynamic BPS models, we take initial priors as $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}, 0.01\boldsymbol{I})$ with $\boldsymbol{m} = (0, \mathbf{1}'/p)'$ and $1/v_0 \sim G(5, 100)$. The discount factor for BPS is based on $(\beta, \delta) = (0.99, 0.95)$.

6.2.3 *Data analysis and forecasting*

The agent models for both applications are analyzed and synthesized following Section 5.1.3.

6.2.4 *Forecast accuracy and comparison*

As with Section 3.1.4, we compare forecast accuracy using mean squared forecast errors (MSFE) and log predictive density ratios (LPDR). The same model using an expanding window (using $1\!: t$) is also compared.

6.2.5 *Bayesian inter-temporal synthesis*

Table 6.1 summarizes the out-of-sample results of the agents and BITS. Looking at the results from the agents with varying temporal specifications, it is evident that longer windows provide better results in terms of point and distribution forecasts. From this we might conclude that longer windows, and perhaps an expanding window using the whole set of data, is better. However, this overview does not give a full picture of the problem and will be explored later in this section in the sequential analysis. Overall, we see that BITS is able to improve over the other agents significantly in both point and distribution forecasts. Comparing the results with an expanding window, BITS outperforms, even though the expanding window analysis performs better than the agents included in the BITS analysis. The results indicate that BITS benefits from synthesizing varying information (inclusion of data) and adapts dynamically to the data to improve accuracy. More crucially, since the temporal space is closed (within the framework used), this demonstrates the ability of BITS to achieve superior accuracy beyond the whole temporal space.

Figs. 6.1-6.4 summarize sequential analysis for 1-step forecasting. Fig. 6.1 shows the 1-step ahead measures MSFE$_{1:t}$ for each time t. BITS uniformly outperforms all window sizes throughout the whole test period. Also notable is how robust the MSFE is for BITS compared to the agents. During crisis periods, particularly the sub-prime mortgage crisis, MSFE jumps for all window sizes, but more so for longer window sizes. This is due to the lack of adaptability in the models with longer window size, even if the overall MSFE is lower than those of shorter window sizes. BITS, on the other hand, stays relatively level during and after the shock, with the majority of relative gains made during these periods.

The distribution forecast results in Fig. 6.2 mirrors the result from the point forecasts. Longer window sizes suffer the most during crisis periods, while shorter window sizes benefit from adaptability, but perform worse in more stable times. The learning-adaptability tradeoff is more prevalent for distribution forecasts, as the measurement of uncertainty is more affected by (over- or under-) learning. As with the point forecasts, BITS is superior compared to all window sizes available, demonstrating the improved uncertainty forecasts and mitigating, or eliminating, temporal uncertainty, at least under the conditions considered.

Table 6.1: US market index forecasting 2007/7-2015/12: Forecast evaluations for monthly US market index, comparing mean squared forecast errors and log predictive density ratios for this $T = 186$ months. Note: LPDR$_{1:T}$ is relative to BITS and Exp. window is expanding window using data from $1:t$.

	MSFE$_{1:T}$	LPDR$_{1:T}$
Window=18	363.6853	-217.9018
Window=24	136.9210	-141.5382
Window=36	90.9580	-90.6934
Window=60	51.5363	-65.9203
Exp. window	24.8914	-22.0135
BITS	21.0956	-

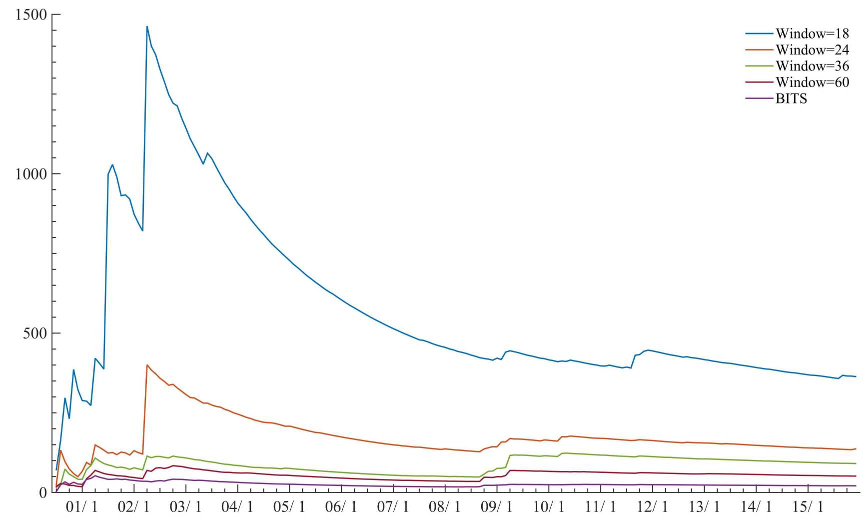

FIGURE 6.1: US market index forecasting 2007/7-2015/12: Mean squared 1-step ahead forecast errors $\mathrm{MSFE}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}186$ months.

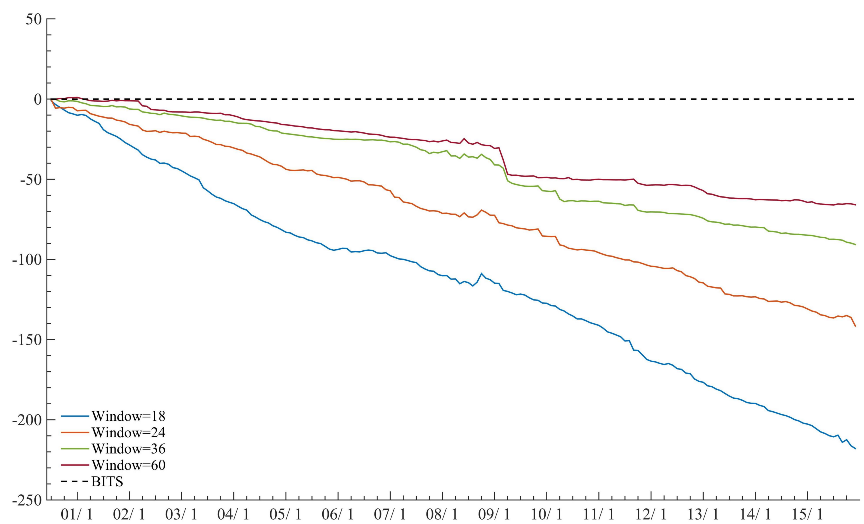

FIGURE 6.2: US market index forecasting 2007/7-2015/12: 1-step ahead log predictive density ratios $\mathrm{LPDR}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}186$ months. The baseline at 0 over all t corresponds to the standard BPS model.

In terms of forecast standard deviations, longer window sizes tend to have smaller uncertainty measures, with large spikes for the shortest window size. The uncertainty measure is tighter for models with more data, as they are able to learn more. This can backfire during structural changes or shocks, as demonstrated by a bump in LPDR in Fig.6.2. On the other hand, shorter windows have spikes that are unreasonable (120+ at around 2002) because they adapt too much. BITS, compared to other results, is relatively stable throughout the estimation period.

The on-line posterior means of BITS coefficients confirm how BITS is able to adapt over time. At the beginning of the analysis, the order of coefficients is in line with the forecast ability of each window size (larger to smaller window size). A gradual decrease in all coefficients up until the sub-prime mortgage crisis can be seen, which is indicative of how these models are misspecified as a whole. In other words, none of the models provide sufficient information. At the sub-prime mortgage crisis, we see a large shock in coefficients, as models with longer window sizes fall off to zero, while shorter window size models rise. This reflects the

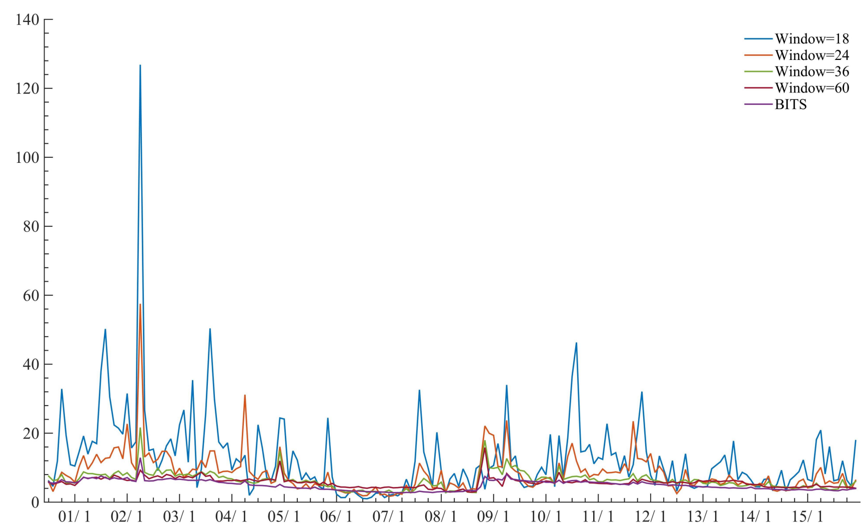

FIGURE 6.3: US market index forecasting 2007/7-2015/12: 1-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}186$ months.

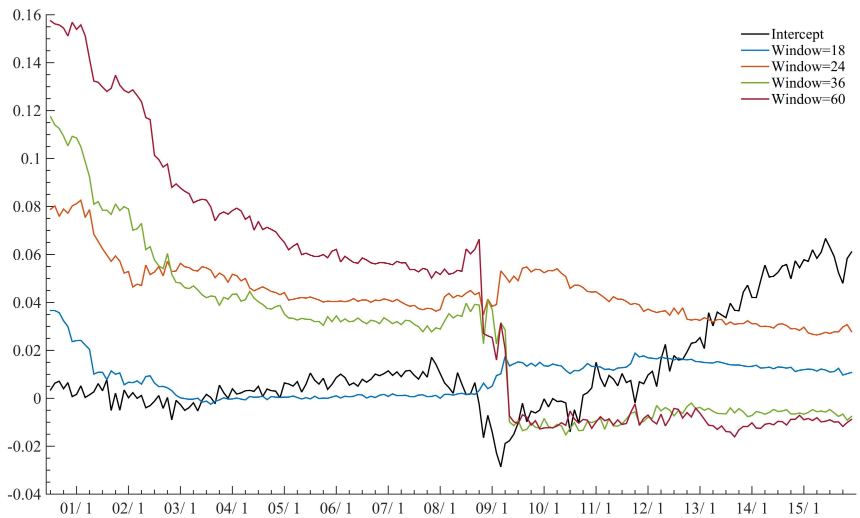

FIGURE 6.4: US market index forecasting 2007/7-2015/12: On-line posterior means of BITS model coefficients sequentially computed at each of the $t = 1{:}186$ months.

change in the market after the crisis: models learnt on pre-crisis data are deemed obsolete, while more adaptive models quickly adjust to the new regime. By shifting the coefficients dynamically, BITS weighs the tradeoff between learning and adaptability, achieving improved forecast accuracy over a long period of time with robust results.

6.3 Frequency selection in long term forecasting

The problem of frequency selection is tackled by applying the framework of BPS with agents varying in data frequency used. For example, if $\mathcal{D}$ is tasked to forecast a month ahead, data will be spliced into multiple frequencies, including monthly, weekly, daily, and everything in between. Each agent will then produce k-step ahead forecasts to forecast the quantity of interest. We can consider each agent as having varying information denseness, referring to how fine or coarse the data is. By synthesizing these agents, the BPS framework allows for dynamic learning of the dependencies and biases of how, for example, daily data provides information on monthly quantities. The resultant method is called frequency auto-splicing synthesis (FASST), as the method auto-splices the data into varying frequencies to be later synthesized.

6.3.1 Data

For this study, we analyze daily US equity market log prices from 1999/3-2014/12. The data is then split up into lower frequencies up until they are monthly data. Long term forecasts are significantly impaired during the sub-prime mortgage crisis at around 2008 due to structural breaks, providing a challenging environment to test the predictive ability of varying temporal specifications.

Accompanying the market log prices, we use the Fama-French three factors (Fama and French, 1993), used often in the finance literature, as predictors. These are; **Mkt-RF**: the performance of the (risky) market as a whole to a risk-free asset (daily returns), **SMB**: Small Minus Big, a factor portfolio that measures the performance of small cap stocks relative to large cap stocks (daily returns), and **HML**: High Minus Low, a factor portfolio that measures the performance of "value" stocks compared to "growth" stocks (daily returns).

6.3.2 Agents and BPS specifications

For the analysis for market log prices, we have $J = 6$ agents. Labeling them M*, the agent models vary in the frequency used: M1- 1-step, monthly frequency; M2- 2-step, interval at every two weeks; M3- 4-step, weekly frequency; M4- 5-step, interval at every four days; M5- 10-step, interval at every other day; M6- 20-step, daily frequency. All models are estimated using an AR(1) term and dynamic regression on the Fama-French three factors as predictors. Each agent model then forecasts the log prices at the end of the month, e.g., M6 forecasts 20-steps ahead using daily data, M3 forecasts 4-steps ahead using weekly data, etc.

FIGURE 6.5: US market index forecasting 2007/2-2014/12: US market index (returns ×100 for % basis).

Prior specifications for the DLM state vector and discount volatility model in each is based on $\boldsymbol{\theta}_0|v_0 \sim N(\mathbf{0}, v_0\boldsymbol{I})$ and $1/v_0 \sim G(1, 0.01)$, using the usual $(\boldsymbol{\theta}, v)$ DLM notation (West and Harrison, 1997, Chap 4). Each agent model uses standard discount factor (β) specification for state evolution variances and discount factor (δ) for residual volatility; we use $(\beta, \delta) = (0.99, 0.95)$ in each of these agent models.

In the dynamic BPS models, we take initial priors as $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}, \boldsymbol{C})$ with $\boldsymbol{m} = (0, \mathbf{1}'/p)'$, $\boldsymbol{C} = diag(0.001, 10\boldsymbol{I})$, and $1/v_0 \sim G(5, 0.01)$. The discount factor for BPS is based on $(\beta, \delta) = (0.95, 0.99)$.

6.3.3 Data analysis and forecasting

The 6 agent models are analyzed and synthesized as follows. First, the agent models are analyzed in parallel over 1999/3-2003/1 as a training period, simply running the DLM forward filtering to the end of that period to calibrate the agent forecasts. This continues over 2003/2-2007/1 now accompanied by the calibration of the other forecast combination methods. Also, at each quarter t during this period, the MCMC-based BPS analysis is run using from 2003/2 data up to time t;

that is, we repeat the analysis with an increasing "moving window" of past data as we move forward in time. This continues over the third period to the end of the series, 2007/2-2014/12; now we also record and compare forecasts as they are sequentially generated. This testing period spans over a quarter century, and we are able to explore predictive performance over periods of drastically varying economic circumstances, check robustness, and compare benefits and characteristics of each strategy. Out-of-sample forecasting is thus conducted and evaluated in a way that mirrors the realities facing portfolio managers.

6.3.4 *Forecast accuracy and comparison*

As with Section 3.1.4, we compare forecast accuracy using mean squared forecast errors (MSFE) and log predictive density ratios (LPDR).

We compare forecasts from FASST with standard Bayesian model uncertainty analysis (i.e. BMA) in which the agent densities are mixed with respect to sequentially updated model probabilities (e.g. Harrison and Stevens, 1976; West and Harrison, 1997, Sect 12.2). In addition, we compare with simpler, equally-weighted averages of agent forecast densities using linear pools (equally-weighted arithmetic means of forecast densities).

6.3.5 *Frequency auto-splicing synthesis*

Table 6.2 summarizes the out-of-sample results of the temporal specifications considered and FASST. First to note is how there is a clear tradeoff between point forecast accuracy and distribution forecast accuracy. Using higher frequency data generally produces more accurate point forecasts, as seen in the MSFE results, with the model using monthly data performing the worst. On the other hand, lower frequency data generally produces better distribution (uncertainty) forecasts, as seen in the LPDR results, with the model using daily data performing the worst. FASST

Table 6.2: US market index forecasting 2007/2-2014/12: Forecast evaluations for monthly US market index, comparing mean squared forecast errors and log predictive density ratios for this $T = 100$ months. Note: $\text{LPDR}_{1:T}$ is relative to FASST.

	$\text{MSFE}_{1:T}$	$\text{LPDR}_{1:T}$
1-step	0.002643	-11.8976
2-step	0.002616	-10.2972
4-step	0.002128	-32.5566
5-step	0.002431	-67.6783
10-step	0.002021	-256.5949
20-step	0.002265	-267.6800
BMA	0.002604	-8.0748
LinP	0.002233	-7.5011
FASST	0.002025	-

mitigates this tradeoff, demonstrated by both point and distribution forecasts being either the best (LPDR) or a close second (MSFE) compared to the other models. By synthesizing information coming from multiple frequencies, FASST essentially takes the better of two worlds, good mean assessments from high frequency data and good uncertainty assessments from low frequency data. BMA, for example, is unable to achieve this, as it measures model probabilities based on distribution forecasts, favoring low frequency data due to its uncertainty measure. Linear pool also mitigates the temporal misspecification, though significantly less compared to FASST.

Figs. 6.6-6.10 summarize sequential analysis for 1-step forecasting. Fig. 6.6 shows the point forecast ability over the testing period. Over the whole period, FASST is almost equivalent, in terms of point forecasts, to the best model (10-step ahead model using data from every other day). The forecast accuracy is almost proportional to the data frequency used. Additionally, BMA degenerates to one of the worst models (Fig. 6.7) and shifts to the worst model at the end. Overall, FASST demonstrates its ability to outperform most models and strategies over the period examined.

Distribution forecasts compared using LPDR (Fig. 6.8) contrasts with the MSFE results. The best models in terms of MSFE are now one of the worst models, consistently underperforming throughout the testing period. BMA benefits here from degenerating to the model with monthly data. However, FASST does outperform all other models throughout most of the testing period, improving on all in terms of distribution forecasts.

The improvement in terms of distribution forecasts is partly due to improved uncertainty forecasts under FASST. Focusing on the forecast standard deviation (Fig. 6.9), models using higher frequency data have significantly lower forecast standard deviations. This is because the models are fitted on data with smaller variability, since daily price variation is much smaller than monthly price variation, and the models fail to adapt uncertainty to long term forecasts. Even for models using low frequency, the discrepancies with the forecast standard deviations from FASST are notable. In particular, forecast standard deviations using FASST

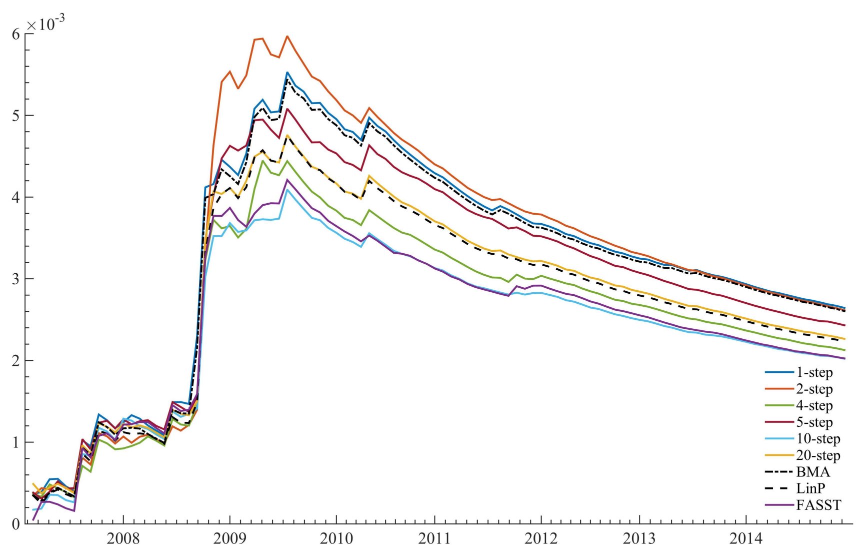

FIGURE 6.6: US market index forecasting 2007/2-2014/12: Mean squared 1-step ahead forecast errors $\text{MSFE}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}100$ months.

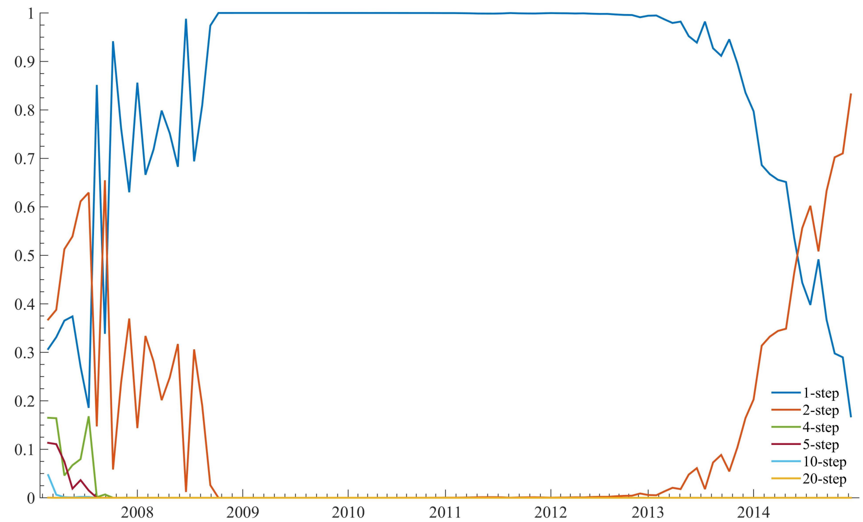

FIGURE 6.7: US market index forecasting 2007/2-2014/12: On-line model probabilities for BMA sequentially computed at each of the $t = 1{:}100$ months.

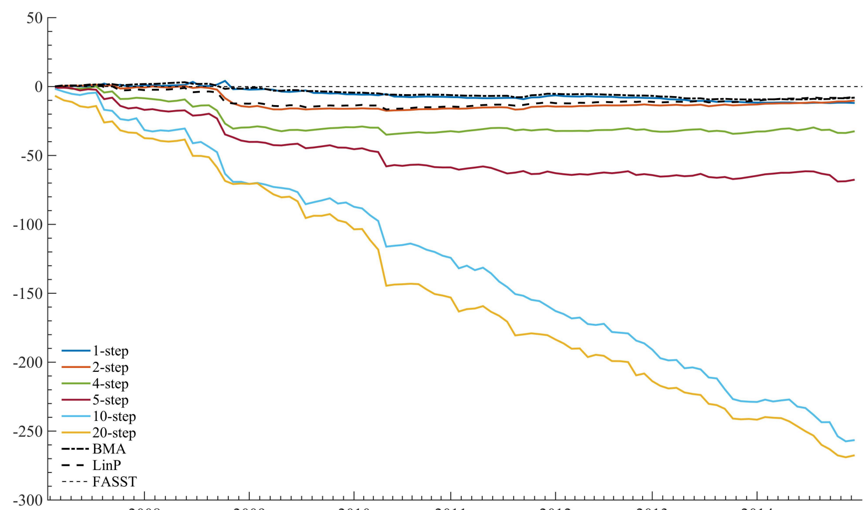

FIGURE 6.8: US market index forecasting 2007/2-2014/12: 1-step ahead log predictive density ratios $\text{LPDR}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}100$ months. The baseline at 0 over all t corresponds to the standard BPS model.

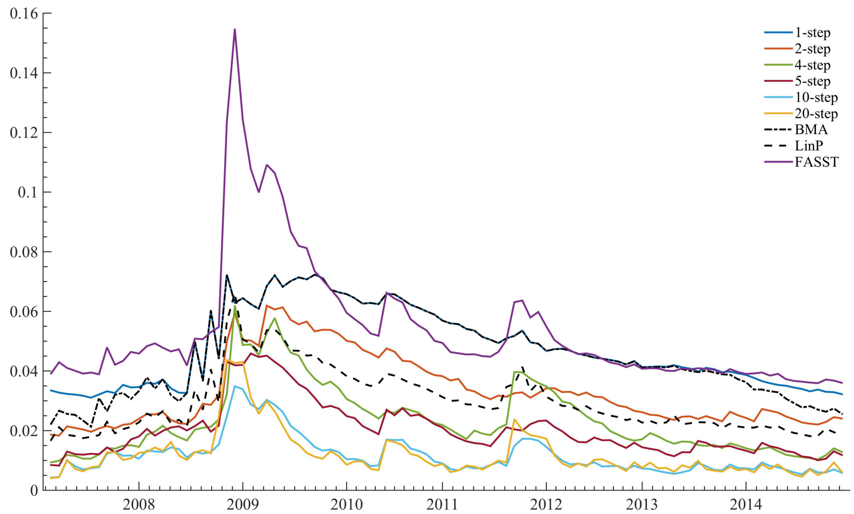

FIGURE 6.9: US market index forecasting 2007/2-2014/12: 1-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}100$ months.

drastically jump during the sub-prime mortgage crisis, doubling that of the largest standard deviation from the models, though quickly converging to the levels of other models. This additional adaptability allows FASST to improve distribution forecasts over all models.

The on-line posterior means of FASST model coefficients are quite different from the coefficients seen in the other studies. While most studies have coefficients summing to roughly one, with all coefficients within $[0, 1]$ for the most part, some coefficients are above one and significantly negative in another. In essence, these two coefficients (models using the highest frequency data) counter-balance each other in order to improve forecasts. This is a major feature of BPS approaches: The analysis properly integrates and responds to time-varying interdependencies among agents. Another interesting point to note is how, apart from the coefficients for 10- and 20-step ahead models that counter-balance each other, the model with the highest coefficient is the 4-step ahead model, which sits in the middle in terms of point and distribution forecasts. From this it can be inferred that FASST bases

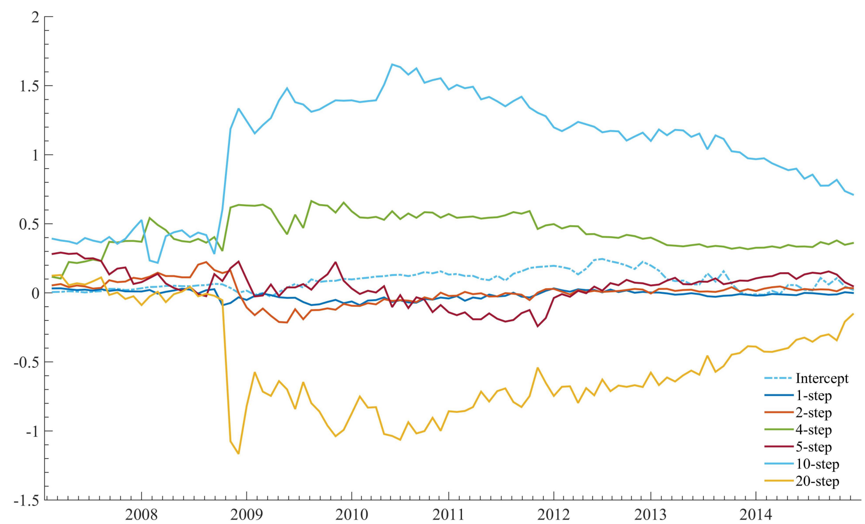

FIGURE 6.10: US market index forecasting 2007/2-2014/12: On-line posterior means of FASST model coefficients sequentially computed at each of the $t = 1{:}100$ months.

its forecasts on the most balanced model, then adapts by changing the coefficients of the other models. Interestingly, the models with the best point forecasts (1- and 2-step) is close to zero. This would indicate that FASST relies less on a model's point forecast ability but more on the uncertainty measurements. However, unlike BMA, FASST does not degenerate to the best distribution forecast model and is able to improve in terms of both point and distribution forecasts due to its dynamic adaptability.

6.4 Summary

Temporal misspecification is a critically important– yet underdeveloped– subject in the analysis and decision making of time series data. The two applications presented in this chapter tackle the two key problems in temporal misspecification: rolling window specification and frequency specification for long terms forecasts. Agents are considered as models with different window sizes or frequency, therefore containing different information. By dynamically synthesizing the set of varying information, BPS improves forecast performances and mitigates the problem of temporal misspecification. The dynamic dependence among these agents (information) are exploited by BPS to dominate the available temporal space and provide insight into the temporal relationship within a time series.

7

BPS for Mixed-Frequency Time Series Forecasting

7.1 Context, literature, and methodology

Time series data are often available at multiple sampling frequencies. For example, US Gross Domestic Product (GDP) is released every quarter, while many other economic variables (including inflation, unemployment, etc.) are released every month, not to mention financial variables released everyday or every micro second. These higher frequency indicators often contain crucial information in understanding the economy. Therefore, it is important to exploit these data in order to make more informed policy decisions.

One approach that has become particularly popular among practitioners and researchers is the mixed data sampling regression (MIDAS) proposed by Ghysels et al. (2005) in the context of utilizing high frequency financial data, and later extended for macroeconomic applications by Clements and Galvao (2008, 2009). The basic idea of the MIDAS regression is to fit some function, often polynomial or some variation, to the higher frequency, which is then included in the lower frequency regression,

$$y_t = \beta_0 + \beta_1 B(L^{1/m};\theta)x_t^{(m)} + \epsilon_t^{(m)}$$

where $B(L^{1/m};\theta)$ is the lag distribution.

While the MIDAS regression provides practical utility, it has several shortcoming. The first, and perhaps most important, is that MIDAS does not provide any theoretical and conceptual foundation to combine information from different frequencies from a Bayesian perspective. The question on how a subjective Bayesian can incorporate information from different frequencies is unanswered. Second, the MIDAS form is heavily restrictive in terms of application, especially in the directions of dynamics and multivariate data.

Taking the approach of incorporating information from multiple frequencies, we propose to utilize the BPS framework in order to achieve potentially superior results and flexibility compared to the MIDAS regression. The proposed Bayesian mixed frequency synthesis (MFS) is a two step framework. The first step projects the high frequency data onto the low frequency of interest. This is done by regressing the high frequency data to the low frequency data using any method $\mathcal{D}$ might find useful in extracting the necessary information. By projecting the data, the resultant forecast distribution from $\mathcal{D}$'s chosen process is essentially filtered information pertaining to the quantity of interest. $\mathcal{D}$ can then generate filtered information for multiple sources of high frequency data. Upon generating multiple sets of information from the high frequency data, $\mathcal{D}$ synthesizes them using the BPS framework. From a Bayesian point of view, the framework utilizes multiple priors drawn from different agents using different data sources. Those priors are then calibrated for dependencies, biases, and miscalibration as they are synthesized to maximize the usage of information provided from these agents (mixed frequency data sources).

Not only does Bayesian mixed frequency synthesis provide a theoretically and philosophically coherent framework for synthesizing data from multiple frequencies, it also grants $\mathcal{D}$ flexibility in terms of modeling. For example, $\mathcal{D}$ can use a linear regression or DLM to project the information, but if the frequency disparity is high (high frequency data sampled at very small intervals; for example, finance data sampled every minute), $\mathcal{D}$ can utilize shrinkage methods (Nakajima and West, 2013a), PCA, or polynomial functions to reduce the dimension. This can be specified for each high frequency $\mathcal{D}$ might want to include to forecast the low frequency data. Additionally, $\mathcal{D}$ can specify the BPS synthesis function, $\alpha_t(y_t|\mathcal{H}_t)$, to express $\mathcal{D}$'s belief on how these information sources should be synthesized.

7.2 Nowcasting GDP

7.2.1 Data

We analyze quarterly US macroeconomic data, focusing on nowcasting GDP with 1-quarter ahead interests using monthly US macroeconomic data. The study involves one quarterly macro series and three monthly macro series (Fig. 4.1): GDP (quarterly), industrial production (IP; monthly), employment (EMP; monthly), and capacity utilization (CU; monthly) in the US economy from 1970/1 to 2015/12, a context of topical interest (Clements and Galvo, 2014; Aastveit et al., 2014, 2016). Quarterly real output growth (GDP) is measured as the log-difference of the real GDP series. We focus on nowcasting GDP using past values of itself and the higher frequency predictors underlying a set of four time series models– the $J = 4$ agents– to be evaluated, calibrated, and synthesized. The time frame includes key periods that warrant special attention: the early 1990s recession, the Asian and Russian financial crises in the late 1990s, the dot-com bubble in the early 2000s, and the sub-prime mortgage crisis and great recession of the late 2000s. Sharp shocks

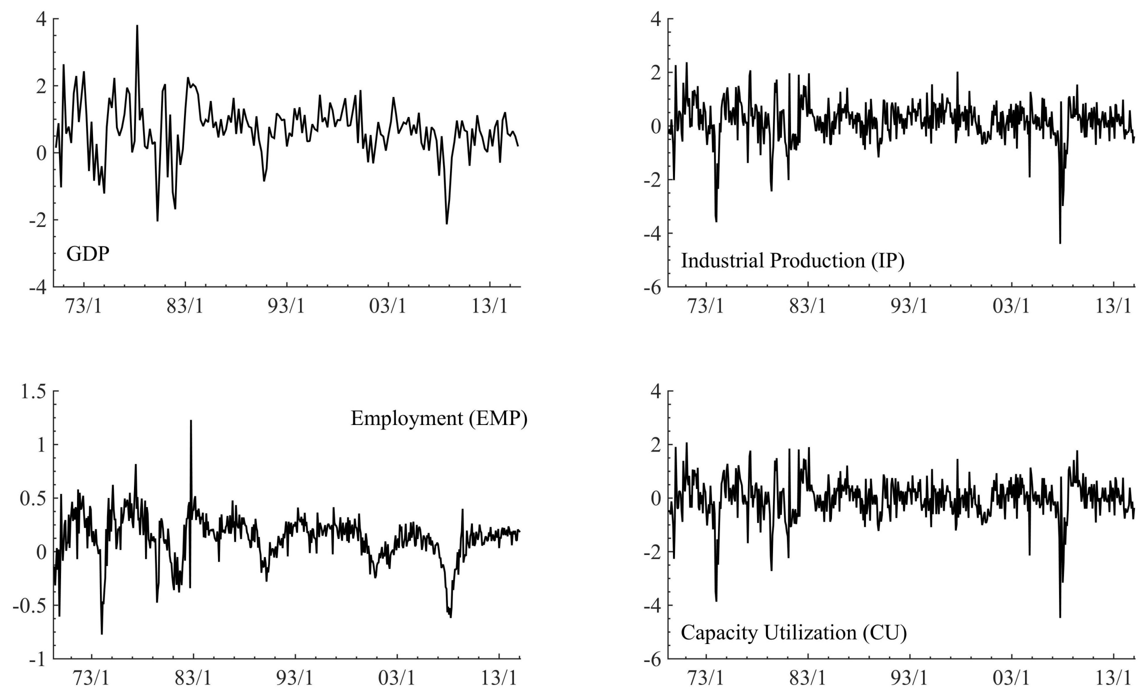

FIGURE 7.1: US macroeconomic series 1991/Q1-2015/Q4: Quarterly and monthly US macroeconomic time series (indices $\times 100$ for $\%$ basis).

to the US economy are evident during this period testing the predictive ability of any models and strategies under pressure. For any mixed frequency strategy to be effective and useful, its predictive performance must be robust under these conditions; most traditional macroeconomic models suffer significant deficiencies in such times.

7.2.2 Agents and BPS specifications

For this analysis of GDP nowcasting, we have $J = 4$ agents. Labeling them M*, the agent models are: M1- AR(3); M2- CU(3); M3- EMP(3); M4- IP(3). M1 is a time-varying AR(3) for GDP and the quarterly period t, which is straightforward. On the other hand, M2-M4 need further discussion. At each quarter t, we denote t/m as the month $m = 1:3$ in that quarter t. Thus, $t/1$ is the first month of the quarter and so on. At any given t/m, M2-M4 forecast GDP at t using lagged monthly data up until t/m as predictors. For example, $t/2$ would be data from [$t/2$ $t/1$ $t-1/3$] if we were to use lags of 3. These are also called *leads*, as these monthly predictors lead into the quarter. A lead of 0 indicates that there are no leads and the predictors are only up until t, while a lead of 2 indicates that the predictors are up until $t+1/2$; the second month into the next quarter.

The crucial idea here is that each M*, except for M1, is a projection of the high frequency information onto the low frequency quantity of interest, GDP. By regressing the higher frequency data onto GDP, we are essentially filtering the monthly information pertaining to GDP. Thus, the forecasts coming from each model can be seen as filtered information to be later synthesized to maximize information gained in terms of the lower frequency quantity of interest.

Each model is fit using a DLM. Prior specifications for the DLM state vector and discount volatility model in each is based on $\boldsymbol{\theta}_0|v_0 \sim N(\mathbf{0}, v_0\boldsymbol{I})$ and $1/v_0 \sim G(1, 0.01)$, using the usual $(\boldsymbol{\theta}, v)$ DLM notation (West and Harrison, 1997, Chap 4). Each agent model uses standard discount factor (β) specification for state evolution variances and discount factor (δ) for residual volatility; we use $(\beta, \delta) = (0.99, 0.95)$ in each of these agent models.

In the dynamic BPS models, we take initial priors as $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}, \boldsymbol{I})$ with $\boldsymbol{m} = (0, \mathbf{1}'/p)'$, and $1/v_0 \sim G(5, 0.01)$. The discount factor for BPS is based on $(\beta, \delta) = (0.99, 0.95)$.

7.2.3 Data analysis and forecasting

The 4 agent models are analyzed and synthesized as follows. First, the agent models are analyzed in parallel over 1970/Q1-1978/Q2 as a training period, simply running the DLM forward filtering to the end of that period to calibrate the agent forecasts. This continues over 1978/Q3-1990/Q4 now accompanied by the calibration of the MIDAS regression using AR(3)-U-MIDAS (Foroni et al., 2015), for each of the three high frequency series (denoted by MIDAS($*$), where * is the high frequency data used). Also, at each quarter t during this period, the MCMC-based BPS analysis is run using from 1978/Q3 data up to time t; that is, we repeat the analysis with an increasing "moving window" of past data as we move forward in time. We do this for the traditional 1-step focused BPS model with no leads (i.e. not using any monthly data between t and $t + 1$), and– separately and in parallel– for a 1-step ahead focused BPS model with a lead of two months, utilizing data two months into the quarter. This continues over the third period to the end of the series, 1991/Q1-2015/Q4; now we also record and compare forecasts as they are sequentially generated. This testing period spans over a quarter century, and we are able to explore predictive performance over periods of drastically varying economic circumstances, check robustness, and compare benefits and characteristics of each strategy. Out-of-sample forecasting is thus conducted and evaluated in a way that mirrors the realities facing decision and policy makers.

7.2.4 Forecast accuracy and comparison

As with Section 3.1.4, we compare forecast accuracy using mean squared forecast errors (MSFE) and log predictive density ratios (LPDR).

7.3 Dynamic MFS and nowcasting

Comparing predictive summaries over the out-of-sample period, BPS improves forecasting accuracy relative to the TVAR(3) model and three MIDAS regressions; see numerical summaries in Table 7.1. Focusing on the point nowcast, excluding MFS, we see that TVAR(3) outperforms MIDAS for no leads but MIDAS(EMP) and MIDAS(IP) outperforming TVAR(3) when we allow leads. This is consistent with previous research, as we expect MIDAS to perform better with leading information into the quarter, getting a better grasp on the economy to come. MFS, on the other

Table 7.1: US GDP nowcasting 1991/Q1-2015/Q4: Nowcast evaluations for quarterly US GDP, comparing mean squared forecast errors and log predictive density ratios for this $T = 100$ quarters. Note: $\text{LPDR}_{1:T}$ is relative to BPS and $t - \#$ is the number of leads into the next quarter.

	$t/0$		$t/2$	
	$\text{MSFE}_{1:T}$	$\text{LPDR}_{1:T}$	$\text{MSFE}_{1:T}$	$\text{LPDR}_{1:T}$
AR(3)	0.3077	-73.0468	0.3077	-82.2589
MIDAS(CU)	0.3420	-25.9444	0.3938	-57.8255
MIDAS(EMP)	0.3669	-46.8859	0.2977	-32.1340
MIDAS(IP)	0.3368	-29.6841	0.3077	-64.3427
MFS	0.2682	-	0.2273	-

hand, significantly outperforms all of the other models. By filtering, calibrating, and synthesizing the information from the higher frequency data, MFS is able to achieve superior point forecasts over MIDAS. Additionally, allowing for leads, MFS outperforms itself with no leads, demonstrating how utilizing high frequency data can improve nowcasting throughout a quarter. These results are echoed for the distribution nowcast as well, as MFS is available to ascertain better measures of uncertainty from the high frequency data. We also see that comparing MFS with and without leads gives us an LPDR of -9.2121, improving over itself using data in between quarters.

We further our analysis by reviewing summary graphs showing aspects of analyses evolving over time during the testing period, a period that includes challenging economic times that impede good predictive performance.

Figs. 7.2 and 7.3 show the 1-step ahead measures $\text{MSFE}_{1:t}(1)$ for each time t for no leads and two month leads. For no leads, we see that TVAR(3) outperforms the rest up until the sub-prime mortgage crisis, but MFS eventually surpasses TVAR(3). With leads, MFS outperforms the rest throughout most of the period of analysis. Note particularly that the shock during the sub-prime mortgage crisis is mitigated in MFS, while the other models suffer greatly as they poorly adapt.

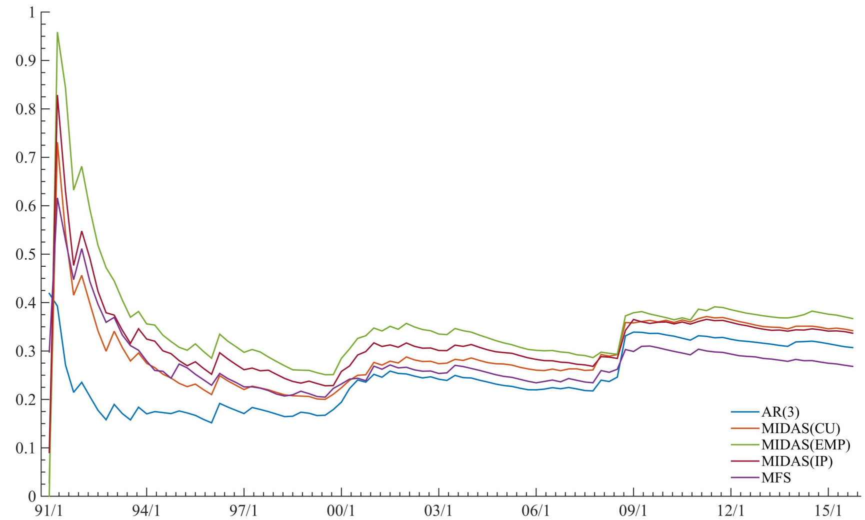

FIGURE 7.2: US GDP nowcasting 1991/Q1-2015/Q4: Mean squared 1-quarter ahead forecast errors $\text{MSFE}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}100$ quarters with no leads.

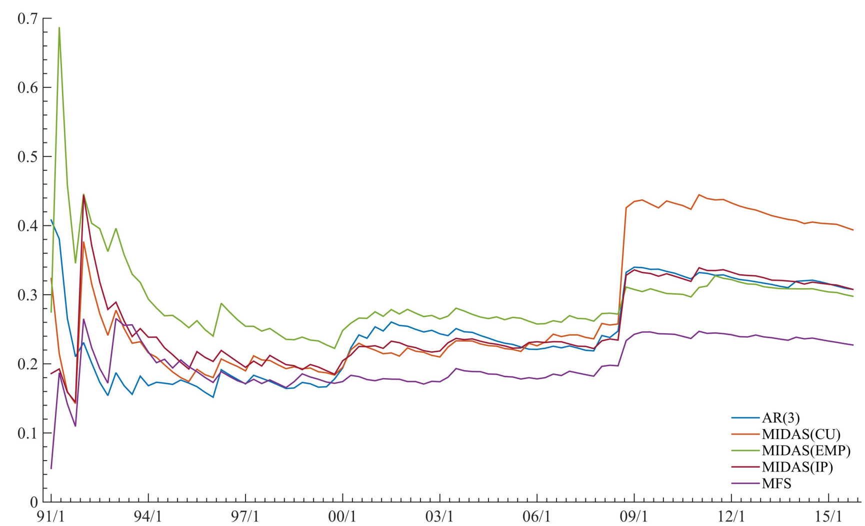

FIGURE 7.3: US GDP nowcasting 1991/Q1-2015/Q4: Mean squared 1-quarter ahead forecast errors $\text{MSFE}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}100$ quarters with leads of two months.

The distribution forecast results, Figs. 7.4 and 7.5, echo the point nowcast results with MFS performing better than MIDAS over almost all periods, except in the very early periods of testing. In particular, during crisis periods, we see significant decrease in LPDR of the TVAR(3) model. This demonstrates how MFS, and to a lesser extent MIDAS, is robust under shocks by utilizing high frequency data that provide crucial, up-to-date information that better represent the current economic situation. This is prevalent for MFS with leads, as shocks can happen within a quarter that cannot be picked up by an TVAR(3) model only using quarterly data.

Figs. 7.6 and 7.7 compare the 1-step ahead standard deviation results. Note that MFS and MIDAS both have lower standard deviations when we include leads, reflecting refined filtered nowcasts throughout a quarter. However, due to the restriction of MIDAS to static volatility, we see that it fails to adapt to large shocks, namely the sub-prime mortgage crisis. MFS, on the other hand, is able to adapt dynamically, as seen in the jump in standard deviation around 2009. This flexibility in volatility is reflected through the LPDR results.

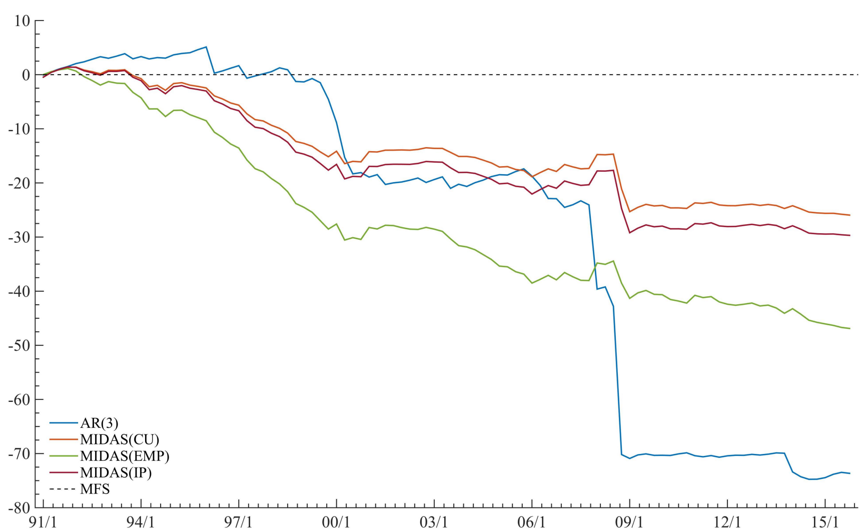

FIGURE 7.4: US GDP nowcasting 1991/Q1-2015/Q4: 1-quarter ahead log predictive density ratios $\text{LPDR}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}100$ quarters with no leads. The baseline at 0 over all t corresponds to the standard BPS model.

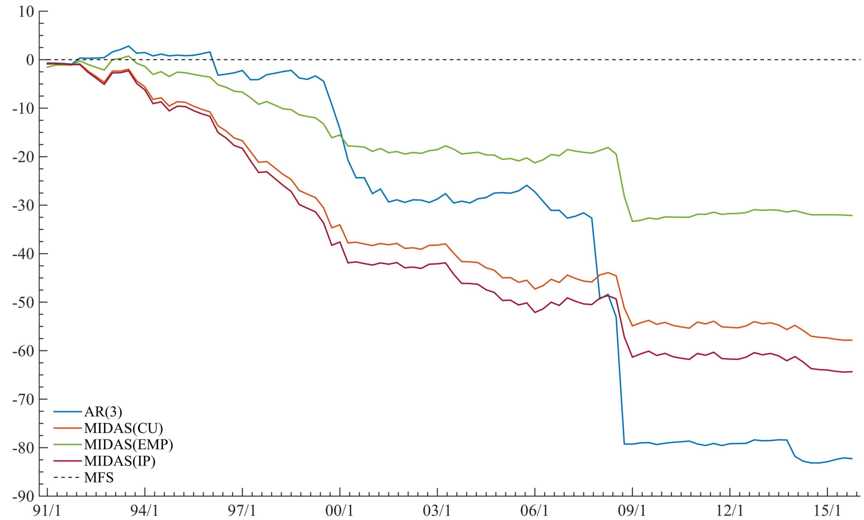

FIGURE 7.5: US GDP nowcasting 1991/Q1-2015/Q4: 1-quarter ahead log predictive density ratios $\text{LPDR}_{1:t}(1)$ sequentially revised at each of the $t = 1{:}100$ quarters with no leads. The baseline at 0 over all t corresponds to the standard BPS model.

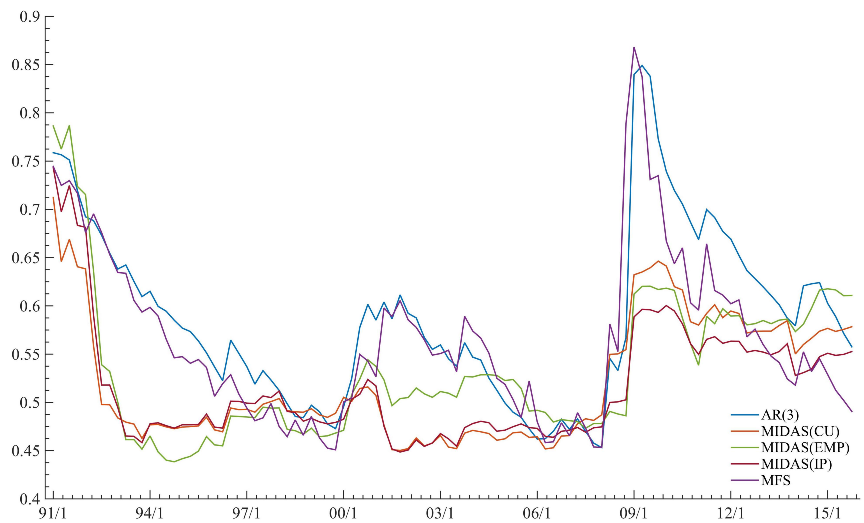

FIGURE 7.6: US GDP nowcasting 1991/Q1-2015/Q4: 1-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}100$ quarters with no leads.

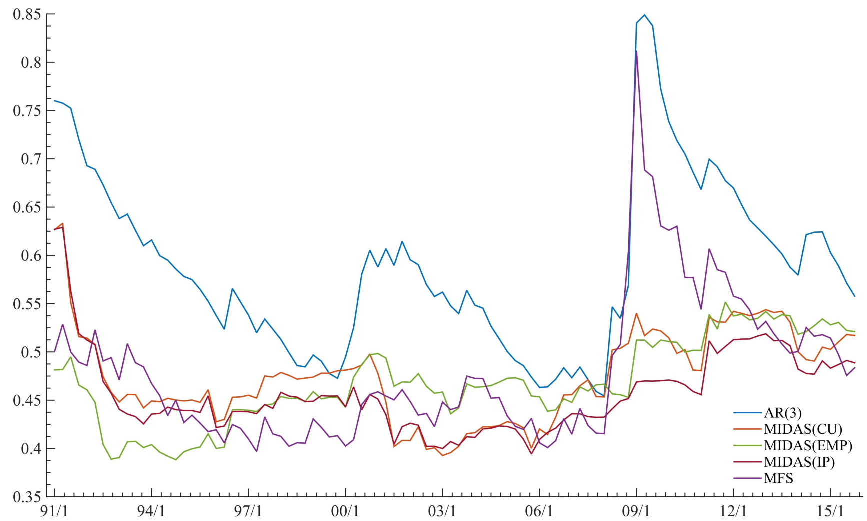

FIGURE 7.7: US GDP nowcasting 1991/Q1-2015/Q4: 1-step ahead forecast standard deviations sequentially computed at each of the $t = 1{:}100$ quarters with leads of two months.

Moving on to MFS on-line posterior means of MFS coefficients (Figs. 7.8 and 7.9), we get a better understanding on how MFS improves nowcasts by dynamically adapting the high frequency information. For the nowcasts with no leads (Fig. 7.8), the information coming from employment is the most significant. During the crisis, there is a significant increase in information provided by industrial production, perhaps because industrial production was quicker to adapt to the shock compared to employment and capacity utilization. Interestingly, GDP (TVAR(3)) provides very little information in terms of model coefficients. This can be seen as a lack of persistence in the lags of GDP, with more information coming from economic indicators that effect GDP rather than GDP of past. This is particularly useful for policy makers making decisions based on metrics they can control, which are often high frequency. For nowcasts with leads (Fig. 7.9), there is less variability during the shock, but more of a persistent structural change over time. Like that of nowcasts with no leads, employment starts out providing the most information, then is later taken over by industrial production at around the dot-com bubble. Capacity

FIGURE 7.8: US GDP nowcasting 1991/Q1-2015/Q4: On-line posterior means of MFS model coefficients sequentially computed at each of the $t = 1{:}100$ quarters with no leads.

utilization decreases over time in an almost inverse proportion to industrial production. This can be seen as the economy shifting from a more manufacturing economy to a more IT based economy, with the persistent decline of the industrial sector having a larger effect on the economy as a whole.

Finally, we note that misspecification (the sum of on-line posterior means of MFS model coefficients) displays more stability in MFS with leads compared to no leads (Fig. 7.10). This is consistent with the results above, as we expect GDP to be better specified when we have leading information. There are also clear trends of misspecification during the dot-com bubble and the sub-prime mortgage crisis. As time progresses, the level of misspecification decreases as MFS learns the data and agents. Misspecification jumps during crises, but the jumps are more severe for MFS with no leads. By incorporating leading information from the high frequency data, misspecification stays level relative to MFS with no leads, demonstrating how high frequency data helps in terms of specification.

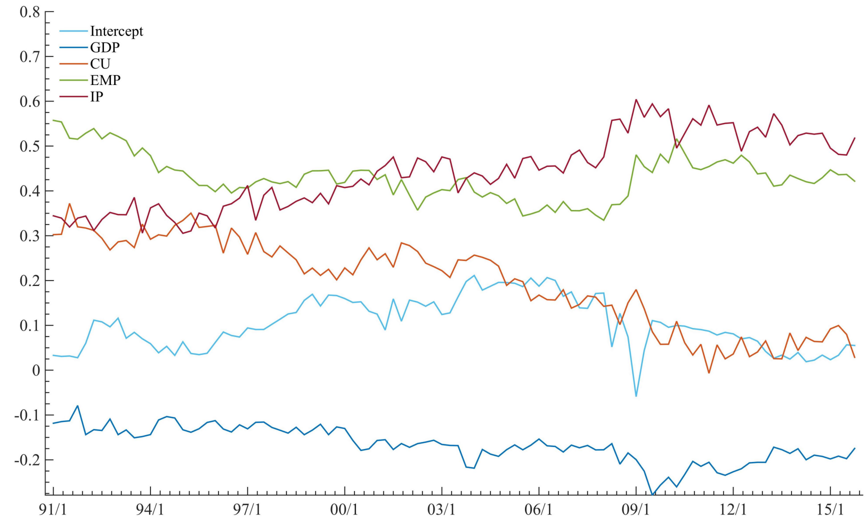

FIGURE 7.9: US GDP nowcasting 1991/Q1-2015/Q4: On-line posterior means of MFS model coefficients sequentially computed at each of the $t = 1{:}100$ quarters with leads of two months.

FIGURE 7.10: US GDP nowcasting 1991/Q1-2015/Q4: On-line posterior means of the sum of MFS model coefficients (misspecification) sequentially computed at each of the $t = 1{:}100$ quarters.

7.4 Summary

The US macroeconomic nowcasting study illustrates how BPS can be effectively extended to mixed frequency data. The developed method, MFS, dynamically synthesizes information coming from multiple frequencies in a practical and flexible way. As a result, MFS improves forecast performance and dominates methods standard in the field, MIDAS, over varying leads and for both point and distribution forecasts. An analysis into the on-line coefficients provides insight into how this information is synthesized over time and how it dynamically changes over different economic regimes. Additionally, uncertainty measures demonstrate the effectiveness of the dynamic nature of MFS under different leading information, providing updated/filtered forecasts throughout a quarter.

8

Concluding Remarks

Drawing on theory of Bayesian agent opinion analysis, BPS provides a theoretically and conceptually sound framework to compare and synthesize density forecasts that has been developed here for dynamic contexts of sequential time series forecasting. With this new framework and extension, decision makers are able to dynamically calibrate, learn, and update coefficients for ranges of forecasts from different agents, including, but not exclusive to, dynamic and static models, experts and professional forecasters, models with varying temporal specifications, and mixed frequency models. The studies explored in this dissertation illustrate the benefits of using BPS for multiple different problems that are topical in economics and finance. BPS improves forecasts and decisions and provides insight into the latent dependencies underlying the economy and market. As BPS provides a unifying framework for Bayesian forecast combination/synthesis, there is great potential in terms of theory, computation, methodology, and application to be further explored.

Apart from the foundational theory of BPS, several theoretical aspects of the synthesized forecasts and understanding of posterior analysis is undeveloped. One theoretical result that should be explored is on the optimality of the BPS synthesized forecasts compared to other restrictive strategies (linear combinations, weights summing to one, etc.). The optimality of combination of forecasts under latent dependencies, the setting and key characteristic of BPS, has not been explored in the literature. Another theoretical question concerns model misspecification. Empirically, through the studies explored, BPS coefficients on models being synthesized roughly sum to one, providing a metric to model misspecification over the whole set of agent models. However, there is no theory that indicate that the coefficients must sum to anything close to one. Further analysis into the theoretical properties of misspecification under BPS is warranted.

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

In terms of methodological development, there are several directions that BPS can take, two of which are noted here. First, the specification of the synthesis function is not fully explored in the studies explored in this dissertation. For example, the synthesis function can learn the discount factors, or incorporate stochastic volatility, in order to improve dynamics over time. If the decision maker believes that there are clear regime changes in the data and agents, a regime switching model can be used as the synthesis function to improve forecasts. In marketing, sales of products differ by weekday/weekend, weather, and so on, and one model might be useful for one characteristic but poor for another (e.g. X model forecasts special events well but performs poorly for regular weekdays). Creative ways to model the synthesis function, including modifying the state equation, can be useful in these cases. Second, the question of agent screening is not addressed. In the studies conducted in this dissertation, the number of agents are relatively small compared to the abundant number of models that a decision maker can construct or consult. There is a clear tradeoff between the number of agents and the efficiency of BPS. As the number of agents increases, the number of redundant agents increases, as it is expected that these agents are similar, increasing noise and inefficiency in BPS. There are two trains of thought to tackle this problem: thresholding and pre-screening. The process of thresholding is simple; the decision maker incorporates a mechanism in the synthesis function to threshold out unnecessary agents (e.g. Nakajima and West, 2013a,b, 2015; Zhou et al., 2014). However, it is unclear whether "unnecessary" can be well defined in the BPS setting, as an agent might be unnecessary in terms of "redundancy" or "low signal". The literature of thresholding and shrinkage focuses on the latter, however, low signal (poor) agents often improve forecasts by providing different information. Thus, thresholding should be defined as screening out redundant forecasts, though it is unclear as to how to do this through the specification of the synthesis function. Pre-screening agents would require the decision maker to screen out agents that are redundant based on some metric. One metric that might be fruitful is to look at BPS as an information maximization problem (note that information appears in the BPS theory as $\mathcal{H}$). A logical way to screen agents would then be to maximize the information provided by the agents up until a certain threshold (marginal increase in information). By doing this, good but redundant forecasts could be effectively screened out based on the information they provide, while poor but unique forecasts may be included.

This dissertation explores multiple facets of problems in time series analysis. However, there are several applications in progress in other areas. For example, using forecasts from actual people (economists, professional forecasters, etc.) is a topical subject in forecasting. One question that is central in the literature is whether forecasters exhibit herding behavior, and when. BPS can address this question through retrospective analysis of the dependence among agents. Another direction that is interesting is applications where competing "theories" that represent the world exist. In economics, finance, and marketing, researchers build structural models based on hypotheses they believe represent the economy, market, or consumer. Researchers are focused on these hypotheses, trying to validate them by demonstrating how they fit to the data. However, reality is most likely beyond these conceptual structures. BPS can provide interesting solutions to reconcile these competing hypotheses by dynamically synthesizing them and learning the latent dependencies. Under the BPS framework, and a broader philosophy of Bayesian thinking, reality is neither light nor shadow, but rather in the twilight zone. BPS thus has the potential to broaden and deepen the understanding of underlying structures of economies, market, and human behavior in a way that has been previously unexplored.

Appendix A

Posterior Computation and MCMC Algorithms

A.1 MCMC algorithm for univariate BPS

At each time, t, the decision maker has the historical information $\{\boldsymbol{y}_{1:t}, \mathcal{H}_{1:t}\}$ and specifies the following priors: $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}_0, \boldsymbol{C}_0)$ and $1/v_0 \sim G(n_0, s_0)$ and discount factors (β, δ).

Step 0. Initialize by sampling $\boldsymbol{F}_t = (1, x_{t1}, ..., x_{tJ})$ from $h_{tj}(x_{tj})$ for all $1 : t$ and $1 : J$.

Step 1. Sample $p(\boldsymbol{\Phi}_{1:t}|\boldsymbol{X}_{1:t}, \boldsymbol{y}_{1:t})$ using FFBS for the following model:

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

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

Step 2. Sample $p(\boldsymbol{X}_{1:t}|\boldsymbol{\Phi}_{1:t}, \boldsymbol{y}_{1:t}, \mathcal{H}_{1:t})$ from

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

Step 2. (a) If the forecasts $h_{tj}(x_{tj})$ are normally distributed (or mixtures thereof) as $h_{tj}(x_{tj}) \sim N(a_{tj}, A_{tj})$, then the posterior distribution for each t is

$$p(\boldsymbol{x}_t|\boldsymbol{\Phi}_t, y_t, \mathcal{H}_t) = N(\boldsymbol{a}_t + \boldsymbol{\Lambda}(y_t - \boldsymbol{a}_t^{'}\boldsymbol{\theta}_t), \boldsymbol{A}_t - \boldsymbol{A}_t\boldsymbol{\theta}_t\boldsymbol{\Lambda}),$$

where

$$\boldsymbol{a}_t = (a_{t1}, a_{t2}, ..., a_{tJ})', \quad \boldsymbol{A}_t = diag(A_{t1}, A_{t2}, ..., A_{tJ}), \quad \text{and} \quad \boldsymbol{\Lambda} = \boldsymbol{\theta}_t^{'}\boldsymbol{A}_t/(\boldsymbol{\theta}_t^{'}\boldsymbol{A}_t\boldsymbol{\theta}_t + v_t),$$

using the properties of conditional normals.

Step 2. (b) For non-normal forecasts, we can either use a Metropolis-Hastings algorithm or importance sampling scheme to sample $p(\boldsymbol{x}_t|\boldsymbol{\Phi}_t, y_t, \mathcal{H}_t)$. Alternatively, a rejection sampling could be used, though not developed here.

Metropolis-Hastings

Step 1. Sample $\boldsymbol{F}_t^* = (1, x_{t1}^*, ..., x_{tJ}^*)'$ from $h_{tj}(x_{tj}^*)$ for all $1:t$ and $1:J$.

Step 2. Accept $\boldsymbol{F}_t^*$, for all $1:t$, with acceptance probability,

$$\alpha(\boldsymbol{F}_t, \boldsymbol{F}_t^*) = \min\left\{1, \frac{p(y_t|\boldsymbol{F}_t^{*\prime}\boldsymbol{\theta}_t, v_t)}{p(y_t|\boldsymbol{F}_t^{\prime}\boldsymbol{\theta}_t, v_t)}\right\}.$$

Importance Sampling

Step 1. Compute $p(y_t|\boldsymbol{F}_t^{(i)\prime}\boldsymbol{\theta}_t, v_t)$ for each $i = 1:I$ samples from $h_{tj}(x_{tj})$.

Step 2. Compute weights,

$$w_t^{(i)} = \frac{p(y_t|\boldsymbol{F}_t^{(i)\prime}\boldsymbol{\theta}_t, v_t)}{\sum_{i=1:I} p(y_t|\boldsymbol{F}_t^{(i)\prime}\boldsymbol{\theta}_t, v_t)}.$$

Step 3. Resample $\boldsymbol{F}_t$ for $1:t$ using weights w_t.

Step 3. Repeat **Step 1.** and **Step 2.** until the samples converge.

A.2 MCMC algorithm for multivariate BPS

At each time, t, the decision maker has the historical information $\{\boldsymbol{Y}_{1:t}, \mathcal{H}_{1:t}\}$ and specifies the following priors: $\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}_0, \boldsymbol{C}_0)$ and $V_0 \sim \boldsymbol{W}^{-1}(n_0, \boldsymbol{D}_0)$ and discount factors (β, δ).

Step 0. Initialize by sampling

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

from $h_{tj}(\boldsymbol{x}_{tj})$ for all $1:t$ and $1:J$.

Step 1. Sample $p(\boldsymbol{\theta}_{1:t}|\boldsymbol{X}_{1:t}, \boldsymbol{Y}_{1:t}, \boldsymbol{V}_{1:t})$ using FFBS for the following model:

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

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

1. Forward filtering

   a. One step ahead prediction of $\boldsymbol{\theta}_t$:

   $$\boldsymbol{a}_t = \boldsymbol{m}_{t-1},$$

   $$\boldsymbol{R}_t = \boldsymbol{C}_{t-1}/\delta.$$

   b. One step ahead prediction of $\boldsymbol{y}_t$:

   $$\boldsymbol{f}_t = \boldsymbol{F}'_t\boldsymbol{a}_t,$$

   $$\boldsymbol{Q}_t = \boldsymbol{F}'_t\boldsymbol{R}_t\boldsymbol{F}_t + \boldsymbol{V}_t.$$

   c. Posterior update:

   $$\boldsymbol{m}_t = \boldsymbol{a}_t + \boldsymbol{A}_t\boldsymbol{e}_t, \quad \boldsymbol{e}_t = \boldsymbol{y}_t - \boldsymbol{f}_t,$$

   $$\boldsymbol{C}_t = \boldsymbol{A}_t\boldsymbol{Q}_t\boldsymbol{A}'_t + \boldsymbol{R}_t, \quad \boldsymbol{A}_t = \boldsymbol{R}_t\boldsymbol{F}_t\boldsymbol{Q}_t^{-1}$$

2. Backward sampling from the normal with:

$$\mathbb{E}[\boldsymbol{\theta}_t] = \boldsymbol{m}_t + \boldsymbol{C}_t \boldsymbol{R}_{t+1}^{-1}(\boldsymbol{\theta}_{t+1} - \boldsymbol{a}_{t+1}),$$

$$\mathbb{V}[\boldsymbol{\theta}_t] = \boldsymbol{C}_t - \boldsymbol{C}_t \boldsymbol{R}_{t+1}^{-1} \boldsymbol{C}_t.$$

Step 2. Individually sample $(\boldsymbol{V}_1, ..., \boldsymbol{V}_T)$ from the following inverse Wishart distributions:

$$\boldsymbol{V}_t | \boldsymbol{X}_t, \boldsymbol{y}_t, \boldsymbol{\theta}_t \sim \boldsymbol{W}^{-1}(n_t, \boldsymbol{D}_t),$$

$$n_t = \beta n_{t-1} + 1, \quad \boldsymbol{D}_t = \beta \boldsymbol{D}_{t-1} + (\boldsymbol{y}_t - \boldsymbol{F}_t' \boldsymbol{\theta}_t)(\boldsymbol{y}_t - \boldsymbol{F}_t' \boldsymbol{\theta}_t)'.$$

Step 3. Sample $p(\boldsymbol{X}_{1:t} | \boldsymbol{\Phi}_{1:t}, \boldsymbol{y}_{1:t}, \mathcal{H}_{1:t})$ from

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

utilizing the properties of conditional normals for variants of normal distribution forecasts and Metropolis-Hastings or importance sampling for more general forecast distributions (see A.1).

Step 4. Repeat **Step 1-Step 3** until convergence.

Bibliography

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. (2016), "Combined density nowcasting in an uncertain economic environment," *Journal of Business & Economic Statistics*, pp. 1–42.

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. 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*, pp. 501–524, Emerald Group Publishing Limited.

Ang, A. and Bekaert, G. (2007), "Stock return predictability: Is it there?" *Review of Financial Studies*, 20, 651–707.

Avramov, D. (2002), "Stock return predictability and model uncertainty," *Journal of Financial Economics*, 64, 423–458.

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

Belmonte, M. A., Koop, G., and Korobilis, D. (2014), "Hierarchical shrinkage in time-varying parameter models," *Journal of Forecasting*, 33(1), 480–94.

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. (2012), "Combination schemes for turning point predictions," *Quarterly Review of Finance and Economics*, 52, 402–412.

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.

Carvalho, C. M., West, M., et al. (2007), "Dynamic matrix-variate graphical models," *Bayesian analysis*, 2, 69–97.

Carvalho, C. M., Chang, J., Lucas, J. E., Nevins, J. R., Wang, Q., and West, M. (2008), "High-dimensional sparse factor modeling: Applications in gene expression genomics," *Journal of the American Statistical Association*, 103, 1438–1456.

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*, 68.

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, C. W., Liu, F. C., and Gerlach, R. (2011), "Bayesian subset selection for threshold autoregressive moving-average models," *Computational Statistics*, 26, 1–30.

Clark, T. E. and McCracken, M. W. (2009), "Improving forecast accuracy by combining recursive and rolling forecasts," *International Economic Review*, 50, 363–395.

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.

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

Clements, M. P. and Galvao, A. B. (2008), "Macroeconomic forecasting with mixed-frequency data," *Journal of Business & Economic Statistics*, 26, 546–554.

Clements, M. P. and Galvao, A. B. (2009), "Forecasting US output growth using leading indicators: an appraisal using MIDAS models," *Journal of Applied Econometrics*, 24, 1187–1206.

Clements, M. P. and Galvo, A. B. (2014), "Measuring macroeconomic uncertainty: US inflation and output growth," ICMA Centre Discussion Papers in Finance 2014/04, Henley Business School, Reading University.

Clyde, M. and George, E. I. (2004), "Model uncertainty," *Statistical science*, pp. 81–94.

Cochrane, J. H. (2008), "The dog that did not bark: A defense of return predictability," *Review of Financial Studies*, 21, 1533–1575.

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.

Cremers, K. M. (2002), "Stock return predictability: A Bayesian model selection perspective," *Review of Financial Studies*, 15, 1223–1249.

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 Schorfheid, F. (2016), "Dynamic prediction pools: an investigation of financial frictions and forecasting performance," *Journal of Econometrics*, 192(2), 391–405.

Fama, E. F. and French, K. R. (1993), "Common risk factors in the returns on stocks and bonds," *Journal of financial economics*, 33, 3–56.

Foroni, C., Marcellino, M., and Schumacher, C. (2015), "U-MIDAS: MIDAS regressions with unrestricted lag polynomials," *Journal of the Royal Statistical Society (Series A: General)*, 29, 57–82.

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.

George, E. I. and McCulloch, R. E. (1993), "Variable selection via Gibbs sampling," *Journal of the American Statistical Association*, 88, 881–889.

George, E. I. and McCulloch, R. E. (1997), "Approaches for Bayesian variable selection," *Statistica Sinica*, pp. 339–373.

George, E. I., Sun, D., and Ni, S. (2008), "Bayesian stochastic search for VAR model restrictions," *Journal of Econometrics*, 142, 553–580.

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.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2005), "There is a risk-return trade-off after all," *Journal of Financial Economics*, 76, 509–548.

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

Harrison, P. J. and Stevens, C. F. (1976), "Bayesian forecasting," *Journal of the Royal Statistical Society (Series B: Methodological)*, 38, 205–247.

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

Inoue, A., Jin, L., and Rossi, B. (2017), "Rolling window selection for out-of-sample forecasting with time-varying parameters," *Journal of Econometrics*, 196, 55–67.

Irie, K. and West, M. (2016), "Bayesian emulation for optimization in multi-step portfolio decisions," *ArXiv*, 1607.01631.

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., 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.

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

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

Lettau, M. and Van Nieuwerburgh, S. (2008), "Reconciling the return predictability evidence," *Review of Financial Studies*, 21, 1607–1652.

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.

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

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

Nakajima, J. and West, M. (2015), "Dynamic network signal processing using latent threshold models," *Digital Signal Processing*, 47, 6–15.

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.

Rossi, B. and Inoue, A. (2012), "Out-of-sample forecast tests robust to the choice of window size," *Journal of Business & Economic Statistics*, 30, 432–453.

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

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

Stella, A. and Stock, J. H. (2012), "A state-dependent model for inflation forecasting," *FRB International Finance Discussion Paper*, 1062.

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.

Stock, J. H. and Watson, M. W. (1999), "Forecasting inflation," *Journal of Monetary Economics*, 44(2), 293–335.

Stock, J. H. and Watson, M. W. (2007), "Why has US inflation become harder to forecast?" *Journal of Money, Credit and Banking*, 39, 3–33.

Terui, N. and van Dijk, H. K. (2002), "Combined forecasts from linear and nonlinear time series models," *International Journal of Forecasting*, 18, 421–438.

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

Wang, H. (2010), "Sparse seemingly unrelated regression modelling: Applications in finance and econometrics," *Computational Statistics & Data Analysis*, 54, 2866–2877.

Welch, I. and Goyal, A. (2008), "A comprehensive look at the empirical performance of equity premium prediction," *The Review of Financial Studies*, 21, 1455–1508.

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

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

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

West, M. (2003), "Bayesian factor regression models in the "large p, small n" paradigm," in *Bayesian Statistics 7*, eds. J. Bernardo, M. Bayarri, J. Berger, A. Dawid, D. Heckerman, A. Smith, and M. West, pp. 723–732, Oxford University Press.

West, M. (2013), "Bayesian dynamic modelling," in *Bayesian Theory and Applications*, eds. P. Damien, P. Dellaportes, N. G. Polson, and D. A. Stephens, chap. 8, pp. 145–166, Clarendon: Oxford University Press.

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 edn.

Yoshida, R. and West, M. (2010), "Bayesian learning in sparse graphical factor models via variational mean-field annealing," *Journal of Machine Learning Research*, 11, 1771–1798.

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.

Biography

Kenichiro Koizumi McAlinn was born in Tokyo, Japan on July 31, 1986 as a synthesis of two priors from Japan and the US. He received his bachelor degree in Economics from Keio University in 2010, studying Bayesian statistics and film theory, writing two theses on spectator engagement of characters in film. He received a dual M.A. degree in Economics from Keio University and Economics and Public Policy from L'Institut d'Etudes Politiques de Paris in 2012 and a M.S. degree in Statistical Science from Duke University in 2015. During his graduate years, he has received support from the Japan service support organization scholarship, French national scholarship, and SAS econometric fellowship. He was the winner of the BEST award in 2013, received two travel awards from ISBA, and nominated for the best TA award in 2015. Ken will continue to update his misspecified prior.