BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel—2024-30

Last updated: August 16, 2024

Decision Synthesis in Monetary Policy

by Tony Chernis,[1] Gary Koop,[2] Emily Tallman[3] and Mike West[3]

[1]Canadian Economic Analysis Department, Bank of Canada
tchernis@bankofcanada.ca

[2]University of Strathclyde

[3]Duke University

DOI: https://doi.org/10.34989/swp-2024-30 | **ISSN** 1701-9397

Acknowledgements

Research of Emily Tallman was partially supported by the US National Science Foundation through NSF Graduate Research Fellowship Program grant DGE~2139754. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Further, the views expressed in this paper are solely those of the authors and may differ from the official views of the Bank of Canada. No responsibility for the views expressed in this paper should be attributed to the Bank of Canada.

Abstract

The macroeconomy is a complicated dynamic system with significant uncertainties that make modelling difficult. Consequently, decision-makers consider multiple models that provide different predictions and policy recommendations and then synthesize that information into a policy decision. We use Bayesian predictive decision synthesis (BPDS) as a way formalize this monetary policy decision-making process. BPDS draws on recent developments in model combination and statistical decision theory that make it possible to combine models in a manner that incorporates decision goals, expectations and outcomes. We develop a BPDS procedure for a case study of monetary policy decision-making with an inflation-targeting central bank and compare the results against standard model-combination approaches.

Topics: Econometric and statistical methods; Economic models; Monetary Policy
JEL codes: C11, C32, C53

Résumé

La macroéconomie est un système dynamique complexe qui comporte de grandes incertitudes, ce qui rend difficile sa modélisation. Par conséquent, les décideurs prennent en compte de multiples modèles qui fournissent différentes prévisions et recommandations de politique, puis ils synthétisent cette information pour formuler une décision de politique monétaire. Nous utilisons la synthèse bayésienne des prévisions et des décisions pour formaliser ce processus de prise de décision sur la politique monétaire. Notre approche s'inspire des récentes avancées dans la combinaison de modèles et la théorie statistique de la décision, qui permet de combiner les modèles de manière à intégrer les objectifs, les attentes et les résultats relatifs aux décisions. Nous élaborons une procédure, fondée sur la synthèse de prévisions bayésienne, afin de réaliser une étude de cas portant sur la prise des décisions de politique monétaire par une banque centrale qui cible l'inflation et de comparer les résultats à ceux d'approches standard de combinaison de modèles.

Sujets : Méthodes économétriques et statistiques ; Modèles économiques ; Politique monétaire
Codes JEL : C11, C32, C53

1 Introduction

Monetary policy makers are tasked with simple, but hard to achieve, objectives. A common objective is to target the future inflation rate, or other macroeconomic outcomes, using the interest rates as the policy instrument. These decisions are made based on uncertain information from many sources. In this paper, these sources are econometric models that generate predictive distributions for the macroeconomic outcomes and the policy instruments over multiple time periods. For a single model, it is straightforward to select an optimal policy instrument using decision analysis and conditional forecasting. Applying standard methods– such as Bayesian model averaging (BMA)– is one way to address the issue of model uncertainty, as routine decision analysis can then be applied to the weighted average as a single model. However, this traditional view ignores the reality that a set of models may each individually recommend very different optimal policy decisions. The question then arises as to how to synthesize this information and, potentially, exploit it in the overall final decision process. This paper addresses and answers this question.

There is an extensive Bayesian econometrics literature on model combination, but discussion of the question that models are built for purposes– specific prediction and decision goals– is very sparse. Traditional BMA analysis weights models according to purely statistical model fit, and in time series explicitly and only scores 1-step ahead forecast outcomes. Extensions and alternatives have arisen to define model weightings based on aspects of past forecast performance with respect to specific forecast goals. Martin et al. (2023) survey Bayesian forecasting in economics and finance and overviews various forecast combination approaches, including some that are more explicitly concerned with goal-focused prediction (e.g. Mitchell and Hall, 2005; Geweke and Amisano, 2011; Conflitti et al., 2015; Kapetanios et al., 2015; Loaiza-Maya et al., 2021; Chernis and Webley, 2022; Aastveit et al., 2023; Bernaciak and Griffin, 2024). Lavine et al. (2021) provide additional perspectives and put many of the earlier approaches in a foundational Bayesian context, justifying model weights based on utilities in forecasting using historical model-specific "scoring" of past forecast outcomes. The underlying theoretical justifications come from Bayesian predictive synthesis (BPS) and the specific class of "mixture BPS" models (McAlinn and West, 2019, section 2.2; Johnson and West, 2022). However, while ultimate decision goals may be implicit in specific applications of model combination, they are rarely if ever taken into account in the analysis and resulting decision-making. This raises questions: A model that has fit or forecast specific outcomes well in the past may be a good bet for use in resulting decision analysis– in our settings, to define optimal decisions about values of policy instruments– but, there is no guarantee that this will be so.

Our view is that models that have recommended optimal policy decisions that turn out to be "good" should be more heavily weighted in looking ahead, just as past statistical predictive performance is– and should be– generally positively weighted. The catch, of course, is defining "good"; econometric models were and are not explicitly used and scored in past policy decisions. The challenge is then to operationalize the concept of "good decision" performance. For example, a vector autoregression (VAR) model can be evaluated on forecast performance using a pseudo real-time forecasting exercise. But, it is not clear how to evaluate such a model when used to advise policy decisions. We can, however, use the VAR model at present to inform near-term decisions and explore how it would have advised on decisions in the past. Evaluations can then compare such analysis to decisions actually made by policy makers in the past (albeit recognizing that past decisions of policy makers were not necessarily correct – rather, just the outcomes of the amorphous reality of monetary policy-making).

Bayesian Predictive Decision Synthesis (BPDS–Tallman and West, 2023; Tallman, 2024) addresses these questions. As part of the theoretical framework of Bayesian predictive synthesis (BPS–McAlinn and West, 2019; Johnson and West, 2022), BPDS explicitly allows and encourages scoring of models based on decision analysis performance as well as statistical predictive accuracy. In addition to reflecting historical outcomes of predictions and decisions, BPDS critically also allows for differential model weighting based on *expected* decision outcomes. This is a complete decision parallel to the proven use of BPS models that incorporate outcome-dependent weights that modify BMA-like mixtures to differentially favour models in different parts of the future outcome space for pure forecasting. The latter concept was introduced by Kapetanios et al. (2015), whose empirically inspired developments recognized, for example, that one model may be better at predicting inflation when inflation is high and rising, while another model may be better when inflation is low and stable. BPS defines the conceptual and theoretical Bayesian bases and a broader methodological framework for this. BPDS goes further by integrating both historical and expected decision outcomes; here we develop, extend, and exemplify BPDS in our central macroeconomic policy context.

BPDS applies the broader Bayesian mixture model approach of BPS using defined utility– or "score"–functions that relate to explicit decision goals. Importantly, this allows for multiple objectives, i.e., multi-attribute decision analysis. For example, a purely predictive vector score function can allow for multiple forecast horizons (e.g. to produce inflation near a target for each of the next eight quarters) and/or multiple outcome criteria (e.g. to separately reflect inflation targeting, interest rate smoothing, and stable growth patterns over coming quarters), among others. For policy makers juggling multiple objectives, this is a key feature; it is rather distinct from conventional approaches that adopt single, scalar criteria for model weighting. For example, a forecast combination approach might choose model weights based on the $h-$step ahead predictive likelihood for a single choice of h, with BMA simply focused on $h = 1$, whereas BPDS can address multi-steps ahead in parallel, along with scoring of outcomes of decision goals that simultaneously target several macroeconomic outcomes.

The opportunities for exploring such practically relevant questions, and the differences relative to traditional single-model and BMA-based analyses, are showcased in our empirical studies. This involves an exploratory case study using US macroeconomic data with multiple-objective score functions to define BPDS model weights, exemplifying the use of BPDS in macroeconomic forecasting and advisory decision-making.

2 BPDS Framework

We present and discuss the structure of BPDS at a particular point in time, ignoring the time dependency and relevance in the notation for clarity in communicating these essentials. Practical implementation in time series is of course sequential, with models at time t depending on all relevant historical data and information.

2.1 Mixture BPDS and Decision Setting

At a given time point, let $\mathbf{y}$ denote the $q-$dimensional outcome variable of interest (e.g. inflation in each of the next q quarters) and $\mathbf{x}$ the vector of control/decision variables (e.g., a target profile of central bank interest/base rates over the next q quarters). Each of a set of J models $\mathcal{M}_j$, $j=1{:}J,$ predicts the outcome $\mathbf{y}$ via a predictive density $p_j(\mathbf{y}|\mathbf{x},\mathcal{M}_j)$ conditional on any considered decision $\mathbf{x}$. The policy-maker responsible for ultimate decisions adopts a general BPDS approach that implies the overall conditional (on $\mathbf{x}$) predictive pdf of the form

$$f(\mathbf{y}|\mathbf{x}) \propto \sum_{j=0:J} \pi_j(\mathbf{x})\alpha_j(\mathbf{y}|\mathbf{x})p_j(\mathbf{y}|\mathbf{x},\mathcal{M}_j) \tag{1}$$

with the following ingredients.

BPDS model probabilities

The *decision dependent model probabilities* $\pi_j(\mathbf{x})$ can differentially weight models j over the decision space of $\mathbf{x}$, incorporating any prior information relevant to model weighting based on past predictive model fit and decision outcomes, and now explicitly allowing for adjustments based on a currently considered decision $\mathbf{x}$. Dependence of $\pi_j(\mathbf{x})$ on $\mathbf{x}$ is simply fundamental and critical in our policy setting.

BPDS calibration functions

The $\alpha_j(\mathbf{y}|\mathbf{x})$ are *calibration functions* that define outcome-dependence of model weights over the outcome space of $\mathbf{y}$ for any chosen $\mathbf{x}$. This defines the opportunity to increase or decrease the model weights differentially over the outcome $\mathbf{y}$ space to address model-specific biases and preferences and address questions of model-specific calibration more generally. The BPDS mixture of eqn. (1) has the equivalent form

$$f(\mathbf{y}|\mathbf{x}) = \sum_{j=0:J} \tilde{\pi}_j(\mathbf{x}) f_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j) \tag{2}$$

where

$$f_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j) = \alpha_j(\mathbf{y}|\mathbf{x}) p_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)/a_j(\mathbf{x}) \quad \text{and} \quad \tilde{\pi}_j(\mathbf{x}) = k(\mathbf{x})\pi_j(\mathbf{x}) a_j(\mathbf{x}) \tag{3}$$

with normalizing terms $k(\mathbf{x})$ and $a_j(\mathbf{x})$ explicitly dependent on $\mathbf{x}$. This form shows how the calibration functions $\alpha_j(\cdot|\cdot)$ modify the initial mixture pdfs $p_j(\cdot|\cdot) \to f_j(\cdot|\cdot)$ with corresponding changes of mixture weights $\pi_j(\mathbf{x}) \to \tilde{\pi}_j(\mathbf{x})$.

Note that the choice of relevant calibration functions $\alpha_j(\mathbf{y}|\mathbf{x})$ will, in any given application, be partly dependent on characteristics of the model pdfs $p_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)$. In particular, the expectation of each $\alpha_j(\mathbf{y}|\mathbf{x})$ under $p_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)$ must be finite in order that eqn. (1) defines a valid BPDS density $f(\mathbf{y}|\mathbf{x})$. Unbounded score functions may sometimes apply, but this point supports the use of bounded scores in general.

Baseline mixture component

The model index $j = 0$ explicitly allows for a *baseline* model component $\mathcal{M}_0$ in the mixture pdf $f(\cdot|\cdot)$ that can, among other things, address the ever-present issue of "model set incompleteness" (Tallman and West, 2023, section 2.2.3). $\mathcal{M}_0$ can be chosen to produce a pdf $f_0(\cdot|\cdot)$ that is over-dispersed relative to the mixture of the initial J models, so supporting outcomes $\mathbf{y}$ that are unusual under the J models; the baseline is then a suitable "fall back" model for times when the other models are forecasting poorly.

Initial mixture

The special case with each $\alpha_j(\mathbf{y}|\mathbf{x}) = 1$ defines the *initial mixture* with no BPDS calibration. We use $p(\mathbf{y}|\mathbf{x})$ in notion, i.e., $p(\mathbf{y}|\mathbf{x}) = \sum_{j=1:J} \pi_j(\mathbf{x}) p_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)$

Special cases fix ideas. First, if $\pi_j(\mathbf{x}) = \pi_j$ with $\pi_0 = 0$ are model probabilities based on historical BMA analysis, and with $\alpha_j(\mathbf{y}|\mathbf{x}) = 1,$ then eqn. (1) specialises to BMA. Thus BMA analyses– with or without this decision dependence in model-specific forecasts– are very special cases of BPDS. Second, again with $\pi_j(\mathbf{x}) = \pi_j$, $\pi_0 = 0$ and $\alpha_j(\mathbf{y}|\mathbf{x}) = 1,$ the decision maker has the freedom to specify the initial mixture probabilities π_j in other ways than with BMA. This includes using historical performance defined by scoring of past forecast outcomes, justifying various approaches to goal-focused model weighting (e.g. Lavine et al., 2021; Loaiza-Maya et al., 2021, and references therein) as special cases of BPDS. Third, mixture BPS (McAlinn and West, 2019; Johnson and West, 2022) is a special case in which models are combined with outcome-dependent weights. In these settings, $\pi_j(\mathbf{x}) = \pi_j$ depends on past predictive performance, $p_j(\mathbf{y}|\mathbf{x}) = p_j(\mathbf{y})$ and $\alpha_j(\mathbf{y}|\mathbf{x}) = \alpha_j(\mathbf{y})$

define outcome-dependent modifications of model probabilities, but there is no decision context so no $\mathbf{x}-$dependence. BPDS critically recognizes that the foundational BPS theory allows explicit incorporation of decision goals– admitting the conditioning on $\mathbf{x}$ throughout all components of eqn. (1)– to extend the foregoing analyses.

With predictions of $(\mathbf{y}|\mathbf{x})$ based on eqn. (2), the Bayesian decision maker acts to identify the optimal decision $\mathbf{x}$ based on a chosen utility function $U(\mathbf{y}, \mathbf{x})$. This involves numerical optimization to maximize the implied expected utility $\bar{U}(\mathbf{x}) = E_f[U(\mathbf{y}, \mathbf{x})|\mathbf{x}]$ over the decision space of $\mathbf{x}$. The notation $E_f[\cdot|\cdot]$ here explicitly represents expectation with respect to the BPDS distribution, and we use $E_p[\cdot|\cdot]$ to denote expectation under the initial mixture.

2.2 Decision-dependent Scores for Calibration Functions

The key step in integrating decision outcomes into relative model weightings is to address the question of how each $\mathcal{M}_j$ would inform decisions if used alone. Given the predictive pdf $p_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)$ and a chosen, potentially model-specific utility function $u_j(\mathbf{y}, \mathbf{x})$, acting based only on $\mathcal{M}_j$ leads to the optimal decision $\mathbf{x}_j$ that maximizes $E_{p_j}[u_j(\mathbf{y}, \mathbf{x})|\mathbf{x}]$ over $\mathbf{x}$. The decision maker has access to this set of model recommendations and is interested in model combination to preferentially weight "good decision models" as well as models that generate good predictions. BPDS formalizes this with specified *score functions* $\mathbf{s}_j(\mathbf{y}, \mathbf{x}_j)$, each being a $k-$vector of utilities that can be chosen to reflect both predictive and decision goals. The use of multi-dimensional scores addresses multiple goals simultaneously.

The Bayesian decision-theoretic development of Tallman and West (2023) generates the resulting functional forms of the BPDS calibration functions as

$$\alpha_j(\mathbf{y}, \mathbf{x}) = \exp\{\boldsymbol{\tau}(\mathbf{x})'\mathbf{s}_j(\mathbf{y}, \mathbf{x}_j)\}, \;\; j = 0{:}J, \tag{4}$$

where $\boldsymbol{\tau}(\mathbf{x})$ is a $k-$vector with elements differentially weighting the multiple utility dimensions of the score vector. The reasoning and theory behind this key result is as follows.

The initial mixture $p(\mathbf{y}|\mathbf{x}) = \sum_{j=1:J} \pi_j(\mathbf{x}) p_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)$ is the $\mathbf{y}-$margin of the joint distribution $p(\mathbf{y}, \mathcal{M}_j|\mathbf{x}) = \pi_j(\mathbf{x}) p_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)$, $(j = 1{:}J)$. Under this initial distribution for any candidate decision $\mathbf{x}$, and with score vectors $\mathbf{s}_j(\mathbf{y}, \mathbf{x}_j)$ defined and evaluated at model-specific optimal decisions $\mathbf{x}_j$, the decision maker has *initial expected score* $\mathbf{m}_p(\mathbf{x}) = \sum_{j=0:J} \pi_j(\mathbf{x}) \mathbf{m}_{jp}(\mathbf{x})$ where $\mathbf{m}_{jp}(\mathbf{x}) = \int_{\mathbf{y}} \mathbf{s}_j(\mathbf{y}, \mathbf{x}_j) p_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j) d\mathbf{y}$. Treating $\mathbf{m}_p(\mathbf{x})$ as a benchmark to improve on in expectation, the BPDS theory enquires about distributions $f(\mathbf{y}, \mathcal{M}_j|\mathbf{x})$ that yield expected scores $\mathbf{m}_f(\mathbf{x}) \geq \mathbf{m}_p(\mathbf{x}) + \boldsymbol{\epsilon}(\mathbf{x})$ for some non-negative $k-$vector (with at least one positive entry) $\boldsymbol{\epsilon}(\mathbf{x})$; this may be chosen to depend on $\mathbf{x}$, or may be a specified constant "decision score improvement". Given $\mathbf{m}_f(\mathbf{x})$, the BPDS theory identifies a unique $f(\cdot, \cdot|\mathbf{x})$ that minimizes the Küllback-Leibler (KL) divergence of $p(\cdot, \cdot|\mathbf{x})$ from $f(\cdot, \cdot|\mathbf{x})$ and has expected score exactly $\mathbf{m}_f(\mathbf{x})$. The theory is that of *relaxed entropy tilting* (Tallman and West, 2022, 2023; West, 2023) and yields $f(\mathbf{y}, \mathcal{M}_j|\mathbf{x}) \propto \pi_j(\mathbf{x}) \alpha_j(\mathbf{y}, \mathbf{x}) p_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)$ with calibration function precisely as in eqn. (4). The *tilting vector* $\boldsymbol{\tau}(x)$ is implicitly defined by the vector of k *target score constraints* $E_f[\mathbf{s}_j(\mathbf{y}, \mathbf{x}_j)|\mathbf{x}] = \mathbf{m}_f(\mathbf{x})$.

BPDS takes the view that the initial mixture is based on past performance, and additional small changes in models and the way they are weighted based on their expected performance may lead to better future decisions. It asks the question as to whether there are perturbations of the mixture based on the initial model probabilities that can lead to improved scores. Consider a stylized example with $J = 2$ models which in the past have forecast equally well. Traditional model averaging methods focused only on past forecasting experience– and BMA in particular– would confer equal weights in the combination. If, however, the models have different expected scores $\mathbf{m}_{jp}(\mathbf{x})$, conferring slightly more weight on the model expected to lead to a higher score makes sense. The entropic (or exponential) tilting theory is general. It is, of course, possible to tilt the initial joint distribution to most targets, so long as they are technically achievable under the initial distribution; but an overly ambitious target score will result in a tilted joint distribution that is empirically unreasonable. Hence we emphasize the importance of selecting $\mathbf{m}_f(\mathbf{x})$ that represents a "small " improvement over the initial benchmark score $\mathbf{m}_p(\mathbf{x})$. This is bolstered by the assumption that the initial model probabilities reflect the empirical plausibility of models as well as any available information about historical predictive and decision performance. Further, as we exemplify in the case study later, aspects of the computational methodology for model fitting in the sequential time series setting naturally inform on, and allow monitoring of, relevant choices of target expected scores.

2.3 BPDS Summary

This section has outlined the main ideas underlying BPDS and the key ingredients of the theory and resulting technical machinery. Specifications of score and utility functions, initial model probabilities, and target scores are all required for implementation and are, of course, application specific. The following section develops full details in the context of the macroeconomic decision-making application. In terms of computation, BPDS requires the use of posterior simulation methods (i.e. draws from conditional predictive densities from each model are required) as well as numerical optimization methods (i.e. to find $\mathbf{x}_j$ or the overall optimal decision $\mathbf{x}$ under the final BPDS analysis).

3 BPDS for Optimal Monetary Policy Decisions

The choice of data and models is inspired by Furlanetto et al. (2019). We use quarterly macroeconomic and financial variables from 1973:Q1 to 2022:Q2 from the FRED-QD database maintained by the Federal Reserve Bank of St. Louis. The data set includes

GDP (the log of real GDP), prices (the log of the GDP deflator), the interest rate (the shadow rate[1] which we treat as the policy rate), investment (the ratio of real gross private domestic investment to GDP), stock prices (log of the S&P500) and the spread (the spread between BAA bonds and the Fed funds rate). Models are run over multiple years, and at the end of each quarter produce forecasts– full predictive distributions in terms of Monte Carlo samples– of outcomes of interest over the following $k = 8$ quarters, conditional on candidate settings of the decision vector which is taken as the trajectory of interest (shadow) rates over those quarters. Within-model decision analysis then delivers model-specific optimal decisions about these rates.

3.1 Models, Forecasts and Model-Specific Decisions

We consider $J = 2$ models: $\mathcal{M}_1$ is a three-variable monetary policy VAR involving GDP, prices, and the interest rate; $\mathcal{M}_2$ is the model of Furlanetto et al. (2019), a VAR with the same variables as $\mathcal{M}_1$ plus investment, stock prices, and the spread. Following the latter, we include 5 lags in the VARs. The two structural VARs are identified using the sign restrictions from Table 1 of Furlanetto et al. (2019). In $\mathcal{M}_1$ these restrictions define supply, demand and monetary policy shocks. In $\mathcal{M}_2$ investment and financial shocks are additionally identified. We condition on a given value of the policy rate and set the monetary policy to be the driving shock. We do this by imposing restrictions on the set of structural shocks underlying the conditional forecasts. Structural shocks other than the monetary policy shock have zero means. We use the asymmetric conjugate prior of Chan (2022), with the advantage that the marginal likelihoods for each can be easily calculated; prior hyperparameter choices are made to maximize the marginal likelihood as in this referenced paper. At each quarter, multi-step ahead predictions are based on simula-

[1]This is the Federal Funds rate when the latter is positive but can go negative when it is at the zero lower bound, taking into account unconventional monetary policy; see Wu and Xia (2016).

tions using the precision-based sampler of Chan et al. (2023). Details on the conditional forecast computations are summarized in Appendix B. In short, this generates $p_j(\mathbf{y}, |\mathbf{x})$ with zero-mean constraints on all shocks apart from the monetary policy variable. We do not restrict the variance (i.e., 'soft' restrictions) such that we also have uncertainty around the path of $\mathbf{x}$. This can be thought of as conditional commitment– we allow the possibility of $\mathbf{x}$ deviating from the proposed policy path with deviations informed by historical uncertainty around forecasts of the interest rate outcomes.

This is a model setting in which the decision variable is also modeled – it is an outcome variable. In the current quarter, $\mathbf{x} = (x_1, \ldots, x_k)'$ is the $k-$vector of interest rate values over the next k quarters, and $\mathbf{y} = (y_1, \ldots, y_k)'$ the corresponding $k-$vector of inflation rates. Whatever other variables are in the VAR model $\mathcal{M}_j$, our interest focuses on the implied $p_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)$ required for BPDS eqn. (1). This conditional predictive is used in decision analysis with the same utility function for each model, namely $u_j(\mathbf{y}, \mathbf{x}) = U(\mathbf{y}, \mathbf{x})$ given by

$$U(\mathbf{y}, \mathbf{x}) = -\sum_{h=1:k} \{\rho^{k-h}(y_h - y^*)^2 + (x_h - x_{h-1})^2\} \tag{5}$$

where $\rho \in (0, 1)$ is a discount factor. This is a conventional quadratic function that reflects the dual goals of inflation rate targeting and interest rate smoothing over the next $k = 8$ periods. The $\mathbf{y}$ terms relates an inflation targeting mandate of $y^* = 2\%$ over the longer run, while the $\mathbf{x}$ terms encourage relatively constrained changes in quarter-to-quarter interest rates (the latter being a "don't rock the boat" consideration, as large swings in interest rates can/will have otherwise unduly effects on the macro-economic system). The terms involving the discount factor ρ represent the fact that monetary policy works with a lag, so it is desirable to less heavily penalize deviations from the target at shorter horizons h. With our two-year horizon ($k = 8$ quarters) our example analysis below adopts $\rho = 0.95$. The model-specific optimal decision vector $\mathbf{x}_j$ then maximizes

$E_{p_j}[U(\mathbf{y},\mathbf{x})|\mathbf{x}]$ over $\mathbf{x}$. Again, this analysis is repeated each quarter over the period of time series of interest, producing rolling updates of the "currently optimal" projections for interest rates over the coming 8 quarters.

3.2 BPDS Model Specification

The BPDS framework requires specification by the decision maker of a relevant class of baseline pdfs $p_0(\mathbf{y}|\mathbf{x})$, the model-specific vector score functions $\mathbf{s}_j(\mathbf{y},\mathbf{x})$, the initial BPDS model probabilities $\pi_j(\mathbf{x})$ as functions of candidate decisions $\mathbf{x}$, and the target expected scores $\mathbf{m}_f(\mathbf{x})$ at any $\mathbf{x}$. These are discussed in turn. In addition to customizing BPDS to the specific application, this section highlights a number of methodological developments relevant to other applications, particularly in: (a) the linkages of the $\pi_j(\mathbf{x})$ to $\mathbf{x}$ that are relevant more generally when $\mathbf{x}$ is both an outcome to be forecast as well as a putative decision variable (highlighted in Section 3.2.3 below); and (b) the relevance of dependence structure among the elements of the vector score under the initial distribution $p(\mathbf{y},\mathcal{M}_j)$ (highlighted in Section 3.2.4 below).

3.2.1 Baseline Distribution

To complete the main BPDS pdf in eqn. (1) requires the baseline $p_0(\mathbf{y}|\mathbf{x})$. This is taken as a multivariate T distribution with 10 degrees of freedom, using the location from the initial mixture $p(\mathbf{y}|\mathbf{x})$ ignoring the baseline (i.e., with $\pi_0(\mathbf{x})=0$) and corresponding variance of that mixture inflated by 4. This defines a relevant, tractable $\mathcal{M}_0$ that can capture outcomes $\mathbf{y}$ that the two VAR models are not predicting well for any $\mathbf{x}$ under consideration, and signal that to the decision maker.

3.2.2 BPDS Score Functions

The considerations of inflation targeting and interest rate smoothing reflected in the model-specific decision analysis in Section 3.1 are relevant to the choices of BPDS score functions. Our example takes $\mathbf{s}_j(\mathbf{y}, \mathbf{x}) = [s_{j1}(y_1, x_1), \ldots, s_{jk}(y_k, x_k)]'$ with elements

$$s_{jh}(y_h, x_h) = \exp\{-(y_h - y^*)^2/(2z_y^2)\} + \exp\{-(x_h - x_{h-1})^2/(2z_x^2)\}, \quad h = 1{:}k, \tag{6}$$

where $y^* = 2\%$ is the inflation target and z_y and z_x are *score bandwidth* parameters. This defines a class of bounded score functions, always relevant in decision analysis and here ensuring that the entropically tilted BPDS pdf of eqn. (3) is always integrable. The score bandwidths are set so that a certain deviation $d_y = (y - y^*)^2$ has a score of ε; given a choice of ε we set $z_y = d_y/\sqrt{-2\log(\varepsilon)}$. Similar considerations apply to choosing z_x. Our analyses use $\varepsilon = 0.4$, $d_y = 2$, and $d_x = 1$ to ensure the score function is dispersed enough to accommodate modest changes in the Federal Funds rate while is more lenient in deviations from the inflation target. Obvious modifications could incorporate horizon $h-$specific inflation targets and differentially weight the two exponential terms, but this form suffices for our main goals in this paper. Note also that, if inflation deviations from target and interest rate changes are "small", then $s_{jh}(y_h, x_h)$ is approximately quadratic in $|y_h - y^*|$ and $|x_h - x_{h-1}|$ for all h, a perhaps more familiar utility form.

3.2.3 Initial Model Probabilities

For clarity in the presentation in this section, we now make explicit the dependency on time, so that the ingredients of the full BPDS predictive pdf in eqn. (1)– with the exponential form of the calibration function of eqn. (4)– are now indexed by current time t, i.e.,

$$f_t(\mathbf{y}_t|\mathbf{x}_t) \propto \sum_{j=0:J} \pi_{tj}(\mathbf{x}_t) \mathrm{e}^{\boldsymbol{\tau}_t(\mathbf{x}_t)'\mathbf{s}_{tj}(\mathbf{y}_t,\mathbf{x}_{tj})} p_{tj}(\mathbf{y}_t|\mathbf{x}_t,\mathcal{M}_j).$$

Bayesian model weighting based on historical predictive performance with respect to define forecast goals, as developed in (Lavine et al., 2021), provides the starting point for specification of the $\pi_{tj}(\mathbf{x}_t)$. The general form adopted is

$$\pi_{tj}(\mathbf{x}_t) \propto \pi_{tj} p_{tj}(\mathbf{x}_t|\mathcal{M}_j), \qquad j = 0{:}J, \tag{7}$$

subject to summing to 1 over $j = 0{:}J$ and with ingredients as follows.

Initial model probabilities

Traditional Bayesian analysis (e.g. West and Harrison, 1997, chapter 12) defines the starting point. Here the time t initial model probabilities are based on standard sequential Bayesian updating from those at $t-1$: That is, $\pi_{tj} \propto \pi_{t-1,j} p_{tj}(\mathbf{z}_{t-1,j}|\mathcal{M}_j)$ where the "marginal model likelihood" term $p_{tj}(\mathbf{z}_{t-1,j}|\mathcal{M}_j)$ is the value of the 1-step ahead predictive pdf under $\mathcal{M}_j$ at the observed values of the last period outcomes $\mathbf{z}_{t-1,j}$ under that model. In our applied setting, this $\mathbf{z}_{t-1,j}$ includes time $t-1$ outcome values of inflation (y), interest rate (x), and other economic indicators modelled and forecast in $\mathcal{M}_j$ in our setting. In general, these can differ across models, but in consideration for the initial weights we restrict to variables common across models.

Then, BPDS allows the decision maker freedom to make alternative choices of the π_{tj}, and the goal and decision focus recommends modification of the standard BMA choice. BMA, after all, only reweights models based on 1-step ahead predictive accuracy. Hence we adopt two modifications, based on recent literature consonant with the goal foci.

First, we use simple power discounting of historically accrued support across models, in which the time $t-1$ to time t evolution is reflected in $\pi_{tj} \propto \pi_{t-1,j}^{\gamma} p_{tj}(\mathbf{z}_{t-1,j}|\mathcal{M}_j)$ where

γ is a discount factor in $(0, 1]$, closer to 1 for most applications. This acts to discount historically accrued support for model j at a per-time unit discount rate γ prior to updating by the time $t-1$ information. Going back at least to Smith (1979) and then, in a formal dynamic model uncertainty context, West and Harrison (1989, chapter 12, p.445), power-discounting has been shown to be of value in empirical studies in implicitly allowing for time-variation in the predictive relevance of different models (e.g. Raftery et al., 2010; Koop and Korobilis, 2013; Zhao et al., 2016). Our case study below uses $\gamma = 0.95$.

Second, reflecting the foci on specific predictive and decision goals, initial model probabilities should also generally be modified based on the recent relative performance of models with respect to the defined goals. This is the premise underlying the specific variants of BPS in the setting of adaptive variable selection, or BPS-AVS, in Lavine et al. (2021), and related developments in Loaiza-Maya et al. (2021), for example. This leads to the immediate BPDS extension of these prior approaches in which the above reasoning is extended to define

$$\pi_{tj} \propto \pi_{t-1,j}^{\gamma} p_{tj}(\mathbf{z}_{t-1,j}|\mathcal{M}_j) \mathrm{e}^{\boldsymbol{\tau}_{t-1}(\mathbf{x}_{t-1})'\mathbf{s}_{t-1,j}(\mathbf{y}_{t-1},\mathbf{x}_{t-1,j})}.$$

Here the discounted Bayesian model probabilities are further updated with AVS-style weights based on the the realized BPDS calibration function based on relative model scores based on the actual decision outcomes at the last time period. As a result, models are initially and naturally reweighted based on both predictive and decision outcome performance at the last time period.

In our applied setting this means that, as a result, models achieving "good" recent trajectories of interest rate smoothness, as well as relatively accurate forecasting performance of realized inflation outcomes, will be rewarded with higher initial BPDS model probabilities in looking forward to the next time point. And we note that the specification here can cut back to define special cases including BPS-AVS (by setting $\boldsymbol{\tau}_{t-1}(\mathbf{x}_{t-1}) = \mathbf{0}$), and within that to traditional BMA (by setting $\gamma = 1$) for comparisons.

Finally, the inclusion in BPDS of the baseline model and its forecast densities leads to a modification of these initial model probabilities to provide a non-zero value π_{t0} for the baseline. We choose a fixed probability– in our analysis $\pi_{t0} = 0.1$ at each time point t– and simply renormalize the π_{tj} above over $j = 1{:}J$ accordingly.

Informative Conditioning on $\mathbf{x}_t$

As noted earlier, in our setting the future values of decision variables are also considered outcomes predicted under the models. The models each forecast the future evolution of interest rates as part of the complex, dynamic macroeconomic system, whereas for decisions we must condition on $\mathbf{x}_t$. This is reflected in the conditional (on $\mathbf{x}_t$) distributions $p_{tj}(\mathbf{y}_t|\mathbf{x}_t, \mathcal{M}_j)$ in BPDS where $\mathbf{x}_t$ is treated as known. The theoretical implication for the BPDS model probabilities is the term $p_{tj}(\mathbf{x}_t|\mathcal{M}_j)$ in eqn. (7)– this is the value of the current marginal predictive pdf of the vector $\mathbf{x}_t$ under $\mathcal{M}_j$. Assuming the prior (to time t) probabilities π_{tj} are specified, this form arises directly via Bayes' theorem; the act of conditioning on $\mathbf{x}_t$ is informative, and the implied update is, simply by Bayes' theorem, that in eqn. (7). Critically, this implies that candidate decision values that are not well-supported under the joint distribution of a model are down-weighted. Conversely, at any candidate decision vector $\mathbf{x}_t$, models that are more predictively supportive of the decision $\mathbf{x}_t$ will be relatively rewarded with higher values of resulting $\pi_{tj}(\mathbf{x}_t)$.

In other applications of BPDS the decision variables may be exogenous, i.e., control variables that are to be chosen by the decision maker but that are not forecast jointly with $\mathbf{y}_t$ in the set of models. In such cases, it will be common to assume that the external choice of $\mathbf{x}_t$ is not informative, and then eqn. (7) results in decision-independent BPDS probabilities $\pi_{tj}(\mathbf{x}_t) = \pi_{tj}$ based only on historical data and information.

3.2.4 BPDS Target Scores

The BPDS target expected score $\mathbf{m}_f(\mathbf{x}) = E_f[\mathbf{s}(\mathbf{y},\mathbf{x})]$ represents a desired improvement over the initial expected score $\mathbf{m}_p(\mathbf{x}) = E_p[\mathbf{s}(\mathbf{y},\mathbf{x})]$. In the multi-objective case the resulting $\boldsymbol{\tau}(\mathbf{x})$ that defines $f(\mathbf{y}|\mathbf{x})$ to satisfy this target expectation is sensitive to both the relative scales and dependence of elements of $\mathbf{s}(\mathbf{y},\mathbf{x})$ under the initial mixture $\mathbf{y} \sim p(\mathbf{y}|\mathbf{x})$ at any candidate decision $\mathbf{x}$. As functions of $\mathbf{y}$ the elements of the random score vector $\mathbf{s}(\mathbf{y},\mathbf{x})$ can be strongly correlated, leading to challenges in specifying relevant targets. This can also complicate the calculation of the implied BPDS tilting vector $\boldsymbol{\tau}(\mathbf{x})$– the vector that is needed to satisfy $\mathbf{m}_f(\mathbf{x}) = E_f[\mathbf{s}(\mathbf{y},\mathbf{x})]$ under the BPDS density of eqn. (1). We address this by explicitly recognizing score dependencies and defining an approach that explicitly incorporates dependence.

Some theoretical intuition is gained by considered cases of "small perturbations" in which $\mathbf{m}_f(\mathbf{x})-\mathbf{m}_p(\mathbf{x})$ has small elements. In this setting, entropic tilting theory in Tallman and West (2022) yields the second-order approximation $\boldsymbol{\tau}(\mathbf{x}) \approx \mathbf{V}_p(\mathbf{x})^{-1}(\mathbf{m}_f(\mathbf{x})-\mathbf{m}_p(\mathbf{x}))$ where $\mathbf{V}_p(\mathbf{x})$ is the variance matrix of $\mathbf{s}(\mathbf{y},\mathbf{x})$ under the initial mixture $p(\mathbf{y}|\mathbf{x})$. This shows that the implied tilting vector will be very sensitive to the initial score scales and dependencies as reflected in $\mathbf{V}_p(\mathbf{x})$, and suggests a prime focus on a *standardized score* scale. That is, define $\mathbf{C}_p(\mathbf{x})$ as the scaled eigenvector matrix such that $\mathbf{V}_p(\mathbf{x}) = \mathbf{C}_p(\mathbf{x})\mathbf{C}_p(\mathbf{x})'$ and set the target score using $\mathbf{m}_f(\mathbf{x}) = \mathbf{m}_p(\mathbf{x}) + \mathbf{C}_p(\mathbf{x})\boldsymbol{\epsilon}(\mathbf{x})$ for a specified *standardized expected score vector* $\boldsymbol{\epsilon}(\mathbf{x})$. The usual convention is taken in which the eigenvector columns of $\mathbf{C}_p(\mathbf{x})$ are ordered according to decreasingly values of the corresponding eigenvalues, so that the first column is "dominant", and so forth. This provides insights into how to practically define target scores related to the absolute standardized scale. As examples of the two extremes, taking $\boldsymbol{\epsilon}(\mathbf{x}) = \epsilon(\mathbf{x})\mathbf{1}$ for some scalar $\epsilon(\mathbf{x})$ represents targets deviating from the initial expected score in equal amounts of $\epsilon(\mathbf{x})$ along each of the standardized eigendimensions. At the other extreme, and most relevant when there are strong score dependencies, taking $\boldsymbol{\epsilon}(\mathbf{x}) = (\epsilon(\mathbf{x}), 0, \ldots, 0)'$ defines the resulting target $\mathbf{m}_f(\mathbf{x})$ based on the major, dominant eigen-dimension alone. The latter is a starting point in general and is taken to define our BPDS case study below. In that setting, we choose $\epsilon(\mathbf{x})$ such that $\min\{(\mathbf{m}_f(\mathbf{x})/\mathbf{m}_p(\mathbf{x}))\} = 0.75$ to define the maximum expected improved score in any dimension.[2] It is obviously straightforward to extend this methodology to define target scores impacted by higher eigen-dimensions, though that is left for future applications.

3.3 BPDS Implementation and Optimal Decisions

The final step in BPDS is to couple the decision maker's utility function with the BPDS predictive equations (1,2) to define the optimal decisions from the model synthesis. The decision maker can adopt any utility function but an initial neutral analysis will be based on using the same form as usual in the model-specific decisions, the function $U(\mathbf{y}, \mathbf{x})$ of eqn. (5). This is used in the example analysis to follow, with the aim of computing $\mathbf{x}$ to maximize the implied expected utility function $\bar{U}(\mathbf{x})$. In the case study analysis we compare decisions recommended by BPDS to those from each of the models and to a traditional BMA-based analysis. On the latter, the BMA mixture uses model weights proportional to the marginal likelihoods of the data which is common to all of the models (which includes inflation, interest rate, and GDP) under each $\mathcal{M}_j$. The BMA mixture naturally involves only the $p_j(\mathbf{y}|\mathbf{x})$ with no BPS/BPDS outcome-dependence, no notion of a baseline model to address model-set incompleteness, and no regard for the decision-focused use of the models. The foundational BPDS framework theoretically allows for these critical considerations as fundamental to the broader subjective Bayesian decision-analytic and goal-focused approach. Then, technically, BMA arises as a special case of the

[2]Additionally, due to the arbitrariness of signs of eigenvectors, we apply a ± 1 multiplier to the first column of $\mathbf{C}_p(\mathbf{x})$ so that the sum of elements are positive, ensuring the target score improves upon $\mathbf{m}_p(\mathbf{x})$.

BPDS analysis as earlier discussed throughout Section 3.

The computation of BPDS involves two key components. First, is the overall optimization over $\mathbf{x}$ which explores potential BPDS decisions and finds the optimizing vector $\mathbf{x}$. This requires an "outer loop" numerical optimization to explore $\mathbf{x}$ space. In our study, this is performed using a trust region method, namely Powell's Derivative Free Optimization Solvers (PDFO–Ragonneau and Zhang, 2023). Due to the possibility of multi-modality in $\bar{U}(\mathbf{x})$, the optimization is run repeatedly (in parallel) from multiple starting values. In some periods over time, we do find evidence of multi-modality, so repeat starts of the optimization routine are mandated. Second, within each evaluation of a potential BPDS decision, it is necessary to compute the tilting vector $\boldsymbol{\tau}(\mathbf{x})$ given the constraint $E_f[\mathbf{s}(\mathbf{y}, \mathbf{x})] = \mathbf{m}_f(\mathbf{x})$ for a target expected score $\mathbf{m}_f(\mathbf{x})$.. The theoretical basis of this is an implicit equation that is solved via standard, generic numerical optimization methods. Relevant details follow Tallman and West (2023, section 4.4) and are summarized in our Appendix A.

The BPDS forecast distributions are evaluated using importance sampling. At any given $\mathbf{x}$, the BPDS predictive distribution in eqn. (1) is simulated by sampling from the $p_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)$ in proportions defined by the BPDS probabilities $\pi_j(\mathbf{x})$; then the resulting importance sampling weights are simply proportional to the realized values of $\alpha_j(\mathbf{y}|\mathbf{x})$. This provides for efficient computation as well as access to traditional methods and metrics– such as importance sampling effective sample sizes (ESS, e.g. Gruber and West, 2016, 2017, in related contexts)– to monitor and evaluate the quality of the resulting Monte Carlo approximations to resulting predictive expectations. Note that this can deliver such metrics to assess "concordance" between the initial densities $p_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)$ and their corresponding BPDS-tilted versions $f_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)$ in eqn. (2), as well as that of the initial mixture $p(\mathbf{y}|\mathbf{x})$ and the resulting $f(\mathbf{y}|\mathbf{x})$. More aggressive BPDS target scores will generally lead to lower concordance, and choices can be partly guided by such empirical evaluations.

4 Case Study

4.1 Overview

The analysis and all empirical results proceed sequentially on an expanding window of data beginning in 1992Q2. Our summaries begin with a comparison of the decisions recommended by BPDS to those suggested by BMA. This is followed by a discussion of the individual models and how they are combined by BPDS and BMA. Additional discussion highlights some operational BPDS details to provide further insights into the resulting decision outcomes.

4.2 Optimal Decisions

Figure 1 shows the actual policy rate each quarter along with the $1{-}8$ quarters ahead policy recommendations that would have been made by BPDS and BMA. In using the shadow rate, the zero lower bound is not in effect and negative values for the policy rate are possible. Recommendations for negative values for the policy rate are not to be taken literally as advising cuts to a negative Federal Funds rate, but rather a suggestion to undertake other forms of monetary easing that would be expected to proxy such cuts.

Since 2014 the optimal policy paths recommended by the two approaches are generally similar, though there are notable differences prior to that time. Some specific periods of interest are now highlighted.

2014 to the Present

During this period, BMA and BPDS provide similar recommendations that are often quite different from the actual policy rate. For almost all of these times, the policy recommendations are to cut interest rates whereas (apart from 2019-2021) the actual policy rate increased. Some differences do arise between BMA and BPDS, for example, during the post-COVID inflation BPDS recommends a higher rate path. BPDS is closer to the decision actually made by the Federal Reserve, although according to BPDS rate cuts should have begun coming down by now.

The Financial Crisis and Subsequent Recession

It is during this period that the differences between BPDS and BMA are most acute. The actual policy rate fell slowly during this period. BPDS recommends rate cuts as well, initially at a more rapid rate than what actually occurred, but as of 2010, its recommendations are similar to the ones the policy makers actually made. In contrast, BMA recommends huge cuts to the policy rate right at the start of the financial crisis but subsequently consistently argues for rate increases.

The First Years of the 21st Century

In the period from 2003 through the beginning of the financial crisis, the actual policy rate was gradually increasing. In this period, BMA consistently recommended rapid rate increases. In contrast, BPDS recommendations were generally similar to what actually transpired, apart from at the beginning of this period where the advice was to raise the policy rate more slowly than occurred.

The 1990s

During this period, the pattern is more mixed. Optimal policy recommendations generated under each of BMA and BPDS often differed from actual decisions, with no consistent pattern; at times the recommended rates were higher than the actual policy rate, and at other times they were lower.

A general pattern, one that occurs throughout the sample period, is that BMA and BPDS typically recommend larger changes in policy rates than were actually implemented by policy makers. Part of this is presumably due to differences between the policy makers utility function and those used in our analyses. Other possible explanations are that policy makers can affect expectations through their communications which is a channel not captured in the model or their models use a much steeper Phillips Curve. Also, we focus only on inflation up to 2 years ahead, without considering the possibility of an over-

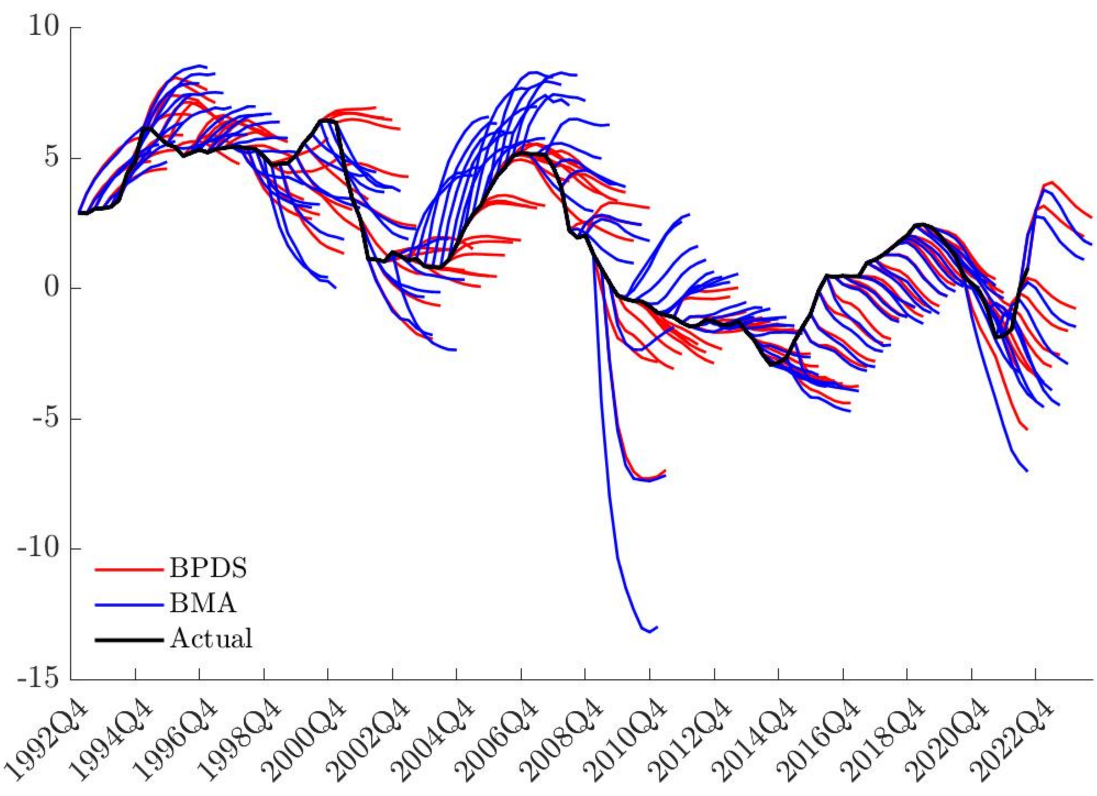

Figure 1: Recursively calculated policy decisions

or under-shoot of inflation after 8 quarters while policy makers would generally desire longer-term targets sustainably.

4.3 Trajectories of BPDS and BMA Predictive Densities

Figures 2–4 shed light on these patterns. These images represent the time trajectories of predictive densities of inflation at each relevant horizon, using BPDS (Fig. 2) and BMA (Fig. 3), as well as their differences (Fig. 4). These indicate that the BPDS mixtures are less dispersed than the BMA mixture for much of the sample period, i.e., BMA predictive distributions are relatively more heavy-tailed, especially at longer horizons. This is partly due to the BPDS score function emphasizing that the policy maker wants to avoid extreme inflation outcomes, and also accounts for why BMA often tells policy makers to make larger changes to the policy rate than BPDS as discussed in Section 4.2. The differences between BPDS and BMA become larger at longer forecast horizons. Medium and long-term macroeconomic forecasting is difficult which leads to standard methods such as BMA producing fairly dispersed predictive densities at longer horizons. BPDS, on the other hand, is reducing this effect, which dampens the BPDS optimal decisions and reduces predictive uncertainty relative to BMA. Then, differences between BPDS and BMA forecast densities are reduced after the financial crisis, which helps account for why their policy recommendations are similar in the last decade of the sample.

4.4 Model Probabilities

Figure 5 shows the trajectories of model probabilities under BPDS and BMA. These are the discounted AVS prior model probabilities π_{tj} over time, the implied initial decision-dependent probabilities $\pi_{tj}(\mathbf{x}_t)$ evaluated at the BPDS-optimal decision $\mathbf{x}_t$ at each time, the resulting BPDS probabilities $\tilde{\pi}_{tj}(\mathbf{x}_t)$ of eqn. (3) at each time, and the standard BMA

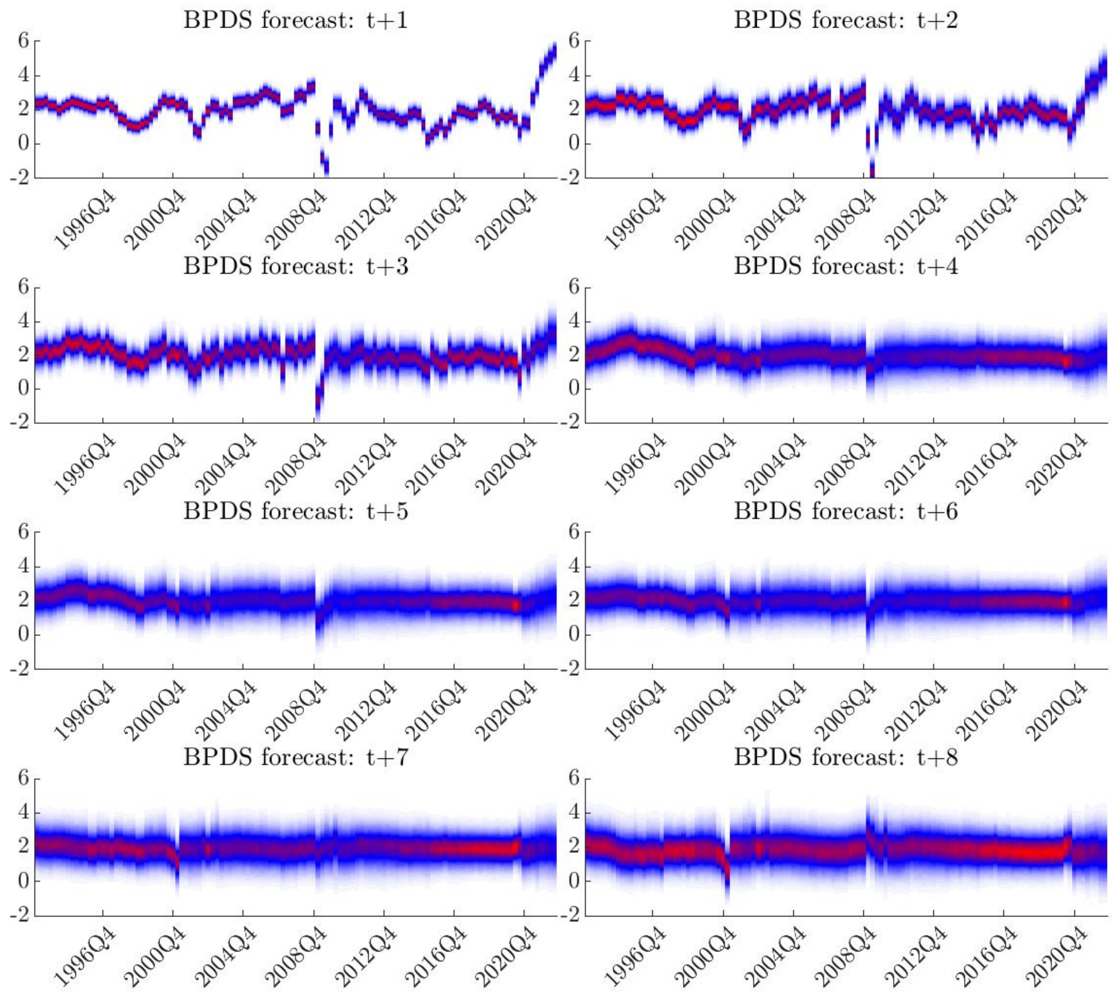

Figure 2: BPDS forecast densities of inflation. The frames represent 1–8 quarter ahead forecasts, reading along the rows from top-left to bottom-right. The colors represent probabilities with the shading from blue, being lower probability, to red being higher probability

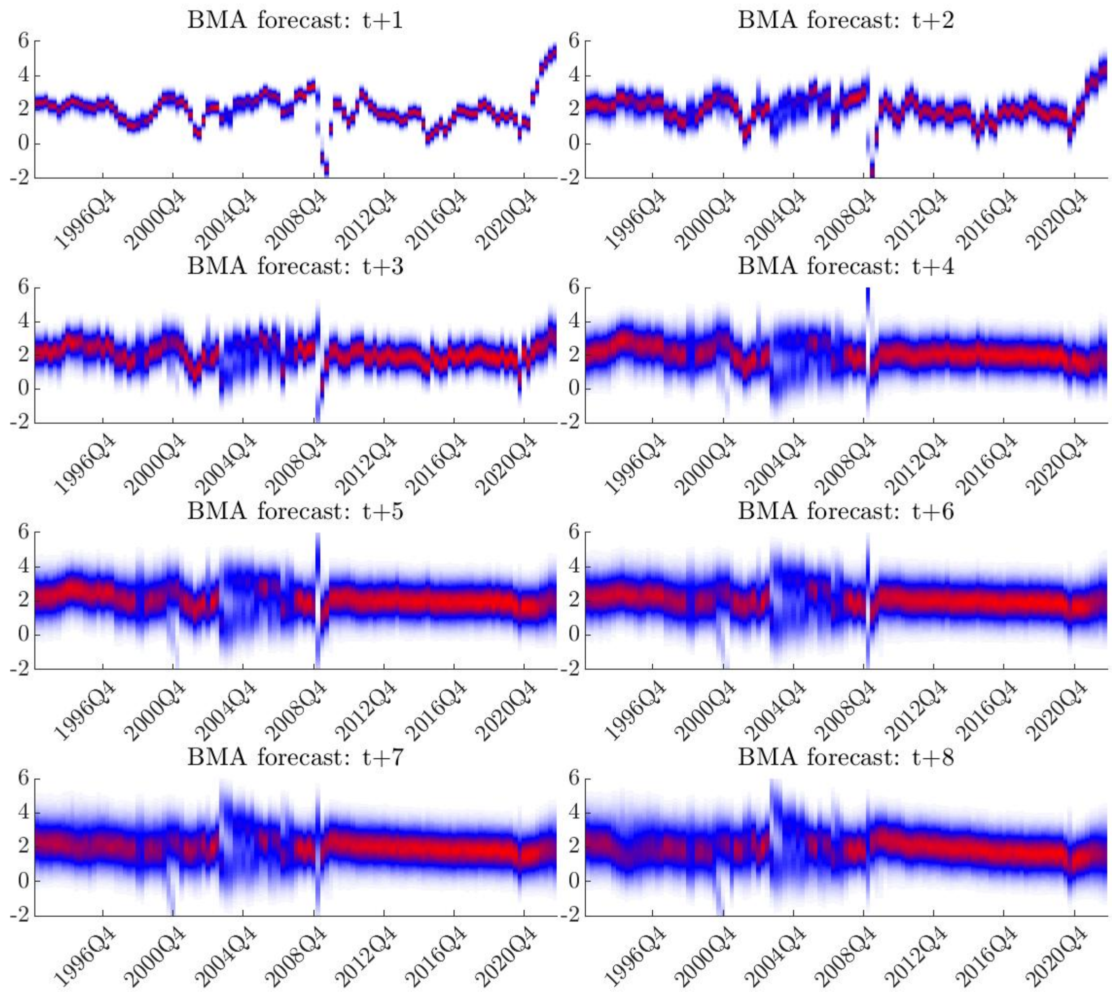

Figure 3: BMA forecast densities of inflation. The frames represent 1–8 quarter ahead forecasts, reading along the rows from top-left to bottom-right. The colors represent probabilities with the shading from blue, being lower probability, to red being higher probability

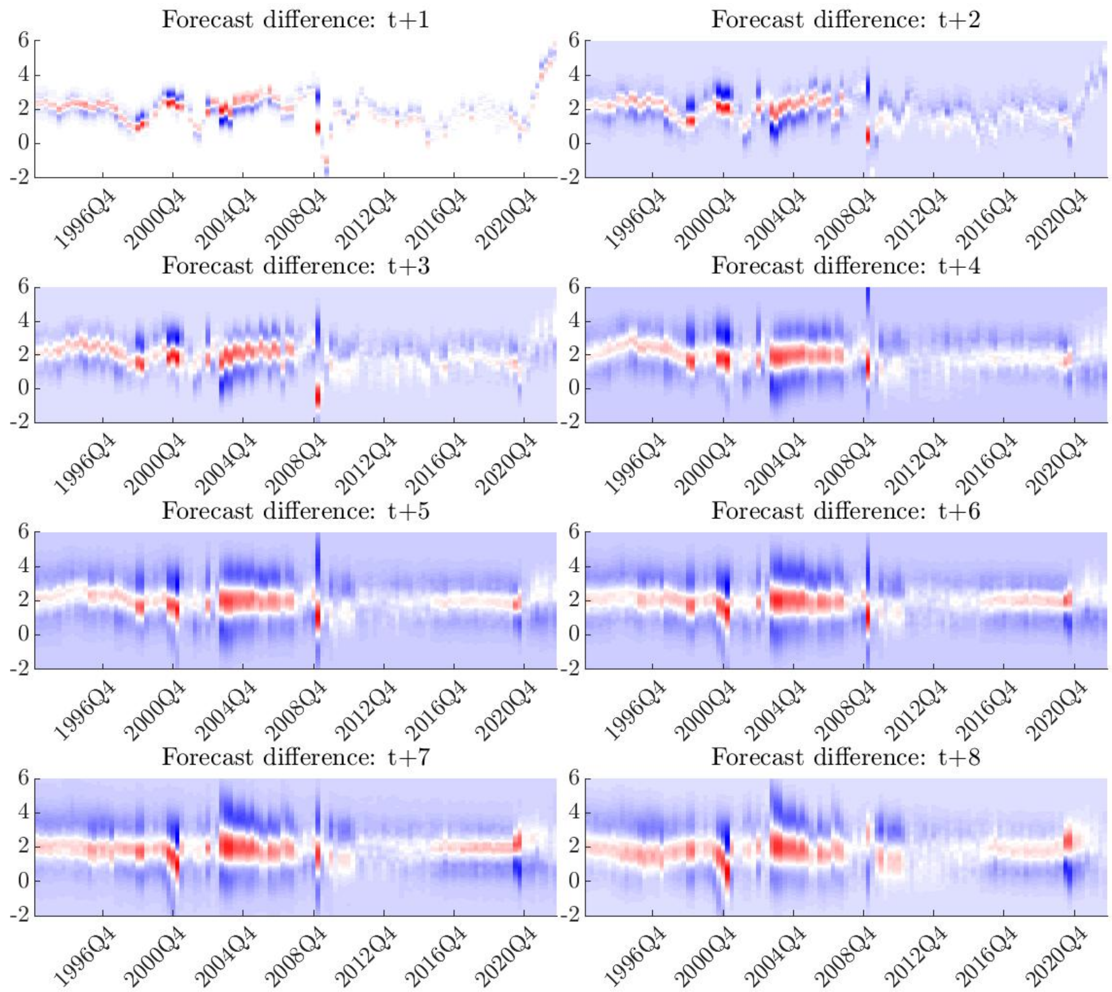

Figure 4: Difference between BPDS and BMA forecast densities of inflation. Red-shaded regions have higher probability under BPDS than under BMA, with blue shading indicating the reverse, and white shading equal probability

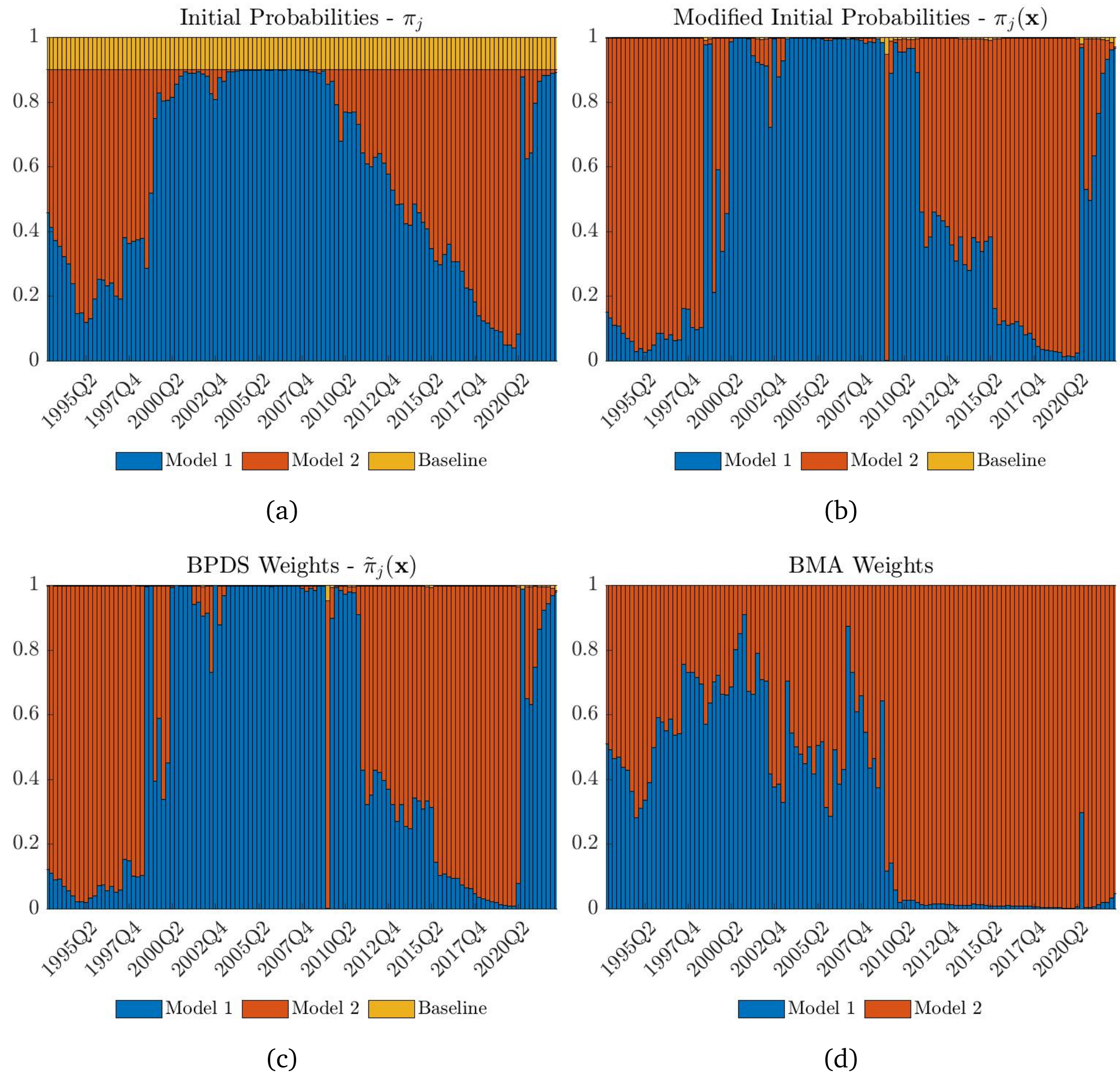

Figure 5: Time trajectories of model probabilities. (a) Prior BPDS probabilities π_{tj} based on discounted AVS with a fixed baseline $\pi_{t0} = 0.1$; (b) BPDS decision-dependent initial probabilities $\pi_{tj}(\mathbf{x}_t)$; (c) Implied BPDS weights $\tilde{\pi}_{tj}(\mathbf{x}_t)$; (d) BMA probabilities.

probabilities over time.

Under traditional BMA, the two model probabilities are appreciable until the financial crisis. After the start of the crisis, the less parsimonious $\mathcal{M}_2$– that includes additional financial variables– receives virtually all the weight. In contrast, BPDS weights vary more over time, allocating most of the weight to the parsimonious $\mathcal{M}_1$ for much of the period (i.e. 1997 through 2017), though $\mathcal{M}_2$ plays more of a role at both the beginning and end of the sample period. That BPDS generally favors the more parsimonious $\mathcal{M}_1$, with less dispersed forecast distributions, partially accounts for why BPDS often dampens extreme recommendations made when using BMA.

BPDS probabilities on the over-dispersed $\mathcal{M}_0$ are generally small, though with notable increases at two critical periods: the start of the financial crisis and the start of the COVID-19 pandemic. In such extreme times, when neither $\mathcal{M}_1$ or $\mathcal{M}_2$ forecasts well, the increased probability on the fall-back $\mathcal{M}_0$– though small– provides an indicator of this.

The BPDS prior model probabilities π_{tj} based on discounted AVS differ noticeably from BMA probabilities (except at the very start of the time period). A big impact then arises from the conditioning on information generated in the decision space to map these π_{tj} to the decision-dependent weights $\pi_{tj}(\mathbf{x}_t)$ at the BPDS optimal decisions $\mathbf{x}_t$ at each time. Recall that this mapping theoretically properly takes into account the likelihood of future, as yet unobserved, interest rate outcomes; this relevant information is not accounted for in the prior weights π_{tj} and is, of course, absent under BMA. The subsequent map from initial probabilities $\pi_{tj}(\mathbf{x}_t)$ to the BPDS weights $\tilde{\pi}_{tj}(\mathbf{x}_t)$ is wholly based on the impact of the entropic tilting towards "more favorable" decisions, in expectation. We see that the impact is rather small over time, and this is to be expected: the BPDS analysis uses "small" perturbations of the initial mixture based on target expected scores that are only modest increases over those under the initial mixture. We expect to see slight tilting towards models that are expected to do well, but not large changes relative to the initial probabilities.

4.5 Additional Insights from BPDS Results

Time trajectories of the evaluated tilting vectors $\boldsymbol{\tau}_t(\mathbf{x}_t)$, evaluated at the optimal decisions $\mathbf{x}_t$, are shown in Figure 6. The values generally tend to increase with horizon h, thus attaching more weight to longer forecasting horizons. This is partly to be expected due to the higher uncertainties at longer forecast horizons.

Figure 7 plots trajectories of several effective sample size (ESS) measures arising from the importance sampling to simulate BPDS predictive distributions as discussed in Section 3.3. This provides a read-out of the extent of tilting the initial mixture $p(\mathbf{y}|\mathbf{x})$ to the BPDS mixture $f(\mathbf{y}|\mathbf{x})$ as well as that for tilting each of the individual model pdfs from $p_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)$ to $f_j(\mathbf{y}|\mathbf{x}, \mathcal{M}_j)$ (again, time-indexed and updated throughout the time series). Until the COVID recession, the ESS of the initial mixture is stable between 90-

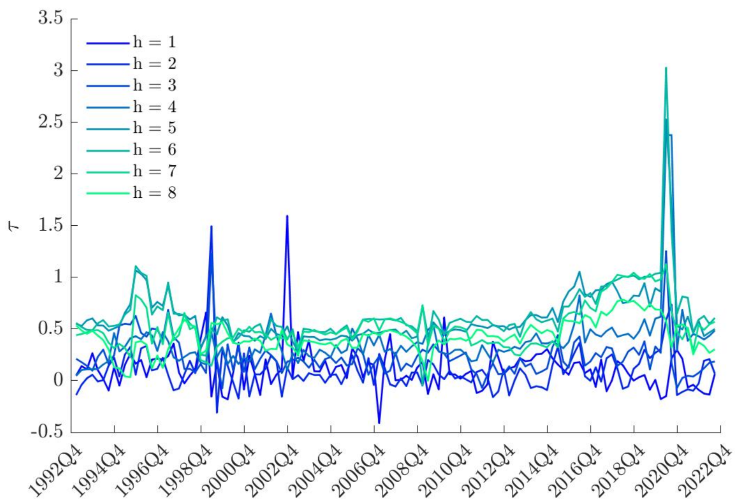

Figure 6: Trajectories of the 8 elements of the evaluated BPDS tilting vector $\boldsymbol{\tau}_t(\mathbf{x}_t)$ at the optimized $\mathbf{x}_t$ at each quarter

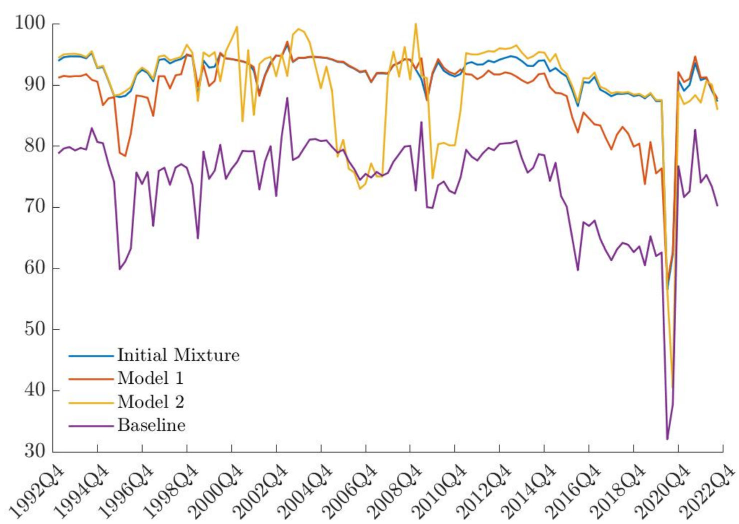

Figure 7: Trajectories of the effective sample size (ESS) metrics for individual models and for the BPDS initial mixture

95% suggesting only a small amount of tilting, as desired. The COVID recession is a period of rapid change as expected, as we see large changes in the initial weights and larger values of $\boldsymbol{\tau}$ required to achieve the desired target. The low value of ESS indicates that at that point the target expected scores are unrealistic given the then current state of the economy. However, the resulting decisions during this time period appear to be rather sensible, meaning we do not need to be too concerned about the low ESS which can, in any case, be redressed by simply increasing the overall Monte Carlo sample size accordingly. The ESS values of individual models are generally lower than that of the overall mixture and somewhat more volatile. One nice point is that, even when one of the models seems to suffer a low ESS, the BPDS mixture ESS is generally maintained at higher values. This indicates that BPDS is able to strike a balance in weighting expected versus historical performance of models on both predictive and decision outcomes.

Finally, Figure 8 compares the realized trajectories of expected utilities under BPDS and BMA. Each uses the same utility function to define the final optimal policy path decision, so these are directly comparable, and the comparison is relevant in terms of the setting of forward, sequential decisions where a change to much lower values at any time point should signal concern to the decision-maker. BPDS is designed to target an expected utility higher than that of the initial mixture, but whether it achieves a higher expected utility than BMA– which has different initial probabilities and lacks outcome-dependent weighting– is a question for empirical study. In this example, as illustrated in the figure, BPDS utility does exceed that of BMA in virtually every period. After the financial crisis, the two are similar, consistent with the earlier finding that they typically produced similar decisions during this time period. However, before the financial crisis, there were several periods during which the BPDS expected utilities were substantially higher than those of BMA. These correspond to times where we see more differences between optimal policy path recommendations, and within which there are some periods of greater concordance

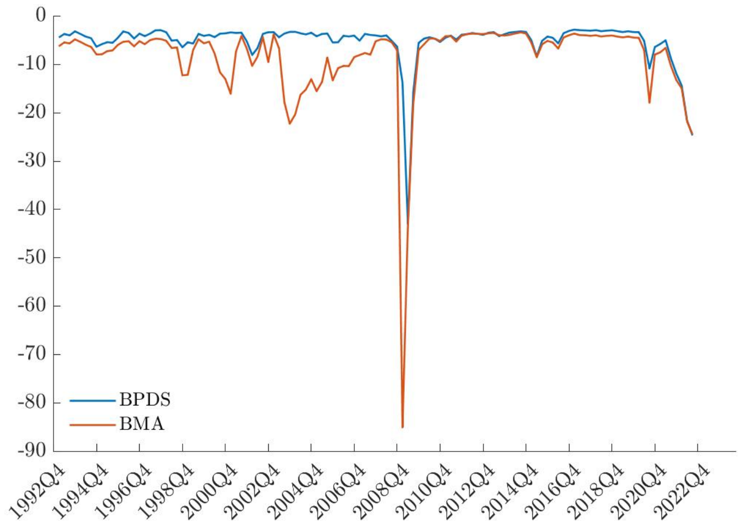

Figure 8: Trajectories of expected utilities comparing BPDS with BMA

between BPDS and actual policy decisions as well as more constrained (i.e., less extreme) recommended decisions under BPDS relative to BMA.

5 Summary Comments

BPDS is the formal, foundational Bayesian framework that extends traditional Bayesian model uncertainty analysis to address explicit use of model-specific decision outcomes as well as purely predictive performance in model comparison and combination. This paper has adapted the BPDS foundations to define implied methodology in formulating macroeconomic decision-making when faced with multiple objectives and multiple outcomes of interest in the monetary policy setting.

Earlier applications of BPDS have focused mainly on financial portfolio forecasting and decisions (Tallman and West, 2023, section 6; Tallman and West, 2024), a setting in which forecasting models do not (generally) depend on the decisions of interest, while utility functions may and often do depend on the models and their predictions. In contrast, the setting of monetary policy analysis is one in which the dependence of models and their forecasts on the decision variables (policy instruments) is simply fundamental.

It is also a setting in which the decision variables are treated simultaneously as outcomes. The future paths of central bank interest rates, for example, are modelled as time series outcomes along with other economic and financial indicators in VAR models. This then leads to conditioning on decision variables to define predictions of other indicators, with consequent implications for relative model weights in the model uncertainty setting. This latter point is critical as it then leads to relatively up- or down-weighting a model based on how well-supported a particular candidate decision is under its predictions; to our knowledge, this is the first time this central question has been formally, statistically addressed. These central features of predictive decision-making in monetary policy contexts are addressed with extensions and customization of the existing theory of BPDS.

The BPDS perspective– of integrating historical and expected decision outcomes with focused aspects of statistical predictive performance into relative model weightings– is new to the policy arena. We argue for this perspective since policy makers are primarily interested in utilizing sets of models for the eventual policy decisions. Pure forecasting exercises– and evaluation and combinations of models for prediction *per se*– is, of course, of parallel interest and importance. We emphasize that BPDS also involves addressing predictive performance on specific, defined outcomes of interest. But most importantly, by putting the spotlight on decision-making, we gain additional insights into policy making that are not possible in exercises that focus solely on predictive performance.

In a recursive real-time decision-making exercise, we find substantial differences at various periods of time between the policy recommendations of BPDS and the traditional Bayesian model averaging approach, though good concordance at other times. When recommended policy decisions differ between the approaches, in most cases the BPDS policy paths are more intuitively sensible and less extreme than under BMA, and more consistent with the actual decisions made by the policy makers at the time. The case study presented investigates and interprets aspects of BPDS– in terms of differential model weights based on historical information alone and then updated based on identified optimal decisions, with consequent insights into how the differences relative to standard BMA arise and are exploited. This case study is a first step towards broader development and evaluation of BPDS in a setting with larger numbers of econometric models. The parallel next steps will naturally include BPDS for scenario forecasting– analyses in collaboration with policy-makers that explore and aim to understand the sensitivity of model-based recommendations relative to chosen potential economic scenarios.

References

Aastveit, K. A., J. L. Cross, and H. K. van Dijk (2023). Quantifying time-varying forecast uncertainty and risk for the real price of oil. *Journal of Business and Economic Statistics 41*, 523–537.

Bernaciak, D. and J. E. Griffin (2024). A loss discounting framework for model averaging and selection in time series models. *International Journal of Forecasting*. doi.org/10.1016/j.ijforecast.2024.03.001.

Chan, J. C. C. (2022). Asymmetric conjugate priors for large Bayesian VARs. *Quantitative Economics 13*, 1145–1169.

Chan, J. C. C., D. Pettenuzzo, A. Poon, and D. Zhu (2023). Conditional forecasts in large Bayesian VARs with multiple soft and hard constraints. *SSRN Electronic Journal*. doi.org/10.2139/ssrn.4358152.

Chernis, T. and T. Webley (2022). Nowcasting Canadian GDP with density combinations. Technical report, Bank of Canada. doi.org/10.34989/sdp-2022-12.

Conflitti, C., C. De Mol, and D. Giannone (2015). Optimal combination of survey forecasts. *International Journal of Forecasting 31*, 1096–1103.

Furlanetto, F., F. Ravazzolo, and S. Sarferaz (2019). Identification of financial factors in economic fluctuations. *Economic Journal 129*, 311–337.

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

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

Gruber, L. F. and M. West (2017). Bayesian forecasting and scalable multivariate volatility analysis using simultaneous graphical dynamic linear models. *Econometrics and Statistics 3,* 3–22.

Johnson, M. C. and M. West (2022). Bayesian predictive synthesis with outcome-dependent pools. *Technical Report, Department of Statistical Science, Duke University*. arXiv:1803.01984.

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

Koop, G. and D. Korobilis (2013). Large time-varying parameter VARs. *Journal of Econometrics 177*, 185–198.

Lavine, I., M. Lindon, and M. West (2021). Adaptive variable selection for sequential prediction in multivariate dynamic models. *Bayesian Analysis 16*, 1059–1083.

Loaiza-Maya, R., G. M. Martin, and D. T. Frazier (2021). Focused Bayesian prediction. *Journal of Applied Econometrics 36*, 517–543.

Martin, G. M., D. T. Frazier, W. Maneesoonthorn, R. Loaiza-Maya, F. Huber, G. Koop, J. Maheu, D. Nibbering, and A. Panagiotelis (2023). Bayesian forecasting in economics and finance: A modern review. *International Journal of Forecasting 40*, 811–839.

McAlinn, K. and M. West (2019). Dynamic Bayesian predictive synthesis in time series forecasting. *Journal of Econometrics 210*, 155–169.

Mitchell, J. and S. G. Hall (2005). Evaluating, comparing and combining density forecasts using the KLIC with an application to the Bank of England and NIESR fan charts of inflation. *Oxford Bulletin of Economics and Statistics 67*, 995–1033.

Raftery, A. E., M. Kárný, and P. Ettler (2010). Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill. *Technometrics 52*, 52–66.

Ragonneau, T. M. and Z. Zhang (2023). PDFO: A cross-platform package for Powell's derivative-free optimization solvers. arxiv.org/abs/2302.13246.

Smith, J. Q. (1979). A generalization of the Bayesian steady forecasting model. *Journal of the Royal Statistical Society (Series B: Methodological) 41*, 375–387.

Tallman, E. (2024). *Bayesian Predictive Decision Synthesis: Methodology and Applications*. Ph. D. thesis, Duke University. stat.duke.edu/alumni/alumni-lists-theses/phd-year.

Tallman, E. and M. West (2022). On entropic tilting and predictive conditioning. *Technical Report, Duke University*. arxiv.org/abs/2207.10013.

Tallman, E. and M. West (2023). Bayesian predictive decision synthesis. *Journal of the Royal Statistical Society Series B: Statistical Methodology 86*, 340–363.

Tallman, E. and M. West (2024). Predictive decision synthesis for portfolios: Betting on better models. *Technical Report, Duke University, Submitted for publication*.

West, M. (2023). Perspectives on constrained forecasting. *Bayesian Analysis*. doi.org/10.1214/23-BA1379.

West, M. and P. J. Harrison (1989). *Bayesian Forecasting and Dynamic Models* (1 ed.). Springer-Verlag.

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

Wu, J. C. and F. D. Xia (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking 48*, 253–291.

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