Real-Time Inflation Forecasting in a Changing World*

Jan J. J. Groen†
Federal Reserve Bank of New York

Richard Paap‡
Erasmus University Rotterdam

Francesco Ravazzolo§
Norges Bank

ECONOMETRIC INSTITUTE REPORT 2009-19

Abstract

This paper revisits inflation forecasting using reduced form Phillips curve forecasts, i.e., inflation forecasts using activity and expectations variables. We propose a Phillips curve-type model that results from averaging across different regression specifications selected from a set of potential predictors. The set of predictors includes lagged values of inflation, a host of real activity data, term structure data, nominal data and surveys. In each of the individual specifications we allow for stochastic breaks in regression parameters, where the breaks are described as occasional shocks of random magnitude. As such, our framework simultaneously addresses structural change and model certainty that unavoidably affects Phillips curve forecasts. We use this framework to describe PCE deflator and GDP deflator inflation rates for the United States across the post-WWII period. Over the full 1960-2008 sample the framework indicates several structural breaks across different combinations of activity measures. These breaks often coincide with, amongst others, policy regime changes and oil price shocks. In contrast to many previous studies, we find less evidence for autonomous variance breaks and inflation gap persistence. Through a *real-time* out-of-sample forecasting exercise we show that our model specification generally provides superior one-quarter and one-year ahead forecasts for quarterly inflation relative to a whole range of forecasting models that are typically used in the literature.

Keywords: Inflation forecasting, Phillips correlations, real-time data, structural breaks, model uncertainty, Bayesian model averaging.
JEL Classification: C11, C22, C53, E31

---

*This paper has benefited from helpful comments by Todd Clark and seminar participants at the Erasmus University Rotterdam, the Federal Reserve Bank of Kansas City, the European Meeting of the Econometric Society in Milan, the Riksbank Workshop on 'Modeling and Forecasting Economic and Financial Time Series with State Space Models' in Stockholm, the CIMF Workshop on 'Forecasting under Model Instability' at Cambridge University, the Australasian Meeting of the Econometric Society in Canberra and the Forecasting and Methods in Macro & Finance Workshop at the NBER Summer Institute in Boston. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of New York, the Federal Reserve System or the Central Bank of Norway.

†International Research, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045, United States; email: jan.groen@ny.frb.org

‡Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR, Rotterdam, Netherlands; email: paap@ese.eur.nl

§Research Department, Norges Bank, Bankplassen 2, 0107 Oslo, Norway; email: francesco.ravazzolo@norges-bank.no

1 Introduction

Control of inflation is at the core of monetary policymaking and, consequently, central bankers have a great interest in reliable inflation forecasts to help them achieving this aim. For other agents in the economy accurate inflation forecasts are likewise of importance, either to be able to assess how policymakers will act in the future or to help them in forming their inflation expectations when negotiating about wages, price contracts and so on. And in the academic literature inflation predictability is assessed to get a gauge on the characteristics of inflation dynamics in general.

The time series properties of inflation measures, however, have changed substantially over time, as shown by Cogley and Sargent (2002, 2005) for the United States, by Benati (2004) for the United Kingdom and by Levin and Piger (2004) for twelve main OECD economies, all of which document significant time-variation in the mean and persistence of inflation. Related to that, Cogley and Sargent (2002) and Haldane and Quah (1999) document substantial shifts in the traditional U.S. and U.K. Phillips curve correlations between inflation and unemployment over the post-WWII period. As Stock and Watson (2007) argue, the observed time-variation in the data generating process of inflation has made it increasingly more difficult to forecast inflation. Next to that, Cogley *et al.* (2009) use, amongst others, time-varying vector autoregressive (VAR) models that exploit the earlier mentioned Phillips curve correlation for several U.S. inflation measures. They show that the resulting R^2-type predictability statistics for inflation have fluctuated substantially over the U.S. post-WWII period and have decreased significantly in the post-1980 years.

Therefore, adding structural change to time series models may help to improve forecasting inflation. Stock and Watson (2007, 2008) show that U.S. inflation is well described by a univariate unobserved component model with a stochastic volatility specification for the disturbances. The out-of-sample performance of this particular model appears to be hard to beat by alternative models, including Phillips curve-type models. More generally, Koop and Potter (2007), through change-point models, and Pesaran *et al.* (2006), through a hierarchical hidden Markov chain model, show that forecast models that incorporate structural breaks exhibit good out-of-sample forecasting performance for a range of macroeconomic series.[1]

Another issue for inflation forecasting is how to choose the predictor variables for future inflation. From a macroeconomic point of view, a reduced form version of the Phillips curve

[1]Clark and McCracken (2008), on the other hand, use VAR models with sequentially updating of lag orders, various windows for parameter estimation, (over-)differencing of variables, intercept corrections, and allowing for discrete breaks in parameters. Their results vary across forecast variables, but in general univariate models seem to be difficult to beat by these VARs that allow for structural changes.

relationship is an obvious choice, as it is a tool often used by macroeconomists to assess how economic fluctuations and expectations impact on inflation dynamics. For forecasting, this framework suggest a model where inflation depends on its lags, a measure of real activity (which approximates the degree of 'economic slack' or excess demand in the economy) and, possibly, a measure of inflation expectations.

Although a number of studies use unemployment as the 'slack measure' in such a Phillips curve forecasting model, there is a lot of uncertainty about the 'appropriate' measure of real activity that can be used in such a forecasting model. Stock and Watson (1999) show that unemployment-based Phillips curve models are frequently outperformed by models using alternative real activity measures. They consider two approaches. One is based on a forecast combination of the different, possible choices of Phillips curve forecasting models. Next, they also consider a single Phillips curve-based model that uses a principal component extracted from all possible 'economic slack' variables as the real activity measure. Stock and Watson (1999) show that the out-of-sample performance of these approaches are favorable compared to traditional Phillips curve specifications, in particular in case of the factor-based approach. Atkeson and Ohanian (2001), on the other hand, apply the Stock and Watson (1999) exercise on a longer U.S. sample, and in their case none of the Phillips curve inflation forecasting models are able to outperform naive random walk forecasts.

Like Stock and Watson (1999), we use in this paper a general version of the reduced form Phillips curve model to forecast inflation, which essentially is an autoregressive model for inflation with added exogenous regressors (an AR-X model). But unlike those papers, we use a framework that allows for both instability in the relationship between inflation and predictor variables as well as uncertainty regarding the inclusion of potential predictors in the Phillips curve-type regression. Bayesian model averaging is used to deal with the latter model uncertainty, where we average over the range of regression models that incorporate all the possible combinations of indicator variables for inflation. To deal with instability, we allow for occasional structural breaks of random magnitude in the regression parameters for each of the regression models that are combined within this model average as well as the error variance. Hence, our forecasting procedure simultaneously incorporates the two major sources of uncertainty, which the literature has shown to be relevant for forecasting and modeling inflation.

Our framework, described above, as well as other more regularly used approaches are used to model different definitions of U.S. inflation on a quarterly sample starting in 1960 and ending in 2008. A range of predictor variables are considered in the modeling exercise, from real variables to nominal and financial variables as well as lags of inflation.

The full sample results show that our methodology identifies several structural breaks in the relationship between the different U.S. inflation rates and potential predictor variables. These changes appear to be caused by important events such as, e.g., the oil crisis and changes in the monetary policy regime. The different specifications are then used to forecast the different inflation measures at both one-quarter ahead and one-year ahead forecast horizons. Where necessary, we use in the out-of-sample forecasting experiments real-time data for inflation and the predictor variables, i.e. the original vintage of data that was available at the time of the forecast. We find that allowing for model uncertainty in combination with structural breaks results in superior forecasts *vis-à-vis* other inflation forecasting approaches.

The remainder of this paper is organized as follows. In Section 2 we introduce our Phillips curve model specification. We discuss the estimation methodology in more detail in Section 3. In Section 4 we apply our model to describe the characteristics of U.S. inflation dynamics in the post-WWII era. Next, we evaluate its real-time forecasting performance in Section 5 by comparing it to other univariate and multivariate model specifications. Finally, in Section 6 we conclude.

2 A Framework for Inflation Modeling

To forecast inflation one can simply suffice by using an autoregressive specification. However, based on economic reasoning, we would expect there to be a set of variables that have predictive power for future inflation over and above contemporaneous and lagged inflation. A framework in which one can think about the role of these predictor variables is spelled out in Section 2.1. As will become clear in that subsection, there are a number of specification issues with such a generalized Phillips curve model of inflation. We therefore propose in Section 2.2 a version of this relationship that potentially can deal with these issues.

2.1 A Reduced Form Generalized Phillips Curve Model

The Phillips curve relationship is originally based on the negative correlation between inflation and unemployment that has been observed over time at varying degrees of strength and significance. Similar relationships between inflation and real activity measures as output growth, detrended output and so on have also been found to be of empirical importance, again at varying degrees of strength and significance. A rationalization for the existence of these relationships is often based on the assumption that there are rigidities in the structure of the economy, such as sticky wages and prices, agents with imperfect information, menu costs and the like. The presence of these rigidities imply, therefore, that there is a set of variables out there, other than inflation, with potential predictive power for future inflation.

Empirical, reduced form Phillips curve models are often explicitly or implicitly based on a traditional 'cost-push' approach to inflation: wage and production costs (the latter amongst others related to energy and imports) drive fluctuations in inflation. The corresponding regression model relates inflation to its own lags, the unemployment gap relative to NAIRU[2] and control variables for supply shocks. Gordon (1997), Stock and Watson (1999) and Atkeson and Ohanian (2001) are examples of empirical applications of this Phillips curve specification on U.S. data.

The modern, New-Keynesian view on the Phillips curve correlation is founded on pricing behavior at the firm level. In each period, only a fraction of firms can reset their prices and they do that in a forward-looking manner such that they maximize their present and future profits. In this framework one ends up with a relationship where inflation depends on either real cost measures, such as the labor share and unit labor costs, or the output gap,[3] plus inflation expectations; see Galí and Gertler (1999). Rule-of-thumb behavior or inflation indexation by firms that cannot change their prices would add lags of inflation to this relationship (see, e.g., Galí and Gertler (1999) and Christiano *et al.* (2005)). Examples of empirical work based on this relationship are Galí and Gertler (1999) and Sbordone (2002). Most of this work, however, entails in-sample studies aimed at uncovering the underlying structural parameters instead of using reduced form representations for the purpose of inflation forecasting.

It is therefore clear that *a priori* the range of potential predictors for inflation is large. Empirically, researchers have ran inflation forecasting regressions using a wide array of predictor variables motivated by the Phillips curve relationship, like unemployment, wages and so on. To deal with this type of uncertainty regarding the specification of this relationship Stock and Watson (1999) use both forecast combinations as well as a factor extracted across 132 explanatory variables. Similarly, Atkeson and Ohanian (2001) run a total of 132 different predictive regressions using comparable indicators as Stock and Watson (1999). Wright (2003) applies Bayesian model averaging across 93 potential specifications, each using one alternative activity measure, to forecast different quarterly U.S. inflation measures out-of-sample. These strategies have mixed success: Stock and

[2]NAIRU stands for *non-accelerating inflation rate of unemployment*, which is the unemployment rate at which the excess demand for labor is such that there is no wage pressure that can result in changes in the inflation rate. The unemployment gap is usually approximated by demeaned unemployment or applying some statistical filter on unemployment.

[3]The labor share and unit labor costs can be seen as proxies for the marginal costs of the representative firm, whereas the output gap reflects the excess demand for goods and is suggestive of the market potential of the goods produced by the representative firm. In both cases, the variables provide an indication of the representative firm's profitability.

Watson (1999) and Wright (2003) are able to beat AR inflation forecasts out-of-sample, but Atkeson and Ohanian (2001) are not able to beat out-of-sample random walk inflation forecasts.

In this paper, we will use the following version of the Stock and Watson (1999) generalized Phillips curve specification as the starting point for modeling inflation dynamics:

$$\begin{aligned} y_{t+h} =& \beta_0 + \sum_{j=1}^{k_1} \beta_j^a a_{jt} + \sum_{j=1}^{k_2} \beta_j^e e_{jt} + \sum_{j=0}^{k_3} \beta_j^y y_{t-j} + \sigma\varepsilon_t \\ =& \beta_0 + \sum_{j=1}^{k} \beta x_{jt} + \sigma\varepsilon_t; \quad t = 1, \dots, T-h, \end{aligned} \tag{1}$$

where T is the total number of time series observations in the sample. Variable y_t in (1) is the inflation measure, defined as $y_t = 100\Delta\ln(P_t) = 100(\ln(P_t) - \ln(P_{t-1}))$ where P_t is a particular price index and $h > 0$ is the forecast horizon with $y_{t+h} = 100\Delta\ln(P_{t+h}) = 100(\ln(P_{t+h}) - \ln(P_{t+h-1}))$. The a_{jt}'s are the k_1 real activity and costs indicator variables and the e_{jt}'s are k_2 proxies of inflation expectations. The model contains k_3 lagged values of y_t and for the disturbance term ε_t we assume that $\varepsilon_t \sim \text{NID}(0,1)$ and $\sigma > 0$. For the ease of notation, we define $(x_{1,t} \cdots x_{k,t})' = (a_{1,t} \cdots a_{k_1,t}\ e_{1,t} \cdots e_{k_2,t}\ y_t \cdots y_{t-k_3})'$ and thus $k = k_1 + k_2 + k_3$. Clearly, the number of predictor variables k in (1) will in practice be large; the aforementioned studies use up to 132 series, whereas we use in this paper up to 14 variables in addition to the lags of inflation. Such a large number for k renders the model inestimable and we therefore have to make a choice about which combination of predictors to include under what circumstances. Hence, we have to adapt (1) such that it incorporates this model uncertainty.

Next, it is not realistic to assume that the relationship between inflation and its potential predictors in (1) has remained stable in our 1960-2008 sample. Different studies for different countries utilizing different techniques univocally document substantial changes in the time series properties of inflation in OECD economies over the post-WWII period. Cogley and Sargent (2002, 2005) for the United States, Benati (2004) for the United Kingdom and Levin and Piger (2004) for twelve OECD economies, for example, observe shifts in the mean and persistence of inflation, and these shifts often coincides with policy regime changes. The changing low frequency behavior of inflation, in turn, will cause time-variation in the Phillips curve relationship. Cogley and Sbordone (2008) and Groen and Mumtaz (2008) show that an empirical New Keynesian Phillips curve model that allows for shifts in the equilibrium inflation rates yields a time-varying reduced form inflation-real activity trade-off, given unchanged 'deep parameters', for a number of G7 economies.

There is also evidence that macroeconomic time series have experienced variance breaks over the post-WWII period that were unrelated to shifts in the mean. See, for example, Cogley and Sargent (2005) who use for the U.S. a VAR model in inflation, unemployment and the interest rate with a stochastic volatility specification for the corresponding disturbance covariance matrix. Also, Sensier and van Dijk (2004) find that for 80% of 214 U.S. macroeconomic time series over 1959-1999 most of the observed reduction in volatility is due to a reduction in conditional volatility rather than breaks in the conditional mean. Sims and Zha (2006) even claim that the observed time-variation in U.S. macroeconomic dynamics are entirely due to breaks in the variance of shocks and not in regression parameters. Thus, next to the uncertainty about the inclusion of predictor variables, we need to account for some form of time-variation in both the regression parameters and error variance of (1).

2.2 Incorporating Model Uncertainty and Structural Breaks

The previous discussion makes it clear that we need to adapt the basic inflation regression model (1) such that it incorporates *model uncertainty* and *structural breaks* as both inflation itself and the Phillips curve correlation between inflation and indicator variables have changed over time.

In our context, model uncertainty reflects the uncertainty about which combination of indicator variables most accurately summarizes the impact of real activity, real costs and expectations on inflation dynamics. To allow for model uncertainty we introduce in our original generalized Phillips curve model (1) k variables $\delta_j \in \{0,1\}$ that describe the inclusion of variable x_{jt} in the regression model for $j = 1, \ldots, k$. This results in

$$y_{t+h} = \beta_0 + \sum_{j=1}^{k} \delta_j \beta_j x_{jt} + \sigma \varepsilon_t; \quad t = 1, \ldots, T-h, \tag{2}$$

where $\varepsilon_t \sim \text{NID}(0,1)$. The vector $D = (\delta_1, \ldots, \delta_k)'$ describes which regressors are included in the regression model. It can take 2^k different values, resulting in 2^k different regression models. Model selection is therefore defined in terms of variable selection, see George and McCulloch (1993) and Kuo and Mallick (1998). We denote each model by the index $i = (\delta_1, \ldots, \delta_k)_2$. Note that the intercept parameter β_0 is always included in the model.

Structural breaks in the regression parameters and the variance are incorporated by introducing time-varying regression parameters β_{jt} and σ_t in (1), that is,

$$y_{t+h} = \beta_{0t} + \sum_{j=1}^{k} \beta_{jt} x_{jt} + \sigma_t \varepsilon_t; \quad t = 1, \ldots, T-h. \tag{3}$$

The structural breaks are described by $k+2$ random variables κ_{jt} which equal 1 in case of a structural break in the jth parameter at time t and 0 otherwise for $j = 0, \ldots, k+1$

and $t = 1, \ldots, T$. We assume that the vector $\kappa_t = (\kappa_{0,t}, \ldots, \kappa_{k,t}, \kappa_{k+1,t})'$ is a sequence of uncorrelated 0/1 processes with

$$\Pr[\kappa_{jt} = 1] = \pi_j; \quad j = 0, \ldots, k+1. \tag{4}$$

The size of the structural breaks is described by an independent random shock η_{jt} with mean zero and variance q_j^2 for $j = 0, \ldots, k+1$. Hence, the time varying parameters are defined as

$$\begin{aligned} \beta_{jt} &= \beta_{j,t-1} + \kappa_{jt}\eta_{jt}; \qquad j = 0, \ldots, k, \\ \ln \sigma_t^2 &= \ln \sigma_{t-1}^2 + \kappa_{k+1,t}\eta_{k+1,t} \end{aligned} \tag{5}$$

with $\eta_t = (\eta_{0,t}, \ldots, \eta_{k+1,t})' \sim \text{NID}(0, Q)$ with $Q = \text{diag}(q_1^2, \ldots, q_{k+1}^2)$.

This specification implies that a regression parameter β_{jt} in (5) remains the same as its previous value $\beta_{j,t-1}$ unless $\kappa_{jt} = 1$ in which case it changes with η_{jt}, see, for example, Koop and Potter (2007) and Giordani *et al.* (2007) for a similar approach. Stochastic structural breaks in the variance parameter $\ln \sigma_t^2$ comply to a similar structure as the β_{jt} parameters. The flexibility of the specification in (5) stems from the fact that the parameters β_{jt} and σ_t^2 are allowed to change every time period, but they are not imposed to change at every point in time. Another attractive property of (5) is that the changes in the individual parameters are not restricted to coincide but are allowed to occur at different points in time.

By combining the two previously discussed extensions of our basic model (1) we obtain a reduced form Phillips curve specification for inflation that simultaneously incorporates model uncertainty and the possibility of structural breaks

$$y_{t+h} = \beta_{0t} + \sum_{j=1}^{k} \delta_j \beta_{jt} x_{jt} + \sigma_t \varepsilon_t; \quad t = 1, \ldots, T-h \tag{6}$$

with $\varepsilon_t \sim \text{NID}(0, 1)$ and (4)–(5).

For parameter inference in (4)–(6), we opt for a Bayesian approach. Such an approach allows us to incorporate parameter uncertainty when forecasting inflation in a natural way. Also, Bayesian inference on $D = (\delta_1, \ldots, \delta_k)$ leads to posterior probabilities for the 2^k possible model specifications. We will use these posterior probabilities for Bayesian model averaging to incorporate model uncertainty into a single inflation forecast. Finally, the approach provides us with the posterior distribution of the unobserved κ_t processes for $t = 1, \ldots, T-h$, which can be used to infer on the timing of structural breaks. By definition, κ_t in (6) does not depend on D which implies that the value of κ_t can be different across different values of D. Hence, structural breaks can occur in different parameters at different time periods across different models, and we average over the latter to obtain our final Phillips curve-type equation.

3 Econometric Methodology

The aim of this section to explain parameter inference in our generalized Phillips curve model under Bayesian model averaging and structural breaks (BMASB), i.e., (4)–(6). Furthermore, we discuss how to obtain forecasts of inflation using this approach. In Section 3.1, we start off with describing the specification of the prior distributions for the parameters and latent variables. We then summarize our Monte Carlo Markov Chain (MCMC) approach used to conduct inference on the model parameters in (4)–(6). Section 3.2 explains how the posterior results can be used as to obtain inflation forecasts.

3.1 Estimation and Inference

The parameters in the model (4)–(6) are the inclusion vector $D = (\delta_1, \ldots, \delta_k)'$, the structural break probabilities $\pi = (\pi_0, \ldots, \pi_{k+1})'$ and the vector of variances of the size of the breaks $q = (q_0^2, \ldots, q_{k+1}^2)'$. We collect the model parameters in a $(3k+4)$-dimensional vector $\theta = (D', \pi', q')'$.

For our Bayesian approach we need to specify the prior distributions for the model parameters. For the variable inclusion parameters we take a Bernoulli distribution with

$$\Pr[\delta_j = 1] = \lambda_j \qquad \text{for } j = 1, \ldots, k. \tag{7}$$

Hence, the parameter λ_j reflect our prior belief about the inclusion of the jth explanatory variable, see George and McCulloch (1993) and Kuo and Mallick (1998). For the structural break probability parameters we take Beta distributions

$$\pi_j \sim \text{Beta}(a_j, b_j) \qquad \text{for } j = 0, \ldots, k+1. \tag{8}$$

The parameters a_j and b_j can be set according to our prior belief about the occurrence of structural breaks. Finally, for the variance parameters we take the inverted Gamma-2 prior

$$q_j^2 \sim \text{IG-2}(\nu_j, \omega_j) \qquad \text{for } j = 0, \ldots, k+1, \tag{9}$$

where ν_j, ω_j, $j = 0, \ldots, k+1$, are parameters which can be chosen to reflect the prior beliefs about the variances. The joint prior specification $p(\theta)$ is given by the product of the prior specifications in (7)–(9).

Posterior results are obtained using the Gibbs sampler of Geman and Geman (1984) combined with the technique of data augmentation of Tanner and Wong (1987). The latent variables $B = \{\beta_t\}_{t=1}^{T-h}$, with $\beta_t = (\beta_{0t}, \beta_{1t}, \ldots, \beta_{kt})'$, $S = \{\sigma_t^2\}_{t=1}^{T-h}$, and $K = \{\kappa_t\}_{t=1}^{T-h}$ are simulated alongside the model parameters θ. To apply the Gibbs sampler we need the complete data likelihood function, that is, the joint density of the data and the latent variables

$$p(y,B,S,K|x,\theta)=\prod_{t=1}^{T-h}p(y_{t+h}|D,x_t,\beta_t,\sigma_t^2)p(\beta_t|\beta_{t-1},\kappa_t,q_0^2,\ldots,q_k^2)$$
$$p(\ln\sigma_t^2|\ln\sigma_{t-1}^2,\kappa_{k+1,t},q_{k+1}^2)\prod_{j=0}^{k+1}\pi_j^{\kappa_{jt}}(1-\pi_j)^{1-\kappa_{jt}},\quad(10)$$

where $y=(y_1,\ldots,y_T)$ and $x=(x_1',\ldots,x_T')'$. The elements $p(y_{t+h}|D,x_t,\beta_t,\sigma_t^2)$, $p(\beta_t|\beta_{t-1},\kappa_t,q_0^2,\ldots,q_k^2)$ and $p(\ln\sigma_t|\ln\sigma_{t-1},\kappa_{k+1,t},q_{k+1}^2)$ are normal density functions, which follow directly from the model specification (5)–(6).

If we combine (10) together with the prior density $p(\theta)$, we obtain the posterior density function

$$p(\theta,B,S,K|y,x)\propto p(\theta)p(y,B,S,K|x,\theta).\quad(11)$$

Our Gibbs sampler is a combination of the Kuo and Mallick (1998) algorithm for variable selection and the efficient sampling algorithm of Gerlach *et al.* (2000) to handle the (occasional) structural breaks. If we define $\theta=(\bar{\theta}',D')'$ with $\bar{\theta}=(\pi',q')'$ and $K_\beta=\{\kappa_{0t},\ldots,\kappa_{kt}\}_{t=1}^{T-h}$ and $K_\sigma=\{\kappa_{k+1,t}\}_{t=1}^{T-h}$, then in each iteration of the sampler we sequentially cycle through the following steps:

1. Draw D conditional on B, S, K, $\bar{\theta}$, y and x.

2. Draw K_β conditional on D, S, K_σ, $\bar{\theta}$, y and x.

3. Draw B conditional on D, S, K, $\bar{\theta}$, y and x.

4. Draw K_σ conditional on D, B, K_β, $\bar{\theta}$, y and x.

5. Draw S conditional on B, D, K, $\bar{\theta}$, y and x.

6. Draw $\bar{\theta}$ conditional on D, B, S, K, y and x.

A more detailed description of this Gibbs sampling algorithm is provided in Appendix A.

3.2 Forecasting

One purpose of model (4)–(6) is to have a generalized, reduced form Phillips curve model for forecasting inflation that incorporates uncertainty about both the appropriate activity variables and the presence of structural breaks. Within our Bayesian framework, it is straightforward to explicitly take into account these two types of uncertainty, as well as parameter uncertainty. For example, the h-step predictive density of y at time T

conditional on past data is given by

$$p(y_{T+h+1}|y,x,y_{T+1},x_{T+1}) = \int \cdots \int \sum_{D} \sum_{K} \sum_{\kappa_{T+1}} p(y_{T+h+1}|D,x_{T+1},\beta_{T+1},\sigma^2_{T+1})$$
$$p(\beta_{T+1}|\beta_T,\kappa_{T+1},q)p(\sigma^2_{T+1}|\sigma^2_T,\kappa_{T+1},q)\prod_{j=0}^{k+1}\pi_j^{\kappa_{j,T+1}}(1-\pi_j)^{1-\kappa_{j,T+1}}$$
$$p(\bar{\theta},D,B,S,K|y,x)d\beta_{T+1}d\sigma^2_{T+1}dBdSd\bar{\theta}, \quad (12)$$

where $p(y_{T+h+1}|D,x_{T+1},\beta_{T+1},\sigma^2_{T+1})$ and $p(\beta_{T+1}|\beta_T,\kappa_{T+1},q)$ and $p(\sigma^2_{T+1}|\sigma^2_T,\kappa_{T+1},q)$ follow directly from (5)–(6), and where $p(\bar{\theta},D,B,S,K|y,x)$ is the posterior density (11) based on the observations until time T. The predictive density (12) consists of a weighted average over all possible model specifications in (6) with weights equal to the posterior model probabilities. Uncertainty regarding the timing of structural breaks is reflected in (12) by the posterior distribution of the in-sample breaks K. Computation of such a predictive distribution is straightforward using the aforementioned Gibbs draws. We simulate in each Gibbs step y_{T+h} using (4)–(6) as the data generating process, where we replace the parameters and the latent variables by the draw from the posterior distribution. As point forecast we use the posterior median of the predictive distribution.

4 (In-)Stability of U.S. Inflation Dynamics?

In this section we apply our framework to model the post-WWII behavior of two U.S. inflation measures. In Section 4.1 we discuss the data we use. Section 4.2 presents and discusses the characterization of U.S. inflation dynamics that results from applying our generalized Phillips curve model (4)–(6) on our data.

4.1 Data

We will consider in this paper two measures of inflation in the United States for a quarterly sample from 1960Q1 to 2008Q4; these are the quarterly log changes in the Personal Consumption Expenditures (PCE) deflator and the Gross Domestic Product (GDP) deflator. Potentially there is wide array of predictors for inflation that can be useful for the analysis in this paper. Atkeson and Ohanian (2001), for example, consider up to 132 potential indicator variables. However, our aim in the next section is to assess the ability of these predictors to forecast inflation in *real-time*. And as both our inflation measures of interest as well as many potential predictor variables are revised over time, it is crucial to be able to use series for which one can get hold of the original data vintages as would have been available at the time of the forecast. We therefore restrict our pool of possible predictor variables for inflation to those for which we have these original vintages, restricting the range to about fourteen series next to the inflation lags.

Both the inflation measures and the set of potential predictors come either directly or are constructed from five data sources. These are the Real-Time Data Set for Macroeconomists (RTDSM) at the Federal Reserve Bank of Philadelphia, the ALFRED® real-time database at the Federal Reserve Bank of St. Louis, the CRSP database from Wharton Research Data Services, the Reuters/University of Michigan Survey of Consumers and data from Global Financial Data. We refer the reader to Appendix B for more details on the data sources and data construction.

Our range of predictor variables can be typified as follows:

- **Real activity and cost indicators:** real GDP in volume terms (ROUTP), real PCE in volume terms (RCONS), real residential investment in volume terms (RINVR), the import deflator (PIMP), the relative unemployment levels (UNEMPL), non-farm pay rolls (NFPR), housing starts (HSTS), real spot price of oil (OIL), real food commodities price index (FOOD) and real raw material commodities price index (RAW).
- **State of the economy:** broad M2 monetary aggregate, level (YL) and slope factors (TS) from the term structure of interest rates.
- **Inflation expectations:** one-year ahead inflation expectations from the Reuters/ University of Michigan Survey of Consumers (MS) as well as the level factor from the term structure of interest rates (YL).

For most of the variables, we use the percentage change of the original series[4] to remove possible stochastic and deterministic trends from the series. Exceptions are the unemployment ratio and housing starts, for which we use the logarithm of the respective levels, as well as the two term structure factors and the inflation expectations survey for which we use the 'raw' levels of the series.

The above mentioned real activity and cost series provide information about either the degree of excess demand in the economy or about the real costs that firms face, which basically are the a_{jt} series in (1). In addition, these a_{jt} series also include a number of nominal variables that are informative about the current and future state of the economy. Of these latter series, the M2 monetary aggregate can either reflect the current stance of monetary policy, if one believes that its growth rate is exogenously determined by the central bank, or it provides information about spending in households and firms (where increased M2 growth indicates increased spending by households and firms). The term

[4]That is, 100 times the quarterly change of the logarithm of the original series.

structure of interest rates contains a lot of forward-looking information about the business cycle, the stance of monetary policy and inflation expectations.

Ang *et al.* (2006) and Diebold *et al.* (2006) argue that at the quarterly frequency term structure dynamics can be efficiently summarized by two factors: level and slope. We approximate the term structure through the 3-month and 6-month Treasury Bill rates plus the 1-year to 5-year Fama and Bliss (1987) zero-coupon bond yields from the CRSP data base, where the level factor is the average across these 7 interest rates and the slope factor is the difference between the 5-year zero-coupon bond yield and the 3-month Treasury bill rate. The level factor can either be interpreted as a market expectation of the long-run level of inflation (Diebold *et al.* 2006) or as the market expectation of the equilibrium level of the central bank policy rate (Ang *et al.* 2006). The slope factor of the term structure is often seen as a good predictor for both turning points in the business cycle (see, for example, Estrella and Hardouvelis 1991) and of the reaction function of the central bank.

Finally, we use one-year ahead inflation expectations from the University of Michigan Survey of Consumers (MS) as one of the expectations measure for our generalized Phillips curve model - the level factor also can be considered as an expectations measure given the aforementioned interpretation of this term structure determinant. Surveys can give potentially a very good steer about agents' expectations and indeed Ang *et al.* (2007) claim that in an out-of-sample context inflation expectation surveys are the most accurate predictors for future U.S. inflation.

4.2 Full-Sample Inflation Characteristics

In this subsection we estimate our generalized Phillips curve model that incorporates model uncertainty and occasional structural breaks, i.e., (4)–(6), over our full 1960-2008 sample for both the PCE deflator and GDP deflator inflation measures. To operationalize the estimation of our BMASB model (4)–(6) we need to take a stand on the values of the prior parameters discussed in Section 3.1. Firstly, we assign high values to ν_j for $j = 0, 1, \dots k+1$ in the prior distribution (9) for the variances q_j^2 of the break magnitudes in (5). This assumption implies that the magnitude of a break at time t when $\Pr[\kappa_{jt} = 1] = 1$ is proportional to the square root ω_j. The values in the prior distributions (8) for the break probabilities can consequently be chosen to limit the number of these breaks. As the posterior probability $\Pr[\kappa_{jt} = 1]$ is lower than 1, our priors are weak on breaks with magnitude lower than a certain proportion of the square root of ω_j or when the probability of a change is absent. More concretely, in (9) we choose ν_j equal to 100 for $j = 0, 1, \dots, k+1$ with the ω_j's fixed on a scale from 0.01 to 0.5 and both α_j and β_j in (8) have a strictly decreasing pattern for the Beta distribution such that we have no more than 3 breaks of maximum magnitude over the full sample. Finally, in the Bernoulli prior distribution for the variable inclusion parameters δ_j we fix λ_j equal to 0.50 for all $j = 1, \ldots, k$. The latter choice of prior values implies that *a priori* all predictor variables are potentially equally important in explaining inflation h steps ahead, so that when we update the model for each forecast we potentially can have a different range of dominant models *a posteriori*. This choice is inspired by existing evidence on the time-varying empirical properties of inflation. For example, in the 1970s inflation is very persistent and both lags and unemployment are dominant explanatory variables, whereas in the 1990s all of this was much less the case; see also the survey of existing empirical findings in Section 2.1.

The aforementioned priors are used in the MCMC algorithm described in Section 3.1 in order to estimate our BMASB Phillips curve model. We run 9,000 Gibbs draw of which the first 1,000 are deleted for burn-in. Of the remaining last 8,000 draws we retain each 2nd draw to obtain a reasonably random sample resulting in 4,000 MCMC draws that can be used for parameter estimation and inference.

The purpose of our full-sample estimation of the BMASB model (4)–(6) for both U.S. inflation measures is to conduct an *ex-post* analysis of the relevance of the different predictor variables for inflation and possible structural breaks in the different regression parameters. By doing that we are able to document how U.S. inflation dynamics has evolved over time from the viewpoint of the Phillips curve trade-off. For these purposes, we can for now suffice with the final, revised, data for all data using the complete sample period from the first quarter of 1960 until the fourth quarter of 2008. We focus on the most frequently used prediction horizons in this literature, i.e., the one-quarter horizon ($h = 1$) and the one-year horizon ($h = 4$), respectively.[5] The different forecast horizons also allows us to explore differences in the lead-lag relationships between inflation and our set potential predictor variables.

Table 1 provides the posterior mean of the inclusion parameters δ_j for all $j = 1, \ldots, k$ in (6) for $h = 1$; essentially these numbers reflect on average the proportion of times a variable is selected across all possible model specifications. The second and third column in Table 1 show that in case of PCE deflator inflation, all lags appear in one or more of the model specifications in the case of $h = 1$ but the one-quarter and three-quarter lags are far more important than the others. For the $h = 4$ case the one-quarter lagged value has the most chance of being selected, whereas the remaining lags have a much lower probability of being included in the models.

Of the real activity and cost indicator variables, the a_{jt}'s in (1), the most frequently

[5] More specifically, this means modeling the quarterly percentage change of the relevant price deflator in the next quarter as well as four quarters from now, respectively.

selected variables to model PCE deflator inflation at $h = 1$ are real raw materials price inflation, real food price inflation, real oil price inflation, real residential investment growth, unemployment rate and real output growth. These variables are also most frequently selected to model one-year ahead PCE deflator inflation, but real output growth is more important and real raw materials price inflation less. The growth in non-farm pay-rolls plays no role in one-quarter ahead prediction, while for one-year ahead prediction this variable becomes much more important. In general, real activity variables, such as consumption growth, are more important determinants of PCE deflator inflation at $h = 4$ than at $h = 1$, whereas for the latter prediction horizon real cost indicators are relatively more important. For the state of the economy variables M2 growth seems to be the most selected variable for $h = 1$, For $h = 4$ the level term structure is also important. The most important variable (apart from lagged inflation) for both horizons is however the Michigan Consumer survey inflation expectations. For the one-quarter ahead horizon, this variable is even included in almost 95% of the cases.

For GDP deflator inflation we obtain the same conclusion for the Michigan Consumer survey inflation expectations, see the final two columns of Table 1. For the state of the economy, we find different results. For one-quarter ahead forecasting, the term structure level factor seems to be the only important variable, although its importance is limited. For one-year ahead prediction M2 growth and the slope of the term structure are most often selected although their posterior inclusions probabilities are quite smaller than the inclusion probabilities of the survey inflation expectations. There are also some differences in the marginal inclusion probabilities of the real activity and cost indicator variables compared to PCE inflation. Most importantly, real output growth is never selected for both horizons. Nonetheless, as was the case for PCE deflator inflation, real activity measures are relatively more important one-year ahead than they are at $h = 1$ and cost indicators relatively less. If we consider the lag selection of inflation, we see that lags 0 to 3 are selected for $h = 4$ and lags 2 to 3 for $h = 1$ although the lags are less important than in case of PCE deflator inflation.

To shed more light on what combinations of explanatory variables dominate the BMASB generalized Phillips curve model (6) one can look at which variable combinations dominate the model average for each inflation measure at each horizon. To that end, Tables 2 and 3 display the top 10 models in terms of their relative posterior probabilities, as selected by our variable selection procedure, for each of the PCE deflator and GDP deflator inflation measures. In general, the conclusions drawn from the results in Table 1 are confirmed by the composition of the dominant models in Tables 2 and 3, i.e., the most selected variables in Table 1 do show up most frequently amongst those top 10 models. For example, for

PCE and GDP deflator inflation at all horizons the inflation expectations of the Michigan Consumer survey is almost always part of the dominating models. And for PCE deflator inflation at $h = 4$ both real output growth and non-farm payroll growth are included in 9 out of the top 10 models, whereas these are only included in, respectively, one and none of the dominating models for $h = 1$. For both inflation measures these examples are very much consistent with the results in Table 1.

A comparison of Tables 2 and 3 highlights a number of differences in how the BMASB Phillips curve model (6) models the dynamics in our two inflation measures. Firstly, the GDP deflator inflation models for both forecast horizons generally consist of less variables than the models selected for PCE deflator inflation. Furthermore, lagged inflation seems to be more important for the latter inflation measure. Next to that, the sum of the posterior probabilities of the top 10 models for the one-quarter horizon is much higher for GDP deflator inflation than for PCE deflator inflation, i.e., approximately 22% for the former and approximately 15% for the latter. At the one-year horizon this difference is slightly smaller, that is 24% and 29%, respectively. This suggests that at the one-quarter horizon the data is more informative when determining which combination of predictor variables is relevant for modeling GDP deflator inflation than in case of PCE deflator inflation.[6] Hence, it seems that the degree of model uncertainty is higher for modeling PCE deflator inflation than for modeling GDP deflator inflation at the one-quarter horizon.

Next, we turn to the posterior results for the regression parameters in (6) to analyze the pattern of parameter estimates for the predictor variables as well as structural breaks in these estimates. For sake of brevity we only focus on the posterior results for variables that are amongst the most regular selected ones, see Table 1, and do not report the remaining posterior regression parameter results.[7] Figures 1 and 2 displays a selection of the posterior medians of β_{jt}, for $j = 0, \ldots, k$, from the BMASB Phillips curve model (6) estimated for PCE deflator inflation at horizons $h = 1$ and $h = 4$. The posterior medians of β_{jt} are conditional on inclusion of the jth variable, that is $\delta_j = 1$. When we focus on Figure 1 a number of interesting patterns emerge. For the more dominant predictor variables, i.e., the Michigan consumer survey inflation expectations, real food price inflation and real raw material input price inflation (see the first column of Table 1), we observe economically plausible parameter estimates. These variables have a positive impact on one-quarter ahead PCE deflator inflation in Figure 1, as these series mainly proxy the impact of both inflation expectations and cost push factors on inflation. The

[6]Note, though, that this does not mean that we are able to select the 'right' predictor variable for GDP deflator inflation as the respective posterior probabilities are probabilities for each of the models relative to all other possible models.

[7]These unreported posterior regression parameter results are available upon request from the authors.

impact of inflation expectations increases after the first oil shock around 1973 and reaches towards the end of the 1970s. During the 1980s and 1990s inflation expectations appear to become less important, but with resurgence of inflation at the end of the sample this trend is reversed. In case of $h = 4$ in Figure 2, we see a similar time-variation albeit that the impact of this expectations measure becomes insignificant at the end of the sample. These observed pattern in the β_{jt} for inflation expectations mimics the time-variation in the mean of inflation: historically high in the stagflation period of the 1970s and very low during the late 1980s up to the early 2000s. In comparison, the time-variation in the corresponding β_{jt}'s for real food price inflation and real raw material input price inflation is relatively subdued during a large part of the sample in Figure 1. However, the impact of these variables on inflation has in increased substantially since 2001 coinciding with a strong upward and subsequent downward trends in global commodity prices.

The remaining group of parameter estimates in Figure 1 relate to variables that are less frequently selected but are still of importance to model PCE deflator inflation at $h = 1$. Generally, the corresponding parameter estimates are economically plausible and exhibit varying degrees of time-variation. For example, in case of real output growth we observe swings in the corresponding β_{jt} during the 1970s and 1980s that involve sign switches, where a positive inflation impact reflects the inflationary impact of higher aggregate demand but a negative sign can approximate the impact of a supply side shock on inflation.[8] From 2001, however, real output growth appears to have an increasingly higher impact on one-quarter ahead PCE deflator inflation. As another example, the pattern observed in the β_{jt} for unemployment resembles those uncovered in other studies (for example, Cogley *et al.* (2009)): in periods of high average inflation (i.e., the mid-1970s) the trade-off between inflation and unemployment is at its strongest and *vice versa* in periods of low inflation (i.e., after 1985). As we saw in Tables 1 and 2, the relative importance for PCE deflator inflation of real activity measures increases at $h = 4$. From Figure 2 it becomes clear that the more dominating activity measures exhibit at this horizon more pronounced time-variations in the corresponding regression parameters than at $h = 1$. In particular, non-farm payrolls and real consumption growth rates have their largest impact in the mid-1970s, when average inflation is high, which then declines in subsequent periods. Again, this suggests that the inflation-activity trade-off for PCE deflator inflation very much varies with shifts in equilibrium inflation.

Taking all of this evidence on parameter time-variation together, we can identify three periods of structural change with tentative evidence of a fourth one. These periods are 1974-1975, 1979-1982, and the period of the 1990s. The oil price crisis of the 1970s and

[8] A negative supply side shock would push down output growth and push up inflation.

the resulting stagflation coincide with the first break period. The second break period relates to the "monetarist experiment" of the Federal Reserve under Chairman Volcker. Note that this period is often identified as the start of a marked structural change in the Fed's monetary policy, see Clarida *et al.* (2000), amongst others. The third break period appears to be related to the widely documented 'Great Moderation' in the volatility of macroeconomic variables; see, e.g., Sensier and van Dijk (2004). The time-variation pattern in some β_{jt}'s suggest that the high inflation volatility period 2006-2008, driven by volatile global commodity prices and the 2007-2008 financial crisis, can be interpreted as a fourth major break point.

We depict similar posterior medians of a selection of the regression parameters in (6) for GDP deflator inflation in Figures 3 and 4. According to the posterior variable selection probabilities in Table 1, models for this inflation rate at horizon $h = 1$ are dominated by the Michigan consumer survey inflation expectations, as well as real raw material input price inflation, real food commodities price inflation and real oil price inflation. This is a similar group of dominant predictor variables as for one-quarter ahead PCE deflator inflation. As in the case of PCE deflator inflation, in Figure 3 the survey-based inflation expectations had their largest, positive impact on one-quarter ahead GDP deflator inflation during the mid-1970s and early 1980s, highlighting the importance of inflation expectations for price setting when inflation is high, hardly any impact during the 1990s with an increased influence since 2001. With respect to the β_{jt} for real raw material input price inflation, we observe a much more gradual time-variation than in case of PCE deflator inflation. Of the group of less frequently selected predictors, unemployment has a similar, time-varying impact on one-quarter ahead GDP deflator inflation as in case of PCE deflator inflation for similar reasons. At the one-year horizon, see Figure 4, we observe more pronounced swings in the β_{jt}'s of the real activity measures than at $h = 1$ where, as was the case for PCE deflator inflation, the impact is the highest when equilibrium inflation was high, i.e., in the mid-1970s. In general, we find in Figure 3 similar periods of structural parameter change as in the PCE deflator inflation case.

The existing literature has focused a lot on documenting the time-variation in the mean, the persistence and the variance of different inflation measures over the post-WWII period. We did survey parts of this literature in earlier sections, so here we suffice with summarizing the general conclusions of the existing literature:

- Both the mean and persistence of U.S. inflation increased during the 1970s, both reaching peaks around 1974-1975 and around 1980, and subsided after 1982-1983. Certainly during the 1970s inflation behaved as a unit root process, whereas from the late 1980s onwards some inflation measures started to behave as quasi-white noise processes; see, e.g., Cogley and Sargent (2005), Cogley *et al.* (2009).

- A majority of studies suggest that the downward shift in U.S. inflation variability around 1985 has been due to an exogenous break in the variance unrelated to breaks in the mean and/or persistence; some studies claim this also happened during the 1970s; see, e.g., Sensier and van Dijk (2004), Sims and Zha (2006).

Do these findings concur with those from our BMASB Phillips curve-type model (6), where one conditions inflation on the (potentially) time-varying impact of an extended range of predictor variable combinations?

Typically one uses in the existing literature (V)AR model-based approaches and thus the usual inflation persistence measures depend on the sum of autoregressive parameters. More precisely, the sum of autoregressive parameters measures the persistence of inflation that is unrelated to the set of conditioning variables (which in case of an AR model only entails an intercept term). Similarly, variance breaks are usually specified in existing studies as either deterministic structural breaks, Markov switching processes or stochastic volatility specifications with time-varying parameters. Our BMASB specification (6) can produce similar measures, as it allows for inflation lags (up to fourth order) to be included in the range of potential model specifications and for stochastic breaks in the disturbance variance. The first columns of Figures 5 and 6 report for (6) the time-variation in the intercept β_{0t}, average persistence and the error variance σ_t^2 for the PCE deflator and GDP deflator inflation rates, respectively, at the one-quarter ahead horizon. In these figures, average persistence is computed by averaging the sum of the included autoregressive parameters across all model specifications using the posterior model probabilities.

For sake of interpretation we report in the second columns of Figures 5 and 6 similar measures based on a time-varying parameter AR (TVP-AR) model and to save space we focus solely on $h = 1$. This TVP-AR model is a version of (6) where we average solely over inflation lags of up to a lag order of 4 quarters, i.e.,

$$y_{t+h} = \beta_{0t} + \sum_{j=0}^{3} \delta_j^* \beta_{jt} y_{t-j} + \sigma_t \varepsilon_t. \tag{13}$$

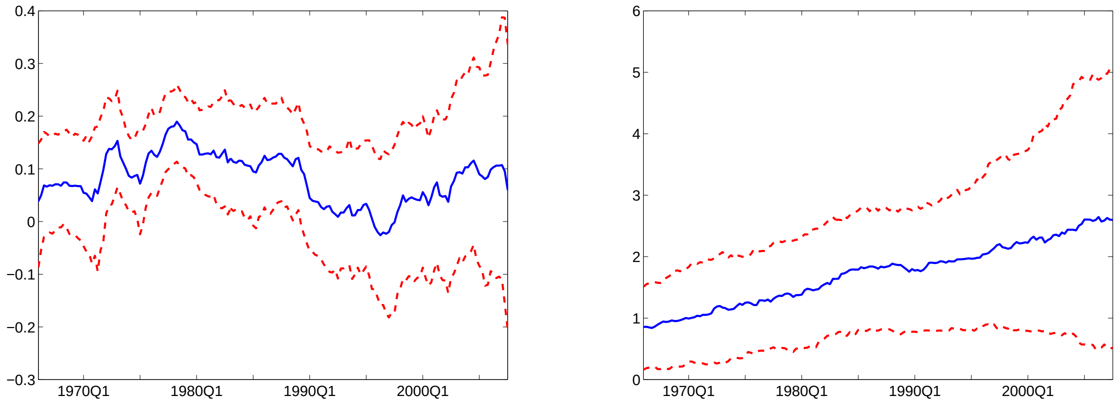
In (13) δ_j^* is a lag order selection parameter similar to the δ_j parameters used in (6). The time-varying intercept, average persistence and error variance terms produced by (13) can be seen as representative of those produced by existing studies, where one usually allows for structural change but does not condition on a large set (of combinations) of additional explanatory variables.

Several conclusions emerge from Figures 5 and 6. Firstly, regardless of the specification we find substantial time-variation in the degree of PCE deflator inflation persistence over our 1960-2008 sample, which seems to peak around the mid-1970s. The degree of persistence implied by the TVP-AR model (13) is very similar to those found in the literature, with the root of inflation at its peak of close to an unit root in the early and late 1970s. For our BMASB Phillips curve model (6), however, the peak in PCE deflator inflation persistence appears to be much lower: around 0.30, with persistence becoming increasingly more negative after the early 1980s. With respect to GDP deflator inflation in Figure 6, persistence as implied by the TVP-AR model (13) behaves similarly as in case of PCE deflator inflation. The inflation persistence implied by our BMASB Phillips curve-type model (6), however, is relatively stable over the whole sample around a low level of approximately 0.10.

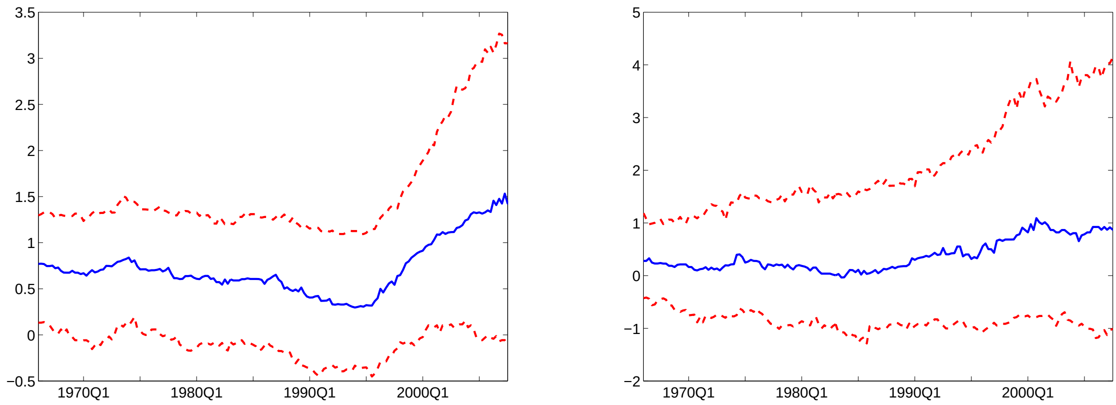
In this context it is worth while to spell out the meaning of these persistence measures for (6) and (13). These measures provide an estimate of the persistence with which inflation on average deviates from either the combined value of the intercept and predictor variables in (6), or from solely the intercept in (13). As it is clear that the intercept in the latter model is relatively stable over the sample, all the low-frequency variation in inflation within (13) will have to come through the persistence terms. There is less of a necessity for this phenomenon in case of our BMASB Phillips curve model, as we have seen in Figures 1 and 3 that the correlation of one-quarter ahead inflation with activity and expectations measures also varies over time.

Next to persistence, we can draw conclusions about autonomous variance breaks from Figures 5 and 6. In case of PCE deflator inflation, essentially none are observed for either specification, suggesting that our particular way of modeling structural change is at the root of this result. For GDP deflator inflation we do observe some time-variation in σ_t^2; under the BMASB Phillips curve specification we notice that the error variance is higher around the mid-1970s than in other periods. For the TVP-AR GDP deflator inflation model we observe a similar pattern in the error variance. Nonetheless, this time-variation in the innovation variance is much less pronounced than that observed in persistence and, in case of (6), in the correlations between inflation and our 14 predictor variables. Overall, the results regarding time-variation in the error variance for both inflation rates suggests that changes in inflation persistence as well as in the persistence and variance of our 14 predictor variables have been the main determinants of changes in the variance of the PCE and GDP inflation rates.

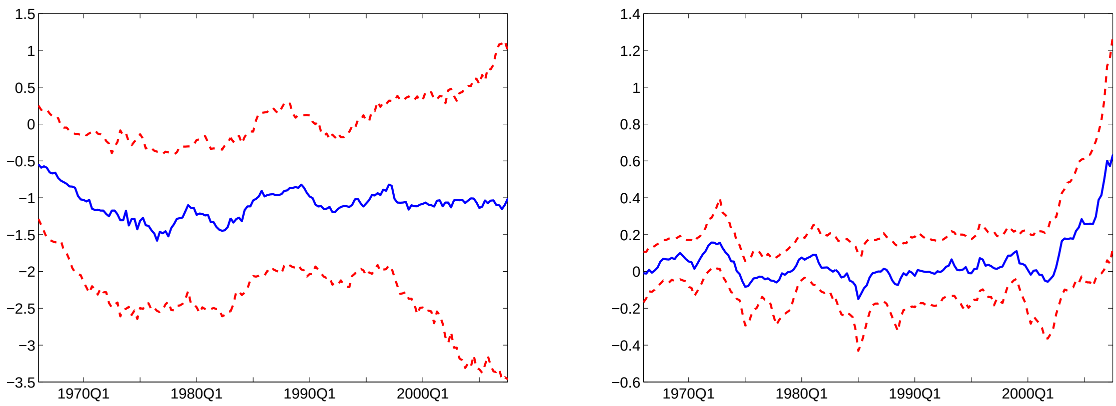
Figure 1: Posterior densities of selected β parameters in BMASB Phillips curve model (6) for $h = 1$ conditional on inclusion: PCE Deflator Inflation



(a) Michigan survey inflation expectations and real raw industrial commodities inflation



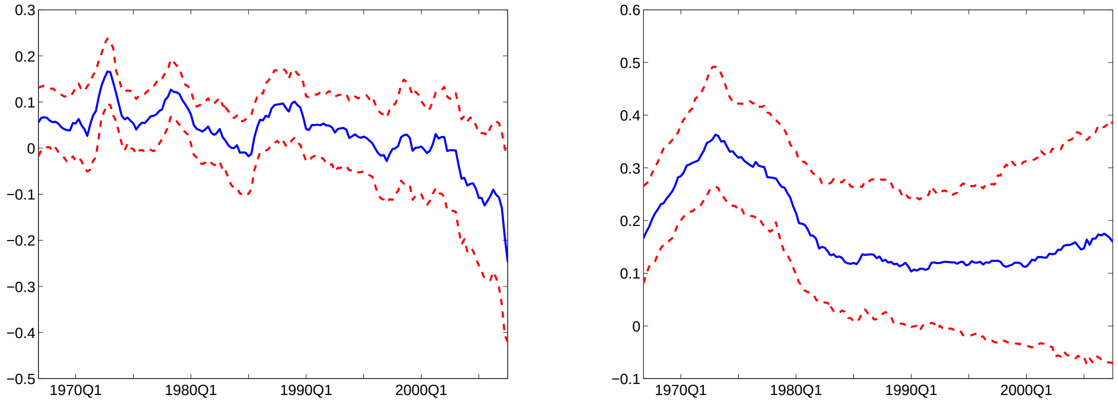
(b) Real food commodities inflation and real residential investment growth



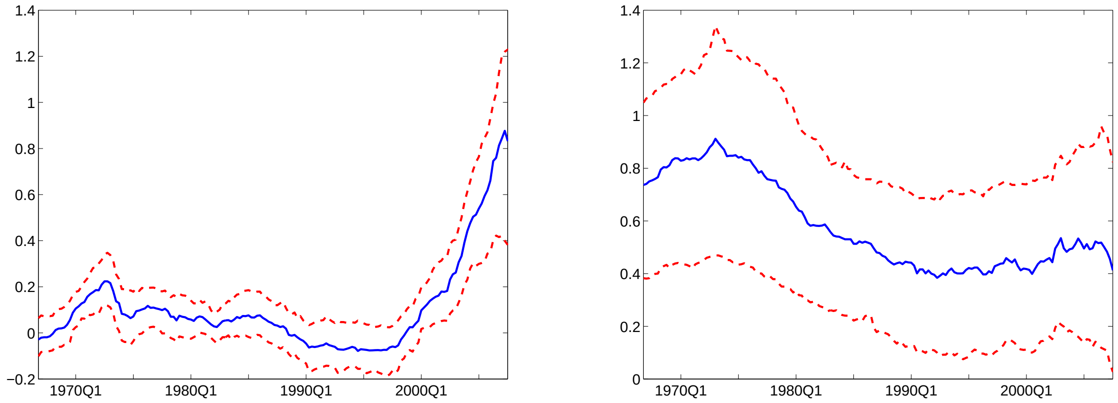
(c) Unemployment ratio and real output growth

Note: The graphs in this figure show the posterior medians of selected β_{jt}'s in (6) for PCE deflator inflation at $h = 1$. The dashed lines in the graphs are the 25th and 75th percentiles of the posterior densities.

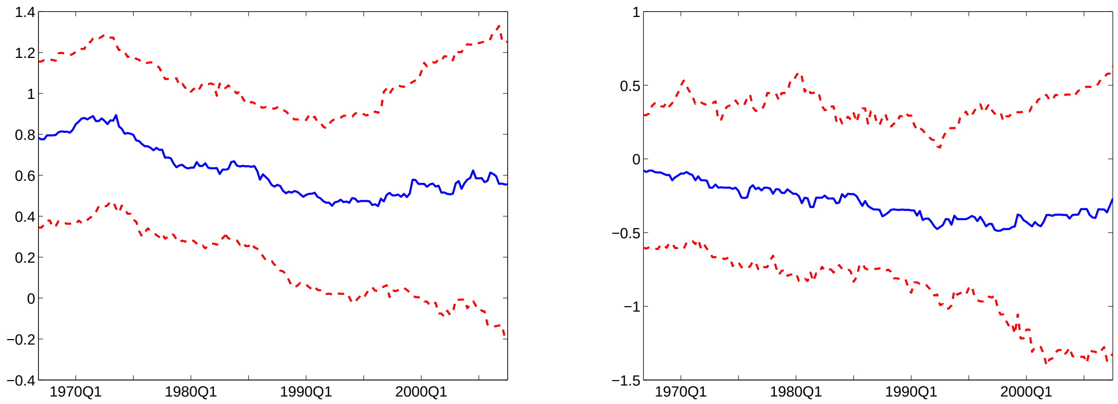
Figure 2: Posterior densities of selected β parameters in BMASB Phillips curve model (6) for $h = 4$ conditional on inclusion: PCE Deflator Inflation



(a) Michigan survey inflation expectations and non-farm payrolls growth rate



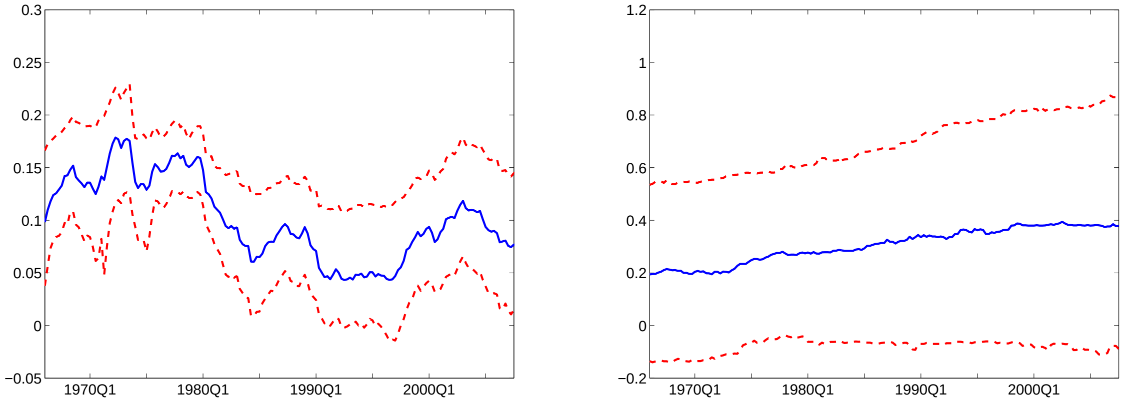
(b) Real output growth and real consumption growth



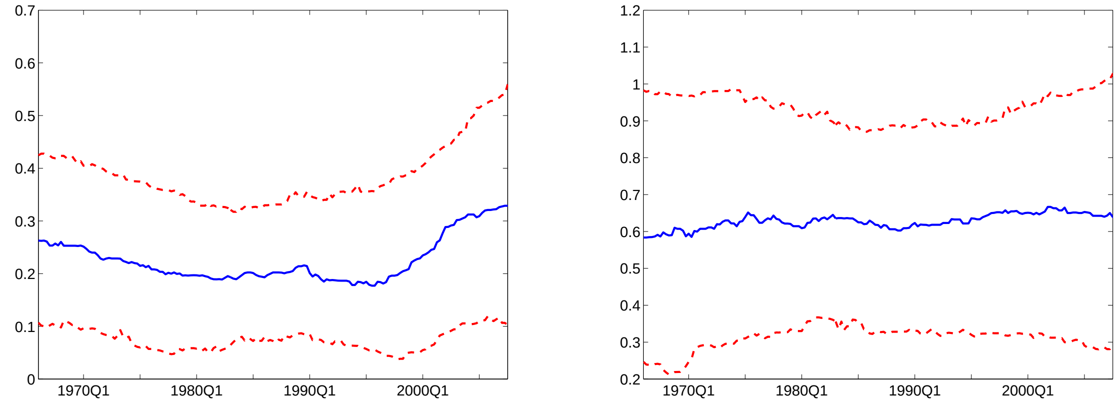
(c) Real food commodities inflation and unemployment ratio

Note: The graphs in this figure show the posterior medians of selected β_{jt}'s in (6) for PCE deflator inflation at $h = 4$. The dashed lines in the graphs are the 25th and 75th percentiles of the posterior densities.

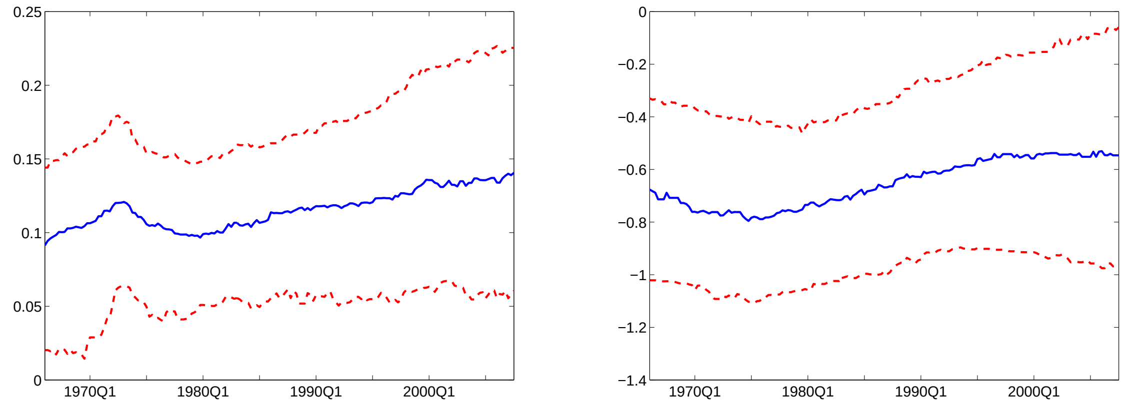
Figure 3: Posterior densities of selected β parameters in BMASB Phillips curve model (6) for $h = 1$ conditional on inclusion: GDP Deflator Inflation



(a) Michigan survey inflation expectations and real raw industrial commodities inflation



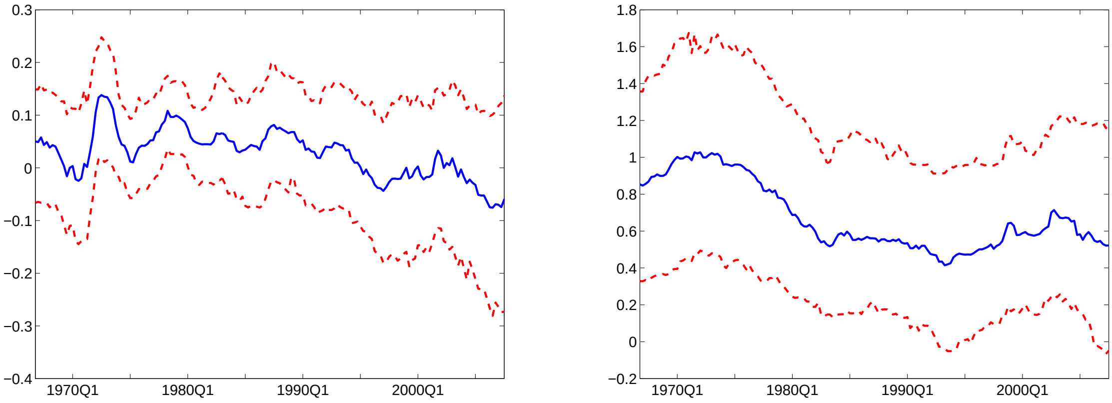
(b) Real oil price inflation and real food commodities inflation



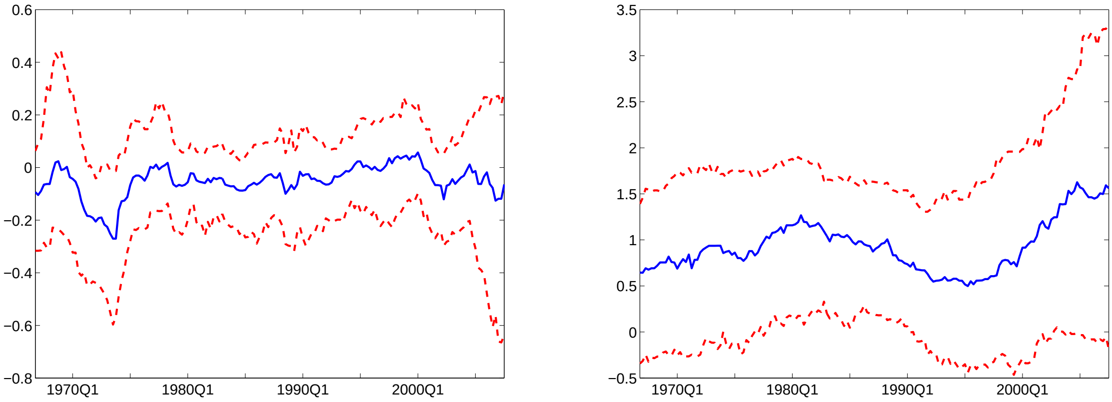
(c) Non-farm payrolls growth rate and unemployment ratio

Note: The graphs in this figure show the posterior medians of selected β_{jt}'s in (6) for GDP deflator inflation at $h = 1$. The dashed lines in the graphs are the 25th and 75th percentiles of the posterior densities.

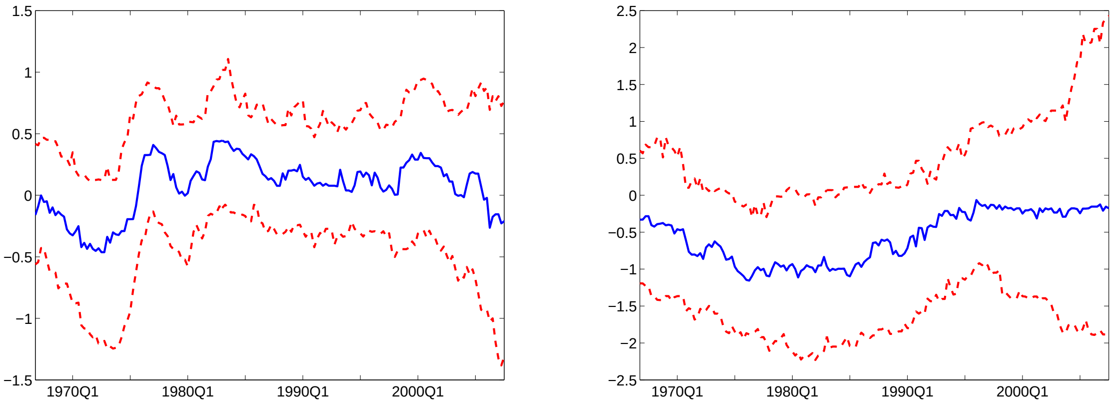
Figure 4: Posterior densities of selected β parameters in BMASB Phillips curve model (6) for $h = 4$ conditional on inclusion: GDP Deflator Inflation



(a) Michigan survey inflation expectations and real consumption growth



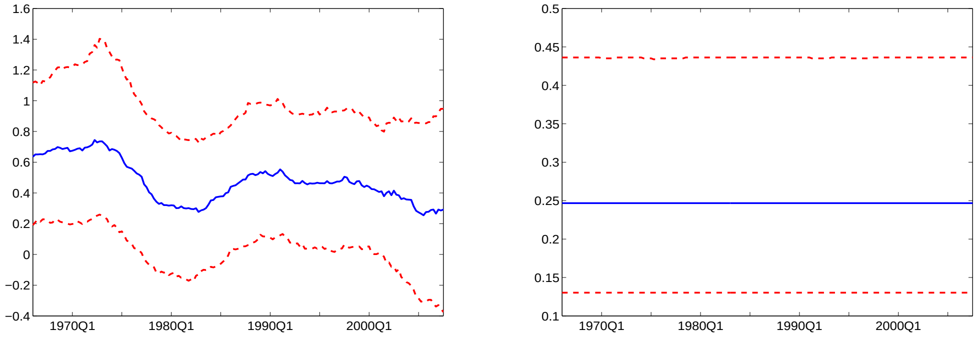
(b) Term structure slope factor and real raw industrial commodities inflation



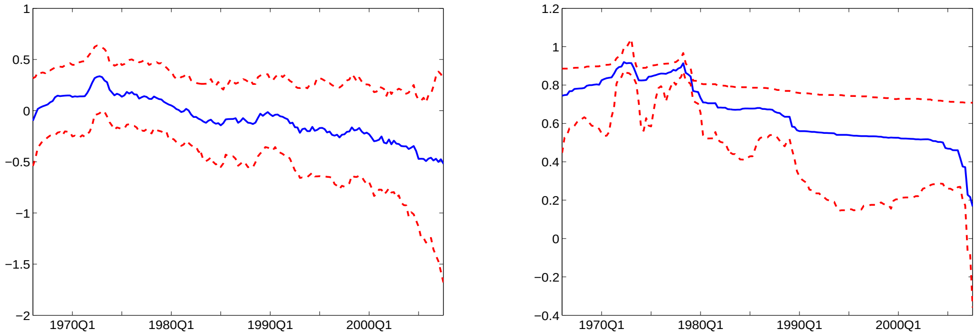
(c) Housing starts growth and unemployment ratio

Note: The graphs in this figure show the posterior medians of selected β_{jt}'s in (6) for GDP deflator inflation at $h = 4$. The dashed lines in the graphs are the 25th and 75th percentiles of the posterior densities.

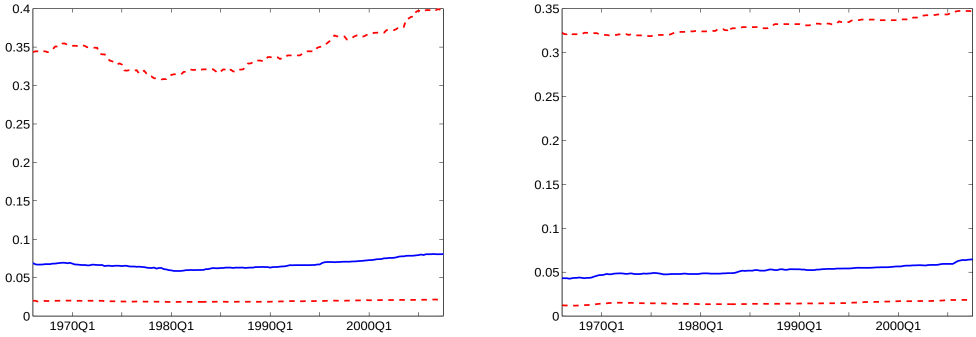
Figure 5: Posterior densities of the intercept, persistence and innovation variance in the TVP-AR model relative to BMASB for $h = 1$: PCE Deflator Inflation



(a) BMASB Intercept – TVP-AR Intercept



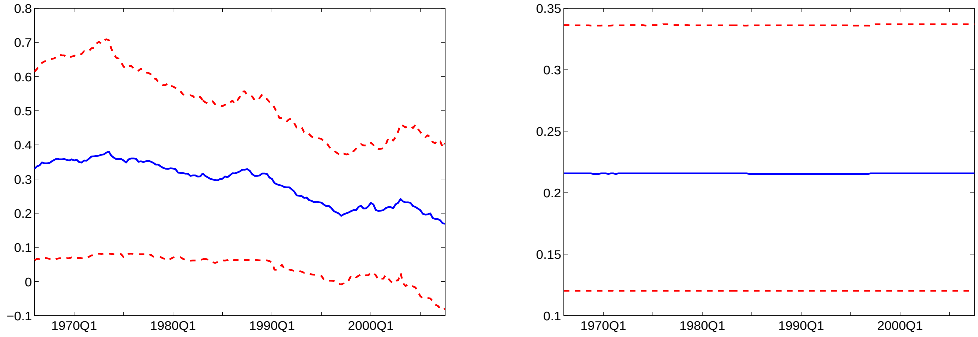
(b) BMASB Persistence – TVP-AR Persistence



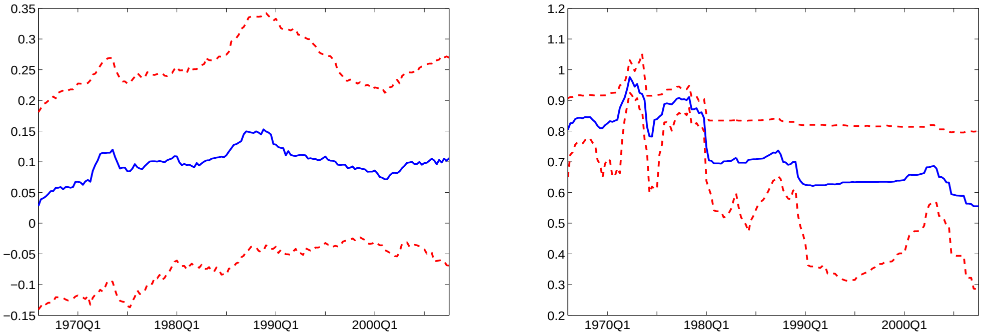
(c) BMASB σ_t^2 – TVP-AR σ_t^2

Note: The graphs in this figure show the posterior medians of the intercept, accumulated persistence and error variance in BMASB model (6) relative to the time-varying AR model (13) for PCE deflator inflation at $h = 1$. Persistence is computed by averaging the sum of the included autoregressive parameters across all model specifications using the posterior model probabilities. The dashed lines in the graphs are the 25th and 75th percentiles of the posterior densities.

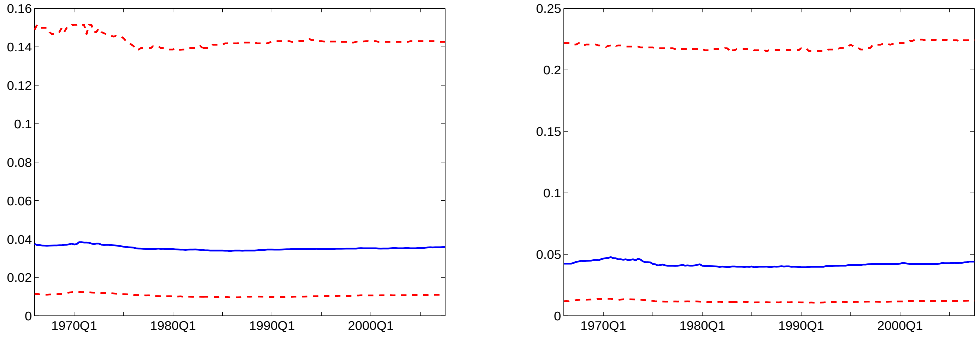
Figure 6: Posterior densities of the intercept, persistence and innovation variance in the TVP-AR model relative to BMASB for $h = 1$: GDP Deflator Inflation



(a) BMASB Intercept – TVP-AR Intercept



(b) BMASB Persistence – TVP-AR Persistence



(c) BMASB σ_t^2 – TVP-AR σ_t^2

Note: The graphs in this figure show the posterior medians of the intercept, accumulated persistence and error variance in BMASB model (6) relative to the time-varying AR model (13) for GDP deflator inflation at $h = 1$. Persistence is computed by averaging the sum of the included autoregressive parameters across all model specifications using the posterior model probabilities. The dashed lines in the graphs are the 25th and 75th percentiles of the posterior densities.

5 Real-Time Prediction of U.S. Inflation Rates

We will be focusing in this section on the out-of-sample forecasting performance of our BMASB Phillips curve model (6) relative to other, often very parsimonious, models that are frequently used for inflation forecasting. Section 5.1 provides an outline of our forecasting exercise, including a description of the alternative models. A discussion of the out-of-sample forecasting results follows in Section 5.2.

5.1 Forecasting procedure

In Section 4.2 we described the full sample developments in inflation dynamics for the U.S. PCE and GDP deflator series through the eyes of our BMASB Phillips curve model (4)–(6). However, the ultimate test for this model is how it competes with alternative specifications in a real-time, out-of-sample context. Hence, the forecasting exercise in this section.

The starting point of our forecasting exercise is the model in (4)–(6), which we have been referring to as the BMASB Phillips curve model. We use the model to obtain and evaluate one-quarter and one-year ahead forecasts for the quarter-on-quarter inflation rate of both the PCE deflator and the GDP deflator in the United States. For computational reasons we obtain the one-year ahead forecasts through direct forecasting.[9] Each forecast is based on a re-estimation of the model using an expanding window of historical data and the MCMC procedure outlined in Section 3.1. For example, suppose the first h-step ahead forecast is produced in quarter t_0 for $h = 1, 4$. As we want to evaluate the forecasts in real-time, we use the original vintage of data available at t_0 to re-estimate the BMASB Phillips curve model on the sample $t = 1, \ldots, t_0$, with the forecast horizon $h = 1$ or 4. The resulting, direct, forecast using data on x_{jt} for $t = 1, \ldots, t_0$ and the posterior draws from the estimation up to t_0 (see Section 3.2) is then evaluated against the vintage of inflation data that is available h quarters ahead, i.e., the vintage at $t_0 + h$. We repeat this process of re-estimation and forecast generation for $t_0 + 1, \ldots, T - h$. This results in a time series of forecast errors for $t = t_0, \ldots, T - h$, which we then use to compute the square root of mean squared forecast errors (RMSE).[10]

To assess how our BMASB Phillips curve model (6) performs in real-time, we need

[9]Whether an iterative procedure provides more accurate forecasts than a direct approach is a matter of ongoing debate, see the discussion in Marcellino *et al.* (2006).

[10]That is, if one defines the out-of-sample forecast error of a model for y_{t+h} as $\hat{\epsilon}_{t+h}$ then

$$\text{RMSE} = \sqrt{\frac{1}{T - t_0 - h} \sum_{s=t_0}^{T-h} \hat{\epsilon}_{s+h}^2}.$$

to compare the corresponding RMSE with those from viable alternative inflation forecast models. These alternatives include univariate models and multivariate models, where our univariate models are summarized in the first panel of Table 4. First amongst these univariate models is the random walk model, which since Atkeson and Ohanian (2001) is seen as one of the hardest models to beat when it comes to out-of-sample inflation prediction. Also, time-invariant autoregressive specifications for inflation, using lag orders between 1 and 4, are considered as parsimonious alternatives to (6).

The two models after those in Table 4 are variations on these parsimonious model specifications that incorporate time-variation in the model structure. The first of these is the TVP-AR model (13) with a maximum lag order of 4 quarters. In this specification we, firstly, allow the intercept, the autoregressive parameters as well as the error variance to break in the same manner as in our BMASB Phillips curve model and, then, construct a BMA across all possible lag order combinations. The other one is an inflation forecast model that has been successfully used by Stock and Watson (2007, 2008) to predict inflation. They propose an observed components model with stochastic volatility specifications for the unobserved component of inflation as well as the temporary deviation from it, i.e.,

$$\begin{aligned} y_t &= \beta_t + \sigma_t \varepsilon_t \\ \beta_t &= \beta_{t-1} + \omega_t \eta_t \\ \ln\sigma_t^2 &= \ln\sigma_{t-1}^2 + u_{1t} \\ \ln\omega_t^2 &= \ln\omega_{t-1}^2 + u_{2t}, \end{aligned} \tag{14}$$

where $\varepsilon_t \sim \text{NID}(0,1)$, $\eta_t \sim \text{NID}(0,1)$ and $u_t = (u_{1t}\ u_{2t})' \sim \text{NID}(\mathbf{0}, \rho I_2)$ with ρ a scalar parameter controlling the smoothness of the stochastic volatility processes, and where ε_t, η_t and u_t are independent. We follow Stock and Watson (2007) and set in (14) $\rho = 0.04$; Stock and Watson (2007) motivate their choice for ρ based on the fit of (14) for U.S. inflation rates over the 1955-2004 sample.

The remaining models in Table 4 all incorporate information from a range of additional regressors. These encompass a simple linear regression of the quarter-to-quarter inflation rate h quarters ahead on all 14 predictor variables described in the previous section plus four inflation lags, that is,

$$y_{t+h} = X_t'\beta + \sigma\varepsilon_t, \tag{15}$$

where $X_t = (1, x_{1t}, \ldots, x_{kt})'$ and $\varepsilon_t \sim \text{NID}(0,1)$. To deal with the curse of dimensionality in such a regression, we also estimate the β parameter in this regression using a ridge regression (shrinkage) estimator:

$$\hat{\beta} = \left(\sum_{t=1}^{T-h} X_t X_t' + \lambda I\right)^{-1} \left(\sum_{t=1}^{T-h} X_t y_{t+h}'\right) \tag{16}$$

and the scalar shrinkage parameter λ. For the latter we choose $\lambda = 10$, as De Mol *et al.* (2006) show that the degree of shrinkage should be proportional to the number of regressors to achieve the best forecasting performance in a data-rich context.

Further, we employ a version of our BMASB Phillips curve model *without* time-varying parameters and variance, i.e, (2) where we construct a Bayesian model average (BMA) across the different selected regressor combinations. Finally, we consider a bivariate VAR model of the inflation rate and real output growth as well as a Bayesian model average across all possible bivariate VAR models of inflation and one of our 14 economic predictor variables.

5.2 Out-of-Sample Results

All of the models discussed in Section 5.1 are used to generate quarter-on-quarter PCE deflator and GDP deflator inflation forecasts one-quarter ahead ($h = 1$) and one-year ahead ($h = 4$). These are evaluated by computing the corresponding RMSEs across three periods: 1980Q1-2008Q4, 1980Q1-1994Q4 and 1995Q1-2008Q4. The evaluation samples span a number of large events that potentially could have caused time-variation in the dynamics of inflation rates, like, e.g., the 'monetarist experiment' by the Federal Reserve under Volcker, the 'Great Moderation' in the mid-1980s and the 9/11 catastrophe in 2001. The forecasts of these models are based on posterior results of the model parameters and, if relevant, the latent variables computed using an expanding window of data, starting with 1960Q1-1979Q4 based on the original data vintages starting from 1979Q4. Finally, the resulting RMSEs based on the corresponding forecast errors are used to compute RMSE ratios relative to our BMASB Phillips curve model to asses how well, or not, they are doing in a real-time out-of-sample setting *vis-à-vis* our model, where a ratio smaller than 1 indicates that a model outperforms (4)–(6) and *vice versa*.

Tables 5 and 6 report in the first line the RMSEs for our BMASB Phillips curve model-based forecasts in case of the PCE and the GDP deflator inflation measures, respectively. Below that line, both tables report the ratio of the RMSE for each of the competing models as discussed in Table 4 relative to the RMSE of our BMASB Phillips curve model. When we focus on PCE deflator inflation first, see Table 5, it becomes quite striking how successful the BMASB Phillips curve forecasts are in comparison with the other models. Over the full 1980-2008 evaluation sample and the first sub-sample none of these can beat our Phillips curve specification (6) at the one-quarter and one-year ahead forecast. In the final sub-sample, only the USCV Stock and Watson (2007) model performs better at both forecast horizons although the difference in performance at one-quarter ahead forecasting is very small. The purely autoregressive and random walk specifications are not performing

Table 1: Marginal posterior probabilities of predictor variable selection

	PCE Deflator Inflation		GDP Deflator Inflation	
	$h = 1$	$h = 4$	$h = 1$	$h = 4$
INFL_t	0.095	0.000	0.000	0.106
INFL_{t-1}	0.638	0.441	0.063	0.194
INFL_{t-2}	0.103	0.132	0.236	0.131
INFL_{t-3}	0.349	0.192	0.225	0.000
ROUTP_t	0.263	0.559	0.000	0.000
RCONS_t	0.125	0.446	0.155	0.357
RINVR_t	0.273	0.397	0.229	0.275
PIMP_t	0.078	0.000	0.000	0.000
UNEMPL_t	0.270	0.288	0.197	0.202
HSTS_t	0.107	0.330	0.289	0.265
NFPR_t	0.000	0.771	0.358	0.113
OIL_t	0.286	0.048	0.367	0.158
FOOD_t	0.409	0.346	0.366	0.218
RAW_t	0.526	0.181	0.368	0.292
$\mathrm{M2}_t$	0.347	0.202	0.005	0.272
TS_t	0.047	0.145	0.000	0.314
YL_t	0.000	0.000	0.154	0.033
MS_t	0.948	0.898	1.000	1.000

Note: The table presents the marginal posterior inclusion probabilities in the predictive regression model (6) for $h = 1$ and $h = 4$ over the full sample, 1960Q1 – 2008Q4.

Variable mnemonics: INFL - PCE or GDP Deflator inflation; ROUTP - percentage quarterly change real GDP; RCONS- percentage quarterly change real personal consumption expenditures; RINVR - percentage quarterly change real residential investment; PIMP - percentage quarterly change import price deflator; UNEMPL - unemployment rate (% labor force); HSTS - log level housing starts; NFPR - percentage quarterly change non-farm payrolls; OIL - percentage quarterly change real oil spot price; FOOD - percentage quarterly change real food commodities price index; RAW - percentage quarterly change real raw materials price index; M2 - percentage quarterly change M2 monetary aggregate; TS - slope term structure level; YL - level term structure factor; MS - one-year ahead inflation expectations from the Michigan Consumer survey.

Table 2: Posterior model probabilities: PCE deflator inflation

Model	Prob. %
Forecast horizon: $h = 1$	
$\mathrm{INFL}_{t-1}, \mathrm{HSTS}_t, \mathrm{TS}_t, \mathrm{RAW}_t, \mathrm{MS}_t$	2.27
$\mathrm{INFL}_{t-1}, \mathrm{RAW}_t, \mathrm{MS}_t$	1.87
$\mathrm{INFL}_{t-1}, \mathrm{RAW}_t$	1.53
$\mathrm{INFL}_{t-1}, \mathrm{INFL}_{t-3}, \mathrm{RINVR}_t, \mathrm{FOOD}_t, \mathrm{RAW}_t, \mathrm{MS}_t$	1.47
$\mathrm{INFL}_{t-1}, \mathrm{INFL}_{t-3}, \mathrm{FOOD}_t, \mathrm{RAW}_t, \mathrm{MS}_t$	1.40
$\mathrm{INFL}_t, \mathrm{INFL}_{t-1}, \mathrm{M2}_t, \mathrm{OIL}_t, \mathrm{FOOD}_t, \mathrm{RAW}_t, \mathrm{MS}_t$	1.40
$\mathrm{INFL}_t, \mathrm{INFL}_{t-1}, \mathrm{PIMP}_t, \mathrm{M2}_t, \mathrm{OIL}_t, \mathrm{FOOD}_t, \mathrm{MS}_t$	1.33
$\mathrm{INFL}_{t-2}, \mathrm{INFL}_{t-3}, \mathrm{ROUTP}_t, \mathrm{OIL}_t, \mathrm{FOOD}_t, \mathrm{MS}_t$	1.27
$\mathrm{INFL}_{t-1}, \mathrm{INFL}_{t-3}, \mathrm{RCONS}_t, \mathrm{UNEMPL}_t, \mathrm{MS}_t$	1.20
$\mathrm{INFL}_{t-2}, \mathrm{INFL}_{t-3}, \mathrm{OIL}_t, \mathrm{MS}_t$	1.20
Forecast horizon: $h = 4$	
$\mathrm{INFL}_{t-3}, \mathrm{ROUTP}_t, \mathrm{RCONS}_t, \mathrm{RINVR}_t, \mathrm{HSTS}_t, \mathrm{NFPR}_t, \mathrm{MS}_t$	3.70
$\mathrm{INFL}_{t-1}, \mathrm{ROUTP}_t, \mathrm{RINVR}_t, \mathrm{NFPR}_t$	3.30
$\mathrm{INFL}_{t-1}\ \mathrm{ROUTP}_t, \mathrm{RCONS}_t, \mathrm{NFPR}_t, \mathrm{MS}_t$	3.20
$\mathrm{INFL}_{t-1}, \mathrm{ROUTP}_t, \mathrm{UNEMPL}_t, \mathrm{NFPR}_t, \mathrm{MS}_t$	2.10
$\mathrm{INFL}_{t-1}, \mathrm{RCONS}_t, \mathrm{UNEMPL}_t, \mathrm{NFPR}_t, \mathrm{FOOD}_t, \mathrm{MS}_t$	2.10
$\mathrm{INFL}_{t-2}, \mathrm{INFL}_{t-3}, \mathrm{ROUTP}_t, \mathrm{RCONS}_t, \mathrm{HSTS}_t, \mathrm{NFPR}_t, \mathrm{TS}_t, \mathrm{FOOD}_t, \mathrm{MS}_t$	2.00
$\mathrm{INFL}_{t-2}, \mathrm{ROUTP}_t, \mathrm{RINVR}_t, \mathrm{UNEMPL}_t, \mathrm{NFPR}_t, \mathrm{MS}_t$	1.80
$\mathrm{INFL}_{t-3}, \mathrm{ROUTP}_t, \mathrm{NFPR}_t, \mathrm{TS}_t, \mathrm{MS}_t$	1.80
$\mathrm{INFL}_{t-3}, \mathrm{ROUTP}_t, \mathrm{NFPR}_t, \mathrm{TS}_t, \mathrm{FOOD}_t, \mathrm{MS}_t$	1.80
$\mathrm{ROUTP}_t, \mathrm{RINVR}_t, \mathrm{HSTS}_t, \mathrm{OIL}_t, \mathrm{FOOD}_t, \mathrm{MS}_t$	1.80

Note: The table lists the ten models with the highest posterior probabilities and their posterior probabilities (%) for the quarterly PCE series, 1960Q1 – 2008Q4. See Table 1 for a description of the predictor variables.

Table 3: Posterior model probabilities: GDP deflator inflation

Model	Prob. %
Forecast horizon: $h = 1$	
$\mathrm{INFL}_{t-3}, \mathrm{ROUTP}_t, \mathrm{RCONS}_t, \mathrm{MS}_t$	4.00
$\mathrm{NFPR}_t, \mathrm{OIL}_t, \mathrm{MS}_t$	3.13
$\mathrm{HSTS}_t, \mathrm{FOOD}_t, \mathrm{MS}_t$	2.53
$\mathrm{FOOD}_t, \mathrm{RAW}_t, \mathrm{MS}_t$	2.40
$\mathrm{NFPR}_t, \mathrm{MS}_t$	2.33
$\mathrm{HSTS}_t, \mathrm{MS}_t$	2.00
$\mathrm{HSTS}_t, \mathrm{NFPR}_t, \mathrm{MS}_t$	1.73
$\mathrm{FOOD}_t, \mathrm{RAW}_t, \mathrm{MS}_t$	1.47
$\mathrm{OIL}_t, \mathrm{MS}_t$	1.40
$\mathrm{UNEMPL}_t, \mathrm{MS}_t$	1.40
Forecast horizon: $h = 4$	
$\mathrm{RINVR}_t, \mathrm{MS}_t$	5.30
MS_t	3.70
$\mathrm{OIL}_t, \mathrm{MS}_t$	3.70
$\mathrm{RINVR}_t, \mathrm{TS}_t, \mathrm{MS}_t$	3.00
$\mathrm{INFL}_{t-1}, \mathrm{INFL}_{t-2}, \mathrm{RAW}_t, \mathrm{MS}_t$	3.00
$\mathrm{HSTS}_t, \mathrm{MS}_t$	2.70
$\mathrm{INFL}_{t-1}, \mathrm{INFL}_{t-2}, \mathrm{M2}_t, \mathrm{MS}_t$	2.00
$\mathrm{RCONS}_t, \mathrm{UNEMPL}_t, \mathrm{HSTS}_t, \mathrm{TS}_t, \mathrm{MS}_t$	1.80
$\mathrm{INFL}_{t-1}, \mathrm{MS}_t$	1.70
$\mathrm{INFL}_{t-1}, \mathrm{INFL}_{t-2}, \mathrm{MS}_t$	1.70

Note: The table lists the ten models with the highest posterior probabilities and their probabilities (%) for quarterly GDP deflator series, 1960Q1 – 2008Q4. See Table 1 for a description of the predictor variables.

Table 4: Alternative univariate and multivariate models for forecasting inflation

name	description	specification
	univariate models	
RW	Random walk	$y_t = y_{t-1} + \varepsilon_t$
AR(1)	Autoregressive model of order 1	$y_t = \mu + \phi_1 y_{t-1} + \varepsilon_t$
AR(2)	Autoregressive model of order 2	$y_t = \mu + \sum_{i=1}^{2} \phi_i y_{t-i} + \varepsilon_t$
AR(3)	Autoregressive model of order 3	$y_t = \mu + \sum_{i=1}^{3} \phi_i y_{t-i} + \varepsilon_t$
AR(4)	Autoregressive model of order 2	$y_t = \mu + \sum_{i=1}^{4} \phi_i y_{t-i} + \varepsilon_t$
TVP-AR(4)	AR(4) with structural instability and BMA	see (13)
UCSV	Unobserved component model with SV	see (14)
Linear	Linear regression with all predictors	$y_{t+h} = X_t'\beta + \sigma\varepsilon_t$
Ridge	Ridge regression with $\lambda = 10$	see (16)
BMA	Bayesian model averaging	(6) with $\beta_t = \beta$, $\sigma_t = \sigma$ $\forall t$
	multivariate models	
VAR(4)	Bivariate VAR(4) with inflation & output growth	$Y_t = \mu + \sum_{i=1}^{4} \Phi_i Y_{t-i} + \epsilon_t$
BMA-VAR(4)	BMA of all possible bivariate VAR(4)	$Y_t = \mu + \sum_{i=1}^{4} \Phi_i Y_{t-1} + \epsilon_t$

Table 5: RMSE - PCE

	Horizon: $h = 1$			Horizon: $h = 4$		
	F	I	II	F	I	II
BMASB	0.39	0.37	0.41	0.43	0.42	0.44
			univariate models			
RW	1.23	1.26	1.18	1.20	1.29	1.12
AR(1)	1.19	1.20	1.17	1.12	1.14	1.10
AR(2)	1.15	1.17	1.14	1.10	1.14	1.07
AR(3)	1.10	1.12	1.09	1.10	1.13	1.07
AR(4)	1.10	1.12	1.09	1.10	1.13	1.07
TVP-AR(4)	1.14	1.16	1.12	1.20	1.29	1.11
UCSV	1.07	1.17	**0.99**	1.07	1.22	**0.95**
Linear	1.04	1.09	1.00	1.18	1.28	1.08
Ridge	1.06	1.06	1.06	1.03	1.06	1.00
BMA	1.05	1.07	1.04	1.26	1.41	1.09
			multivariate models			
VAR(4)	1.08	1.10	1.07	1.05	1.03	1.07
VAR-BMA(4)	1.05	1.05	1.06	1.11	1.20	1.01

Note: The table presents root mean square prediction error (RMSE) of the BMASB Phillips curve-type model (6), the first line, as well as RMSE ratios relative to it for different univariate and multivariate models, see Table 4, for the full 1980Q1-2008Q4 evaluation sample (F) and two subsamples (I: 1980Q1-1994Q4, II: 1995Q1-2008Q4) at one-quarter ($h = 1$) and one-year ($h = 4$) ahead forecasting horizons for inflation. **Bold** indicates when our BMASB Phillips curve forecasts are outperformed by any of the forecasts from the competing models.

Table 6: RMSE - GDP deflator

	Horizon: $h = 1$			Horizon: $h = 4$		
	F	I	II	F	I	II
BMASB	0.27	0.31	0.22	0.32	0.36	0.27
			univariate models			
RW	1.25	1.19	1.36	1.10	1.11	1.08
AR(1)	1.21	1.15	1.34	1.10	1.09	1.13
AR(2)	1.15	1.11	1.23	1.09	1.09	1.10
AR(3)	1.08	1.04	1.17	1.09	1.10	1.09
AR(4)	1.09	1.04	1.18	1.09	1.11	1.07
TVP-AR(4)	1.11	1.09	1.14	1.22	1.09	1.41
UCSV	1.11	1.16	1.03	1.10	1.23	**0.90**
Linear	1.09	1.04	1.19	1.23	1.26	1.18
Ridge	1.04	1.01	1.12	1.01	1.04	**0.97**
BMA	1.13	1.08	1.23	1.30	1.36	1.20
			multivariate models			
VAR(4)	1.07	1.05	1.13	1.02	1.02	1.02
VAR-BMA(4)	1.14	1.11	1.21	1.16	1.17	1.13

Note: See the notes for Table 5.

well, as our model clearly outperforms these in terms of RMSE. Furthermore, we see that models containing explanatory variables perform in general better than models which only use lagged inflation information for prediction.

The results for GDP deflator inflation are quite similar to those for PCE deflator inflation, see Table 6. BMASB Phillips curve model (4)–(6) is only outperformed by two model specification in the final sub-sample for one-year ahead forecasts. These two model specifications are the Ridge estimator approach and again the USCV model.

The general conclusion from Tables 5 and 6 is that the BMASB Phillips curve-type model (4)–(6) does really well for predicting different inflation series at different forecasts horizons. Only in the sample 1995-2008 the model is outperformed by the USCV specification of Stock and Watson (2007) at one-year ahead forecasting but in the sample 1980-1994 the BMASB Phillips curve-type model performs clearly better than this model. Therefore, our BMASB Phillips curve-type specification does capture very well the time-variation in both the correlation between inflation and activity measures (the 'Phillips curve correlation') as well as inflation dynamics itself. Several studies have shown that by 'sucking' in a lot of data in an efficient way, model averaging and ridge regression can be simple and effective ways to face future instability of unknown form. Our forecasting results, however, indicate that accounting for structural instability may improve forecast performance. The BMASB model (4)–(6), which allows simultaneously for model uncertainty and structural instability, overall has the best out-of-sample performance, stressing the roles of both kinds of uncertainty.

6 Conclusion

Forecasting inflation has become much more difficult over the last decades. As a consequence, Phillips curve forecasts, i.e., inflation forecasts using an economic activity variable, have not fared well in several empirical studies and hardly ever improve upon simple univariate forecasts. Nonetheless, Phillips curve-type of relationships remain the backbone of many macroeconomic models and are important to understand policy discussions about the business cycle and inflation.

The failure of Phillips curve forecasts has several sources. Firstly, there is uncertainty about which set of activity measures best describes the Phillips correlation at a particular time. Also, inflation dynamics have changed over time resulting in breaks in the mean and variance of inflation, which in return would have caused breaks in Phillips curve-type relationships. In this paper we have introduced a generalized, reduced form Phillips curve-type model that attempts to incorporate uncertainty about the above two elements. It allows for uncertainty in the inclusion of relevant predictor variables (model uncertainty), the estimation uncertainty in the model parameters (parameter uncertainty) and finally the stability in the value of the model parameters (structural instability).

We apply our approach to model and forecast PCE and GDP deflator inflation in the U.S. between 1960 and 2008, where the forecasts are for two forecast horizons, one-quarter ahead and one-year ahead. When we use our framework to model the post-WWII inflation dynamics in the U.S. we do find some interesting empirical facts. First, over the period 1960-2008 several structural breaks occurred in the relationship between US inflation and predictor variables which include its own lags, real activity and cost measures, and other macroeconomic indicators. These changes appear to coincide with important events such as the oil crises in the 1970s, changes in the monetary policy regime, and the economic recession at the beginning of 1990s. Next, we find less evidence for exogenous breaks in the variance of inflation than what usually is found in the literature. And by conditioning on a vast range of potential combinations of activity measures, our framework finds substantially lower degrees of, time-varying, persistence in the inflation deviations from its mean than in other studies.

Finally, we find that allowing for model uncertainty and structural breaks at the same time results in superior inflation forecasts. Our Phillips curve-type specification provides very accurate forecasts of U.S. inflation for the 1980-2008 period compared to a set of competing linear models and nonlinear models including the random walk. Only in the latter half of our forecast evaluation period, i.e., 1995-2008, the UCSV model of Stock and Watson (2007) seems to be a good alternative for one-year ahead forecasting.

References

Ang, A., G. Bekaert and M. Wei, 2007, Do Macro Variables, Asset Markets or Surveys Forecast Inflation Better?, *Journal of Monetary Economics* **54**, 1163–1212.

Ang, A., M. Piazzesi and M. Wei, 2006, What Does the Yield Curve Tell us about GDP Growth?, *Journal of Econometrics* **131**, 359–403.

Atkeson, A. and L. E. Ohanian, 2001, Are Phillips Curves Useful for Forecasting Inflation?, *Federal Reserve Bank of Minneapolis Quarterly Review* **25**, 2–11.

Benati, L., 2004, Evolving Post-World War II U.K. Economic Performance, *Journal of Money, Credit, and Banking* **36**, 691–718.

Carter, C. and R. Kohn, 1994, On Gibbs Sampling for State-Space Models, *Biometrika* **81**, 541–553.

Carter, C. and R. Kohn, 1997, Semiparametric Bayesian Inference for Time Series with Mixed Spectra, *Journal of the Royal Statistical Society, Series B* **B**, 255–268.

Christiano, L., M. Eichenbaum and C. Evans, 2005, Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy, *Journal of Political Economy* **113**, 1–45.

Clarida, R., J. Galí and M. Gertler, 2000, Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory, *Quarterly Journal of Economics* **115**, 147–180.

Clark, T. and M. W. McCracken, 2008, Averaging Forecasts from VARs with Uncertainty Instabilities, *Journal of Applied Econometrics* . Forthcoming.

Cogley, T. and A. M. Sbordone, 2008, Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve, *American Economic Review* **98**, 2101–2126.

Cogley, T., G. E. Primiceri and T. J. Sargent, 2009, Inflation-Gap Persistence in the U.S., *American Economic Journal: Macroeconomics* . Forthcoming.

Cogley, T. and T. J. Sargent, 2002, Evolving Post World War II U.S. Inflation Dynamics, *in* B. S. Bernanke and K. Rogoff (editors), *NBER Macroeconomics Annual 2001*, MIT Press, Cambridge, U.S.

Cogley, T. and T. J. Sargent, 2005, Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S., *Review of Economic Dynamics* **8**, 262–302.

De Mol, C., D. Giannone and L. Reichlin, 2006, Forecasting with a Large Number of Predictors: Is Bayesian Regression a Valid Alternative to Principal Components?, *Discussion Paper 5829*, Centre for Economic Policy Research.

Diebold, F. X., G. D. Rudebusch and B. Aruoba, 2006, The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach, *Journal of Econometrics* **131**, 309–338.

Estrella, A. and G. A. Hardouvelis, 1991, The Term Structure as a Predictor of Real Economic Activity, *Journal of Finance* **46**, 555–576.

Fama, E. F. and R. R. Bliss, 1987, The Information in Long-Maturity Forward Rates, *American Economic Review* **77**, 680–692.

Galí, J. and M. Gertler, 1999, Inflation Dynamics: A Structural Econometric Analysis, *Journal of Monetary Economics* **44**, 195–222.

Geman, S. and D. Geman, 1984, Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images, *IEEE Transaction on Pattern Analysis and Machine Intelligence* **6**, 721–741.

George, E. I. and R. E. McCulloch, 1993, Variable Selection Via Gibbs Sampling, *Journal of the American Statistical Association* **88**, 881–889.

Gerlach, R., C. Carter and R. Kohn, 2000, Efficient Bayesian Inference for Dynamic Mixture Models, *Journal of the American Statistical Association* **95**, 819–828.

Giordani, P. and R. Kohn, 2007, Efficient Bayesian Inference for Multiple Change-Point and Mixture Innovation Models, *Sveriges Riksbank Working paper* .

Giordani, P., R. Kohn and D. van Dijk, 2007, A Unified Approach to Nonlinearity, Outliers and Structural Breaks, *Journal of Econometrics* **137**, 112–137.

Gordon, R. J., 1997, The Time-Varying NAIRU and its Implications for Economic Policy, *Journal of Economic Perspectives* **11**, 11–32.

Groen, J. J. J. and H. Mumtaz, 2008, Investigating the Structural Stability of the Phillips Curve Relationship, *Working Paper 350*, Bank of England.

Haldane, A. and D. Quah, 1999, UK Phillips Curves and Monetary Policy, *Journal of Monetary Economics* **44**, 259–278.

Kim, S., S. Shephard and S. Chib, 1998, Stochastic volatility: Likelihood inference and comparison with ARCH models, *Review of Economic Studies* **65**, 361–393.

Koop, G. and S. Potter, 2007, Estimation and forecasting in models with multiple breaks, *Review of Economic Studies* pp. 763–789.

Kuo, L. and B. Mallick, 1998, Variable Selection for Regression Models, *The Indian Journal of Statistics* **60**(Series B), 65–81.

Levin, A. T. and J. M. Piger, 2004, Is Inflation Persistence Intrinsic in Industrial Economies?, *Working Paper 334*, European Central Bank.

Marcellino, M., J. H. Stock and M. W. Watson, 2006, A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series, *Journal of Econometrics* **127**(1-2), 499–526.

Pesaran, M.H., D. Pettenuzzo and A. Timmermann, 2006, Forecasting Time Series Subject to Multiple Structural Breaks, *Review of Economic Studies* **73(4)**, 1057–1084.

Sbordone, A. M., 2002, Prices and Unit Labor Costs: A New Test of Price Stickiness, *Journal of Monetary Economics* **49**, 265–292.

Sensier, M. and D. J. C. van Dijk, 2004, Testing for Volatility Changes in U.S. Macroeconomic Time Series, *Review of Economics and Statistics* **86**, 833–839.

Shephard, N.E., 1994, Partial non-Gaussian state-space models, *Biometrika* **81**, 115–131.

Sims, C. A. and T. Zha, 2006, Were there Regime Switches in U.S. Monetary Policy?, *American Economic Review* **96**, 54–81.

Stock, J.H. and M.W. Watson, 1999, Forecasting Inflation, *Journal of Monetary Economics* **44**, 293–335.

Stock, J.H. and M.W. Watson, 2007, Has Inflation Become Harder to Forecast?, *Journal of Money, Credit and Banking* **39**, 3–33.

Stock, J.H. and M.W. Watson, 2008, Phillips Curve Inflation Forecasts, *Working Paper 14322*, National Bureau of Economic Research.

Tanner, M.A. and W.H. Wong, 1987, The Calculation of Posterior Distributions by Data Augmentation, *Journal of the American Statistical Association* **82**, 528–550.

Wright, J. H., 2003, Forecasting U.S. Inflation by Bayesian Model Averaging, *International Finance Discussion Paper 780*, Board of Governors of the Federal Reserve System.

Appendices

A Gibbs Sampling Algorithm

Following the scheme in Section 3.1, the Gibbs sampler for BMASB Phillips Curve model (4)-(6) sequentially goes through the following steps:

Step 1: Sampling the variable selection parameters in $\boldsymbol{D}$

We follow Kuo and Mallick (1998), which is a simplified version of the George and McCulloch (1993) algorithm. Starting from the previous iteration, the variable D is drawn from its full conditional posterior distribution. We compute the value of the posterior density (11) for $\delta_j = 0$ and $\delta_j = 1$ given the value of the other parameters which results in p_{j0} and p_{j1}, respectively. The full conditional posterior is then given by

$$\Pr[\delta_j = 1|\bar{\theta}, B, S, K, D_{-j}, y, x] = \frac{p_{j1}}{p_{j0} + p_{j1}}, \tag{A.1}$$

for $j = 1, \ldots, k$, where $D_{-j} = (\delta_1, \ldots, \delta_{j-1}, \delta_{j+1}, \ldots, \delta_k)'$.

Step 2: Sampling $\boldsymbol{K_\beta}$

The (occasional) structural breaks in the regression parameters B, measured by the latent variable κ_{jt}, are drawn using the algorithm of Gerlach *et al.* (2000, Section 3), which derives its efficiency from generating κ_{jt} without conditioning on the states β_{jt}. The conditional posterior density for κ_{jt}, $t = 1, \ldots, T$, $j = 0, \ldots, k$ unconditional on B is

$$\begin{aligned}
&p(\kappa_{0t}, \ldots, \kappa_{kt}|K_{\beta,-t}, K_\sigma, S, \theta, y, x) \\
&\quad \propto p(y|K, S, \theta, x)p(\kappa_{0t}, \ldots, \kappa_{kt}|K_{\beta,-t}, K_\sigma, S, \theta, x) \\
&\quad \propto p(y_{t+h+1}, \ldots, y_{T-h}|y_{h+1}, \ldots, y_{t+h}, K, S, \theta, x) \\
&\qquad\quad p(y_{t+h}|y_{h+1}, \ldots, y_{t+h-1}, \kappa_1, \ldots, \kappa_t, K_\sigma, S, \theta, x)p(\kappa_{0t}, \ldots, \kappa_{kt}|K_{\beta,-t}, K_\sigma, S, \theta, x),
\end{aligned} \tag{A.2}$$

where $K_{\beta,-t} = \{\{\kappa_{js}\}_{j=0}^{k}\}_{s=1,s\neq t}^{T-h}$. The density $p(\kappa_{0t}, \ldots, \kappa_{kt}|K_{\beta,-t}, K_\sigma, S, \theta, x)$ is equal to $\prod_{j=0}^{k} \pi_j^{\kappa_{jt}}(1-\pi_j)^{1-\kappa_{jt}}$ since κ_{jt} does not depend on δ_j. The two remaining densities $p(y_{t+h+1}, \ldots, y_{T-h}|y_{h+1}, \ldots, y_{t+h}, K, S, \theta, x)$ and $p(y_{t+h}|y_{h+1}, \ldots, y_{t+h-1}, \kappa_1, \ldots, \kappa_t, K_\sigma, S, \theta, x)$ can easily be evaluated as shown in Gerlach *et al.* (2000, Section 3). Because κ_t can take a finite number of values, the integrating constant can easily be computed by normalization.

Step 3: Sampling the regression parameters in $\boldsymbol{B}$

The full conditional posterior density for the latent regression parameters B is computed using a simulation smoother. We follow Carter and Kohn (1994). The Kalman smoother is applied to derive the conditional mean and variance of the latent factors; for the initial value a multivariate normal prior with mean 0 is chosen. Note that in case the variable x_j is not selected, the full conditional distributions of κ_{jt} and β_{jt} for $t = 1, \dots, T - h$ do not depend on the data y and x. Hence, in this case we sample unconditionally from (4) and (5).

Steps 4 and 5: Sampling the variance parameters K_σ and S

To draw K_σ and S we want to follow a similar approach as above. As the model for $\ln \sigma_t^2$ does not result in a linear state space model the Kalman filter cannot be applied. Therefore, we apply the approach of Giordani and Kohn (2007) and rewrite the model (5)–(6) as

$$\begin{aligned} \ln(y_{t+h} - \beta_{0t} - \sum_{j=1}^{k} \delta_j \beta_{jt} x_{jt})^2 &= \ln \sigma_t^2 + u_t \\ \ln \sigma_t^2 &= \ln \sigma_{t-1}^2 + \kappa_{k+1,t} \eta_{k+1,t}, \end{aligned} \tag{A.3}$$

where $u_t = \ln \varepsilon_t^2$ has a log χ^2 distribution with 1 degree of freedom. We follow Carter and Kohn (1994, 1997), Shephard (1994) and Kim *et al.* (1998) and approximate the $\ln \chi^2(1)$ distribution by a finite mixture of normal distributions. We consider a mixture of five normal distributions such that the density of u_t is given by

$$f(u_t) = \sum_{s=1}^{5} \varphi_s \frac{1}{\omega_s} \phi((u_t - \mu_s)/\omega_s) \tag{A.4}$$

with $\sum_{s=1}^{5} \varphi_s = 1$. The appropriate values for μ_s, ω_s^2 and φ_s can be found in Carter and Kohn (1997, Table 1). In each step of the Gibbs sampler we simulate a component of the mixture distribution from the distribution of the mixing distribution. Given the value of the mixture component we can apply standard Kalman filter techniques. Hence, the variables K_σ and S can be sampled in a similar way as K_β and B in step 2 and 3.

Step 6: Sampling $\bar{\theta}$

Finally, to sample the parameters $\bar{\theta}$ we can use standard results in Bayesian inference. Hence, the probabilities π_j are sampled from Beta distributions and the variance parameters q_j^2 are sampled from inverted Gamma-2 distributions.

B Data Sources and Construction

Inflation rates

Our two dependent variables are inflation rates based on the gross domestic product (GDP) deflator as well as the personal consumption expenditures (PCE) deflator. Both measures get revised on a regular basis and we therefore do not retrieve our data from the usual data sources. Instead, we get the original vintages of the underlying data from the 'Real-Time Data Set for Macroeconomists' (RTDSM) at the Federal Reserve Bank of Philadelphia (http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data). The RTDSM proxies the original vintages for each quarter by selecting the data that was originally available around the middle of that quarter (as close as possible to the 15th day of the middle month of a quarter). Vintages of inflation rates are then constructed as the percentage quarterly changes of the respective deflator series.[11]

Explanatory variables

We use in this paper an extensive set of activity and expectations measures to model inflation dynamics. Like the aforementioned inflation rates, the bulk of these variables gets revised so we strive to use as much as possible the original vintages of underlying data. Some of the measures can be directly retrieved from the respective real-time databases, others need to be constructed.

Real output growth - ROUTP We take the original quarterly data vintages for GDP in volume terms from the RTDSM at the Federal Reserve Bank of Philadelphia. Based on these we construct real output growth rates, i.e., the percentage quarterly change in real GDP.

Real consumption growth - RCONS We take the original quarterly data vintages for real personal consumption expenditures (PCE) from the RTDSM at the Federal Reserve Bank of Philadelphia. Based on these we construct real consumption growth rates, i.e., percentage quarterly change in real PCE.

Real residential investment growth - RINVR We take the original quarterly data vintages for real residential investment from the RTDSM at the Federal Reserve Bank of Philadelphia. Based on these we construct real residential investment growth rates, i.e., percentage quarterly change in the real residential investment level.

Import price inflation - PIMP We take the original quarterly data vintages for the imports deflator from the RTDSM at the Federal Reserve Bank of Philadelphia. Based on these we construct import price inflation, i.e., percentage quarterly change in the imports deflator.

Non-farm payrolls growth rate - NFPR From the ALFRED® real-time database, we take as quarterly vintages those monthly data vintages of non-farm payrolls employment

[11] We define percentage quarterly change as 100 times the quarterly change of the logarithm of the original series.

that are closest to the middle of quarter. Then, we transform these data to the quarterly frequency through averaging; finally, the non-farm payrolls growth rate is constructed as the percentage quarterly change in non-farm payrolls.

Housing starts growth rate - HSTS We take the original quarterly data vintages of monthly housing starts from the RTDSM at the Federal Reserve Bank of Philadelphia. Then, we transform these data to the quarterly frequency through averaging; finally, the housing starts growth rate is constructed as the percentage quarterly change in housing starts.

M2 growth rate - M2 From the ALFRED® real-time database at the Federal Reserve Bank of St. Louis, we take as quarterly vintages those monthly data vintages of the M2 monetary aggregate that are closest to the middle of quarter. Then, we transform these data to the quarterly frequency through averaging; finally, the M2 growth rate is constructed as the percentage quarterly change in the M2 level.

Unemployment ratio - UNEMPL We take the original quarterly data vintages for unemployment as a percentage of the labor force (UNEMPL) from the RTDSM at the Federal Reserve Bank of Philadelphia.

Level term structure factor - YL This is a proxy for the level factor describing the dynamics in the term structure of interest rates. The term structure is approximated by seven interest rates: the 3-month Treasury bill rate, the 6-month Treasury bill rate, both from Global Financial Data (https://www.globalfinancialdata.com/), as well as the Fama and Bliss (1987) 1-year, 2-year, 3-year, 4-year and 5-year zero-coupon bond yields from the CRSP database at Wharton Research Data Services. These are monthly data, which are not revised as they are financial data. In order to get quarterly data we select the aforementioned interest rates at the end of the first month of a quarter. The level term structure factor equals the cross-sectional average across the above seven interest rates for each quarter.

Slope term structure factor - TS This is a proxy for the slope factor describing the dynamics in the term structure of interest rates. We use the same interest rates as for the level term structure factor - see above. These are monthly data; in order to get quarterly data we select the aforementioned interest rates at the end of the first month of a quarter. The slope term structure factor equals the spread between the 5-year zero-coupon bond yield and the 3-month T-bill rate for each quarter.

Real oil price inflation - OIL To construct real oil prices, we first retrieve nominal oil prices - for this we use the West Texas Intermediate oil spot price from Global Financial Data. Quarterly observations result by selecting in each quarter the observed oil spot price closest to the middle of the quarter; as these data are market prices they are not prone to revisions. Quarterly data vintages of real oil prices are then constructed by deflating the aforementioned oil spot price, which is unrevised, by either the GDP deflator or PCE deflator for that vintage, depending on which inflation rate one wants to model. Vintages of real oil price inflation are then equal to the percentage quarterly change in the constructed real oil price level.

Real food commodities inflation - FOOD Vintages of real food commodities inflation are constructed in a similar manner as those for real oil price inflation - see above. Only now the construction is based on the Commodities Research Bureau (CRB) Index of Foodstuffs commodity prices, which is based on the spot prices for butter, cocoa beans, corn, cottonseed oil, hogs, lard, steers, sugar and wheat. The CRB Foodstuffs price index is acquired through Global Financial Data.

Real raw industrial commodities inflation - RAW Vintages of real raw industrial commodities inflation are constructed in a similar manner as those for real oil price inflation - see above. Only now the construction is based on the CRB Index of Raw Industrials commodity prices, which is based on the spot prices for burlap, copper scrap, cotton, hides, lead scrap, print cloth, rosin, rubber, steel scrap tallow, tin, wool tops and zinc. The CRB Raw Industrials price index is acquired through Global Financial Data.

Reuters/University of Michigan Survey of Consumers' inflation expectations - MS The Reuters/University of Michigan Survey of Consumers asks members of the general public, amongst other, to give a quantitative assessment of expected inflation in a year's time. As this is a one-year ahead measure, we lag these series, which are never revised, with four-quarters as to make them properly real-time. The quarterly data are retrieved from http://www.sca.isr.umich.edu/main.php at the University of Michigan.