Real-Time Inflation Forecasting in a Changing World*

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

Richard Paap‡
Econometric Institute, Erasmus University Rotterdam

Francesco Ravazzolo§
Norges Bank

Preliminary Draft

Abstract

This paper revisits real-time forecasting of U.S. inflation based on Phillips curve-inspired linear regression models. Our innovation is to allow for both structural breaks in the regression parameters and the variance as well as uncertainty regarding which set of predictor variables one can include in these regressions ('model uncertainty'). Structural breaks are described by occasional shocks of random magnitude. The set of potential predictors includes lagged values of inflation, output series, interest rate series and money. Parameter estimation and forecasting are performed using a Gibbs sampling approach with Bayesian model averaging. We compare our approach with many alternative univariate and multivariate model specifications including a random walk model. Posterior results show that our model specification provides superior 1-step ahead and 4-step ahead forecasts for both CPI and GDP deflator inflation rates in terms of root mean squared prediction error. Also, the common finding of autonomous inflation volatility breaks is rejected by our approach: breaks in the conditional mean of inflation drive structural breaks in U.S. inflation measures.

Keywords: Real-time forecasting, Inflation, Structural Breaks, Model Uncertainty, Bayesian Model Averaging.
JEL Classification: C11, C22, C53, E31

---

*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@few.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 around the world 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. For example, Stock and Watson (1999) test the forecasting ability of several reduced form Phillips curve relationships for a number of U.S. inflation measures.

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.* (2007) 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) show that U.S. inflation is well described by a integrated moving average process with time-varying parameters and the out-of-sample performance of this model is hard to beat by alternative 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]

[1] Clark and McCracken (2007a,b), 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

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 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 predictor variables.

Given the multitude of potentially useful activity measures, approaches that compress the information in these variables in a parsimonious way could be useful in dealing with this type of model uncertainty. Stock and Watson (1999) consider two approaches: one is a forecast combination, or model average, of the different, possible choices of Phillips curve forecasting models. Another approach is based on the Stock and Watson (2002a,b) dynamic factor framework and uses a principal component extracted from all possible 'economic slack' variables as the real activity measure in a single Phillips curve-based model. The Stock and Watson (1999) out-of-sample inflation forecasting results based on the aforementioned approaches are favorable compared to traditional specifications, in particular in case of the factor-based approach. As a consequence, they claim that for post-WWII U.S. inflation forecasting based on Phillips curve relationships, the uncertainty due to structural breaks is second order to the uncertainty about the choice for real activity measure. 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 no-change forecasts.

Like Stock and Watson (1999) and Atkeson and Ohanian (2001), 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

variables, but in general univariate models seem to be difficult to beat by these VARs that allow for structural changes.

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 and error variance for each of the regression models that are combined within this model average. Hence, our forecasting procedure simultaneously incorporates the two major sources of uncertainty, which the literature has shown to be relevant for forecasting 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 2006. A range of predictor variables are considered in the modeling exercise, from real variables like, e.g., real GDP and consumption growth rates to nominal and financial variables, such as M1 growth and the slope of the term structure, 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 the oil crisis, changes in the monetary policy regime, and the economic recession at the beginning of 1990s. 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 model specification. We consider prior specification and discuss posterior simulation and forecasting. In Section 3 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 4 by comparing it to other univariate and multivariate model specifications. Finally, in Section 5 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.

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

[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.

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: Stock and Watson (1999) use both forecast combinations as well as a factor extracted from 132 indicator series, and Atkeson and Ohanian (2001) run a total of 132 different predictive regressions, in both cases with mixed success. In this paper, we will apply the following generalized Phillips curve specification as our predictive regression:

$$\begin{aligned} y_t =& \beta_0 + \sum_{i=1}^{k1} \beta_i^a a_{i,t} + \sum_{j=1}^{k2} \beta_j^e e_{j,t} + \sum_{l=1}^{L} \beta_l^y y_{t-l} + \sigma\varepsilon_t; \quad t = 1, \ldots, T \\ =& \beta_0 + \sum_{i=1}^{k} \beta x_{i,t} + \sigma\varepsilon_t; \quad i = 1, \ldots, k \end{aligned} \tag{1}$$

with $(x_{1,t} \cdots x_{k,t})' = (a_{1,t} \cdots a_{k1,t}\ e_{1,t} \cdots e_{k2,t}\ y_{t-1} \cdots y_{t-L})'$ and thus $k = k1 + k2 + L$. In (1), y_t is the inflation measure, defined as $y_t = \Delta \ln(P_t) = \ln(P_t) - \ln(P_{t-1})$ where P_t is a particular price index, the $a_{i,t}$'s are the $k1$ real activity and costs indicator variables, the $e_{j,t}$'s are $k2$ proxies of inflation expectations and ε_t is regression error term with $\varepsilon_t \sim \mathrm{N}(0, 1)$ and $\sigma > 0$. Clearly, the number of predictor variables in (1), k, will in practice be large; the aforementioned studies use up to 132 series, whereas we use in this paper up to 17 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 of our 1960-2006 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. Thus, next to the uncertainty about the inclusion of predictor variables, we need to include some form of time-variation in (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; this reflects uncertainty about which combination of indicator variables most accurately summarizes the impact of real activity, real costs and expectations on inflation dynamics.

- Structural breaks; both inflation itself and the Phillips curve correlation between inflation and indicator variables have changed over time.

To allow for model uncertainty we can expand our original generalized Phillips curve model (1) by introducing into it k binary variables s_j that determine the inclusion of variable x_{jt} in the regression model for $j = 1, \ldots, k$. This results in:

$$y_t = \beta_0 + \sum_{j=1}^{k} s_j \beta_j x_{jt} + \sigma_t \varepsilon, \quad t = 1, \ldots, T. \tag{2}$$

where y_t, $x_t = (x_{1t} \cdots x_{kt})'$ and ε_t are similar as in (1), and $S = (s_1 \ldots s_k)'$, $\quad s_j \in \{0, 1\} \quad \forall \quad j = 1, \ldots, k$, with $Pr[s_j = 1] = \lambda_j$ for $j = 1, \ldots, k$. The vector S 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 $i = (s_1, \ldots, s_k)_2$, and the intercept parameter β_0 is always included in each possible model.

Structural breaks, on the other hand, can simply be incorporated in (1) through time-variation in the regression parameters:

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

where the time varying parameters are defined as

$$\begin{aligned} \beta_{jt} &= \beta_{j,t-1} + \kappa_{jt} \eta_{jt}, \qquad j = 0, \ldots, k, \\ \ln \sigma_t^2 &= \ln \sigma_{t-1}^2 + \kappa_{k+1,t} \eta_{k+1,t}. \end{aligned} \tag{4}$$

with

$$\eta_t = \begin{pmatrix} \eta_{0,t} \\ \eta_{1,t} \\ \vdots \\ \eta_{k,t} \\ \eta_{k+1,t} \end{pmatrix} \sim \mathrm{N}(0, Q); \quad Q = \begin{pmatrix} q_0^2 & 0 & 0 & \cdots & 0 \\ 0 & q_1^2 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & q_{k+1}^2 \end{pmatrix}.$$

In (4), the vector $\kappa_t = (\kappa_{0,t}\ \kappa_{1,t} \cdots \kappa_{k+1,t})'$ is a $((k+2) \times 1)$ vector of unobserved and uncorrelated 0/1 processes with $\Pr[\kappa_{jt} = 1] = \pi_j$ for $j = 0, ..., k+1$. This specification implies that a regression parameter β_{jt} in (4) remains the same as its previous value $\beta_{j,t-1}$ unless $\kappa_{jt} = 1$ in which case it changes with η_{jt}, see, e.g., Koop and Potter (2007) and Giordani *et al.* (2007) for a similar approach. Stochastic structural breaks in the variance parameter $\ln \sigma_t$ comply to a similar structure as the β_{jt} parameters. The flexibility of the specification in (4) 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 on time. The occurrence of a change is described by the latent binary random variable κ_{jt}, while the magnitude of the change is determined by η_{jt}, which is assumed to be normally distributed with mean zero. Another attractive property of (4) is that the changes in the individual parameters are not restricted to coincide but are allowed to occur at different points in time.

Allowing in particular for the possibility of separate variance breaks seems important, as there is some evidence that macroeconomic time series have experienced over the post-WWII period variance breaks 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 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. Finally, Sims and Zha (2006) 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.

Ultimately, we want to use a reduced form Phillips curve specification for inflation that simultaneously incorporates model uncertainty and the possibility of structural breaks. By combining the two previously discussed extensions of our basic model (1) we can easily do that, i.e.,

$$y_t = \beta_{0t} + \sum_{j=1}^{k} s_j \beta_{jt} x_{jt} + \sigma_t \varepsilon_t, \quad t = 1, \ldots, T. \tag{5}$$

where $S = (s_1 \dots s_k)'$, $\quad s_j \in \{0,1\} \quad \forall \quad j = 1, \dots, k$ as in (3), and the parameter time-variation has a set-up identical to (4). For parameter inference in (5), we opt for a Bayesian approach. Firstly, such an approach will provide us with the posterior distribution of the unobserved κ_t processes for $t = 1, \dots, T$, which can be used to infer on the timing of structural breaks. Also, Bayesian inference on S leads to posterior probabilities of the 2^k possible models that can be used for Bayesian model averaging to incorporate model uncertainty into a single inflation forecast. By definition, κ_t in (5) does not depend on S. Thus, the value of κ_t can be different across different values of S and thus breaks can occur in different parameters and at different time periods across different models.

2.3 Prior Specification and Posterior Simulation

The parameters in the model (5)–(4) are the inclusion variable $S = (s_1, \dots, s_k)'$, the structural break probabilities $\pi = (\pi_0, \dots, \pi_{k+1})'$ and the vector of variances of the size of the breaks $q = (q_0^2, \dots, q_{k+1}^2)'$. We collect the model parameters in a $(3k+4)$-dimensional vector $\theta = (S, \pi, q)'$.

To facilitate the posterior simulation we make use of independent conjugate priors. For the variable inclusion parameters we take the following prior distribution

$$\Pr[s_j = 1] = \lambda_j \qquad \text{for } j = 1, \dots, k. \tag{6}$$

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

$$\pi_j \sim \text{Beta}(a_j, b_j) \qquad \text{for } j = 0, \dots, k+1. \tag{7}$$

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

$$q_j^2 \sim \text{IG-2}(\nu_j, \delta_j) \qquad \text{for } j = 0, \dots, k+1, \tag{8}$$

where ν_j, δ_j, $j = 0, \dots, k+1$, are parameters which can be chosen to reflect the prior beliefs about the variances. Realistic values of the parameters in the different prior distributions depend on the problem at hand.

The joint prior specification $p(\theta)$ is given by the product of the priors specification in (6)–(8).

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 $B = \{\beta_t\}_{t=1}^T$, with $\beta_t = (\beta_{0t}, \beta_{1t}, \ldots, \beta_{kt})'$, $R = \{\sigma_t^2\}_{t=1}^T$, and $K = \{\kappa_t\}_{t=1}^T$ 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

$$p(y, B, R, K|x, \theta) = \prod_{t=1}^{T} p(y_t|S, x_t, \beta_t, \sigma_t^2) p(\beta_t|\beta_{t-1}, \kappa_t, q_0^2, \ldots, q_k^2)$$
$$p(\ln \sigma_t^2 | \ln \sigma_{t-1}^2, \kappa_{k+1,t}, q_{k+1}^2) \prod_{j=0}^{k+1} \pi_j^{\kappa_{jt}} (1-\pi_j)^{1-\kappa_{jt}}, \quad (9)$$

where $y = (y_1, \ldots, y_T)$ and $x = (x_1', \ldots, x_T')'$. The elements $p(y_t|S, x_t, \beta_t, \sigma_t^2)$, $p(\beta_t|\beta_{t-1}, \kappa_t, q_0^2, \ldots, q_k^2)$, $p(\ln \sigma_t| \ln \sigma_{t-1}, \kappa_{k+1,t}, q_{k+1}^2)$ are normal density functions, which follow directly from the model specification (5)-(4).

If we combine (9) together with the prior density $p(\theta)$, we obtain the posterior density function

$$p(\theta, B, R, K|y, x) \propto p(\theta) p(y, B, R, K|x, \theta). \quad (10)$$

To derive the Gibbs sampler we combine 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 $\theta = (S, \bar{\theta})$ with $\bar{\theta} = (\pi, q)'$ and $K_\beta = \{\kappa_{0t}, \ldots, \kappa_{kt}\}_{t=1}^T$ and $K_\sigma = \{\kappa_{k+1,t}\}_{t=1}^T$, the sampling scheme can be summarized as follows:

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

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

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

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

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

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

Step 1 is done similarly to and Kuo and Mallick (1998), which is a simplified version of the George and McCulloch (1993) algorithm. Starting from the previous iteration, the variable S is drawn from its full conditional posterior distribution. We compute the value of the posterior distribution (10) for $s_j = 0$ and $s_j = 1$ given the value of the other parameters which results in in p_{j0} and p_{j1}, respectively. The full conditional posterior is then given by

$$\Pr[s_j = 1 | y, x, \bar{\theta}, B, R, K, S_{-j}] = \frac{p_{j1}}{p_{j0} + p_{j1}}, \tag{11}$$

for $j = 1, \ldots, k$, where $S_{-j} = (s_1, \ldots, s_{j-1}, s_{j+1}, \ldots, s_k)'$.

The (occasional) structural breaks, measured by the latent variable κ_{jt}, are drawn in step 2 using the algorithm of Gerlach *et al.* (2000), 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

$$\begin{aligned} & p(\kappa_{0t}, \ldots, \kappa_{kt} | K_{\beta,-t}, K_\sigma, S, R, \bar{\theta}, y, x) \\ & \quad \propto p(y | K, S, R, \bar{\theta}, x) p(\kappa_{0t}, \ldots, \kappa_{kt} | K_{\beta,-t}, K_\sigma, S, R, \bar{\theta}, x) \\ & \quad \propto p(y_{t+1}, \ldots, y_T | y_1, \ldots, y_t, K, S, R, \bar{\theta}, x) \\ & \qquad\quad p(y_t | y_1, \ldots, y_{t-1}, \kappa_1, \ldots, \kappa_t, K_\sigma, S, R, \bar{\theta}, x) p(\kappa_{0t}, \ldots, \kappa_{kt} | K_{\beta,-t}, K_\sigma, S, R, \bar{\theta}, x), \end{aligned} \tag{12}$$

where $K_{\beta,-t} = \{\{\kappa_{js}\}_{j=0}^k\}_{s=1, s\neq t}^T$. The density $p(\kappa_{0t}, \ldots, \kappa_{kt} | K_{\beta,-t}, K_\sigma, S, R, \bar{\theta}, x)$ is equal to $\prod_{j=0}^k \pi_j^{\kappa_{jt}} (1 - \pi_j)^{1-\kappa_{jt}}$ since κ_{jt} does not depend on s_j. The two remaining densities $p(y_{t+1}, \ldots, y_T | y_1, \ldots, y_t, K, S, R, \bar{\theta}, x)$ and $p(y_t | y_1, \ldots, y_{t-1}, \kappa_1, \ldots, \kappa_t, K_\sigma, S, R, \bar{\theta}, x)$ can easily be evaluated as shown in Gerlach *et al.* (2000). Because κ_t can take a finite number of values, the integrating constant can easily be computed by normalization.

The full conditional posterior density for the latent regression parameters B in step 3 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 $t = 1, \ldots, T$ do not depend on the data y and x. Hence, in this case we sample unconditionally from the generating process and the binary random process for κ_{jt}.

To draw K_σ and R in steps 4 and 5 we want to follow the same approach. As the model for $\ln \sigma_t$ 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)–(4) as

$$\begin{aligned} \ln(y_t - \beta_{0t} - \sum_{j=1}^k s_j \beta_{jt} x_{jt})^2 &= \ln \sigma_t^2 + u_t \\ \beta_{jt} &= \beta_{j,t-1} + \kappa_{jt} \eta_{jt}, \qquad j = 0, \ldots, k, \\ \ln \sigma_t^2 &= \ln \sigma_{t-1}^2 + \kappa_{k+1,t} \eta_{k+1,t} \end{aligned} \tag{13}$$

where $u_t = \ln \varepsilon_t^2$ has a log χ^2 distribution with 1 degree of freedom. We follow Carter and Kohn (1994), Carter and Kohn (1997), Shephard (1994) and Kim *et al.* (1998) who show that the $\ln \chi^2(1)$ distribution can be approximated very accurately 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

$$f(u_t) = \sum_{s=1}^{5} \varphi_s \frac{1}{\omega_s} \phi((u_t - \mu_s)/\omega_s). \tag{14}$$

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 R can be sampled in a similar way as K_β and B in step 2 and 3.

Finally, to sample the parameters $\bar{\theta}$ in step 6 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 and λ^2 are sampled from inverted Gamma-2 distributions.

The purpose of model (5) is to have a generalized, reduced form Phillips curve model for forecasting inflation that incorporates uncertainty about both the appropriate indicator 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 one-step ahead predictive density of y_{T+1} at time T conditional on y, x and x_{T+1} is given by

$$\begin{aligned} p(y_{T+1}|y, x, x_{T+1}) = \iiint \sum_{S} \sum_{K} \sum_{\kappa_{T+1}} & p(y_{T+1}|S, x_{T+1}, \beta_{T+1}, \sigma_{T+1}) \\ & p(\beta_{T+1}|\beta_T, \kappa_{T+1}, q) p(\sigma_{T+1}|\sigma_T, \kappa_{T+1}, q) \prod_{j=0}^{k+1} \pi_j^{\kappa_{j,T+1}} (1-\pi_j)^{1-\kappa_{j,T+1}} \\ & p(B, K, R, S, \bar{\theta}|y, x) dB dR d\bar{\theta}, \end{aligned} \tag{15}$$

where $p(y_{T+1}|S, x_{T+1}, \beta_{T+1}, \sigma_{T+1})$ and $p(\beta_{T+1}|\beta_T, \kappa_{T+1}, q)$ and $p(\sigma_{T+1}|\sigma_T, \kappa_{T+1}, q)$ follow directly from (4) and where $p(B, K, R, S, \bar{\theta}|y, x)$ is the simulated posterior density. In (15), we essentially compute a weighted average over all possible model specifications in (5) by averaging out over the posterior distribution of S, with weights equal to the posterior model probabilities. Uncertainty regarding the timing of structural breaks is reflected in (15) by the posterior distribution of the in-sample breaks K. Computation of such a predictive density is straightforward using the aforementioned Gibbs draws. So in case of (15), we simulate in each Gibbs step y_{T+1} using (5)–(4) as the data generating process, where we replace the parameters and the latent variables by the draw from the posterior distribution. As point estimate we then use the resulting posterior median.

3 The (In-)Stability of U.S. Inflation Dynamics

In this section we apply our framework to model the post-WWII behavior of a number of U.S. inflation measures. In Section 3.1 we discuss the data we use. Section 3.2 presents and discusses the characterization of U.S. inflation dynamics that results from applying our generalized Phillips curve model (5) on our data.

3.1 Data

We will consider in this paper two measures of inflation in the United States for a quarterly sample from 1960Q1 to 2006Q3; these are the quarterly log changes in the seasonally adjusted Consumer Price Index (CPI) and the Gross Domestic Product (GDP) deflator. Potentially there is wide array of potential predictors for inflation that would 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 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 ten series.

The set of potential predictors contains quarterly log changes of real GDP (ROUTPUT), real consumption (RCONS), real residential investment (RINVRESID), the import deflator (PIMPORT), the industrial production index (IP), the quarterly rate of unemployment (EMPLOY) and house starts (HSTARTS). These series provide information about either the degree of excess demand in the economy or about the real costs that firms face. In addition, we use two nominal series: the M1 money aggregate (M1) growth rate and the slope of the term structure (TS) measured as the difference between the 10-year Treasure Bond yield and the 3-month Treasury Bill interest rate at the end of quarter. Both variables can be seen as informative about the current and future state of the economy; M1 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 M1 growth indicates increased spending by households and firms). The slope of the term structure is often seen as a good predictor for turning points in the business cycle (see, for example, Estrella and Hardouvelis (1991)). In addition to these real activity and costs indicator series, basically the $a_{i,t}$ series in (1), we use one-year ahead inflation expectations from the University of Michigan Survey of Consumers (MS) as the expectations measure for our generalized Phillips curve model. 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.

The CPI price index, the 10-year Treasury Bond yield and the 3-month Treasury Bill interest rate are obtained from the Federal Reserve Bank of St. Louis' FRED® database, whereas the survey-based inflation expectations are obtained from the University of Michigan. These series are by definition real-time in nature. All the remaining series are periodically revised and thus we need to retrieve the original data vintages. The M1 growth rate is based on the original M1 data retrieved from Federal Reserve Bank of St. Louis' ALFRED® real-time database. The other variables, including the GDP deflator data, we got from the Federal Reserve Bank of Philadelphia's Real-Time Data Set for Macroeconomists (RTDSM).

3.2 Full-Sample Inflation Characteristics

In this subsection we conduct an *ex-post* analysis of the relevance of the different predictor variables and possible structural breaks in the different regression parameters within our generalized Phillips curve model that incorporates model uncertainty and occasional structural breaks, i.e., (5)–(4). 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 those purposes, we can for now suffice with the final, revised, data using the complete sample period from the first quarter 1960 until third quarter 2006 for the CPI series.

Table 1 provides the posterior mean of the inclusion parameters $s_j \forall j = 1, \ldots, k$ in (5); essentially these numbers reflect on average the proportion of times a variable is selected. In case of CPI inflation, see the first column in Table 1, the first four lags are always included in the range of potential specifications. The real activity and cost indicator variables are in most frequently selected to model CPI inflation,in particular real consumption, unemployment, house starts and the term structure slope have posterior inclusion probabilities close to one. Industrial production and real residential investment, however, cannot really be considered as important drivers of CPI inflation. Finally, inflation expectations are also important for understanding CPI inflation dynamics.

The posterior probabilities for predictor variable inclusion in case of GDP deflator inflation are slightly different; see the second column in Table 1. In this case, the impact of real activity on GDP deflator dynamics seems to be best captured by real GDP growth, whose inclusion probability is close to one. Real residential investment, M1 growth, unemployment and housing starts only appear to play a role in a minority of the range of specifications. As in the case of CPI inflation, lags are important to model GDP deflator inflation, in particular the first three lags. Also, inflation expectations appear again to be a major driver of inflation dynamics.

Next, we turn to the posterior results for the regression parameters to analyze the occurrence of structural breaks. Figure 1 displays for CPI inflation the posterior means of the latent binary variable κ_{jt} governing the occurrence of changes in the regression parameters, together with the associated posterior mean of β_{jt}, for $j = 0, \ldots, k$. The posterior mean of β_{jt} is conditional on inclusion of the jth variable, that is $s_j = 1$. We also show the 25th and 75th percentiles of the posterior distributions in the graphs. Several conclusions can be drawn from these graphs. First of all, the posterior means of κ_{jt} often show quite erratic, 'spiky' behavior in some subperiods, suggesting that the probabilities of structural breaks in the parameters vary considerably from one period to the next.[4] This occurs for two reasons. On the one hand, κ_{jt} can be different across different values of S, such that breaks can occur at different times across models. On the other hand, in case a break is estimated to have occurred in a certain month, the posterior probability of a break in the next month will be much lower.

The second conclusion which can be drawn from the graphs in Figure 1 is that despite the volatile behavior of the break probabilities, three periods with considerable probability mass for structural change can be identified. These periods are 1974-1975, 1979-1982, and the period after 1991. The oil price crisis of the 1970s and resulting polotical problems provide a possible explanation for the first break period. The second break period coincides with 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

[4]Recall that the posterior mean of κ_{jt} is identical to the posterior probability of a break occurring in the regression parameter for the jth variable x_{jt} at time t.

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 Dijk (2004).

The third observation follows from the patterns of the individual posterior means of the regression parameters in Figure 1. The intercept parameter β_{0t} has a very unstable pattern which closely follows CPI inflation patterns of those years. The posterior probability of breaks κ_{0t} is high and persistent over this period. This is not a surprise as the intercept is likely to change if one of the regression parameters changes. The unemployment rate is the only other variable that shows high instability during the "monetarist experiment" period, suggesting that a large part of the instability of this period can only be modeled by a time varying constant. The lags of CPI inflation show the most instability around the oil crises. Also note that the stock market crash in October 1987 gives rise to an isolated jump in the break probability for these variables. Instability in real activity and costs indicator variables is higher at beginning of 1990s, suggesting that the 'Great Moderation' coincided with a change in the slope of the Phillips curve. The Michigan survey-based inflation expectations measure appears to have had a stable impact on CPI inflation dynamics. Finally, Figure 2 shows the posterior probabilities of a break in the disturbance variance of (5) for CPI inflation together with the posterior mean of this variance over time. The posterior probabilities are rather small. Hence, allowing for instability in the regression coefficients seem to be more important than instability in the volatility, which seems to be quite stable.

For GDP deflator inflation, the posterior means of the break parameter κ_{jt} and the corresponding posteriors means of β_{jt} for $j = 0, \ldots, k$ are reported in Figure 3. As in the case of CPI inflation, the bulk of the probability mass for a structural break is concentrated in the 1974-1975 and 1979-1982 periods. However, after 1982 there is not a lot of evidence for a structural break in the 1990s, which is in contrast to the results for CPI inflation. Another difference relative to the CPI inflation case, is that the intercept term in (5) for GDP deflator inflation has been very stable over the sample. This might indicate that any instability in this inflation measure has not so much been due to level shifts in any of the series, but rather due to changes in the correlation of inflation with the potential indicator variables (the 'Phillips curve slope'). More specifically, a lot of the change parameters appear to have taken place in a smoother manner than in case of CPI inflation, as witnessed, for example, by the posterior means for the autoregressive parameters in Figure 3. Also of interest is the difference in the impact of the term structure slope and inflation expectations. The term structure slope seems to have been a highly instable predictor for GDP deflator inflation, much more than for CPI inflation. And in case of inflation expectations, the impact on GDP deflator inflation has fluctuated over time, which is opposite to what we found for CPI inflation.

Figure 4 shows similar posterior probabilities for GDP deflator inflation as for CPI inflation in Figure 2. So despite the inflation measure, our generalized Phillips curve model (5) does not provide the same evidence for exogenous variance breaks in inflation as in other a studies, such as in Cogley and Sargent (2005). This might suggest that this common finding of autonomous inflation volatility breaks is likely to be the result of considering a too narrow range of potential conditioning variables for the conditional mean of inflation.

4 Real Time Prediction of U.S. Inflation Rates

4.1 Forecasting procedure

The starting point of our forecasting exercise is the model in (5)–(4), which we will refer to as BMASB. The y_t variable is either the quarterly log change of seasonally adjusted CPI or the GDP deflator. We use as potential explanatory variables x_t the first 4 lags of the respective inflation measure and all the other predictors discussed in the previous section.

We first compute posterior results of the model parameters and the latent variables for the sample 1960Q1-1979Q4. The forecast period runs from 1980Q1 through 2006Q3. Therefore, we use the original data vintages starting from 1979Q4. The vintages of the RTDSM are dated to reflect the information available around the middle of each quarter. In each vintage for macroeconomic variables in t the observations run up to period $t-1$. The Michigan survey is lagged with four quarters, as it concerns a one-year ahead inflation expectation. In the estimation, we update the posterior estimates of model parameters for each period using an expanding window. For computational reasons we obtain multi-step ahead forecasts through direct forecasting.[5]

We consider two forecast horizons: the next quarter ($h = 1$) and four quarters ahead ($h = 4$). For $h = 4$ the forecasting sample is reduced with four observations and the first forecast concerns 1981Q1. As stated before we consider a quadratic loss function and use posterior predictive means as point forecasts. To evaluate the

[5]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).

out-of-sample point forecasts we compute the Root Mean Squared Prediction Error (RMSPE) over the full forecasting period, and two subsamples running from 1980Q1 (1981Q1 for $h = 4$) to 1994Q4 and from 1995Q1 to 2006Q3, respectively.

To evaluate the prediction accuracy of our BMASB model, we compare it with several alternatives. These alternatives include univariate models and multivariate models. The univariate models are given in the first panel of Table 2. The IMA(1,1) is also used by Stock and Watson (2007) to predict inflation. The multivariate models we use are summarized in the second panel of Table 2.

We use weak informative priors with mean value around zero for the parameters of all the models to shrink parameters to zero. If possible we use analytical expression for marginal likelihoods, and posterior and predictive densities.

4.2 Out-of-Sample Results

We apply the methodology in Section 2 to forecast quarterly log changes for US CPI and GDP deflator over the sample period from 1980Q1 to 2006Q3 and two forecasting horizons, one-quarter ahead ($h = 1$) and four-quarter ahead ($h = 4$). We also consider two subperiods, 1980Q1–1994Q4 and 1995Q1–2006Q3. The initial three observations of the evaluation samples are discarded in the analysis of the 4-step ahead forecast horizon.

Tables 3 and 4 report the root mean squared prediction errors based on posterior mean forecasts for the CPI and the GDP deflator inflation measures, respectively. The first line show the results for our model (BMASB), while the remaining lines report the results for the competing models as discussed in Table 2.

The general conclusion is that BMASB does quite well for different inflation series and different forecasts horizons. When we focus on CPI inflation first, BMASB is the best model for 1-step ahead prediction. BMASB gives forecasts which are 20% more accurate than the random walk over the full sample in term of Root Mean Square Prediction Error (RMSPE), and around 5% better than any other competitors, the Linear Model, the regression model with a Ridge estimator, the BMA and the BMA-VAR(4) approach. Splitting the sample in two subperiods, we see that some models are doing quite well in the first subperiod. Similar to Stock and Watson (2007) we find that the predictability is much lower in the second part of the sample. The Ridge regression gives the most accurate forecasts over this subperiod, but the BMASB approach provides very similar results.

For the 4-step ahead forecasts, results are similar. BMA is now the best approach over the full sample, but BMASB is very close. BMA does very well on the first subperiod, but relatively poorly in the second subperiod which was also the case for the 1-step ahead predictions. The BMASB approach provides the smallest RMSPE in the second subperiod, where the other models do not seem to be very accurate.

When we decompose our BMASB in model averaging (BMA) and structural instability (SB), we find that SB does very poorly. SB is not accurate in the first sample, and only at the end of the second sample period its forecast accuracy improves. This may be explained by the complexity of the models and the small sample we use for parameter estimation. As initial in-sample period we use only 87 observations, which is probably not enough to estimate a model with 16 time-varying parameters. It is difficult to distinguish between between noise and signal. When the sample size increases, structural instability can be modeled more adequately. Note that allowing for variable selection helps to reduce the number of time-varying parameters as not all predictors are selected in all models.

Adding multivariate information helps, but it is not the best strategy. The VAR(4) performs similarly to AR models. The BMA-VAR(4) approach gives sometimes better results than using an individual VAR(4), and it is quite accurate in the first subperiod for 1-step ahead forecasting. It is however outperformed by the BMASB approach.

Results are similar when we focus on the GDP deflator. BMASB is the best model for both forecast horizons and all samples. The linear regression model, the univariate and multivariate BMA approach and the Ridge regression are close competitors for 1-step ahead forecasting, especially for the full sample and the first subsample. When we focus on 4-step ahead forecasting results are a bit different. All the models are less accurate, and the random walk becomes more difficult to beat. BMASB give statistical gains of about 10% over a random walk, while other models have lower gains. As in Clark and McCracken (2007a) model averaging of multivariate models is quite accurate, but gains are very low compared to the use of the random walk specification.

In sum, we find evidence that the predictability of inflation is lower in the last years of our sample. We associate this deterioration to the change in behavior of inflation over time due to a high degree of structural instability. Therefore, constructing models that cope with this unknown form of instability seem to be very useful for forecasting. Both model averaging and Ridge regression which are shown in several studies to be simple and effective ways to face future instability provide accurate results. On contrary, models that try to explicitly model instability, such as the IMA of Stock and Watson (2007) or our SB approach, perform poorly in our forecast exercises. We expect these models to have problems to distinguish sign and noise when the estimation sample is limited. However, a model as (5)-(4), which allows simultaneously for model uncertainty and structural instability, gives the best out-of-sample forecasts, stressing the roles of both kinds of uncertainty.

5 Conclusion

Forecasting inflation has become much more difficult over the last decades. To forecast inflation one has to deal with different sources of uncertainty. The succes of a forecasting exercise depends on how well a forecaster can deal with these sources of uncertainty. In this paper we have introduced a forecasting framework which incorporate three sources of uncertainty at the same time. 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).

The model consists of a linear regression model where the regression parameters and the variance are subject to structural change. Furthermore, we allow for uncertainty about which predictors to include in the model. Finally, our Bayesian approach deals with model and parameter uncertainty in the forecasting exercise. We apply our approach to forecast two measure of US inflation, CPI and GDP deflator with two forecast horizons, 1 quarter ahead and 4 quarters ahead.

Several conclusions can be drawn from the empirical results. First, over the period 1960-2006 several structural breaks occurred in the relationship between US inflation and predictor variables which include its own lags, growth measures and other macroeconomic indicators. These changes appear to be caused by important events such as the oil crisis, changes in monetary policy, and the economic recession at the beginning of 1990s. Second, we find that allowing for model uncertainty and structural breaks at the same time results in superior forecasts. We find that our specification provide very accurate forecasts of US inflation over the last 25 years compared to a set of competing linear models and nonlinear models including the random walk. The evidence is robust to different measures of US inflation and forecast horizons.

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.

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, 2007a, Averaging Forecasts from VARs with Uncertainty Instabilities, *Working paper* .

Clark, T. and M. W. McCracken, 2007b, Forecasting with small macroeconomic VARs in the presence of instability, *in* M. Wohar and D. Rapach (editors), *Forecasting in the Presence of Structural Breaks and Model Uncertainty*, Elsevier.

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

Cogley, T., G. E. Primiceri and T. J. Sargent, 2007, Inflation-Gap Persistence in the U.S., *Mimeo*, University of California at Davis, Northwestern University and NYU.

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.

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

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*, Bank of England. Forthcoming.

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, 2002a, Forecasting Using Principal Components from a Large Number of Predictors, *Journal of the American Statistical Association* **97**, 1167–1179.

Stock, J.H. and M.W. Watson, 2002b, Macroeconomic Forecasting Using Diffusion Indexes, *Journal of Business & Economic Statistics* **20**, 147–162.

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

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.

Table 1: Posterior probability of predictor variable selection

Variable	Marginal posterior inclusion probability	
	CPI	GDP DEFL
AR_{-1}	1.000	0.503
AR_{-2}	1.000	1.000
AR_{-3}	1.000	1.000
AR_{-4}	1.000	0.124
$ROUTPUT_{-1}$	0.709	1.000
$RCONS_{-1}$	1.000	0.069
$RINVRESID_{-1}$	0.207	0.222
$PIMPORT_{-1}$	0.403	0.035
$M1_{-1}$	0.527	0.236
$EMPLOY_{-1}$	1.000	0.246
$HSTARTS_{-1}$	1.000	0.286
IP_{-1}	0.023	0.036
TS_{-1}	1.000	0.003
MS_{-1}	0.879	0.937

Note: The table presents the marginal posterior probabilities of variables to be selected in the predictive regression model for quarterly CPI.

Table 2: Alternative univariate and multivariate models to forecast inflation

name	description	specification
	univariate models	
RW	Random walk	$\Delta y_t = \mu + \varepsilon_t$
IMA(1,1)	Integrated moving average	$\Delta y_t = \mu + \varepsilon_t + \psi \varepsilon_{t-1}$
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	$y_t = \mu + \sum_{i=1}^{4} \phi_i y_{t-i} + \varepsilon_t$
Linear	Linear regression with all predictors	$y_t = \beta_0 + x_t' \beta + \varepsilon_t$
Ridge	Ridge regression with $\lambda = 10$	$y_t = \beta_0 + x_t' \beta + \varepsilon_t$
BMA	Bayesian model averaging	(5) with $\beta_t = \beta$, $\sigma_t = \sigma \ \forall t$
SB	Structural instability	(5)–(4) with $s_j = 1 \ \forall j$
	multivariate models	
VAR(4)	Bivariate VAR(4) with CPI & real output	$Y_t = \mu + \sum_{i=1}^{4} \Phi_i Y_{t-i} + \epsilon_t$
BMA-VAR(4)	BMA of all possible bivariate VAR(4)	$Y_t = \mu + \sum_{i=1}^{4} \Phi_i Y_{t-1} + \epsilon_t$

Note: The y_t and Y_t variable contains the quarterly log change of seasonally adjusted CPI. Furthermore, $\varepsilon_t \sim N(0, \sigma^2)$ and $\epsilon_t \sim N(0, \Sigma)$.

Table 3: RMSPE - CPI

	1-step ahead			4-step ahead		
BMASB	**0.38**	**0.42**	0.31	0.49	0.60	**0.33**
			univariate models			
RW	1.25	1.28	1.17	1.17	1.17	1.20
IMA	1.16	1.21	1.03	1.09	1.10	1.06
AR(2)	1.18	1.23	1.06	1.06	1.05	1.11
AR(3)	1.11	1.15	1.02	1.07	1.06	1.10
AR(4)	1.11	1.12	1.06	1.08	1.05	1.17
Linear	1.07	1.01	1.18	1.07	1.04	1.19
Ridge	1.08	1.11	**1.00**	1.00	0.98	1.05
BMA	1.08	1.03	1.18	**0.97**	**0.93**	1.13
SB	1.50	1.57	1.32	1.18	1.18	1.20
			multivariate models			
VAR(4)	1.11	1.13	1.05	1.09	1.06	1.20
VAR-BMA	1.06	1.05	1.06	1.09	1.07	1.16

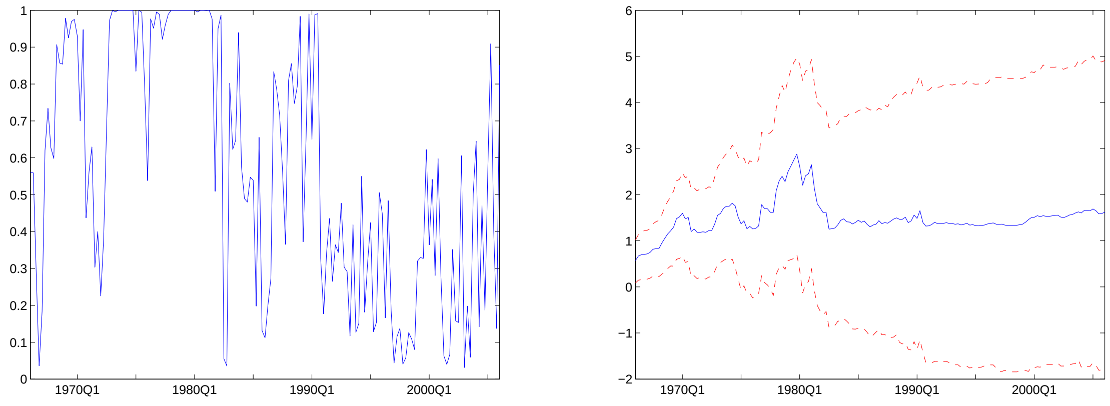
Note: The table presents Root Mean Square Prediction Error (RMSPE) of the BMASB model and relative to the BMASB for different univariate and multivariate models for the full sample (F) and two subsamples (I: 1980Q1:1994Q4, II: 1995Q1:2006Q3), 1-step and 4-step ahead forecasting horizons.

Table 4: RMSPE - GDP DEFLATOR

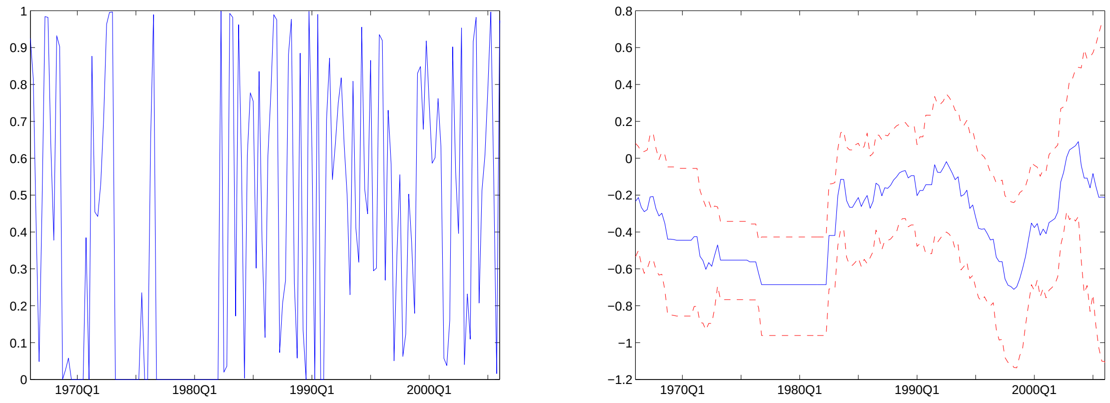
	1-step ahead			4-step ahead		
BMASB	**0.25**	**0.30**	**0.18**	**0.31**	**0.36**	**0.23**
			univariate models			
RW	1.22	1.22	1.19	1.11	1.10	1.11
IMA	1.08	1.08	1.07	1.04	1.05	1.01
AR(2)	1.12	1.12	1.12	1.09	1.08	1.13
AR(3)	1.06	1.06	1.07	1.09	1.08	1.12
AR(4)	1.06	1.06	1.07	1.09	1.09	1.11
Linear	1.05	1.05	1.02	1.08	1.07	1.12
Ridge	1.04	1.03	1.06	1.05	1.05	1.05
BMA	1.03	1.01	1.06	1.08	1.05	1.18
SB	1.42	1.47	1.23	1.45	1.48	1.35
			multivariate models			
VAR(4)	1.05	1.06	1.04	1.04	1.02	1.10
VAR-BMA	1.03	1.03	1.03	1.03	1.02	1.06

Note: The table presents Root Mean Square Prediction Error (RMSPE) of the BMASB model and relative to the BMASB for different univariate and multivariate models for the full sample (F) and two subsamples (I: 1980Q1:1994Q4, II: 1995Q1:2006Q3), 1-step and 4-step ahead forecasting horizons.

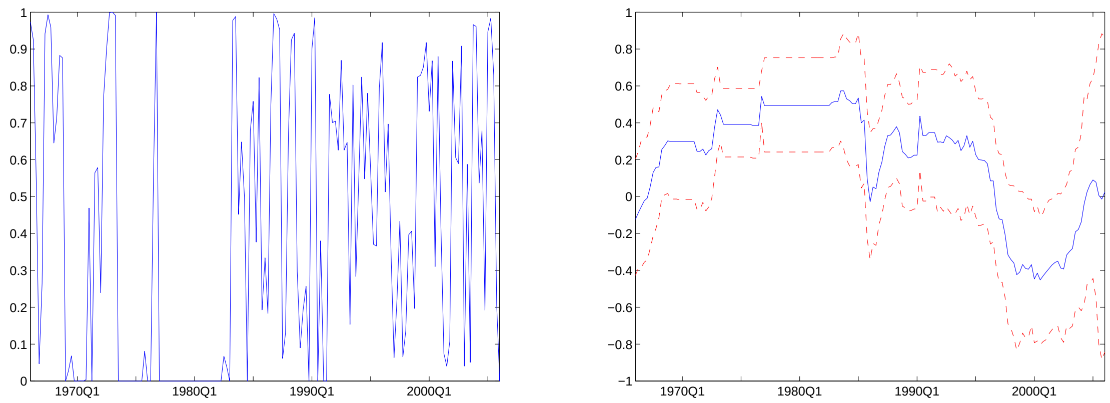
Figure 1: Posterior densities of the breaks and β parameters conditional on inclusion: CPI



(a) Intercept

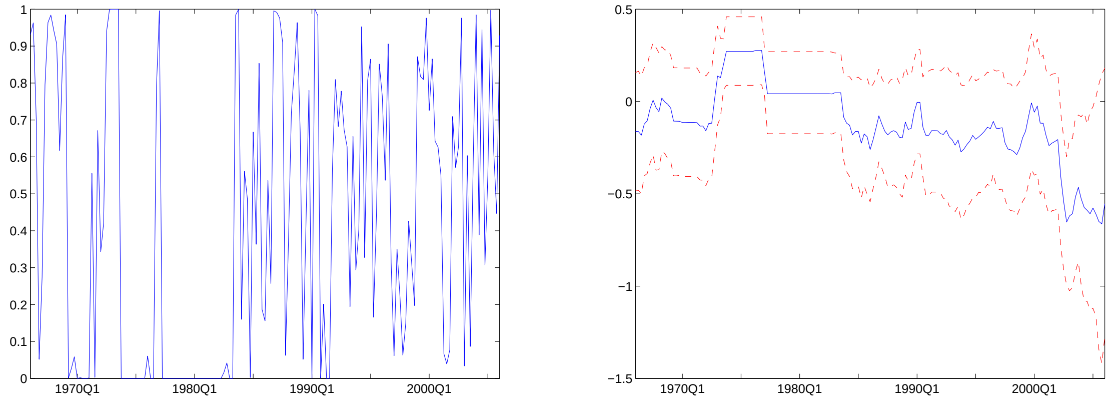


(b) CPI_{-1}

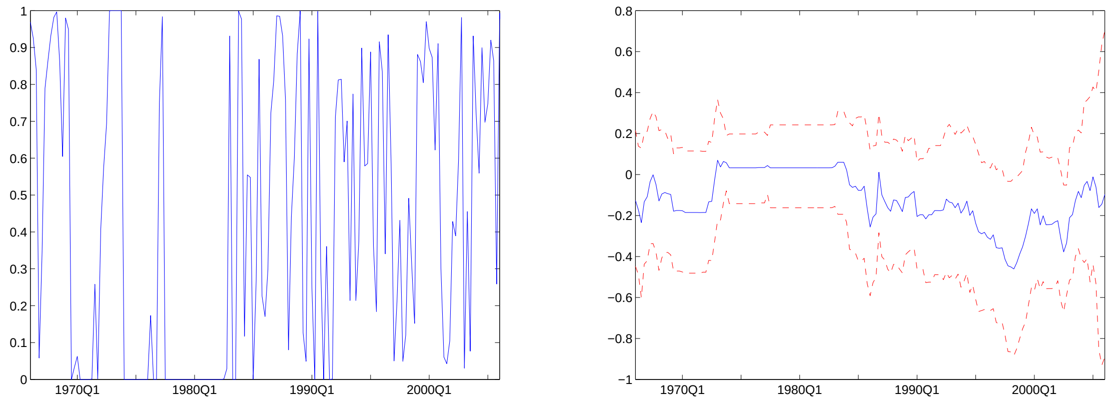


(c) CPI_{-2}

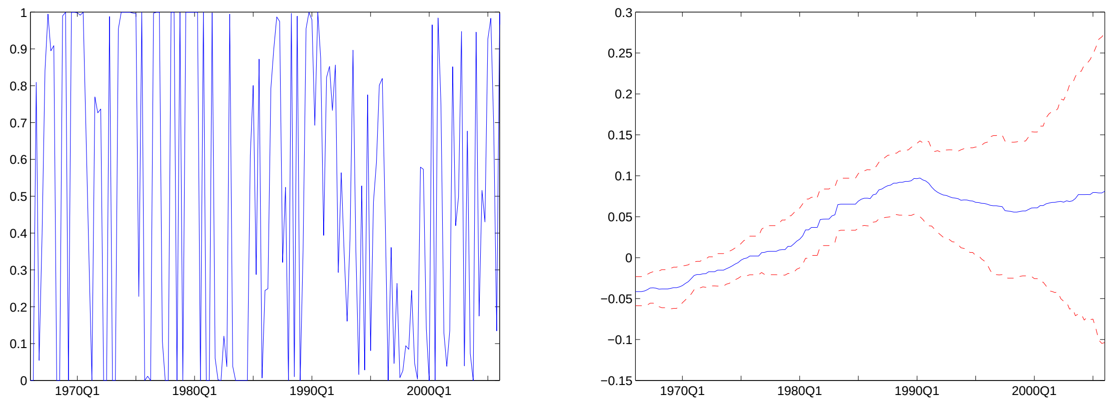
(continued on next page)



(d) CPI_{-3}

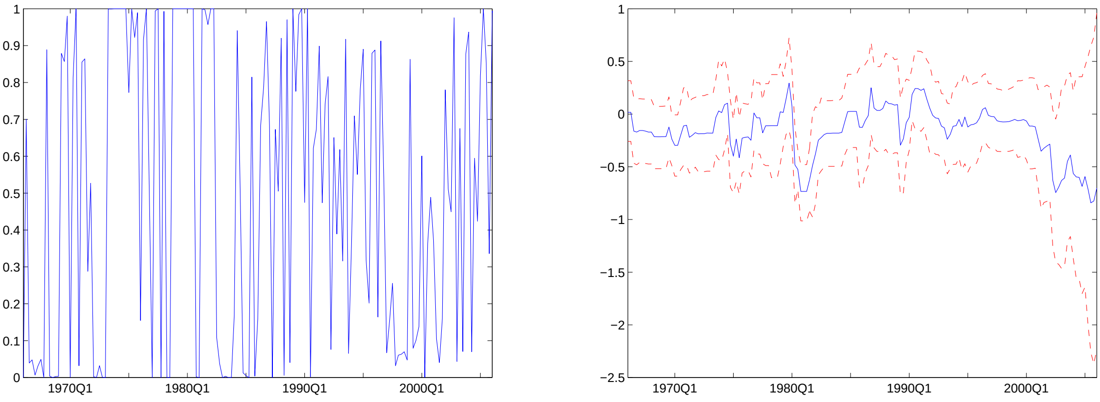


(e) CPI_{-4}

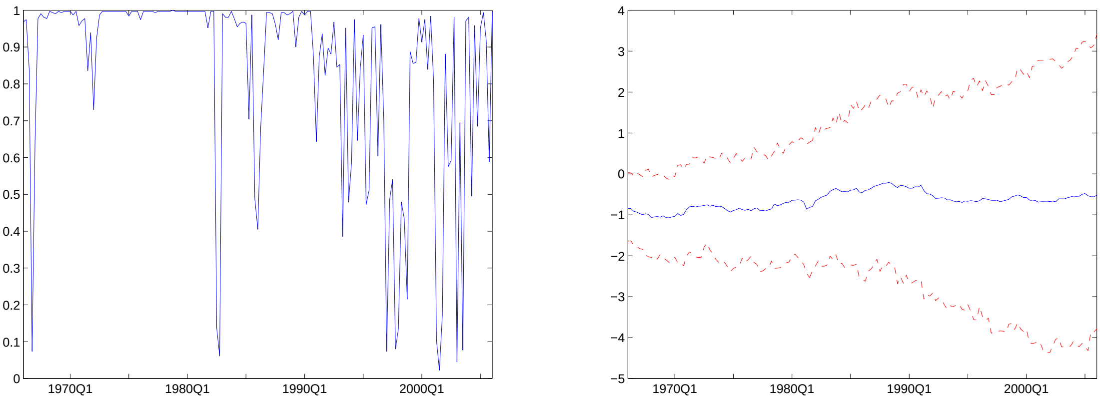


(f) $ROUTPUT_{-1}$

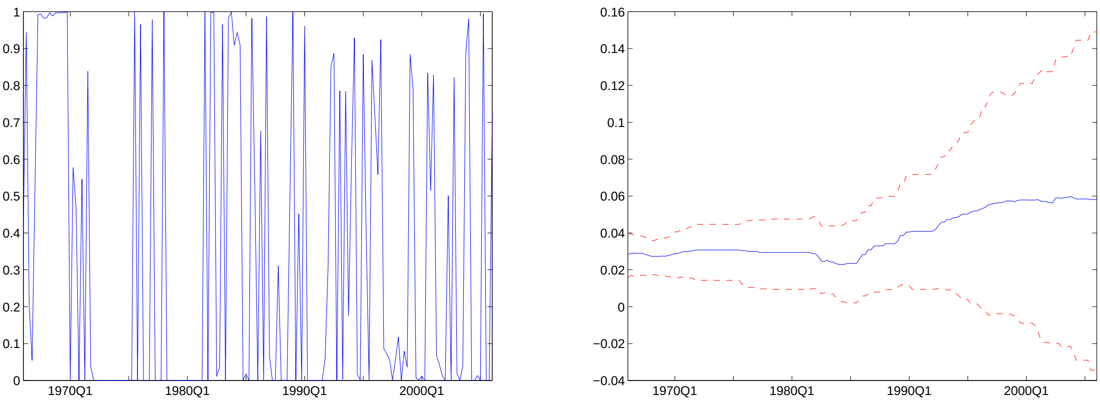
(continued on next page)



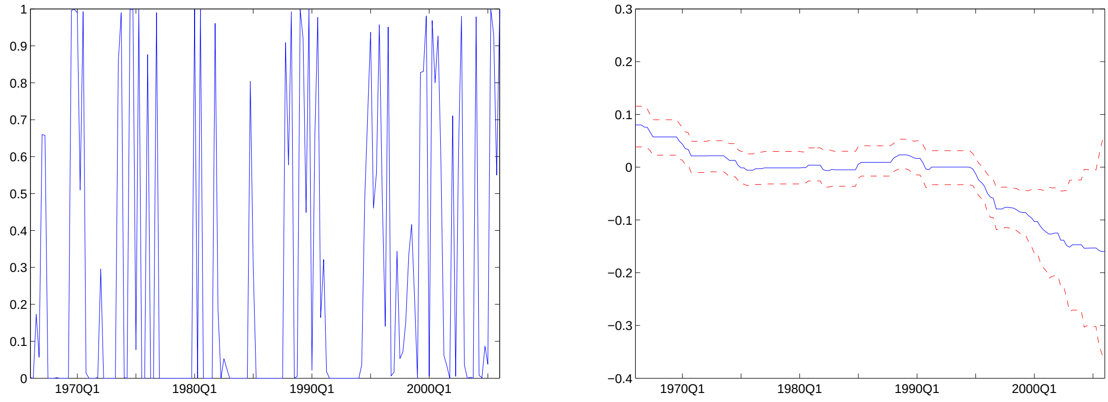
(g) $RCONS_{-1}$



(h) $RINVRESID_{-1}$

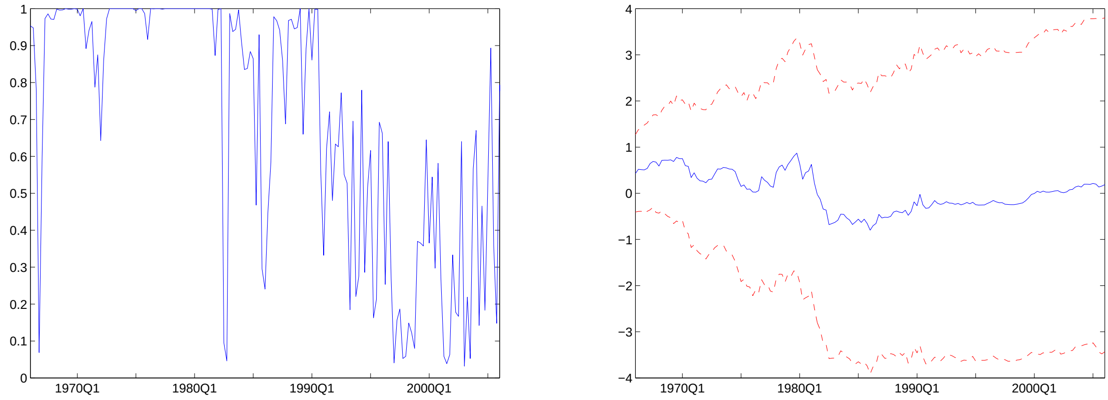


(i) $PIMPORT_{-1}$

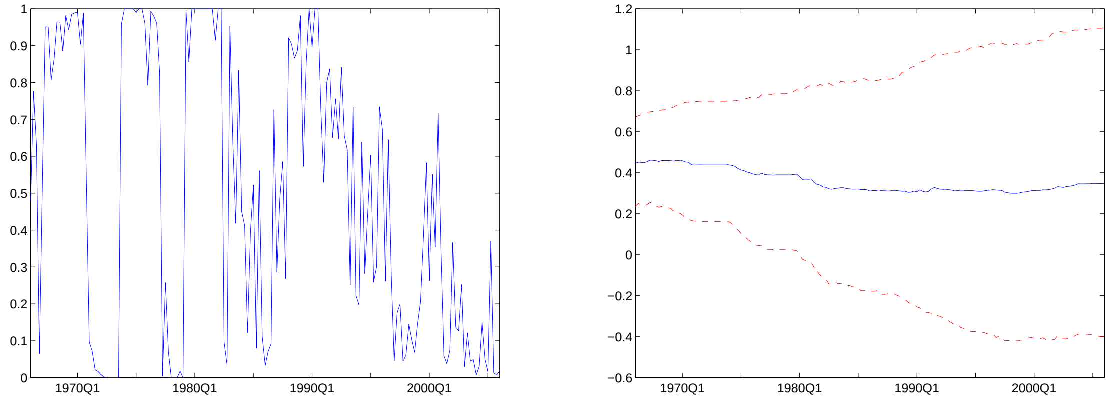


(j) $M1_{-1}$

(continued on next page)

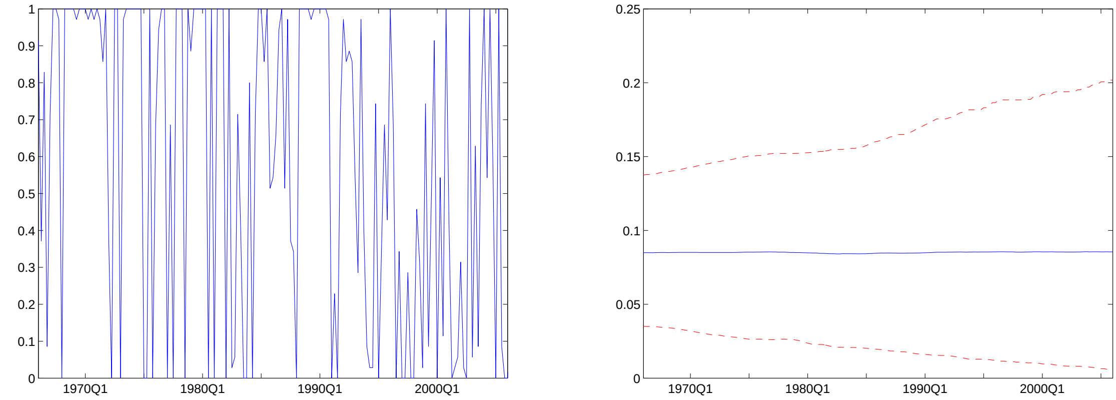


(k) $EMPLOY_{-1}$

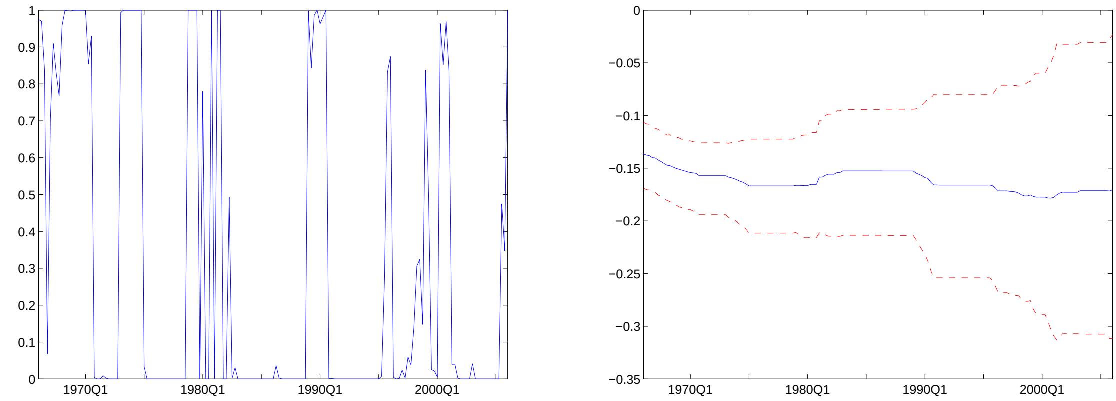


(l) $HSTARTS_{-1}$

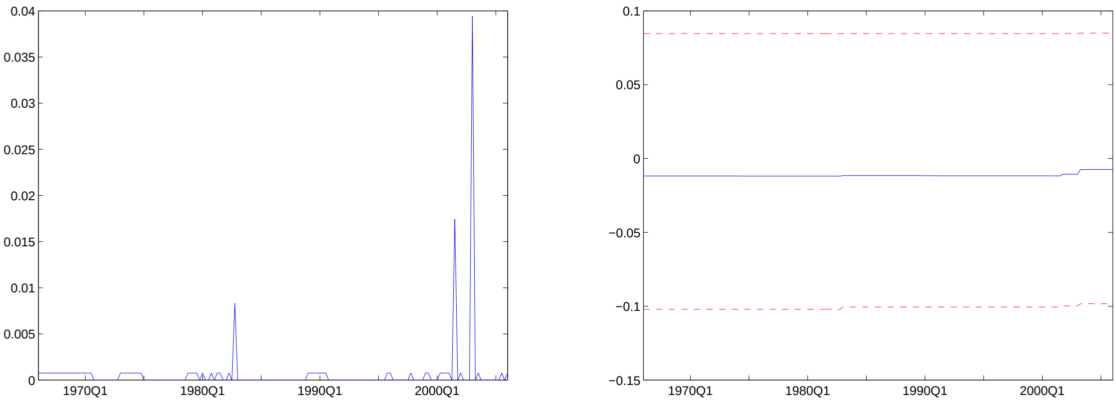
(continued on next page)



(m) IP_{-1}



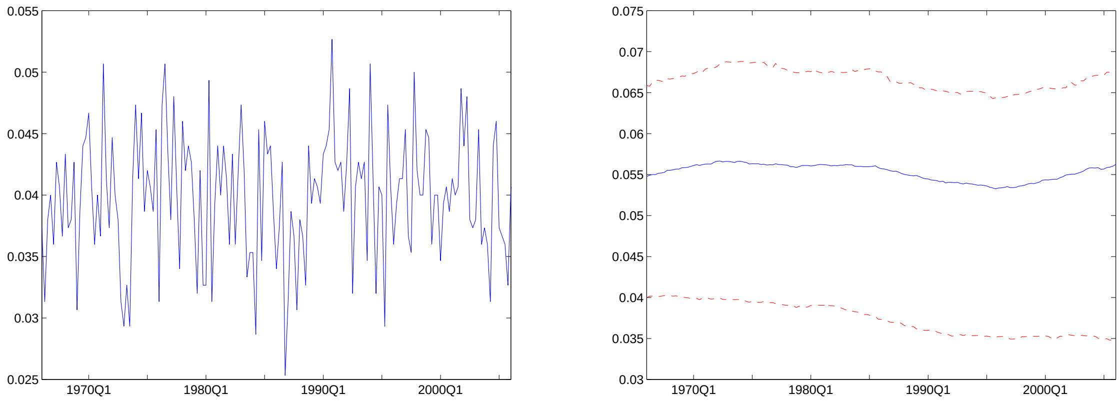
(n) TS_{-1}



(o) MS_{-1}

Note: The graphs in this figure show the posterior means (solid line) of κ_{jt} on the left side and β_{jt} on the right side. The dashed lines in the graphs for the coefficients are the 25th and 75th percentiles of the posterior densities.

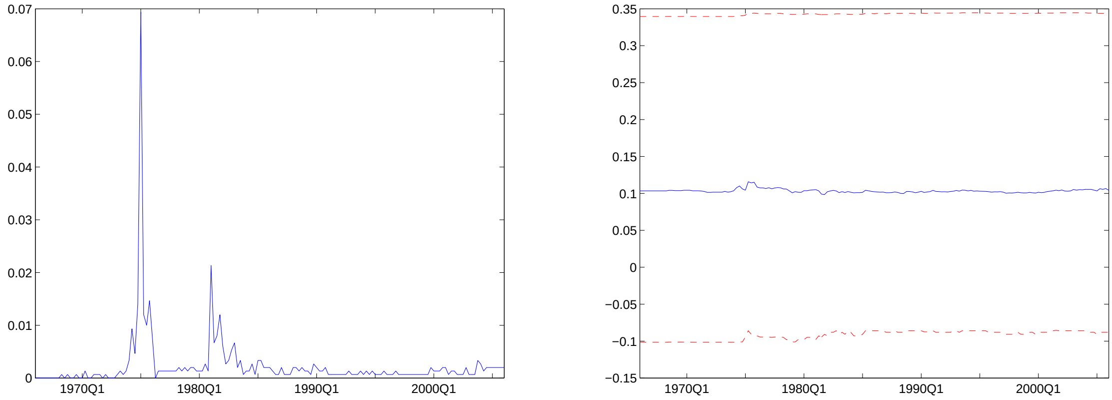
Figure 2: Posterior densities of the breaks and σ_t^2 parameter: CPI



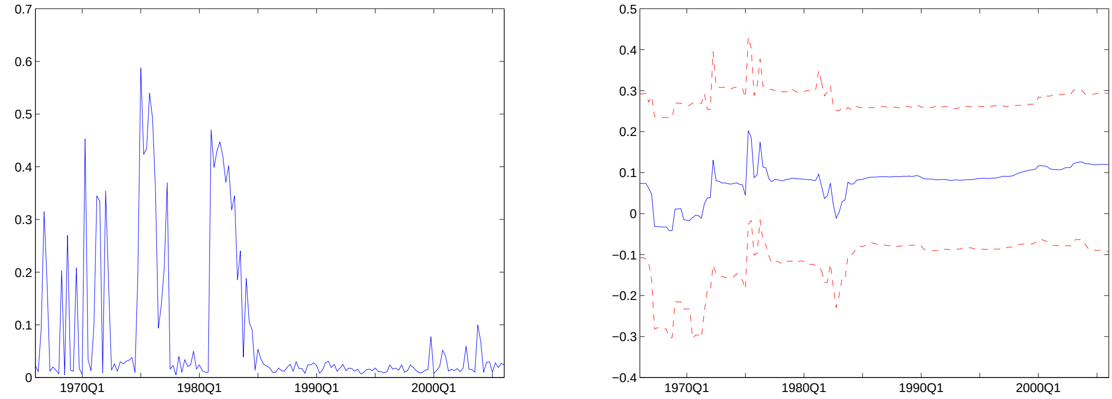
(a) σ_t^2

Note: The graphs in this figure show the posterior means (solid line) of $\kappa_{\sigma,t}$ on the left side and σ_t^2 on the right side. The dashed lines in the graph for σ_t^2 are the 25th and 75th percentiles of the posterior densities.

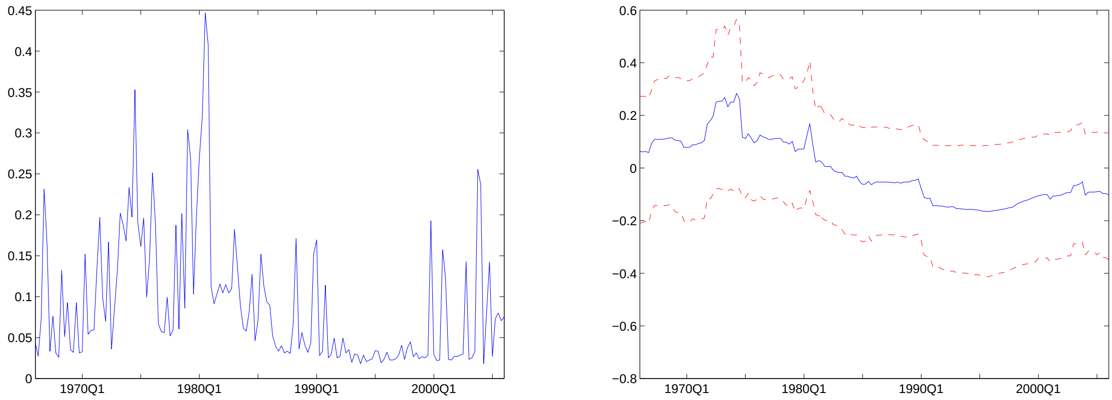
Figure 3: Posterior densities of the breaks and β parameters conditional on inclusion: GDP DEFLATOR



(a) Intercept

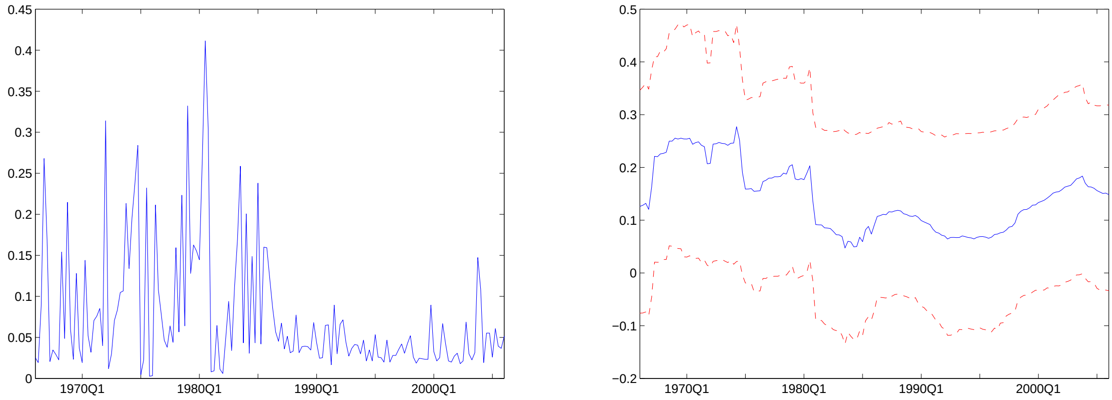


(b) $GDPDEFL_{-1}$

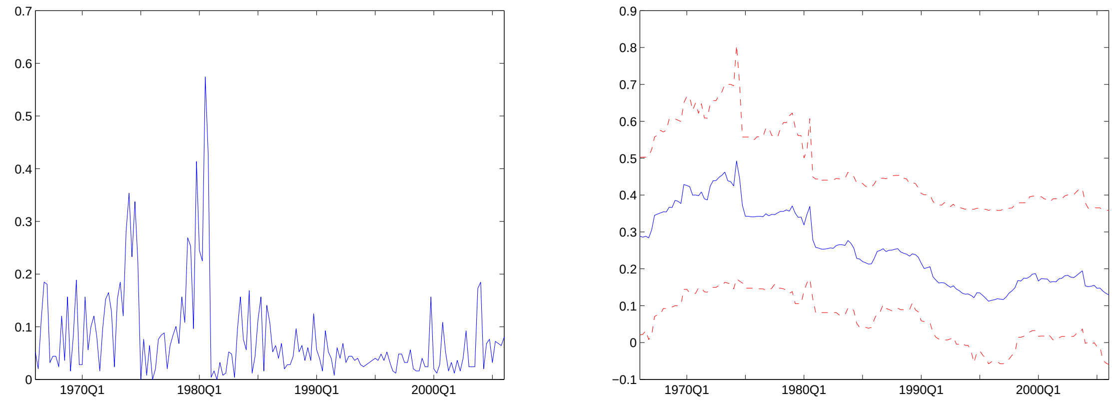


(c) $GDPDEFL_{-2}$

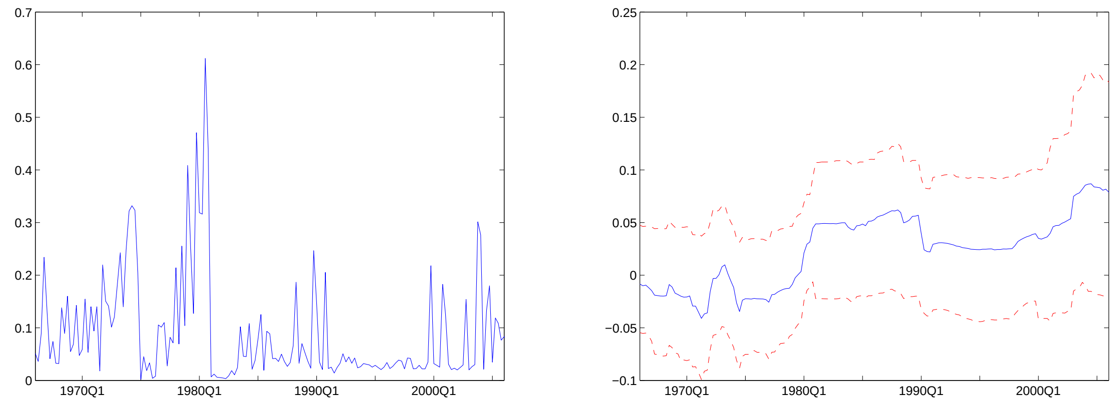
(continued on next page)



(d) $GDPDEFL_{-3}$

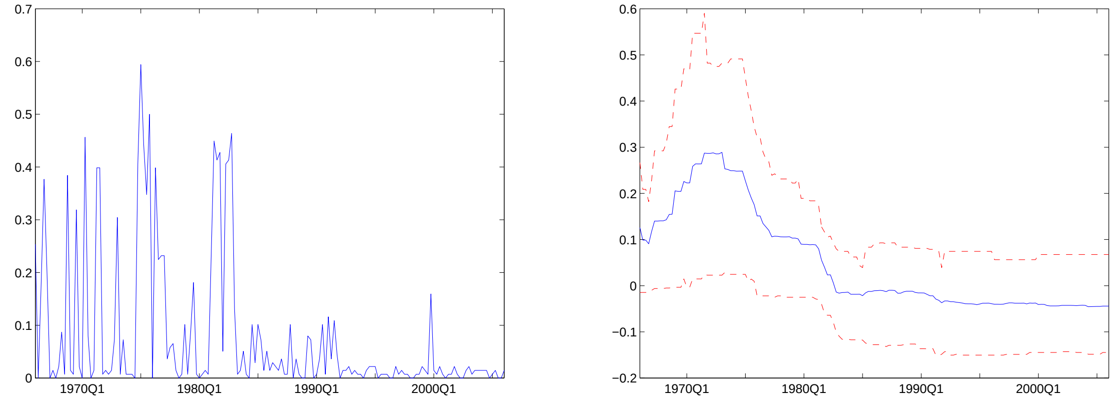


(e) $GDPDEFL_{-4}$

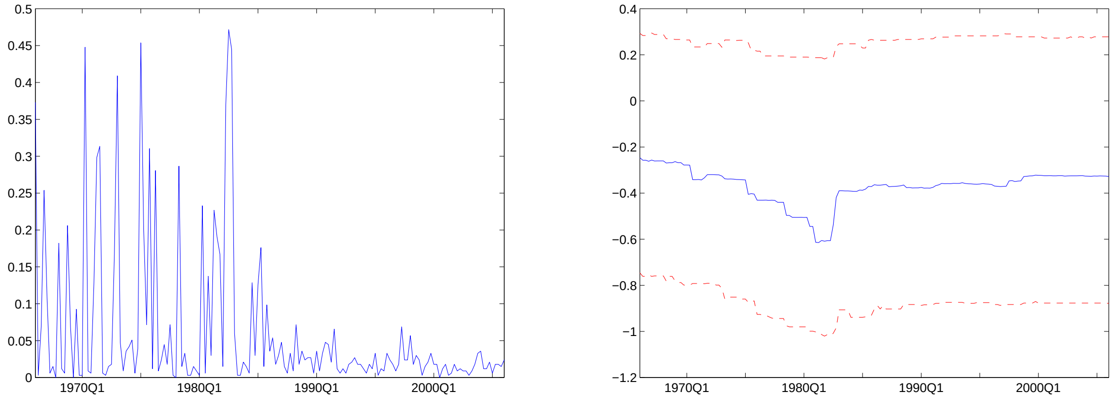


(f) $ROUTPUT_{-1}$

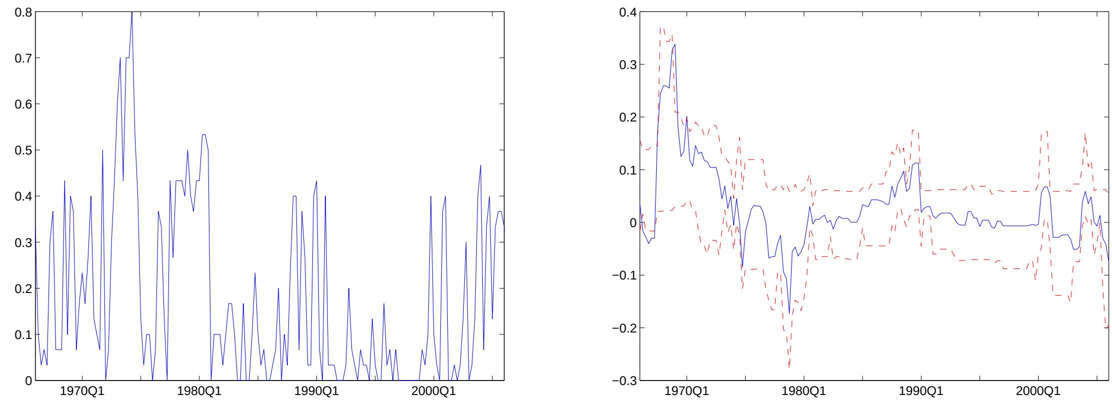
(continued on next page)



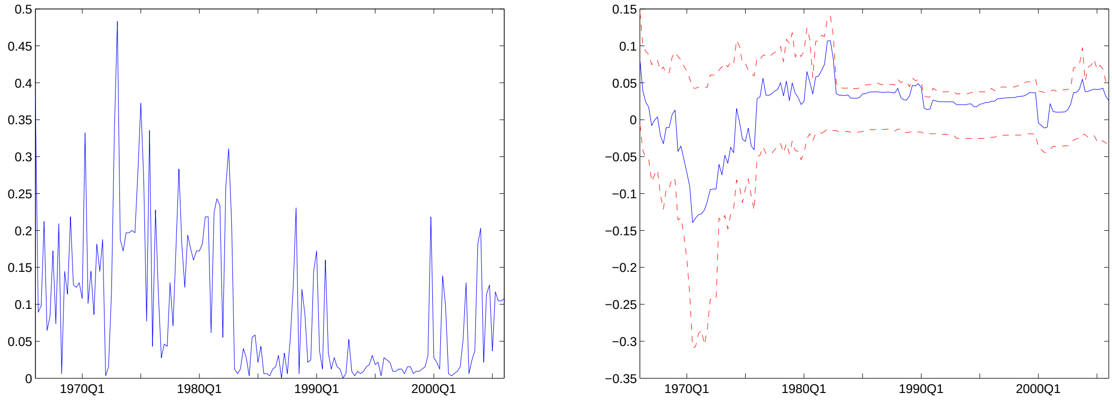
(g) $RCONS_{-1}$



(h) $RINVRESID_{-1}$

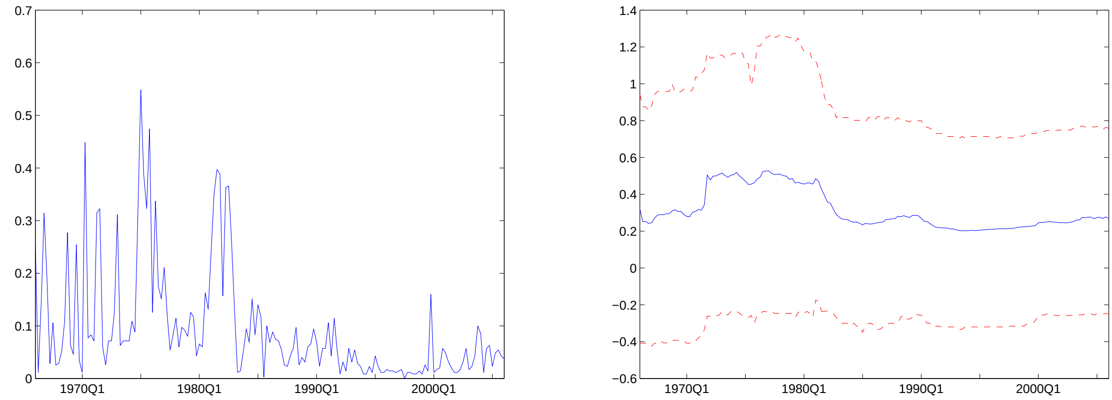


(i) $PIMPORT_{-1}$

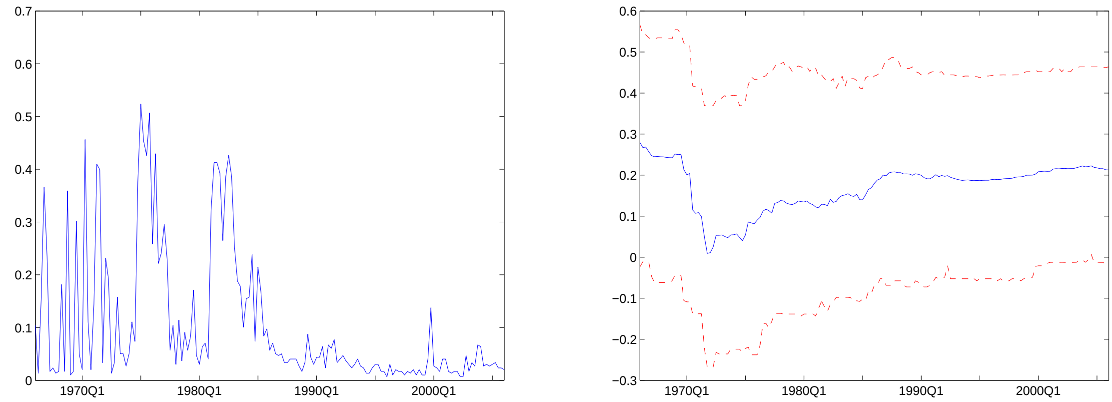


(j) $M1_{-1}$

(continued on next page)

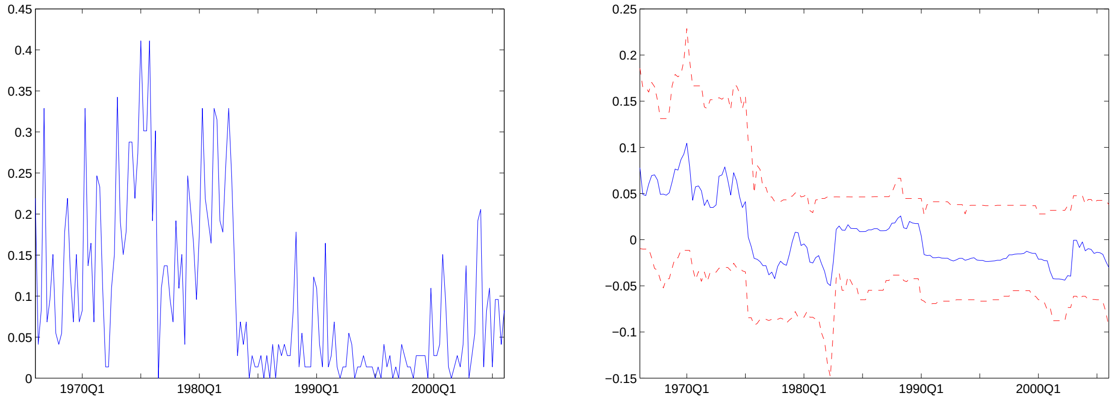


(k) $EMPLOY_{-1}$

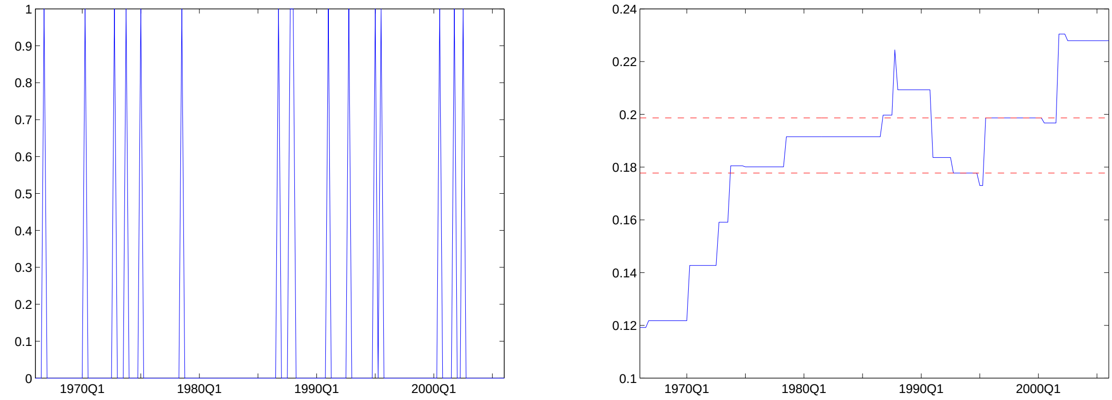


(l) $HSTARTS_{-1}$

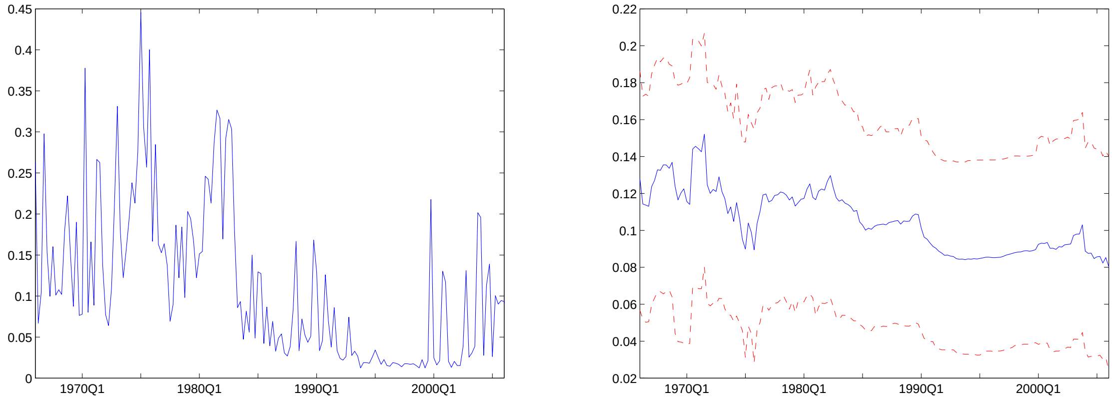
(continued on next page)



(m) IP_{-1}



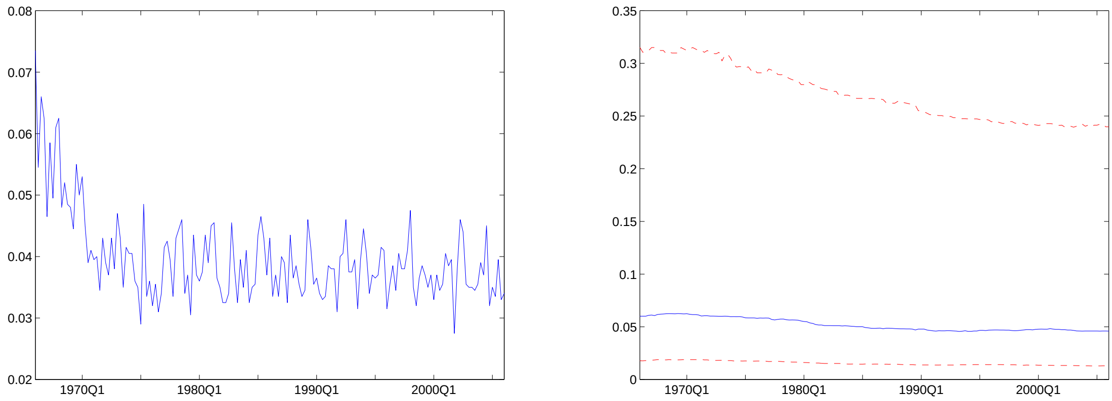
(n) TS_{-1}



(o) MS_{-1}

Note: The graphs in this figure show the posterior means (solid line) of κ_{jt} on the left side and β_{jt} on the right side. The dashed lines in the graphs for the coefficients are the 25th and 75th percentiles of the posterior densities.

Figure 4: Posterior densities of the breaks and σ_t^2 parameter: GDP DEFLATOR



(a) σ_t^2

Note: The graphs in this figure show the posterior means (solid line) of $\kappa_{\sigma,t}$ on the left side and σ_t^2 on the right side. The dashed lines in the graph for σ_t^2 are the 25th and 75th percentiles of the posterior densities.