Testing for time variation in an unobserved components model for the U.S. economy

Tino Berger$^1$, Gerdie Everaert$^2$, and Hauke Vierke$^3$

$^1$University of Göttingen
$^2,^3$Ghent University

May 17, 2016

Abstract

This paper analyzes the amount of time variation in the parameters of a reduced-form empirical macroeconomic model for the U.S. economy. We decompose output, inflation and unemployment in their stochastic trend and business cycle gap components, with the latter linked through the Phillips curve and Okun’s law. A novel Bayesian model selection procedure is used to test which parameters are vary over time and which components exhibit stochastic volatility. Using data from 1959Q2 to 2014Q3 we find substantial time variation in Okun’s law, while the Phillips curve slope appears to be stable. Stochastic volatility is found to be important for cyclical shocks to the economy, while the volatility of permanent shocks remains stable.

JEL: C32, E24, E31

1 Introduction

Over the last decades the U.S. economy has experienced a number of notable structural changes. Well documented are the productivity slowdown in the early 1970s and the reduction in the volatility of key macroeconomic variables in the mid 1980s, known as the Great Moderation. More recently, due to the experience of the 2001 recession and the Great Recession, the interest in the academic literature in analyzing structural changes has been renewed. In particular, during the Great Recession, with unemployment being very high, most Phillips curve estimates imply that prices should have fallen much more than what the actual data show. This case of missing deflation has cast doubt on the stability of the Phillips curve. Furthermore, in the aftermath of the last two recessions, job growth appeared substantially lower than what the level of output growth would have implied. These episodes, known as ‘jobless recoveries’, have led many observers to conclude that the trade-off between unemployment and output has changed. Finally, the severity of the Great Recession and the related increases in the volatility of key macroeconomic variables may herald the end of the Great Moderation.
A growing literature investigates time variation in macroeconomic relationships. First, the necessity for empirical models to account for changes in the volatility of macroeconomic variables has been emphasized by Hamilton (2008) and Fernández-Villaverde and Rubio-Ramírez (2010), the former showing that not accounting for volatility changes can lead to biased estimates and misleading inference. Second, regarding the relationship between inflation and real economic activity, the literature has collected growing evidence for a change in the slope of the Phillips curve. Ball and Mazumder (2011) forecast inflation over the period 2008-2010 using backward-looking Phillips curve estimates for the period 1960-2007. The model predicts substantial deflation, which is not in line with the slightly positive actual inflation rate observed over this period. Hall (2011) also emphasizes the case of missing deflation during the Great Recession and notes that inflation remained remarkably stable at a small but positive rate despite the large and persistent slack in real activity. Roberts (2006) analyzes U.S. data prior to the Great Recession and finds that the reduced-form Phillips curve slope fell by nearly half between the periods 1960-1983 and 1984-2002. Similar results can be found in Atkeson and Ohanian (2001) and Mishkin (2007). Regarding the relationship between unemployment and real economic activity, a different strand of the literature investigates the stability of Okun’s law. Daly et al. (2012) note that if Okun’s law had held in 2009, the U.S. unemployment rate would only have risen by about half of the observed rise. Owyang and Sekhposyan (2012) conclude that the relationship between unemployment and output fluctuations changes significantly during the most recent recession periods. Lee (2000) reports international evidence for structural breaks in the Okun coefficient during the 1970s. Contradicting evidence is given by Ball et al. (2013), who find that Okun’s law is a ‘strong and stable’ relationship.

Measuring these various types of structural change is challenging as it relates to variables that are not directly observed. The Phillips curve links inflation to expected inflation and to a measure for the deviation of real economic activity from its potential, such as the output gap or the unemployment gap. Each of these determinants is unobserved. The same argument holds for Okun’s law, which models the interaction between the output gap and the unemployment gap. To proxy these unobserved factors, many studies rely on purely statistical trend-cycle decompositions based on filtering techniques - such as the Hodrick-Prescott filter - or use external estimates provided by a statistical bureau - such as the Congressional Budget Office’s (CBO) series for the U.S. economy. The first approach suffers from a lack of structural interpretation while the second entails the risk of falling into an endogeneity trap. The CBO for instance follows a growth model for calculating potential output thereby relying on constant values for the slope of the Phillips curve and the Okun’s law coefficient. As such, these slopes and their stability are artificially imposed on the data from the outset.

In this paper, we set up and estimate a multivariate unobserved components model for the U.S. economy to jointly estimate a time-varying NAIRU, trend inflation, potential output, and the respective gaps. Important model parameters are allowed to change over time. Specifically, we allow the forward-looking New Keynesian Phillips curve slope, the Okun’s law coefficient, the growth rate of potential output and the variances of the innovations to all unobserved components

---

1 An alternative version of Okun’s law relates the change in the unemployment rate to output growth. This framework, however, rests on the restrictive assumption of a constant natural rate of unemployment and a constant growth rate of potential output.
to vary over time.

The model in our paper is most closely related to the following recent papers. First, Stella and Stock (2012) estimate the time-varying trend inflation and the NAIRU using a bivariate unobserved components (UC) model with stochastic volatility (SV). While the Phillips curve slope is treated as constant in the forward-looking inflation equation, the implied backward-looking Phillips curve has a time-varying slope parameter which is found to vary considerably. Second, Chan et al. (2016) build on this model and use a bounded random walk specification for the trend components. However, their analysis can be understood as a forecasting exercise as less emphasis is put on time variation in the parameters. They stick to a bivariate model of inflation and unemployment. Third, Kim et al. (2014) allow for two structural breaks in the slope of the U.S. New Keynesian Phillips curve. The sensitivity of inflation to the CBO output gap is found to be small but significant prior to 1971, while being insignificant from 1971 onwards.

Another strand of the literature analyses time variation in the parameters of vector autoregressive (VAR) models. Well known contributions, focussing on potential changes in the conduct of monetary policy, include Sims and Zha (2006) who estimate a VAR with Markov-switching parameters and Cogley and Sargent (2005), Cogley et al. (2010), Primiceri (2005) who estimate VARs with time-varying parameters (TVP-VAR) that evolve as random walks. TVP-VARs are a very flexible tool since they allow for different types of structural change, i.e. time variation in the persistence, in the correlation structure and in the conditional variances. However, they suffer from an over-parameterization problem that can seriously impact estimation accuracy (see e.g. Chan et al., 2012). To deal with this problem, the TVP-VAR literature uses small dimensional models with a low number of lags (Cogley and Sargent, 2005, for instance estimate a trivariate VAR with two lags) and typically sets tight priors for the time-varying parameters. Hence, the amount of time variation in the posterior estimates may be largely driven by the priors (see e.g. Reusens and Croux, 2015). Our UC model outlined in the next section has a reduced form VARMA representation and thus relates to the TVP-VAR literature. In contrast to the TVP-VAR literature we will use a small scale macroeconomic model to decompose output, inflation and unemployment into their stochastic trend and business cycle gap components, with the latter being linked through the Phillips curve and Okun’s law. This allows us to analyze more directly potential structural changes in the innovation variances of unobserved variables such as the output gap or potential output and potential time variation in the parameters governing the relation among them.

A common feature of the literature investigating structural change is that model uncertainty is mostly ignored. Time variation is typically modeled through discrete breaks or allowing parameters to change gradually by specifying them as stochastic random walks. Hence, structural change is assumed from the outset, without testing whether it is actually relevant. Kim et al. (2014), for instance, allow the slope of the U.S. Phillips curve to change over three states of the economy but do not test whether the obtained slopes are significantly different. In contrast, we explicitly address model uncertainty. Key parameters and components are modeled as random walks, but instead of specifying which components to include and deciding whether they are constant or vary over time, we will test which features of the model are relevant and fall back to a more parsimonious specification when appropriate. This not only avoids over-parameterization but will also provide
us with information on which components are relevant and which parameters actually vary over time. Model selection for UC models is a challenge, though, as it leads to testing problems that are non-regular from a classical point of view. We will therefore use the Bayesian stochastic model specification search outlined in Frühwirth-Schnatter and Wagner (2010). In our baseline model, we consider eight potentially time-varying parameters such that a total of $2^8$ different models are considered in the model selection procedure. To the best of our knowledge, this is the first study to allow and explicitly test for such a wide range of time-varying parameters in a macroeconomic UC model. As such, our results will provide new evidence on the form and the degree of structural change in the U.S. economy.

Our main findings can be summarized as follows. First, the correlation between cyclical unemployment and cyclical output varies over time and appears to be less pronounced in recessions. Second, the slope of the Phillips curve is constant over time. This finding is robust over a forward and backward-looking specification. Third, the growth rate of potential output has decreased from a quarterly growth rate of 1% in the 1960s to 0.4% in the 2000s. The most substantial decreases are observed over the 1970s and 2000s. Fourth, shocks to the output gap and to the transitory inflation component exhibit stochastic volatility while shocks to the NAIRU, potential output and trend inflation appear to be homoskedastic.

The remainder of the paper is structured as follows: The next section introduces our empirical model and explains how we test for time variation. Results are presented in Section 3. In Section 4 we perform several robustness checks and discuss model extensions. The final section concludes.

## 2 Empirical approach

This section explains our econometric approach. We first lay out a multivariate unobserved components model with time-varying parameters and stochastic volatilities, designed to fit U.S. macroeconomic data. We then turn to the Bayesian stochastic model selection approach as well as a description of the Markov Chain Monte Carlo (MCMC) algorithm employed to estimate the model.

### 2.1 An unobserved components model

**Output: trend/cycle decomposition**

Consider a decomposition of real GDP $y_t$ into a stationary cycle $y^c_t$ and a non-stationary trend $y^\tau_t$ referred to as potential output

$$y_t = y^\tau_t + y^c_t + \varepsilon^y_t,$$

$$\varepsilon^y_t \sim i.i.d. \mathcal{N}(0, \sigma^2_{\varepsilon,y}),$$

(1)
where $\varepsilon^y_t$ is included to capture measurement error and non-persistent shocks. Potential output is modeled as a random walk process with stochastic drift $\kappa_t$

$$y_{t+1}^y = \kappa_t + y_t^y + \exp \{ h_t^y \} \psi_t^y, \quad \psi_t^y \sim i.i.d. N(0, 1),$$  \hspace{1cm} (2)

$$\kappa_{t+1} = \kappa_t + \psi_t^\kappa, \quad \psi_t^\kappa \sim i.i.d. N(0, \sigma_{\psi,\kappa}^2).$$  \hspace{1cm} (3)

The stochastic drift is included to capture permanent changes in the growth rate of potential output. The productivity slowdown in the early 1970s for instance is likely to have lowered the growth rate of potential output. Demographic changes as well as potential long-run effects of the Great Recession are other potential drivers of $\kappa_t$. The output gap $y_c^t$ is modeled as a stationary autoregressive (AR) process of order two

$$y_{t+1}^c = \rho_1 y_t^c + \rho_2 y_{t-1}^c + \exp \{ h_t^c \} \psi_t^c, \quad \psi_t^c \sim i.i.d. N(0, 1).$$  \hspace{1cm} (4)

This AR(2) specification allows the output gap to exhibit the standard hump-shaped pattern. The stochastic volatility terms $\exp \{ h_t^y \}$ and $\exp \{ h_t^c \}$ in the innovations to the trend and the cycle are included to account for changes in macroeconomic volatility such as the Great Moderation or the recent increase in volatility due to the financial crises. These components are specified below.

**Inflation: a time-varying New-Keynesian Phillips curve**

In contemporary macroeconomic models, the New-Keynesian Phillips Curve (NKPC) relates actual inflation to expected inflation and some measure for excess demand. It can be derived from a micro-founded theoretical model with Calvo (1983) pricing in which firms seek to set their price as a mark-up over marginal costs but are only randomly allowed to change their prices. However, in its pristine form the empirical performance of the NKPC is disappointing as the slope of the NKPC is often found to be small and insignificant. Moreover, it fails to match important stylized facts of inflation dynamics. The purely forward-looking specification implies that current inflation is the discounted present value of expected future activity gaps. As the activity gap is a stationary process inflation should be stationary as well. This is at odds with the observed high degree of persistence in inflation, which is typically found to be non-stationary. Fuhrer and Moore (1995), Mankiw (2001), Rudd and Whelan (2005, 2007) and Mavroeidis et al. (2014) discuss these failures in greater detail.

An appealing way to match the NKPC with the data is the introduction of stochastic trend inflation as in Kim et al. (2014); Morley et al. (2013); Stella and Stock (2012). Cogley and Sbordone (2008) derive a NKPC that allows for a time-varying trend inflation rate. By incorporating trend inflation their purely forward-looking NKPC fits the data well without the need to include backward-looking components. We follow this literature and use an inflation gap NKPC in which inflation is modeled in deviation from trend inflation $\pi^T_t$ and the output gap is used as a measure.

---

2The trend may be attributed to shifts in monetary policy (see e.g. Woodford, 2008; Cogley and Sbordone, 2008; Goodfriend and King, 2012).
of real activity, i.e.

\[ \pi_t - \pi^r_t = \omega E_t(\pi_{t+1} - \pi^r_{t+1}) + \beta_t^\pi y^c_t + \tilde{\zeta}_t, \]  

(5)

where \( \omega \) is a discount factor. As shown by Beveridge and Nelson (1981), in the presence of a zero-mean transitory component the trend component \( \pi^T_t \) equals the long-run inflation forecast \( \lim_{h \to \infty} E(\pi_{t+h}) \). Following Cogley and Sbordone (2008) and Kim et al. (2014), the term \( \tilde{\zeta}_t \) is included to capture variation in the inflation gap that is not explained by the conventional forward-looking NKPC. According to Mavroeidis et al. (2014) this term can be interpreted as a combination of cost-push shocks, such as shocks to the markup or to input (e.g. oil) prices. As this term is potentially serially correlated, it allows for an additional source of inflation persistence not related to expectations or real activity. Hence, our model resembles alternative hybrid NKPC models that explicitly add lagged inflation or supply shock variables. As we want to analyze whether the slope of the Phillips curve varies over time, we allow \( \beta_t^\pi \) to be a time-varying parameter. Our specification of the Phillips curve is most closely related to that of Kim et al. (2014) who allow \( \beta_t^\pi \) to vary using a three-state Markov switching model. Iterating equation (5) forward and rearranging yields

\[
\pi_t = \pi^r_t + \lim_{j \to \infty} \omega^j E_t(\pi_{t+j} - \pi^r_{t+j}) + \sum_{j=0}^{\infty} \omega^j E_t(\beta^\pi_{t+j} y^c_{t+j}) + \zeta_t, \\
= \pi^r_t + \sum_{j=0}^{\infty} \omega^j E_t(\beta^\pi_{t+j} y^c_{t+j}) + \zeta_t, 
\]

(6)

with \( \zeta_t = \sum_{j=0}^{\infty} E_t \left( \tilde{\zeta}_{t+j} \right) \) and \( \lim_{j \to \infty} \omega^j = 0 \). Equation (6) implies that inflation has a trend/cycle representation, i.e.

\[
\pi_t = \pi^r_t + \pi^c_t + \varepsilon^\pi_t, \quad \varepsilon^\pi_t \sim i.i.d. N(0, \sigma^2_{\varepsilon, \pi}), 
\]

(7)

where \( \pi^c_t \) is the inflation gap given by

\[
\pi^c_t = \sum_{j=0}^{\infty} \omega^j E_t(\beta^\pi_{t+j} y^c_{t+j}) + \zeta_t. 
\]

(8)

The idiosyncratic term \( \varepsilon^\pi_t \) is added in equation (7) to capture measurement error and non-persistent shocks. Trend inflation \( \pi^r_t \) is modeled as a driftless random walk

\[
\pi^r_{t+1} = \pi^r_t + \exp \{ h^\pi_t \} \psi^\pi_t, \quad \psi^\pi_t \sim i.i.d. N(0, 1), 
\]

(9)

where the innovations \( \psi^\pi_t \) are allowed to exhibit stochastic volatility to capture changes in the dynamics of long-run inflation, possibly driven by different monetary policy regimes (see e.g. Stock and Watson, 2007, for a similar specification). The slope of the Phillips curve \( \beta^\pi_t \) is allowed
to change over time according to a random walk

$$\beta_{t+1} = \beta_t + \eta_t^\pi,$$

$$\eta_t^\pi \sim i.i.d. \mathcal{N}(0, \sigma_{\eta,\pi}^2). \quad (10)$$

We model the temporary inflation component $\zeta_t$ in equation (8) as an AR(1) process

$$\zeta_t + 1 = \phi \zeta_t + \exp \{ h_t^\zeta \} \psi_t^\zeta, \quad \psi_t^\zeta \sim i.i.d. \mathcal{N}(0, 1). \quad (11)$$

Given the DGPs of $y_c^t$ and $\beta_t^\pi$ in equations (4) and (10) and the assumption that $\psi_t^\zeta$ and $\eta_t^\pi$ are mutually uncorrelated error terms, the inflation gap term $\sum_{j=0}^{\infty} \omega_j E_t(\beta_t^\pi + j y_c^t + j)$ in equation (8) can be expressed as

$$\pi_t^c = \beta_t^\pi \left[ \begin{array}{c} 1 \\ 0 \end{array} \right] \left( \left[ \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right] - \omega \left[ \begin{array}{cc} \rho_1 & \rho_2 \\ 1 & 0 \end{array} \right] \right)^{-1} \left[ \begin{array}{c} y_c^t \\ y_c^{t-1} \end{array} \right], \quad (12)$$

$$= \frac{\beta_t^\pi}{1 - \omega \rho_1 - \omega^2 \rho_2} \left( y_c^t + \omega \rho_2 y_c^{t-1} \right). \quad (13)$$

Hence, the model for inflation in equation (7) can be rewritten as

$$\pi_t = \pi_t^c + \beta_t^\pi y_c^t + \zeta_t + \varepsilon_t^\pi,$$  

where $y_c^t = \frac{1}{1 - \omega \rho_1 - \omega^2 \rho_2} \left( y_c^t + \omega \rho_2 y_c^{t-1} \right)$.

**Unemployment: a time-varying Okun’s law relation**

We assume that the unemployment rate $u_t$ has the following trend/cycle representation

$$u_t = u_t^\tau + \beta_t^u y_c^t + \varepsilon_t^u,$$

$$\varepsilon_t^u \sim i.i.d. \mathcal{N}(0, \sigma_{\varepsilon,\pi}^2), \quad (15)$$

where $\varepsilon_t^u$ captures measurement error and non-persistent shocks. Following, among others, Staiger et al. (1997) and Laubach (2001) we model trend unemployment $u_t^\tau$ as a random walk process

$$u_{t+1}^\tau = u_t^\tau + \exp \{ h_t^u \} \psi_t^u, \quad \psi_t^u \sim i.i.d. \mathcal{N}(0, 1). \quad (16)$$

We give this component a NAIRU interpretation, i.e. as long as the observed unemployment rate equals this long-run trend, no inflationary pressure emanates from the labor market. Again, we allow for stochastic volatility in the trend component so that the variance of permanent shocks to the labor market can differ over time. The strength of Okun’s law $\beta_t^u$ is allowed to change over time according to a random walk process

$$\beta_{t+1}^u = \beta_t^u + \eta_t^u,$$

$$\eta_t^u \sim i.i.d. \mathcal{N}(0, \sigma_{\eta,\pi}^2). \quad (17)$$

7
Stochastic volatilities

All stochastic volatilities are modeled as random walks

\[ h_{k,t+1} = h_{k,t} + \gamma_{k,t}, \quad \gamma_{k,t} \sim i.i.d. \mathcal{N}(0, \sigma_{\gamma,k}^2), \tag{18} \]

for \( k = y, \pi, u, c, \zeta \). A key feature of the stochastic volatility components \( \exp \{ h_{k,t} \} \psi_{k,t} \) is that they are nonlinear but can be transformed into linear components by taking the logarithm of their squares

\[ \ln \left( \exp \{ h_{k,t} \} \psi_{k,t} \right)^2 = 2h_{k,t} + \ln \left( \psi_{k,t} \right)^2, \tag{19} \]

where \( \ln \left( \psi_{k,t} \right)^2 \) is log-chi-square distributed with expected value \(-1.2704\) and variance 4.93. Following Kim et al. (1998), we approximate the linear model in (19) by an offset mixture time series model as

\[ g_{k,t} = 2h_{k,t} + \epsilon_{k,t}, \tag{20} \]

where \( g_{k,t} = \ln \left( \left( \exp \{ h_{k,t} \} \psi_{k,t} \right)^2 + c \right) \) with \( c = .001 \) being an offset constant, and the distribution of \( \epsilon_{k,t} \) being the following mixture of normals

\[ f \left( \epsilon_{k,t} \right) = \sum_{i=1}^{M} q_i f_N \left( \epsilon_{k,t} | m_i - 1.2704, \nu_i^2 \right), \tag{21} \]

with component probabilities \( q_i \), means \( m_i - 1.2704 \) and variances \( \nu_i^2 \). Equivalently, this mixture density can be written in terms of the component indicator variable \( \iota_{k,t} \) as

\[ \epsilon_{k,t} | \left( \iota_{k,t} = i \right) \sim \mathcal{N} \left( m_i - 1.2704, \nu_i^2 \right), \quad \text{with} \quad \Pr \left( \iota_{k,t} = i \right) = q_i. \tag{22} \]

Following Omori et al. (2007), we use a mixture of \( M = 10 \) normal distributions to make the approximation to the log-chi-square distribution sufficiently good. Values for \( \{ q_i, m_i, \nu_i^2 \} \) are provided by Omori et al. in their Table 1.

2.2 Stochastic model specification search

The empirical model outlined in Section 2.1 nests a number of model specifications used in the recent literature. The univariate unobserved components model for inflation examined by Stock and Watson (2007) can for instance be obtained by restricting \( \beta_{\pi,t}, \varrho \) and \( \sigma_{\epsilon,\pi}^2 \) to zero. The bivariate unobserved components specification for inflation and unemployment of Stella and Stock (2012) is nested when we replace the output gap by the unemployment gap, set \( \varrho \) and \( \sigma_{\epsilon,x}^2 \) to zero and restrict \( \beta_7^T \) to be constant.

A key question therefore is which model components are relevant and which can be excluded. However, model specification for state space models is a difficult task as this leads to non-regular testing problems. Consider for instance the question whether the slope of the Phillips curve should
be modeled as constant or varying over time. This implies testing \( \sigma_{\eta,\pi}^2 = 0 \) against \( \sigma_{\eta,\pi}^2 > 0 \), which is a non-regular testing problem as the null hypothesis lies on the boundary of the parameter space. A similar problem arises when testing whether the temporary component \( \zeta_t \) should be included in equation (14) or whether the stochastic volatilities are relevant.

As an alternative, we use a Bayesian stochastic model specification search. The Bayesian approach is well-suited to deal with non-regular testing problems by computing posterior probabilities for each of the candidate models. In particular, Frühwirth-Schnatter and Wagner (2010) show how to extend Bayesian variable selection in standard regression models to state space models. Their approach relies on a non-centered parameterization of the state space model in which (i) binary stochastic indicators for each of the model components are sampled together with the parameters and (ii) the standard inverse Gamma prior for the variances of innovations to the components is replaced by a Gaussian prior centered at zero for the square root of these variances. The exact implementation applied to our state space model is outlined below.

**Non-centered parameterization**

Frühwirth-Schnatter and Wagner (2010) argue that a first piece of information on the hypothesis whether a variance parameter in a state space model is zero or not can be obtained by considering a non-centered parameterization. For the variances of the innovations to the slope of the Phillips curve and Okun’s law, i.e. \( \sigma_{\eta,\pi}^2 \) and \( \sigma_{\eta,u}^2 \), this implies rearranging equations (10) and (17) to

\[
\begin{align*}
\beta_{j,t+1} &= \beta_{j,t} + \sigma_{\eta,j} \tilde{\beta}_{j,t+1}, \\
\text{with} \quad \tilde{\beta}_{j,t+1} &= \tilde{\beta}_{j,t} + \tilde{\eta}_t, \quad \tilde{\beta}_{1,t} = 0, \quad \tilde{\eta}_t \sim \text{i.i.d.} \mathcal{N}(0,1),
\end{align*}
\]

for \( j = \pi, u \) and where \( \beta_{j,t} \) is the initial value of the level of \( \beta_{j,t} \). A crucial aspect of the non-centered parameterization is that it is not identified, i.e. the signs of \( \sigma_{\eta,j} \) and \( \tilde{\beta}_{j,t} \) can be changed by multiplying both with -1 without changing their product in equation (23). As a result of the non-identification, the likelihood function is symmetric around 0 along the \( \sigma_{\eta,j} \) dimension and therefore multimodal. If the slope of the Phillips curve varies over time, i.e. \( \sigma_{\eta,j}^2 > 0 \), then the likelihood function will concentrate around the two modes \(-\sigma_{\eta,j} \) and \( \sigma_{\eta,j} \). For \( \sigma_{\eta,j}^2 = 0 \) the likelihood function will become unimodal around zero. As such, allowing for non-identification of \( \sigma_{\eta,j} \) provides useful information on whether \( \sigma_{\eta,j}^2 > 0 \).

Likewise, the non-centered parameterization of the stochastic volatility terms in equation (18) is given by

\[
\begin{align*}
h_{t+1}^k &= h_{t}^k + \sigma_{\gamma,k} \tilde{h}_{t+1}^k, \\
\text{with} \quad \tilde{h}_{t+1}^k &= \tilde{h}_t^k + \tilde{\gamma}_t, \quad \tilde{h}_1^k = 0, \quad \tilde{\gamma}_t \sim \text{i.i.d.} \mathcal{N}(0,1),
\end{align*}
\]

for \( k = y, \pi, u, c, \zeta \) and where \( h_0^k = 0 \) is the initial value of the level of \( h_t^k \).
Finally, the non-centered parameterization of the time-varying drift in equation (3) is given by

$$\kappa_{t+1} = \kappa_0 + \sigma_{\psi,\kappa} \tilde{\kappa}_{t+1},$$  \hspace{1cm} (27)

with

$$\tilde{\kappa}_{t+1} = \tilde{\kappa}_t + \tilde{\psi}_t, \quad \tilde{\kappa}_1 = 0, \quad \tilde{\psi}_t \sim \text{i.i.d.} \mathcal{N}(0,1),$$  \hspace{1cm} (28)

and where $\kappa_0$ is the initial value of the level of $\kappa_t$.

**Parsimonious specification**

A second advantage of the non-centered parameterization is that when e.g. $\sigma_{\eta,\pi}^2 = 0$ the transformed component $\tilde{\beta}_t^j$, in contrast to $\beta_t^j$, does not degenerate to the time-invariant slope of the Phillips curve as this is now represented by $\beta_0^j$. As such, the question whether the slopes of the Phillips curve and Okun’s law are varying over time or not can be expressed as a variable selection problem in equation (23). To this aim Frühwirth-Schnatter and Wagner (2010) introduce the parsimonious specification

$$\beta_t^j = \beta_0^j + \delta_j \sigma_{\eta,j} \tilde{\beta}_t^j,$$  \hspace{1cm} (29)

for $j = \pi, u$ and where $\delta_j$ is a binary indicator which is either 0 or 1. If $\delta_j = 0$, the component $\tilde{\beta}_t^j$ drops from the model such that $\beta_0^j$ represents the constant slope parameter. If $\delta_j = 1$ then $\tilde{\beta}_t^j$ is included in the model and $\sigma_{\eta,j}$ is estimated from the data. In this case $\beta_0^j$ is the initial value of the slope parameter.

Likewise, the parsimonious non-centered parameterization of the stochastic volatility terms in equation (25) is given by

$$h_t^k = h_0^k + \theta_k \sigma_{\gamma,k} \tilde{h}_t^k,$$  \hspace{1cm} (30)

for $k = y, \pi, u, c, \zeta$ and where $\theta_k$ is again a binary indicator that is either 0 or 1. If $\theta_k = 0$, the component $\tilde{h}_t^k$ drops from the model such that $(\exp\{h_0^k\})^2$ is the constant variance of $\psi_t^k$. If $\theta_k = 1$ then $\tilde{h}_t^k$ is included in the model and $\sigma_{\gamma,j}$ is estimated from the data. In this case $(\exp\{h_0^k\})^2$ is the initial value of the time-varying variance of $\psi_t^k$.

Finally, the parsimonious non-centered parameterization of the time-varying drift term in equation (27) is given by

$$\kappa_t = \kappa_0 + \lambda_{\kappa} \sigma_{\psi,\kappa} \tilde{\kappa}_t,$$  \hspace{1cm} (31)

where $\lambda_{\kappa}$ is a binary indicator that is either 0 or 1. If $\lambda_{\kappa} = 0$, the component $\tilde{\kappa}_t$ drops from the model such that $\kappa_0$ is the constant drift in potential output. If $\lambda_{\kappa} = 1$ then $\tilde{\kappa}_t$ is included in the model and $\sigma_{\psi,\kappa}$ is estimated from the data. In this case $\kappa_0$ is the initial value of the time-varying drift $\kappa_t$.

Collecting the binary indicators in the vector $\mathcal{M} = (\delta_\pi, \delta_u, \theta_y, \theta_\pi, \theta_u, \theta_c, \theta_\zeta, \lambda_\kappa)$, each model is indicated by a value for $\mathcal{M}$, e.g. $\mathcal{M} = (0,1,0,0,0,1,0,1)$ is a model with a constant Phillips curve slope, a time-varying Okun’s law coefficient, stochastic volatility in the innovations to the
output gap component, a constant variance for the innovations to the trend components in output, inflation and unemployment as well as to the AR(1) inflation gap component and a time-varying drift in potential output.

**Gaussian prior centered at zero**

Our Bayesian estimation approach requires choosing prior distributions for the model parameters \( \rho = (\rho_1, \rho_2) \), \( \varrho, \beta_0 = (\beta^\pi_0, \beta^u_0) \) and \( h_0 = (h^y_0, h^\pi_0, h^u_0, h^c_0) \), for the binary indicators \( M \) and for the variances of the idiosyncratic factors \( \sigma^2 = (\sigma^2_{\varepsilon, y}, \sigma^2_{\varepsilon, \pi}, \sigma^2_{\varepsilon, u}) \), the innovations to the drift component \( \sigma^2_{\psi, \kappa} \), the time-varying parameters \( \sigma^2_\eta = (\sigma^2_{\eta, \pi}, \sigma^2_{\eta, u}) \) and the stochastic volatility components \( \sigma^2_\gamma = (\sigma^2_{\gamma, y}, \sigma^2_{\gamma, \pi}, \sigma^2_{\gamma, u}, \sigma^2_{\gamma, c}, \sigma^2_{\gamma, c}) \).

It is well-known that when using an inverse Gamma prior distribution for the variance parameters, the choice of the shape and scale hyperparameters that define this distribution have a strong influence on the posterior when the true value of the variance is close to zero. More specifically, as the inverse Gamma does not have probability mass at zero, using it as a prior distribution tends to push the posterior density away from zero. This is of particular importance when estimating the variances of the innovations to the time-varying parameters, to the drift in potential output and to the stochastic volatilities, as for these components we want to decide whether they are relevant or not. A further important advantage of the non-centered parameterization is therefore that it allows us to replace the standard inverse Gamma prior on a variance parameter \( \sigma^2 \) by a Gaussian prior centered at zero on \( \sigma \). Centering the prior distribution at zero makes sense as for both \( \sigma^2 = 0 \) and \( \sigma^2 > 0 \), \( \sigma \) is symmetric around zero. Frühwirth-Schnatter and Wagner (2010) show that, compared to using an inverse Gamma prior for \( \sigma^2 \), the posterior density of \( \sigma \) is much less sensitive to the hyperparameters of the Gaussian distribution and is not pushed away from zero when \( \sigma^2 = 0 \).

As such we choose a Gaussian prior distribution centered at zero for \( \sigma_\eta, \sigma_{\psi, \kappa} \) and \( \sigma_\gamma \), which are the standard deviations of the innovations to the time-varying parameters, to the drift in potential output and to the stochastic volatilities. For the variances of the idiosyncratic factors \( \sigma^2_{\varepsilon} \), which are always included in the model, we choose the standard inverse Gamma prior distribution. For each of the model parameters in \( \rho, \varrho \) and \( \beta \) we assume a normal prior distribution. Details on the prior distributions are presented in Section 3.2 below. For the binary indicators \( M \) we choose a uniform prior distribution over all combinations of the indicators such that each model has the same prior probability, i.e. \( p(M) = 2^{-8} \), and each model component has a prior probability \( p_0 = 0.5 \) of being included in the model.

**2.3 MCMC algorithm**

In a standard linear Gaussian state space model, the Kalman filter can be used to filter the unobserved states from the data and to construct the likelihood function such that the unknown parameters can be estimated using maximum likelihood. However, the inclusion of the time-varying parameters \( \beta^\pi_t \) and \( \beta^u_t \) on the unobserved output gap \( y^*_t \) and the stochastic volatilities \( h^k_t \) in the state space model given in eq. (1) - (18) and the use of the stochastic model specification search outlined in Section 2.2 imply a highly non-linear estimation problem for which the standard
approach via the Kalman filter and maximum likelihood is not feasible. Instead we use the Gibbs sampler which is a MCMC method to simulate draws from the intractable joint and marginal posterior distributions of the unknown parameters and the unobserved states using only tractable conditional distributions. Intuitively, this amounts to reducing the complex non-linear model into a sequence of blocks for subsets of parameters/states that are tractable conditional on the other blocks in the sequence.

For notational convenience, define a state vector $\alpha_t = (y_t^\tau, \pi_t^\tau, u_t^\tau, y_t^c, \zeta_t)$, a time-varying parameter vector $\beta_t = (\beta_\pi^\tau, \beta_u^\tau)$, a stochastic volatilities vector $h_t = (h_y^\tau, h_\pi^\tau, h_u^\tau, h_c^\tau, h_\zeta^\tau)$ and an indicator vector $\iota_t = (\iota_y^\tau, \iota_\pi^\tau, \iota_u^\tau, \iota_c^\tau, \iota_\zeta^\tau)$. The unknown parameters are collected in the vector $\phi = (\rho, \varrho, \beta_0, \sigma, \sigma_2^\epsilon)$, with $\sigma = (\sigma_\eta, \sigma_\psi, \kappa, \sigma_\gamma)$. Finally, let $x_t = (y_t, \pi_t, u_t)$ be the data vector. Stacking observations over time, we denote $x_t = \{x_t\}_{t=1}^T$ and similarly for $\alpha, \beta, h$ and $\iota$. The posterior density of interest is then given by $f(\alpha, \beta, h, \iota, \phi, M|x)$. Following Frühwirth-Schnatter and Wagner (2010) our MCMC scheme is as follows:

1. Sample the binary indicators in $M$ from $f(M|\alpha, \beta, h, x)$ marginalizing over the parameters $\phi$ and sample the unrestricted parameters in $\phi$ from $f(\phi|\alpha, \beta, h, M, x)$ while setting the restricted parameters, i.e. the elements in $\sigma$ for which the corresponding component is not included in the model, equal to 0.

2. Sample the trend and temporary components $\alpha$ from $f(\alpha|\beta, h, \phi, M, x)$, the time-varying parameters $\beta$ from $f(\beta|\alpha, h, \phi, M, x)$, the mixture indicators $\iota$ from $f(\iota|\alpha, \beta, h, \phi, M, x)$ and the stochastic volatilities $h$ from $f(h|\alpha, \beta, \iota, \phi, M, x)$.

3. Perform a random sign switch for $\sigma_{\eta,j}$ and $\{\tilde{\beta}_j^\pi\}_{t=1}^T$; for $\sigma_{\psi,\kappa}$ and $\{\tilde{\kappa}_t\}_{t=1}^T$ and for $\sigma_{\gamma,k}$ and $\{\tilde{h}_k^\tau\}_{t=1}^T$, e.g. $\sigma_{\eta,\pi}$ and $\{\tilde{\beta}_\pi^\tau\}_{t=1}^T$ are left unchanged with probability 0.5 while with the same probability they are replaced by $-\sigma_{\eta,\pi}$ and $\{-\tilde{\beta}_\pi^\tau\}_{t=1}^T$.

Given an arbitrary set of starting values, sampling from these blocks is iterated $J$ times and, after a sufficiently long burn-in period $B$, the sequence of draws $(B+1, \ldots, J)$ approximates a sample from the virtual posterior distribution $f(\alpha, \beta, h, \iota, \phi, M|x)$. Details on the exact implementation of each of the blocks can be found in Appendix A. The results reported below are based on 70,000 Gibbs sampler iterations, with the first 20,000 discarded as a burn-in period. We store 10,000 useful draws.

3 Estimation results

3.1 Data

We estimate the model using quarterly U.S. data from 1959Q2 - 2014Q3. Inflation is measured by the annualized quarterly change in the core personal consumption expenditures (PCE) index. For unemployment we use the civilian unemployment rate as collected by the Bureau of Labor Statistics. Output is measured by the log ($\times100$) of real GDP. All series are taken from St. Louis Federal Reserve Economic Data.
3.2 Prior choice

Table 1 reports summary information on our prior distributions for the unknown parameters. For the variance parameters of the idiosyncratic factors $\sigma^2 = \left( \sigma^2_{\varepsilon,y}, \sigma^2_{\varepsilon,\pi}, \sigma^2_{\varepsilon,u} \right)$ we use the inverse Gamma prior $IG(c_0, C_0)$ where the shape $c_0 = \nu_0 T$ and scale $C_0 = s_0 \sigma^2_0$ parameters are calculated from the prior belief $\sigma^2_0$ about the variance parameter and the prior strength $\nu_0$ which is expressed as a fraction of the sample size $T$. Following the notation in Frühwirth-Schnatter and Wagner (2010), for the remaining parameters we use a Gaussian prior $N(a_0, A_0 \sigma^2_0)$ in a regression with homoskedastic errors and $N(a_0, A_0)$ when the errors exhibit stochastic volatility. Details on the notation are given in Appendix A. Each of the prior choices is discussed below. Note that in Table 1 and in the text we report and discuss standard deviations rather than variances as the former are easier to interpret.

- **Idiosyncratic components**, $\sigma^2_{\varepsilon,y}, \sigma^2_{\varepsilon,\pi}, \sigma^2_{\varepsilon,u}$: We set the prior beliefs to $\sigma_{\varepsilon,y} = 0.1$, $\sigma_{\varepsilon,\pi} = 1.0$, and $\sigma_{\varepsilon,u} = 0.5$. The strength of all three priors is 0.1. The larger value for $\sigma_{\varepsilon,\pi}$ is in line with the literature, which usually finds relative large measurement errors in inflation.

- **Volatility of trend components**, $\exp \{ h^y \}, \exp \{ h^\pi \}, \exp \{ h^u \}$: The prior beliefs $a_0$ for the constant volatility part $h_0$ of the level shocks to potential output, trend inflation, and the NAIRU are set to $\ln(0.1)$, $\ln(0.2)$, and $\ln(0.01)$ respectively with the prior standard deviation $\sqrt{A_0}$ set to 0.1. Note that the prior belief $\ln(0.1)$ for potential output implies that, if there is no time-varying volatility, 95% of the innovations lie between $-0.2$ and $+0.2$ per quarter. For inflation and unemployment the 95% interval ranges from $-0.4$ to $+0.4$ and $-0.1$ to $+0.1$ respectively. These values are within the range of previous estimates and are of an economically reasonable value. The prior distribution on the time-varying part of the volatility of the trends is uninformative and centered at zero: $\sigma_{\gamma,k} \sim N(0,1)$, for $k = y, \pi, u$.

- **Potential output growth**, $\kappa$: In line with existing estimates, our prior belief about the time-invariant part of the output drift is given by $\kappa \sim N(0.75, 0.1^2)$. For example, Morley et al. (2003), Sinclair (2009) and Mitra and Sinclair (2012) find values between 0.79 and 0.86 for quarterly U.S. postwar data. Similar to the volatilities of the trends, we set the time-varying part of potential output growth to $\sigma_{\psi,\kappa} \sim N(0,1)$.

- **Output gap**, $\rho$, $\exp \{ h^\psi \}$: While the output gap is stationary by assumption, it is often found to be a very persistent process (see e.g. Morley et al., 2003; Kim et al., 2014). In order to ensure stationarity, we find it useful to impose prior information on the sum of the AR(2) parameters instead of restricting each parameter separately. Hence, we use an informative prior on the sum $(\rho_1 + \rho_2) \sim N(0.9, 0.015^2)$ and a much less informative prior on the first lag $\rho_1 \sim N(1.25, 0.5^2)$. The prior belief of 0.90 for $(\rho_1 + \rho_2)$ is an average of values typically found in the literature on trend-cycle decomposition of U.S. GDP (see e.g. Kuttner, 1994; Kuttner, 1999; Kuttner, 2001).

---

3. Since this prior is conjugate, $\nu_0 T$ can be interpreted as the number of fictitious observations used to construct the prior belief $\sigma^2_0$.

4. See for instance Morley et al. (2013); Kim et al. (2014); Stock and Watson (2007). Regarding the smoothness of the NAIRU, we are close to Fleischman and Roberts (2011) who estimate the NAIRU’s standard deviation around 0.1.
Table 1: Prior distributions of model parameters

| Inverse Gamma priors: $IG(c_0, C_0) = IG(\nu_0T, \nu_0T\sigma^2_\epsilon)$ | Percentiles |
|---|---|---|---|
| $\sigma_0$ | $\nu_0$ | 2.5% | 97.5% |
| idiosyncratic component output | $\sigma_{x,y}$ | 0.10 | 0.10 | 0.080 | 0.141 |
| idiosyncratic component inflation | $\sigma_{x,\pi}$ | 1.00 | 0.10 | 0.777 | 1.408 |
| idiosyncratic component unemployment | $\sigma_{x,u}$ | 0.50 | 0.10 | 0.387 | 0.701 |

| Gaussian priors homoskedastic errors: $N(a_0, A_0\sigma^2_\epsilon)$ | Percentiles |
|---|---|---|---|
| Regression parameters | $a_0$ | $\sqrt{A_0} \times \sigma_\epsilon$ | 2.5% | 97.5% |
| const. Phillips curve slope | $\beta_0^{\pi}$ | 0.20 | 0.25 | -0.290 | 0.690 |
| const. Okun coefficient | $\beta_0^{a}$ | -0.50 | 0.25 | -0.745 | -0.255 |
| Non-centered components | | | | |
| std. of time-varying Phillips curve | $\sigma_{t,\pi}$ | 0.00 | 1.00 | -1.960 | 1.960 |
| std. of time-varying Okun coefficient | $\sigma_{t,u}$ | 0.00 | 1.00 | -0.980 | 0.980 |

| Gaussian priors SV errors: $N(a_0, A_0)$ | Percentiles |
|---|---|---|---|
| Regression parameters | $a_0$ | $\sqrt{A_0}$ | 2.5% | 97.5% |
| 1st AR lag: output gap | $\rho_1$ | 1.25 | 0.50 | 0.270 | 2.230 |
| sum of AR lags: output gap | $\rho_1 + \rho_2$ | 0.90 | 0.015 | 0.871 | 0.930 |
| AR lag: AR(1) inflation component | $\varphi$ | 0.70 | 0.05 | 0.602 | 0.798 |
| const. output drift | $\kappa_0$ | 0.75 | 0.10 | 0.554 | 0.946 |
| Stochastic volatility parameters | | | | |
| const. volatility of potential output | $h_0^\gamma$ | ln (0.10) | 0.10 | ln (0.082) | ln (0.122) |
| const. volatility of trend inflation | $h_0^\pi$ | ln (0.20) | 0.10 | ln (0.164) | ln (0.243) |
| const. volatility of NAIRU | $h_0^{\pi}$ | ln (0.05) | 0.10 | ln (0.041) | ln (0.061) |
| const. volatility of output gap | $h_0^u$ | ln (0.60) | 0.10 | ln (0.493) | ln (0.730) |
| const. volatility of temporary inflation | $h_0^\pi$ | ln (0.70) | 0.10 | ln (0.575) | ln (0.852) |
| Non-centered components | | | | |
| std. of SV: potential output | $\sigma_{\gamma,y}$ | 0.00 | 1.00 | -1.960 | 1.960 |
| std. of SV: trend inflation | $\sigma_{\gamma,\pi}$ | 0.00 | 1.00 | -1.960 | 1.960 |
| std. of SV: NAIRU | $\sigma_{\gamma,u}$ | 0.00 | 1.00 | -1.960 | 1.960 |
| std. of SV: output gap | $\sigma_{\gamma,c}$ | 0.00 | 1.00 | -1.960 | 1.960 |
| std. of SV: AR(1) inflation component | $\sigma_{\gamma,\zeta}$ | 0.00 | 1.00 | -1.960 | 1.960 |
| std. of time-varying output drift | $\alpha_{\psi,\kappa}$ | 0.00 | 1.00 | -1.960 | 1.960 |

Notes: We set IG priors on the variance parameters $\sigma^2$ but in the top panel of this table we report details on the implied prior distribution for the standard deviations $\sigma$ as these are easier to interpret. Likewise, in the bottom panel of the table we report $\sqrt{A_0}$ instead of $A_0$. For the stochastic volatility parameters $h_0$, we report a logarithm expression for the mean and percentiles as the arguments can then easily be interpreted as the mean and percentiles of $exp(h_0)$.

Morley et al., 2003; Luo and Startz, 2014). The small prior standard deviation of 0.015 is to ensure that the output gap is stationary. Setting a lower belief together with a higher
standard deviations results in a similar posterior, though. The prior belief of 1.25 for $\rho_1$, which implies a prior belief of -0.35 for $\rho_2$, is in line with the typical hump-shaped pattern in response to cyclical shocks. Nevertheless, with a prior standard deviation of 0.5 we are very uninformative on these individual parameters. The prior distribution for the time-invariant part of the cyclical volatility is given by $h_0^c \sim \mathcal{N}(\ln(0.6), 0.1^2)$, implying that 95% of the shocks lie between $-1.2$ and $+1.2$. Again, an uninformative prior for the time-varying volatility part is used, i.e. $\gamma_{c,e} \sim \mathcal{N}(0, 1)$.

- **AR(1) inflation component**, $\varrho$, $\exp\{h_0^c\}$: We set the prior distribution for the autoregressive coefficient of the AR(1) inflation component to $\varrho \sim \mathcal{N}(0.7, 0.05^2)$. The relative small standard deviation ensures that $\varrho$ lies within a region of medium persistence. With values too close to one, the AR(1) component becomes highly persistent and soaks up all variation in trend inflation. If $\varrho$ becomes too small, $\zeta$ is indistinguishable from the white noise inflation component $\varepsilon_\pi^t$. The prior distribution of the time-invariant part of the volatility component is set to $h_0^\gamma \sim \mathcal{N}(\ln(0.7), 0.1^2)$. A loose prior is used for the standard deviation of the time-varying component: $\gamma_{c,e} \sim \mathcal{N}(0, 1)$, allowing for a high degree of time variation in $\zeta$ as found in Kim et al. (2014).

- **Slope of Phillips curve**, $\beta^\pi$: The literature provides a wide range of estimates for the slope of the Phillips curve, depending on the specification. Backward-looking models typically yield the highest estimates, while in forward-looking specifications $\beta^\pi$ is often found small and statistically insignificant (see e.g. Kim et al., 2014). A potential explanation for these findings can be found in Lee and Nelson (2007). They show that Phillips curve slope estimates can vary greatly, depending on the inflation expectations horizon and the persistence of the cyclical driving variable. Our prior choice reflects the wide range of estimates found in the literature. For the time-invariant part $\beta_0^\pi$ we set a prior distribution of $\beta_0^\pi \sim \mathcal{N}(0.2, 0.25^2)$, which implies that the 95% interval of our prior belief ranges from -0.3 to 0.7. Our prior belief about the degree of time variation in the Phillips curve is uninformative, i.e. $\eta_{\pi,e} \sim \mathcal{N}(0, 1)$.

- **Okun coefficient**, $\beta^u$: According to Lee (2000) and Reifschneider et al. (2013) the impact of the unemployment gap on the output gap is close to $-2$ for the U.S., which would correspond to a value of $-0.5$ in our model as we express this relationship in reverse. Owyang and Sekhposyan (2012) estimate a rolling regression and find a very similar value on average. Thus, we set the prior distribution to $\beta_0^u \sim \mathcal{N}(-0.5, 0.125^2)$. The prior on the degree of time variation in Okun’s law is set to $\eta_{u,e} \sim \mathcal{N}(0, 1)$.

3.3 Results stochastic model specification search

We first estimate an unrestricted model with all binary indicators set to one to generate posterior distributions for the standard deviations ($\sigma$) of the innovations to the 8 non-centered components of interest. If these distributions are bimodal, with low or no probability mass at zero, this can be taken as a first indication of time variation in the considered component. Results are shown in Figure 1. Clear-cut bimodality is found in the posterior distribution of the standard deviation of the innovations to the Okun’s law parameter ($\eta_{u,e}$), to the volatility of the output gap ($\gamma_{c,e}$)
and the temporary inflation component ($\sigma_{\gamma,\zeta}$) and to the drift in potential output ($\sigma_{\psi,\kappa}$). For the stochastic volatility of trend inflation evidence is less clear. While the distribution $\sigma_{\gamma,\pi}$ appears to have two modes, it also has a considerable probability mass at zero. For the innovations to the Phillips curve parameter and to the stochastic volatility components in trend output and trend unemployment, the posterior distributions of $\sigma_{\eta,\pi}$, $\sigma_{\gamma,y}$ and $\sigma_{\gamma,u}$ are clearly unimodal at zero. This suggests that these components are stable over time.

As a more formal test for time variation, we next sample the stochastic binary indicators together with the other parameters in the model. Table 2 displays the individual posterior probabilities for the binary indicators being one. These probabilities are calculated as the average selection frequencies over all iterations of the Gibbs sampler. The second row shows results for our benchmark case $A_0 = 1$. This implies a relatively loose prior on the degree of time variation $\sigma$. To check robustness, the other rows show results over alternative values for $A_0$. The first row shows results for the case where $A_0 = 0.1$. This corresponds to a relatively stronger prior that allows for less time variation. The third and fourth row show results for diffuse prior distributions that allow for large variances on the time-varying components. The following conclusions can be drawn. First, the stochastic model specification search rejects time variation in the slope of the Phillips curve but clearly favors time variation in the Okun’s law parameter. Over all four prior settings, the posterior probability of time variation is below 1% for the Phillips curve slope while being 100% for the Okun’s law parameter. Second, the selection procedure clearly favors a model with a time-varying drift in potential output. Third, stochastic volatility is selected for the cyclical but not for the trend components in output, inflation and unemployment. In our benchmark case ($A_0 = 1$) the posterior probabilities of a stochastic volatility component in the trend components varies between 9% and 17%, while these probabilities fall below 5% when more diffuse priors are used. When the prior distribution allows for less time variation ($A_0 = 0.1$), the inclusion prob-
abilities of the stochastic volatility components increase, but remain below 0.5. This finding is in line with the results in Kim et al. (2004), who find that U.S. aggregate real GDP underwent a volatility reduction in the early 1980s in its cyclical component but not in its trend component.

Table 2: Posterior inclusion probabilities for the binary indicators over different prior variances $A_0$

<table>
<thead>
<tr>
<th>Prior</th>
<th>Time-varying parameters</th>
<th>Stochastic volatilities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Phillips curve</td>
<td>Okun’s law</td>
</tr>
<tr>
<td>$p_0$</td>
<td>$A_0$</td>
<td>$\delta_\pi$</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.0078</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>0.0031</td>
</tr>
<tr>
<td>0.5</td>
<td>10</td>
<td>0.0009</td>
</tr>
<tr>
<td>0.5</td>
<td>100</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

In the baseline specification, we assign a 0.5 prior probability to each of the binary indicators being one. As noted by Scott and Berger (2010), this prior choice does not provide multiplicity control for the Bayesian variable selection. When the number of possible variables is very large and each of the binary indicators has a prior probability of 0.5, the fraction of selected variables will very likely be around 0.5. Our findings appear to be unaffected by this issue, though. First, the number of variables to be selected is only 8 in this paper. Second, we re-estimate the (unrestricted) model with different priors. Specifically, the prior inclusion probability on each of the 8 components is set to 0.1 and 0.9 respectively. The resulting posterior probabilities are reported in Table 3. For all prior choices the same model is selected, i.e. the indicators $\delta_\pi, \theta_\pi, \theta_\zeta$ and $\lambda_\kappa$ have inclusion probabilities of 100%, while the indicators $\delta_u, \theta_u, \theta_\kappa$ and $\theta_\zeta$ are excluded in the majority of all draws.

Besides inference on the importance of time variation in the individual components, the model selection search also allows to compute overall model probabilities. The introduction of 8 binary indicators leads to $2^8$ possible models. As 4 out of the 8 binary indicators have low individual probabilities, most models have a probability of zero. As a result, in the benchmark case where $A_0 = 1$ only 7 models are selected in more than 1% of the Gibbs iterations. The posterior probabilities for these models are reported in Table 4. The favored model has a time-varying Okun’s law parameter and stochastic volatility in the output gap and the transitory inflation component, while the Phillips curve slope is constant and there is no stochastic volatility in the three trend components. This model choice is robust to different prior specification, i.e. the model in row one has the highest probability for each of the four considered values of $A_0$. In the case of strict priors ($A_0 = 0.1$), this model has a posterior probability of 39%, which rises to 90% when diffuse priors ($A_0 = 100$) are used. The three other models with notable probabilities larger than

---

5The increase in the posterior probability may appear counter intuitive, but is due to the fact that by restricting the amount of time variation the competing models become similar in their marginal likelihoods and thus the posterior probability shrinks towards the prior probability $p_0 = 0.5$. 

17
Table 3: Posterior inclusion probabilities for the binary indicators over different prior probabilities $p_0$

<table>
<thead>
<tr>
<th>Prior</th>
<th>Time-varying parameters</th>
<th>Stochastic volatilities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Phillips curve</td>
<td>Okun’s law</td>
</tr>
<tr>
<td>$p_0$</td>
<td>$A_0$</td>
<td>$\delta_\pi$</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>0.0031</td>
</tr>
<tr>
<td>0.9</td>
<td>1</td>
<td>0.0222</td>
</tr>
<tr>
<td>0.1</td>
<td>1</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

5% include stochastic volatility in either potential output, trend inflation or in the NAIRU. As the variance of the prior $A_0$ increases, the probabilities of these models shrink towards zero.

Table 4: Posterior model probabilities over different prior variances $A_0$ (with $p_0 = 0.5$)

<table>
<thead>
<tr>
<th>Model</th>
<th>Posterior probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_\pi$</td>
<td>$\delta_u$</td>
</tr>
<tr>
<td>$A_0 = 0.1$</td>
<td>$A_0 = 1$</td>
</tr>
<tr>
<td>0 1 1 0 0 0 1 1</td>
<td>0.3907</td>
</tr>
<tr>
<td>0 1 1 0 0 1 1 1</td>
<td>0.1721</td>
</tr>
<tr>
<td>0 1 1 0 1 0 1 1</td>
<td>0.1033</td>
</tr>
<tr>
<td>0 1 1 0 1 1 1 1</td>
<td>0.0401</td>
</tr>
<tr>
<td>0 1 1 1 0 0 1 1</td>
<td>0.1468</td>
</tr>
<tr>
<td>0 1 1 1 0 1 1 1</td>
<td>0.0673</td>
</tr>
<tr>
<td>0 1 1 1 1 0 1 1</td>
<td>0.0497</td>
</tr>
</tbody>
</table>

For completeness, Figure 2 shows the evolution of the four components for which the time variation does not show up as relevant using the model selection. These components will be restricted to be constant in the remainder of this paper. The evolution of the significant time-varying components is discussed more in detail in below.

3.4 Parameter estimates and unobserved components

In this section we present the results of the model that is favored by the stochastic model selection. We will refer to this as the parsimonious model. As a convergence check we plot the 20th-order autocorrelations for all parameter and component draws in Figure 3. This diagnostic has been used before in Primiceri (2005) and Liu and Morley (2014). The majority of autocorrelations lie well below 0.1, while for a few parameters we find values between 0.2 and 0.3. Only one value is as high as 0.5. We take this as evidence for satisfactory convergence of the Markov-Chain.

The posterior distributions of the parsimonious model’s time-invariant parameters are plotted
**Figure 2:** Evolution of the time-varying components not selected by the model search (all binary indicators set to 1)

(a) Phillips curve
(b) SV potential output
(c) SV trend inflation
(d) SV NAIRU

**Figure 3:** 20th-order autocorrelations of all parameter and component draws (parsimonious model)
in Figure 4.\textsuperscript{6} Descriptive statistics are given in Table 5. For the standard deviations of the non-centered variables the posterior distributions are bimodal. Thus, we report descriptive statistics on the unimodal posterior of the respective squared standard deviation parameters. The evolution of the unobserved components is shown in Figures 5-9 and discussed more in detail below.

**Inflation**

Figure 5 plots actual inflation against the median of the posterior distribution of trend inflation and its 90\% highest posterior density (HPD) interval for the parsimonious model. Trend inflation evolves smoothly and tracks the low-frequency movements in observed inflation. It steadily rises over the Great Inflation period from the late 1960s until the late 1970s and then falls back during the disinflation period of the 1980s and 1990s. Since the late 1990s trend inflation remains low and stable at around 2\%. Our estimated trend inflation series is very similar to those reported by Cogley and Sbordone (2008), Kim et al. (2014) and Stella and Stock (2012). The standard deviation of innovations to trend inflation was found to be constant over time by the model selection procedure and is estimated with a posterior median of 0.21. This result is consistent with Kim et al. (2014) who find similar values over three distinct regimes and could not reject

\textsuperscript{6}Note that the posterior distributions of the standard deviations for the non-centered variables in the unrestricted model are reported in Figure 1. As there is no noticeable difference in these distributions in the parsimonious model, they are not included in Figure 4.
Table 5: Posterior distributions of model parameters (parsimonious model)

<table>
<thead>
<tr>
<th>Inverse Gamma</th>
<th>Percentiles</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>median</td>
<td>2.5%</td>
<td>97.5%</td>
</tr>
<tr>
<td>idiosyncratic component output $\sigma_{\varepsilon,y}$</td>
<td>0.154</td>
<td>0.097</td>
<td>0.217</td>
</tr>
<tr>
<td>idiosyncratic component inflation $\sigma_{\varepsilon,\pi}$</td>
<td>0.547</td>
<td>0.482</td>
<td>0.620</td>
</tr>
<tr>
<td>idiosyncratic component unemployment $\sigma_{\varepsilon,u}$</td>
<td>0.239</td>
<td>0.212</td>
<td>0.269</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gaussian</th>
<th>Percentiles</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression parameters</td>
<td>median</td>
<td>2.5%</td>
<td>97.5%</td>
</tr>
<tr>
<td>const. Phillips curve slope $\beta_0^\pi$</td>
<td>0.077</td>
<td>-0.094</td>
<td>0.262</td>
</tr>
<tr>
<td>const. Okun coefficient $\beta_0^u$</td>
<td>-0.447</td>
<td>-0.557</td>
<td>-0.343</td>
</tr>
<tr>
<td>1st AR lag: output gap $\rho_1$</td>
<td>1.377</td>
<td>1.214</td>
<td>1.536</td>
</tr>
<tr>
<td>sum of AR lags: output gap $\rho_1 + \rho_2$</td>
<td>0.942</td>
<td>0.918</td>
<td>0.964</td>
</tr>
<tr>
<td>AR lag: AR(1) inflation component $\varrho$</td>
<td>0.760</td>
<td>0.662</td>
<td>0.862</td>
</tr>
<tr>
<td>const. output drift $\kappa_0$</td>
<td>1.035</td>
<td>0.926</td>
<td>1.144</td>
</tr>
</tbody>
</table>

Stochastic volatility parameters

| const. volatility of potential output $\exp\{h_0^y\}$ | 0.131 | 0.110 | 0.158 |
| const. volatility of trend inflation $\exp\{h_0^\pi\}$ | 0.207 | 0.174 | 0.246 |
| const. volatility of NAIRU $\exp\{h_0^u\}$ | 0.078 | 0.068 | 0.091 |
| const. volatility of output gap $\exp\{h_0^c\}$ | 0.544 | 0.463 | 0.632 |
| const. volatility of temporary inflation $\exp\{h_0^z\}$ | 0.507 | 0.432 | 0.594 |

Non-centered components

| variance of time-varying Okun coefficient $\sigma_{\eta,u}^2$ | 0.0019 | 0.0008 | 0.0035 |
| variance of SV: output gap $\sigma_{\gamma,c}^2$ | 0.0387 | 0.0108 | 0.0890 |
| variance of SV: AR(1) inflation component $\sigma_{\gamma,z}^2$ | 0.0174 | 0.0038 | 0.0436 |
| variance of time-varying output drift $\sigma_{\psi,n}^2$ | 0.0003 | 0.0001 | 0.0007 |

the null of constant volatility for trend inflation. It contrasts with Stock and Watson (2007) who find considerable variability in the variance of innovations to trend inflation. However, they do not allow for a persistent transitory component ($\zeta_t$ in our model) in the inflation gap such that trend inflation has to incorporate this component. In fact when we drop $\zeta_t$, the model selection procedure selects a specification with stochastic volatility for trend inflation (results not reported).
Morley et al. (2013) who find the real activity gap, as measured by the unemployment gap, to be an important driver of the inflation gap. Kim et al. (2014), however, find that the slope coefficient in a three regime model is only significant until the early 1970s and insignificant thereafter.

Our estimates shed light on a number of important episodes of U.S. monetary economic history. First, the Great Inflation of the late 1960s and the 1970s is reflected in a prolonged rise in trend inflation combined with an increase in the level and volatility of the temporary inflation component $\zeta_t$. In our model, the latter captures the variation in inflation that is not explained by the conventional forward-looking Phillips curve. From our estimates, this component mainly seems to capture the extent to which the oil price shocks of 1973-74 and 1979-80 drove up inflation without increasing inflation expectations or being reflected in the output gap. Second, the aggressive disinflation strategy pursued by Paul Volcker when he became chairman of the Federal Reserve in the early 1980s resulted in a steady but strong decline in trend inflation. Together with the sudden drop in the temporary inflation component, due to a drop in oil prices, this resulted in a sharp decline in realized inflation. The impact of the disinflation strategy on output depends on the credibility of monetary policy (see e.g. Ball, 1994). Imperfect credibility raises the output cost of reducing inflation. Our results point to a large output gap in the beginning of the disinflation period. In line with the small and stable slope of the Phillips curve, this is accompanied by only moderate negative deviations of realized inflation from its trend. This pattern changes during the second half of the disinflationary period, where the output gap decreases and realized inflation tracks trend inflation more closely. We take this as evidence that the credibility of the FED improved over time. This explanation is in line with the findings of Goodfriend and King (2005), who build a model with imperfect credibility, i.e. the FED acquires credibility over time as agents
change their beliefs about whether the new policy regime is permanent. According to the authors, the initial real effects of the Volcker disinflation were mainly due to its imperfect credibility. Third, our results also contribute to the discussion on the missing deflation puzzle during the Great Recession. Specifically, this paper casts doubt on the existence of such a puzzle as the link between inflation and real activity is weak over the full sample. The fact that actual inflation does not deviate substantially from trend inflation is therefore consistent with a relatively large output gap.

**Figure 6:** AR(1) inflation component (parsimonious model)

![Figure 6: AR(1) inflation component (parsimonious model)](image)

**Output**

Figure 8 plots the posterior results for the various components in output. Potential output, depicted in panel (a), is estimated as a smooth upward trend that tracks the low frequency movements in U.S. real GDP. The constant volatility of shocks to the level of potential output is found to be small with a posterior median of 0.13, while the drift in potential output, depicted in panel (b), exhibit substantial time variation. The downward trend in the drift term implies potential output growth to slow down from around 4% on an annual base in the early 1960s to about 1.6% at the end of the sample. This overall movement is highly consistent with the CBO’s estimates for potential output growth, although this series is more volatile. The first sizable drop in potential output growth is in the early to mid 1970s. This is a well-known feature of the data generally referred to as the great productivity slowdown. From the late 1970s to the early 2000s potential output growth varies around an annual rate of 3%. The second sizable drop occurred during the 2000s with the most recent estimates pointing to a pessimistic scenario where slow growth is the ‘new normal’. Our results support the argument in Perron and Wada (2009) that it is important to account for changes in potential output growth in trend-cycle decompositions of output.
Panel (c) of Figure 8 shows the estimated output gap together with the CBO gap. Both series evolve very similar and are able to identify the recession periods as dated by the NBER. A somewhat sizable difference in the level of the two series is observed during the 1980s. This is due to the fact that our model attributes most of the variation in real GDP during the early 1980s to cyclical shocks while the CBO assigns a larger fraction to potential output growth-related shocks as visualized by the sharp drop in the CBO potential growth series displayed in panel (b) in that period. The Great Moderation shows up in panel (d) as a considerable drop in the stochastic volatility of innovations to the output gap in the 1980s and a low volatility period that continued until 2007. During the Great Recession, volatility increases considerably but by 2014 it has almost returned to its pre-crisis level.

Unemployment

The NAIRU is shown in Figure 9 along with the CBO’s NAIRU estimate and the actual unemployment rate. We find that the NAIRU evolves very smoothly over time which implies that most of the variations in unemployment are assigned to cyclical (demand-related) factors. However, the recent decline in the unemployment rate may also partially be driven by people exiting the labor force, possibly discouraged jobless workers. Our NAIRU estimate is consistent with Laubach (2001) and Basistha and Startz (2008). Similar to the latter study, our multivariate model results in a relatively narrow 90% HPD interval for the NAIRU.

Regarding the relation between the output gap and the unemployment gap, we find substantial time variation in Okun’s law parameter as displayed in Figure 12, panel (b). The time variation captures both changes at business cycle frequency and long-run changes, which is similar to the findings of Knotek (2007). We find that the sensitivity of the unemployment rate to cyclical output is higher during recessions than during recoveries. This asymmetric pattern holds for most
of the postwar business cycles, except for the period after the 2001 recession, over which labor market sensitivity continues to increase. Before this turning point, the Okun coefficient fluctuates around a value of $-0.4$, suggesting that a positive output gap of 1% is associated with a negative deviation of the unemployment rate from the NAIRU of $-0.4\%$. Since the 2001 recession, the Okun coefficient has decreased to roughly $-0.7$ in the Great Recession. Recently the Okun coefficient has quickly returned back to the historical average. According to the reasoning in Daly and Hobijn (2010), the spike in the Okun coefficient around the year 2009 can be explained by a surge in labor productivity, accompanied by a reduction in employment and hours worked which led to a break in the pattern between unemployment and output as observed over the past 60 years.
There exist several explanations for why the correlation between output and unemployment may depend on the business cycle stance. Starting from a microeconomic model, Campbell and Fisher (2000) explain how asymmetries in firms’ adjustment costs can lead to asymmetric job creation and destruction rates at the macro level. Palley (1993) focuses on the aggregate labor market and explains the negative excess sensitivity of cyclical unemployment to cyclical output with sectoral shifts and changing behavior of female labor force participants. Silvapulle et al. (2004) offer an explanation based on over-pessimistic firm behavior. If bad news is believed more quickly than good news, firms tend to adjust the workforce relatively quick in recessions, but are reluctant to hire during recoveries. The authors argue that such behavior leads to asymmetry in the Okun coefficient typically found in U.S. data. Moreover, our findings are in line with the literature on insider-outsider models pioneered by Lindbeck and Snower (1988) and Blanchard and Summers (1986). After a cyclical rise in unemployment, the remaining workers (so-called insiders) may demand higher wages during the following recovery due to labor turnover costs. Instead of creating new jobs for the unemployed workers (so-called outsiders), economic recovery translates into higher insider wages. Such behavior gives rise to asymmetry in the Okun coefficient, leading to persistent cyclical unemployment.

Our estimates also contribute to the discussion on jobless recoveries in the United States. We do not find that the business cycle sensitivity of the Okun coefficient has changed in the 1990s. Rather, our results suggest that recoveries have always been ‘jobless’ in the sense that the unemployment rate adjusts faster during recessions than during recoveries. The argument that the unemployment rate has become less sensitive to output growth over time is not supported by our model. However, the notion of ‘jobless recoveries’ is typically related to job growth and thus makes a statement about employment dynamics. The estimated Okun’s law coefficient reflects sensitivity of the unemployment rate and is sensitive to changes in labor force participation. Slower than
average job growth could be counteracted by decreasing labor force participation, leaving the unemployment rate and therefore also the Okun’s law coefficient unchanged.

In sum, we find substantial time variation in various model’s parameters. There is a sizable reduction in the volatility of output gap shocks and inflation gap shocks. We also find a significant decline in potential output growth in the 1970s and even more pronounced in the 2000s. Moreover, there is time variation in the Okun’s law parameter with unemployment being more sensitive to the output gap in recessions than in expansions.

4 Model extensions and robustness checks

In this section we check the robustness of our results along several dimensions. First, we replace the forward-looking NKPC by a backward-looking version to see if our findings regarding the inflation dynamics and the stability of the slope of the Phillips curve depend on its forward-looking specification. Second, we estimate the model for different inflation measures. Third, we use the unemployment gap, instead of the output gap, as a measure of real activity in the Phillips curve. Fourth, we test whether the persistence of the temporary inflation component $\zeta_t$ has fallen over time. Fifth, we allow cyclical unemployment to be more sluggish than the output gap. Finally, upon imposing that there is no stochastic volatility in innovations to the level of potential output, we test whether these level shocks are relevant. We have also experimented with using higher AR orders for the output gap $y_c^t$ and the transitory inflation component $\zeta_t$. The deeper lags were however found to be irrelevant and did not change $y_c^t$ and $\zeta_t$. These results are not reported.

4.1 Backward-looking Phillips curve

As described in Section 3.1, the literature provides mixed support for the NKPC in empirical applications. In contrast to the theoretical foundations, some studies find an important backward-looking component in inflation dynamics, i.e. inflation depends on its own lagged values. We therefore check for the robustness of our findings by replacing equation (14) by the following backward-looking Phillips curve specification

$$\pi_t = \sum_{p=1}^{4} b_p \pi_{t-p} + \beta_1 y_c^t + \varepsilon_1^\pi,$$

where the sum of the coefficients on lagged inflation is assumed to be one.\(^8\) The model is identical to the baseline model except for the absence of a stochastic trend in inflation and the temporary inflation component $\zeta$. This specification matches standard backward-looking Phillips curve models in the literature, except the use of the output gap instead of the unemployment gap as a real-activity measure.\(^9\)

Figure 10 plots the posterior distribution of the standard deviation $\sigma_{\eta,\pi}$ together with the time variation in the Phillips curve slope when setting $\delta_\pi$ (together with all the other binary

\(^8\)Results are nearly identical when the unit-root assumption is relaxed. The sum of the coefficients in the unrestricted case is close to one.

\(^9\)Among many others see e.g. Rudd and Whelan (2005) and Ball and Mazumder (2011).
indicators) equal to 1. In line with the literature, we find that the sensitivity of inflation to real economic activity has decreased over time, from around 0.15 in the 1970s to below 0.05 in the 2000s. Different to the existing literature we proceed and test whether this time variation is relevant or not. Although the posterior distribution of $\sigma_{\eta,\pi}$ plotted in panel (a) shows some signs of bimodality, the large probability mass at zero suggests that the time variation in the Phillips curve is probably not a relevant feature of the model. This is confirmed by the posterior inclusion probabilities reported in Table 6. Similar to the results for the New-Keynesian specification, the backward-looking Phillips curve is clearly found to be stable over time. Again, this finding is robust over different priors as in all settings the posterior probability of time variation is below 1%. Figure 11 plots the posterior distributions of the time-invariant Phillips curve slope parameter in the parsimonious model for both the forward and backward-looking specification. The probability mass of the coefficient in the backward-looking model is strictly positive and has a slightly higher median. This is in line with the literature that usually finds a steeper and significant Phillips curve slope in backward compared to purely forward-looking models (see Lee and Nelson, 2007, for a potential explanation).

4.2 Alternative inflation measures

In the empirical literature on the Phillips curve, no single preferred inflation measure has emerged. This paper focuses on the core PCE inflation series, since it eliminates large outliers associated with energy price fluctuations as pointed out by Stock and Watson (2010). However, other inflation measures have been used repeatedly in the literature such as core and headline CPI inflation or the implicit GDP deflator. We check the robustness of our results by estimating the model for different inflation measures but leaving all prior distributions unchanged. Columns 2-4 in Table 7
Table 6: Stability of Phillips curve for different models

<table>
<thead>
<tr>
<th>Prior</th>
<th>Posterior probabilities of $\delta_\pi = 1$</th>
<th>New-Keynesian</th>
<th>Backward-looking</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_0$</td>
<td>$A_0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.0078</td>
<td>0.0062</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>0.0031</td>
<td>0.0016</td>
</tr>
<tr>
<td>0.5</td>
<td>10</td>
<td>0.0009</td>
<td>0.0004</td>
</tr>
<tr>
<td>0.5</td>
<td>100</td>
<td>0.0002</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

Figure 11: Posterior distribution of $\beta_\pi^0$: backward versus forward-looking Phillips curve (parsimonious model)

report the estimated posterior distribution of the constant Phillips curve slope. The median slope estimates range between 0.067 for the GDP deflator and 0.110 for headline CPI and headline PCE inflation. In all five specifications the 95% credible interval covers positive and negative values. Thus, the slope of the Phillips curve is found to be more or less flat regardless of the inflation measure used. Columns 5-8 report the posterior inclusion probabilities of a time-varying Phillips curve and the time-varying volatilities in trend and temporary inflation. Again, findings are robust to the different inflation measures. The posterior probability of time variation in the Phillips curve is below 1% in all cases. Results differ more for the stochastic volatility component in trend inflation. However, probabilities remain below 5% except for the baseline measure and thus no evidence of time-varying shocks to trend inflation is found. Finally, stochastic volatility is always included in the temporary inflation component. We conclude that our findings are robust to alternative measures of price inflation.\textsuperscript{10}

\textsuperscript{10}Estimates for the output and unemployment components do not change notably, but are available on request.
Table 7: Robustness of parameter estimates to different inflation measures

<table>
<thead>
<tr>
<th></th>
<th>Posterior parameter distribution (parsimonious model)</th>
<th>Posterior inclusion probability (sampling the binary indicators)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope coefficient $\beta_0^\pi$</td>
<td>Phillips curve $\delta_\pi$</td>
</tr>
<tr>
<td></td>
<td>median</td>
<td>2.5%</td>
</tr>
<tr>
<td>CPI</td>
<td>0.104</td>
<td>-0.129</td>
</tr>
<tr>
<td>CPI excl. F&amp;E</td>
<td>0.103</td>
<td>-0.067</td>
</tr>
<tr>
<td>PCE</td>
<td>0.104</td>
<td>-0.095</td>
</tr>
<tr>
<td>PCE excl. F&amp;E</td>
<td>0.077</td>
<td>-0.094</td>
</tr>
<tr>
<td>GDP deflator</td>
<td>0.073</td>
<td>-0.115</td>
</tr>
</tbody>
</table>

Note: Priors are set to $p_0 = 0.5$ and $A_0 = 1$.

4.3 Unemployment gap instead of output gap

The baseline model finds a constant Phillips curve slope and a time-varying Okun’s law parameter. Consequently, when we replace the output gap with the unemployment gap in the Phillips curve, the impact of unemployment on inflation varies over time. However, when we estimate the model with the unemployment gap in the Phillips curve, the model selection procedure rejects a time-varying slope parameter. We believe that this is due to the fact that the real activity measure, independent of whether we proxy it by the output gap or the unemployment gap, has little impact on inflation. Thus, all conclusions drawn remain unchanged when replacing the output gap with the unemployment gap.

4.4 Time-varying persistence in the temporary inflation component

Cogley et al. (2010) and Kim et al. (2014) show that the persistence of transitory inflation has fallen over time. To allow for this possibility we have extended the model by specifying the AR parameter $\varrho$ of the temporary inflation component $\zeta_t$ as a random walk process

$$\varrho_{t+1} = \varrho_t + \eta_t^\varrho, \quad \eta_t^\varrho \sim i.i.d. N(0, \sigma_{\eta,\varrho}^2).$$

(33)

To enable model selection on $\varrho_t$, we use the non-centered parameterization

$$\varrho_{t+1} = \varrho_0 + \delta_\varrho \eta_t^\varrho \tilde{\varrho}_{t+1},$$

(34)

with

$$\tilde{\varrho}_{t+1} = \tilde{\varrho}_t + \tilde{\eta}_t^\varrho, \quad \tilde{\varrho}_0 = 0, \quad \tilde{\eta}_t^\varrho \sim i.i.d. N(0, 1),$$

(35)

and where $\varrho_0$ is the initial value of the level of $\varrho_t$ and $\delta_\varrho$ is a binary indicator that is either 0 or 1. We have imposed stationarity on $\zeta_t$ by setting a reflecting barrier on $\varrho_t$ at -1 and 1, i.e. the previous Gibbs draw for the entire series $\varrho_t$ is used when the new draw includes any values for $\varrho_t$ that are not in the (-1,1) interval.
In line with the priors set in Table 1, we use $\mathcal{N}(0.70, 0.05^2)$ and $\mathcal{N}(0, 1)$ as prior distributions for $\theta_0$ and $\sigma_{\eta,\varrho}$ respectively, while the binary indicator $\delta_{\varrho}$ again has a prior probability $p_0 = 0.5$ of being one. Results for $\varrho_t$ and $\sigma_{\eta,\varrho}$ when setting the binary indicator $\delta_{\varrho} = 1$ are reported in Figure 12. The plot of $\varrho_t$ shows a slight decrease in persistence from the early 1980s onwards. There is however huge uncertainty around these estimates, which goes hand in hand with the much lower variance of $\zeta_t$ in the second half of the sample (see Figure 6, panel b). The posterior distribution of $\sigma_{\eta,\varrho}$ shows very little signs of bimodality and has considerable probability mass around zero. When sampling the binary indicator $\delta_{\varrho}$, the posterior probability that it is one is only 5.19% (see Table 8, first row). Hence, $\varrho_t$ is clearly selected to be constant by the stochastic model specification search.

**Figure 12**: Time-varying persistence $\varrho_t$ in the temporary inflation component $\zeta_t$ (all binary indicators set to 1)

![Figure 12](image)

**Table 8**: Posterior inclusion probabilities for the binary indicators over different model extensions

<table>
<thead>
<tr>
<th>Model extension</th>
<th>Time-varying parameters/components</th>
<th>Stochastic volatilities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Phillips curve</td>
<td>Okun’s law</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>$\varrho_{t+1} = \varrho_t + \eta_{t}^\varrho$</td>
<td>0.0027</td>
<td>1.0000</td>
</tr>
<tr>
<td>$\xi_t = \zeta_t + \psi_{t}^\xi$</td>
<td>0.0011</td>
<td>1.0000</td>
</tr>
<tr>
<td>$\sigma_{t}^\varrho = \varrho_{t-1} + \rho_{t}^\sigma \eta_{t}$</td>
<td>0.0022</td>
<td>0.2917</td>
</tr>
<tr>
<td>$\gamma_{t+1} = \gamma_{t} + \kappa_{t} \sigma_{t}$</td>
<td>0.0022</td>
<td>1.0000</td>
</tr>
<tr>
<td>$+ \lambda_{y} \sigma_{v,y} \gamma_{t+1} + \lambda_{c} \sigma_{v,c} \tilde{K}_{t+1}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.5 Sluggishness in unemployment

As cyclical unemployment may be more persistent than the output gap, the time variation in Okun’s law presented in Section 3.4 could be the result of a sluggish reaction of cyclical unemployment to the output gap. To allow for this potential sluggishness in unemployment, we first extend equation (15) to

\[ u_t = u_t^\tau + \beta_t u_t^\gamma + \xi_t + \epsilon_t^u, \]  

(36)

with \( \xi_t = \varsigma \xi_{t-1} + \psi_t \) a persistent temporary unemployment component that is not related to the business cycle, i.e. exists apart from Okun’s law. The non-centered parameterization for \( \xi_t \) is given by

\[ \xi_{t+1} = \xi_0 + \delta_{\xi} \sigma_{\eta,\xi} \tilde{\xi}_{t+1}, \]  

(37)

with \( \tilde{\xi}_{t+1} = \tilde{\xi}_t + \tilde{\eta}_t, \quad \tilde{\xi}_0 = 0, \quad \tilde{\eta}_t \sim i.i.d. \mathcal{N}(0,1), \]  

(38)

and where \( \xi_0 \) is the initial value of the level of \( \xi_t \) and \( \delta_{\xi} \) is a binary indicator that is either 0 or 1. We use \( \mathcal{N}(0,1) \) as a prior distribution for both \( \xi_0 \) and \( \sigma_{\eta,\xi} \) and \( \mathcal{N}(0.5,0.25^2) \) as a prior distribution for the AR parameter \( \varsigma \). The binary indicator \( \delta_{\xi} \) has a prior probability \( p_0 = 0.5 \) of being one. The posterior distributions of \( \xi_t \) and \( \sigma_{\eta,\xi} \) when setting the binary indicator \( \delta_{\xi} \) (and all other indicators) to one, plotted in Figure 13, show that the temporary unemployment component is clearly not a relevant model extension. When sampling the binary indicator \( \delta_{\xi} \), the posterior probability that it is one is only 2.9% (see Table 8, second row).

Figure 13: Transitory unemployment component \( \xi_t \) (all binary indicators set to 1)

A more sluggish response of unemployment may however also mean that there are additional dynamics in the unemployment gap, i.e. transitory unemployment is related to the output gap (through Okun’s law) but is relatively more persistent. We therefore further allow unemployment
to be more sluggish compared to the output gap by extending the unemployment equation (15) to

\[ u_t = u^e_t + u^c_t + \varepsilon^u_t, \]

(39)

with \( u^c_t = \varsigma u^c_{t-1} + \beta^u u^c_{t-1} + \beta^u y^c_{t} \). By allowing the unemployment gap \( u^c_t \) to depend on its own lag \( u^c_{t-1} \) we allow for additional persistence compared to the output gap \( y^c_{t} \). We use \( \mathcal{N}(0.5, 0.25^2) \) as a prior distribution for the AR parameter \( \varsigma \). Figure 14 plots the time-varying Okun’s law parameter \( \beta^u \) together with the posterior distribution of the standard deviation \( \sigma_{n,u} \) of its innovations obtained setting all binary indicators to 1. Interestingly, \( \beta^u \) is now much more stable, i.e. the variation at the business cycle frequency that was clearly present in the baseline model has now completely vanished. This stability is confirmed by the posterior distribution of \( \sigma_{n,u} \) which shows some bimodality but also has considerable probability mass at zero. The third row in Table 8 reports the posterior inclusion probabilities when sampling the binary indicators. With a probability that \( \delta_u = 1 \) of 29.17%, there is no longer convincing evidence that the relationship between cyclical unemployment and output is unstable over time. Figure 15 shows the sluggishness in unemployment and its implication for the unemployment gap measure \( u^c_t \) in the parsimonious model, which compared to the benchmark case now also sets the binary indicator \( \delta_u = 0 \). The posterior distribution of the AR parameter \( \varsigma \) plotted in panel (a) shows that there is moderate sluggishness in the reaction of unemployment to the output gap. The mean of the posterior distribution of \( \varsigma \) is 0.53 with 5 and 95 percentiles being 0.47 and 0.59 respectively. The impact of this sluggishness on the unemployment gap can be seen in panel (b) where we plot the full unemployment gap \( u^c_t = \varsigma u^c_{t-1} + \beta^u y^c_{t} \) together with the contemporaneous reaction \( \beta^u y^c_{t} \) of unemployment to the output gap. When comparing these two gaps, the slower recovery of unemployment in the aftermath of the NBER recession periods becomes particularly apparent. Figure 15, panel (b) further shows that the evolution in the sluggish unemployment gap \( u^c_t \) from the extended model - where \( \beta^u \) is constant - is highly similar to that of the unemployment gap \( \beta^u y^c_{t} \) from the baseline model reported in Section 3.4 - where there is no sluggishness (\( \varsigma = 0 \)) but \( \beta^u \) varies over time. The time variation in \( \beta^u \) in the baseline model thus seems to capture the sluggish reaction of unemployment to the output gap. This is in line with the potential explanations (i.e. over-pessimistic firms, insider-outsider effects) for the cyclical behavior of the Okun’s law parameter put forward in Section 3.4.

### 4.6 Trend stationarity of potential output

Perron and Wada (2009) argue that U.S. real GDP over the period 1947Q1 to 1998Q2 is adequately modeled using a non-stochastic trend with a (smooth) break in its slope around 1973. The results of our baseline model in Section 3 confirm this reduction in trend growth in the 1970s - and add a second one in the 2000s - but do not confirm that the real output trend \( y^\tau_t \) is non-stochastic, i.e. the variance of shocks to the level of the trend is estimated to be non-zero. However, we did not use model selection to test whether this variance is zero since we allow and test for stochastic volatility \( h^y_t \) in this component. In the non-centered parsimonious specification, testing whether the variance of the disturbances to \( y^\tau_t \) is zero would imply a product of binary indicators and hence non-identification, i.e. the binary indicator \( \theta_y \) on the stochastic volatility component would then
Figure 14: Okun’s law parameter in the extended model \( u_t^c = \varsigma u_{t-1}^c + \beta_u^c y_t^c \) (all binary indicators set to 1)

![Graph showing the Okun's law parameter in the extended model](image)

Figure 15: Sluggish unemployment gap in the extended model \( u_t^c = \varsigma u_{t-1}^c + \beta_u^c y_t^c \) (parsimonious model)

![Graph showing the sluggish unemployment gap](image)

Note: The parsimonious model sets \( \delta_\pi = \delta_u = \theta_y = \theta_\pi = \theta_u = 0 \) and \( \theta_\varsigma = \theta_\zeta = \lambda_\varsigma = 1 \). Panel (b) plots three unemployment gap measures: (i) \( u_t^c = \varsigma u_{t-1}^c + \beta_u^c y_t^c \) and \( \beta_u^c y_t^c \) are both calculated from the extended model’s results allowing for a sluggish reaction of unemployment to the output gap but the Okun’s law parameter \( \beta_u^c \) being constant over time with \( u_t^c \) the full (sluggish) unemployment gap and \( \beta_u^c y_t^c \) the contemporaneous reaction only; (ii) \( \beta_u^c y_t^c \) is calculated from the baseline model results presented in Section 3.4 allowing for a time-varying Okun’s law parameter but no sluggishness (\( \varsigma = 0 \)).

Only be identified when the new binary indicator on the variance of the innovations to \( y_t^c \) is one. Note that if the stochastic model specification search selects stochastic volatility, this automatically
implies that the variance of the innovations to $y_t^\tau$ is, at least over some period, different from zero. However, the stochastic volatility component $h_t^y$ was not selected by the procedure. Hence, we can drop it from the model and proceed to test whether the variance of the disturbances to the trend component $y_t^\tau$ is zero. To this end, we rewrite equations (2)-(3) to

$$y_{t+1}^\tau = y_0^\tau + \kappa_0 t + \lambda_y \sigma_{\psi,y} \tilde{y}_{t+1}^\tau + \lambda_\kappa \sigma_{\psi,\kappa} \tilde{K}_{t+1},$$

with

$$\tilde{y}_{t+1}^\tau = \tilde{y}_t^\tau + \tilde{\psi}_t^\psi, \quad \tilde{y}_0^\tau = 0, \quad \tilde{\psi}_t^\psi \sim i.i.d. \mathcal{N}(0,1),$$

$$\tilde{\kappa}_{t+1} = \tilde{\kappa}_t + \tilde{\psi}_t^\kappa, \quad \tilde{\kappa}_0 = 0, \quad \tilde{\psi}_t^\kappa \sim i.i.d. \mathcal{N}(0,1),$$

$$\tilde{K}_{t+1} = \tilde{K}_t + \tilde{\kappa}_t, \quad \tilde{K}_0 = 0,$$

and where $\lambda_y$ and $\lambda_\kappa$ are binary indicators that are either 0 or 1. In this non-centered parameterization, $\tilde{y}_t^\tau$ captures the time-varying level part of potential output while $\tilde{K}_t$ represents the time-varying part of the drift component. In addition to the priors set in Table 1, we use $\mathcal{N}(y_1,10^2)$ and $\mathcal{N}(0,1)$ as prior distributions for $y_0^\tau$ and $\sigma_{\psi,y}$ respectively. The binary indicator $\lambda_y$ has a prior probability $p_0 = 0.5$ of being one. The posterior distributions of $\sigma_{\psi,y}$ and $\sigma_{\psi,\kappa}$ (setting all binary indicators to 1) reported in Figure 16 both show clear bimodality. This further supports our finding that shocks to both the level and the drift of potential output are relevant. This is confirmed by the posterior inclusion probability of 100% for $\lambda_y$ and $\lambda_\kappa$ (see last row of Table 8) when sampling the binary indicators.

**Figure 16:** Posterior distribution of the standard deviations of the innovations to the level and the drift of potential output (all binary indicators set to 1)
5 Conclusion

We have investigated the degree of time variation in the parameters of a multivariate unobserved components model designed for the U.S. economy over the period 1959 to 2014. The empirical model decomposes real GDP, inflation and the unemployment rate into a common stochastic cyclical factor and their respective stochastic trends and idiosyncratic components. Key parameters such as the growth rate of potential output, the slope of the Phillips curve, the Okun’s law coefficient as well as all variance parameters are allowed to vary over time. Importantly, while allowing for time variation the priors in the Bayesian estimation strategy are set to nest the case that the parameters are actually constant. In a first estimation step, a stochastic model selection procedure is employed to test which parameters vary over time. We find that potential output growth, the Okun’s law coefficient, the variance of innovations to the output gap and to a persistent inflation gap component vary over time while the slope of the Phillips curve and the variances of innovations to all trend components are found to be constant. Our estimation result show a clear decrease in potential output growth, which is characterized by substantial drops in the 1970s and the 2000s. The Okun’s law coefficient is found to be lower in recessions than in expansions, i.e. unemployment is more sensitive to the output gap in a downturn and reacts less sensitively in a recovery. Once a sluggish reaction of unemployment to output is added to the model, Okun’s law is found to be stable again. With regard to the dynamics of inflation, we find that the inflation gap and idiosyncratic shocks are the major determinants of inflation changes. The inflation gap is not very sensitive to the output gap but mainly driven by a persistent AR(1) component. The latter component exhibits stochastic volatility and principally captures the large and persistent swings in inflation during the inflationary period in the 1970s and 1980s.
References


Appendix A  
Gibbs sampling algorithm

In this appendix we provide details on the Gibbs sampling algorithm used in Section 2.3 to jointly sample the binary indicators $M$, the hyperparameters $\phi$, the trend and temporary components $\alpha$, the time-varying parameters $\beta$, the mixture indicators $\iota$ and the stochastic volatilities $h$. The structure of our Gibbs sampling approach is based on Frühwirth-Schnatter and Wagner (2010).

Block 1: Sampling the binary indicators $M$ and the parameters $\phi$

For notational convenience, let us define a general regression model

$$w = z^M b^M + e, \quad e \sim N(0, \Sigma),$$

(A-1)

with $w$ a vector including observations on a dependent variable $w_t$ and $z$ an unrestricted predictor matrix with rows $z_t$ that contain the state processes from the vectors $\alpha_t$, $\beta_t$ and $h_t$ that are relevant for explaining $w_t$. The corresponding unrestricted parameter vector with the relevant elements from $\phi$ is denoted $b$. $z^M$ and $b^M$ are then the restricted predictor matrix and restricted parameter vector that exclude those elements in $z$ and $b$ for which the corresponding indicator in $M$ is 0. Furthermore, $\Sigma$ is a diagonal matrix with elements $\sigma^2_{e,t}$ that may vary over time to allow for heteroskedasticity of a known form.

A naive implementation of the Gibbs sampler would be to sample $M$ from $f(M | \alpha, \beta, h, \phi, w)$ and $\phi$ from $f(\phi | \alpha, \beta, h, M, w)$. However, this approach does not result in an irreducible Markov chain as whenever an indicator in $M$ equals zero, the corresponding coefficient in $\phi$ is also zero which implies that the chain has absorbing states. Therefore, as in Frühwirth-Schnatter and Wagner (2010) we marginalize over the parameters $\phi$ when sampling $M$ and next draw the parameters $\phi$ conditional on the indicators $M$. The posterior distribution $f(M | \alpha, \beta, h, w)$ can be obtained using Bayes’ Theorem as

$$f(M | \alpha, \beta, h, w) \propto f(w | M, \alpha, \beta, h) p(M),$$

(A-2)

with $p(M)$ being the prior probability of $M$ and $f(w | M, \alpha, \beta, h)$ being the marginal likelihood of the regression model (A-1) where the effect of the parameters $b^M$ and $\sigma^2_e$ has been integrated out. The closed form solution of the marginal likelihood depends on whether the error term $e_t$ is homoskedastic or heteroskedastic. More specifically:

- In the homoskedastic case $\Sigma = \sigma^2_e I_T$, under the normal-inverse gamma conjugate prior
  
  $$b^M \sim N(a^M_0, A^M_0 \sigma^2_e), \quad \sigma^2_e \sim IG(c_0, C_0),$$

  (A-3)

  the closed form solution for $f(w | M, \alpha, \beta, h)$ is

  $$f(w | M, \alpha, \beta, h) \propto \left| A^M_0 \right|^{0.5} \Gamma(c_T) C_0^{c_0} \left| A^M_T \right|^{0.5} \Gamma(c_0) \left( C^M_T \right)^{c_T},$$

  (A-4)
In this block we sample the binary indicators $\delta$.

Block 1(a): Sampling the binary indicators $\delta$ is therefore split up in the following subblocks:

1. Regression format of (A-1) as $\sigma$ is no binary indicator in equation (1), $\sigma_f M$ the elements in the binary indicators $\eta = (\eta, \pi)$, $\eta, \pi$ are sampled simultaneously, we use a single-move sampler in which each of them is sampled from $IG(c_T, C_T)$ with $c_T$ as in equation (A-7) and $C_T = C_0 + 0.5(\epsilon^{xy}_t \epsilon^y_t)$ with $\epsilon^y$ calculated from $\epsilon^y_t = y_t - y^c_t - y^c_t$.

Following George and McCulloch (1993), instead of using a multi-move sampler in which all the elements in $M$ are sampled simultaneously, we use a single-move sampler in which each of the binary indicators $\delta_j$ (for $j = \pi, u$), $\theta_k$ (for $k = y, \pi, u, c, \zeta$) and $\lambda_n$ in $M$ is sampled from $f(\delta_j|\delta_j, \theta, \lambda_n, \alpha, \beta, h, x)$, $f(\theta_k|\delta, \theta, \lambda_n, \alpha, \beta, h, x)$ and $f(\lambda_n|\delta, \theta, \alpha, \beta, h, x)$ respectively. Block 1 is therefore split up in the following subblocks:

**Block 1(a): Sampling the binary indicators $\delta$ and the parameters $\beta, \sigma_\eta$ and $\sigma^2_\epsilon$**

In this block we sample the binary indicators $\delta = (\delta_\pi, \delta_u)$ and the parameters $\beta = (\beta_0^\pi, \beta_0^u)$, $\sigma_\eta = (\sigma_{\eta, \pi}, \sigma_{\eta, u})$ and $\sigma^2_\epsilon = (\sigma^2_{\epsilon, \pi}, \sigma^2_{\epsilon, y}, \sigma^2_{\epsilon, u})$ conditional on the states $\alpha, \beta$ and $h$.

As there is no binary indicator in equation (1), $\sigma^2_{\epsilon, y}$ can be sampled directly from $IG(c_T, C_T)$ with $c_T$ as in equation (A-7) and $C_T = C_0 + 0.5(\epsilon^{xy}_t \epsilon^y_t)$ with $\epsilon^y$ calculated from $\epsilon^y_t = y_t - y^c_t - y^c_t$.

Next, using equation (29), equations (14) and (15) can be rewritten in the general linear regression format of (A-1) as:

$$w_t \pi_t - \pi^c_t - \zeta_t = \left[ \begin{array}{c} \tilde{y}^c_t \\ \tilde{y}^c_t \end{array} \right] \left[ \begin{array}{c} \beta_0^\pi \\ \sigma_{\eta, \pi} \end{array} \right] + \tilde{\epsilon}_t,$$

$$u_t - u^c_t = \left[ \begin{array}{c} \tilde{y}^c_t \\ \tilde{y}^c_t \end{array} \right] \left[ \begin{array}{c} \beta_0^u \\ \sigma_{\eta, u} \end{array} \right] + \tilde{u}_t,$$
where in both the restricted vector $z^M_i$ and the restricted parameter vector $b^M$ the second term is excluded when $\delta_j = 0$ (for $j = \pi, u$). Note that, next to the parameters in $b^M$ and $\sigma^2_j$, each of the specifications (A-12) and (A-13) depends only on the data $w_t$, on some of the states in $\alpha_t$ and $\beta_t$ and on $\delta_j$. As such, we can simplify the specification of the posterior from $f(\delta_j | \theta, \lambda, \alpha, \beta, h, x)$ to $f(\delta_j | \alpha, \beta, w)$ for which we have $f(\delta_j | \alpha, \beta, w) \propto f(w | \delta_j, \alpha, \beta) p(\delta_j)$. As the error terms $\epsilon^T_t$ in the inflation equation and $\epsilon^u_t$ in the employment equation are homoskedastic, we have $\Sigma = \sigma^2 \mathbf{I}_T$ in the general notation of equation (A-1) such that the marginal likelihood $f(w | \delta_j, \alpha, \beta)$ can be calculated as in equation (A-4). The binary indicator $\delta_j$ can then be sampled from the Bernoulli distribution with probability

$$p(\delta_j = 1 | \alpha, \beta, w) = \frac{f(\delta_j = 1 | \alpha, \beta, w)}{f(\delta_j = 0 | \alpha, \beta, w) + f(\delta_j = 1 | \alpha, \beta, w)},$$

while $\sigma^2_j$ can be sampled from $\text{IG}(c_T, C_T^M)$ and, conditionally on $\sigma^2_j$, $b^M$ from $\mathcal{N}(a^M_t, A^M_t \sigma^2_j)$, for $j = \pi, u$ and with $a^M_t, A^M_t, c_T$ and $C_T^M$ as defined in equations (A-5)-(A-8). Note that $b^M = (b^*_0, \sigma_{\eta,j})'$ when $\delta_j = 1$ and $b^M = b^*_0$ when $\delta_j = 0$. In the former case $\sigma_{\eta,j}$ is sampled from the posterior while in the latter case we set $\sigma_{\eta,j} = 0$.

Block 1(b): Sampling the binary indicators $\theta$ and the parameters $h_0$ and $\sigma_{\gamma}$

In this block we sample the binary indicators $\theta = (\theta_y, \theta_\pi, \theta_u, \theta_c, \theta_\zeta)$ and the parameters $h_0 = (h^*_0, h^*_\pi, h^*_u, h^*_c, h^*_\zeta)$ and $\sigma_{\gamma} = (\sigma_{\gamma,y}, \sigma_{\gamma,\pi}, \sigma_{\gamma,u}, \sigma_{\gamma,c}, \sigma_{\gamma,\zeta})$ conditional on the states $\alpha$, $\beta$ and $h$. Using equation (30), equation (20) can be rewritten in the general linear regression format of (A-1) as

$$g^k_t = \frac{w_t}{y^k_t - (m^k_t - 1.2704)} = 2 \left[ \frac{z^M_t}{1 - \theta^k h^k_t} \right] \left[ \frac{b^M}{\sigma_{\gamma,k}} \right] + \frac{\epsilon^k_t}{\epsilon^k_t},$$

for $k = y, \pi, u, c, \zeta$, with $\epsilon^k_t = \epsilon^k_t - (m^k_t - 1.2704)$ is $\epsilon^k_t$ recentered around zero and where using equations (2), (9), (16), (4) and (11), $g^k_t = \ln \left( (\exp(h^k_t \psi^k_t)^2 + .001) \right)$ can be calculated as

$$g^k_t = \ln \left( (y^k_t - y^k_{t-1} - \kappa_k)^2 + .001 \right),$$

$$g^k_t = \ln \left( (\pi^k_t - \pi^k_{t-1})^2 + .001 \right),$$

$$g^k_t = \ln \left( (u^k_t - u^k_{t-1})^2 + .001 \right),$$

$$g^k_t = \ln \left( (g^k_t - \rho_1 y^k_{t-1} - \rho_2 y^k_{t-2})^2 + .001 \right),$$

$$g^k_t = \ln \left( (\zeta_t - \psi^k_{t-1})^2 + .001 \right).$$

As specification (A-15) depends only on the data $w_t$, on the stochastic volatility term $h^k_t$ and on $\theta_k$, we can simplify the specification of the posterior from $f(\theta_k | \delta, \theta_k, \lambda, \alpha, \beta, h, x)$ to $f(\theta_k | h, w)$. Using Bayes’ Theorem, we have $f(\theta_k | h, w) \propto f(w | \theta_k, h) p(\theta_k)$. Given the mixture distribution

44
of \( \epsilon_t^k \) defined in equation (22), the error term \( \tilde{z}_t^k \) in equation (A-15) has a heteroskedastic variance \( \nu_t^2 \) such that \( \Sigma = \text{diag} \left( \nu_1^2, \ldots, \nu_T^2 \right) \) in the general notation of equation (A-1). In this case, the marginal likelihood \( f(w | \theta_k, h) \) can be calculated as in equation (A-9). The binary indicator \( \theta_k \) can then be sampled from the Bernoulli distribution with probability \( p(\theta_k = 1 | h, w) \) calculated from an equation similar to (A-14). Next, \( b^M \) can be sampled from \( \mathcal{N}(a_T^M, A_T^M) \) for \( k = y, \pi, u, c, \zeta \) and with \( a_T^M \) and \( A_T^M \) as defined in equations (A-10) and (A-11). Note that \( b^M = (h_0^k, \sigma_{\gamma,k})^T \) when \( \theta_k = 1 \) and \( b^M = h_0^k \) when \( \theta_k = 0 \). In the latter case, we set \( \sigma_{\gamma,k} = 0 \).

**Block 1(c): Sampling the binary indicator \( \lambda_\kappa \) and the parameters \( \kappa_0 \) and \( \sigma_{\psi,\kappa} \)**

In this block, we sample the binary indicator \( \lambda_\kappa \) and the parameters \( \kappa_0 \) and \( \sigma_{\psi,\kappa} \) conditional on the states \( \alpha, \beta \) and \( h \). Using equation (31), equation (2) can be rewritten in the general linear regression format of (A-1) as

\[
\begin{bmatrix}
  \frac{w_t}{\gamma_t} \\
  y_t - \gamma_{t-1}
\end{bmatrix} =
\begin{bmatrix}
  1 & \lambda_\kappa \kappa_t \\
  \sigma_{\psi,\kappa}
\end{bmatrix}
\begin{bmatrix}
  \kappa_0 \\
  \exp \{ h_t^T \psi_T^k \}
\end{bmatrix} + \epsilon_t,
\]

(A-21)

with \( \Sigma = \text{diag} \left( \exp \{ h_t^T \}^2, \ldots, \exp \{ h_T^T \}^2 \right) \). The indicator \( \lambda_\kappa \) can then be sampled from the posterior distribution \( f(\lambda_\kappa | \alpha, w) \propto f(w | \lambda_\kappa, \alpha) p(\lambda_\kappa) \) with the marginal likelihood \( f(w | \lambda_\kappa, \alpha) \) calculated from equation (A-9). Next, \( b^M \) can be sampled from \( \mathcal{N}(a_T^M, A_T^M) \) with \( a_T^M \) and \( A_T^M \) as defined in equations (A-10) and (A-11). Note that \( b^M = (\kappa_0, \sigma_{\psi,\kappa})^T \) when \( \lambda_\kappa = 1 \) and \( b^M = \kappa_0 \) when \( \lambda_\kappa = 0 \). In the latter case, we set \( \sigma_{\psi,\kappa} = 0 \).

**Block 1(d): Sampling the parameters \( \rho \) and \( \varrho \)**

For sampling \( \rho = (\rho_1, \rho_2) \) conditional on the states \( \alpha, \beta \) and \( h \), equation (4) can be written in the general notation of equation (A-1) as: \( w_t = y_t^\gamma, z_t = (y_{t-1}^\gamma, y_{t-2}^\gamma), b = (\rho_1, \rho_2)^T \) and \( \epsilon_t = \exp \{ h_t^\gamma \} \psi_T^\gamma \), such that \( \Sigma = \text{diag} \left( \exp \{ h_t^\gamma \}^2, \ldots, \exp \{ h_T^\gamma \}^2 \right) \). Under the normal prior distribution \( \mathcal{N}(a_T, A_T) \) with \( a_T \) and \( A_T \) as in equations (A-10) and (A-11).

Likewise, for filtering \( \varrho \) conditional on the states \( \alpha, \beta \) and \( h \), equation (11) can be written in the general notation of equation (A-1) as: \( w_t = \zeta_t, z_t = \zeta_{t-1}, b = \varrho \) and \( \epsilon_t = \exp \{ h_t^{\zeta} \} \psi_T^{\zeta} \), such that \( \Sigma = \text{diag} \left( \exp \{ h_t^{\zeta} \}^2, \ldots, \exp \{ h_T^{\zeta} \}^2 \right) \). Under the normal prior distribution \( \mathcal{N}(a_T, A_T) \), \( \varrho \) can again be sampled from the posterior \( \mathcal{N}(a_T, A_T) \) with \( a_T \) and \( A_T \) as in equations (A-10) and (A-11).

**Block 2: Sampling the state vectors \( \alpha, \beta \) and \( h \) and mixture indicators \( \iota \)**

In this block, we use a forward-filtering and backward-sampling approach of Carter and Kohn (1994) and De Jong and Shephard (1995) to sample the states \( \alpha, \beta \) and \( h \) based on a general state
space model of the form

\[
 w_t = Z_t^M s_t^M + e_t, \quad e_t \sim iid \mathcal{N}(0, H_t), \quad (A-22)
\]

\[
 s_{t+1} = R_0 + R_1 s_t + K_t v_t, \quad v_t \sim iid \mathcal{N}(0, Q_t), \quad s_1 \sim iid \mathcal{N}(a_1, A_1), \quad (A-23)
\]

where \( w_t \) is now a vector of observations and \( s_t \) an unobserved state vector. The matrices \( Z_t, R_0, R_1, K_t, H_t, Q_t \) and the expected value \( a_1 \) and variance \( A_1 \) of the initial state vector \( s_1 \) are assumed to be known (conditioned upon). The vector \( s_t^M \) and the matrix \( Z_t^M \) are again restricted versions of \( s_t \) and \( Z_t \) with the elements excluded depending on the model indicators \( M \). The error terms \( e_t \) and \( v_t \) are assumed to be serially uncorrelated and independent of each other at all points in time. As equations (A-22)-(A-23) constitute a linear Gaussian state space model, the unknown state variables in \( s_t \) can be filtered using the standard Kalman filter. Sampling \( s = [s_1, \ldots, s_T] \) from its conditional distribution can then be done using the multimove simulation smoother of Carter and Kohn (1994) and De Jong and Shephard (1995).

Block 2(a) Sampling the trend and temporary components \( \alpha \)

We first filter and draw the state vector \( \alpha = (y^\tau, \pi^\tau, u^\tau, \kappa, y^c, \zeta) \) conditionally on the time-varying parameters \( \beta \), the stochastic volatilities \( h \) and the hyperparameters \( \phi \). More specifically, using the general notation in equations (A-22)-(A-23), the unrestricted (i.e. \( \lambda = 1 \)) conditional state space representation is given by

\[
 \begin{bmatrix}
 w_t \\
 y_t \\
 \pi_t \\
 u_t \\
 y^c_t \\
 y^c_{t-1}
\end{bmatrix} = \begin{bmatrix}
 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
 0 & 1 & 0 & 1 & a & b & \beta^u_t \\
 0 & 0 & 1 & 0 & \beta^u_t & 0 & 0 \\
 \end{bmatrix}
 \begin{bmatrix}
 y^f_t \\
 \pi^f_t \\
 u^f_t \\
 \kappa_t \\
 \zeta_t \\
 y^c_t \\
 y^c_{t-1}
\end{bmatrix}
 + \begin{bmatrix}
 e_t \\
 e^y_t \\
 e_u^f_t \\
 e^\kappa_t \\
 e^\zeta_t \\
 e^u_t \\
 e^c_t
\end{bmatrix}, \quad (A-24)
\]

\[
 \begin{bmatrix}
 y^f_{t+1} \\
 \pi^f_{t+1} \\
 u^f_{t+1} \\
 \kappa_{t+1} \\
 \zeta_{t+1} \\
 y^c_{t+1} \\
 y^c_{t+1}
\end{bmatrix} = \begin{bmatrix}
 \kappa_0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0
\end{bmatrix}
 + \begin{bmatrix}
 1 & 0 & 0 & \sigma_{\psi, \kappa} & 0 & 0 & 0 \\
 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & \rho_1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
 \begin{bmatrix}
 y^f_t \\
 \pi^f_t \\
 u^f_t \\
 \kappa_t \\
 \zeta_t \\
 y^c_t \\
 y^c_{t-1}
\end{bmatrix}
 + \begin{bmatrix}
 \psi^y_t \\
 \psi^\pi_t \\
 \psi^u_t \\
 \psi^\kappa_t \\
 \psi^\zeta_t \\
 \psi^c_t \\
 \psi^{c_{t-1}}
\end{bmatrix}, \quad (A-25)
\]
unconditional distributions. Note that using \( \zeta \) each of these components while the stationary components \( \gamma \) are initialized by setting \( a_1 = 0 \) and \( A_1 = 1000 \) for each of these components while the stationary components \( \zeta \) and \( \gamma \) are initialized from their unconditional distributions. Note that using \( \kappa_0, \sigma_{\psi, \kappa} \) and \( \tilde{\kappa}, \kappa_t \) can easily be reconstructed from equation (27).

In the restricted model (i.e. \( \lambda_\kappa = 0 \)) \( \tilde{\kappa} \) is excluded from \( s_t^M \), with appropriate adjustment of the other matrices. In this case, no forward-filtering and backward-sampling is needed and \( \tilde{\kappa}_t \) can be sampled directly from its prior using equation (28).

**Block 2(b): Sampling the time-varying parameters \( \beta \)**

We next filter and draw the time-varying parameters \( \beta = (\beta^\pi, \beta^u) \) conditionally on the state vector \( \alpha \), the stochastic volatilities \( h \), the hyperparameters \( \phi \) and the binary indicators \( M \). More specifically, using equation (29) in (14) and (15), the unrestricted (i.e. \( \delta_j = 1 \)) conditional state space representations for the time-varying parameters \( \tilde{\beta}_t^\pi \) and \( \tilde{\beta}_t^u \) are given by

\[
\begin{bmatrix}
\pi_t - \pi_t^1 - \zeta_t - \beta^\pi_0 \tilde{y}_t^\pi \\
\tilde{\beta}_{t+1}^\pi \\
\end{bmatrix} = \begin{bmatrix} \sigma_{\eta, \pi \tilde{y}_t^\pi} & \tilde{\beta}_t^\pi \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_t^\pi \\ \varepsilon_t^\pi \\
\end{bmatrix},
\]

(A-26)

\[
\begin{bmatrix}
\sigma_t - \sigma_t^1 - \zeta_t - \beta^\pi_0 \tilde{y}_t^\pi \\
\tilde{\beta}_{t+1}^\pi \\
\end{bmatrix} = \begin{bmatrix} \sigma_{\eta, \pi \tilde{y}_t^\pi} & \tilde{\beta}_t^\pi \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_t^\pi \\ \varepsilon_t^\pi \\
\end{bmatrix}.
\]

(A-27)

with \( H_t = \sigma_t^2_{\epsilon, \pi} \) and \( Q_t = 1 \), and

\[
\begin{bmatrix}
u_t - \nu_t^1 - \beta^u_0 \tilde{y}_t^u \\
\tilde{\beta}_{t+1}^u \\
\end{bmatrix} = \begin{bmatrix} \sigma_{\eta, \pi \tilde{y}_t^u} & \tilde{\beta}_t^u \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_t^u \\ \varepsilon_t^u \\
\end{bmatrix},
\]

(A-28)

\[
\begin{bmatrix}
u_t - \nu_t^1 - \beta^u_0 \tilde{y}_t^u \\
\tilde{\beta}_{t+1}^u \\
\end{bmatrix} = \begin{bmatrix} \sigma_{\eta, \pi \tilde{y}_t^u} & \tilde{\beta}_t^u \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_t^u \\ \varepsilon_t^u \\
\end{bmatrix}.
\]

(A-29)

with \( H_t = \sigma_t^2_{\epsilon, u} \) and \( Q = 1 \). Both random walk components \( \tilde{\beta}_t^\pi \) and \( \tilde{\beta}_t^u \) are initialized by setting \( a_1 = 0 \) and \( A_1 = 1000 \).

In the restricted model (i.e. \( \delta_j = 0 \)), \( Z^M \) and \( s^M \) are empty. In this case, no forward-filtering and backward-sampling is needed and \( \tilde{\beta}_t^\pi \) can be sampled directly from its prior using equation (24). Note that the sampling of the state vector \( \alpha \) in block 2(a) depends on \( \beta_t^\pi \) rather than on \( \tilde{\beta}_t^\pi \). Using \( \beta_t^\pi, \sigma_{\psi, \kappa} \) and \( \tilde{\beta}_t^\pi, \beta_t^\pi \) can easily be reconstructed from equation (23).
Block 2(c): Sampling the mixture indicators $\iota$ and the stochastic volatilities $h$

In this block we draw the mixture indicators $\iota = (\iota^y, \iota^\pi, \iota^u, \iota^c, \iota^\zeta)$ and the stochastic volatilities $h = (h^y, h^\pi, h^u, h^c, h^\zeta)$ conditionally on the state vector $\alpha$, the time-varying parameters $\beta$, the hyperparameters $\phi$ and the binary indicators $M$. Following Del Negro and Primiceri (2014), in this block we first sample the mixture indicator $\iota_k^t$ (for $k = y, \pi, u, c, \zeta$) from its conditional probability mass

$$p(\iota_k^t = i | h_k^t, \epsilon_k^t) \propto q_i f_{N}(\epsilon_k^t | 2h_k^t + m_i - 1.2704, \nu_i^2),$$

(A-30)

with values for $\{q_i, m_i, \nu_i^2\}$ taken from Table 1 in Omori et al. (2007).

Next, we filter and sample the stochastic volatility terms $\hat{h}_k^t$ (for $k = y, \pi, u, c, \zeta$) conditioning on the transformed states $g_k^t$ defined in equations (A-16)-(A-20), on the mixture indicators $\iota_k^t$ and on the parameters $\phi$. More specifically, the unrestricted (i.e. $\theta_k = 1$) conditional state space representation is given by

$$w_t \begin{bmatrix} g_t^k - (m_{i_k^t} - 1.2704) - 2h_0^k \\ \tilde{h}_k^{t+1} \end{bmatrix} = \begin{bmatrix} 2\theta^k \sigma_{\gamma,k} \\ 1 \end{bmatrix} \begin{bmatrix} Z_t^M \\ s_t^M \end{bmatrix} + \begin{bmatrix} \epsilon_t \\ \gamma_t^k \end{bmatrix},$$

(A-31)

$$s_{t+1} = \begin{bmatrix} 1 \\ R_1 \end{bmatrix} \begin{bmatrix} g_t^k \\ \tilde{h}_k^t \end{bmatrix} + \begin{bmatrix} 1 \\ K_t \end{bmatrix} \begin{bmatrix} \gamma_t^k \end{bmatrix}.$$