FACULTEIT ECONOMIE
EN BEDRIJFSKUNDE

TWEEKERKENSTRAAT 2
B-9000 GENT
Tel. : 32 - (0)9 – 264.34.61
Fax. : 32 - (0)9 – 264.35.92

WORKING PAPER

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

Tino Berger
Gerdie Everaert
Hauke Vierke

April 2015
2015/903

D/2015/7012/05

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

Tino Berger[1], Gerdie Everaert[2], and Hauke Vierke[3]

[1,3]University of Göttingen
[2,3]Ghent University

April 6, 2015

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 set up an unobserved components model to decompose output, inflation and unemployment in their stochastic trend and business cycle gap components. The latter are related through the Phillips curve and Okun's Law. Key parameters such as the potential output growth rate, the slope of the Phillips curve and the strength of Okun's Law, are allowed to change over time in order to account for potential structural changes in the U.S. economy. Moreover, stochastic volatility is added to all components to account for shifts in macroeconomic volatility. A Bayesian stochastic model specification search is employed to test which parameters are time-varying and which unobserved components exhibit stochastic volatility. Using quarterly data from 1959Q2 to 2014Q3 we find substantial time variation in Okun's Law, while the Phillips curve slope appears to be stable. The potential output growth rate exhibits a drastic and persistent decline. 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 Phillip 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. Moreover, in the aftermath of the last two recessions, job growth was substantially lower than what the level of output growth would

have implied. These episodes, known as 'jobless recoveries', have let many observers to conclude that the trade-off between unemployment and output has changed. Finally, the severeness 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 false hypothesis testing. Second, regarding the relation 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 data prior to the Great recession and finds that the Phillips curve slope of a reduced-form equation for the U.S. 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). Third, regarding the relationship between unemployment and real economic activity, a related 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 actual rise. Owyang and Sekhposyan (2012) conclude that the relation between unemployment and output fluctuations changes significantly during the most recent recessions. 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 time variation 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. These determinants are unobserved. The same argument holds for Okun's Law which models the interaction between the output gap and the unemployment gap.[1] 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

[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.

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 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 Philips curve has a time-varying slope parameter which is found to vary considerably. Second, Chan et al. (2012) 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.

We contribute to this literature in the following way. Our unrestricted model nests important empirical models with time-varying parameters. In contrast to the existing literature, we start from a more general framework which allows most of the model's parameters to vary according to a random walk process. This allows for a very flexible evolution over time. We then select a parsimonious model by testing the relevance of the estimated time variation in each of the model's components. To this end, we use the Bayesian stochastic model specification search for state space models as outlined in Frühwirth-Schnatter and Wagner (2010). The Bayesian approach is well-suited to deal with the non-regular testing problem of deciding whether a component is fixed or time-varying. To the best of our knowledge, this is the first study to allow and explicitely test for a wide range of time-varying parameters in a macroeconomic time-series model. A 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. First, it lays out a multivariate unobserved components model with time varying parameters and stochastic volatilities, designed to fit U.S. macroeconomic data. Second, the Bayesian stochastic model selection approach is explained followed by 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_t^c and a non-stationary trend y_t^τ referred to as potential output

$$y_t = y_t^\tau + y_t^c + \varepsilon_t^y, \qquad \varepsilon_t^y \sim i.i.d.\mathcal{N}(0, \sigma_{\varepsilon,y}^2), \tag{1}$$

where ε_t^y is included to capture measurement error and non-persistent shocks. Potential output is modeled as a random walk process with stochastic drift κ_t

$$y_{t+1}^\tau = \kappa_t + y_t^\tau + \exp\{h_t^y\}\,\psi_t^y, \qquad \psi_t^y \sim i.i.d.\mathcal{N}(0,1), \tag{2}$$

$$\kappa_{t+1} = \kappa_t + \psi_t^\kappa, \qquad \psi_t^\kappa \sim i.i.d.\mathcal{N}(0, \sigma_\kappa^2). \tag{3}$$

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 κ_t. The output gap y_t^c is modeled as a stationary autoregressive (AR) process of order two

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

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).[2] 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 π_t^τ and the output gap is used as a measure of real activity, i.e.

$$\pi_t - \pi_t^\tau = \omega E_t(\pi_{t+1} - \pi_{t+1}^\tau) + \beta_t^\pi y_t^c + \widetilde{\zeta}_t, \tag{5}$$

where ω is a discount factor. As shown by Beveridge and Nelson (1981), in the presence of a zero-mean transitory component the trend component π_t^τ 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 $\widetilde{\zeta}_t$ 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. A s we want to analyze whether the slope of the Phillips curve is time-varying we allow β_t^π 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 β_t^π to vary using a three-state Markov switching model. Iterating equation (5) forward and rearranging yields

$$\begin{aligned}\pi_t &= \pi_t^\tau + \lim_{j\to\infty} \omega^j E_t(\pi_{t+j} - \pi_{t+j}^\tau) + \sum_{j=0}^{\infty} \omega^j E_t(\beta_{t+j}^\pi y_{t+j}^c) + \zeta_t, \\ &= \pi_t^\tau + \sum_{j=0}^{\infty} \omega^j E_t(\beta_{t+j}^\pi y_{t+j}^c) + \zeta_t, \end{aligned} \tag{6}$$

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

$$\pi_t = \pi_t^\tau + \pi_t^c + \varepsilon_t^\pi, \qquad \varepsilon_t^\pi \sim i.i.d.\mathcal{N}(0, \sigma_{\varepsilon,\pi}^2), \tag{7}$$

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

where π_t^c is the inflation gap given by

$$\pi_t^c = \sum_{j=0}^{\infty} \omega^j E_t(\beta_{t+j}^{\pi} y_{t+j}^c) + \zeta_t. \tag{8}$$

The idiosyncratic term ε_t^{π} is added in equation (7) to capture measurement error and non-persistent shocks. Trend inflation π_t^{τ} is modeled as a driftless random walk

$$\pi_{t+1}^{\tau} = \pi_t^{\tau} + \exp\{h_t^{\pi}\}\psi_t^{\pi}, \qquad \psi_t^{\pi} \sim i.i.d.\mathcal{N}(0,1), \tag{9}$$

where the innovations ψ_t^{π} 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 β_t^{π} is allowed to change over time according to a random walk

$$\beta_{t+1}^{\pi} = \beta_t^{\pi} + \eta_t^{\pi}, \qquad \eta_t^{\pi} \sim i.i.d.\mathcal{N}(0, \sigma_{\eta,\pi}^2). \tag{10}$$

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

$$\zeta_{t+1} = \varrho\zeta_t + \exp\left\{h_t^{\zeta}\right\}\psi_t^{\zeta}, \qquad \psi_t^{\zeta} \sim i.i.d.\mathcal{N}(0,1). \tag{11}$$

Given the DGPs of y_t^c and β_t^{π} in equations (4) and (10) and the assumption that ψ_t^c and η_t^{π} are mutually uncorrelated error terms, the output gap term $\sum_{j=0}^{\infty} \omega^j E_t(\beta_{t+j}^{\pi} y_{t+j}^c)$ in equation (8) can be expressed as

$$\pi_t^c = \beta_t^{\pi} \begin{bmatrix} 1 & 0 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \omega \begin{bmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} y_t^c \\ y_{t-1}^c \end{bmatrix}, \tag{12}$$

$$= \frac{\beta_t^{\pi}}{1 - \omega\rho_1 - \omega^2\rho_2} \left(y_t^c + \omega\rho_2 y_{t-1}^c\right). \tag{13}$$

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

$$\pi_t = \pi_t^{\tau} + \beta_t^{\pi}\widetilde{y}_t^c + \zeta_t + \varepsilon_t^{\pi}, \tag{14}$$

where $\widetilde{y}_t^c = \frac{1}{1-\omega\rho_1-\omega^2\rho_2}\left(y_t^c + \omega\rho_2 y_{t-1}^c\right)$.

Unemployment: a time-varying Okun's Law relation

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

$$u_t = u_t^{\tau} + \beta_t^u y_t^c + \varepsilon_t^u, \qquad \varepsilon_t^u \sim i.i.d.\mathcal{N}(0, \sigma_{\varepsilon,u}^2), \tag{15}$$

where ε_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^τ as a random walk process

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

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 β_t^u is allowed to change over time according to a random walk process

$$\beta_{t+1}^u = \beta_t^u + \eta_t^u, \qquad \eta_t^u \sim i.i.d.\mathcal{N}(0,\sigma_{\eta,u}^2). \tag{17}$$

Stochastic volatilities

All stochastic volatilities are modeled as random walks

$$h_{t+1}^k = h_t^k + \gamma_t^k, \qquad \gamma_t^k \sim i.i.d.\mathcal{N}(0,\sigma_{\gamma,k}^2), \tag{18}$$

for $k = y, \pi, u, c, \zeta$. A key feature of the stochastic volatility components $\exp\left\{h_t^k\right\}\psi_t^k$ is that they are nonlinear but can be transformed into linear components by taking the logarithm of their squares

$$\ln\left(\exp\left\{h_t^k\right\}\psi_t^k\right)^2 = 2h_t^k + \ln\left(\psi_t^k\right)^2, \tag{19}$$

where $ln\left(\psi_t^k\right)^2$ 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

$$g_t^k = 2h_t^k + \epsilon_t^k, \tag{20}$$

where $g_t^k = \ln\left(\left(\exp\left\{h_t^k\right\}\psi_t^k\right)^2 + c\right)$ with $c = .001$ being an offset constant, and the distribution of ϵ_t^k being the following mixture of normals

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

with component probabilities q_i, means $m_i - 1.2704$ and variances ν_i^2. Equivalently, this mixture density can be written in terms of the component indicator variable ι_t^k as

$$\epsilon_t^k | \left(\iota_t^k = i\right) \sim \mathcal{N}\left(m_i - 1.2704, \nu_i^2\right), \qquad \text{with} \qquad Pr\left(\iota_t^k = i\right) = q_i. \tag{22}$$

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 the Subsection 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 β_t^π, ϱ and $\sigma^2_{\epsilon,\pi}$ 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 ϱ and $\sigma^2_{\epsilon,\pi}$ to zero and restrict β_t^π 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 time-varying. This implies testing $\sigma^2_{\eta,\pi} = 0$ against $\sigma^2_{\eta,\pi} > 0$, 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 ζ_t should be included in equation (14) or whether the stochastic volatilities are relevant.

In principle we could derive the reduced form VARMA representation of our model and apply standard structural break tests for mean and variances. However, by testing for time-variation in the UC model instead of the the reduced form VARMA model, we can distinguish between changes in the volatilities to permanent versus transitory shocks. Moreover, deriving the reduced form VARMA representation requires assumptions regarding the order of integration of all variables. However, for output the order of integration is an outcome of our testing procedure as we allow potential output growth to be either constant or evolve as a random walk in which case output is I(2).

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. $\sigma^2_{\eta,\pi}$ and $\sigma^2_{\eta,u}$, this implies rearranging equations (10) and (17) to

$$\beta^j_{t+1} = \beta^j_0 + \sigma_{\eta,j}\widetilde{\beta}^j_{t+1}, \tag{23}$$

$$\text{with} \quad \widetilde{\beta}^j_{t+1} = \widetilde{\beta}^j_t + \widetilde{\eta}^j_t, \qquad \widetilde{\beta}^j_0 = 0, \qquad \widetilde{\eta}^j_t \sim i.i.d.\mathcal{N}(0,1), \tag{24}$$

for $j = \pi, u$ and where β^j_0 is the initial value of the level of β^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 $\widetilde{\beta}^j_t$ 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 is time-varying, i.e. $\sigma^2_{\eta,j} > 0$, then the likelihood function will concentrate around the two modes $-\sigma_{\eta,j}$ and $\sigma_{\eta,j}$. For $\sigma^2_{\eta,j} = 0$ the likelihood function will become unimodal around zero. As such, allowing for non-identification of $\sigma_{\eta,j}$ provides useful information on whether $\sigma^2_{\eta,j} > 0$.

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

$$h^k_{t+1} = h^k_0 + \sigma_{\gamma,k}\widetilde{h}^k_{t+1}, \tag{25}$$

$$\text{with} \quad \widetilde{h}^k_{t+1} = \widetilde{h}^k_t + \widetilde{\gamma}^k_t, \qquad \widetilde{h}^k_0 = 0, \qquad \widetilde{\gamma}^k_t \sim i.i.d.\mathcal{N}(0,1), \tag{26}$$

for $k = y, \pi, u, c, \zeta$ and where $h^k_0 = 0$ is the initial value of the level of h^k_t.

Finally, the non-centered parameterization of the time-varying drift in equation (3) is given by

$$\kappa_{t+1} = \kappa_0 + \sigma_\kappa\widetilde{\kappa}_{t+1}, \tag{27}$$

$$\text{with} \quad \widetilde{\kappa}_{t+1} = \widetilde{\kappa}_t + \widetilde{\psi}^\kappa_t, \qquad \widetilde{\kappa}_0 = 0, \qquad \widetilde{\psi}^\kappa_t \sim i.i.d.\mathcal{N}(0,1), \tag{28}$$

and where $\kappa_0 = 0$ is the initial value of the level of κ_t.

Parsimonious specification

A second advantage of the non-centered parameterization is that when e.g. $\sigma^2_{\eta,\pi} = 0$ the transformed component $\widetilde{\beta}^\pi_t$, in contrast to β_t, does not degenerate to the time-invariant slope of the Phillips curve as this is now represented by β^π_0. As such, the question whether the slopes of the Phillips curve and Okun's Law are time-varying 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

$$\beta^j_t = \beta^j_0 + \delta_j\sigma_{\eta,j}\widetilde{\beta}^j_t, \tag{29}$$

for $j = \pi, u$ and where δ_j is a binary indicator which is either 0 or 1. If $\delta_j = 0$, the component $\widetilde{\beta}^j_t$ drops from the model such that β^j_0 represents the constant slope parameter. If $\delta_j = 1$ then $\widetilde{\beta}^j_t$ is included in the model and $\sigma_{\eta,j}$ is estimated from the data. In this case β^j_0 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

$$h_t^k = h_0^k + \theta_k \sigma_{\gamma,k} \widetilde{h}_t^k, \tag{30}$$

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

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

$$\kappa_t = \kappa_0 + \lambda \sigma_\kappa \widetilde{\kappa}_t, \tag{31}$$

where λ is a binary indicator that is either 0 or 1. If $\lambda = 0$, the component $\widetilde{\kappa}_t$ drops from the model such that κ_0 is the constant drift in potential output. If $\lambda = 1$ then $\widetilde{\kappa}_t$ is included in the model and σ_κ is estimated from the data. In this case κ_0 is the initial value of the time-varying drift κ_t.

Collecting the binary indicators in the vector $\mathcal{M} = (\delta_\pi, \delta_u, \theta_y, \theta_\pi, \theta_u, \theta_c, \theta_\zeta, \lambda)$, 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 $\rho = (\rho_1, \rho_2)$, ϱ, $\beta_0 = (\beta_0^\pi, \beta_0^u)$ and $h_0 = (h_0^y, h_0^\pi, h_0^u, h_0^c, h_0^\zeta)$, for the binary indicators $\mathcal{M}$ and for the variances of the idiosyncratic factors $\sigma_\varepsilon^2 = \left(\sigma_{\varepsilon,y}^2, \sigma_{\varepsilon,\pi}^2, \sigma_{\varepsilon,u}^2\right)$, the innovations to the drift component σ_κ^2, the time-varying parameters $\sigma_\eta^2 = \left(\sigma_{\eta,\pi}^2, \sigma_{\eta,u}^2\right)$ and the stochastic volatility components $\sigma_\gamma^2 = \left(\sigma_{\gamma,y}^2, \sigma_{\gamma,\pi}^2, \sigma_{\gamma,u}^2, \sigma_{\gamma,c}^2, \sigma_{\gamma,\zeta}^2\right)$.

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 σ^2 by a Gaussian prior centered at zero on σ. Centering the prior distribution at zero makes sense as for both $\sigma^2 = 0$ and $\sigma^2 > 0$, σ is symmetric around zero. Frühwirth-Schnatter and Wagner (2010) show that, compared to using an inverse Gamma prior for σ^2, the posterior density of σ 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 σ_η, σ_κ and σ_γ, 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 σ_ε^2, which are always included in the model, we choose the standard inverse Gamma prior distribution. For each of the model parameters in ρ, ϱ and β we assume a normal prior distribution. Details on the prior distributions are presented in Subsection 3.2 below. For the binary indicators $\mathcal{M}$ we choose a uniform prior distribution over all combinations of the indicators such that each model has the same prior probability, i.e. $p(\mathcal{M}) = 2^{-8}$, 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 β_t^π and β_t^u on the unobserved output gap y_t^c and the stochastic volatilities h_t^k in the state space model given in eq. (1) - (18) and the use of the stochastic model specification search outlined in Subsection 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 $\alpha_t = (y_t^\tau, \pi_t^\tau, u_t^\tau, y_t^c, \zeta_t, \kappa_t)$, a time-varying parameter vector $\beta_t = (\beta_t^\pi, \beta_t^u)$, a stochastic volatilities vector $h_t = \left(h_t^y, h_t^\pi, h_t^u, h_t^c, h_t^\zeta\right)$ and an indicator vector $\iota_t = \left(\iota_t^y, \iota_t^\pi, \iota_t^u, \iota_t^c, \iota_t^\zeta\right)$. The unknown parameters are collected in the vector $\phi = \left(\rho, \varrho, \beta_0, \sigma, \sigma_\varepsilon^2\right)$, with $\sigma = (\sigma_\eta, \sigma_\kappa, \sigma_\gamma)$. Finally, let $x_t = (y_t, \pi_t, u_t)$ be the data vector. Stacking observations over time, we denote $x = \{x_t\}_{t=1}^T$ and similarly for α, β, h and ι. The posterior density of interest is then given by $f(\alpha, \beta, h, \iota, \phi, \mathcal{M}|x)$. Following Frühwirth-Schnatter and Wagner (2010) our MCMC scheme is as follows:

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

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

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

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 $(B+1, ..., J)$ approximates a sample from the virtual posterior distribution $f(\alpha, \beta, h, \iota, \phi, \mathcal{M}|x)$. Details on the exact implementation of each of the blocks can be found in Appendix A. The results reported below are based on 35,000 Gibbs sampler iterations, with the first 10,000 discarded as a burn-in period. We store every 5^{th} of the remaining 25,000 iterations, leaving 5,000 draws for inference.

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 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 $\sigma_\varepsilon^2 = \left(\sigma_{\varepsilon,y}^2, \sigma_{\varepsilon,\pi}^2, \sigma_{\varepsilon,u}^2\right)$ we use the inverse Gamma prior $IG(c_0, C_0)$ where the shape $c_0 = \nu_0 T$ and scale $C_0 = s_0\sigma_0^2$ parameters are calculated from the *prior belief* σ_0^2 about the variance parameter and the *prior strength* ν_0 which is expressed as a fraction of the sample size T.[3] Following the notation in Frühwirth-Schnatter and Wagner (2010), for the remaining parameters we use a Gaussian prior $\mathcal{N}(a_0, A_0\sigma_e^2)$ in a regression with homoskedastic errors and $\mathcal{N}(a_0, A_0)$ 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, $\sigma_{\varepsilon,g}^2$, $\sigma_{\varepsilon,g}^2$, $\sigma_{\varepsilon,\pi}^2$:** 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,

[3] Since this prior is conjugate, $\nu_0 T$ can be interpreted as the number of fictitious observations used to construct the prior belief σ_0^2.

Table 1: Prior distributions of model parameters

Inverse Gamma priors: $IG(c_0, C_0) = IG(\nu_0 T, \nu_0 T \sigma_0^2)$				Percentiles	
		σ_0	ν_0	2.5%	97.5%
idiosyncratic component output	$\sigma_{\varepsilon,y}$	0.10	0.10	0.080	0.141
idiosyncratic component inflation	$\sigma_{\varepsilon,\pi}$	1.00	0.10	0.777	1.408
idiosyncratic component unemployment	$\sigma_{\varepsilon,u}$	0.50	0.10	0.387	0.701
Gaussian priors homoskedastic errors: $\mathcal{N}(a_0, A_0 \sigma_e^2)$				Percentiles	
Regression parameters		a_0	$\sqrt{A_0} \times \sigma_e$	2.5%	97.5%
const. Phillips curve slope	β_0^π	0.20	0.25×1.0	-0.290	0.690
const. Okun coefficient	β_0^u	-0.50	0.25×0.5	-0.745	-0.255
Non-centered components					
std. of time-varying Phillips curve	$\sigma_{\eta,\pi}$	0.00	1.00×1.0	-1.960	1.960
std. of time-varying Okun coefficient	$\sigma_{\eta,u}$	0.00	1.00×0.5	-0.980	0.980
Gaussian priors SV errors: $\mathcal{N}(a_0, A_0)$				Percentiles	
Regression parameters		a_0	$\sqrt{A_0}$	2.5%	97.5%
1st AR lag: output gap	ρ_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	ϱ	0.70	0.05	0.602	0.798
const. output drift	κ_0	0.75	0.10	0.554	0.946
Stochastic volatility parameters					
const. volatility of potential output	h_0^y	$\ln(0.10)$	0.10	$\ln(0.082)$	$\ln(0.122)$
const. volatility of trend inflation	h_0^π	$\ln(0.20)$	0.10	$\ln(0.164)$	$\ln(0.243)$
const. volatility of NAIRU	h_0^u	$\ln(0.05)$	0.10	$\ln(0.041)$	$\ln(0.061)$
const. volatility of output gap	h_0^c	$\ln(0.60)$	0.10	$\ln(0.493)$	$\ln(0.730)$
const. volatility of temporary inflation	h_0^ζ	$\ln(0.70)$	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	σ_κ	0.00	1.00	-1.960	1.960

Notes: We set IG priors on the variance parameters σ^2 but in the top panel of this table we report details on the implied prior distribution for the standard deviations σ 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 $exp\{h_0\}$.

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.[4] The prior distribution on the time-varying part of the volatility of the trends is uninformative and centered at zero: $\sigma_{\gamma,k} \sim \mathcal{N}(0,1)$, for $k = y, \pi, u$.

- **Potential output growth,** κ: In line with existing estimates, our prior belief about the time-invariant part of the output drift is given by $\kappa \sim \mathcal{N}(0.75, 0.1^2)$. 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 $\sigma_\kappa \sim \mathcal{N}(0,1)$.

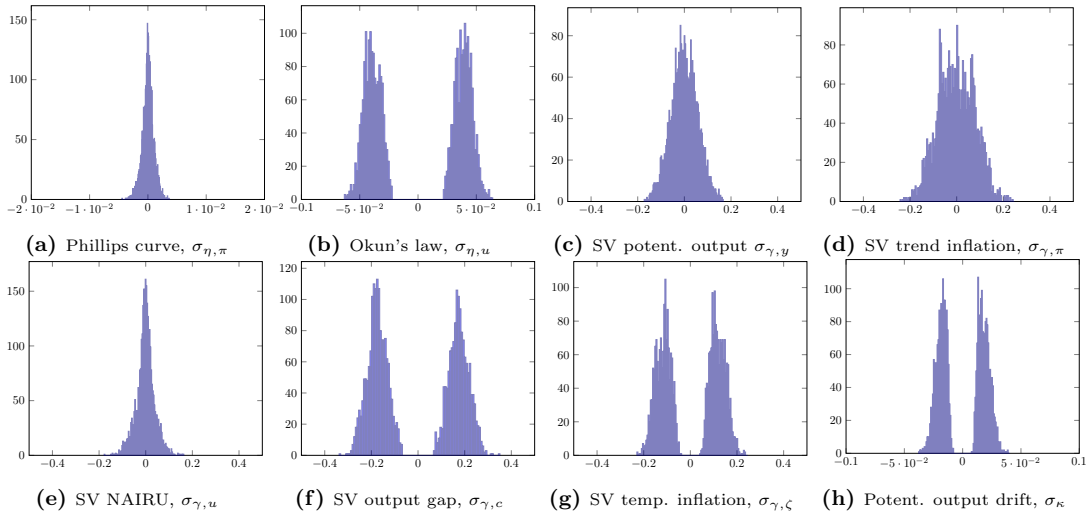
- **Output gap,** ρ**,** $\exp\{h^c\}$: While the output gap is stationary by assumption, it is often found to be a very persistent process (see e.g. Morley et al., 2003; Kim et al., 2014). In order to ensure stationarity, we find it useful to impose prior information on the sum of the AR(2) parameters instead of restricting each parameter separately. Hence, we use an informative prior on the sum $(\rho_1 + \rho_2) \sim \mathcal{N}(0.9, 0.015^2)$ and a much less informative prior on the first lag $\rho_1 \sim \mathcal{N}(1.25, 0.5^2)$. 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; 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 ρ_1, which implies a prior belief of -0.35 for ρ_2, is in line with the typical hump-shaped pattern in response to cyclical shocks. With a prior standard deviation of 0.5 we are very uninformative on these individual parameters, though. 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. $\sigma_{\gamma,c} \sim \mathcal{N}(0,1)$.

- **AR(1) inflation component,** ϱ**,** $\exp\{h^\zeta\}$: We set the prior distribution for the autoregressive coefficient of the AR(1) inflation component to $\varphi \sim \mathcal{N}(0.7, 0.05^2)$. The relative small standard deviation ensures that ϱ 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 ϱ becomes too small, ζ is indistinguishable from the white noise inflation component ε_t^π. The prior distribution of the time-invariant part of the volatility component is set to $h_0^\zeta \sim \mathcal{N}(\ln(0.7), 0.1^2)$. A loose prior is used for the standard deviation of the time-varying component: $\sigma_{\gamma,\zeta} \sim \mathcal{N}(0,1)$, allowing for a high degree of time variation in ζ_t as found in Kim et al. (2014).

- **Slope of Phillips curve,** β^π: Estimates for the slope of the Phillips curve in the literature differ depending on whether a forward or backward-looking curve is modeled. In forward-looking specifications, β^π is often found small and statistically insignificant (see e.g. Kim et al., 2014). For the time-invariant part β_0^π we set a prior distribution of $\beta_0^\pi \sim \mathcal{N}(0.2, 0.25^2)$.

[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.

Figure 1: Posterior distributions of the standard deviations for the non-centered variables in the unrestricted model (all binary indicators set to 1)



(a) Phillips curve, $\sigma_{\eta,\pi}$ **(b)** Okun's law, $\sigma_{\eta,u}$ **(c)** SV potent. output $\sigma_{\gamma,y}$ **(d)** SV trend inflation, $\sigma_{\gamma,\pi}$

(e) SV NAIRU, $\sigma_{\gamma,u}$ **(f)** SV output gap, $\sigma_{\gamma,c}$ **(g)** SV temp. inflation, $\sigma_{\gamma,\zeta}$ **(h)** Potent. output drift, σ_{κ}

Our prior belief about the degree of time variation in the Phillips curve is uninformative, i.e. $\sigma_{\eta,\pi} \sim \mathcal{N}(0, 1)$.

- **Okun coefficient,** β^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 $\sigma_{\eta,u} \sim \mathcal{N}(0, 1)$.

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 (σ) 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 ($\sigma_{\eta,u}$), to the volatility of the output gap ($\sigma_{\gamma,c}$) and the temporary inflation component ($\sigma_{\gamma,\zeta}$) and to the drift in potential output (σ_{κ}). 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 σ. 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 model selection rejects time variation in the slope of the Phillips curve. Over all four prior specifications, the posterior probability for a model with a time-varying Phillips curve slope is either far below or just above one percent. Second, the data clearly favor time variation in the Okun's Law parameter. Third, for the trend components in output, inflation and unemployment, a model with a constant volatility fits the data best. In our benchmark case ($A_0 = 1$) the posterior probabilities of a stochastic volatility component in the trend components varies between 8 and 18%, while the probabilities fall well below 5% when more diffuse priors are used. When the prior distribution allows for little time variation ($A_0 = 0.1$), the inclusion probabilities of the stochastic volatility components increase, but remain below 0.5.[5]

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

Prior		Posterior							
		Time-varying parameter			Stochastic volatility				
		Phillips curve	Okun's law	Output drift	Potential output	Trend inflation	NAIRU	Output gap	Temp. inflation
p_0	A_0	δ_π	δ_u	λ	θ_y	θ_π	θ_u	θ_c	θ_ζ
0.5	0.1	0.0110	1.0000	1.0000	0.2150	0.1704	0.3084	1.0000	1.0000
0.5	1	0.0026	1.0000	1.0000	0.1456	0.0828	0.1992	1.0000	1.0000
0.5	10	0.0007	1.0000	1.0000	0.0247	0.0263	0.0923	1.0000	1.0000
0.5	100	0.0000	1.0000	1.0000	0.0219	0.0206	0.0429	1.0000	1.0000

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 $\delta_u, \theta_c, \theta_\zeta$ and λ have inclusion probabilities of ≥ 0.5, while the indicators $\delta_\pi, \theta_y, \theta_\pi$ and θ_u are excluded in the majority of all draws.

Besides inference on the importance of time variation in the individual components, the model

[5]The 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 $p_0 = 0.5$.

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

Prior		Posterior							
		Time-varying parameter			Stochastic volatility				
		Phillips curve	Okun's law	Output drift	Potential output	Trend inflation	NAIRU	Output gap	Temp. inflation
p_0	A_0	δ_π	δ_u	λ	θ_y	θ_π	θ_u	θ_c	θ_ζ
0.5	1	0.0026	1.0000	1.0000	0.1456	0.0828	0.1992	1.0000	1.0000
0.9	1	0.0192	1.0000	1.0000	0.4606	0.4478	0.4484	1.0000	1.0000
0.1	1	0.0004	1.0000	1.0000	0.0174	0.0302	0.0496	1.0000	1.0000

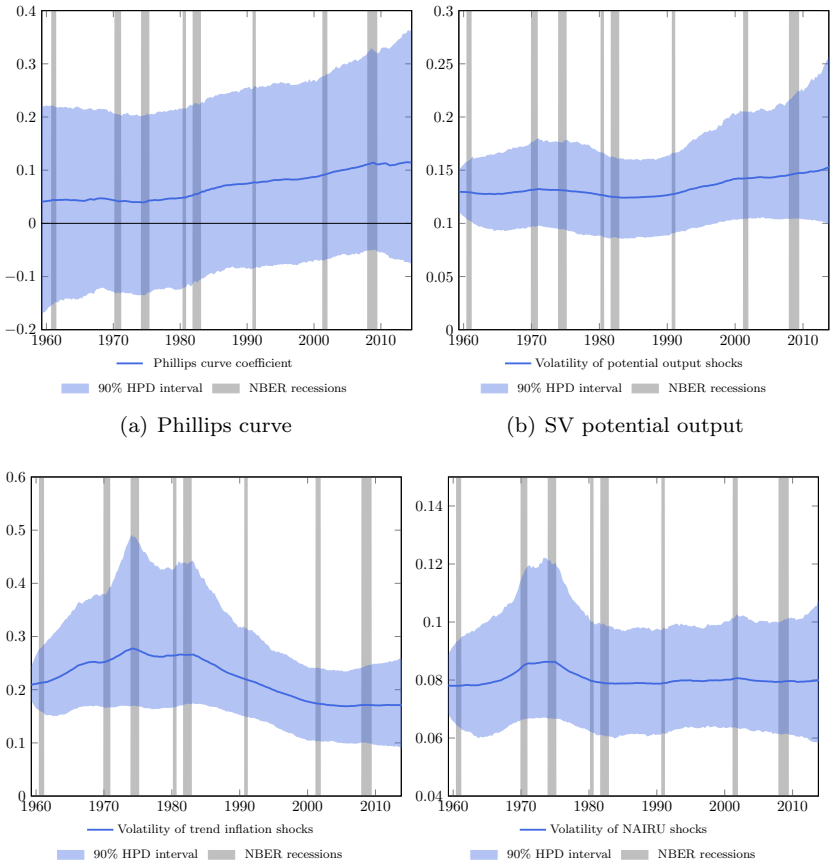
selection search also allows to compute overall model probabilities. The introduction of 8 binary indicators leads to 2^8 possible models. As 4 out of the 8 binary indicators have low individual probabilities, most models have a probability of zero. As a result, in the benchmark case where $A_0 = 1$ 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 45%, which rises to 92% when diffuse priors ($A_0 = 100$) are used. The three other models with notable probabilities larger than 10% 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 $p_0 = 0.5$)

Model								Posterior probability			
δ_π	δ_u	λ	θ_y	θ_π	θ_u	θ_c	θ_ζ	$A_0 = 0.1$	$A_0 = 1$	$A_0 = 10$	$A_0 = 100$
0	1	1	0	0	0	1	1	0.4454	0.6415	0.8620	0.9180
0	1	1	0	0	1	1	1	0.1976	0.1334	0.0864	0.0395
0	1	1	0	1	0	1	1	0.0888	0.1080	0.0228	0.0183
0	1	1	1	0	0	1	1	0.1290	0.0555	0.0223	0.0208
0	1	1	0	1	1	1	1	0.0438	0.0342	0.0034	0.0023
0	1	1	1	1	0	1	1	0.0192	0.0121	0.0001	0.0000
0	1	1	1	0	1	1	1	0.0502	0.0111	0.0023	0.0011

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.

Figure 2: Evolution of the time-varying components not selected by the model search



(a) Phillips curve

(b) SV potential output

(c) SV trend inflation

(d) SV NAIRU

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 20^{th}-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 in Figure 4.[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.

[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.

Figure 3: 20^{th}-order autocorrelations of all parameter and component draws

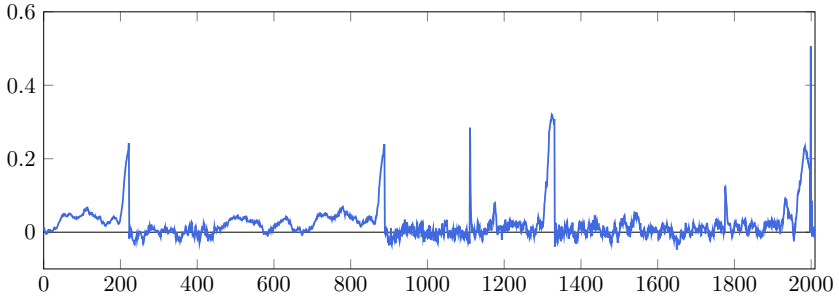
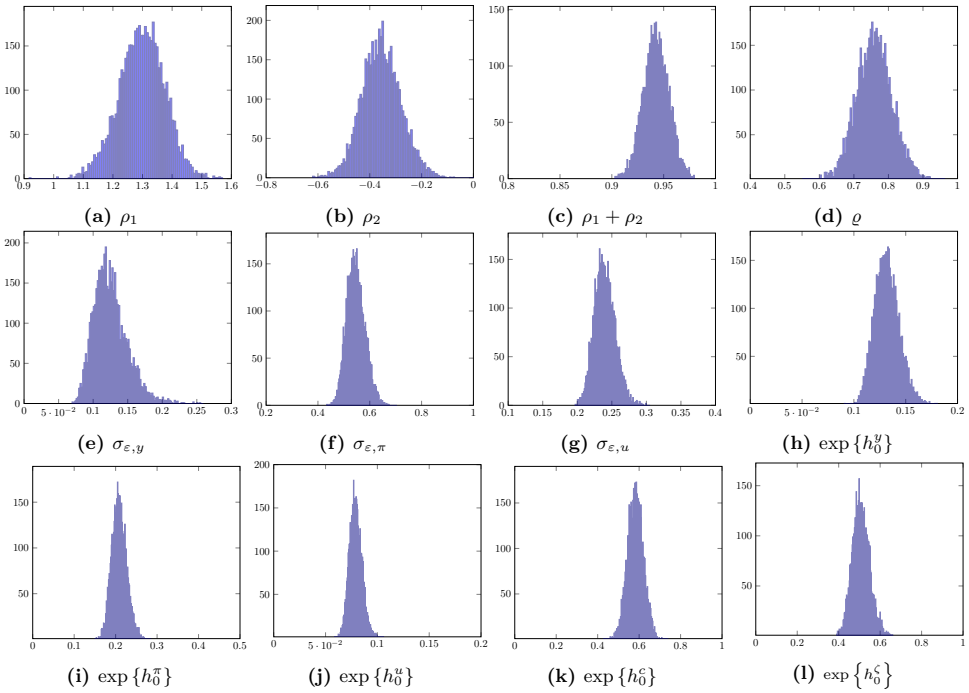


Figure 4: Posterior distributions of parameters (parsimonious model)



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 variance 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 the null of constant volatility for trend inflation. It contrasts with Stock and Watson (2007) who find considerable

Table 5: Posterior distributions of model parameters (parsimonious model)

Inverse Gamma			Percentiles	
		median	2.5%	97.5%
idiosyncratic component output	$\sigma_{\varepsilon,y}$	0.122	0.091	0.168
idiosyncratic component inflation	$\sigma_{\varepsilon,\pi}$	0.544	0.492	0.607
idiosyncratic component unemployment	$\sigma_{\varepsilon,u}$	0.239	0.217	0.266
Gaussian			Percentiles	
Regression parameters		median	2.5%	97.5%
const. Phillips curve slope	β_0^{π}	0.067	−0.061	0.196
const. Okun coefficient	β_0^{u}	−0.438	−0.525	−0.350
1st AR lag: output gap	ρ_1	1.299	1.166	1.418
sum of AR lags: output gap	$\rho_1 + \rho_2$	0.942	0.922	0.963
AR lag: AR(1) inflation component	ϱ	0.759	0.676	0.846
const. output drift	κ_0	1.023	0.933	1.118
Stochastic volatility parameters				
const. volatility of potential output	$\exp\{h_0^y\}$	0.131	0.113	0.152
const. volatility of trend inflation	$\exp\{h_0^{\pi}\}$	0.208	0.181	0.239
const. volatility of NAIRU	$\exp\{h_0^{u}\}$	0.079	0.070	0.089
const. volatility of output gap	$\exp\{h_0^{c}\}$	0.582	0.519	0.646
const. volatility of temporary inflation	$\exp\left\{h_0^{\zeta}\right\}$	0.505	0.441	0.579
Non-centered components				
variance of time-varying Okun coefficient	$\sigma^2_{\eta,u}$	0.0016	0.0009	0.0029
variance of SV: output gap	$\sigma^2_{\gamma,c}$	0.0350	0.0131	0.0757
variance of SV: AR(1) inflation component	$\sigma^2_{\gamma,\zeta}$	0.0155	0.0053	0.0397
variance of time-varying output drift	σ^2_{κ}	0.0003	0.0001	0.0007

variability in the variance of innovations to trend inflation. However, they do not allow for a persistent transitory component (ζ_t in our model) in the inflation gap such that trend inflation has to incorporate this component.[7] Similar to the recent literature, we find that the inflation gap and idiosyncratic shocks are the most important driver of inflation. On average, they account for more than 90% of the variance of inflation changes at the one-quarter horizon. The inflation gap itself is driven by the output gap y_t^c but more importantly by the persistent AR(1) component ζ_t. As can be seen from Figure 6, ζ_t and its stochastic volatility peak in the 1970s. Inflation is not very sensitive to the output gap. As shown in panel (a) of Figure 7, we estimate the slope of the (time-invariant) Phillips curve to be very small with a posterior median of 0.07 and a 90% HPD interval ranging from -0.06 to 0.20. This finding confirms Kim et al. (2014) but contrasts with Morley et al. (2013) who find the real activity gap, as measured by the unemployment gap, to be

[7] In fact when we drop ζ_t, the model selection procedure selects a specification with stochastic volatility for trend inflation (results not reported).

an important driver of the inflation gap.

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 ζ_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 crop 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 week 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 5: Trend inflation (parsimonious model)

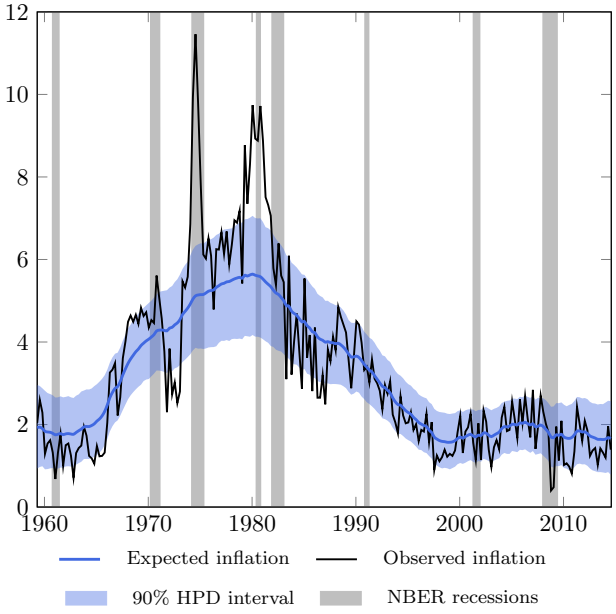
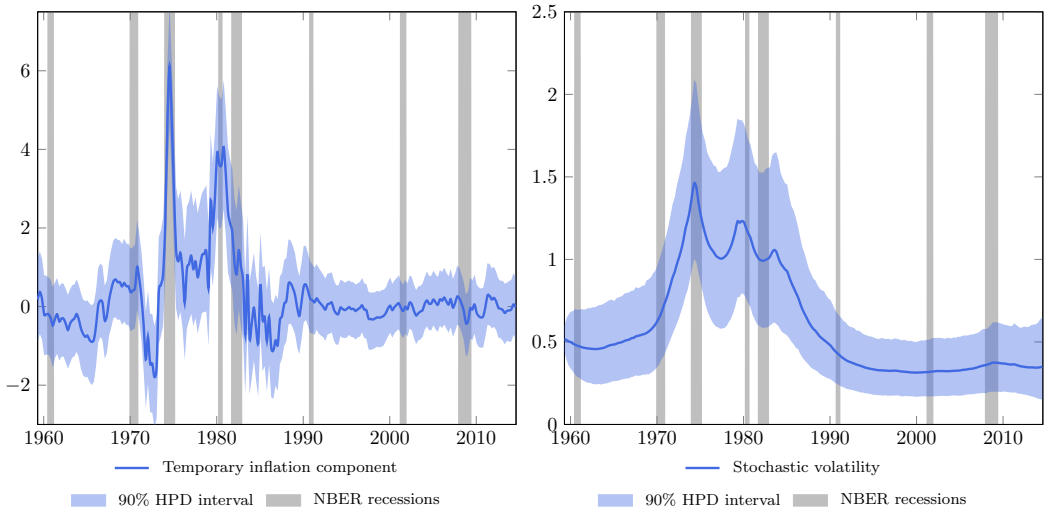


Figure 6: AR(1) inflation component (parsimonious model)



(a) AR(1) inflation component (ζ_t)

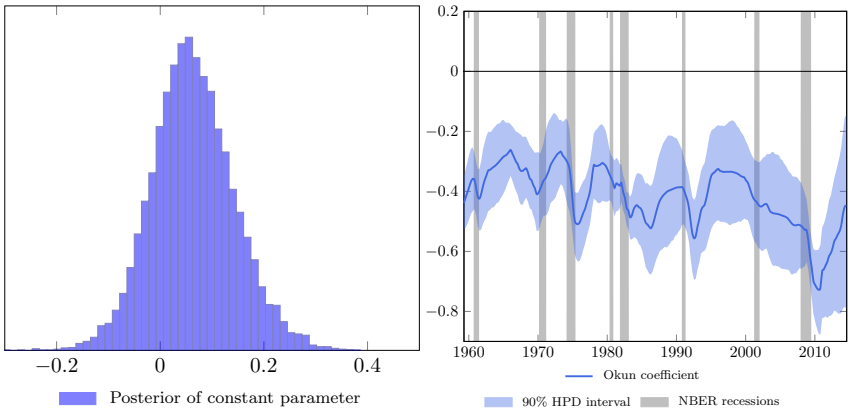
(b) Stochastic volatility of ζ_t

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 Perron and Wada (2009) who highlight the importance of accounting for breaks in potential output growth for UC models. By analyzing data from 1974 to 1998 they find one break in 1973.

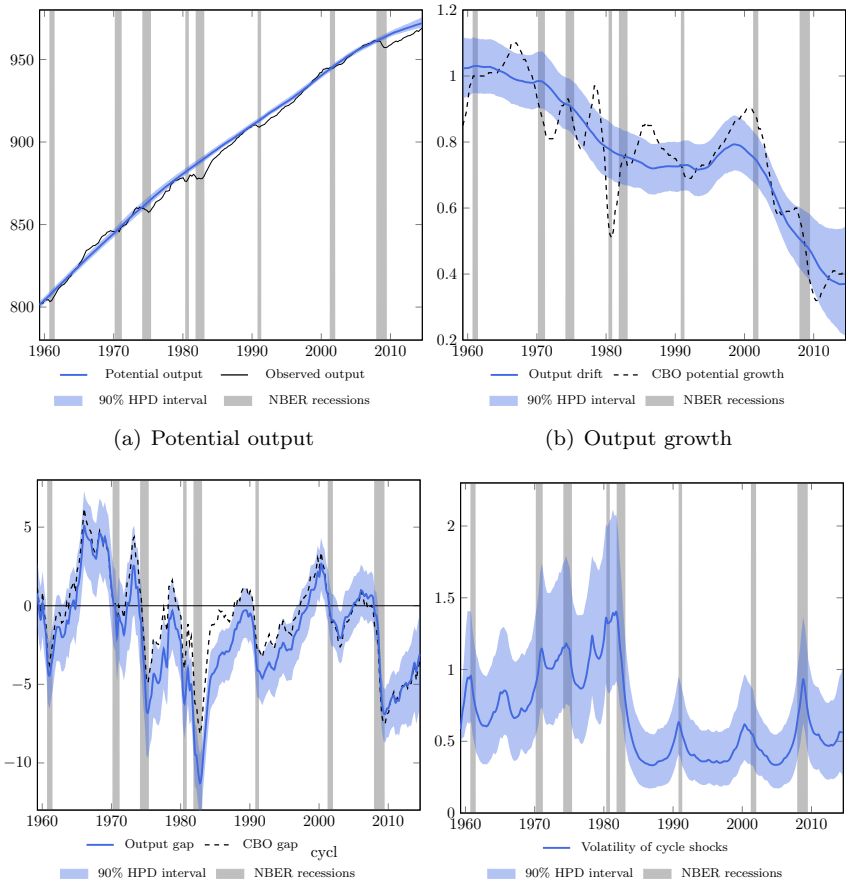
Figure 7: Phillips curve and Okun's Law (parsimonious model)



(a) Phillips curve β_t^{π} (b) Okun's Law β_t^{u}

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 has not led to a permanent increase as it returns almost to its pre-crisis level in 2009.

Figure 8: Output components (parsimonious model)



(a) Potential output

(b) Output growth

(c) Output gap

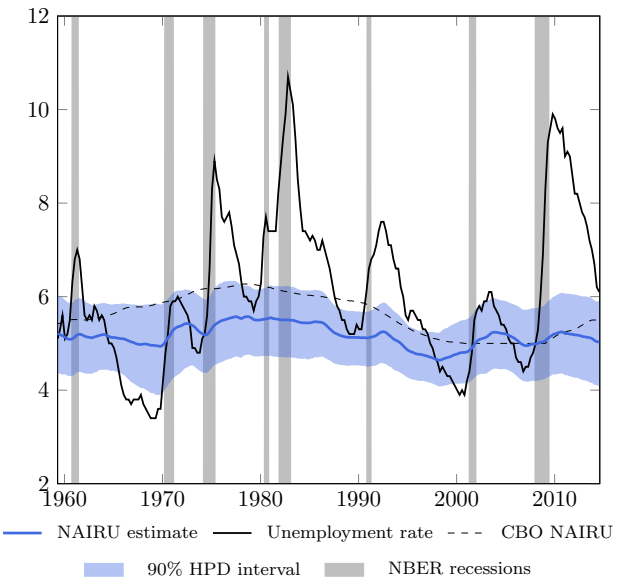
(d) Stochastic volatility of output gap

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 7, 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

Figure 9: NAIRU (parsimonious model)



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 sensitivty 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 NKPC by a backward-looking Phillips curve in order to see if our findings regarding the inflation dynamics and the stability of β^π depend on the forward-looking specification of the Phillips curve. Second, we use the unemployment gap, instead of the output gap, as a measure of real activity in the Phillips curve.

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

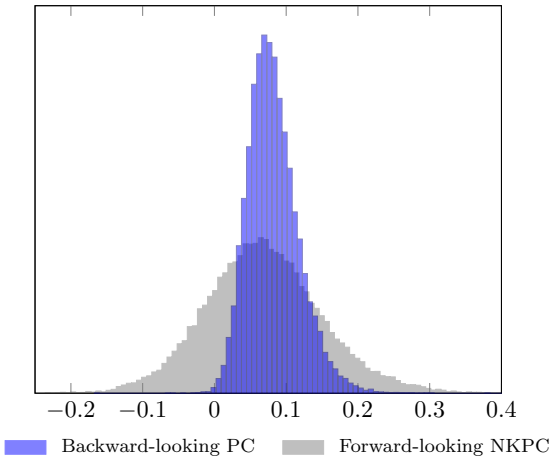
$$\pi_t = \sum_{p=1}^{4} b_p \pi_{t-p} + \beta_t^\pi \widetilde{y}_t^c + \varepsilon_t^\pi, \tag{32}$$

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 ζ. This specification matches standard backward-looking Phillips curve models in the literature.[9] Table 7 gives the posterior probability of time variation in the slope of the Phillips curve for both the baseline and the backward-looking model. The Phillips curve is found to be stable in both specifications. Moreover, the finding is robust to different prior distributions for the degree of time variation. In all settings, the posterior probability of a time-varying Phillips

[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).

Figure 10: Posterior distributions of β_0^π: Backward and Forward-looking



curve are in the area of 1% or below. However, the estimated slope coefficients differ between the forward and the backward-looking model. Figure 10 plots the posterior distributions of the time-invariant slope parameter for both models. 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 bigger and more significant Phillips curve slope in backward compared to purely forward-looking models.

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 6 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.[10]

[10]Estimates for the output and unemployment components do not change notably, but are available on request.

Table 6: Robustness of parameter estimates to different inflation measures

	Posterior parameter distribution			Posterior inclusion probability		
	Slope coefficient β_0^π			Phillips curve	SV trend inflation	SV temp. inflation
	median	2.5%	97.5%	δ_π	θ_π	θ_ζ
CPI	0.110	-0.080	0.303	0.0070	0.0303	1.0000
CPI excl. F&E	0.094	-0.032	0.245	0.0080	0.0430	1.0000
PCE	0.110	-0.043	0.272	0.0026	0.0378	1.0000
PCE excl. F&E	0.067	-0.061	0.196	0.0026	0.0828	1.0000
GDP deflator	0.076	-0.072	0.235	0.0024	0.0476	1.0000

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 by the unemployment gap in the Phillips curve, the impact of unemployment on inflation is time-varying. 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, independently on weather 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 by the unemployment gap.

Table 7: Stability of Phillips curve for different models

Prior		Posterior probability of $\delta_\pi = 1$	
p_0	A_0	New-Keynesian	Backward-looking
0.5	0.1	0.0110	0.0040
0.5	1	0.0026	0.0012
0.5	10	0.0007	0.0008
0.5	100	0.0000	0.0004

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, 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 time-invariant. In a first estimation step, a stochastic model selection procedure is employed to test which parameters are time-varying. We find that potential output growth, Okun's Law coefficient, the variance of innovations to the output gap and to a persistent inflation gap component are time-varying while the slope of the Phillips curve and the variances of innovations to all trend components are time-invariant. Our estimation result show a clear decrease in potential output growth, which can be characterized by substantial drops in the 1970s and the 2000s. 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 sensitive in a recovery. With regard to the dynamics of inflation, we find that the inflation gap and idiosyncratic shocks are the major determinants of inflation changes. However, the inflation gap is not very sensitive to the output gap but is driven by a persistent AR(1) component. The latter component exhibits stochastic volatility and mainly captures the large and persistent swings in inflation during the inflationary period in the 1970s and 1980s.

6 Acknowledgements

The authors would like to thank Freddy Heylen, Thomas Kneib, James Morley and Joris Wauters for their constructive comments and suggestions. The paper has also benefited from discussions with participants at the 22nd and 23rd Symposium of the Society for Nonlinear Dynamics and Econometrics and various seminars and workshops.

References

Atkeson, Andrew, Ohanian, Lee E. 2001. Are phillips curves useful for forecasting inflation? *Quarterly Review* **Winter**: 2–11.

Ball, Laurence 1994. Credible Disinflation with Staggered Price-Setting. *American Economic Review* **84**(1): 282–89.

Ball, Laurence M., Leigh, Daniel, Loungani, Prakash 2013. Okun's law: Fit at fifty? NBER Working Papers 18668.

Ball, Laurence, Mazumder, Sandeep 2011. Inflation dynamics and the great recession. *Brookings Papers on Economic Activity* **42**(1): 337–405.

Basistha, Arabinda, Startz, Richard 2008. Measuring the nairu with reduced uncertainty: A multiple-indicator common-cycle approach. *The Review of Economics and Statistics* **90**(4): 805–811.

Beveridge, Stephen, Nelson, Charles R. 1981. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics* **7**(2): 151–174.

Blanchard, Olivier J., Summers, Lawrence H. 1986. Hysteresis and the european unemployment problem. In NBER Macroeconomics Annual 1986, Volume 1 NBER Chapters. National Bureau of Economic Research, Inc pp. 15–90.

Calvo, Guillermo A. 1983. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics* **12**(3): 383–398.

Campbell, Jeffrey R., Fisher, Jonas D. M. 2000. Aggregate employment fluctuations with microeconomic asymmetries. *The American Economic Review* **90**(5): pp. 1323–1345.

Chan, Joshua C C, Koop, Gary, Potter, Simon M 2012. A bounded model of time variation in trend inflation, nairu and the phillips curve. ANU Working Papers in Economics and Econometrics 2012-590, Australian National University, College of Business and Economics, School of Economics.

Cogley, Timothy, Sbordone, Argia M. 2008. Trend inflation, indexation, and inflation persistence in the new keynesian phillips curve. *American Economic Review* **98**(5): 2101–26.

Daly, Mary C., Hobijn, Bart, Şahin, Ayşegul, Valletta, Robert G. 2012. A search and matching approach to labor markets: Did the natural rate of unemployment rise? *Journal of Economic Perspectives* **26**(3): 3–26.

Daly, Mary, Hobijn, Bart 2010. Okun's law and the unemployment surprise of 2009. *FRBSF Economic Letter.*

Del Negro, M., Primiceri, G. 2014. Time varying structural vector autoregressions and monetary policy: a corrigendum. Staff Reports 619, Federal Reserve Bank of New York.

Fernández-Villaverde, Jesús, Rubio-Ramírez, Juan 2010. Macroeconomics and Volatility: Data, Models, and Estimation. NBER Working Papers 16618, December.

Fleischman, Charles A., Roberts, John M. 2011. From many series, one cycle: improved estimates of the business cycle from a multivariate unobserved components model. Technical Report 2011-46, Board of Governors of the Federal Reserve System (U.S.).

Frühwirth-Schnatter, Sylvia, Wagner, Helga 2010. Stochastic model specification search for gaussian and partial non-gaussian state space models. *Journal of Econometrics* **154**(1): 85–100.

Fuhrer, Jeff, Moore, George 1995. Inflation persistence. *The Quarterly Journal of Economics* **110**(1): 127–59.

George, E.I., McCulloch, R. 1993. Variable selection via gibbs sampling. *Journal of the American Statistical Association* **88**: 881–889.

Goodfriend, Marvin, King, Robert G. 2005. The incredible Volcker disinflation. *Journal of Monetary Economics* **52**(5): 981–1015.

Goodfriend, Marvin, King, Robert G. 2012. The great inflation drift. In *The Great Inflation: The Rebirth of Modern Central Banking,* ed. Michael D. Bordo and Athanasios Orphanides. University of Chicago Press pp. 181–216.

Hall, Robert E 2011. The long slump. Technical Report, National Bureau of Economic Research.

Hamilton, James D. 2008. Macroeconomics and arch. NBER Working Papers 14151, June.

Kim, Chang-Jin, Manopimoke, Pym, Nelson, Charles R 2014. Trend inflation and the nature of structural breaks in the New Keynesian Phillips curve. *Journal of money, credit and banking* **46**: 253–266.

Kim, Sangjoon, Shephard, Neil, Chib, Siddhartha 1998. Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *Review of Economic Studies* **65**(3): 361–93.

Knotek, Edward S. 2007. How useful is okun's law? Technical Report Q IV, Federal Reserve Bank of Kansas City.

Kuttner, Kenneth N. 1994. Estimating potential output as a latent variable. *Journal of Business & Economic Statistics* **12**(3): pp. 361–368.

Laubach, Thomas 2001. Measuring the nairu: Evidence from seven economies. *The Review of Economics and Statistics* **83**(2): pp. 218–231.

Lee, Jim 2000. The robustness of okun's law: Evidence from oecd countries. *Journal of Macroeconomics* **22**(2): 331–356.

Lindbeck, Assar, Snower, Dennis J 1988. Cooperation, Harassment, and Involuntary Unemployment: An Insider-Outsider Approach. *American Economic Review* **78**(1): 167–88.

Liu, Yuelin, Morley, James 2014. Structural evolution of the postwar U.S. economy. *Journal of Economic Dynamics and Control* **42**(C): 50–68.

Luo, Sui, Startz, Richard 2014. Is it one break or ongoing permanent shocks that explains U.S. real GDP? *Journal of Monetary Economics* **66**(C): 155–163.

Mankiw, N. Gregory 2001. The inexorable and mysterious tradeoff between inflation and unemployment. *The Economic Journal* **111**(471): pp. C45–C61.

Mavroeidis, Sophocles, Plagborg-Møller, Mikkel, Stock, James H. 2014. Empirical evidence on inflation expectations in the new keynesian phillips curve. *Journal of Economic Literature* **52**(1): 124–88.

Mishkin, Frederic S. 2007. Inflation dynamics. *International Finance* **10**(3): 317–334.

Mitra, Sinchan, Sinclair, Tara M. 2012. Output Fluctuations In The G-7: An Unobserved Components Approach. *Macroeconomic Dynamics* **16**(03): 396–422.

Morley, James C, Nelson, Charles R, Zivot, Eric 2003. Why are the beveridge-nelson and unobserved-components decompositions of gdp so different? *The Review of Economics and Statistics* **85**(2): 235–243.

Morley, James, Piger, Jeremy M., Rasche, Robert H. 2013. Inflation in the g7: Mind the gap(s)? *Macroeconomic Dynamics* pp. 1–30.

Omori, Yasuhiro, Chib, Siddhartha, Shephard, Neil, Nakajima, Jouchi 2007. Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics* **140**(2): 425–449.

Owyang, Michael T., Sekhposyan, Tatevik 2012. Okun's Law over the Business Cycle: Was the Great Recession All That Different? *Federal Reserve Bank of St. Louis Review* (September): 399–418.

Palley, Thomas I 1993. Okun's law and the asymmetric and changing cyclical behaviour of the usa economy. *International Review of Applied Economics* **7**(2): 144–62.

Perron, Pierre, Wada, Tatsuma 2009. Let's take a break: Trends and cycles in US real GDP. *Journal of Monetary Economics* **56**(6): 749–765.

Primiceri, Giorgio E. 2005. Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies* **72**(3): 821–852.

Reifschneider, D., Wascher, W.L., Wilcox, D. 2013. Aggregate supply in the united states: Recent developments and implications for the conduct of monetary policy. Finance and Economics Discussion Series 2013-77, Board of Governors of the Federal Reserve System (U.S.).

Roberts, John M. 2006. Monetary policy and inflation dynamics. *International Journal of Central Banking.*

Rudd, Jeremy, Whelan, Karl 2005. New tests of the new-keynesian phillips curve. *Journal of Monetary Economics* **52**(6): 1167–1181.

Rudd, Jeremy, Whelan, Karl 2007. Modeling inflation dynamics: A critical review of recent research. *Journal of Money, Credit and Banking* **39**(s1): 155–170.

Scott, James G., Berger, James O. 2010. Bayes and empirical-bayes multiplicity adjustment in the variable-selection problem. *Ann. Statist.* **38**(5): 2587–2619.

Shephard 1994. Partial non-Gaussian state space. *Biometrika* **81**: 115–131.

Silvapulle, Paramsothy, Moosa, Imad, Silvapulle, Mervyn 2004. Asymmetry in Okun's law. *Canadian Journal of Economics* **37**(2): 353–374.

Sinclair, Tara M. 2009. The relationships between permanent and transitory movements in u.s. output and the unemployment rate. *Journal of Money, Credit and Banking* **41**(2-3): 529–542.

Staiger, Douglas, Stock, James H, Watson, Mark W 1997. The nairu, unemployment and monetary policy. *Journal of Economic Perspectives* **11**(1): 33–49.

Stella, Andrea, Stock, James H. 2012. A state-dependent model for inflation forecasting. International Finance Discussion Papers 2012-1062, Board of Governors of the Federal Reserve System (U.S.).

Stock, James H, Watson, Mark W 2007. Why has us inflation become harder to forecast? *Journal of Money, Credit and banking* **39**(s1): 3–33.

Stock, James H., Watson, Mark W. 2010. Modeling inflation after the crisis. NBER Working Papers 16488, National Bureau of Economic Research, Inc, October.

Woodford, Michael 2008. How important is money in the conduct of monetary policy? *Journal of money, credit and banking* **40**: 1561–1598.

Appendix A Gibbs sampling algorithm

In this appendix we provide details on the Gibbs sampling algorithm used in Subsection 2.3 to jointly sample the binary indicators $\mathcal{M}$, the hyperparameters ϕ, the trend and temporary components α, the time-varying parameters β, the mixture indicators ι 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 $\mathcal{M}$ and the parameters ϕ

For notational convenience, let us define a general regression model

$$w = z^{\mathcal{M}} b^{\mathcal{M}} + e, \qquad e \sim \mathcal{N}(0, \Sigma), \tag{A-1}$$

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 α_t, β_t and h_t that are relevant for explaining w_t. The corresponding unrestricted parameter vector with the relevant elements from ϕ is denoted b. $z^{\mathcal{M}}$ and $b^{\mathcal{M}}$ are then the restricted predictor matrix and restricted parameter vector that exclude those elements in z and b for which the corresponding indicator in $\mathcal{M}$ is 0. Furthermore, Σ 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 $\mathcal{M}$ from $f(\mathcal{M}|\alpha, \beta, h, \phi, w)$ and ϕ from $f(\phi|\alpha, \beta, h, \mathcal{M}, w)$. However, this approach does not result in an irreducible Markov chain as whenever an indicator in $\mathcal{M}$ equals zero, the corresponding coefficient in ϕ is also zero which implies that the chain has absorbing states. Therefore, as in Frühwirth-Schnatter and Wagner (2010) we marginalize over the parameters ϕ when sampling $\mathcal{M}$ and next draw the parameters ϕ conditional on the indicators $\mathcal{M}$. The posterior distribution $f(\mathcal{M}|\alpha, \beta, h, w)$ can be obtained using Bayes' Theorem as

$$f(\mathcal{M}|\alpha, \beta, h, w) \propto f(w|\mathcal{M}, \alpha, \beta, h)\, p(\mathcal{M}), \tag{A-2}$$

with $p(\mathcal{M})$ being the prior probability of $\mathcal{M}$ and $f(w|\mathcal{M}, \alpha, \beta, h)$ being the marginal likelihood of the regression model (A-1) where the effect of the parameters $b^{\mathcal{M}}$ and σ^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

$$b^{\mathcal{M}} \sim \mathcal{N}\left(a_0^{\mathcal{M}}, A_0^{\mathcal{M}} \sigma^2_e\right), \qquad \sigma^2_e \sim IG(c_0, C_0), \tag{A-3}$$

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

$$f(w|\mathcal{M}, \alpha, \beta, h) \propto \frac{\left|A_T^{\mathcal{M}}\right|^{0.5}}{\left|A_0^{\mathcal{M}}\right|^{0.5}} \frac{\Gamma(c_T)\, C_0^{c_0}}{\Gamma(c_0)\left(C_T^{\mathcal{M}}\right)^{c_T}}, \tag{A-4}$$

and the posterior moments $a_T^{\mathcal{M}}$, $A_T^{\mathcal{M}}$, c_T and $C_T^{\mathcal{M}}$ of $b^{\mathcal{M}}$and σ_e^2 can be calculated as

$$a_T^{\mathcal{M}} = A_T^{\mathcal{M}} \left(\left(z^{\mathcal{M}}\right)' w + \left(A_0^{\mathcal{M}}\right)^{-1} a_0^{\mathcal{M}} \right), \tag{A-5}$$

$$A_T^{\mathcal{M}} = \left(\left(z^{\mathcal{M}}\right)' z^{\mathcal{M}} + \left(A_0^{\mathcal{M}}\right)^{-1} \right)^{-1}, \tag{A-6}$$

$$c_T = c_0 + T/2\,, \tag{A-7}$$

$$C_T^{\mathcal{M}} = C_0 + 0.5 \left(w'w + \left(a_0^{\mathcal{M}}\right)' \left(A_0^{\mathcal{M}}\right)^{-1} a_0^{\mathcal{M}} - \left(a_T^{\mathcal{M}}\right)' \left(A_T^{\mathcal{M}}\right)^{-1} a_T^{\mathcal{M}} \right). \tag{A-8}$$

- In the heteroskedastic case $\Sigma = \text{diag}\left(\sigma_{e,1}^2, ..., \sigma_{e,T}^2\right)$, under the normal conjugate prior $b^{\mathcal{M}} \sim \mathcal{N}\left(a_0^{\mathcal{M}}, A_0^{\mathcal{M}}\right)$ the closed form solution for the marginal likelihood $f\left(w \,|\mathcal{M}, \alpha, \beta, h\right)$ is

$$f\left(w \,|\mathcal{M}, \alpha, \beta, h\right) \propto \frac{|\Sigma|^{-0.5} \left|A_T^{\mathcal{M}}\right|^{0.5}}{\left|A_0^{\mathcal{M}}\right|^{0.5}} \exp\left(-\frac{1}{2} \left(w'\Sigma^{-1}w + \left(a_0^{\mathcal{M}}\right)' \left(A_0^{\mathcal{M}}\right)^{-1} a_0^{\mathcal{M}} - \left(a_T^{\mathcal{M}}\right)' \left(A_T^{\mathcal{M}}\right)^{-1} a_T^{\mathcal{M}} \right) \right), \tag{A-9}$$

with

$$a_T^{\mathcal{M}} = A_T^{\mathcal{M}} \left(\left(z^{\mathcal{M}}\right)' \Sigma^{-1} w + \left(A_0^{\mathcal{M}}\right)^{-1} a_0^{\mathcal{M}} \right), \tag{A-10}$$

$$A_T^{\mathcal{M}} = \left(\left(z^{\mathcal{M}}\right)' \Sigma^{-1} z^{\mathcal{M}} + \left(A_0^{\mathcal{M}}\right)^{-1} \right)^{-1}. \tag{A-11}$$

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

Block 1(a): Sampling the binary indicators δ and the parameters β, σ_η and σ_ε^2

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 $\sigma_\varepsilon^2 = \left(\sigma_{\varepsilon,y}^2, \sigma_{\varepsilon,\pi}^2, \sigma_{\varepsilon,u}^2\right)$ conditional on the states α, β and h. First, as there is no binary indicator in equation (1), $\sigma_{\varepsilon,y}^2$ can be sampled directly from $IG\left(c_T, C_T\right)$ with c_T as in equation (A-7) and $C_T = C_0 + 0.5(\varepsilon^{y'}\varepsilon^y)$ with ε^y calculated from $\varepsilon_t^y = y_t - y_t^\tau - y_t^c$.

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

$$\overbrace{\pi_t - \pi_t^\tau - \zeta_t}^{w_t} = \overbrace{\begin{bmatrix} \widetilde{y}_t^c & \delta_\pi \widetilde{\beta}_t^\pi \widetilde{y}_t^c \end{bmatrix}}^{z_t^{\mathcal{M}}} \overbrace{\begin{bmatrix} \beta_0^\pi \\ \sigma_{\eta,\pi} \end{bmatrix}}^{b^{\mathcal{M}}} + \overbrace{\varepsilon_t^\pi}^{e_t}, \tag{A-12}$$

$$\underbrace{u_t - u_t^\tau}_{w_t} = \underbrace{\begin{bmatrix} y_t^c & \delta_u \widetilde{\beta}_t^u y_t^c \end{bmatrix}}_{z_t^{\mathcal{M}}} \underbrace{\begin{bmatrix} \beta_0^u \\ \sigma_{\eta,u} \end{bmatrix}}_{b^{\mathcal{M}}} + \underbrace{\varepsilon_t^u}_{e_t}, \tag{A-13}$$

where in both the restricted vector $z_t^{\mathcal{M}}$ and the restricted parameter vector $b^{\mathcal{M}}$ the second term is excluded when $\delta_j = 0$ (for $j = \pi, u$). Note that, next to the parameters in $b^{\mathcal{M}}$ and σ_e^2, each of the specifications (A-12) and (A-13) depends only on the data w_t, on some of the states in α_t and β_t and on δ_j. As such, we can simplify the specification of the posterior from $f\left(\delta_j|\delta_{\setminus j}, \theta, \lambda, \alpha, \beta, h, x\right)$ to $f\left(\delta_j\,|\alpha, \beta, w\right)$ for which we have $f\left(\delta_j\,|\alpha, \beta, w\right) \propto f\left(w\,|\delta_j, \alpha, \beta\right) p\left(\delta_j\right)$. As the error terms ε_t^{π} in the inflation equation and ε_t^{u} in the employment equation are homoskedastic, we have $\Sigma = \sigma_e^2 I_T$ in the general notation of equation (A-1) such that the marginal likelihood $f\left(w\,|\delta_j, \alpha, \beta\right)$ can be calculated as in equation (A-4). The binary indicator δ_j can then be sampled from the Bernoulli distribution with probability

$$p\left(\delta_j = 1\,|\alpha, \beta, w\right) = \frac{f\left(\delta_j = 1\,|\alpha, \beta, w\right)}{f\left(\delta_j = 0\,|\alpha, \beta, w\right) + f\left(\delta_j = 1\,|\alpha, \beta, w\right)}, \tag{A-14}$$

while $\sigma_{\varepsilon,j}^2$ can be sampled from $IG\left(c_T, C_T^{\mathcal{M}}\right)$ and, conditionally on $\sigma_{\varepsilon,j}^2$, $b^{\mathcal{M}}$ from $\mathcal{N}\left(a_T^{\mathcal{M}}, A_T^{\mathcal{M}}\sigma_{\varepsilon,j}^2\right)$, for $j = \pi, u$ and with $a_T^{\mathcal{M}}$, $A_T^{\mathcal{M}}$, c_T and $C_T^{\mathcal{M}}$ as defined in equations (A-5)-(A-8). Note that $b^{\mathcal{M}} = \left(\beta_0^j, \sigma_{\eta,j}\right)'$ when $\delta_j = 1$ and $b^{\mathcal{M}} = \beta_0^j$ 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 θ and the parameters h_0 and σ_γ

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

$$\overbrace{g_t^k - \left(m_{\iota_t^k} - 1{,}2704\right)}^{w_t} = \overbrace{2\begin{bmatrix} 1 & \theta_k \widetilde{h}_t^k \end{bmatrix}}^{z_t^{\mathcal{M}}} \overbrace{\begin{bmatrix} h_0^k \\ \sigma_{\gamma,k} \end{bmatrix}}^{b^{\mathcal{M}}} + \overbrace{\widetilde{\epsilon}_t^k}^{e_t}, \tag{A-15}$$

for $k = y, \pi, u, c, \zeta$, with $\widetilde{\epsilon}_t^k = \epsilon_t^k - \left(m_{\iota_t^k} - 1{,}2704\right)$ is ϵ_t^k recentered around zero and where using equations (2), (9), (16), (4) and (11), $g_t^k = \ln\left(\left(\exp\{h_t^k\}\psi_t^k\right)^2 + .001\right)$ can be calculated as

$$g_t^y = \ln\left(\left(y_t^\tau - y_{t-1}^\tau - \kappa_t\right)^2 + .001\right), \tag{A-16}$$

$$g_t^\pi = \ln\left(\left(\pi_t^\tau - \pi_{t-1}^\tau\right)^2 + .001\right), \tag{A-17}$$

$$g_t^u = \ln\left(\left(u_t^\tau - u_{t-1}^\tau\right)^2 + .001\right), \tag{A-18}$$

$$g_t^c = \ln\left(\left(y_t^c - \rho_1 y_{t-1}^c - \rho_2 y_{t-2}^c\right)^2 + .001\right), \tag{A-19}$$

$$g_t^\zeta = \ln\left(\left(\zeta_t - \varrho\zeta_{t-1}\right)^2 + .001\right). \tag{A-20}$$

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

of ϵ_t^k defined in equation (22), the error term $\widetilde{\epsilon}_t^k$ in equation (A-15) has a heteroskedastic variance $v^2_{\iota_t^k}$ such that $\Sigma = \text{diag}\left(v^2_{\iota_1^k}, ..., v^2_{\iota_T^k}\right)$ in the general notation of equation (A-1). In this case, the marginal likelihood $f\left(w\,|\theta_k, h\right)$ can be calculated as in equation (A-9). The binary indicator θ_k can then be sampled from the Bernoulli distribution with probability $p\left(\theta_k = 1\,|h, w\right)$ calculated from an equation similar to (A-14). Next, $b^{\mathcal{M}}$ can be sampled from $\mathcal{N}\left(a_T^{\mathcal{M}}, A_T^{\mathcal{M}}\right)$ for $k = y, \pi, u, c, \zeta$ and with $a_T^{\mathcal{M}}$ and $A_T^{\mathcal{M}}$ as defined in equations (A-10) and (A-11). Note that $b^{\mathcal{M}} = \left(h_0^k, \sigma_{\gamma,k}\right)'$ when $\theta_k = 1$ and $b^{\mathcal{M}} = h_0^k$ when $\theta_k = 0$. In the latter case, we set $\sigma_{\gamma,k} = 0$.

Block 1(c): Sampling the binary indicator λ and the parameters κ_0 and σ_κ

In this block we sample the binary indicator λ and the parameters κ_0 and σ_κ conditional on the states α, β and h. Using equation (31), equation (2) can be rewritten in the general linear regression format of (A-1) as

$$\overbrace{y_t^\tau - y_{t-1}^\tau}^{w_t} = \overbrace{\begin{bmatrix} 1 & \lambda\widetilde{\kappa}_t \end{bmatrix}}^{z_t^{\mathcal{M}}} \overbrace{\begin{bmatrix} \kappa_0 \\ \sigma_\kappa \end{bmatrix}}^{b^{\mathcal{M}}} + \overbrace{\exp\{h_t^y\}\psi_t^y}^{e_t}, \tag{A-21}$$

with $\Sigma = \text{diag}\left(\exp\{h_1^y\}^2, \ldots, \exp\{h_T^y\}^2\right)$. The indicator λ can then be sampled from the posterior distribution $f\left(\lambda\,|\alpha, w\right) \propto f\left(w\,|\lambda, \alpha\right) p\left(\lambda\right)$ with the marginal likelihood $f\left(w\,|\lambda, \alpha\right)$ calculated from equation (A-9). Next, $b^{\mathcal{M}}$ can be sampled from $\mathcal{N}\left(a_T^{\mathcal{M}}, A_T^{\mathcal{M}}\right)$ with $a_T^{\mathcal{M}}$ and $A_T^{\mathcal{M}}$ as defined in equations (A-10) and (A-11). Note that $b^{\mathcal{M}} = \left(\kappa_0, \sigma_\kappa\right)'$ when $\lambda = 1$ and $b^{\mathcal{M}} = \kappa_0$ when $\lambda = 0$. In the latter case, we set $\sigma_\lambda = 0$.

Block 1(d): Sampling the parameters ρ and ϱ

For filtering $\rho = (\rho_1, \rho_2)$ conditional on the states α, β and h, equation (4) can be written in the general notation of equation (A-1) as: $w_t = y_t^c$, $z_t = \left(y_{t-1}^c, y_{t-2}^c\right)$, $b = (\rho_1, \rho_2)'$ and $e_t = \exp\{h_t^c\}\psi_t^c$, such that $\Sigma = \text{diag}\left(\exp\{h_1^c\}^2, \ldots, \exp\{h_T^c\}^2\right)$. Under the normal prior distribution $\mathcal{N}\left(a_0, A_0\right)$, ρ can then be sampled from the posterior $\mathcal{N}\left(a_T, A_T\right)$ with a_T and A_T as in equations (A-10) and (A-11).

Likewise, for filtering ϱ conditional on the states α, β 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 $e_t = \exp\{h_t^\zeta\}\psi_t^\zeta$, such that $\Sigma = \text{diag}\left(\exp\{h_1^\zeta\}^2, \ldots, \exp\{h_T^\zeta\}^2\right)$. Under the normal prior distribution $\mathcal{N}\left(a_0, A_0\right)$, ϱ can again be sampled from the posterior $\mathcal{N}\left(a_T, A_T\right)$ with a_T and A_T as in equations (A-10) and (A-11).

Block 2: Sampling the state vectors α, β and h and mixture indicators ι

In this block we use a forward-filtering and backward-sampling approach for the states α, β and h based on a general state space model of the form

$$w_t = Z_t^{\mathcal{M}} s_t^{\mathcal{M}} + e_t, \qquad e_t \sim iid\mathcal{N}\left(0, H_t\right), \tag{A-22}$$

$$s_{t+1} = R_0 + R_1 s_t + K_t v_t, \qquad v_t \sim iid\mathcal{N}\left(0, Q_t\right), \qquad s_1 \sim iid\mathcal{N}\left(a_1, A_1\right), \tag{A-23}$$

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 $s_t^{\mathcal{M}}$ and the matrix $Z_t^{\mathcal{M}}$ are again restricted versions of s_t and Z_t with the elements excluded depending on the model indicators $\mathcal{M}$. 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 Gibbs sampler of Shephard (1994).

Block 2(a) Sampling the trend and temporary components α

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 β, the stochastic volatilities h and the hyperparameters ϕ. 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

$$\overbrace{\begin{bmatrix} y_t \\ \pi_t \\ u_t \end{bmatrix}}^{w_t} = \overbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & a\beta_t^\pi & b\beta_t^\pi \\ 0 & 0 & 1 & 0 & 0 & \beta_t^u & 0 \end{bmatrix}}^{Z_t^{\mathcal{M}}} \overbrace{\begin{bmatrix} y_t^\tau \\ \pi_t^\tau \\ u_t^\tau \\ \widetilde{\kappa}_t \\ \zeta_t \\ y_t^c \\ y_{t-1}^c \end{bmatrix}}^{s_t^{\mathcal{M}}} + \overbrace{\begin{bmatrix} \varepsilon_t^y \\ \varepsilon_t^\pi \\ \varepsilon_t^u \end{bmatrix}}^{e_t}, \tag{A-24}$$

$$\underbrace{\begin{bmatrix} y_{t+1}^\tau \\ \pi_{t+1}^\tau \\ u_{t+1}^\tau \\ \widetilde{\kappa}_{t+1} \\ \zeta_{t+1} \\ y_{t+1}^c \\ y_t^c \end{bmatrix}}_{s_{t+1}} = \underbrace{\begin{bmatrix} \kappa_0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{R_0} + \underbrace{\begin{bmatrix} 1 & 0 & 0 & \sigma_\kappa & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \varrho & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho_1 & \rho_2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}}_{R_1} \underbrace{\begin{bmatrix} y_t^\tau \\ \pi_t^\tau \\ u_t^\tau \\ \widetilde{\kappa}_t \\ \zeta_t \\ y_t^c \\ y_{t-1}^c \end{bmatrix}}_{s_t} + K_t \underbrace{\begin{bmatrix} \psi_t^y \\ \psi_t^\pi \\ \psi_t^u \\ \psi_t^\kappa \\ \psi_t^\zeta \\ \psi_t^c \end{bmatrix}}_{\nu_t}, \tag{A-25}$$

$$\text{with} \quad K_t = \begin{bmatrix} \exp\{h_t^y\} & 0 & 0 & 0 & 0 & 0 \\ 0 & \exp\{h_t^\pi\} & 0 & 0 & 0 & 0 \\ 0 & 0 & \exp\{h_t^u\} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\psi,\kappa} & 0 & 0 \\ 0 & 0 & 0 & 0 & \exp\{h_t^\zeta\} & 0 \\ 0 & 0 & 0 & 0 & 0 & \exp\{h_t^c\} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and where $a = \left(1 - \omega\rho_1 - \omega^2\rho_2\right)^{-1}$, $b = a\omega\rho_2$, $H_t = diag\left(\sigma^2_{\varepsilon,y}, \sigma^2_{\varepsilon,\pi}, \sigma^2_{\varepsilon,u}\right)$ and $Q_t = I_6$. The random walk components y^τ_t, π^τ_t, u^τ_t and $\widetilde{\kappa}_t$ are initialized by setting $a_1 = 0$ and $A_1 = 1000$ for each of these components while the stationary components ζ_t and y^c_t are initialized from their unconditional distributions. Note that using κ_0, σ_κ and $\widetilde{\kappa}_t$, κ_t can easily be reconstructed from equation (27).

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

Block 2(b): Sampling the time-varying parameters β

We next filter and draw the time-varying parameters $\beta = (\beta^\pi, \beta^u)$ conditionally on the state vector α, the stochastic volatilities h, the hyperparameters ϕ and the binary indicators $\mathcal{M}$. 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 $\widetilde{\beta}^\pi_t$ and $\widetilde{\beta}^u_t$ are given by

$$\overbrace{\begin{bmatrix} \pi_t - \pi^\tau_t - \zeta_t - \beta^\pi_0 \widetilde{y}^c_t \end{bmatrix}}^{w_t} = \overbrace{\begin{bmatrix} \sigma_{\eta,\pi}\widetilde{y}^c_t \end{bmatrix}}^{Z^{\mathcal{M}}_t} \overbrace{\begin{bmatrix} \widetilde{\beta}^\pi_t \end{bmatrix}}^{s^{\mathcal{M}}_t} + \overbrace{\begin{bmatrix} \varepsilon^\pi_t \end{bmatrix}}^{e_t}, \tag{A-26}$$

$$\underbrace{\begin{bmatrix} \widetilde{\beta}^\pi_{t+1} \end{bmatrix}}_{s_{t+1}} = \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_{R_1} \underbrace{\begin{bmatrix} \widetilde{\beta}^\pi_t \end{bmatrix}}_{s_t} + \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_{K_t} \underbrace{\begin{bmatrix} \widetilde{\eta}^\pi_t \end{bmatrix}}_{\nu_t}, \tag{A-27}$$

with $H_t = \sigma^2_{\varepsilon,\pi}$ and $Q_t = 1$, and

$$\overbrace{\begin{bmatrix} u_t - u^\tau_t - \beta^u_0 y^c_t \end{bmatrix}}^{w_t} = \overbrace{\begin{bmatrix} \sigma_{\eta,u} y^c_t \end{bmatrix}}^{Z^{\mathcal{M}}_t} \overbrace{\begin{bmatrix} \widetilde{\beta}^u_t \end{bmatrix}}^{s^{\mathcal{M}}_t} + \overbrace{\begin{bmatrix} \varepsilon^u_t \end{bmatrix}}^{e_t}, \tag{A-28}$$

$$\underbrace{\begin{bmatrix} \widetilde{\beta}^u_{t+1} \end{bmatrix}}_{s_{t+1}} = \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_{R_1} \underbrace{\begin{bmatrix} \widetilde{\beta}^u_t \end{bmatrix}}_{s_t} + \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_{K_t} \underbrace{\begin{bmatrix} \widetilde{\eta}^u_t \end{bmatrix}}_{\nu_t}, \tag{A-29}$$

with $H_t = \sigma^2_{\varepsilon,u}$ and $Q = 1$. Both random walk components $\widetilde{\beta}^\pi_t$ and $\widetilde{\beta}^u_t$ are initialized by setting $a_1 = 0$ and $A_1 = 1000$.

In the restricted model (i.e. $\delta_j = 0$), $Z^{\mathcal{M}}$ and $s^{\mathcal{M}}$ are empty. In this case, no forward-filtering and backward-sampling is needed and $\widetilde{\beta}^j_t$ can be sampled directly from its prior using equation (24). Note that the sampling of the state vector α in block 2(a) depends on β^j_t rather than on $\widetilde{\beta}^j_t$. Using β^j_0, $\sigma_{\eta,j}$ and $\widetilde{\beta}^j_t$, β^j_t can easily be reconstructed from equation (23).

Block 2(c): Sampling the mixture indicators ι and the stochastic volatilities h

In this block we draw the mixture indicators $\iota = \left(\iota^y, \iota^\pi, \iota^u, \iota^c, \iota^\zeta\right)$ and the stochastic volatilities $h = \left(h^y, h^\pi, h^u, h^c, h^\zeta\right)$ conditionally on the state vector α, the time-varying parameters β, the hyperparameters ϕ and the binary indicators $\mathcal{M}$. Following Del Negro and Primiceri (2014), in this block we first sample the mixture indicator ι^k_t (for $k = y, \pi, u, c, \zeta$) from its conditional probability

mass

$$p\left(\iota_t^k = i | h_t^k, \epsilon_t^k\right) \propto q_i f_{\mathcal{N}}\left(\epsilon_t^k | 2h_t^k + m_i - 1.2704, \nu_i^2\right), \tag{A-30}$$

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 $\widetilde{h}_t^k$ (for $k = y, \pi, u, c, \zeta$) conditioning on the transformed states g_t^k defined in equations (A-16)-(A-20), on the mixture indicators ι_t^k and on the parameters ϕ. More specifically, the unrestricted (i.e. $\theta_k = 1$) conditional state space representation is given by

$$\overbrace{\left[\ g_t^k - \left(m_{\iota_t^k} - 1,2704\right) - 2h_0^k\ \right]}^{w_t} = \overbrace{\left[\ 2\theta^k \sigma_{\gamma,k}\ \right]}^{Z_t^{\mathcal{M}}} \overbrace{\left[\ \widetilde{h}_t^k\ \right]}^{s_t^{\mathcal{M}}} + \overbrace{\left[\ \widetilde{\epsilon}_t^k\ \right]}^{e_t}, \tag{A-31}$$

$$\underbrace{\left[\ \widetilde{h}_{t+1}^k\ \right]}_{s_{t+1}} = \underbrace{[1]}_{R_1} \underbrace{\left[\ \widetilde{h}_t^k\ \right]}_{s_t} + \underbrace{\left[\ 1\ \right]}_{K_t} \underbrace{\left[\ \widetilde{\gamma}_t^k\ \right]}_{\nu_t}, \tag{A-32}$$

with $H_t = v_{\iota_t^k}^2$, $Q_t = 1$ and where $\widetilde{\epsilon}_t^k = \epsilon_t^k - \left(m_{\iota_t^k} - 1,2704\right)$ is ϵ_t^k recentered around zero. The random walk components $\widetilde{h}_t^k$ are initialized by setting $a_1 = 0$ and $A_1 = 1000$.

In the restricted model (i.e. $\theta_k = 0$), $Z^{\mathcal{M}}$ and $s^{\mathcal{M}}$ are empty. In this case, no forward-filtering and backward-sampling is needed and $\widetilde{h}_t^k$ can be sampled directly from its prior using equation (26). Note that the sampling of the state vector α in block 2(a) depends on h_t^k rather than on $\widetilde{h}_t^k$. Using h_0^k, $\sigma_{\gamma,k}$ and $\widetilde{h}_t^k$, h_t^k can easily be reconstructed from equation (25).