Measuring the Output Gap Using Stochastic Model Specification Search

Angelia L. Grant
Economics Discipline Group,
University of Technology Sydney

Joshua C.C. Chan*
Economics Discipline Group,
University of Technology Sydney

May 2017

Abstract

It is well known that different specification choices can give starkly different output gap estimates. To account for model uncertainty, we use stochastic model specification search to average estimates over a wide variety of popular specifications—a total of more than 2000 possible combinations. In particular, we consider three types of specification choices: sets of variables used in the analysis, output trend specifications and distributional assumptions. Using US data, we find that the unemployment gap is useful in estimating the output gap, but conditional on the unemployment gap, the inflation gap no longer depends on the output gap. Our results show a steady decline in trend output growth throughout the sample, and the estimate at the end of our sample is only about 1.2%. Moreover, data favor t over Gaussian distributed innovations, suggesting the relatively frequent occurrence of extreme events.

Keywords: model averaging, trend inflation, potential output, NAIRU, Okun's law, Phillips curve

JEL classification: C11, C52, E32

1 Introduction

How do we measure the overall slack in the economy? This is an important problem for both academics and policymakers and has generated a large empirical literature. One popular approach measures the overall slack by estimating the so-called output gap— the deviation of the output of an economy from its potential or trend output. Early contributions to this approach include the Beveridge-Nelson decomposition (Beveridge and Nelson, 1981), the Hodrick-Prescott filter (Hodrick and Prescott, 1980, 1997), and the unobserved components models of Harvey (1985), Watson (1986) and Clark (1987). These are all univariate approaches that use only output data.

Later work extends these univariate approaches to incorporate more information—they seek to exploit the effects of the output gap on other macroeconomic variables. For example, Kuttner (1994) studies a bivariate unobserved components model of output and inflation through a Phillips curve relationship; Sinclair (2009) considers an alternative bivariate model of output and unemployment; and Berger and Kempa (2011) propose a trivariate system that includes output, inflation and the exchange rate. Many other papers use different sets of macroeconomic variables, and different specifications tend to give markedly different estimates.

Given the sensitivity of output gap estimates to the exact specification used in the analysis, here we take a Bayesian model averaging approach. By including a wide variety of specifications used in the literature, the model-average estimates explicitly take into account model uncertainty and are therefore more robust with respect to different modeling choices. In addition, the results from this model averaging exercise also tell us about features that are supported by the data—should we include inflation or unemployment in estimating the output gap? Should trend output be modeled as a deterministic or a stochastic process?

We consider three types of specification choices. The first is the set of variables to include and how the relationships between the variables are specified. We consider systems with output, inflation and unemployment—our specifications include univariate models of only output, bivariate models of output and inflation or unemployment, and trivariate models with all three variables.

The second model choice concerns the specification of the trend output. This is a key specification because the output gap is the difference between actual output and the unobserved trend output. Again, there is a wide range of specifications used in the

literature. For instance, a popular trend specification is a random walk with a constant drift, as in, e.g., Watson (1986) and Morley, Nelson, and Zivot (2003). Luo and Startz (2014) and Grant and Chan (2017b) consider breaks in the random walk drift, whereas a stochastic drift process is considered in Harvey and Jaeger (1993) and Harvey, Trimbur, and Van Dijk (2007). In contrast, Perron and Wada (2009) favor a linear deterministic trend but with a break in the growth rate. We consider a class of specifications that include stochastic and deterministic trends, and the trend growth rate can either be constant, have a small number of breaks or be time-varying.

The third model choice concerns the distributional assumptions of the innovations. In a recent paper, Creal, Koopman, and Zivot (2010) show that allowing for stochastic volatility helps more precisely estimate the output gap, but they only consider Gaussian innovations. In addition to allowing for stochastic volatility, we consider a framework to select between Gaussian and t distributed innovations in a data-driven manner. This is motivated by a few recent papers, such as Cúrdia, Del Negro, and Greenwald (2014) and Chiu, Mumtaz, and Pinter (2015), which highlight the importance of rare large shocks in driving business cycles.

Given the large number of specification choices, a brute-force approach that enumerates and estimates all possible combinations would be time-consuming. In addition, computing the marginal likelihood for models with time-varying coefficients and stochastic volatility is a difficult problem (e.g., Chan and Eisenstat, 2015). Instead, we construct a suitable model space that includes all relevant models under consideration. We then navigate this large model space using an algorithm based on the stochastic model specification search for time-varying parameter models proposed in Frühwirth-Schnatter and Wagner (2010). Using the proposed approach, not only can we compute the posterior model probabilities for all the specifications, we can also estimate the output gap and other quantities of interest by averaging the corresponding estimates over all the specifications.[1] Our methodology is similar to that used in Berger, Everaert, and Vierke (2016), but their focus is on testing time variation in the model parameters, whereas we are interested in

---

[1]Alternative ways to estimate the output gap using Bayesian model averaging have been considered. For example, Morley and Piger (2012) perform model averaging of output gaps using the Bayesian information criteria for a class of univariate models. Garratt, Mitchell, and Vahey (2014) combine output gaps from bivariate models of output and inflation using a rolling window of predictive likelihoods. Grant and Chan (2017b) conduct a model comparison exercise that includes various univariate models and two bivariate models. Relatedly, Kleijn and Van Dijk (2006) consider Bayesian model averaging of various univariate decompositions for US industrial production and unemployment, while Strachan and Van Dijk (2013) conduct a business cycle analysis by averaging models with different cointegrating relationships, overidentifying restrictions, lag structures, etc.

assessing a range of specification choices for estimating the output gap.

Using US data, our results suggest that the unemployment gap is useful in determining the output gap. In contrast, we find evidence that the inflation gap does not depend on the output gap once the unemployment gap is taken into account. This is consistent with the results in Morley, Piger, and Rasche (2015), where the unemployment gap explains a large fraction of the variation in the inflation gap in G7 countries. Our estimates indicate a steady decline in the trend output growth rate: from about 4% in the 1960s to about 3% in the early 1980s to 2000s, followed by a substantial drop after the Great Recession to about 1.2%. Lastly, we find evidence in favor of t distributed innovations, suggesting the relatively frequent occurrence of extreme events.

2 Stochastic Model Specification Search

In this section we develop a general model space that nests a wide range of common specifications. This is constructed along three dimensions. In Section 2.1 we first consider models that include different sets of variables and various ways these variables are related to each other. Section 2.2 investigates a variety of trend output specifications, including random walks, integrated random walks and a linear deterministic trend specification. Lastly, Section 2.3 considers specifications with Gaussian or t distributed innovations.

2.1 Model Dimension Search

As mentioned in the introduction, univariate specifications that include only output—such as the Beveridge-Nelson decomposition (Beveridge and Nelson, 1981), the Hodrick-Prescott filter (Hodrick and Prescott, 1980, 1997), and unobserved components models (Morley, Nelson, and Zivot, 2003; Grant and Chan, 2017b)—are widely used to estimate trend output. Bivariate models of output and inflation (Kuttner, 1994; Basistha, 2007; Basistha and Nelson, 2007) or output and unemployment (Sinclair, 2009) are also common.

Following Berger, Everaert, and Vierke (2016), we consider a trivariate unobserved components model of (log) output y_t, the inflation rate π_t and the unemployment rate u_t.

4

Specifically,

$$\pi_t - \pi_t^* = \rho^\pi(\pi_{t-1} - \pi_{t-1}^*) + \delta_{\lambda^u}\lambda^u(u_t - u_t^*) + \delta_{\lambda^y}\lambda^y(y_t - y_t^*) + \varepsilon_t^\pi, \tag{1}$$

$$y_t - y_t^* = \delta_{\gamma^u}\gamma^u(u_t - u_t^*) + c_t, \tag{2}$$

$$u_t - u_t^* = \rho_1^u(u_{t-1} - u_{t-1}^*) + \rho_2^u(u_{t-2} - u_{t-2}^*) + \varepsilon_t^u, \tag{3}$$

$$c_t = \rho_1^c c_{t-1} + \rho_2^c c_{t-2} + \varepsilon_t^c, \tag{4}$$

where π_t^* is trend inflation, y_t^* is trend output and u_t^* is the non-accelerating inflation rate of unemployment, or NAIRU.

The inflation equation (1) states that the deviation of inflation from its underlying trend, $\pi_t - \pi_t^*$, depends on both the unemployment gap $u_t - u_t^*$ and the output gap $y_t - y_t^*$. When the unemployment rate is at the NAIRU and output is at its trend, inflation equals its underlying trend plus a transitory AR(1) component.

The output equation (2) relates the output gap to the unemployment gap, where γ^u is the Okun's coefficient. Following Morley, Nelson, and Zivot (2003), we assume that c_t, the cyclical component of output, follows a stationary AR(2) process. Finally, the unemployment gap in (3) is modeled as a stationary AR(2) process. Since the model (1)–(4) is a triangular system, we assume the innovations ε_t^π, ε_t^c and ε_t^u to be independent at all leads and lags. For now we leave their distributions unspecified; they are discussed in Section 2.3.

In the model above we introduce three binary indicators δ_{λ^u}, δ_{λ^y} and δ_{γ^u}, i.e., each takes values in $\{0, 1\}$. This setup allows us to consider a range of models with different dimensions for estimating the output gap. For instance, when all the indicators are zero, this trivariate model reduces to the univariate unobserved components model of output in Watson (1986).[2] When $\delta_{\lambda^u} = 0$, $\delta_{\gamma^u} = 0$ and $\delta_{\lambda^y} = 1$, it becomes a bivariate unobserved components model of output and inflation. When $\delta_{\lambda^u} = 0$, $\delta_{\lambda^y} = 0$ and $\delta_{\gamma^u} = 1$, the model becomes a bivariate unobserved components model of output and unemployment. In addition, by allowing both the output and unemployment gaps to enter the inflation equation, we can investigate which gap is a better proxy for real activity, or if both are informative for modeling inflation.

---

[2]Technically, the dimension of the trivariate model remains the same. But since the three equations (1)-(3) are now unrelated to each other, it becomes the same as estimating three univariate models separately. The output gap estimates from this model with all the indicators set to zero are the same as those from a univariate model of output.

2.2 Trend Output Specification Search

There are many different specifications for modeling the trend output y_t^*. For example, Perron and Wada (2009) favor a linear deterministic trend with a break in the trend growth rate in 1973. In contrast, using the marginal likelihood as the model selection criterion, Luo and Startz (2014) and Grant and Chan (2017b) find evidence in support of a random walk trend where the drift has a break in 2006-2007.

More importantly, how one models y_t^* turns out to have a critical impact on the output gap and the trend output growth estimates. For instance, it is well known that the correlated unobserved components model of Morley, Nelson, and Zivot (2003) yields output gap estimates that are small in amplitude, whereas the Hodrick-Prescott (HP) filter generates large and persistent cycles. Grant and Chan (2017a) show that this difference arises entirely due to the different ways the trend output is modeled. In particular, the correlated unobserved components model assumes a random walk trend with a constant drift (or a small number of breaks in the drift), whereas the HP filter implicitly models the output trend as a particular random walk with a stochastic drift (see also Harvey and Jaeger, 1993).

Given the importance of modeling the trend output y_t^* appropriately, we adopt a flexible data-driven approach based on the stochastic model specification search for time-varying parameter models proposed in Frühwirth-Schnatter and Wagner (2010). This approach constructs a general model that nests all the alternative models of interest. By introducing suitable binary indicators and treating these indicators as model parameters, one can then navigate the model space in a data-driven manner.

To illustrate the key ideas, below we consider two representations of a stochastic trend model. We will then introduce the general trend specification we use in the application. The first is the following centered parameterization:

$$
\begin{aligned}
y_t^* &= y_{t-1}^* + a_t + \varepsilon_t^{y^*}, & \varepsilon_t^{y^*} &\sim \mathcal{N}(0, \sigma_{y^*}^2), \\
a_t &= a_{t-1} + \varepsilon_t^a, & \varepsilon_t^a &\sim \mathcal{N}(0, \sigma_a^2),
\end{aligned}
$$

where the initial condition a_0 is treated as an unknown parameter. Here a_t can be interpreted as the time-varying trend growth rate at time t. This setup can be used to nest a number of trend specifications. For example, if both variances $\sigma_{y^*}^2$ and σ_a^2 are zero, then y_t^* follows a linear deterministic trend with slope a_0. If $\sigma_{y^*}^2 > 0$ and $\sigma_a^2 = 0$, then y_t^* follows a random walk with drift a_0.

One complication of this approach is that zero lies at the boundary of the parameter space for variances $\sigma_{y^*}^2$ and σ_a^2. To get around this technical difficulty, we consider the second representation—the non-centered parameterization proposed in Frühwirth-Schnatter and Wagner (2010). Specifically, let

$$y_t^* = y_0^* + a_0 t + \sigma_{y^*} \widetilde{y}_t^* + \sigma_a \sum_{i=1}^t \widetilde{a}_i,$$

where

$$\widetilde{y}_t^* = \widetilde{y}_{t-1}^* + \varepsilon_t^{\widetilde{y}^*}, \qquad\qquad \varepsilon_t^{\widetilde{y}^*} \sim \mathcal{N}(0,1), \qquad (5)$$
$$\widetilde{a}_t = \widetilde{a}_{t-1} + \varepsilon_t^{\widetilde{a}}, \qquad\qquad \varepsilon_t^{\widetilde{a}} \sim \mathcal{N}(0,1), \qquad (6)$$

and the state equations are initialized with $\widetilde{y}_0^* = \widetilde{a}_0 = 0$.[3] Here σ_{y^*} and σ_a are the standard deviations which are defined to have support in the real line.[4]

To get back to the centered parameterization, let $a_t = a_0 + \sigma_a \widetilde{a}_t$, $\zeta_t^a = \sigma_a \varepsilon_t^{\widetilde{a}}$ and $\zeta_t^{y^*} = \sigma_{y^*} \varepsilon_t^{\widetilde{y}^*}$. Then, we have

$$a_t - a_{t-1} = \sigma_a \varepsilon_t^{\widetilde{a}} = \zeta_t^a, \qquad\qquad \zeta_t^a \sim \mathcal{N}(0, \sigma_a^2),$$
$$y_t^* - y_{t-1}^* = a_0 + \delta_{y^*} \sigma_{y^*} \varepsilon_t^{\widetilde{y}^*} + \sigma_a \widetilde{a}_t,$$
$$= a_t + \zeta_t^{y^*}, \qquad\qquad \zeta_t^{y^*} \sim \mathcal{N}(0, \sigma_{y^*}^2).$$

Hence, the two parameterizations are equivalent.

Next, we use the non-centered parameterization and enrich it by allowing breaks in the random walk drift or deterministic growth rate a_0. In addition, we introduce binary indicators to select between the alternative trend specifications. More specifically, consider the following trend output specification:

$$y_t^* = y_0^* + \delta_{a_0} a_0 t + \delta_{y^*} \sigma_{y^*} \widetilde{y}_t^* + \delta_a \sigma_a \sum_{i=1}^t \widetilde{a}_i + \delta_{d_1} d_1 t D_{1973,t} + \delta_{d_2} d_2 t D_{2007,t} \qquad (7)$$

where $\delta_{a_0}, \delta_{y^*}, \delta_a, \delta_{d_1}$ and δ_{d_2} are binary indicators, and $D_{1973,t}$ and $D_{2007,t}$ are dummy

---

[3]Berger, Everaert, and Vierke (2016) also use this approach to test for time variation in the trend growth rate. We build upon their work and consider a larger class of trend specifications. Consequently, our empirical results are somewhat different from theirs, which are discussed in Section 3.

[4]Note that the signs of σ_{y^*} and σ_a are not identified. We refer the readers to Frühwirth-Schnatter and Wagner (2010) for a more detailed discussion.

variables—$D_{1973,t}$ takes value 1 if time t is 1973Q1 or later, and 0 otherwise; $D_{2007,t}$ is 1 if time t is 2007Q1 or later, and 0 otherwise. This trend specification includes many models previously considered in the literature. Below we give a few examples.

If all the indicators are zero except δ_{a_0}, then the trend output y_t^* has a linear deterministic trend with growth rate a_0. If in addition $\delta_{d_1} = 1$, then this trend specification is the same as in Perron and Wada (2009)—a deterministic trend with a break in the growth rate in 1973Q1. Similarly, if δ_{d_2} is also 1, then the deterministic trend growth has an additional break in 2007Q1, as considered in Grant and Chan (2017b).

Next, if $\delta_a = \delta_{d_1} = \delta_{d_2} = 0$ and $\delta_{y^*} = \delta_{a_0} = 1$, then y_t^* follows a random walk with drift a_0. If in addition $\delta_{d_2} = 1$, then the drift has a break in 2007Q1, similar to the specification in Luo and Startz (2014).

If $\delta_a = \delta_{a_0} = 1$ and $\delta_{y^*} = \delta_{d_1} = \delta_{d_2} = 0$, then y_t^* has the same trend specification implicitly assumed by the Hodrick-Prescott filter (see, e.g, Harvey and Jaeger, 1993; Grant and Chan, 2017a). If both $\delta_a = \delta_{y^*} = 1$, then y_t^* has a stochastic trend with a time-varying growth rate. Hence, the framework in (7) nests many popular trend specifications in the literature, and we are able to average the output gap estimates across all these different specifications.

Finally, we assume that trend inflation and the NAIRU follow independent driftless random walks. Unlike output that is expected to have a positive long-run growth, the long-run growth for both inflation and unemployment is expected to be zero—at least for developed countries like the US. More specifically, consider

$$\pi_t^* = \pi_{t-1}^* + \varepsilon_t^{\pi^*}, \tag{8}$$
$$u_t^* = u_{t-1}^* + \varepsilon_t^{u^*}, \tag{9}$$

where $\varepsilon_t^{\pi^*} \sim \mathcal{N}(0, \sigma_{\pi^*}^2)$ and $\varepsilon_t^{u^*} \sim \mathcal{N}(0, \sigma_{u^*}^2)$.

2.3 Distribution Search

A large and growing literature has highlighted the importance of allowing for time-varying volatility in modeling macroeconomic time series, especially output and inflation. See, for example, Cogley and Sargent (2005), Primiceri (2005), Stock and Watson (2007), Clark (2011) and Chan (2013). However, most of these papers assume Gaussian innovations—

consequently these models assign relatively low probability to extreme events. Using a dynamic stochastic general equilibrium model with stochastic volatility and t innovations, Cúrdia, Del Negro, and Greenwald (2014) demonstrate that rare large shocks play an important role in driving US business cycles. More recently, Chiu, Mumtaz, and Pinter (2015) consider a vector autoregression (VAR) with stochastic volatility and t innovations, and find that it fits the data better than alternative models. Cross and Poon (2016) show that these models also forecast better than their counterparts with Gaussian innovations.

In light of this discussion, we allow ε_t^π, ε_t^c and ε_t^u—the innovations to the inflation equation (1), the output equation (4) and the unemployment equation (3), respectively—to have stochastic volatility. In addition, we implement a distribution search to choose between a Gaussian and a t distribution. Specifically, we introduce three binary variables δ_{z^π}, δ_{z^c} and δ_{z^u} as follows. If $\delta_{z^\pi} = 0$, then ε_t^π follows a Gaussian distribution with variance $e^{h_t^\pi}$, i.e., $\varepsilon_t^\pi \sim \mathcal{N}(0, e^{h_t^\pi})$. If $\delta_{z^\pi} = 1$, then ε_t^π follows a t distribution with the following latent variable representation (see, e.g., Geweke, 1993; Koop, Poirier, and Tobias, 2007): $(\varepsilon_t^\pi \mid z_t^\pi) \sim \mathcal{N}(0, z_t^\pi e^{h_t^\pi})$ and $z_t^\pi \sim \mathcal{IG}(\nu_\pi/2, \nu_\pi/2)$, where $\mathcal{IG}(b_1, b_2)$ denotes the inverse-gamma distribution with mean $b_2/(b_1 - 1)$ when $b_1 > 1$. Similarly we define δ_{z^c} and δ_{z^u} to choose between a Gaussian and a t distribution for ε_t^c and ε_t^u, respectively.

Finally, the log volatilities h_t^π and h_t^c are assumed to follow random walk processes as follows:

$$h_t^\pi = h_{t-1}^\pi + \varepsilon_t^{h^\pi}, \tag{10}$$

$$h_t^c = h_{t-1}^c + \varepsilon_t^{h^c}, \tag{11}$$

$$h_t^u = h_{t-1}^u + \varepsilon_t^{h^u}, \tag{12}$$

where $\varepsilon_t^{h^\pi} \sim \mathcal{N}(0, \sigma_{h^\pi}^2)$, $\varepsilon_t^{h^c} \sim \mathcal{N}(0, \sigma_{h^c}^2)$ and $\varepsilon_t^{h^u} \sim \mathcal{N}(0, \sigma_{h^u}^2)$.

2.4 Model Summary

To summarize, the observation equations of the full model are specified in (1)–(4), where the innovations are given as follows: if $\delta_{z^\pi} = 1$, then $\varepsilon_t^\pi \sim \mathcal{N}(0, z_t^\pi e^{h_t^\pi})$ with $z_t^\pi \sim \mathcal{IG}(\nu_\pi/2, \nu_\pi/2)$; if $\delta_{z^\pi} = 0$, $\varepsilon_t^\pi \sim \mathcal{N}(0, e^{h_t^\pi})$ instead. Similarly, for $\delta_{z^c} = 1$, $\varepsilon_t^c \sim \mathcal{N}(0, z_t^c e^{h_t^c})$ with $z_t^c \sim \mathcal{IG}(\nu_c/2, \nu_c/2)$; if $\delta_{z^c} = 0$, $\varepsilon_t^c \sim \mathcal{N}(0, e^{h_t^c})$. And for $\delta_{z^u} = 1$, $\varepsilon_t^u \sim \mathcal{N}(0, z_t^u e^{h_t^u})$ with $z_t^u \sim \mathcal{IG}(\nu_u/2, \nu_u/2)$; if $\delta_{z^u} = 0$, $\varepsilon_t^u \sim \mathcal{N}(0, e^{h_t^u})$. Finally, the state equations are given in (5)–(12).

Table 1 provides a summary of the binary indicators. Given that we have eleven indicators, there are a total of $2^{11} = 2048$ models to consider. Estimating each of these models and averaging the corresponding output gap estimates would simply be impractical. In the next section we outline an algorithm to efficiently navigate this large model space and obtain an output gap estimate that averages across these models.

Table 1: Summary of the binary indicators.

δ_{λ^u}	determines if the unemployment gap enters the inflation equation
δ_{γ^u}	determines if the unemployment gap enters the output equation
δ_{λ^y}	determines if the output gap enters the inflation equation
δ_{a_0}	determines if the trend output growth has a drift when $\delta_a = 0$
δ_{y^*}	determines if permanent level shocks affect the trend output
δ_a	determines if the trend output has a time-varying growth rate
δ_{d_1}	determines if the drift in trend output growth has a break in 1973
δ_{d_2}	determines if the drift in trend output growth has a break in 2007
δ_{z^π}	determines if the innovations to inflation equation follow a t distribution
δ_{z^c}	determines if the innovations to output gap equation follow a t distribution
δ_{z^u}	determines if the innovations to unemployment equation follow a t distribution

2.5   Priors and Bayesian Estimation

In this section we give an overview of the prior distributions and the posterior sampler to simulate from the joint posterior distribution. Details of the priors and the derivations of the posterior sampler are provided in the Appendix. Most of the priors used are standard; here we briefly discuss the priors on the variances of the innovations in the state equations.

Following Frühwirth-Schnatter and Wagner (2010), we consider normal priors with zero mean for the (unsigned) standard deviations σ_{y^*} and σ_a: $\sigma_{y^*} \sim \mathcal{N}(0, V_{\sigma_{y^*}})$ and $\sigma_a \sim \mathcal{N}(0, V_{\sigma_a})$. By a change of variable (see, e.g., Kroese and Chan, 2014, Section 3.5), it can be shown that the implied prior for the variance σ_a^2 is $\mathcal{G}(0.5, 0.5/V_{\sigma_a})$, where $\mathcal{G}(b_1, b_2)$ denotes the Gamma distribution with mean b_1/b_2. Compared to the conventional inverse-gamma prior, this gamma prior has more mass concentrated around small values of σ_a^2.

Figure 1 plots the gamma and inverse gamma density functions with the same mean and variance. It is clear that the shapes of the two densities are rather different around zero: while the mode of the inverse gamma density is away from zero, most of the mass of the gamma density is near zero. Hence, the gamma prior provides shrinkage—*a priori*

it favors the more parsimonious model in which y_t^* does not have a time-varying growth rate.
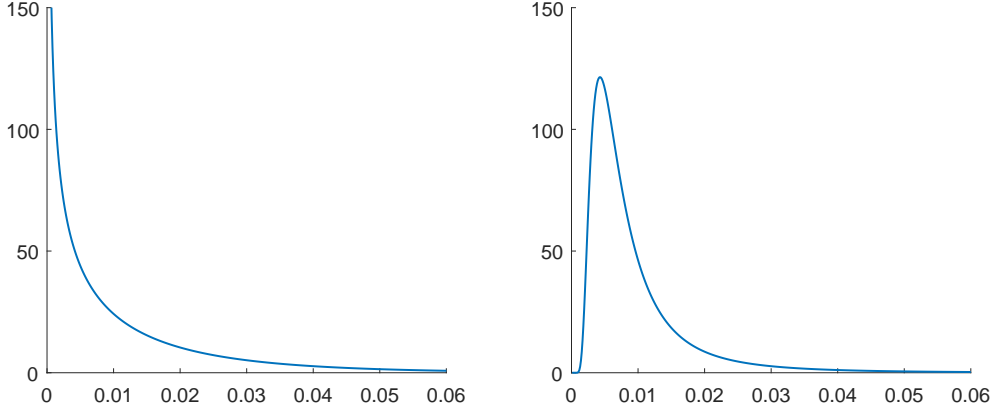

Figure 1: Probability density functions of gamma (left panel) and inverse gamma (right panel) with mean 0.1^2 and standard deviation $\sqrt{2} \times 0.01^2$.

Our posterior sampler consists of 11 blocks and uses elements of the algorithms in Frühwirth-Schnatter and Wagner (2010) and Chan, Koop, and Potter (2016). In particular, a key novel feature of our approach is that it builds upon the band matrix algorithms developed in Chan and Jeliazkov (2009) and Chan (2013). It is shown in McCausland, Miller, and Pelletier (2011) that this band matrix approach is more efficient compared to the conventional Kalman filter-based algorithms. In addition, we follow the recommendation of Frühwirth-Schnatter and Wagner (2010) where we integrate out the relevant regression coefficients when we sample the model indicators to improve efficiency.

Below we provide a brief outline of the sampler. The derivation details are given in the Appendix.

1. Sample $\boldsymbol{\pi}^*$ from its full conditional distribution. This is done using the precision sampler of Chan and Jeliazkov (2009).

2. Sample $\mathbf{u}^*$ from its full conditional distribution. This is again implemented by using the precision sampler of Chan and Jeliazkov (2009).

3. Sample $\widetilde{\mathbf{y}}^*$ and $(y_0^*, a_0, \sigma_{y^*}, \sigma_a, d_1, d_2)$ from their full conditional distributions, and randomly permute the signs of $\widetilde{\mathbf{y}}^*$ and σ_{y^*} as suggested in Frühwirth-Schnatter and Wagner (2010).

4. Sample $\widetilde{\mathbf{a}}$ and $(y_0^*, a_0, \sigma_{y^*}, \sigma_a, d_1, d_2)$ from their full conditional distributions, and randomly permute the signs of $\widetilde{\mathbf{a}}$ and σ_a.

5. Sample $(\delta_{\lambda^u}, \delta_{\lambda^y}, \delta_{\gamma^u})$ and $(\rho^\pi, \lambda^u, \lambda^y, \gamma^u)$ jointly. This is achieved by first drawing the indicators marginally of $(\rho^\pi, \lambda^u, \lambda^y, \gamma^u)$, followed by sampling the latter from their joint conditional distribution.

6. Sample $(\delta_{a_0}, \delta_{y^*}, \delta_a, \delta_{d_1}, \delta_{d_2})$ and $(y_0^*, a_0, \sigma_{y^*}, \sigma_a, d_1, d_2)$ jointly. Again, this is achieved by first drawing the indicators marginally of $(y_0^*, a_0, \sigma_{y^*}, \sigma_a, d_1, d_2)$, followed by sampling the latter from their full conditional distribution.

7. Sample $(\delta_{z^\pi}, \delta_{z^c}, \delta_{z^u})$, $\mathbf{z}^\pi$, $\mathbf{z}^c$ and $\mathbf{z}^u$ jointly: first sample the indicators marginally of $\mathbf{z}^\pi$, $\mathbf{z}^c$ and $\mathbf{z}^u$, followed by drawing $\mathbf{z}^\pi$, $\mathbf{z}^c$ and $\mathbf{z}^u$ from their full conditional distributions.

8. Sample $\boldsymbol{\rho}^u$ and $\boldsymbol{\rho}^c$ from their full conditional distributions.

9. Sample $\mathbf{h}^\pi$, $\mathbf{h}^c$ and $\mathbf{h}^u$ from their full conditional distributions using the auxiliary mixture sampler of Kim, Shepherd, and Chib (1998) as implemented in Chan and Hsiao (2014).

10. Sample the variances $\sigma_{u^*}^2$, $\sigma_{\pi^*}^2$, $\sigma_{h^\pi}^2$, $\sigma_{h^c}^2$ and $\sigma_{h^u}^2$ from their respectively full conditional distributions. The conditional distributions of these variances are non-standard due to the gamma priors. We draw each of these variances using an independence-chain Metropolis-Hastings step.

11. Sample ν_π, ν_c and ν_u using an independence-chain Metropolis-Hastings step as implemented in Chan and Hsiao (2014).

3   Empirical Results

The data for our analysis are US quarterly real GDP, CPI inflation and the civilian unemployment rate from 1948Q1 to 2015Q4. All series are sourced from the Federal Reserve Bank of St. Louis economic database. The real GDP series is transformed by taking logs and multiplying by 100; the CPI index is transformed to an annualized growth rate. The posterior analysis presented below is based on 100000 posterior draws after a burn-in period of 10000.

3.1 Posterior Model Probabilities

Table 2 reports the posterior means of the eleven indicator variables. Since these are binary variables, the posterior means can be interpreted as inclusion probabilities of various features. First, in terms of model dimension, the results suggest that the unemployment gap is directly useful for estimating the output gap—the posterior mean of δ_{γ^u} is 1, and the unemployment gap is always included in the output equation. This therefore supports the joint modeling of output and unemployment, as is done in Sinclair (2009) and Mertens (2014).

However, the estimate of δ_{λ^y} is only 0.25, indicating that the output gap is relatively unimportant in determining inflation once we take into account the unemployment gap. This is in line with the finding in Morley, Piger, and Rasche (2015), who show that the unemployment gap explains a large fraction of the variation in the inflation gap in G7 countries, but in contrast to earlier work that emphasizes the direct connection between the output gap and inflation, such as Kuttner (1994) and Gerlach and Smets (1999). We emphasize that our results do *not* imply that inflation is not useful for estimating the output gap. Rather the information is channeled through the unemployment gap: inflation helps refine the estimate on the unemployment gap, which in turn helps pin down the output gap.

Table 2: Posterior estimates of the indicator variables.

δ_{λ^u}	δ_{γ^u}	δ_{λ^y}	δ_{a_0}	δ_{y^*}	δ_a	δ_{d_1}	δ_{d_2}	δ_{z^π}	δ_{z^c}	δ_{z^u}
0.98	1.00	0.25	1.00	0.55	1.00	0.15	0.00	0.99	0.57	0.89

Second, in terms of trend output specifications, our results strongly indicate that trend output has a time-varying growth rate. Recall that if $\delta_{d_1} = \delta_{d_2} = 1$ and $\delta_a = 0$, then the trend growth rate has breaks in 1973 and 2007—and nowhere else. By contrast, if $\delta_a = 1$, then the trend growth rate changes every period. The posterior mean of δ_a is 1, whereas those of δ_{d_1} and δ_{d_2} are 0.15 and 0, respectively. This suggests that the breaks in 1973 and 2007 are not needed if we allow the trend output growth to gradually drift every period. In contrast to earlier work, such as Perron and Wada (2009), Luo and Startz (2014) and Grant and Chan (2017b), that finds one or two breaks in trend output growth using univariate models, the results from our trivariate model favor specifying trend output growth as a stochastic process. Finally, the posterior mean of δ_{y^*} is estimated to be 0.55, showing mixed evidence on the role of the permanent level shocks to the trend output.

We take this as partial support for the trend output specification in Grant and Chan (2017a) that turns off these permanent level shocks.

Third, in terms of selecting between Gaussian and t innovations, the data generally favor the latter in both the inflation and unemployment equations, whereas the evidence is mixed for the output equation. In other words, for inflation and unemployment, even after allowing for stochastic volatility, the magnitude of some innovations is still larger than that is typical under the Gaussian distribution. These results support the claim that rare large shocks play an important role in driving US business cycles.

Next, we report in Table 3 the posterior model probabilities of the top ten models. These ten models account for a majority of the posterior model probabilities (the sum of the probabilities is 0.79), but none of the models dominate. This illustrates a substantial degree of model uncertainty and highlights the importance of averaging across models.

Table 3: Posterior model probabilities of the top ten models.

δ_{λ^u}	δ_{γ^u}	δ_{λ^y}	δ_{a_0}	δ_{y^*}	δ_a	δ_{d_1}	δ_{d_2}	δ_{z^π}	δ_{z^c}	δ_{z^u}	Prob
1	1	0	1	1	1	0	0	1	1	1	0.172
1	1	0	1	0	1	0	0	1	1	1	0.143
1	1	0	1	1	1	0	0	1	0	1	0.125
1	1	0	1	0	1	0	0	1	0	1	0.111
1	1	1	1	1	1	0	0	1	1	1	0.059
1	1	1	1	1	1	0	0	1	0	1	0.046
1	1	1	1	0	1	0	0	1	1	1	0.041
1	1	1	1	0	1	0	0	1	0	1	0.035
1	1	0	1	1	1	1	0	1	1	1	0.034
1	1	0	1	0	1	1	0	1	1	1	0.028

The most likely model is one in which the unemployment gap is included in both the inflation and output equations, and the output gap is left out of the inflation equation; trend output has a time-varying growth rate and a permanent level shock component; and the innovations in all equations are t distributed. This model is closely followed by two variants: one in which the permanent level shock component is turned off; in the other the distribution of the output equation is Gaussian instead of t. Even though these three models receive the most support from the data, model uncertainty is substantial. In what follows, we report various features of interest by averaging across all the different models rather than showing only estimates from the most likely model.

3.2   Parameter Estimates

In this section we discuss the parameter estimates from the stochastic model specification search exercise—these estimates are computed by averaging across all the different models. We focus on the relationships among the inflation gap, the unemployment gap and the output gap, how persistent they are, and how often large shocks impact the inflation and output equations.

Table 4 reports the posterior means and 2.5 and 97.5 percentiles for selected parameters. The posterior mean of λ^u is -0.57, indicating a downward sloping Phillips curve—in fact, a 1 percentage point unemployment gap is associated with an inflation rate that is on average 0.57 percentage points below its trend. Counterintuitively, the coefficient associated with the output gap in the inflation equation, λ^y, is negative, but its magnitude is small and the corresponding 95% credible interval includes 0. This is consistent with the posterior estimates of the indicator variables reported in Table 2, which indicate that one needs not include the output gap in the inflation equation in the presence of the unemployment gap. Finally, the Okun's coefficient γ^u—i.e., the coefficient associated with the unemployment gap in the output equation—is estimated to be -1.71, which is only slightly different from the widely accepted value of -2 (Ball, Leigh, and Loungani, 2013).

Table 4: Posterior means and percentiles for selected parameters.

	posterior mean	2.5 percentile	97.5 percentile
λ^u	-0.57	-1.30	-0.01
γ^u	-1.71	-1.95	-1.47
λ^y	-0.04	-0.39	0.09
ρ^π	0.33	0.16	0.50
ρ_1^c	0.90	0.73	1.06
ρ_2^c	0.01	-0.13	0.15
ρ_1^u	1.64	1.53	1.75
ρ_2^u	-0.67	-0.78	-0.56
ν_π	7.7	2.8	30.3
ν_c	27.5	5.3	48.9
ν_u	13.5	2.4	46.0

Second, the AR(1) coefficient for the inflation gap, ρ^π, is estimated to be 0.33, indicating that the inflation gap is moderately persistent. In stark contrast, both the output and unemployment gaps are substantially more persistent—the sums of the AR coefficients

are 0.91 and 0.97, respectively.

Third, the degree of freedom parameters associated with the t distributions in the inflation, output and unemployment equations are estimated to be 7.7, 27.5 and 13.5, respectively. This implies that there are more large, rare shocks to the inflation and unemployment equations than allowed under the Gaussian assumption, whereas the innovations in the output equation are more similar to the those under the Gaussian distribution. These results are again consistent with the posterior estimates of the indicator variables reported in Table 2.

Next, we plot in Figure 2 the histograms of the posterior draws of σ_{y^*} and σ_a, the standard deviations of the shocks to trend output and trend output growth, respectively. It is clear that while the density of σ_{y^*} has a point mass at 0, the density of σ_a has two distinct modes and no mass near 0. This is in line with earlier results on posterior model probabilities—there is mixed evidence on the role of permanent level shocks to trend output, but the data strongly favor a time-varying growth rate for trend output.
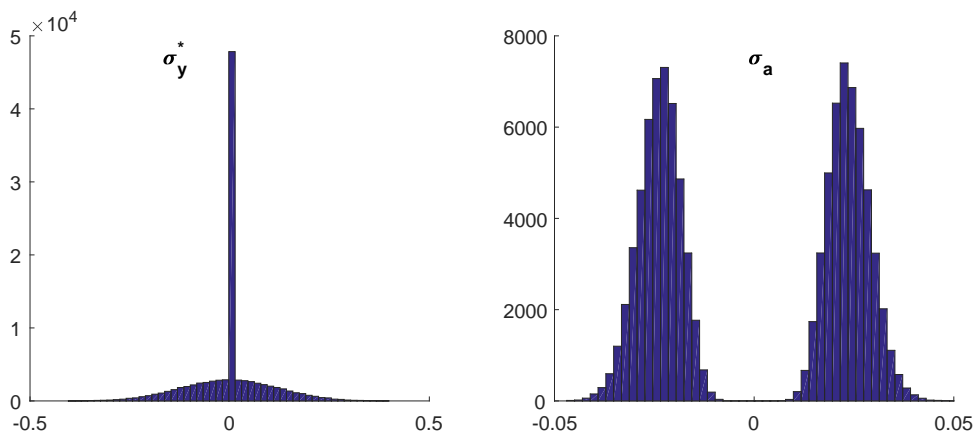

Figure 2: Histograms of the posterior draws of σ_{y^*} (left panel) and σ_a (right panel).

3.3 Output Gap and Trend GDP Growth Estimates

In this section we report the estimates of the output gap and the trend output growth rate from the proposed model. For comparison, we also report the corresponding estimates produced by the US Congressional Budget Office (CBO), as well as estimates from the univariate correlated unobserved components model (UCUR) of Morley, Nelson, and Zivot (2003) with two breaks in the trend growth rate at 1973Q1 and 2007Q1.

Figure 3 depicts the output gap estimates. It is evident that all three output gaps generally coincide with the recession periods dated by the National Bureau of Economic Research (NBER). However, there are substantial differences between the estimates from the new trivariate model and those from the univariate UCUR model. For example, under the new model, the trough of the recession in the early 1980s is estimated to be about -11%, whereas the estimate under the UCUR is only about -3%. In addition, under the UCUR, the output gap essentially closed in around 2011, whereas the new model indicates that the output gap has remained large in magnitude since the Great Recession. This might reflect the fact that the estimates from our trivariate model take into account the large unemployment gap, which is not captured in the univariate UCUR model.

Despite using very different methodologies, the output gap estimates from the new model and the CBO are remarkably similar—the correlation between the two measures is about 0.95. However, our estimates tend to be larger in magnitude, at least up until the early 1990s. For example, the CBO estimates an output gap of about -8% at the trough of the 1980s recession, compared to -11% from our model. This might reflect the typically higher natural rate of unemployment estimated by the CBO up until the 1990s (see Figure 6 below)—a higher NAIRU often implies a smaller unemployment gap, and therefore a smaller output gap in magnitude through Okun's law.
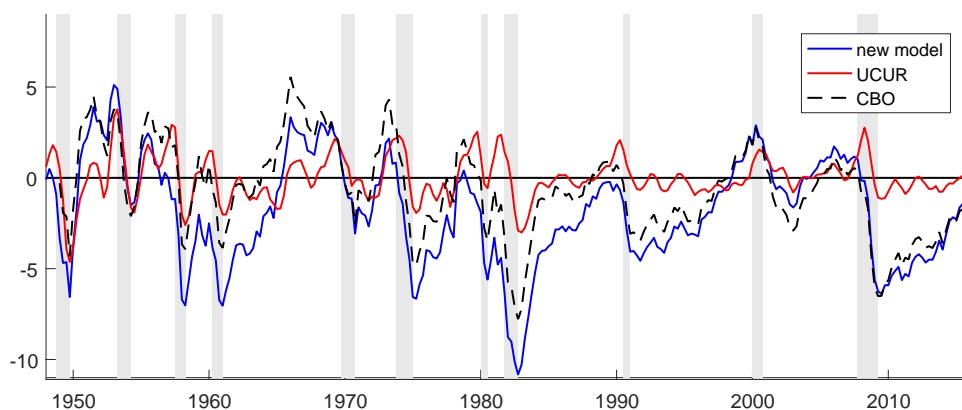

Figure 3: Estimates of output gap from the proposed model, the CBO and the correlated UC model with two breaks in the trend growth rate at 1973Q1 and 2007Q1. The shaded regions are the NBER recession dates.

Figure 4 plots the estimates of the annualized trend output growth. All three estimates show an overall steady decline: from between 4% and 5% in the 1960s to about 3%

from the early 1980s to 2000s. After that all estimates drop substantially to between 1% and 2% before the Great Recession. The general decline in trend output growth is consistent with the results in Gordon (2012), who argues that US labor productivity has slowed markedly after 1972 compared to the previous eight decades, because the main ideas of the Second Industrial Revolution which started in early 1900s had mostly been implemented by then.
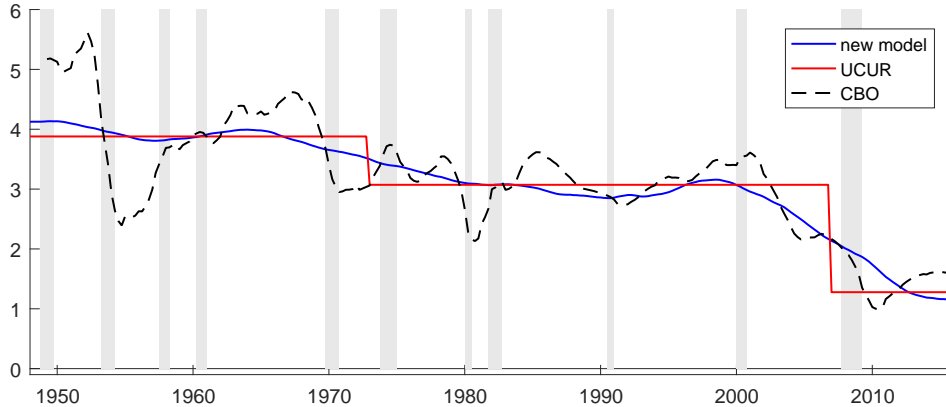

Figure 4: Estimates of annualized trend output growth from the proposed model, the CBO and the correlated UC model with two breaks in the trend growth rate at 1973Q1 and 2007Q1.

The UCUR and our proposed models give strikingly similar estimates; the only difference lies in the nature of the breaks—whether the decrease in trend output is gradual in our model or abrupt in the UCUR model. As the results on posterior model probabilities in Section 3.1 show, there is strong evidence in support of specifying a gradual change in output trend growth. The estimates from the CBO are also generally similar, though they seem to be more volatile. At the end of the sample at 2015Q4, the estimates from the proposed model, the UCUR and the CBO are 1.2%, 1.3% and 1.6%, respectively.

3.4   Trend Inflation, NAIRU and SV Estimates

Next, we plot trend inflation from the proposed model in Figure 5. For comparison, we also report the estimates from the univariate unobserved components model with stochastic volatility (UCSV) of Stock and Watson (2007). Overall the two series are broadly similar, showing double peaks in mid-1970s and early 1980s, and a marked decline since mid-1980s. The trend inflation estimates from the proposed model, however, are

generally less volatile.

It is also interesting to note that in contrast to the results in Berger, Everaert, and Vierke (2016), our trend inflation estimates increase from about 2.5% in mid-2000s to about 5% in 2010. Coibion and Gorodnichenko (2015) argue that the absence of disinflation during the Great Recession—i.e., that inflation was much higher than expected given the high unemployment rate—can be attributed to the rise in firms' inflation expectations between 2009 and 2011. Our results support this explanation. In particular, the proposed model indicates that in order to explain both the high unemployment rate and the higher than expected inflation rate during the Great Recession given the Phillips curve relationship, trend inflation needs to be higher.
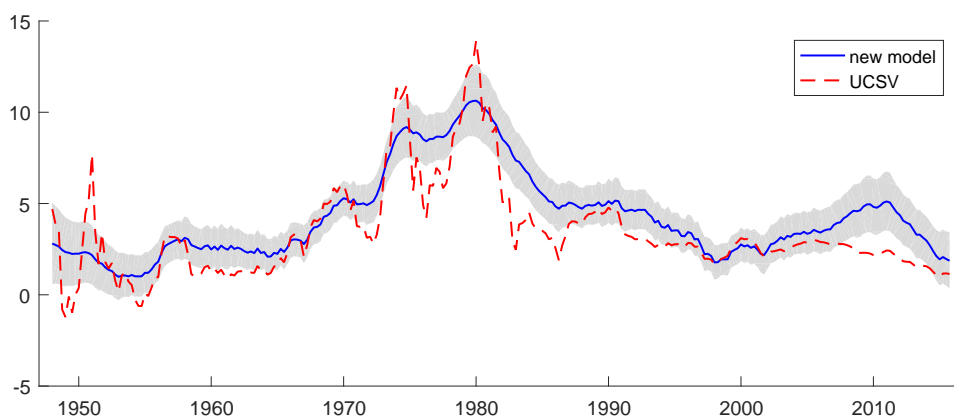

Figure 5: Estimates of trend inflation from the proposed model and the UCSV model. The shaded region represents the 10% and 90% quantiles.

Figure 6 depicts the NAIRU estimates from the proposed model and the natural rate of unemployment produced by the CBO. Both estimates are smooth and exhibit only very low frequency movements. However, as mentioned previously, the CBO estimates tend to be larger than those from the proposed model up until late 1990s—from then onward they are very similar. At the end of the sample, the estimates from the proposed model and the CBO are 4.7% and 4.8%, respectively.
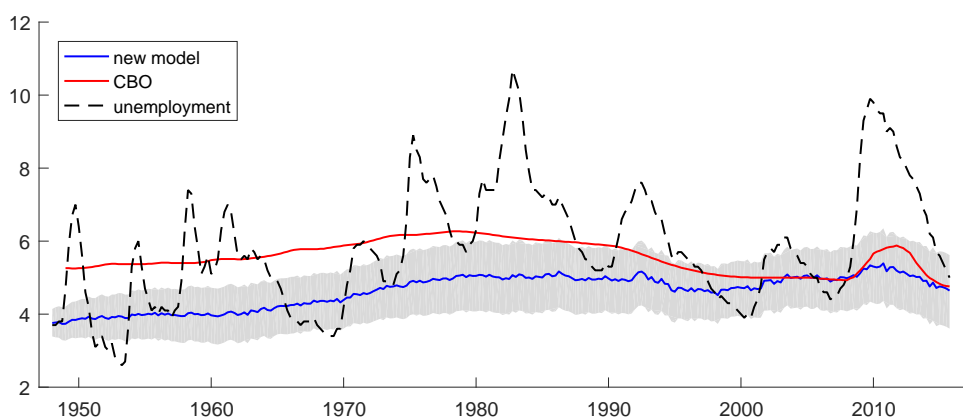
19

Figure 6: Estimates of the NAIRU from the proposed model as well as the natural rate of unemployment produced by the CBO. The shaded region represents the 10% and 90% quantiles.

Figure 7 reports the stochastic volatility estimates for the inflation, output and unemployment equations. The three series of volatility estimates are clearly time-varying, but they show different patterns. Both the volatilities of the inflation and unemployment innovations are episodic—they are high at the beginning of the sample; they drop substantially during the 1960s; they increase again during the Great Inflation, and subside during the Great Moderation, until they peak again following the aftermath of the Great Recession. In contrast, the volatility of the output innovations appears to experience only a one-off drop in the early 1980s, which coincides with the usual timing of the Great Moderation. These estimates highlight the empirical relevance of allowing stochastic volatility in modeling the three variables.
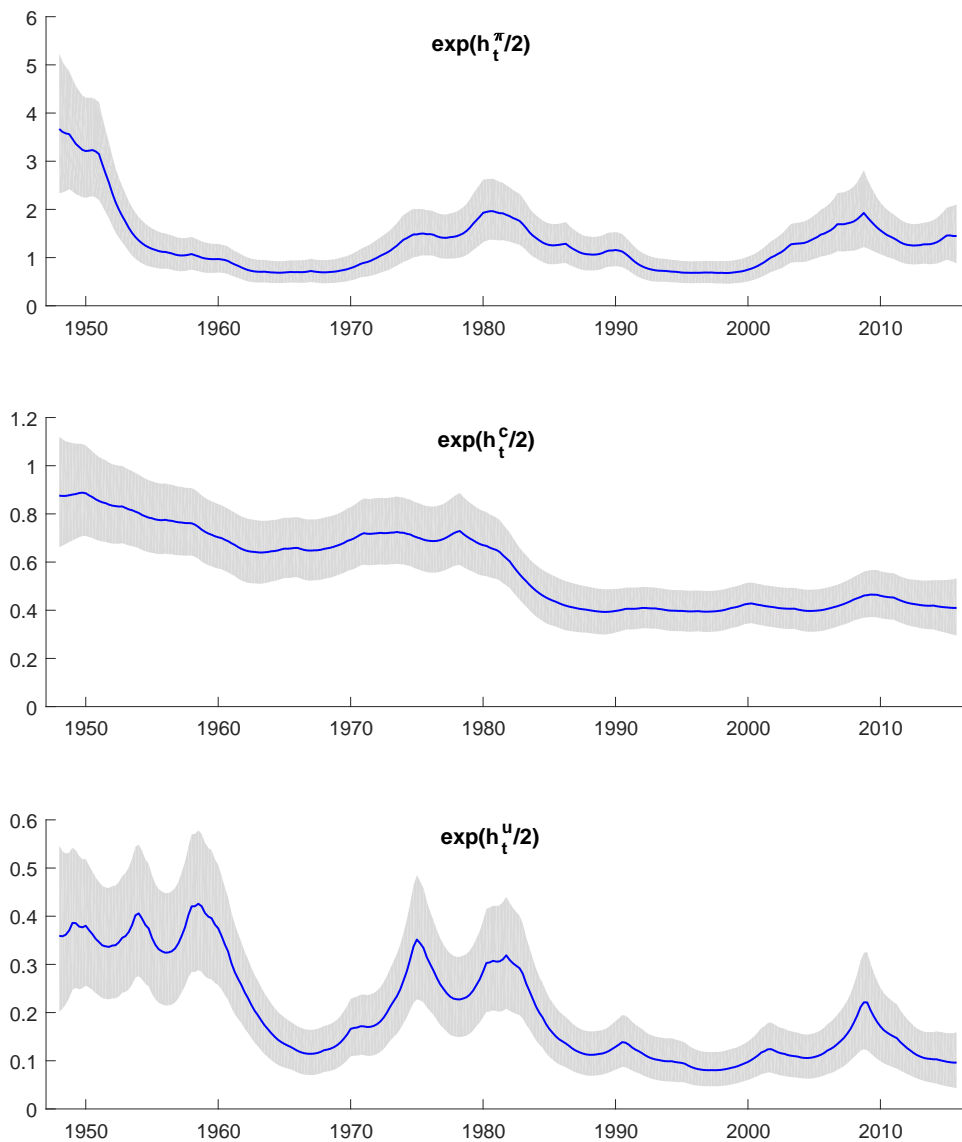
Figure 7: Estimates of the stochastic volatility for the inflation equation (top panel), the output equation (middle panel) and the unemployment equation (bottom panel). The estimates are expressed in standard deviations. The shaded region represents the 10% and 90% quantiles.

4 Concluding Remarks and Future Research

This paper accounts for the sensitivity of the output gap to different specification choices by taking a Bayesian model averaging approach that averages estimates across a wide variety of specifications proposed in the literature. In particular, we consider three types of specification choices—sets of variables, trend output specifications and distributional assumptions—and construct a suitable model space that nests all relevant models. We then navigate this large model space using an algorithm based on the stochastic model specification search for time-varying parameter models proposed in Frühwirth-Schnatter and Wagner (2010).

Our results indicate that the unemployment gap is useful in determining the output gap, but conditional on the unemployment gap, the inflation gap does not depend on the output gap. We find a steady decline in trend output growth throughout the sample; the estimate at the end of the sample is only about 1.2%. Moreover, data favor t distributed innovations, suggesting the relatively frequent occurrence of extreme events.

For future research, it is worthwhile to investigate if other macroeconomic variables provide additional information about the output gap. One promising approach is to embed the proposed trivariate system in a large Bayesian vector autoregression such as those in Banbura, Giannone, and Reichlin (2010) and Koop (2013). Using this framework, we can also assess the impacts of the output gap on other macroeconomic variables.

Appendix: Estimation Details

This appendix describes the priors and provides the estimation details of the proposed unobserved components model. For convenience we reproduce the trivariate UC model below:

$$\pi_t - \pi_t^* = \rho^\pi(\pi_{t-1} - \pi_{t-1}^*) + \delta_{\lambda^u}\lambda^u(u_t - u_t^*) + \delta_{\lambda^y}\lambda^y(y_t - y_t^*) + \varepsilon_t^\pi, \tag{13}$$

$$y_t - y_t^* = \delta_{\gamma^u}\gamma^u(u_t - u_t^*) + c_t, \tag{14}$$

$$u_t - u_t^* = \rho_1^u(u_{t-1} - u_{t-1}^*) + \rho_2^u(u_{t-2} - u_{t-2}^*) + \varepsilon_t^u, \tag{15}$$

$$c_t = \rho_1^c c_{t-1} + \rho_2^c c_{t-2} + \varepsilon_t^c, \tag{16}$$

where the initial conditions are given as $c_0 = c_{-1} = 0$, $\pi_0^* = \pi_0$, $u_0^* = u_0$, $u_{-1}^* = u_{-1}$ and $y_0^* \sim \mathcal{N}(y_{0,0}^*, V_{y_0^*})$. The error terms are drawn from either the Gaussian distribution or the t distribution: $\varepsilon_t^\pi \sim \mathcal{N}(0, z_t^\pi e^{h_t^\pi})$ with $z_t^\pi \sim \mathcal{IG}(\nu_\pi/2, \nu_\pi/2)$ if $\delta_{z^\pi} = 1$; otherwise we have $\varepsilon_t^\pi \sim \mathcal{N}(0, e^{h_t^\pi})$ instead. Similarly, for $\delta_{z^c} = 1$, $\varepsilon_t^c \sim \mathcal{N}(0, z_t^c e^{h_t^c})$ with $z_t^c \sim \mathcal{IG}(\nu_c/2, \nu_c/2)$; if $\delta_{z^c} = 0$, $\varepsilon_t^c \sim \mathcal{N}(0, e^{h_t^c})$. Finally, $\varepsilon_t^u \sim \mathcal{N}(0, z_t^u e^{h_t^u})$ with $z_t^u \sim \mathcal{IG}(\nu_u/2, \nu_u/2)$ if $\delta_{z^u} = 1$; otherwise $\varepsilon_t^u \sim \mathcal{N}(0, e^{h_t^u})$.

We consider the following noncentered parameterization of the trend output y_t^*:

$$y_t^* = y_0^* + \delta_{a_0}a_0 t + \delta_{y^*}\sigma_{y^*}\widetilde{y}_t^* + \delta_a\sigma_a\sum_{i=1}^{t}\widetilde{a}_i + \delta_{d_1}d_1 tD_{1973,t} + \delta_{d_2}d_2 tD_{2007,t}, \tag{17}$$

where

$$\widetilde{y}_t^* = \widetilde{y}_{t-1}^* + \varepsilon_t^{\widetilde{y}^*}, \tag{18}$$

$$\widetilde{a}_t^* = \widetilde{a}_{t-1}^* + \varepsilon_t^{\widetilde{a}^*}. \tag{19}$$

The state equations are initialized with $\widetilde{y}_0^* = \widetilde{a}_0^* = 0$ and the innovations have standard normal distributions: $\varepsilon_t^{\widetilde{y}^*} \sim \mathcal{N}(0,1)$ and $\varepsilon_t^{\widetilde{a}^*} \sim \mathcal{N}(0,1)$. The trend inflation π_t^* and NAIRU u_t^* follow random walk processes:

$$\pi_t^* = \pi_{t-1}^* + \varepsilon_t^{\pi^*}, \tag{20}$$

$$u_t^* = u_{t-1}^* + \varepsilon_t^{u^*}, \tag{21}$$

where $\varepsilon_t^{\pi^*} \sim \mathcal{N}(0, \sigma_{\pi^*}^2)$ and $\varepsilon_t^{u^*} \sim \mathcal{N}(0, \sigma_{u^*}^2)$, and the state equations are initialized with $\pi_1^* \sim (\pi_0^*, V_{\pi^*})$ and $u_1^* \sim (u_0^*, V_{u^*})$.

Finally, the log-volatilities follow random walk processes:

$$h_t^{\pi} = h_{t-1}^{\pi} + \varepsilon_t^{h^{\pi}}, \tag{22}$$

$$h_t^c = h_{t-1}^c + \varepsilon_t^{h^c}, \tag{23}$$

$$h_t^u = h_{t-1}^u + \varepsilon_t^{h^u}, \tag{24}$$

where $\varepsilon_t^{h^{\pi}} \sim \mathcal{N}(0, \sigma_{h^{\pi}}^2)$, $\varepsilon_t^{h^c} \sim \mathcal{N}(0, \sigma_{h^c}^2)$ and $\varepsilon_t^{h^u} \sim \mathcal{N}(0, \sigma_{h^u}^2)$. These are initialized as $h_1^{\pi} \sim (h_0^{\pi}, V_{h^{\pi}})$, $h_1^c \sim (h_0^c, V_{h^c})$ and $h_1^u \sim (h_0^u, V_{h^u})$ with $h_0^{\pi} = h_0^c = h_0^u = 0$ and $V_{h^{\pi}} = V_{h^c} = V_{h^u} = 100$.

We assume proper but relatively noninformative priors for the model parameters. In particular, we assume independent normal priors for the coefficients: $\rho^{\pi} \sim \mathcal{N}(\rho_0^{\pi}, V_{\rho^{\pi}})$, $\lambda^u \sim \mathcal{N}(\lambda_0^u, V_{\lambda^u})$, $\lambda^y \sim \mathcal{N}(\lambda_0^y, V_{\lambda^y})$, $\gamma^u \sim \mathcal{N}(\gamma_0^u, V_{\gamma^u})$, $a_0 \sim \mathcal{N}(a_{0,0}, V_{a_0})$, $d_1 \sim \mathcal{N}(d_{1,0}, V_{d_1})$, $d_2 \sim \mathcal{N}(d_{2,0}, V_{d_2})$, $\boldsymbol{\rho}^c \sim \mathcal{N}(\boldsymbol{\rho}_0^c, \mathbf{V}_{\boldsymbol{\rho}^c})1(\boldsymbol{\rho}^c \in \mathbf{R})$ and $\boldsymbol{\rho}^u \sim \mathcal{N}(\boldsymbol{\rho}_0^u, \mathbf{V}_{\boldsymbol{\rho}^u})1(\boldsymbol{\rho}^u \in \mathbf{R})$, where $\mathbf{R}$ is the stationarity region. We set all prior means to be zero and set the prior variances to be relatively large, with $V_{\rho^{\pi}} = V_{\lambda^u} = V_{\lambda^y} = V_{\gamma^u} = V_{a_0} = V_{d_1} = V_{d_2} = 1$ and $\mathbf{V}_{\boldsymbol{\rho}^c} = \mathbf{V}_{\boldsymbol{\rho}^u} = \mathbf{I}_2$.

For each of the binary indicators δ_i, we assume an independent Bernoulli prior with success probability p_i, where $i = \lambda^u, \gamma^u, \lambda^y, a_0, y^*, a, d_1, d_2, z^{\pi}, z^c, z^u$. The priors for the degree of freedom parameters are uniform: $\nu_{\pi} \sim \mathcal{U}(2, \bar{b}_{\nu_{\pi}})$, $\nu_c \sim \mathcal{U}(2, \bar{b}_{\nu_c})$ and $\nu_u \sim \mathcal{U}(2, \bar{b}_{\nu_u})$, where we set $\bar{b}_{\nu_{\pi}} = \bar{b}_{\nu_c} = \bar{b}_u = 50$. The lower bounds are set at 2 to ensure the first two moments of the t distribution exist. The upper bounds are set at 30 since a t distribution with degree of freedom larger than 50 is essentially indistinguishable from normal for our purpose.

Following Frühwirth-Schnatter and Wagner (2010), we assume gamma priors for the error variances in the state equations, or equivalently normal prior with zero mean for the (unsigned) standard deviations: $\sigma_{y^*} \sim \mathcal{N}(0, V_{\sigma_{y^*}})$, $\sigma_a \sim \mathcal{N}(0, V_{\sigma_a})$,

$$\sigma_{u^*}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2V_{\sigma_{u^*}}}\right), \sigma_{\pi^*}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2V_{\sigma_{\pi^*}}}\right), \sigma_{h^{\pi}}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2V_{\sigma_{h^{\pi}}}}\right),$$

$$\sigma_{h^c}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2V_{\sigma_{h^c}}}\right), \sigma_{h^u}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2V_{\sigma_{h^u}}}\right).$$

We set $V_{\sigma_{\pi^*}} = 0.2^2$, $V_{\sigma_a} = 0.01^2$, $V_{\sigma_{u^*}} = V_{\sigma_{y^*}} = V_{\sigma_{h^{\pi}}} = V_{\sigma_{h^c}} = V_{\sigma_{h^u}} = 0.1^2$. These priors imply that the prior means are $\mathbb{E}\sigma_{\pi^*}^2 = 0.2^2$, $\mathbb{E}\sigma_a^2 = 0.01^2$, $\mathbb{E}\sigma_{u^*}^2 = \mathbb{E}\sigma_{y^*}^2 = \mathbb{E}\sigma_{h^{\pi}}^2 = \mathbb{E}\sigma_{h^c}^2 = \mathbb{E}\sigma_{h^u}^2 = 0.1^2$.

Posterior draws can be obtained by following the steps below:

1. **Sample $\boldsymbol{\pi}^*$ from its full conditional distribution**. Since (13) and (20) define a standard linear Gaussian state space model, standard algorithms can be applied to sample $\boldsymbol{\pi}^*$. To derive the conditional density, first write (13) as

$$\mathbf{H}_{\rho^\pi}\boldsymbol{\pi} = \mathbf{H}_{\rho^\pi}\boldsymbol{\pi}^* + \delta_{\lambda^u}\lambda^u(\mathbf{u} - \mathbf{u}^*) + \delta_{\lambda^y}\lambda^y(\mathbf{y} - \mathbf{y}^*) + \boldsymbol{\varepsilon}^\pi,$$

where

$$\mathbf{H}_{\rho^\pi} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\rho^\pi & 1 & 0 & \cdots & 0 \\ 0 & -\rho^\pi & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -\rho^\pi & 1 \end{pmatrix},$$

and $\boldsymbol{\varepsilon}^\pi \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_\pi)$. If $\delta_{z^\pi} = 1$, $\mathbf{S}_\pi = \operatorname{diag}(z_1^\pi \mathrm{e}^{h_1^\pi}, \ldots, z_T^\pi \mathrm{e}^{h_T^\pi})$; otherwise $\mathbf{S}_\pi = \operatorname{diag}(\mathrm{e}^{h_1^\pi}, \ldots, \mathrm{e}^{h_T^\pi})$.

Since $|\mathbf{H}_{\rho^\pi}| = 1$ for any ρ^π, the log conditional likelihood for $\boldsymbol{\pi}$ is given by (ignoring terms not involving $\boldsymbol{\pi}^*$):

$$-\frac{1}{2}(\boldsymbol{\pi} - \boldsymbol{\pi}^* - \mathbf{m}_{\pi^*})'\mathbf{H}_{\rho^\pi}'\mathbf{S}_\pi^{-1}\mathbf{H}_{\rho^\pi}(\boldsymbol{\pi} - \boldsymbol{\pi}^* - \mathbf{m}_{\pi^*}),$$

where $\mathbf{m}_{\pi^*} = \mathbf{H}_{\rho^\pi}^{-1}(\delta_{\lambda^u}\lambda^u(\mathbf{u} - \mathbf{u}^*) + \delta_{\lambda^y}\lambda^y(\mathbf{y} - \mathbf{y}^*))$ Next, rewrite (20) as

$$\mathbf{H}\boldsymbol{\pi}^* = \widetilde{\boldsymbol{\mu}}_{\pi^*} + \boldsymbol{\varepsilon}^{\pi^*},$$

where $\boldsymbol{\varepsilon}^{\pi^*} \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_{\pi^*})$ with $\mathbf{S}_{\pi^*} = \operatorname{diag}(V_{\pi^*}, \sigma_{\pi^*}^2, \ldots, \sigma_{\pi^*}^2)$, $\widetilde{\boldsymbol{\mu}}_{\pi^*} = (\pi_0^*, 0, \ldots, 0)'$, and

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

Since $|\mathbf{H}| = 1$, the log prior density of $\boldsymbol{\pi}^*$ is (again ignoring terms not involving $\boldsymbol{\pi}^*$):

$$-\frac{1}{2}(\boldsymbol{\pi}^* - \boldsymbol{\mu}_{\pi^*})'\mathbf{H}'\mathbf{S}_{\pi^*}^{-1}\mathbf{H}(\boldsymbol{\pi}^* - \boldsymbol{\mu}_{\pi^*}),$$

where $\boldsymbol{\mu}_{\pi^*} = \mathbf{H}^{-1}\widetilde{\boldsymbol{\mu}}_{\pi^*}$. Then, by standard linear regression results (see, e.g., Kroese and

Chan, 2014, p.237-240), the conditional distribution of $\boldsymbol{\pi}^*$ is $\mathcal{N}(\widehat{\boldsymbol{\pi}}^*, \mathbf{K}_{\boldsymbol{\pi}^*}^{-1})$, where

$$\mathbf{K}_{\boldsymbol{\pi}^*} = \mathbf{H}'\mathbf{S}_{\pi^*}^{-1}\mathbf{H} + \mathbf{H}'_{\rho^{\pi}}\mathbf{S}_{\pi}^{-1}\mathbf{H}_{\rho^{\pi}}, \quad \widehat{\boldsymbol{\pi}}^* = \mathbf{K}_{\boldsymbol{\pi}^*}^{-1}\left(\mathbf{H}'\mathbf{S}_{\pi^*}^{-1}\mathbf{H}\boldsymbol{\mu}_{\pi^*} + \mathbf{H}'_{\rho^{\pi}}\mathbf{S}_{\pi}^{-1}\mathbf{H}_{\rho^{\pi}}(\boldsymbol{\pi} - \mathbf{m}_{\pi^*})\right).$$

Since $\mathbf{H}$ and $\mathbf{H}_{\rho^*}$ are band matrices, so is the precision matrix $\mathbf{K}_{\boldsymbol{\pi}^*}$. As such, the precision sampler of Chan and Jeliazkov (2009) can be used to sample $\boldsymbol{\pi}^*$ efficiently.

2. Sample $\mathbf{u}^*$ from its full conditional distribution. Information of $\mathbf{u}^*$ comes from four sources: the measurement equations (13), (14) and (15), as well as the state equation (21). As before, if $\delta_{z^c} = 1$, define $\mathbf{S}_c = \text{diag}(z_1^c e^{h_1^c}, \dots, z_T^c e^{h_T^c})$; otherwise let $\mathbf{S}_c = \text{diag}(e^{h_1^c}, \dots, e^{h_T^c})$. Moreover, set $\mathbf{S}_u = \text{diag}(z_1^u e^{h_1^u}, \dots, z_T^u e^{h_T^u})$ if $\delta_{z^u} = 1$; otherwise let $\mathbf{S}_u = \text{diag}(e^{h_1^u}, \dots, e^{h_T^u})$. Then, rewrite these equations in matrix form:

$$\mathbf{H}_{\rho^{\pi}}(\boldsymbol{\pi} - \boldsymbol{\pi}^*) = \delta_{\lambda^y}\lambda^y(\mathbf{y} - \mathbf{y}^*) + \delta_{\lambda^u}\lambda^u(\mathbf{u} - \mathbf{u}^*) + \boldsymbol{\varepsilon}^{\pi},$$
$$\mathbf{y} - \mathbf{y}^* = \delta_{\gamma^u}\gamma^u(\mathbf{u} - \mathbf{u}^*) + \mathbf{c},$$
$$\mathbf{H}_{\rho^u}(\mathbf{u} - \mathbf{u}^*) = \boldsymbol{\varepsilon}^u,$$
$$\mathbf{H}_{\rho^c}\mathbf{c} = \boldsymbol{\varepsilon}^c,$$
$$\mathbf{H}\mathbf{u}^* = \mathbf{H}\boldsymbol{\mu}_{u^*} + \boldsymbol{\varepsilon}^{u^*},$$

where $\boldsymbol{\varepsilon}^c \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_c)$, $\boldsymbol{\varepsilon}^u \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_u)$, $\boldsymbol{\varepsilon}^{u^*} \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_{u^*})$ with $\mathbf{S}_{u^*} = \text{diag}(V_{u^*}, \sigma_{u^*}^2, \dots, \sigma_{u^*}^2)$, $\boldsymbol{\mu}_{u^*} = \mathbf{H}^{-1}(u_0^*, 0, \dots, 0)'$, and

$$\mathbf{H}_{\rho^u} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -\rho_1^u & 1 & 0 & 0 & \cdots & 0 \\ -\rho_2^u & -\rho_1^u & 1 & 0 & \cdots & 0 \\ 0 & -\rho_2^u & -\rho_1^u & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\rho_2^u & -\rho_1^u & 1 \end{pmatrix} \quad \mathbf{H}_{\rho^c} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -\rho_1^c & 1 & 0 & 0 & \cdots & 0 \\ -\rho_2^c & -\rho_1^c & 1 & 0 & \cdots & 0 \\ 0 & -\rho_2^c & -\rho_1^c & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\rho_2^c & -\rho_1^c & 1 \end{pmatrix}.$$

Using a similar derivation as above, it can be shown that the conditional distribution of $\mathbf{u}^*$ is $\mathcal{N}(\widehat{\mathbf{u}}^*, \mathbf{K}_{\mathbf{u}^*}^{-1})$, where

$$\mathbf{K}_{\mathbf{u}^*} = \mathbf{H}'\mathbf{S}_{u^*}^{-1}\mathbf{H} + \mathbf{H}'_{\rho^u}\mathbf{S}_u^{-1}\mathbf{H}_{\rho^u} + (\delta_{\gamma^u}\gamma^u)^2\mathbf{H}'_{\rho^c}\mathbf{S}_c^{-1}\mathbf{H}_{\rho^c} + (\delta_{\lambda^u}\lambda^u)^2\mathbf{S}_{\pi}^{-1},$$
$$\widehat{\mathbf{u}}^* = \mathbf{K}_{\mathbf{u}^*}^{-1}\left(\mathbf{H}'\mathbf{S}_{u^*}^{-1}\mathbf{H}\boldsymbol{\mu}_{u^*} + \mathbf{H}'_{\rho^u}\mathbf{S}_u^{-1}\mathbf{H}_{\rho^u}\mathbf{u} - \delta_{\gamma^u}\gamma^u\mathbf{H}'_{\rho^c}\mathbf{S}_c^{-1}\mathbf{H}_{\rho^c}(\mathbf{y} - \mathbf{y}^* - \delta_{\gamma^u}\gamma^u\mathbf{u})\right.$$
$$\left. -\delta_{\lambda^u}\lambda^u\mathbf{S}_{\pi}^{-1}(\mathbf{H}_{\rho^{\pi}}(\boldsymbol{\pi} - \boldsymbol{\pi}^*) - \delta_{\lambda^y}\lambda^y(\mathbf{y} - \mathbf{y}^*) - \delta_{\lambda^u}\lambda^u\mathbf{u})\right).$$

Once again the precision matrix $\mathbf{K}_{\mathbf{u}^*}$ is banded and the precision sampler of Chan and

Jeliazkov (2009) can be used to sample $\mathbf{u}^*$ efficiently.

3. Sample $\widetilde{\mathbf{y}}^*$ and $(y_0^*, a_0, \sigma_{y^*}, \sigma_a, d_1, d_2)$ from their full conditional distributions, and randomly permute the signs of $\widetilde{\mathbf{y}}^*$ and σ_{y^*}. First, we derive the full conditional distribution of $\widetilde{\mathbf{y}}^*$. To that end, rewrite (17) in the following matrix notations:

$$\mathbf{y}^* = y_0^* \mathbf{1}_T + \delta_{a_0} a_0 \mathbf{1}_{1:T} + \delta_{y^*} \sigma_{y^*} \widetilde{\mathbf{y}} + \delta_a \sigma_a \widetilde{\mathbf{A}} + \delta_{d_1} d_1 \widetilde{\mathbf{D}}_{1973} + \delta_{d_2} d_2 \widetilde{\mathbf{D}}_{2007}, \qquad (25)$$

where $\mathbf{1}_T$ is a $T \times 1$ column of ones, $\mathbf{1}_{1:T} = (1, 2, 3, \ldots, T)'$ and $\widetilde{\mathbf{A}} = (\widetilde{a}_1, \widetilde{a}_1 + \widetilde{a}_2, \ldots, \sum_{t=1}^{T} \widetilde{a}_t)'$. Let $\odot$ denote the component-wise multiplication. Then, $\widetilde{\mathbf{D}}_{1973} = \mathbf{D}_{1973} \odot \mathbf{1}_{1:T}$ and $\widetilde{\mathbf{D}}_{2007} = \mathbf{D}_{2007} \odot \mathbf{1}_{1:T}$.

Next, rewrite (13) and (14) in terms of $\widetilde{\mathbf{y}}$:

$$\mathbf{H}_{\rho^\pi} \boldsymbol{\pi} = \mathbf{H}_{\rho^\pi} \boldsymbol{\pi}^* + \delta_{\lambda^u} \lambda^u (\mathbf{u} - \mathbf{u}^*) + \delta_{\lambda^y} \lambda^y (\mathbf{y} - \mathbf{m}_{\widetilde{y}^*} - \delta_{y^*} \sigma_{y^*} \widetilde{\mathbf{y}}) + \boldsymbol{\varepsilon}^\pi,$$

$$\mathbf{y} = \mathbf{m}_{\widetilde{y}^*} + \delta_{y^*} \sigma_{y^*} \widetilde{\mathbf{y}} + \delta_{\gamma^u} \gamma^u (\mathbf{u} - \mathbf{u}^*) + \mathbf{c},$$

where $\mathbf{m}_{\widetilde{y}^*} = y_0^* \mathbf{1}_T + \delta_{a_0} a_0 \mathbf{1}_{1:T} + \delta_a \sigma_a \widetilde{\mathbf{A}} + \delta_{d_1} d_1 \widetilde{\mathbf{D}}_{1973} + \delta_{d_2} d_2 \widetilde{\mathbf{D}}_{2007}$, $\boldsymbol{\varepsilon}^\pi \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_\pi)$ and $\mathbf{c} \sim \mathcal{N}(\mathbf{0}, (\mathbf{H}_{\rho^c}' \mathbf{S}_c^{-1} \mathbf{H}_{\rho^c})^{-1})$. The state equation (18) can be written as

$$\mathbf{H} \widetilde{\mathbf{y}} = \boldsymbol{\varepsilon}^{\widetilde{y}^*},$$

where $\boldsymbol{\varepsilon}^{\widetilde{y}^*} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_T)$. Hence, using a similar derivation as above, the full conditional distribution of $\widetilde{\mathbf{y}}^*$ is $\mathcal{N}(\widehat{\widetilde{\mathbf{y}}}^*, \mathbf{K}_{\widetilde{\mathbf{y}}^*}^{-1})$, where

$$\mathbf{K}_{\widetilde{\mathbf{y}}^*} = \mathbf{H}'\mathbf{H} + (\delta_{\lambda^y} \lambda^y \delta_{y^*} \sigma_{y^*})^2 \mathbf{S}_\pi^{-1} + (\delta_{y^*} \sigma_{y^*})^2 \mathbf{H}_{\rho^c}' \mathbf{S}_c^{-1} \mathbf{H}_{\rho^c},$$

$$\widehat{\widetilde{\mathbf{y}}}^* = \mathbf{K}_{\widetilde{\mathbf{y}}^*}^{-1} \left(-\delta_{\lambda^y} \lambda^y \delta_{y^*} \sigma_{y^*} \mathbf{S}_\pi^{-1} (\mathbf{H}_{\rho^\pi} (\boldsymbol{\pi} - \boldsymbol{\pi}^*) - \delta_{\lambda^y} \lambda^y (\mathbf{y} - \mathbf{m}_{\widetilde{y}^*}) - \delta_{\lambda^u} \lambda^u (\mathbf{u} - \mathbf{u}^*)) \right.$$

$$\left. + \delta_{y^*} \sigma_{y^*} \mathbf{H}_{\rho^c}' \mathbf{S}_c^{-1} \mathbf{H}_{\rho^c} (\mathbf{y} - \mathbf{m}_{\widetilde{y}^*} - \delta_{\gamma^u} \gamma^u (\mathbf{u} - \mathbf{u}^*)) \right).$$

Again the precision matrix $\mathbf{K}_{\widetilde{\mathbf{y}}^*}$ is banded and we sample $\widetilde{\mathbf{y}}^*$ using the algorithm in Chan and Jeliazkov (2009).

Now, we derive the conditional distribution of $\boldsymbol{\beta}_{y^*} = (y_0^*, a_0, \sigma_{y^*}, \sigma_a, d_1, d_2)'$. First, note that the prior on $\boldsymbol{\beta}_{y^*}$ is given by $\boldsymbol{\beta}_{y^*} \sim \mathcal{N}(\boldsymbol{\beta}_{y^*,0}, \mathbf{V}_{\boldsymbol{\beta}_{y^*}})$, where $\boldsymbol{\beta}_{y^*,0} = (y_{0,0}^*, a_{0,0}, 0, 0, d_{1,0}, d_{2,0})'$ and $\mathbf{V}_{\boldsymbol{\beta}_{y^*}} = \text{diag}(V_{y^*}, V_{a_0}, V_{\sigma_{y^*}}, V_{\sigma_a}, V_{d_1}, V_{d_2})$. Next, we rewrite (25) as $\mathbf{y}^* = \mathbf{X}_{y^*} \boldsymbol{\beta}_{y^*}$, where

$\mathbf{X}_{y^*} = (\mathbf{1}_T, \delta_{a_0}\mathbf{1}_{1:T}, \delta_{y^*}\widetilde{\mathbf{y}}, \delta_a\widetilde{\mathbf{A}}, \delta_{d_1}\widetilde{\mathbf{D}}_{1973} + \delta_{d_2}\widetilde{\mathbf{D}}_{2007})$. Then, (13) and (14) become:

$$\mathbf{H}_{\rho^\pi}(\boldsymbol{\pi} - \boldsymbol{\pi}^*) = \delta_{\lambda^y}\lambda^y(\mathbf{y} - \mathbf{X}_{y^*}\boldsymbol{\beta}_{y^*}) + \delta_{\lambda^u}\lambda^u(\mathbf{u} - \mathbf{u}^*) + \boldsymbol{\varepsilon}^\pi,$$
$$\mathbf{y} = \mathbf{X}_{y^*}\boldsymbol{\beta}_{y^*} + \delta_{\gamma^u}\gamma^u(\mathbf{u} - \mathbf{u}^*) + \mathbf{c}.$$

Using standard results from linear regression, the full conditional distribution of $\boldsymbol{\beta}_{y^*}$ is $\mathcal{N}(\widehat{\boldsymbol{\beta}}_{y^*}, \mathbf{K}_{\boldsymbol{\beta}_{y^*}}^{-1})$, where

$$\mathbf{K}_{\boldsymbol{\beta}_{y^*}} = \mathbf{V}_{\boldsymbol{\beta}_{y^*}}^{-1} + (\delta_{\lambda^y}\lambda^y)^2\mathbf{X}'_{y^*}\mathbf{S}_\pi^{-1}\mathbf{X}_{y^*} + \mathbf{X}'_{y^*}\mathbf{H}'_{\boldsymbol{\rho}^c}\mathbf{S}_c^{-1}\mathbf{H}_{\boldsymbol{\rho}^c}\mathbf{X}_{y^*},$$
$$\widehat{\boldsymbol{\beta}}_{y^*} = \mathbf{K}_{\boldsymbol{\beta}_{y^*}}^{-1}\left(\mathbf{V}_{\boldsymbol{\beta}_{y^*}}^{-1}\boldsymbol{\beta}_{y^*,0} - \delta_{\lambda^y}\lambda^y\mathbf{X}'_{y^*}\mathbf{S}_\pi^{-1}(\mathbf{H}_{\rho^\pi}(\boldsymbol{\pi} - \boldsymbol{\pi}^*) - \delta_{\lambda^y}\lambda^y\mathbf{y} - \delta_{\lambda^u}\lambda^u(\mathbf{u} - \mathbf{u}^*))\right.$$
$$\left. + \mathbf{X}'_{y^*}\mathbf{H}'_{\boldsymbol{\rho}^c}\mathbf{S}_c^{-1}\mathbf{H}_{\boldsymbol{\rho}^c}(\mathbf{y} - \delta_{\gamma^u}\gamma^u(\mathbf{u} - \mathbf{u}^*)))\right).$$

As discussed in Frühwirth-Schnatter and Wagner (2010), the signs of $\widetilde{y}^*$ and σ_{y^*} are not identifiable. To improve the efficiency of the sampler, we randomly permute the signs of $\widetilde{y}^*$ and σ_{y^*}. Specifically, let U be a discrete random variable that takes values in $\{-1, 1\}$ with equal probabilities. Given the current draws $\widetilde{y}^*$ and σ_{y^*}, we return $U\widetilde{y}^*$ and $U\sigma_{y^*}$.

4. Sample $\widetilde{\mathbf{a}}$ and $(y_0^*, a_0, \sigma_{y^*}, \sigma_a, d_1, d_2)$ from their full conditional distributions, and randomly permute the signs of $\widetilde{\mathbf{a}}$ and σ_a. Since $\mathbf{H}\widetilde{\mathbf{A}} = \widetilde{\mathbf{a}}$, we can sample $\widetilde{\mathbf{A}}$ and transform the draw to get $\widetilde{\mathbf{a}}$. Now, it follows from (25) that

$$\mathbf{y}^* = \mathbf{m}_{\widetilde{A}} + \delta_a\sigma_a\widetilde{\mathbf{A}},$$

where $\mathbf{m}_{\widetilde{A}} = y_0^*\mathbf{1}_T + \delta_{a_0}a_0\mathbf{1}_{1:T} + \delta_{y^*}\sigma_{y^*}\widetilde{\mathbf{y}} + \delta_{d_1}d_1\widetilde{\mathbf{D}}_{1973} + \delta_{d_2}d_2\widetilde{\mathbf{D}}_{2007}$. Next, rewrite (13) and (14) in terms of $\widetilde{\mathbf{A}}$ as follows:

$$\mathbf{H}_{\rho^\pi}(\boldsymbol{\pi} - \boldsymbol{\pi}^*) = \delta_{\lambda^u}\lambda^u(\mathbf{u} - \mathbf{u}^*) + \delta_{\lambda^y}\lambda^y(\mathbf{y} - \mathbf{m}_{\widetilde{A}} - \delta_a\sigma_a\widetilde{\mathbf{A}}) + \boldsymbol{\varepsilon}^\pi,$$
$$\mathbf{y} = \mathbf{m}_{\widetilde{A}} + \delta_a\sigma_a\widetilde{\mathbf{A}} + \delta_{\gamma^u}\gamma^u(\mathbf{u} - \mathbf{u}^*) + \mathbf{c},$$

where $\boldsymbol{\varepsilon}^\pi \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_\pi)$ and $\mathbf{c} \sim \mathcal{N}(\mathbf{0}, (\mathbf{H}'_{\boldsymbol{\rho}^c}\mathbf{S}_c^{-1}\mathbf{H}_{\boldsymbol{\rho}^c})^{-1})$. The state equation for $\widetilde{\mathbf{a}}$ is given by

$$\mathbf{H}\widetilde{\mathbf{a}} = \boldsymbol{\varepsilon}^{\widetilde{a}}, \quad \boldsymbol{\varepsilon}^{\widetilde{a}} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_T),$$

which implies that $\widetilde{\mathbf{A}} = \mathbf{H}^{-1}\widetilde{\mathbf{a}} \sim \mathcal{N}(\mathbf{0}, ((\mathbf{H}^2)'\mathbf{H}^2)^{-1})$. Following a similar derivation as

before, one can show that the full conditional distribution of $\widetilde{\mathbf{A}}$ is $\mathcal{N}(\widehat{\widetilde{\mathbf{A}}}, \mathbf{K}_{\widetilde{\mathbf{A}}}^{-1})$, where

$$\mathbf{K}_{\widetilde{\mathbf{A}}} = (\mathbf{H}^2)'\mathbf{H}^2 + (\delta_{\lambda^y}\lambda^y\delta_a\sigma_a)^2\mathbf{S}_\pi^{-1} + (\delta_a\sigma_a)^2\mathbf{H}_{\rho^c}'\mathbf{S}_c^{-1}\mathbf{H}_{\rho^c},$$

$$\widehat{\widetilde{\mathbf{A}}} = \mathbf{K}_{\widetilde{\mathbf{A}}}^{-1}\left(-\delta_{\lambda^y}\lambda^y\delta_a\sigma_a\mathbf{S}_\pi^{-1}(\mathbf{H}_{\rho^\pi}(\boldsymbol{\pi} - \boldsymbol{\pi}^*) - \delta_{\lambda^y}\lambda^y(\mathbf{y} - \mathbf{m}_{\widetilde{A}}) - \delta_{\lambda^u}\lambda^u(\mathbf{u} - \mathbf{u}^*))\right.$$
$$\left. + \delta_a\sigma_a\mathbf{H}_{\rho^c}'\mathbf{S}_c^{-1}\mathbf{H}_{\rho^c}(\mathbf{y} - \mathbf{m}_{\widetilde{A}} - \delta_{\gamma^u}\gamma^u(\mathbf{u} - \mathbf{u}^*)))\right).$$

Since $\mathbf{K}_{\widetilde{\mathbf{A}}}$ is again a band matrix, we can sample $\widetilde{\mathbf{A}}$ efficiently using the algorithm in Chan and Jeliazkov (2009). Then, $(y_0^*, a_0, \sigma_{y^*}, \sigma_a, d_1, d_2)$ can be sampled exactly as before. Finally, we randomly permute the signs of $\widetilde{\mathbf{a}}$ and σ_a.

5. Sample $(\delta_{\lambda^u}, \delta_{\lambda^y}, \delta_{\gamma^u})$ **and** $(\rho^\pi, \lambda^u, \lambda^y, \gamma^u)$ **jointly**. This is achieved by first drawing the indicators marginally of $(\rho^\pi, \lambda^u, \lambda^y, \gamma^u)$, followed by sampling the latter from their joint conditional distribution.

We first derive the full conditional distribution of $(\rho^\pi, \lambda^u, \lambda^y, \gamma^u)$. Note that $\boldsymbol{\beta}_\pi = (\rho^\pi, \lambda^u, \lambda^y)'$ and γ^u are conditionally independent given other parameters and states, and we sample them in turn. To that end, rewrite (13) as

$$\boldsymbol{\pi} - \boldsymbol{\pi}^* = \mathbf{X}_\pi\boldsymbol{\beta}_\pi + \boldsymbol{\varepsilon}^\pi,$$

where $\boldsymbol{\varepsilon}^\pi \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_\pi)$ and

$$\mathbf{X}_\pi = \begin{pmatrix} 0 & \delta_{\lambda^u}(u_1 - u_1^*) & \delta_{\lambda^y}(y_1 - y_1^*) \\ \pi_1 - \pi_1^* & \delta_{\lambda^u}(u_2 - u_2^*) & \delta_{\lambda^y}(y_2 - y_2^*) \\ \vdots & \vdots & \vdots \\ \pi_{T-1} - \pi_{T-1}^* & \delta_{\lambda^u}(u_T - u_T^*) & \delta_{\lambda^y}(y_T - y_T^*). \end{pmatrix}$$

Recall that the prior for $\boldsymbol{\beta}_\pi$ is given by $\boldsymbol{\beta}_\pi \sim \mathcal{N}(\boldsymbol{\beta}_{\pi,0}, \mathbf{V}_{\boldsymbol{\beta}_\pi})$, where $\boldsymbol{\beta}_{\pi,0} = (\rho_0^\pi, \lambda_0^u, \lambda_0^y)'$ and $\mathbf{V}_{\boldsymbol{\beta}_\pi} = \text{diag}(V_{\rho^\pi}, V_{\lambda^u}, V_{\lambda^y})$. Hence, the full conditional distribution of $\boldsymbol{\beta}_\pi$ is $\mathcal{N}(\widehat{\boldsymbol{\beta}}_\pi, \mathbf{K}_{\boldsymbol{\beta}_\pi}^{-1})$, where

$$\mathbf{K}_{\boldsymbol{\beta}_\pi} = \mathbf{V}_{\boldsymbol{\beta}_\pi}^{-1} + \mathbf{X}_\pi'\mathbf{S}_\pi^{-1}\mathbf{X}_\pi, \quad \widehat{\boldsymbol{\beta}}_\pi = \mathbf{K}_{\boldsymbol{\beta}_\pi}^{-1}(\mathbf{V}_{\boldsymbol{\beta}_\pi}^{-1}\boldsymbol{\beta}_{\pi,0} + \mathbf{X}_\pi'\mathbf{S}_\pi^{-1}(\boldsymbol{\pi} - \boldsymbol{\pi}^*)).$$

Similarly, the full conditional distribution of γ^u is $\mathcal{N}(\widehat{\lambda}^y, K_{\gamma^u}^{-1})$, where

$$K_{\gamma^u} = V_{\gamma^u}^{-1} + \delta_{\gamma^u}^2(\mathbf{u} - \mathbf{u}^*)'\mathbf{H}_{\rho^c}'\mathbf{S}_c^{-1}\mathbf{H}_{\rho^c}(\mathbf{u} - \mathbf{u}^*),$$
$$\widehat{\lambda}^y = K_{\gamma^u}^{-1}(V_{\gamma^u}^{-1}\gamma_0^u + \delta_{\gamma^u}(\mathbf{u} - \mathbf{u}^*)'\mathbf{H}_{\rho^c}'\mathbf{S}_c^{-1}\mathbf{H}_{\rho^c}(\mathbf{y} - \mathbf{y}^*)).$$

Now, we derive the conditional distribution of the indicators $(\delta_{\lambda^u}, \delta_{\lambda^y}, \delta_{\gamma^u})'$ marginally of the coefficients $\boldsymbol{\beta}_\pi$ and γ^u. The prior is simply the product of Bernoulli distributions:

$$p(\delta_{\lambda^u}, \delta_{\lambda^y}, \delta_{\gamma^u}) = \prod_{i=\lambda^u,\lambda^y,\gamma^u} p_i^{\delta_i}(1-p_i)^{1-\delta_i}.$$

Next, using a similar derivation in Chan and Grant (2016), one can show that the joint density function of $\boldsymbol{\pi}$ and $\mathbf{y}$ marginally of $\boldsymbol{\beta}_\pi$ and γ^u is proportional to (omitting any constants not depending on the indicators):

$$p(\boldsymbol{\pi}, \mathbf{y} \mid \delta_{\lambda^u}, \delta_{\lambda^y}, \delta_{\gamma^u}, \cdot) \propto |\mathbf{K}_{\boldsymbol{\beta}_\pi}|^{-\frac{1}{2}} \mathrm{e}^{\frac{1}{2}\widehat{\boldsymbol{\beta}}_\pi' \mathbf{K}_{\boldsymbol{\beta}_\pi} \widehat{\boldsymbol{\beta}}_\pi} |K_{\gamma^u}|^{-\frac{1}{2}} \mathrm{e}^{\frac{1}{2}(\widehat{\lambda}^y)^2 K_{\gamma^u}}.$$

We can therefore evaluate the conditional distribution of the indicators marginally of $\boldsymbol{\beta}_\pi$ and γ^u. Finally, $(\delta_{\lambda^u}, \delta_{\lambda^y}, \delta_{\gamma^u})$ can be sampled using the inverse-transform method (see, e.g. Kroese and Chan, 2014, Algorithm 2.3).

6. Sample $(\delta_{a_0}, \delta_{y^*}, \delta_a, \delta_{d_1}, \delta_{d_2})$ **and** $(y_0^*, a_0, \sigma_{y^*}, \sigma_a, d_1, d_2)$ **jointly**. Again, this is achieved by first drawing the indicators marginally of $(y_0^*, a_0, \sigma_{y^*}, \sigma_a, d_1, d_2)$, followed by sampling the latter from their full conditional distribution.

The full conditional distribution of $\boldsymbol{\beta}_{y^*} = (y_0^*, a_0, \sigma_{y^*}, \sigma_a, d_1, d_2)'$ is $\mathcal{N}(\widehat{\boldsymbol{\beta}}_{y^*}, \mathbf{K}_{\boldsymbol{\beta}_{y^*}}^{-1})$, which is derived in Step 3. Now, using a similar derivation in Step 4, one can show that the distribution of $(\delta_{a_0}, \delta_{y^*}, \delta_a, \delta_{d_1}, \delta_{d_2})$ marginally of $\boldsymbol{\beta}_{y^*}$ is given by

$$p(\delta_{a_0}, \delta_{y^*}, \delta_a, \delta_{d_1}, \delta_{d_2} \mid \boldsymbol{\pi}, \mathbf{y}, \cdot) \propto \prod_{i=a_0,y^*,a,d_1,d_2} p_i^{\delta_i}(1-p_i)^{1-\delta_i} \times |\mathbf{K}_{\boldsymbol{\beta}_{y^*}}|^{-\frac{1}{2}} \mathrm{e}^{\frac{1}{2}\widehat{\boldsymbol{\beta}}_{y^*}' \mathbf{K}_{\boldsymbol{\beta}_{y^*}} \widehat{\boldsymbol{\beta}}_{y^*}}.$$

7. Sample $(\delta_{z^\pi}, \delta_{z^c}, \delta_{z^u})$, $\mathbf{z}^\pi$, $\mathbf{z}^c$ **and** $\mathbf{z}^u$ **jointly**: first sample the indicators marginally of $\mathbf{z}^\pi$, $\mathbf{z}^c$ and δ_{z^u}, followed by drawing $\mathbf{z}^\pi$, $\mathbf{z}^c$ and $\mathbf{z}^u$ from their full conditional distributions. By our assumptions, the joint conditional density of $\boldsymbol{\varepsilon}^\pi = (\varepsilon_1^\pi, \ldots, \varepsilon_T^\pi)'$ is given by:

$$p(\boldsymbol{\varepsilon}^\pi \mid \mathbf{h}^\pi, \delta_{z^\pi} = 0) = \prod_{t=1}^{T} (2\pi \mathrm{e}^{h_t^\pi})^{-\frac{1}{2}} \mathrm{e}^{-\frac{1}{2}\mathrm{e}^{-h_t^\pi}(\varepsilon_t^\pi)^2},$$

$$p(\boldsymbol{\varepsilon}^\pi \mid \mathbf{h}^\pi, \delta_{z^\pi} = 1) = \left(\frac{\Gamma(\frac{\nu_\pi+1}{2})}{\sqrt{\nu_\pi \pi} \Gamma(\frac{\nu_\pi}{2})}\right)^T \prod_{t=1}^{T} \mathrm{e}^{-\frac{1}{2}h_t^\pi} \left(1 + \frac{1}{\nu_\pi} \mathrm{e}^{-h_t^\pi}(\varepsilon_t^\pi)^2\right)^{-\frac{\nu_\pi+1}{2}}.$$

Hence, the conditional density of δ_{z^π} marginally of $\mathbf{z}^\pi$ is simply

$$p(\delta_{z^\pi} \mid \boldsymbol{\varepsilon}^\pi, \mathbf{h}^\pi) \propto p_{z^\pi}^{\delta_{z^\pi}}(1-p_{z^\pi})^{1-\delta_{z^\pi}} p(\boldsymbol{\varepsilon}^\pi \mid \mathbf{h}^\pi, \delta_{z^\pi}).$$

Similar expression can be derived for δ_{z^c} and δ_{z^u}. Next, given δ_{z^π} and other parameters, $z_1^\pi, \ldots, z_T^\pi$ are independent inverse-gamma random variables:

$$
\begin{aligned}
\left(z_t^\pi \mid \nu_\pi, \delta_{z^\pi} = 0\right) &\sim \mathcal{IG}\left(\frac{\nu_\pi}{2}, \frac{\nu_\pi}{2}\right), \\
\left(z_t^\pi \mid \nu_\pi, \delta_{z^\pi} = 1\right) &\sim \mathcal{IG}\left(\frac{\nu_\pi + 1}{2}, \frac{\nu_\pi + \mathrm{e}^{-h_t^\pi}(\varepsilon_t^\pi)^2}{2}\right).
\end{aligned}
$$

Finally, $z_1^c, \ldots, z_T^c$ as well as $z_1^u, \ldots, z_T^u$ can be sampled similarly.

8. Sample $\boldsymbol{\rho}^u$ and $\boldsymbol{\rho}^c$ from their full conditional distributions. Since $\boldsymbol{\rho}^u$ and $\boldsymbol{\rho}^c$ are conditionally independent given the latent states and other parameters, we can sample them jointly. First, let

$$
\mathbf{X}_{\boldsymbol{\rho}^u} = \begin{pmatrix} u_0 - u_0^* & u_{-1} - u_{-1}^* \\ u_1 - u_1^* & u_0 - u_0^* \\ \vdots & \vdots \\ u_{T-1} - u_{T-1}^* & u_{T-2} - u_{T-2}^* \end{pmatrix}, \quad \mathbf{X}_{\boldsymbol{\rho}^c} = \begin{pmatrix} c_0 & c_{-1} \\ c_1 & c_0 \\ \vdots & \vdots \\ c_{T-1} & c_{T-2} \end{pmatrix}.
$$

Then, we can rewrite (15) and (16) as:

$$
\begin{aligned}
\mathbf{u} - \mathbf{u}^* &= \mathbf{X}_{\boldsymbol{\rho}^u} \boldsymbol{\rho}^u + \boldsymbol{\varepsilon}^u, \\
\mathbf{c} &= \mathbf{X}_{\boldsymbol{\rho}^c} \boldsymbol{\rho}^c + \boldsymbol{\varepsilon}^c,
\end{aligned}
$$

where $\boldsymbol{\varepsilon}^u \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_u)$ and $\boldsymbol{\varepsilon}^c \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_c)$. Hence, the conditional distributions of $\boldsymbol{\rho}^u$ and $\boldsymbol{\rho}^c$ are, respectively, $\mathcal{N}(\widehat{\boldsymbol{\rho}}^u, \mathbf{K}_{\boldsymbol{\rho}^u}^{-1}) 1(\boldsymbol{\rho}^u \in \mathbf{R})$ and $\mathcal{N}(\widehat{\boldsymbol{\rho}}^c, \mathbf{K}_{\boldsymbol{\rho}^c}^{-1}) 1(\boldsymbol{\rho}^c \in \mathbf{R})$, where $\mathbf{R}$ is the stationarity region and

$$
\begin{aligned}
\mathbf{K}_{\boldsymbol{\rho}^u} &= \mathbf{V}_{\boldsymbol{\rho}^u}^{-1} + \mathbf{X}_{\boldsymbol{\rho}^u}' \mathbf{S}_u^{-1} \mathbf{X}_{\boldsymbol{\rho}^u}, & \widehat{\boldsymbol{\rho}}^u &= \mathbf{K}_{\boldsymbol{\rho}^u}^{-1}\left(\mathbf{V}_{\boldsymbol{\rho}^u}^{-1} \boldsymbol{\rho}_0^u + \mathbf{X}_{\boldsymbol{\rho}^u}' \mathbf{S}_u^{-1}(\mathbf{u} - \mathbf{u}^*)\right), \\
\mathbf{K}_{\boldsymbol{\rho}^c} &= \mathbf{V}_{\boldsymbol{\rho}^c}^{-1} + \mathbf{X}_{\boldsymbol{\rho}^c}' \mathbf{S}_c^{-1} \mathbf{X}_{\boldsymbol{\rho}^c}, & \widehat{\boldsymbol{\rho}}^c &= \mathbf{K}_{\boldsymbol{\rho}^c}^{-1}\left(\mathbf{V}_{\boldsymbol{\rho}^c}^{-1} \boldsymbol{\rho}_0^c + \mathbf{X}_{\boldsymbol{\rho}^u}' \mathbf{S}_c^{-1} \mathbf{c}\right).
\end{aligned}
$$

9. Sample $\mathbf{h}^\pi$, $\mathbf{h}^c$ and $\mathbf{h}^u$ from their full conditional distributions. The log-volatilities $\mathbf{h}^\pi$, $\mathbf{h}^c$ and $\mathbf{h}^u$ can be sampled easily using the auxiliary mixture sampler of Kim, Shepherd, and Chib (1998); see Chan and Hsiao (2014) for a textbook treatment. For example, if $\delta_{z^\pi} = 1$, define $w_t^\pi = \log((\varepsilon_t^\pi)^2 / z_t^\pi)$; otherwise, let $w_t^\pi = \log((\varepsilon_t^\pi)^2)$. Then, we feed $w_1^\pi, \ldots, w_T^\pi$ into the auxiliary mixture sampler as data.

10. Sample the variances $\sigma_{u^*}^2$, $\sigma_{\pi^*}^2$, $\sigma_{h^\pi}^2$, $\sigma_{h^c}^2$ and $\sigma_{h^u}^2$ from their respectively full conditional distributions. The variances are conditionally independent given the

states and other parameters. Hence, we can sample them individually. The conditional distributions of the variances are non-standard due to the gamma priors. However, each can be sampled using a Metropolis-Hastings step with an inverse-gamma proposal density. For example, to sample $\sigma_{u^*}^2$, we can first obtain a candidate draw s^2 from $\mathcal{IG}(T/2 - 1, \sum_{t=2}^{T}(u_t^* - u_{t-1}^*)^2/2)$. Given the current draw $\sigma_{u^*}^2$, we accept the candidate draw s^2 with probability

$$\min\left\{1, \exp\left(-\frac{1}{2V_{\sigma_{u^*}}}(s^2 - \sigma_{u^*}^2)\right)\right\};$$

otherwise, we keep σ_{u^*}. The other variances can be sampled similarly.

11. Sample ν_π, ν_c and ν_u .

The conditional distributions of ν_π, ν_c and ν_u are non-standard, but each can be sampled using a Metropolis-Hastings step. For example, to sample e ν_π, we first maximize $\log p(\nu_\pi \,|\, \mathbf{z}_\pi$ using the Newton-Raphson method to obtain the mode and the negative Hessian evaluated at the mode, denoted as $\widehat{\nu}_\pi$ and K_{ν_π}, respectively. Then, we implement an independence-chain Metropolis-Hastings step with proposal density $\mathcal{N}(\widehat{\nu}_\pi, K_{\nu_\pi}^{-1})$. For implementation details, see Chan and Hsiao (2014).

References

L. M. Ball, D. Leigh, and P. Loungani. Okun's law: fit at fifty? Technical report, National Bureau of Economic Research, 2013.

M. Banbura, D. Giannone, and L. Reichlin. Large Bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1):71–92, 2010.

A. Basistha. Trend-cycle correlation, drift break and the estimation of trend and cycle in Canadian GDP. *Canadian Journal of Economics*, 40(2):584–606, 2007.

A. Basistha and C. R. Nelson. New measures of the output gap based on the forward-looking new Keynesian Phillips curve. *Journal of Monetary Economics*, 54(2):498–511, 2007.

T. Berger and B. Kempa. Bayesian estimation of the output gap for a small open economy: The case of Canada. *Economics Letters*, 112(1):107–112, 2011.

T. Berger, G. Everaert, and H. Vierke. Testing for time variation in an unobserved components model for the U.S. economy. *Journal of Economic Dynamics and Control*, 2016. Forthcoming.

S. Beveridge and C. R. Nelson. 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, 1981.

J. C. C. Chan. Moving average stochastic volatility models with application to inflation forecast. *Journal of Econometrics*, 176(2):162–172, 2013.

J. C. C. Chan and E. Eisenstat. Bayesian model comparison for time-varying parameter VARs with stochastic volatility. *CAMA Working Paper*, 32/2015, 2015.

J. C. C. Chan and A. L. Grant. Fast computation of the deviance information criterion for latent variable models. *Computational Statistics and Data Analysis*, 100:847–859, 2016.

J. C. C. Chan and C. Y. L. Hsiao. Estimation of stochastic volatility models with heavy tails and serial dependence. In I. Jeliazkov and X.-S. Yang, editors, *Bayesian Inference in the Social Sciences*. John Wiley & Sons, Hoboken, 2014.

J. C. C. Chan and I. Jeliazkov. Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation*, 1(1/2):101–120, 2009.

J. C. C. Chan, G. Koop, and S. M. Potter. A bounded model of time variation in trend inflation, NAIRU and the Phillips curve. *Journal of Applied Econometrics*, 31 (3):551–565, 2016.

C. J. Chiu, H. Mumtaz, and G. Pinter. Forecasting with VAR models: Fat tails and stochastic volatility. Technical report, CReMFi, School of Economics and Finance, QMUL, 2015.

P. K. Clark. The cyclical component of US economic activity. *The Quarterly Journal of Economics*, 102(4):797–814, 1987.

T. E. Clark. Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility. *Journal of Business and Economic Statistics*, 29(3):327–341, 2011.

T. Cogley and T. J. Sargent. Drifts and volatilities: Monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics*, 8(2):262–302, 2005.

O. Coibion and Y. Gorodnichenko. Is the Phillips curve alive and well after all? Inflation expectations and the missing disinflation. *American Economic Journal: Macroeconomics*, 7(1):197–232, 2015.

D. Creal, S. J. Koopman, and E. Zivot. Extracting a robust US business cycle using a time-varying multivariate model-based bandpass filter. *Journal of Applied Econometrics*, 25(4):695–719, 2010.

J. Cross and A. Poon. Forecasting structural change and fat-tailed events in Australian macroeconomic variables. *Economic Modelling*, 58:34–51, 2016.

V. Cúrdia, M. Del Negro, and D. L. Greenwald. Rare shocks, great recessions. *Journal of Applied Econometrics*, 29(7):1031–1052, 2014.

S. Frühwirth-Schnatter and H. Wagner. Stochastic model specification search for Gaussian and partial non-Gaussian state space models. *Journal of Econometrics*, 154:85–100, 2010.

A. Garratt, J. Mitchell, and S. P. Vahey. Measuring output gap nowcast uncertainty. *International Journal of Forecasting*, 30(2):268 – 279, 2014.

S. Gerlach and F. Smets. Output gaps and monetary policy in the EMU area. *European Economic Review*, 43(46):801–812, 1999.

J. Geweke. Bayesian treatment of the independent Student-t linear model. *Journal of Applied Econometrics*, 8:S19–S40, 1993.

R. J. Gordon. Is US economic growth over? Faltering innovation confronts the six headwinds. Technical report, National Bureau of Economic Research, 2012.

A. L. Grant and J. C. C. Chan. Reconciling output gaps: Unobserved components model and Hodrick-Prescott filter. *Journal of Economic Dynamics and Control*, 75:114–121, 2017a.

A. L. Grant and J. C. C. Chan. A Bayesian model comparison for trend-cycle decompositions of output. *Journal of Money, Credit and Banking*, 49(2-3):525–552, 2017b.

A. C. Harvey. Trends and cycles in macroeconomic time series. *Journal of Business and Economic Statistics*, 3(3):216–227, 1985.

A. C. Harvey and A. Jaeger. Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics*, 8(3):231–231, 1993.

A. C. Harvey, T. M. Trimbur, and H. K. Van Dijk. Trends and cycles in economic time series: A Bayesian approach. *Journal of Econometrics*, 140:618–649, 2007.

R. J. Hodrick and E. C. Prescott. Postwar US business cycles: An empirical investigation. *Carnegie Mellon University discussion paper, no. 451.*, 1980.

R. J. Hodrick and E. C. Prescott. Postwar US business cycles: An empirical investigation. *Journal of Money, Credit, and Banking*, pages 1–16, 1997.

S. Kim, N. Shepherd, and S. Chib. Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies*, 65(3):361–393, 1998.

R. Kleijn and H. K. Van Dijk. Bayes model averaging of cyclical decompositions in economic time series. *Journal of Applied Econometrics*, 21(2):191–212, 2006.

G. Koop. Forecasting with medium and large Bayesian VARs. *Journal of Applied Econometrics*, 28(2):177–203, 2013.

G. Koop, D. J. Poirier, and J. L. Tobias. *Bayesian Econometric Methods.* Cambridge University Press, 2007.

D. P. Kroese and J. C. C. Chan. *Statistical Modeling and Computation.* Springer, New York, 2014.

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

S. Luo and R. Startz. Is it one break or ongoing permanent shocks that explains US real GDP? *Journal of Monetary Economics*, 66:155–163, 2014.

W. J. McCausland, S. Miller, and D. Pelletier. Simulation smoothing for state-space models: A computational efficiency analysis. *Computational Statistics and Data Analysis*, 55(1):199–212, 2011.

E. Mertens. On the reliability of output gap estimates in real time. *Unpublished manuscript, Federal Reserve Board*, 2014.

J. Morley, J. Piger, and R. Rasche. Inflation in the G7: Mind the gap(s). *Macroeconomic Dynamics*, 19(4):883–912, 2015.

J. C. Morley and J. Piger. The asymmetric business cycle. *Review of Economics and Statistics*, 94(1):208–221, 2012.

J. C. Morley, C. R. Nelson, and E. Zivot. Why are the Beveridge-Nelson and unobserved-components decompositions of GDP so different? *Review of Economics and Statistics*, 85(2):235–243, 2003.

P. Perron and T. Wada. Let's take a break: Trends and cycles in US real GDP. *Journal of Monetary Economics*, 56(6):749–765, 2009.

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

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

J. H. Stock and M. W. Watson. Why has U.S. inflation become harder to forecast? *Journal of Money Credit and Banking*, 39(s1):3–33, 2007.

R. W. Strachan and H. K. Van Dijk. Evidence on features of a DSGE business cycle model from Bayesian model averaging. *International Economic Review*, 54(1):385–402, 2013.

M. W. Watson. Univariate detrending methods with stochastic trends. *Journal of Monetary Economics*, 18(1):49–75, 1986.