Online Appendix for: A Bounded Model of Time Variation in Trend Inflation, NAIRU and the Phillips Curve

Joshua C.C. Chan
Australian National University

Gary Koop
University of Strathclyde

Simon M. Potter
Federal Reserve Bank of New York

November 20, 2013

1 Summary of Contents of Online Appendix

This appendix contains two sections. The first is a technical appendix which describes the prior and MCMC algorithm used to estimate our bivariate model of inflation and unemployment (labelled **Bi-UC** in the paper). It also provides additional estimation details about the other models used for comparison. The second is section contains a prior predictive analysis. Papers cited in this appendix are listed in the references in the paper itself.

2 Technical Appendix

We remind the reader that our model is defined by:

$$\begin{array}{l}
(\pi_t - \tau_t^\pi) = \rho_t^\pi \left(\pi_{t-1} - \tau_{t-1}^\pi\right) + \lambda_t \left(u_t - \tau_t^u\right) + \varepsilon_t^\pi \\
(u_t - \tau_t^u) = \rho_1^u \left(u_{t-1} - \tau_{t-1}^u\right) + \rho_2^u \left(u_{t-2} - \tau_{t-2}^u\right) + \varepsilon_t^u \\
\tau_t^\pi = \tau_{t-1}^\pi + \varepsilon_t^{\tau\pi} \\
\tau_t^u = \tau_{t-1}^u + \varepsilon_t^{\tau u} \\
\rho_t^\pi = \rho_{t-1}^\pi + \varepsilon_t^{\rho\pi} \\
\lambda_t = \lambda_{t-1} + \varepsilon_t^\lambda
\end{array}, \qquad (1)$$

and

$$
\begin{aligned}
\varepsilon_t^{\pi} &\sim N(0, \mathrm{e}^{h_t}) \\
h_t &= h_{t-1} + \varepsilon_t^h, \\
\varepsilon_t^h &\sim N(0, \sigma_h^2), \\
\varepsilon_t^u &\sim N(0, \sigma_u^2).
\end{aligned} \tag{2}
$$

Trends are bounded through:

$$
\begin{aligned}
\varepsilon_t^{\tau\pi} &\sim TN(a_\pi - \tau_{t-1}^{\pi}, b_\pi - \tau_{t-1}^{\pi}; 0, \sigma_{\tau\pi}^2) \\
\varepsilon_t^{\tau u} &\sim TN(a_u - \tau_{t-1}^{u}, b_u - \tau_{t-1}^{u}; 0, \sigma_{\tau u}^2)
\end{aligned} \tag{3}
$$

and time varying parameters through

$$
\begin{aligned}
\varepsilon_t^{\rho\pi} &\sim TN(-\rho_{t-1}^{\pi}, 1 - \rho_{t-1}^{\pi}; 0, \sigma_{\rho\pi}^2) \\
\varepsilon_t^{\lambda} &\sim TN(-1 - \lambda_{t-1}, 0 - \lambda_{t-1}; 0, \sigma_{\lambda}^2).
\end{aligned} \tag{4}
$$

We also impose the stationary condition on the unemployment equation and assume $\rho_1^u + \rho_2^u < 1$, $\rho_2^u - \rho_1^u < 1$ and $|\rho_2^u| < 1$.

We use notation where $\pi = (\pi_1, \ldots, \pi_T)'$, $u = (u_1, \ldots, u_T)'$, $y = (\pi', u')'$, $\tau^\pi = (\tau_1^\pi, \ldots, \tau_T^\pi)'$, $\tau^u = (\tau_1^u, \ldots, \tau_T^u)'$, $\rho^\pi = (\rho_1^\pi, \ldots, \rho_T^\pi)'$, $\lambda = (\lambda_1, \ldots, \lambda_T)'$, $h = (h_1, \ldots, h_T)'$, and define $\varepsilon^\pi, \varepsilon^u, \varepsilon^{\tau\pi}, \varepsilon^{\tau u}, \varepsilon^h, \varepsilon^{\rho\pi}, \varepsilon^\lambda$ similarly. In addition, let θ denote the model parameters, i.e., $\theta = (\sigma_u^2, \sigma_{\tau\pi}^2, \sigma_{\tau u}^2, \sigma_h^2, \sigma_{\rho\pi}^2, \sigma_\lambda^2, a_\pi, b_\pi, a_u, b_u, \rho_1^u, \rho_2^u)'$.

2.1 The Prior

We require a prior for the initial condition in every state equation and these are:

$$
\begin{aligned}
\tau_1^\pi &\sim TN(a_\pi, b_\pi; \tau_0^\pi, \omega_{\tau\pi}^2), \\
\tau_1^u &\sim TN(a_u, b_u; \tau_0^u, \omega_{\tau u}^2), \\
\rho_1^\pi &\sim TN(0, 1; \rho_0^\pi, \omega_{\rho\pi}^2), \\
\lambda_1 &\sim TN(-1, 0; \lambda_0, \omega_\lambda^2), \\
h_1 &\sim TN(h_0, \omega_h^2),
\end{aligned}
$$

where τ_0^π, $\omega_{\tau\pi}^2$, τ_0^u, τ_{-1}^u, $\omega_{\tau u}^2$, ρ_0^π, $\omega_{\rho\pi}^2$, λ_0, $\omega_{\rho u}^2$, h_0, and ω_h^2 are known constants. In particular we choose the relatively non-informative values of $\tau_0^\pi = 3$, $\tau_0^u = \tau_{-1}^u = 5$, $h_0 = \rho_0^\pi = \lambda_0 = 0$, $\omega_{\tau\pi}^2 = \omega_{\tau u}^2 = \omega_h^2 = 5$ and $\omega_{\rho\pi}^2 = \omega_{\rho u}^2 = \omega_\lambda^2 = 1$.

Note that the need for τ^u_{-1} arises from the use of an AR(2) specification in the unemployment equation.

The prior for the model parameters is specified as

$$p(\theta) = p(\sigma^2_u)p(\sigma^2_h)p(\sigma^2_{\tau\pi})p(\sigma^2_{\tau u})p(\sigma^2_{\rho\pi})p(\sigma^2_\lambda)p\left(a_\pi\right)p\left(b_\pi\right)p\left(a_u\right)p\left(b_u\right)p\left(\rho^u_1\right)p\left(\rho^u_2\right),$$

where $\sigma^2_u \sim IG(\underline{\nu}_u, \underline{S}_u)$, $\sigma^2_h \sim IG(\underline{\nu}_h, \underline{S}_h)$, $\sigma^2_{\tau\pi} \sim IG(\underline{\nu}_{\tau\pi}, \underline{S}_{\tau\pi})$, $\sigma^2_{\tau u} \sim IG(\underline{\nu}_{\tau u}, \underline{S}_{\tau u})$, $\sigma^2_{\rho\pi} \sim IG(\underline{\nu}_{\rho\pi}, \underline{S}_{\rho\pi})$, $\sigma^2_\lambda \sim IG(\underline{\nu}_\lambda, \underline{S}_\lambda)$, and $IG(\cdot,\cdot)$ denotes the inverse-Gamma distribution. We choose relatively small values for the degrees of freedom parameters, which imply large prior variances, i.e., the priors are relatively non-informative. Specifically, we set $\underline{\nu}_u = \underline{\nu}_h = \underline{\nu}_{\tau\pi} = \underline{\nu}_{\tau u} = \underline{\nu}_{\rho\pi} = \underline{\nu}_{\rho u} = \underline{\nu}_\lambda = 10$. We then choose values for the scale parameters so that the parameters have the desired prior means. We set $\underline{S}_u = \underline{S}_h = 0.9$, which imply prior means $E(\sigma^2_u) = E(\sigma^2_h) = 0.1$. Next, we set $\underline{S}_{\tau\pi} = 0.18$ and $\underline{S}_{\tau u} = 0.09$, which imply $E(\sigma^2_{\tau\pi}) = 0.02$ and $E(\sigma^2_{\tau u}) = 0.01$. These values are chosen to reflect the desired smoothness of the corresponding state transition. For example, the prior mean for $\sigma^2_{\tau\pi}$ implies that with high probability the difference between consecutive trend inflation, $\tau^\pi_t - \tau^\pi_{t-1}$, lies within the values -0.3 and 0.3. We set $\underline{S}_{\rho\pi} = \underline{S}_\lambda = 0.018$, which imply prior means $E(\sigma^2_{\rho\pi}) = E(\sigma^2_{\rho u}) = E(\sigma^2_\lambda) = 0.002$. These values imply a relatively smooth transition for the relevant states.

For the bounds we use uniform priors: $a_\pi \sim U(0,2), b_\pi \sim U(3,5), a_u \sim U(3,5), b_u \sim U(6,8)$. The priors for ρ^u_1 and ρ^u_2 are jointly normal with mean $(1.8, -0.8)'$ and covariance matrix $5I_2$.

2.2 MCMC Algorithm

We extend the MCMC sampler developed in Chan, Koop and Potter (2013) which in turn is an adaptation of the algorithm introduced in Chan and Strachan (2012).

Specifically, we sequentially draw from (we suppress the dependence on π_0, u_0 and u_{-1}):

1. $p(\tau^\pi \,|\, y, \tau^u, \rho^\pi, \lambda, h, \theta)$
2. $p(\tau^u \,|\, y, \tau^\pi, \rho^\pi, \lambda, h, \theta)$
3. $p(\rho^\pi \,|\, y, \tau^\pi, \tau^u, \lambda, h, \theta)$
4. $p(\lambda \,|\, y, \tau^\pi, \tau^u, \rho^\pi, h, \theta)$

5. $p(h \,|\, y, \tau^\pi, \tau^u, \rho^\pi, \lambda, \theta)$

6. $p\left(\theta | y, \tau^\pi, \tau^u, \rho^\pi, \lambda, h\right)$

Step 1: To derive the conditional distribution $p(\tau^\pi \,|\, y, \tau^u, \rho^\pi, \lambda, h, \theta)$, we first rewrite the inflation equation as

$$K_\pi \pi = \mu_\pi + K_\pi \tau^\pi + \varepsilon^\pi, \quad \varepsilon^\pi \sim N(0, \Omega_\pi),$$

where $\Omega_\pi = \mathrm{diag}(\mathrm{e}^{h_1}, \ldots, \mathrm{e}^{h_T})$ and

$$\mu_\pi = \begin{pmatrix} \rho_1^\pi(\pi_0 - \tau_0^\pi) + \lambda_1(u_1 - \tau_1^u) \\ \lambda_2(u_2 - \tau_2^u) \\ \lambda_3(u_3 - \tau_3^u) \\ \vdots \\ \lambda_T(u_T - \tau_T^u) \end{pmatrix}, \quad K_\pi = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\rho_2^\pi & 1 & 0 & \cdots & 0 \\ 0 & -\rho_3^\pi & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & -\rho_T^\pi & 1 \end{pmatrix}.$$

Since $|K_\pi| = 1$ for any ρ^π, K_π is invertible. Therefore, we have

$$(\pi \,|\, u, \tau^u, \rho^\pi, \lambda, h, \theta) \sim N(K_\pi^{-1}\mu_\pi + \tau^\pi, (K_\pi' \Omega_\pi^{-1} K_\pi)^{-1}),$$

i.e.,

$$\log p(\pi \,|\, u, \tau^u, \rho^\pi, \lambda, h, \theta) \propto -\frac{1}{2}\iota_T' h - \frac{1}{2}(\pi - K_\pi^{-1}\mu_\pi - \tau^\pi)' K_\pi' \Omega_\pi^{-1} K_\pi (\pi - K_\pi^{-1}\mu_\pi - \tau^\pi), \quad (5)$$

where ι_T is a $T \times 1$ column of ones. Similarly, rewrite the state equation for τ^π as

$$H\tau^\pi = \alpha_\pi + \varepsilon^{\tau\pi},$$

where

$$\alpha_\pi = \begin{pmatrix} \tau_0^\pi \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{pmatrix}.$$

That is, the prior density for τ^π is given by

$$\log p(\tau^\pi \,|\, \sigma^2_{\tau\pi}) \propto \frac{1}{2}(\tau^\pi - H^{-1}\alpha_\pi)' H' \Omega^{-1}_{\tau\pi} H(\tau^\pi - H^{-1}\alpha_\pi) + g_{\tau\pi}(\tau^\pi, \sigma^2_{\tau\pi}), \quad (6)$$

where $a_\pi < \tau^\pi_t < b_\pi$ for $t = 1, \ldots, T$, $\Omega_{\tau\pi} = \mathrm{diag}(\omega^2_{\tau\pi}, \sigma^2_{\tau\pi}, \ldots, \sigma^2_{\tau\pi})$ and

$$\begin{aligned} g_{\tau\pi}(\tau^\pi, \sigma^2_{\tau\pi}) = & -\log\left(\Phi\left(\frac{b_\pi}{\omega_{\tau\pi}}\right) - \Phi\left(\frac{a_\pi}{\omega_{\tau\pi}}\right)\right) \\ & - \sum_{t=2}^{T} \log\left(\Phi\left(\frac{b_\pi - \tau^\pi_{t-1}}{\sigma_{\tau\pi}}\right) - \Phi\left(\frac{a_\pi - \tau^\pi_{t-1}}{\sigma_{\tau\pi}}\right)\right). \end{aligned}$$

Combining (5) and (6), we obtain

$$\begin{aligned} & \log p(\tau^\pi \,|\, y, \tau^u, \rho^\pi, \rho^u, \lambda, h, \theta) \\ & \quad \propto -\frac{1}{2}(\pi - K^{-1}_\pi \mu_\pi - \tau^\pi)' K'_\pi \Omega^{-1}_\pi K_\pi (\pi - K^{-1}_\pi \mu_\pi - \tau^\pi) \\ & \quad\quad - \frac{1}{2}(\tau^\pi - H^{-1}\alpha_\pi)' H' \Omega^{-1}_{\tau\pi} H(\tau^\pi - H^{-1}\alpha_\pi) + g_{\tau\pi}(\tau^\pi, \sigma^2_{\tau\pi}), \\ & \quad \propto -\frac{1}{2}(\tau^\pi - \hat{\tau}^\pi)' D^{-1}_{\tau\pi}(\tau^\pi - \hat{\tau}^\pi) + g_{\tau\pi}(\tau^\pi, \sigma^2_{\tau\pi}), \end{aligned}$$

where $a_\pi < \tau^\pi_t < b_\pi$ for $t = 1, \ldots, T$, and

$$\begin{aligned} D_{\tau\pi} &= \left(H' \Omega^{-1}_{\tau\pi} H + K'_\pi \Omega^{-1}_\pi K_\pi\right)^{-1}, \\ \hat{\tau}^\pi &= D_{\tau\pi}(H' \Omega^{-1}_{\tau\pi} \alpha_\pi + K'_\pi \Omega^{-1}_\pi K_\pi (\pi - K^{-1}_\pi \mu_\pi)). \end{aligned}$$

Since this conditional density is non-standard, we sample τ^π via an independence-chain Metropolis-Hastings (MH) step. Specifically, candidate draws are first obtained from the $N(\hat{\tau}^\pi, D_{\tau\pi})$ distribution with the precision-based algorithm discussed in Chan and Jeliazkov (2009), and they are accepted or rejected via an acceptance-rejection Metropolis-Hastings (ARMH) step.

Step 2: To implement Step 2, note that information about τ^u comes from three sources: the two measurement equations for inflation and unemployment and the state equation for τ^u. We derive an expression for each component in turn. First, write the inflation equation as:

$$z = \Lambda \tau^u + \varepsilon^\pi, \quad \varepsilon^\pi \sim N(0, \Omega_\pi),$$

where $z_t = (\pi_t - \tau^\pi_t) - \rho^\pi_t(\pi_{t-1} - \tau^\pi_{t-1}) - \lambda_t u_t$, $z = (z_1, \dots, z_T)'$, and $\Lambda = \mathrm{diag}(-\lambda_1, \dots, -\lambda_T)$. Hence, ignoring any terms not involving τ^u, we have

$$\log p(\pi \mid u, \tau^u, \rho^\pi, \lambda, h, \theta) \propto -\frac{1}{2}(z - \Lambda\tau^u)'\Omega_\pi^{-1}(z - \Lambda\tau^u). \tag{7}$$

The second component comes from the unemployment equation, which can be written as:

$$K_u u = \mu_u + K_u\tau^u + \varepsilon^u, \quad \varepsilon^u \sim N(0, \Omega_u),$$

where $\Omega_u = I_T \otimes \sigma^2_u$ and

$$\mu_u = \begin{pmatrix} \rho^u_1(u_0 - \tau^u_0) + \rho^u_2(u_{-1} - \tau^u_{-1}) \\ \rho^u_2(u_0 - \tau^u_0) \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad K_u = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -\rho^u_1 & 1 & 0 & 0 & \cdots & 0 \\ -\rho^u_2 & -\rho^u_1 & 1 & 0 & \cdots & 0 \\ 0 & -\rho^u_2 & -\rho^u_1 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -\rho^u_2 & -\rho^u_1 & 1 \end{pmatrix}.$$

Thus, ignoring any terms not involving τ^u, we have

$$\log p(u \mid \tau^u, \theta) \propto -\frac{1}{2}(u - K_u^{-1}\mu_u - \tau^u)'K_u'\Omega_u^{-1}K_u(u - K_u^{-1}\mu_u - \tau^u). \tag{8}$$

The third component is contributed by the state equation for τ^u :

$$\log p(\tau^u \mid \sigma^2_{\tau u}) \propto -\frac{1}{2}(\tau^u - H^{-1}\alpha_u)'H'\Omega_{\tau u}^{-1}H(\tau^u - H^{-1}\alpha_u) + g_{\tau u}(\tau^u, \sigma^2_{\tau u}), \tag{9}$$

where $\alpha_u = (\tau^u_0, 0, \dots, 0)'$, $a_u < \tau^u_t < b_u$ for $t = 1, \dots, T$, $\Omega_{\tau u} = \mathrm{diag}(\omega^2_{\tau u}, \sigma^2_{\tau u}, \dots, \sigma^2_{\tau u})$ and

$$\begin{aligned} g_{\tau u}(\tau^u, \sigma^2_{\tau u}) = &- \log\left(\Phi\left(\frac{b_u}{\omega_{\tau u}}\right) - \Phi\left(\frac{a_u}{\omega_{\tau u}}\right)\right) \\ &- \sum_{t=2}^{T} \log\left(\Phi\left(\frac{b_u - \tau^u_{t-1}}{\sigma_{\tau u}}\right) - \Phi\left(\frac{a_u - \tau^u_{t-1}}{\sigma_{\tau u}}\right)\right). \end{aligned}$$

Combining (7), (8) and (9), we obtain

$$
\begin{aligned}
&\log p(\tau^u \,|\, y, \tau^\pi, \rho_1^u, \rho_2^u, \lambda, h, \theta) \\
&\qquad \propto -\frac{1}{2}(z - \Lambda\tau^u)'\Omega_\pi^{-1}(z - \Lambda\tau^u) \\
&\qquad\quad - \frac{1}{2}(u - K_u^{-1}\mu_u - \tau^u)'K_u'\Omega_u^{-1}K_u(u - K_u^{-1}\mu_u - \tau^u) \\
&\qquad\quad - \frac{1}{2}(\tau^u - H^{-1}\alpha_u)'H'\Omega_{\tau u}^{-1}H(\tau^u - H^{-1}\alpha_u) + g_{\tau u}(\tau^u, \sigma_{\tau u}^2), \\
&\qquad \propto -\frac{1}{2}(\tau^u - \hat{\tau}^u)'D_{\tau u}^{-1}(\tau^u - \hat{\tau}^u) + g_{\tau u}(\tau^u, \sigma_{\tau u}^2),
\end{aligned}
$$

where $a_u < \tau_t^u < b_u$ for $t = 1, \ldots, T$, and

$$
\begin{aligned}
D_{\tau u} &= \left(H'\Omega_{\tau u}^{-1}H + K_u'\Omega_u^{-1}K_u + \Lambda'\Omega_\pi^{-1}\Lambda\right)^{-1}, \\
\hat{\tau}^u &= D_{\tau u}(H'\Omega_{\tau u}^{-1}\alpha_u + K_u'\Omega_u^{-1}K_u(u - K_u^{-1}\mu_u) + \Lambda'\Omega_\pi^{-1}z).
\end{aligned}
$$

Again, we sample τ^u via an ARMH step with candidate draws obtained from $N(\hat{\tau}^u, D_{\tau u})$.

Step 3: Next, we derive an expression for $p(\rho^\pi \,|\, y, \tau^\pi, \tau^u, \lambda, h, \theta)$. First, let $\pi_t^* = \pi_t - \tau_t^\pi$, $u_t^* = u_t - \tau_t^u$, $\pi^* = (\pi_1^*, \ldots, \pi_T^*)'$, and $u^* = (u_1^*, \ldots, u_T^*)'$. Then the measurement equation for inflation can be rewritten as

$$
\pi^* + \Lambda u^* = X_\pi \rho^\pi + \varepsilon^\pi, \quad \varepsilon^\pi \sim N(0, \Omega_\pi),
$$

where $X_\pi = \text{diag}(\pi_0^*, \ldots, \pi_{T-1}^*)$ and $\Lambda = \text{diag}(-\lambda_1, \ldots, -\lambda_T)$. From the state equation for ρ^π we also have

$$
H\rho^\pi = \varepsilon^{\rho\pi}.
$$

Therefore, using a similar argument as before, we have

$$
p(\rho^\pi \,|\, y, \tau^\pi, \tau^u, \rho^u, \lambda, h, \theta) \propto -\frac{1}{2}(\rho^\pi - \hat{\rho}^\pi)'D_{\rho\pi}^{-1}(\rho^\pi - \hat{\rho}^\pi) + g_{\rho\pi}(\rho^\pi, \sigma_{\rho\pi}^2),
$$

where $0 < \rho_t^\pi < 1$ for $t = 1, \ldots, T$,

$$
g_{\rho\pi}(\rho^\pi, \sigma_{\rho\pi}^2) = -\sum_{t=2}^{T} \log\left(\Phi\left(\frac{1 - \rho_{t-1}^\pi}{\sigma_{\rho\pi}}\right) - \Phi\left(\frac{-\rho_{t-1}^\pi}{\sigma_{\rho\pi}}\right)\right),
$$

$$
D_{\rho\pi} = \left(H'\Omega_{\rho\pi}^{-1}H + X_\pi'\Omega_\pi^{-1}X_\pi\right)^{-1}, \quad \hat{\rho}^\pi = D_{\rho\pi}X_\pi'\Omega_\pi^{-1}(\pi^* + \Lambda u^*),
$$

and $\Omega_{\rho\pi} = \text{diag}(\omega^2_{\rho\pi}, \sigma^2_{\rho\pi}, \ldots, \sigma^2_{\rho\pi})$. As before, we implement an ARMH step with approximating density $N(\hat{\rho}^\pi, D_{\rho\pi})$.

Step 4: Using the same argument as before, we have

$$p(\lambda \,|\, y, \tau^\pi, \tau^u, \rho^\pi, h, \theta) \propto -\frac{1}{2}(\lambda - \hat{\lambda})' D_\lambda^{-1}(\lambda - \hat{\lambda}) + g_\lambda(\lambda, \sigma^2_\lambda),$$

where $-1 < \lambda_t < 0$ for $t = 1, \ldots, T$,

$$g_\lambda(\lambda, \sigma^2_\lambda) = -\sum_{t=2}^{T} \log\left(\Phi\left(\frac{-\lambda_{t-1}}{\sigma_\lambda}\right) - \Phi\left(\frac{-1-\lambda_{t-1}}{\sigma_\lambda}\right)\right),$$

$$D_\lambda = \left(H'\Omega_\lambda^{-1}H + X_u'\Omega_\pi^{-1}X_u\right)^{-1}, \quad \hat{\lambda} = D_\lambda X_u'\Omega_\pi^{-1}w,$$

$X_u = \text{diag}(u^*_0, \ldots, u^*_{T-1})$, $w = (\pi^*_1 - \rho^\pi_1\pi^*_0, \ldots, \pi^*_T - \rho^\pi_T\pi^*_{T-1})'$, and $\Omega_\lambda = \text{diag}(\omega^2_\lambda, \sigma^2_\lambda, \ldots, \sigma^2_\lambda)$. As before, we implement an ARMH step with approximating density $N(\hat{\rho}^u, D_{\rho u})$.

Step 5: For Step 5, we use the algorithm in Chan and Strachan (2012) to sample from $p(h \,|\, y, \tau^\pi, \tau^u, \rho^\pi, \lambda, \theta)$.

Step 6: We draw from θ in separate blocks, mainly using standard results for the regression model. We use notation where θ_{-x} for all parameters in θ except for x.

Using standard linear regression results, it can be shown that $\rho^u = (\rho^u_1, \rho^u_2)'$ is a bivariate truncated normal:

$$p(\rho^u \,|\, y, \tau^\pi, \tau^u, \rho^\pi, \lambda, h, \theta_{-\rho^u}) \propto -\frac{1}{2}(\rho^u - \hat{\rho}^u)' D_{\rho u}^{-1}(\rho^u - \hat{\rho}^u) + g_{\rho u}(\rho^u),$$

with the stationarity constraints $\rho^u_1 + \rho^u_2 < 1$, $\rho^u_2 - \rho^u_1 < 1$ and $|\rho^u_2| < 1$, where

$$D_{\rho u} = \left(V_{\rho u}^{-1} + X_u'X_u/\sigma^2_u\right)^{-1}, \quad \hat{\rho}^u = D_{\rho u}X_u'^{*}/\sigma^2_u, \quad X_u = \begin{pmatrix} u^*_0 & u^*_{-1} \\ u^*_1 & u^*_0 \\ \vdots & \vdots \\ u^*_{T-1} & u^*_{T-2} \end{pmatrix}.$$

A draw from this truncated normal distribution can be obtained via acceptance-rejection sampling with proposal from $N(\hat{\rho}^u, D_{\rho u})$.

To sample from the error variances, first note that they are conditionally independent given the data and the states. Hence, we can sample each element one by one.

Now, both $p(\sigma_u^2 \,|\, y, \tau^\pi, \tau^u, \rho^\pi, \lambda, h, \theta_{-\sigma_u^2})$ and $p(\sigma_h^2 \,|\, y, \tau^\pi, \tau^u, \rho^\pi, \lambda, h, \theta_{-\sigma_h^2})$ are standard inverse-Gamma densities, respectively:

$$(\sigma_u^2 \,|\, y, \tau^\pi, \tau^u, \rho^\pi, \lambda, h, \theta_{-\sigma_u^2}) \sim IG\left(\underline{\nu}_u + \frac{T}{2}, \underline{S}_u + \frac{1}{2}\sum_{t=1}^T (u_t^* - \rho_1^u u_{t-1}^* - \rho_2^u u_{t-2}^*)^2\right)$$

$$(\sigma_h^2 \,|\, y, \tau^\pi, \tau^u, \rho^\pi, \rho^u, \lambda, h, \theta_{-\sigma_h^2}) \sim IG\left(\underline{\nu}_h + \frac{T-1}{2}, \underline{S}_h + \frac{1}{2}\sum_{t=2}^T (h_t - h_{t-1})^2\right).$$

Next, the log conditional density for $\sigma_{\tau\pi}^2$ is given by

$$\log p(\sigma_{\tau\pi}^2 \,|\, y, \tau^\pi, \tau^u, \rho^\pi, \lambda, h, \theta_{-\sigma_{\tau\pi}^2}) \propto$$
$$-(\underline{\nu}_{\tau\pi} + 1)\log\sigma_{\tau\pi}^2 - \frac{\underline{S}_{\tau\pi}}{\sigma_{\tau\pi}^2} - \frac{T-1}{2}\log\sigma_{\tau\pi}^2 - \frac{1}{2\sigma_{\tau\pi}^2}\sum_{t=2}^T (\tau_t^\pi - \tau_{t-1}^\pi)^2 + g_{\tau\pi}(\tau^\pi, \sigma_{\tau\pi}^2),$$

which is a non-standard density. To proceed, we implement an MH step with the proposal density

$$IG\left(\underline{\nu}_{\tau\pi} + \frac{T-1}{2}, \underline{S}_{\tau\pi} + \frac{1}{2}\sum_{t=2}^T (\tau_t^\pi - \tau_{t-1}^\pi)^2\right).$$

Similarly, the log conditional density for $\sigma_{\tau u}^2$ is given by

$$\log p(\sigma_{\tau u}^2 \,|\, y, \tau^\pi, \tau^u, \rho^\pi, \lambda, h, \theta_{-\sigma_{\tau u}^2}) \propto$$
$$-(\underline{\nu}_{\tau u} + 1)\log\sigma_{\tau u}^2 - \frac{\underline{S}_{\tau u}}{\sigma_{\tau u}^2} - \frac{T-1}{2}\log\sigma_{\tau u}^2 - \frac{1}{2\sigma_{\tau u}^2}\sum_{t=2}^T (\tau_t^u - \tau_{t-1}^u)^2 + g_{\tau u}(\tau^u, \sigma_{\tau u}^2).$$

Again, a draw from $p(\sigma_{\tau u}^2 \,|\, y, \tau^\pi, \tau^u, \rho^\pi, \lambda, h, \theta_{-\sigma_{\tau u}^2})$ is obtained via an MH step with the proposal density

$$IG\left(\underline{\nu}_{\tau u} + \frac{T-1}{2}, \underline{S}_{\tau u} + \frac{1}{2}\sum_{t=2}^T (\tau_t^u - \tau_{t-1}^u)^2\right).$$

The remaining error variances, $\sigma_{\rho\pi}^2$ and σ_λ^2, are sampled analogously.

To draw from the bounds a_π, b_π, a_u and b_u, we use a Griddy-Gibbs sampler which is the same as the one used in Chan, Koop and Potter (2013). Specifically, since each of the bounds has finite support, we can evaluate its conditional density on a fine grid, which can then be used to construct the associated cumulative distribution function. Finally, a draw from the conditional density can be obtained via the inverse-transform method. The reader is referred to our earlier paper for details.

2.3 Specification and Estimation Details for Other Models

The other models used in our forecast comparisons are mostly restricted special cases of **Bi-UC** and all specification and prior details are identical to **Bi-UC** except that the relevant restriction is imposed. The exceptions to this are discussed in this sub-section.

The VAR(2) is specified as:

$$y_t = \mu + B_1 y_{t-1} + B_2 y_{t-2} + \varepsilon_t,$$

where $y_t = (\pi_t, u_t)'$ and $\varepsilon_t = (\varepsilon_t^\pi, \varepsilon_t^u)' \sim N(0, \Sigma)$. We use a relatively noninformative prior. For μ the prior is $N(0, 100)$. For the VAR coefficients, we assume each is N(0,1) and all are, a priori, uncorrelated with one another. The prior for Σ is $IW\left(10, \begin{bmatrix} 1.4 & 0 \\ 0 & 0.7 \end{bmatrix}\right)$ so that the prior mean of the error variances in the two equations are 0.2 and 0.1, respectively.

We also use a VAR(2) with stochastic volatility:

$$y_t = \mu + B_1 y_{t-1} + B_2 y_{t-2} + A^{-1}\varepsilon_t,$$

where $\varepsilon_t^\pi \sim N(0, \mathrm{e}^{h_t})$, $\varepsilon_t^u \sim N(0, \sigma_u^2)$ and

$$A = \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}.$$

The VAR coefficients have the same prior as the VAR(2) without stochastic volatility. All details (including prior) relating to h_t and σ_u^2 are exactly as in **Bi-UC** and a has a $N(0, 10)$ prior.

The VARs are estimated using MCMC methods as outlined, e.g., in Koop and Korobilis (2009).

For **Bi-UC-TVP-**ρ^u, all details are identical to **Bi-UC** except that we need to specify the initial conditions for ρ_{1t}^u and ρ_{2t}^u which are now time-varying. These are both assumed to be $N(0, 5)$. The priors for the two error variances in the two state equations are both $IG(10, 0.009)$, a relatively noninformative choice implying prior means of 0.001.

The VAR(2) with Minnesota Prior model is specified as follows. Let $\beta = \mathrm{vec}((\mu, B_1, B_2)')$ and $X_t = I_2 \otimes (1, y_{t-1}', y_{t-2}')$, and let $\widehat{b}$ denote the least squares estimator of β. We fix $\Sigma = \widehat{\Sigma}$, where $\widehat{\Sigma} = T^{-1}\sum_{t=1}^{T}(y_t - X_t\widehat{b})'(y_t - X_t\widehat{b})$. For β, we consider the prior $\beta \sim N(0, \underline{V}_{\text{mn}})$, where

$$\underline{V}_{\text{mn}} = \text{diag}\left(\underline{a}_3 s_1^2, \underline{a}_1, \frac{\underline{a}_2 s_1^2}{s_2^2}, \frac{\underline{a}_1}{4}, \frac{\underline{a}_2 s_1^2}{4s_2^2}, \underline{a}_3 s_2^2, \frac{\underline{a}_2 s_2^2}{s_1^2}, \underline{a}_1, \frac{\underline{a}_2 s_2^2}{4s_1^2}, \frac{\underline{a}_1}{4}\right)$$

and s_i^2 is the i-th diagonal element of $\widehat{\Sigma}$, $i = 1, 2$. We set $\underline{a}_1 = 0.1$, $\underline{a}_2 = 0.05$, and $\underline{a}_3 = 0.1$.

The bivariate random walk model is:

$$\pi_t = \pi_{t-1} + \epsilon_t^\pi,$$
$$u_t = u_{t-1} + \epsilon_t^u,$$

where $\epsilon_t^\pi \sim N(0, \sigma_\pi^2)$ and $\epsilon_t^u \sim N(0, \sigma_u^2)$ are independent.

This model is a special case of **Bi-UC**, where $\rho_1^u = 1$, $\rho_2^u = 0$, $\tau_t^\pi = \tau_t^u = \lambda_t = 0$, $\rho_t^\pi = 1$ for $t = 1, \ldots, T$, and the errors are independent and homoskedastic.

The UCSV-AR(2) model is:

$$\pi_t = \tau_t^\pi + \epsilon_t^\pi,$$
$$u_t = \rho_1^u u_{t-1} + \rho_2^u u_{t-2} + \epsilon_t^u,$$

where $\epsilon_t^\pi \sim N(0, e^{h_t})$ and $\epsilon_t^u \sim N(0, \sigma_u^2)$ are independent. The inflation trend τ_t^π and log-volatility h_t follow independent random walks. This model is a special case of **Bi-UC**, where $\tau_t^u = \lambda_t = \rho_t^\pi = 0, = 1$ for $t = 1, \ldots, T$.

3 Prior Predictive Analysis

To convince the reader of the sensibility of our model and prior, this subsection presents results from a prior predictive analysis.

We begin by computing the predictive densities for future trend inflation and future NAIRU, τ_{T+k}^π and τ_{T+k}^u respectively, with $k = 20$ under the **Bi-UC** model. The results are reported in Figure 3 and 3. These figures should that the predictive densities produced by our model are sensible and show the role of the bounds tightening these densities in a sensible manner.

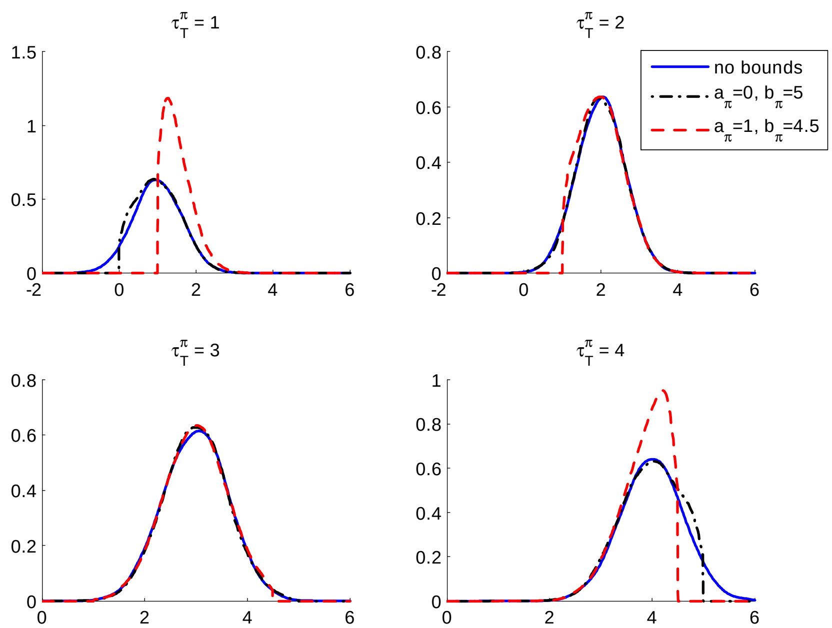

Figure 1: Predictive densities for τ^{π}_{T+k} under the **Bi-UC** model where $k = 20$.

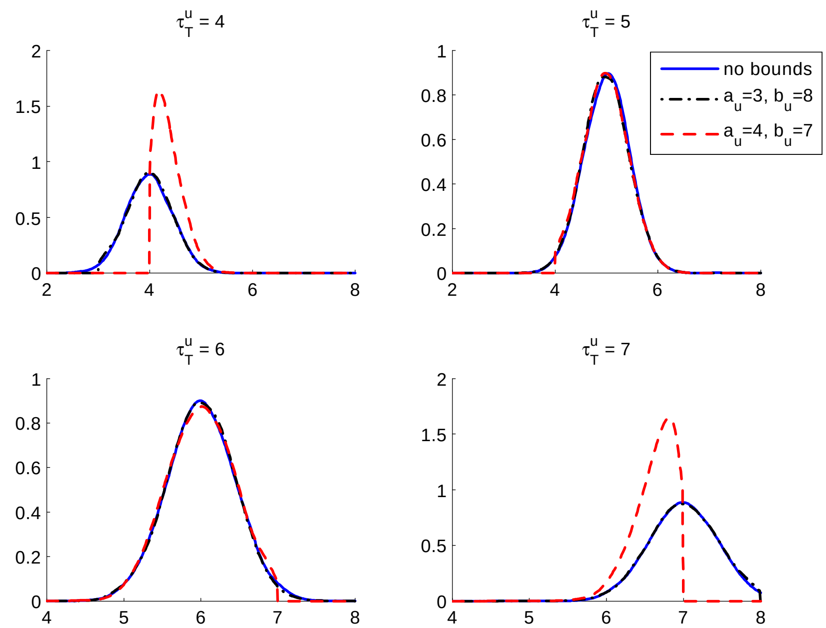

Figure 2: Predictive densities for τ^u_{T+k} under the **Bi-UC** model where $k = 20$.

Next, we preform a prior predictive analysis. This involves taking a draw from the prior (using the prior described in the Technical Appendix) and then simulating from the state equations. Given the drawn parameters, states and an initial value for inflation and the unemployment rate (we set $\pi_0 = 3$ and $u_0 = u_{-1} = 5.5$), an artificial dataset of inflation and unemployment can be generated. This is repeated 10^4 times and, for each generated data set, we compute various features of interest such as quantiles, variance, autocorrelations, etc. in order to build up the prior cumulative distribution function (cdf) for each of these features. Tables 1 and 2 present these cdf's evaluated at the feature of interest calculated for the observed data set. It can be seen that all of the features of the data can be well explained by our model. Our model does worst at explaining the autocorrelation patterns in the unemployment rate series. However, even for this case, it does as well as an unbounded version of our model.

To formally compare **Bi-UC** and **Bi-UC-NoBound**, we compute the log

Table 1: Prior cdfs of features for π_t.

Feature	**Bi-UC**	**Bi-UC-NoBound**
16%-tilde	0.576	0.574
median	0.557	0.547
84%-tilde	0.663	0.585
variance	0.598	0.520
fraction of $\pi_t < 0$	0.431	0.417
fraction of $\pi_t > 10$	0.687	0.620
lag 1 autocorrelation	0.634	0.706
lag 4 autocorrelation	0.726	0.581
MA coefficient	0.539	0.713

Table 2: Prior cdfs of features for u_t.

Feature	**Bi-UC**	**Bi-UC-NoBound**
16%-tilde	0.694	0.708
median	0.715	0.721
84%-tilde	0.725	0.703
variance	0.626	0.587
fraction of $u_t < 4$	0.273	0.261
fraction of $u_t > 8$	0.746	0.719
lag 1 autocorrelation	0.965	0.965
lag 4 autocorrelation	0.922	0.903
MA coefficient	0.765	0.764

Bayes factors using the prior predictive densities for various combinations of the features of interest considered in Tables 1 and 2. In particular, we divide the features into three groups: “Quantile" includes the first three features of interest (16%-tilde, median and 84%-tilde), “Spread and Drift" includes the next three (variance, fraction of $y_t < 0$, and fraction of $y_t > 10$ for $y_t = \pi_t$ or u_t), and “Dynamics" include the last three features of interest (lag 1 autocorrelation, lag 4 autocorrelation and MA coefficient). The results are reported in Table 3. The fact that the inclusion of bounds leads to a more parsimonious model (without causing the fit of the model to deteriorate), leads to log Bayes factors strongly in favor of our bounded model for inflation. For the unemployment rate, the bounds play less of a role and our model is performing roughly as well as its unbounded version.

Table 3: Log Bayes factors in favor of **Bi-UC** against **Bi-UC-NoBound**.

	Quantile	Spread and drift	Dynamics	All
π_t	286.67	328.76	1.45	613.85
u_t	0.47	0.34	-0.53	0.15