Federal Reserve Bank of Cleveland Working Paper Series

Online Appendix

A Unified Framework to Estimate Macroeconomic Stars

Saeed Zaman

Working Paper No. 21-23R

August 2022

Suggested citation: Zaman, Saeed. 2022. "Online Appendix to A Unified Framework to Estimate Macroeconomic Stars." Working Paper No. 21-23R. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202123r.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

Online Appendix

A Unified Framework to Estimate Macroeconomic Stars

Saeed Zaman*
Federal Reserve Bank of Cleveland, USA

First version: June 30, 2021
This version: July 31, 2022

I am extremely grateful to Joshua Chan, Todd Clark, Julia Darby, Ana Galvão, Gary Koop (my primary PhD advisor), Edward Knotek, Kurt Lunsford, Elmar Mertens, James Mitchell, Aubrey Poon, Ellis Tallman, Willem Van Zandweghe, Randal Verbrugge, and Ben Wong for valuable comments on the first draft. This research has benefited from discussions with Fabio Canova, Siddhartha Chib, Olivier Coibion, Sharada Davidson, Jesus Fernandez-Villaverde, Garo Garabedian, Paolo Gelain, Yuriy Gorodnichenko, Stuart McIntyre, Eric Sims, and Ping Wu. I also thank participants at the Society for Computational Economics 27th Conference 2021, NBER-NSF SBIES 2021, European Seminar on Bayesian Econometrics 2021, Computational and Financial Econometrics Conference 2021, 2022 North American Summer Meeting of the Econometric Society, 5th Vienna Workshop on High-Dimensional Time Series in Macroeconomics and Finance, Midwest Macroeconomics Spring 2022, Society for Nonlinear Dynamics and Econometrics 2022, ESCoE Conference on Economic Measurement 2022, and the Cleveland Fed Brown Bag. The views expressed herein are those of the author and do not necessarily represent the views of the Federal Reserve Bank of Cleveland or the Federal Reserve System.

*Research department, Federal Reserve Bank of Cleveland, Ohio, USA; email: saeed.zaman@clev.frb.org

Contents

A1. Bayesian Estimation Details

A1.a. Base model equations

For convenience, we list all model equations keeping the numbering as in the main text.

$$U_t = U_t^* + U_t^c \tag{6}$$

$$U_t - U_t^* = \rho_1^u(U_{t-1} - U_{t-1}^*) + \rho_2^u(U_{t-2} - U_{t-2}^*) + \phi^u ogap_t + \varepsilon_t^u,\ \ \varepsilon_t^u \sim N(0, e^{h_t^u}) \tag{7}$$

where, $\rho_1^u + \rho_2^u < 1$, $\rho_2^u - \rho_1^u < 1$, and $|\rho_2^u| < 1$; $\phi^u < 0$

$$U_t^* = U_{t-1}^* + \varepsilon_t^{u*},\ \ \varepsilon_t^{u*} \sim TN(a_u - U_{t-1}^*, b_u - U_{t-1}^*; 0, \sigma_{u*}^2) \tag{8}$$

$$Z_t^u = C_t^u + \beta^u U_t^* + \varepsilon_t^{zu},\ \ \varepsilon_t^{zu} \sim N(0, \sigma_{zu}^2) \tag{9}$$

$$C_t^u = C_{t-1}^u + \varepsilon_t^{cu},\ \ \varepsilon_t^{cu} \sim N(0, \sigma_{cu}^2) \tag{10}$$

$$gdp_t = gdp_t^* + ogap_t \tag{11}$$

$$gdp_t^* = 2gdp_{t-1}^* - gdp_{t-2}^* + \varepsilon_t^{gdp*},\ \ \varepsilon_t^{gdp*} \sim N(0, \sigma_{gdp*}^2) \tag{12}$$

$$g_t^* \equiv \triangle gdp_t^*$$

$$g_t^* = g_{t-1}^* + \varepsilon_t^{gdp*} \tag{13}$$

$$ogap_t = \rho_1^g(ogap_{t-1}) + \rho_2^g(ogap_{t-2}) + a^r(r_t^L - r_t^* - tp_t^*) + \lambda^g(U_t - U_t^*) + \varepsilon_t^{ogap} \tag{14}$$

where, $\varepsilon_t^{ogap} \sim N(0, e^{h_t^o})$, $\rho_1^g + \rho_2^g < 1$, $\rho_2^g - \rho_1^g < 1$, and $|\rho_2^g| < 1$; $\lambda^g < 0$

$$Z_t^g = C_t^g + \beta^g * 4 * g_t^* + \varepsilon_t^{zg},\ \ \varepsilon_t^{zg} \sim N(0, \sigma_{zg}^2) \tag{15}$$

$$C_t^g = C_{t-1}^g + \varepsilon_t^{cg},\ \ \varepsilon_t^{cg} \sim N(0, \sigma_{cg}^2) \tag{16}$$

$$P_t - P_t^* = \rho^p(P_{t-1} - P_{t-1}^*) + \lambda_t^p(U_t - U_t^*) + \varepsilon_t^p,\ \ \varepsilon_t^p \sim N(0, e^{h_t^p}) \tag{17}$$

where, $|\rho^p| < 1$

$$\lambda_t^p = \lambda_{t-1}^p + \varepsilon_t^{\lambda p},\ \ \varepsilon_t^{\lambda p} \sim N(0, \sigma_{\lambda p}^2) \tag{18}$$

$$P_t^* = P_{t-1}^* + \varepsilon_t^{p*},\ \ \varepsilon_t^{p*} \sim N(0, \sigma_{p*}^2) \tag{19}$$

$$\pi_t - \pi_t^* = \rho_t^\pi(\pi_{t-1} - \pi_{t-1}^*) + \lambda_t^\pi(U_t - U_t^*) + \varepsilon_t^\pi,\ \ \varepsilon_t^\pi \sim N(0, e^{h_t^\pi}) \tag{20}$$

$$\rho_t^\pi = \rho_{t-1}^\pi + \varepsilon_t^{\rho\pi},\ \ \varepsilon_t^{\rho\pi} \sim TN(0 - \rho_{t-1}^\pi, 1 - \rho_{t-1}^\pi; 0, \sigma_{\rho\pi}^2) \tag{21}$$

where, ρ^π is truncated so that $0 < \rho_t^\pi < 1$.

$$\lambda_t^\pi = \lambda_{t-1}^\pi + \varepsilon_t^{\lambda\pi},\ \ \varepsilon_t^{\lambda\pi} \sim TN(-1 - \lambda_{t-1}^\pi, 0 - \lambda_{t-1}^\pi; 0, \sigma_{\lambda\pi}^2) \tag{22}$$

λ^π is the slope of the price Phillips curve and is constrained in the interval (-1,0).

$$\pi_t^* = \pi_{t-1}^* + \varepsilon_t^{\pi*},\ \ \varepsilon_t^{\pi*} \sim N(0, \sigma_{\pi*}^2) \tag{23}$$

$$Z_t^\pi = C_t^\pi + \beta^\pi \pi_t^* + \varepsilon_t^{z\pi},\ \ \varepsilon_t^{z\pi} \sim N(0, \sigma_{z\pi}^2) \tag{24}$$

$$C_t^\pi = C_{t-1}^\pi + \varepsilon_t^{c\pi},\ \ \varepsilon_t^{c\pi} \sim N(0, \sigma_{c\pi}^2) \tag{25}$$

$$W_t^* = \pi_t^* + P_t^* + Wedge_t + \varepsilon_t^{w*},\ \ \varepsilon_t^{w*} \sim N(0, \sigma_{w*}^2) \tag{26}$$

$$Wedge_t = Wedge_{t-1} + \varepsilon_t^{wlr},\ \ \varepsilon_t^{wlr} \sim N(0, \sigma_{wlr}^2) \tag{27}$$

$$W_t - W_t^* = \rho_t^w(W_{t-1} - W_{t-1}^*) + \lambda_t^w(U_t - U_t^*) + \kappa_t^w(\pi_t - \pi_t^*) + \varepsilon_t^w,\ \ \varepsilon_t^w \sim N(0, e^{h_t^w}) \tag{28}$$

$$\rho_t^w = \rho_{t-1}^w + \varepsilon_t^{\rho w},\ \ \varepsilon_t^{\rho w} \sim TN(0 - \rho_{t-1}^w, 1 - \rho_{t-1}^w; 0, \sigma_{\rho w}^2) \tag{29}$$

$$\lambda_t^w = \lambda_{t-1}^w + \varepsilon_t^{\lambda w},\ \ \varepsilon_t^{\lambda w} \sim TN(-1 - \lambda_{t-1}^w, 0 - \lambda_{t-1}^w; 0, \sigma_{\lambda w}^2) \tag{30}$$

λ^w is the slope of the wage Phillips curve and is constrained in the interval (-1,0).

$$\kappa_t^w = \kappa_{t-1}^w + \varepsilon_t^{\kappa w},\ \ \varepsilon_t^{\kappa w} \sim N(0, \sigma_{\kappa w}^2) \tag{31}$$

$$i_t - \pi_t^* - r_t^* = \rho^i(i_{t-1} - \pi_{t-1}^* - r_{t-1}^*) + \lambda^i(U_t - U_t^*) + \kappa^i(\pi_t - \pi_t^*) + \varepsilon_t^i,\ \ \varepsilon_t^i \sim N(0, e^{h_t^i}) \tag{32}$$

where, ρ^i is truncated so that $0 < \rho^i < 1$.

$$r_t^* = \zeta g_t^* + D_t. \tag{33}$$

$$D_t = D_{t-1} + \varepsilon_t^d,\ \ \varepsilon_t^d \sim N(0, \sigma_d^2) \tag{34}$$

$$Z_t^r = C_t^r + \beta^r r_t^* + \varepsilon_t^{zr},\ \ \varepsilon_t^{zr} \sim N(0, \sigma_{zr}^2) \tag{35}$$

$$C_t^r = C_{t-1}^r + \varepsilon_t^{cr},\ \ \varepsilon_t^{cr} \sim N(0, \sigma_{cr}^2) \tag{36}$$

$$h_t^{id} = h_{t-1}^{id} + \varepsilon_t^j,\ \ \varepsilon_t^j \sim N(0, \sigma_j^2) \tag{37}$$

where $id = \{u, ogap, p, \pi, w, i\}$, and $j = \{hu, ho, hp, h\pi, hw, hi\}$

A1.b. Prior elicitation

Our prior settings are similar to those used in Chan, Koop, and Potter (2016) [CKP], Chan, Clark, and Koop (2018) [CCK], and Gonzalez-Astudillo and Laforte (2020). As discussed in CCK, UC models with several unobserved variables, such as the one developed in this paper, require informative priors. That said, our priors settings for most variables are only slightly informative. The use of inequality restrictions on some parameters such as the Phillips curve, persistence, and bounds on u-star could be viewed as additional sources of information that eliminates the need for tight priors, something also noted by CKP. For the parameters for which there is strong agreement in the empirical literature on their values, such as the Taylor-rule equation parameters, we use relatively tight priors, such that prior distributions are centered on prior means with small variance.

In the table below, the notation $N(a, b)$ denotes a Normal distribution with mean a and variance b; and $IG(\nu, S)$ denotes an Inverse Gamma distribution with degrees of freedom parameter ν and scale parameter S.

Table A1: Prior settings

Parameter	**Parameter Description**	**Prior**
a^r	Coefficient on interest rate gap in output gap equation	$N(0, 1)$
ρ_1^g	Persistence in output gap: lag 1	$N(1.3, 0.1^2)$
ρ_2^g	Persistence in output gap: lag 2	$N(-0.5, 0.1^2)$
ρ_1^u	Persistence in UR gap: lag 1	$N(1.3, 0.1^2)$
ρ_2^u	Persistence in UR gap: lag 2	$N(-0.5, 0.1^2)$
ρ^p	Persistence in productivity gap	$N(0.1, 1)$
ζ	Relationship between r* and g*	$N(1, 0.1)$
ρ^i	Persistence in interest rate gap	$N(0.85, 0.1^2)$
λ^i	Interest rate sensitivity to UR gap: $(-2 * (1 - \rho^i))$	$N(-0.3, 0.1^2)$
κ^i	Interest rate sensitivity to inflation: $(0.5 * (1 - \rho^i))$	$N(0.075, 0.1^2)$
λ^g	Output gap response to UR gap	$N(-0.02, 1)$
ϕ^u	UR gap response to output gap	$N(-0.02, 1)$
β^g	Link between g* and survey	$N(1, 0.1^2)$
β^u	Link between u* and survey	$N(1, 0.05^2)$
β^r	Link between r* and survey	$N(1, 0.1^2)$
β^π	Link between π^* and survey	$N(1, 0.05^2)$
$\sigma^2_{\pi^*}$	Var. of the shocks to π^*	$IG(10, 0.1^2 \times 9)$
$\sigma^2_{p^*}$	Var. of the shocks to p^*	$IG(10, 0.142^2 \times 9)$
$\sigma^2_{u^*}$	Var. of the shocks to u^*	$IG(10, 0.1^2 \times 9)$
$\sigma^2_{gdp^*}$	Var. of the shocks to gdp^*	$IG(10, 0.01^2 \times 9)$
σ^2_d	Var. of the shocks to d	$IG(10, 0.1^2 \times 9)$
$\sigma^2_{w^*}$	Var. of the shocks to w^*	$IG(10, 0.03^2 \times 9)$
σ^2_{ho}	Var. of the volatility – Ogap eq.	$IG(10, 0.707^2 \times 9)$
σ^2_{hu}	Var. of the volatility – UR gap eq.	$IG(10, 0.707^2 \times 9)$
σ^2_{hp}	Var. of the volatility – Productivity eq.	$IG(10, 0.316^2 \times 9)$
σ^2_h	Var. of the volatility – Price Inf. eq.	$IG(10, 0.316^2 \times 9)$
σ^2_{hw}	Var. of the volatility – Wage Inf. eq.	$IG(10, 0.316^2 \times 9)$

Continued on next page

Table A1 – continued from previous page

Parameter	Parameter Description	Prior
σ^2_{hi}	Var. of the volatility – Interest rate eq.	$IG(10, 0.316^2 \times 9)$
$\sigma^2_{\lambda\pi}$	Var. of the shocks to TVP λ^π, Price Phillips curve	$IG(10, 0.04^2 \times 9)$
$\sigma^2_{\lambda w}$	Var. of the shocks to TVP λ^w, Wage Phillips curve	$IG(10, 0.04^2 \times 9)$
$\sigma^2_{\lambda p}$	Var. of the shocks to TVP λ^p, Cyc. Productivity	$IG(10, 0.04^2 \times 9)$
$\sigma^2_{\kappa w}$	Var. of the shocks to TVP κ^w, PT: π to Wages	$IG(10, 0.04^2 \times 9)$
$\sigma^2_{\rho w}$	Var. of the shocks to TVP ρ^w, Persist. Wage-gap	$IG(10, 0.04^2 \times 9)$
$\sigma^2_{\rho\pi}$	Var. of the shocks to TVP ρ^π, Persist. Inflation-gap	$IG(10, 0.04^2 \times 9)$
C^π_0	Time-varying Intercept in eq. linking survey to pi-star	$N(0, 0.1)$
C^u_0	Time-varying Intercept in eq. linking survey to u-star	$N(0, 0.1)$
C^g_0	Time-varying Intercept in eq. linking survey to g-star	$N(0, 0.1)$
C^r_0	Time-varying Intercept in eq. linking survey to r-star	$N(0, 0.1)$
$\sigma^2_{c\pi}$	Var. of the shocks to TVP C^π	$IG(10, 0.1^2 \times 9)$
σ^2_{cu}	Var. of the shocks to TVP C^u	$IG(10, 0.1^2 \times 9)$
σ^2_{cg}	Var. of the shocks to TVP C^g	$IG(10, 0.1^2 \times 9)$
σ^2_{cr}	Var. of the shocks to TVP C^r	$IG(10, 0.1^2 \times 9)$
σ^2_{wlr}	Var. of the shocks to TVP $Wedge$	$IG(10, 0.1^2 \times 9)$
$\sigma^2_{z\pi}$	Var. of the shocks in measurement eq. Z^π,	$IG(10, 0.2 \times 9)$
σ^2_{zu}	Var. of the shocks in measurement eq. Z^u,	$IG(10, 0.3 \times 9)$
σ^2_{zg}	Var. of the shocks in measurement eq. Z^g,	$IG(10, 0.1 \times 9)$
σ^2_{zr}	Var. of the shocks in measurement eq. Z^r,	$IG(10, 0.2 \times 9)$
π^*_0	Initial value of pi-star	$N(3, 5^2)$
u^*_0	Initial value of u-star, $t = 0$	$N(5, 5^2)$
u^*_{-1}	Initial value of u-star, $t = -1$	$N(5, 5^2)$
p^*_0	Initial value of p-star	$N(3, 5^2)$
w^*_0	Initial value of w-star, $E(p^*_0) + E(\pi^*_0) = 6$	$N(6, 5^2)$
D_0	Initial value of D, "catch-all" component of r-star	$N(0, 0.3162^2)$
gdp^*_0	Initial value of gdp-star, $t = 0$	$N(750, 10^2)$
gdp^*_{-1}	Initial value of gdp-star, $t = -1$	$N(750, 10^2)$

A1.c. MCMC algorithm

The estimation of our complex UC model and sampling from its joint posterior distribution reduces to sequentially drawing from a set of conditional posterior densities, some of which are standard and some that are non-standard.

Collect all the time-invariant model parameters into θ:
$\theta = (\rho_1^u, \rho_2^u, \sigma_{hu}^2, \phi_u, \sigma_{u*}^2, \beta^u, \sigma_{zu}^2, \sigma_{cu}^2, \sigma_{gdp*}^2, \rho_1^g, \rho_2^g, a^r, \lambda^g, \sigma_{ho}^2, \sigma_{zg}^2, \sigma_{cg}^2, \beta^g, \rho^p, \sigma_{hp}^2, \sigma_{p*}^2, \sigma_{\lambda\pi}^2, ...$
$\sigma_{\rho\pi}^2, \sigma_{h\pi}^2, \sigma_{\pi*}^2, \sigma_{z\pi}^2, \sigma_{c\pi}^2, \beta^\pi, \sigma_{w*}^2, \sigma_{hw}^2, \sigma_{\rho w}^2, \sigma_{\lambda w}^2, \sigma_{\kappa w}^2, \rho^i, \lambda^i, \kappa^i, \sigma_{hi}^2, \sigma_{zr}^2, \sigma_{cr}^2, \sigma_{wlr}^2, \beta^r, \sigma_d^2)$

We denote $\bullet$ as representing all other model parameters.
1. $p(U^*|Y,\bullet)$ 2. $p(gdp^*|Y,\bullet)$ 3. $p(P^*|Y,\bullet)$ 4. $p(\pi^*|Y,\bullet)$ 5. $p(w^*|Y,\bullet)$ 6. $p(r^*|Y,\bullet)$
7. $p(\lambda^p|Y,\bullet)$ 8. $p(\rho^\pi|Y,\bullet)$ 9. $p(\lambda^\pi|Y,\bullet)$ 10. $p(\rho^w|Y,\bullet)$ 11. $p(\lambda^w|Y,\bullet)$ 12. $p(\kappa^w|Y,\bullet)$
13. $p(h^p, h^\pi, h^w, h^i|Y,\bullet)$ 14. $p(C^u, C^g, C^\pi, C^r, Wedge|Y,\bullet)$ 15. $p(D|Y,\bullet)$ 16. $p(\theta|Y,\bullet)$

Step 1. Derive the conditional distribution $p(U^*|Y,\bullet)$

The derivation of this distribution is most complex because the information about U^* comes from eight sources (i.e., model equations). Below, we derive an expression for each of the eight sources.

The first source is the state equation of U^*. We rewrite it in a matrix notation as follows,

$$HU^* = \alpha_u + \varepsilon^{u*} \qquad \varepsilon^{u*} \sim N(0, \Omega_{u*}), \quad where\ \ \Omega_{u*} = diag(\omega_{u*}^2, \sigma_{u*}^2, ..., \sigma_{u*}^2) \tag{38}$$

where,

$$\alpha_u = \begin{pmatrix} U_0^* \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{pmatrix}$$

That is, the prior density for U^* is given by

$$p(U^*|\sigma_{U*}^2) \propto -\tfrac{1}{2}(U^* - H^{-1}\alpha_u)' H' \Omega_{u*}^{-1} H (U^* - H^{-1}\alpha_u) + g_{u*}(U^*, \sigma_{u*}^2)$$

where,
$a_u < U^* < b_u$ for $t = 1, ..., T$, and

$$g_{u*}(U^*, \sigma_{u*}^2) = -log\left(\Phi\left(\frac{b_u}{\omega_{u_*}}\right) - \Phi\left(\frac{a_u}{\omega_{u_*}}\right)\right) - \sum_{t=2}^{T} log\left(\Phi\left(\frac{b_u - U_{t-1}^*}{\sigma_{u_*}}\right) - \Phi\left(\frac{a_u - U_{t-1}^*}{\sigma_{u_*}}\right)\right)$$

The second source of information comes from the unemployment measurement equation. Rewrite

the equation in a matrix notation,

$$K_u U = \mu^u + K_u U^* + \varepsilon^u \qquad \varepsilon^u \sim N(0, \Omega_u), \quad where\ \ \Omega_u = diag(e^{h_1^u}, e^{h_2^u}, ..., e^{h_T^u}) \tag{39}$$

and,

$$\mu_u = \begin{pmatrix} \rho_1^u(U_0 - U_0^*) + \rho_2^u(U_{-1} - U_{-1}^*) \\ \rho_2^u(U_0 - U_0^*) \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad K_u = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\rho_1^u & 1 & 0 & \cdots & 0 \\ -\rho_2^u & -\rho_1^u & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & -\rho_2^u & -\rho_1^u & 1 \end{pmatrix}$$

Ignoring any terms not involving U^*, we have

$$log\ \ p(U|U^*, \bullet) \propto -\tfrac{1}{2}(U - K_u^{-1}\mu_u - U^*)' K_u' \Omega_u^{-1} K_u (U - K_u^{-1}\mu_u - U^*)$$

The third source of information comes from the inflation measurement equation. Rewrite the equation in a matrix notation,

$$Z = \Lambda^\pi U^* + \varepsilon^\pi \qquad \varepsilon^\pi \sim N(0, \Omega_\pi), \quad where\ \ \Omega_\pi = diag(e^{h_1^\pi}, e^{h_2^\pi}, ..., e^{h_T^\pi}) \tag{40}$$

where,

$z_t = (\pi_t - \pi_t^*) - \rho_t^\pi(\pi_{t-1} - \pi_{t-1}^*) - \lambda_t^\pi U_t,$

$Z = (z_1, ..., z_T)'$ and $\Lambda^\pi = diag(-\lambda_1^\pi, ..., -\lambda_T^\pi)$

Ignoring any terms not involving U^*, we have

$$log\ \ p(\pi|U^*, U, \pi^*, h^\pi, \rho^p, \bullet) \propto -\tfrac{1}{2}(Z - \Lambda^\pi U^*)' \Omega_\pi^{-1} (Z - \Lambda^\pi U^*)$$

The fourth source of information comes from the productivity measurement equation. Rewrite the equation in a matrix notation,

$$M^P = \Lambda^P U^* + \varepsilon^P \qquad \varepsilon^P \sim N(0, \Omega_P), \quad where\ \ \Omega_P = diag(e^{h_1^p}, e^{h_2^p}, ..., e^{h_T^p}) \tag{41}$$

where,

$m_t = (P_t - P_t^*) - \rho^P(P_{t-1} - P_{t-1}^*) - \lambda_t^P U_t,$

$M^P = (m_1, ..., m_T)'$ and $\Lambda^P = diag(-\lambda_1^P, ..., -\lambda_T^P)$

Ignoring any terms not involving U^*, we have

$$log\ \ p(P|U^*, U, P^*, h^p, \rho^p, \bullet) \propto -\tfrac{1}{2}(M^P - \Lambda^P U^*)' \Omega_P^{-1} (M^P - \Lambda^P U^*)$$

The fifth source of information comes from the wage measurement equation. Rewrite the equation in a matrix notation,

$$M^w = \Lambda^w U^* + \varepsilon^w \qquad \varepsilon^w \sim N(0, \Omega_w), \quad where\ \ \Omega_w = diag(e^{h_1^w}, e^{h_2^w}, ..., e^{h_T^w}) \tag{42}$$

where,

$m_t^w = (W_t - W_t^*) - \rho_t^W (W_{t-1} - W_{t-1}^*) - \lambda_t^W U_t - \kappa_t^W (\pi_t - \pi_t^*)$,

$M^w = (m_1^w, ..., m_T^w)'$ and $\Lambda^w = diag(-\lambda_1^W, ..., -\lambda_T^W)$

Ignoring any terms not involving U^*, we have

$log\ \ p(W|U^*, W, W^*, h^w, \rho^W, \bullet) \propto -\frac{1}{2}(M^w - \Lambda^w U^*)' \Omega_w^{-1} (M^w - \Lambda^w U^*)$

The sixth source of information comes from the output gap measurement equation. Rewrite the equation in a matrix notation,

$$M^g = \Lambda^g U^* + \varepsilon^g \qquad \varepsilon^g \sim N(0, \Omega_{ogap}), \quad where\ \ \Omega_{ogap} = diag(e^{h_1^o}, e^{h_2^o}, ..., e^{h_T^o}) \tag{43}$$

where,

$m_t^g = ogap_t - \rho_1^g (ogap_{t-1}) - \rho_2^g (ogap_{t-2}) - \lambda^g U_t - a^r (r_t - r_t^*)$,

$M^g = (m_1^g, ..., m_T^g)'$ and $\Lambda^g = diag(-\lambda^g, ..., -\lambda^g)$

Ignoring any terms not involving U^*, we have

$log\ \ p(ogap|U^*, U, \bullet) \propto -\frac{1}{2}(M^g - \Lambda^g U^*)' \Omega_{ogap}^{-1} (M^g - \Lambda^g U^*)$

The seventh source of information comes from the Taylor-type rule measurement equation. Rewrite the equation in a matrix notation,

$$M^{ui} = \Gamma^{ui} U^* + \varepsilon^i \qquad \varepsilon^i \sim N(0, \Omega_i), \quad where\ \ \Omega_i = diag(e^{h_1^i}, e^{h_2^i}, ..., e^{h_T^i}) \tag{44}$$

where,

$m_t^{ui} = i_t - \pi_t^* - r_t^* - \rho^i (i_{t-1} - \pi_{t-1}^* - r_{t-1}^*) - \kappa^i (\pi_t - \pi_t^*) - \lambda^i U_t$,

$M^{ui} = (m_1^{ui}, ..., m_T^{ui})'$ and $\Gamma^{ui} = diag(-\lambda^i, ..., -\lambda^i)$

Ignoring any terms not involving U^*, we have

$log\ \ p(i|U^*, U, \pi, \bullet) \propto -\frac{1}{2}(M^{ui} - \Gamma^{ui} U^*)' \Omega_i^{-1} (M^{ui} - \Gamma^{ui} U^*)$

The eighth source of information comes from the measurement equation that links surveys

to U^*. Rewrite the equation in a matrix notation,

$$F^u = \beta^u U^* + \varepsilon^{zu} \qquad \varepsilon^{zu} \sim N(0, \Omega_{zu}), \quad where \ \Omega_{zu} = diag(\sigma^2_{zu}, ..., \sigma^2_{zu}) \tag{45}$$

where,

$f^u_t = Z^u_t - C^u_t,$

$F^u = (f^u_1, ..., f^u_T)'$

Ignoring any terms not involving U^*, we have

$log \ p(Z^u|U^*, U, \pi, \bullet) \propto -\frac{1}{2}(F^u - \beta^u U^*)' \Omega^{-1}_{zu}(F^u - \beta^u U^*)$

Combining the above eight conditional densities we obtain,

$log \ p(U^*|Y, \bullet) \propto -\frac{1}{2}(U^* - \hat{U}^*)' D^{-1}_{U^*}(U^* - \hat{U}^*) + g_{u*}(U^*, \sigma^2_{u*})$

where,

$D_{U^*} = (H'\Omega^{-1}_{U^*}H + K'_u\Omega^{-1}_u K_u + \Lambda^{\pi'}\Omega^{-1}_\pi \Lambda^\pi + \Lambda^{w'}\Omega^{-1}_w \Lambda^w + \Lambda^{g'}\Omega^{-1}_{ogap}\Lambda^g + \Gamma^{ui'}\Omega^{-1}_i \Gamma^{ui} + \Lambda^{P'}\Omega^{-1}_P \Lambda^P + (\beta^u)^2 \Omega^{-1}_{zu})^{-1}$

$\hat{U}^* = D_{U*}(H'\Omega^{-1}_{U^*}\alpha_u + K'_u\Omega^{-1}_u K_u(U - K^{-1}_u \mu_u) + \Lambda^{\pi'}\Omega^{-1}_\pi Z + \Lambda^{w'}M^w + \Lambda^w + \Lambda^{g'}\Omega^{-1}_{ogap}M^g + \Gamma^{ui'}\Omega^{-1}_i M^{ui} + \Lambda^{P'}\Omega^{-1}_P M^P + \beta^u \Omega^{-1}_{zu} F^u)$

The addition of the term $g_{u*}(U^*, \sigma^2_{u*})$ leads to a non-standard density. Accordingly, we sample U^* using an independence-chain Metropolis-Hastings (MH) procedure. This involves first generating candidate draws from $N(\hat{U}^*, D_{U*})$ using the precision-based algorithm (of Chan and Jeliazkov, 2009) that are then accepted or rejected based on the accept-reject Metropolis-Hastings (ARMH) algorithm (discussed in Chan and Strachan, 2012).

Step 2. Derive the conditional distribution $p(gdp^*|Y, \bullet)$

The information about gdp^* comes from five sources. Below, we derive an expression for each of these sources.

The first source is the state equation of gdp^*. We rewrite it in a matrix notation as follows,

$$H_2 gdp^* = \alpha_{gdp*} + \varepsilon^{gdp*} \qquad \varepsilon^{gdp*} \sim N(0, \Omega_{gdp*}), \quad where \ \Omega_{gdp*} = diag(\omega^2_{gdp*}, \sigma^2_{gdp*}, ..., \sigma^2_{gdp*}) \tag{46}$$

where,

$$\alpha_{gdp*} = \begin{pmatrix} gdp_0^* + \triangle gdp_0^* \\ -gdp_0^* \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad H_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -2 & 1 & 0 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -2 & 1 \end{pmatrix}$$

H_2 is a band matrix with unit determinant and hence is invertible.

The prior density for gdp^* is given by

$$p(gdp^*|\sigma^2_{gdp*}) \propto -\tfrac{1}{2}(gdp^* - H_2^{-1}\alpha_{gdp*})' H_2' \Omega^{-1}_{gdp*} H_2 (gdp^* - H_2^{-1}\alpha_{gdp*})$$

The second source of information about gdp^* is from the output gap measurement equation. Rewrite in matrix form,

$$H_{rhog} gdp = H_{rhog} gdp^* + a^r \tilde{r} + \lambda^g \tilde{u} + \alpha_{gmore} + \varepsilon^{ogap} \qquad \varepsilon^{ogap} \sim N(0, \Omega_{ogap}), \quad where \ \Omega_{ogap} = diag(e^{h_1^o}, e^{h_2^o}, ..., e^{h_T^o}) \tag{47}$$

where,

$$\alpha_{gmore} = \begin{pmatrix} \rho_1^g(gdp_0 - gdp_0^*) + \rho_2^g(gdp_{-1} - gdp_{-1}^*) \\ \rho_2^g(gdp_0 - gdp_0^*) \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad H_{rhog} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -\rho_1^g & 1 & 0 & 0 & \cdots & 0 \\ -\rho_2^g & -\rho_1^g & 1 & 0 & \cdots & 0 \\ 0 & -\rho_2^g & -\rho_1^g & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\rho_2^g & -\rho_1^g & 1 \end{pmatrix},$$

$$\tilde{r} = \begin{pmatrix} r_1 - r_1^* \\ r_2 - r_2^* \\ r_3 - r_3^* \\ \vdots \\ r_T - r_T^* \end{pmatrix} \quad \tilde{u} = \begin{pmatrix} U_1 - U_1^* \\ U_2 - U_2^* \\ U_3 - U_3^* \\ \vdots \\ U_T - U_T^* \end{pmatrix}$$

$log\ \ p(gdp|gdp^*, \bullet) \propto -\frac{1}{2}(gdp - H^{-1}_{rhog}(H_{rhog} gdp^* + a^r \tilde{r} + \lambda^g \tilde{u} + \alpha_{gmore}))' H'_{rhog} \Omega^{-1}_{ogap} H_{rhog} (gdp - H^{-1}_{rhog}(H_{rhog} gdp^* + a^r \tilde{r} + \lambda^g \tilde{u} + \alpha_{gmore}))$

The third source of information comes from the unemployment gap measurement equation. Rewrite that equation in matrix notation,

$$Y^{ugdp} = \Gamma^u gdp^* + \varepsilon^u \qquad \varepsilon^u \sim N(0, \Omega_u), \quad where \ \Omega_u = diag(e^{h_1^u}, e^{h_2^u}, ..., e^{h_T^u}) \tag{48}$$

where,

$y_t^{ugdp} = \tilde{u_t} - \rho_1^u \tilde{u_{t-1}} - \rho_2^u \tilde{u_{t-2}} - \phi^u gdp$, where $\tilde{u_t} = (U_t - U_t^*)$

$Y^{ugdp} = (y_1^{ugdp}, ..., y_T^{ugdp})'$

Ignoring any terms not involving gdp^*, we have

$$log\ \ p(U|gdp^*, \bullet) \propto -\tfrac{1}{2}(Y^{ugdp} - \Gamma^u gdp^*)'\Omega_u^{-1}(Y^{ugdp} - \Gamma^u gdp^*)$$

The fourth source of information comes from the equation linking r-star to g-star, i.e.,

$$r_t^* = \zeta(gdp_t^* - gdp_{t-1}^*) + D_t \tag{49}$$

Rewrite this equation in matrix notation,

$$r^* = \zeta H gdp^* + \alpha_{gr} + D \tag{50}$$

where,

$\alpha_{gr} = (-\zeta gdp_0^*, 0, 0, ..., 0)'$

Ignoring any terms not involving gdp^*, we have

$$log\ \ p(r^*|gdp^*, D, \bullet) \propto -\tfrac{1}{2}(r^* - (\zeta H gdp^* + \alpha_{gr} + D))'(r^* - (\zeta H gdp^* + \alpha_{gr} + D))$$

The fifth source of information comes from the measurement equation that links surveys to g^*. Rewrite the equation in a matrix notation,

$$F^g = \beta^g(H gdp^* - \alpha_g) + \varepsilon^{zg} \qquad \varepsilon^{zg} \sim N(0, \Omega_{zg}), \quad where\ \ \Omega_{zg} = diag(\sigma_{zg}^2, ..., \sigma_{zg}^2) \tag{51}$$

where,

$f_t^g = Z_t^g - C_t^g$, $F^g = (f_1^g, ..., f_T^g)'$

$\alpha_g = (gdp_0^*, 0, 0, ..., 0)'$ is a $T \times 1$ vector.

Ignoring any terms not involving gdp^*, we have

$$log\ \ p(Z^g|gdp^*, \bullet) \propto -\tfrac{1}{2}(F^g - \beta^g(H gdp^* - \alpha_g))'\Omega_{zg}^{-1}(F^g - \beta^g(H gdp^* - \alpha_g))$$

Combining the above five conditional densities we obtain,

$$log\ \ p(gdp^*|Y, \bullet) \propto -\tfrac{1}{2}(gdp^* - \hat{gdp}^*)' D_{gdp*}^{-1}(gdp^* - \hat{gdp}^*)$$

where,

$D_{gdp*} = (H_2'\Omega_{gdp*}^{-1}H_2 + H_{rhog}'\Omega_{ogap}^{-1}H_{rhog} + \Gamma^{u'}\Omega_u^{-1}\Gamma^u + (\zeta H)'(\zeta H) + \beta^g H'\Omega_{zg}^{-1}\beta^g H)^{-1}$

$\hat{gdp}^* = D_{gdp*}(H_2'\Omega_{gdp*}^{-1}H_2\alpha_{gdp*} + H_{rhog}'\Omega_{ogap}^{-1}(H_{rhog}gdp - a^r\tilde{r} - \lambda^g\tilde{u} - \alpha_{gmore}) + \Gamma^{u'}\Omega_u^{-1}Y^{ugdp} + (\zeta H)'(r^* - \alpha_{gr} + D) + \beta^g H'\Omega_{zg}^{-1}F^g)$

Step 3. Derive the conditional distribution $p(P^*|Y, \bullet)$

First, rewrite the productivity measurement eq. as

$$K_p P = \mu_p + K_p P^* + \varepsilon^P \qquad \varepsilon^P \sim N(0, \Omega_P), \quad where \ \ \Omega_P = diag(e^{h_1^p}, e^{h_2^p}, ..., e^{h_T^p}) \tag{52}$$

$$\mu_p = \begin{pmatrix} \rho_1^P(P_0 - P_0^*) + \lambda_1^P(U_1 - U_1^*) \\ \lambda_2^P(U_2 - U_2^*) \\ \lambda_3^P(U_3 - U_3^*) \\ \vdots \\ \lambda_T^P(U_T - U_T^*) \end{pmatrix}, \quad K_P = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\rho_2^P & 1 & 0 & \cdots & 0 \\ 0 & -\rho_3^P & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & -\rho_T^P & 1 \end{pmatrix}, \quad P^* = \begin{pmatrix} P_1^* \\ P_2^* \\ P_3^* \\ \vdots \\ P_T^* \end{pmatrix}$$

Since $\mid K_P \mid = 1$ for any ρ_P, K_P is invertible. Therefore, we have likelihood

$$p(P|P^*, U, \bullet) \sim N(K_P^{-1}\mu_P + P^*, \ \ (K_P^{'}\Omega_P^{-1}K_P)^{-1})$$

i.e.,

$$log \ \ p(P|U, \bullet) \propto -\tfrac{1}{2}\iota_T h^P - \tfrac{1}{2}(P - K_P^{-1}\mu_P - P^*)^{'} K_P^{'}\Omega_P^{-1}K_P(P - K_P^{-1}\mu_P - P^*),$$

where ι_T is a $T \times 1$ column of ones.

Similarly, rewrite the state equation for P^* as

$$HP^* = \alpha_p + \varepsilon^{P*} \qquad \varepsilon^{P*} \sim N(0, \Omega_{P*}), \quad where \ \ \Omega_{P*} = diag(\omega_{P*}^2, \sigma_{P*}^2, ..., \sigma_{P*}^2) \tag{53}$$

where,

$$\alpha_p = \begin{pmatrix} P_0^* \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad K_P = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{pmatrix}$$

That is, the prior density for P^* is given by

$$p(P^*|\sigma_{P*}^2) \propto -\tfrac{1}{2}(P^* - H^{-1}\alpha_p)^{'} H^{'}\Omega_{P*}^{-1}H(P^* - H^{-1}\alpha_p)$$

Now account for the third source of information about P* in the equation $W^* = P^* + \pi^* + Wedge + \varepsilon^{w*}$,

$$p(P^*|W^*, \pi^*, \sigma_{W*}^2) \propto -\tfrac{1}{2}(P^* - (W^* - \pi^* - Wedge))^{'}\Omega_{W*}^{-1}(P^* - (W^* - \pi^* - Wedge))$$

where,

$$\Omega_{W*} = diag(\sigma^2_{W*}, \sigma^2_{W*}, ..., \sigma^2_{W*}),\ \ W^* = (W^*_1, ..., W^*_T)',\ \ \pi^* = (\pi^*_1, ..., \pi^*_T)',\ \ Wedge = (Wedge_1, ..., Wedge_T)'$$

Combining the above three conditional densities we obtain,

$$log\ \ p(P^*|Y, \bullet) \propto -\tfrac{1}{2}(P^* - \hat{P}^*)' D^{-1}_{P*}(P^* - \hat{P}^*)$$

where,

$$D_{P*} = (H' \Omega^{-1}_{P*} H + K'_P \Omega^{-1}_P K_P + \Omega^{-1}_{W*})^{-1}$$

$$\hat{P}^* = D_{P*}(H^{-1}\Omega^{-1}_{P*}\alpha_p + K'_P \Omega^{-1}_P K_P(P - K^{-1}_P \mu_P) + \Omega^{-1}_{W*}(W^* - \pi^*))$$

The candidate draws are sampled from $N(\hat{P}^*, D_{P*})$ using the precision-based algorithm.

Step 4. Derive the conditional distribution $p(\pi^*|Y, \bullet)$

The information about π^* comes from six sources. Below, we derive an expression for each of these sources.

The first source is the inflation measurement equation. Rewrite it in a matrix notation as,

$$K_\pi \pi = \mu_\pi + K_\pi \pi^* + \varepsilon^\pi \qquad \varepsilon^\pi \sim N(0, \Omega_\pi), \quad where\ \ \Omega_\pi = diag(e^{h^\pi_1}, e^{h^\pi_2}, ..., e^{h^\pi_T}) \tag{54}$$

where,

$$\mu_\pi = \begin{pmatrix} \rho^\pi_1(\pi_0 - \pi^*_0) + \lambda^\pi_1(U_1 - U^*_1) \\ \lambda^\pi_2(U_2 - U^*_2) \\ \lambda^\pi_3(U_3 - U^*_3) \\ \vdots \\ \lambda^\pi_T(U_T - U^*_T) \end{pmatrix}, \quad K_\pi = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\rho^\pi_2 & 1 & 0 & \cdots & 0 \\ 0 & -\rho^\pi_3 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & -\rho^\pi_T & 1 \end{pmatrix}$$

Since $\mid K_\pi \mid = 1$ for any ρ_π, K_π is invertible. Therefore, we have likelihood

$$log\ \ p(\pi|U, U^*, \bullet) \propto -\tfrac{1}{2}\iota_T h^\pi - \tfrac{1}{2}(\pi - (K^{-1}_\pi \mu_\pi + \pi^*))' K'_\pi \Omega^{-1}_\pi K_\pi (\pi - (K^{-1}_\pi \mu_\pi + \pi^*))$$

The second source of information is from the state equation of π^*. Rewrite it in a matrix notation,

$$H\pi^* = \alpha_\pi + \varepsilon^{\pi*} \qquad \varepsilon^{\pi*} \sim N(0, \Omega_{\pi*}), \quad where\ \ \Omega_{\pi*} = diag(\omega^2_{\pi*}, \sigma^2_{\pi*}, ..., \sigma^2_{\pi*}) \tag{55}$$

where,

$$\alpha_\pi = \begin{pmatrix} \pi_0^* \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

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

$$p(\pi^*|\sigma^2_{\pi*}) \propto -\tfrac{1}{2}(\pi^* - H^{-1}\alpha_\pi)' H' \Omega^{-1}_{\pi*} H(\pi^* - H^{-1}\alpha_\pi)$$

Now account for the third source of information about π^* in the equation $W^* = P^* + \pi^* + Wedge + \varepsilon^{w*}$,

$$p(\pi^*|W^*, P^*, \sigma^2_{W*}) \propto -\tfrac{1}{2}(\pi^* - (W^* - P^* - Wedge))' \Omega^{-1}_{W*}(\pi^* - (W^* - P^* - Wedge))$$

where,

$$\Omega_{W*} = diag(\sigma^2_{W*}, \sigma^2_{W*}, ..., \sigma^2_{W*}),\ \ W^* = (W_1^*, ..., W_T^*)',\ \ P^* = (P_1^*, ..., P_T^*)',\ \ Wedge = (Wedge_1, ..., Wedge_T)'$$

The fourth source of information is from the wage measurement equation. Rewrite in matrix notation,

$$M^{w\pi} = X_{w\pi}\pi^* + \varepsilon^w \qquad \varepsilon^w \sim N(0, \Omega_w), \quad where\ \ \Omega_w = diag(e^{h_1^w}, e^{h_2^w}, ..., e^{h_T^w}) \tag{56}$$

where,

$$m_t^{w\pi} = w_t - w_t^* - \rho_t^w(w_{t-1} - w_{t-1}^*) - \lambda_t^w(U_t - U_t^*) - \kappa_t^w \pi_t$$
$$M^{w\pi} = (m_1^{w\pi}, m_2^{w\pi}, ..., m_T^{w\pi})$$

$$X_{w\pi} = \begin{pmatrix} -\kappa_1^w & 0 & 0 & \cdots & 0 \\ 0 & -\kappa_2^w & 0 & \cdots & 0 \\ 0 & 0 & -\kappa_3^w & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & -\kappa_T^w \end{pmatrix}$$

$$log\ \ p(W|\pi^*, \bullet) \propto -\tfrac{1}{2}(M^{w\pi} - X_{w\pi}\pi^*)' \Omega_w^{-1}((M^{w\pi} - X_{w\pi}\pi^*)$$

The fifth source is the Taylor-rule equation. Rewrite the equation in the matrix notation,

$$M^{\pi i} = \alpha_{\pi i} + (K_{\pi i} + \Gamma_\pi)\pi^* + \varepsilon^i \qquad \varepsilon^i \sim N(0, \Omega_i), \quad where\ \ \Omega_i = diag(e^{h_1^i}, e^{h_2^i}, ..., e^{h_T^i}) \tag{57}$$

where,

$m_t^{\pi i} = i_t - \rho^i i_{t-1} - r_t^* + \rho^i r_{t-1}^* - \lambda^i(U_t - U_t^*) - \kappa^i \pi_t$
$M^{\pi i} = (m_1^{\pi i}, m_2^{\pi i}, ..., m_T^{\pi i})'$

$$K_{\pi i} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\rho^i & 1 & 0 & \cdots & 0 \\ 0 & -\rho^i & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & -\rho^i & 1 \end{pmatrix}, \quad \Gamma_\pi = \begin{pmatrix} -\kappa^i & 0 & 0 & \cdots & 0 \\ 0 & -\kappa^i & 0 & \cdots & 0 \\ 0 & 0 & -\kappa^i & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & -\kappa^i \end{pmatrix}, \quad \alpha_{\pi i} = \begin{pmatrix} -\rho^i \pi_0^* \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$log\ \ p(i|\pi^*, \pi, \bullet) \propto -\tfrac{1}{2}(M^{\pi i} - (\alpha_{\pi i} + (K_{\pi i} + \Gamma_\pi)\pi^*))'\Omega_i^{-1}(M^{\pi i} - (\alpha_{\pi i} + (K_{\pi i} + \Gamma_\pi)\pi^*))$$

The sixth source of information comes from the measurement equation that links surveys to π^*. Rewrite the equation in a matrix notation,

$$F^\pi = \beta^\pi \pi^* + \varepsilon^{z\pi} \qquad \varepsilon^{z\pi} \sim N(0, \Omega_{z\pi}), \quad where\ \ \Omega_{z\pi} = diag(\sigma_{z\pi}^2, ..., \sigma_{z\pi}^2) \tag{58}$$

where,

$f_t^\pi = Z_t^\pi - C_t^\pi,$

$F^\pi = (f_1^\pi, ..., f_T^\pi)'$

Ignoring any terms not involving π^*, we have

$$log\ \ p(Z^\pi|\pi^*, \pi, \bullet) \propto -\tfrac{1}{2}(F^\pi - \beta^\pi \pi^*)'\Omega_{z\pi}^{-1}(F^\pi - \beta^\pi \pi^*)$$

Combining the above six conditional densities we obtain,

$$log\ \ p(\pi^*|Y, \bullet) \propto -\tfrac{1}{2}(\pi^* - \hat{\pi}^*)' D_{\pi^*}^{-1}(\pi^* - \hat{\pi}^*)$$

where,
$D_{\pi^*} = (H'\Omega_{\pi^*}^{-1}H + K_\pi'\Omega_\pi^{-1}K_\pi + \Omega_{w^*}^{-1} + X_{w\pi}'\Omega_w^{-1}X_{w\pi} + (K_{\pi i}' + \Gamma_\pi)')\Omega_i^{-1}(K_{\pi i}' + \Gamma_\pi)' + (\beta^\pi)^2\Omega_{zr}^{-1})^{-1}$

$\hat{\pi}^* = D_{\pi^*}(H'\Omega_{\pi^*}^{-1}\alpha_\pi + K_\pi'\Omega_\pi^{-1}K_\pi(\pi - K_\pi^{-1}\mu_\pi) + \Omega_{w^*}^{-1}(W^* - P^*) + X_{w\pi}'\Omega_w^{-1}M^{w\pi} + (K_{\pi i}' + \Gamma_\pi)'\Omega_i^{-1}(M^{\pi i} - \alpha_{\pi i}) + \beta^\pi\Omega_{zr}^{-1}F^\pi)$

The candidate draws are sampled from $N(\hat{\pi}^*, D_{\pi^*})$ using the precision-based algorithm.

Step 5. Derive the conditional distribution $p(w^*|Y, \bullet)$

The information about w^* comes from two sources. Below, we derive an expression for each of these sources.

The first source is the nominal wage measurement equation. Rewrite it in a matrix notation as,

$$K_w W = \mu_w + K_w W^* + \varepsilon^w \qquad \varepsilon^w \sim N(0, \Omega_w), \quad where \ \ \Omega_w = diag(e^{h_1^w}, e^{h_2^w}, ..., e^{h_T^w}) \tag{59}$$

where,

$$\mu_w = \begin{pmatrix} \rho_1^w(W_0 - W_0^*) + \lambda_1^w(U_1 - U_1^*) + \kappa_1^w(\pi_1 - \pi_1^*) \\ \lambda_2^w(U_2 - U_2^*) + \kappa_2^w(\pi_2 - \pi_2^*) \\ \lambda_3^w(U_3 - U_3^*) + \kappa_3^w(\pi_3 - \pi_3^*) \\ \vdots \\ \lambda_T^w(U_T - U_T^*) + \kappa_T^w(\pi_T - \pi_T^*) \end{pmatrix}, \quad K_w = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\rho_2^w & 1 & 0 & \cdots & 0 \\ 0 & -\rho_3^w & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & -\rho_T^w & 1 \end{pmatrix}$$

Since $\mid K_w \mid = 1$ for any ρ_w, K_w is invertible.

Ignoring any terms not involving w^*, we have the likelihood

$$log \ \ p(W|W^*, \bullet) \propto -\tfrac{1}{2}\iota_T h^w - \tfrac{1}{2}(W - (K_w^{-1}\mu_w + W^*))' K_w^{'} \Omega_w^{-1} K_w (W - (K_w^{-1}\mu_w + W^*))$$

The second source is the state equation of W^*, which describes W^* as the sum of P^* and π^*. This equation can be thought of as describing the prior density for W^*. Rewrite it in a matrix form.

$$W^* = P^* + \pi^* + Wedge + \varepsilon^{w*} \qquad \varepsilon^{w*} \sim N(0, \Omega_{w*}) \tag{60}$$

$$p(W^*|P^*, \pi^*, \sigma_{w*}^2) \propto -\tfrac{1}{2}(W^* - (P^* + \pi^* + Wedge))' \Omega_{w*}^{-1} (W^* - (P^* + \pi^* + Wedge))$$

Combining the above two conditional densities we obtain,

$$log \ \ p(W^*|Y, \bullet) \propto -\tfrac{1}{2}(W^* - \hat{W}^*)' D_{W*}^{-1} (W^* - \hat{W}^*)$$

where,

$$D_{W*} = (K_w^{'} \Omega_w^{-1} K_w + \Omega_{W*}^{-1})^{-1}$$

$$\hat{W}^* = D_{W*}(K_w^{'} \Omega_w^{-1}(K_w W - \mu_w) + \Omega_{w*}^{-1}(P^* + \pi^*))$$

The candidate draws are sampled from $N(\hat{W}^*, D_{W*})$ using the precision-based algorithm.

Step 6. Derive the conditional distribution $p(r^*|Y, \bullet)$

The information about r^* comes from four sources. Below, we derive an expression for each of these sources.

The first source is the output gap measurement equation. We rewrite it in a matrix notation as follows,

$$H_{rhog}ogap = \alpha_{ogap} - a^r r^* + \varepsilon^{ogap} \qquad \varepsilon^{ogap} \sim N(0, \Omega_{ogap}) \tag{61}$$

where,

$$\alpha_{ogap} = \begin{pmatrix} \rho_1^g(ogap_0) + \rho_2^g(ogap_{-1}) + a^r r_1 + \lambda^g(U_1 - U_1^*) \\ \rho_2^g(ogap_0) + a^r r_2 + \lambda^g(U_2 - U_2^*) \\ a^r r_3 + \lambda^g(U_3 - U_3^*) \\ \vdots \\ a^r r_T + \lambda^g(U_T - U_T^*) \end{pmatrix}$$

Ignoring any terms not involving r^*, we have
$log\ \ p(ogap|r^*, \bullet) \propto -\frac{1}{2}(ogap - H_{rhog}^{-1}(\alpha_{ogap} - a^r r^*))' H_{rhog}^{'} \Omega_{ogap}^{-1} H_{rhog}(ogap - H_{rhog}^{-1}(\alpha_{ogap} - a^r r^*))$

The second source is the state equation linking r^* to g^*. We rewrite it in a matrix notation as follows,

$$r^* = \zeta \triangle gdp^* + H^{-1}\varepsilon^d \qquad \varepsilon^d \sim N(0, \Omega_d), \quad where\ \ \Omega_d = diag(\omega_d^2, \sigma_d^2, ..., \sigma_d^2) \tag{62}$$

Ignoring any terms not involving r^*, the prior density for r^* is given by
$log\ \ p(r^*|gdp^*, \sigma_d^2, \bullet) \propto -\frac{1}{2}(r^* - \zeta \triangle gdp^*)' H' \Omega_d^{-1} H(r^* - \zeta \triangle gdp^*)$

The third source is the Taylor-type rule equation. We rewrite it in a matrix notation as follows,

$$M^{ri} = \alpha_{ri} + K_{\pi i} r^* + \varepsilon^i \quad \varepsilon^i \sim N(0, \Omega_i), \quad where\ \ \Omega_i = diag(e^{h_1^i}, e^{h_2^i}, ..., e^{h_T^i}) \tag{63}$$

where,

$$m_t^{ri} = i_t - \rho^i i_{t-1} - \pi_t^* + \rho^i \pi_{t-1}^* - \lambda^i(U_t - U_t^*) - \kappa^i(\pi_t - \pi_t^*),$$

$$M^{ri} = (m_1^{ri}, m_2^{ri}, ..., m_T^{ri})'$$

$$\alpha_{ri} = \begin{pmatrix} -\rho^i r_0^* \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad K_{\pi i} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\rho^i & 1 & 0 & \cdots & 0 \\ 0 & -\rho^i & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & -\rho^i & 1 \end{pmatrix}$$

Ignoring any terms not involving r^*, we have

$$log\ \ p(i|r^*,\bullet) \propto -\tfrac{1}{2}\iota_T h^i - \tfrac{1}{2}(M^{ri} - (\alpha_{ri} + K_{\pi i} r^*))' \Omega_i^{-1} (M^{ri} - (\alpha_{ri} + K_{\pi i} r^*))$$

The fourth source of information comes from the measurement equation that links surveys to r^*. Rewrite the equation in a matrix notation,

$$F^r = \beta^r r^* + \varepsilon^{zr} \qquad \varepsilon^{zr} \sim N(0, \Omega_{zr}), \quad where\ \ \Omega_{zr} = diag(\sigma_{zr}^2, ..., \sigma_{zr}^2) \tag{64}$$

where,

$f_t^r = Z_t^r - C_t^r,$

$F^r = (f_1^r, ..., f_T^r)'$

Ignoring any terms not involving r^*, we have

$$log\ \ p(Z^r|r^*,\bullet) \propto -\tfrac{1}{2}(F^r - \beta^r r^*)' \Omega_{zr}^{-1} (F^r - \beta^r r^*)$$

Combining the above four conditional densities we obtain,

$$log\ \ p(r^*|Y,\bullet) \propto -\tfrac{1}{2}(r^* - \hat{r}^*)' D_{r*}^{-1} (r^* - \hat{r}^*)$$

where,

$D_{r*} = ((-a^r)^2 \Omega_{ogap}^{-1} + H' \Omega_d^{-1} H + K_{\pi i}' \Omega_i^{-1} K_{\pi i} + (\beta^r)(2) \Omega_{zr}^{-1})^{-1}$

$\hat{r}^* = D_{r*}(-a^r \Omega_{ogap}^{-1} (H_{rhog} ogap - \alpha_{ogap}) + H' \Omega_d^{-1} H \zeta \triangle gdp^* + K_{\pi i}' \Omega_i^{-1} (M^{ri} - \alpha_{ri}) + \beta^r \Omega_{zr}^{-1} F^r)$

The candidate draws are sampled from $N(\hat{r}^*, D_{r*})$ using the precision-based algorithm.

Step 7. Derive the conditional distribution $p(\lambda^p|Y,\bullet)$

The information about λ^p comes from two sources. Below, we derive an expression for each of these two sources.

The first source is the productivity measurement equation. Rewrite it in a matrix notation,

$$B = X_u \lambda^p + \varepsilon^p \qquad \varepsilon^p \sim N(0, \Omega_p) \tag{65}$$

where,

$B = (\tilde{p}_1 - \rho^p \tilde{p}_0, ..., \tilde{p}_T - \rho^p \tilde{p}_{T-1})$
$\tilde{p}_t = p_t - p_t^*$
$\tilde{u}_t = U_t - U_t^*$
$X_u = diag(\tilde{u}_1, ..., \tilde{u}_T)$

Ignoring any terms not involving λ^p, we have the likelihood

$$log\ \ p(p|\lambda^p, \bullet) \propto -\tfrac{1}{2}(B - X_u\lambda^p)'\Omega_p^{-1}(B - X_u\lambda^p)$$

The second source of information comes from the state equation for λ^p. We rewrite it in a matrix notation as follows,

$$H\lambda^p = \varepsilon^{\lambda p} \qquad \varepsilon^{\lambda p} \sim N(0, \Omega_{\lambda p}), \quad where\ \ \Omega_{\lambda p} = diag(\omega_{\lambda p}^2, \sigma_{\lambda p}^2, ..., \sigma_{\lambda p}^2) \tag{66}$$

Ignoring any terms not involving λ^p, the prior density for λ^p is given by

$$log\ \ p(\lambda^p|\sigma_{\lambda p}^2, \Omega_{\lambda p}) \propto -\tfrac{1}{2}(\lambda^p)'H'\Omega_{\lambda p}^{-1}H(\lambda^p)$$

Combining the above two conditional densities we obtain,

$$log\ \ p(\lambda^p|Y, \bullet) \propto -\tfrac{1}{2}(\lambda^p - \hat{\lambda}^p)'D_{\lambda^p}^{-1}(\lambda^p - \hat{\lambda}^p)$$

where,

$D_{\lambda^p} = (H'\Omega_{\lambda p}^{-1}H + X_u'\Omega_p^{-1}X_u)^{-1}$

$\hat{\lambda}^p = D_{\lambda^p}(X_u'\Omega_p^{-1}B)$

The candidate draws are sampled from $N(\hat{\lambda}^p, D_{\lambda^p})$ using the precision-based algorithm.

Step 8. Derive the conditional distribution $p(\rho^\pi|Y, \bullet)$

The information about ρ^π comes from two sources. Below, we derive an expression for each of these two sources.

First, we define some notation,

$\tilde{\pi}_t = \pi_t - \pi_t^*$
$\tilde{u}_t = U_t - U_t^*$
$\tilde{\Pi} = (\tilde{\pi}_1, ..., \tilde{\pi}_T)'$
$\tilde{u} = (\tilde{u}_1, ..., \tilde{u}_T)'$

The first source is the price inflation measurement equation. Rewrite it in a matrix notation,

$$\tilde{\Pi} + \Lambda\tilde{u} = X_\pi\rho^\pi + \varepsilon^\pi \qquad \varepsilon^\pi \sim N(0, \Omega_\pi) \tag{67}$$

where,

$X_\pi = diag(\tilde{\pi}_0, ..., \tilde{\pi}_{T-1})$
$\Lambda = diag(-\lambda_1^\pi, ..., -\lambda_T^\pi)$

Ignoring any terms not involving ρ^π, we have the likelihood

$log\ \ p(\pi|\rho^{\pi}, \bullet) \propto -\frac{1}{2}(\tilde{\Pi} - (X_{\pi}\rho^{\pi} - \Lambda\tilde{u}))'\Omega_{\pi}^{-1}(\tilde{\Pi} - (X_{\pi}\rho^{\pi} - \Lambda\tilde{u}))$

The second source comes from the state equation for ρ^{π}. We rewrite it in a matrix notation as follows,

$$H\rho^{\pi} = \varepsilon^{\rho\pi} \qquad \varepsilon^{\rho\pi} \sim N(0, \Omega_{\rho\pi}), \quad where\ \ \Omega_{\rho\pi} = diag(\omega_{\rho\pi}^2, \sigma_{\rho\pi}^2, ..., \sigma_{\rho\pi}^2) \tag{68}$$

$0 < \rho_t^{\pi} < 1$ for t=1,....,T

Ignoring any terms not involving ρ^{π}, the prior density for ρ^{π} is given by

$log\ \ p(\rho^{\pi}|\sigma_{\rho\pi}^2, \Omega_{\rho\pi}) \propto -\frac{1}{2}(\rho^{\pi})'H'\Omega_{\rho\pi}^{-1}H(\rho^{\pi}) + g_{\rho\pi}(\rho^{\pi}, \sigma_{\rho\pi}^2)$

where,

$$g_{\rho\pi}(\rho^{\pi}, \sigma_{\rho\pi}^2) = -\sum_{t=2}^{T} log\left(\Phi\left(\frac{1-\rho_{t-1}^{\pi}}{\sigma_{\rho\pi}}\right) - \Phi\left(\frac{0-\rho_{t-1}^{\pi}}{\sigma_{\rho\pi}}\right)\right)$$

Combining the above two conditional densities we obtain,

$log\ \ p(\rho^{\pi}|Y, \bullet) \propto -\frac{1}{2}(\rho^{\pi} - \hat{\rho}^{\pi})' D_{\rho^{\pi}}^{-1}(\rho^{\pi} - \hat{\rho}^{\pi}) + g_{\rho\pi}(\rho^{\pi}, \sigma_{\rho\pi}^2)$

where,

$D_{\rho^{\pi}} = (H'\Omega_{\rho\pi}^{-1}H + X_{\pi}'\Omega_{\pi}^{-1}X_{\pi})^{-1}$

$\hat{\rho}^{\pi} = D_{\rho^{\pi}}(X_{\pi}'\Omega_{\pi}^{-1}(\tilde{\Pi} + \Lambda\tilde{u}))$

The addition of the term $g_{\rho\pi}(\rho^{\pi}, \sigma_{\rho\pi}^2)$ leads to a non-standard density. Accordingly, we sample ρ^{π} using an independence-chain Metropolis-Hastings (MH) procedure. This involves first generating candidate draws from $N(\hat{\rho}^{\pi}, D_{\rho^{\pi}})$ using the precision-based algorithm that are then accepted or rejected based on the accept-reject Metropolis-Hastings (ARMH) algorithm (discussed in Chan and Strachan, 2012).

Step 9. Derive the conditional distribution $p(\lambda^{\pi}|Y, \bullet)$

The information about λ^{π} comes from two sources. Below, we derive an expression for each of these two sources.

First, we define some notation,

$\tilde{\pi}_t = \pi_t - \pi_t^*$
$\tilde{u}_t = U_t - U_t^*$
$NW = (\tilde{\pi}_1 - \rho_1^{\pi}\tilde{\pi}_0, ..., \tilde{\pi}_T - \rho_T^{\pi}\tilde{\pi}_{T-1})'$

The first source is the price inflation measurement equation. Rewrite it in a matrix notation,

$$NW = X_u\lambda^\pi + \varepsilon^\pi \qquad \varepsilon^\pi \sim N(0, \Omega_\pi) \tag{69}$$

where,

$X_u = diag(\tilde{u}_1, ..., \tilde{u}_T)$

Ignoring any terms not involving λ^π, we have the likelihood

$$log\ \ p(\pi|\lambda^\pi, \bullet) \propto -\tfrac{1}{2}(NW - X_u\lambda^\pi)'\Omega_\pi^{-1}(NW - X_u\lambda^\pi)$$

The second source comes from the state equation for λ^π. We rewrite it in a matrix notation as follows,

$$H\lambda^\pi = \varepsilon^{\lambda\pi} \qquad \varepsilon^{\lambda\pi} \sim N(0, \Omega_{\lambda\pi}), \quad where\ \ \Omega_{\lambda\pi} = diag(\omega^2_{\lambda\pi}, \sigma^2_{\lambda\pi}, ..., \sigma^2_{\lambda\pi}) \tag{70}$$

$-1 < \lambda^\pi_t < 0$ for t=1,....,T

Ignoring any terms not involving λ^π, the prior density for λ^π is given by

$$log\ \ p(\lambda^\pi|\sigma^2_{\lambda\pi}, \Omega_{\lambda\pi}) \propto -\tfrac{1}{2}(\lambda^\pi)'H'\Omega^{-1}_{\lambda\pi}H(\lambda^\pi) + g_{\lambda\pi}(\lambda^\pi, \sigma^2_{\lambda\pi})$$

where,

$$g_{\lambda\pi}(\lambda^\pi, \sigma^2_{\lambda\pi}) = -\sum_{t=2}^{T} log\left(\Phi\left(\frac{0 - \lambda^\pi_{t-1}}{\sigma_{\lambda\pi}}\right) - \Phi\left(\frac{-1 - \lambda^\pi_{t-1}}{\sigma_{\lambda\pi}}\right)\right)$$

Combining the above two conditional densities we obtain,

$$log\ \ p(\lambda^\pi|Y, \bullet) \propto -\tfrac{1}{2}(\lambda^\pi - \hat{\lambda}^\pi)'D^{-1}_{\lambda^\pi}(\lambda^\pi - \hat{\lambda}^\pi) + g_{\lambda\pi}(\lambda^\pi, \sigma^2_{\lambda\pi})$$

where,

$D_{\lambda^\pi} = (H'\Omega^{-1}_{\lambda\pi}H + X'_u\Omega^{-1}_\pi X_u)^{-1}$

$\hat{\lambda}^\pi = D_{\lambda^\pi}(X'_u\Omega^{-1}_\pi NW)$

The addition of the term $g_{\lambda\pi}(\lambda^\pi, \sigma^2_{\lambda\pi})$ leads to a non-standard density. Accordingly, we sample λ^π using an independence-chain Metropolis-Hastings (MH) procedure. This involves first generating candidate draws from $N(\hat{\lambda}^\pi, D_{\lambda^\pi})$ using the precision-based algorithm that are then accepted or rejected based on the accept-reject Metropolis-Hastings (ARMH) algorithm (discussed in Chan and Strachan, 2012).

Step 10. Derive the conditional distribution $p(\rho^w|Y,\bullet)$

The information about ρ^w comes from two sources. Below, we derive an expression for each of these two sources.

First, we define some notation,

$\tilde{w}_t = w_t - w_t^*$
$\tilde{u}_t = U_t - U_t^*$
$\tilde{w} = (\tilde{w}_1, ..., \tilde{w}_T)'$
$\tilde{u} = (\tilde{u}_1, ..., \tilde{u}_T)'$
$\tilde{\pi}_t = \pi_t - \pi_t^*$
$\tilde{\pi} = (\tilde{\pi}_1, ..., \tilde{\pi}_T)'$

The first source is the wage inflation measurement equation. Rewrite it in a matrix notation,

$$\tilde{w} + \Lambda^w \tilde{u} + \Lambda^{w\pi}\tilde{\pi} = X_w \rho^w + \varepsilon^{\rho w} \qquad \varepsilon^{\rho w} \sim N(0, \Omega_w) \tag{71}$$

where,

$X_w = diag(\tilde{w}_0, ..., \tilde{w}_{T-1})$
$\Lambda^w = diag(-\lambda_1^w, ..., -\lambda_T^w)$
$\Lambda^{w\pi} = diag(-\kappa_1^w, ..., -\kappa_T^w)$

Ignoring any terms not involving ρ^w, we have the likelihood

$$log\ \ p(w|\rho^w, \bullet) \propto -\tfrac{1}{2}(\tilde{w} - (X_w\rho^w - \Lambda^w\tilde{u} - \Lambda^{w\pi}\tilde{\pi}))'\Omega_w^{-1}(\tilde{w} - (X_w\rho^w - \Lambda^w\tilde{u} - \Lambda^{w\pi}\tilde{\pi}))$$

The second source comes from the state equation for ρ^w. We rewrite it in a matrix notation as follows,

$$H\rho^w = \varepsilon^{\rho w} \qquad \varepsilon^{\rho w} \sim N(0, \Omega_{\rho w}), \quad where\ \ \Omega_{\rho w} = diag(\omega_{\rho w}^2, \sigma_{\rho w}^2, ..., \sigma_{\rho w}^2) \tag{72}$$

$0 < \rho_t^w < 1$ for t=1,....,T

Ignoring any terms not involving ρ^w, the prior density for ρ^w is given by

$log\ \ p(\rho^w|\sigma_{\rho w}^2, \Omega_{\rho w}) \propto -\tfrac{1}{2}(\rho^w)'H'\Omega_{\rho w}^{-1}H(\rho^w) + g_{\rho w}(\rho^w, \sigma_{\rho w}^2)$

where,

$$g_{\rho w}(\rho^w, \sigma^2_{\rho w}) = -\sum_{t=2}^{T} log\left(\Phi\left(\frac{1-\rho^w_{t-1}}{\sigma_{\rho w}}\right) - \Phi\left(\frac{0-\rho^w_{t-1}}{\sigma_{\rho w}}\right)\right)$$

Combining the above two conditional densities we obtain,

$$log\ \ p(\rho^w|Y,\bullet) \propto -\tfrac{1}{2}(\rho^w - \hat{\rho}^w)' D^{-1}_{\rho^w}(\rho^w - \hat{\rho}^w) + g_{\rho w}(\rho^w, \sigma^2_{\rho w})$$

where,

$D_{\rho^w} = (H'\Omega^{-1}_{\rho w} H + X'_w \Omega^{-1}_w X_w)^{-1}$

$\hat{\rho}^w = D_{\rho^w}(X'_w \Omega^{-1}_w(\tilde{w} + \Lambda^w \tilde{u} + \Lambda^{w\pi}\tilde{\pi}))$

The addition of the term $g_{\rho\pi}(\rho^\pi, \sigma^2_{\rho\pi})$ leads to a non-standard density. Accordingly, we sample ρ^π using an independence-chain Metropolis-Hastings (MH) procedure. This involves first generating candidate draws from $N(\hat{\rho}^\pi, D_{\rho^\pi})$ using the precision-based algorithm that are then accepted or rejected based on the accept-reject Metropolis-Hastings (ARMH) algorithm (discussed in Chan and Strachan, 2012).

Step 11. Derive the conditional distribution $p(\lambda^w|Y,\bullet)$

The information about λ^w comes from two sources. Below, we derive an expression for each of these two sources.

First, we define some notation,

$\tilde{w}_t = w_t - w^*_t$
$\tilde{u}_t = U_t - U^*_t$
$\tilde{\pi}_t = \pi_t - \pi^*_t$
$B^w = (\tilde{w}_1 - \rho^w_1 \tilde{w}_0 - \kappa^w_1 \tilde{\pi}_1, ..., \tilde{w}_T - \rho^w_T \tilde{w}_{T-1} - \kappa^w_{T-1}\tilde{\pi}_T)'$

The first source is the wage inflation measurement equation. Rewrite it in a matrix notation,

$$B^w = X_u \lambda^w + \varepsilon^w \qquad \varepsilon^w \sim N(0, \Omega_w) \tag{73}$$

where,

$X_u = diag(\tilde{u}_1, ..., \tilde{u}_T)$

Ignoring any terms not involving λ^w, we have the likelihood

$$log\ \ p(w|\lambda^w,\bullet) \propto -\tfrac{1}{2}(B^w - X_u\lambda^w)'\Omega^{-1}_w(B^w - X_u\lambda^w)$$

The second source comes from the state equation for λ^w. We rewrite it in a matrix notation as

follows,

$$H\lambda^w = \varepsilon^{\lambda w} \qquad \varepsilon^{\lambda w} \sim N(0, \Omega_{\lambda w}), \quad where\ \ \Omega_{\lambda w} = diag(\omega^2_{\lambda w}, \sigma^2_{\lambda w}, ..., \sigma^2_{\lambda w}) \tag{74}$$

$-1 < \lambda^w_t < 0$ for t=1,....,T

Ignoring any terms not involving λ^w, the prior density for λ^w is given by

$log\ \ p(\lambda^w|\sigma^2_{\lambda w}, \Omega_{\lambda w}) \propto -\frac{1}{2}(\lambda^w)' H' \Omega^{-1}_{\lambda w} H(\lambda^w) + g_{\lambda w}(\lambda^w, \sigma^2_{\lambda w})$

where,

$$g_{\lambda w}(\lambda^w, \sigma^2_{\lambda w}) = -\sum_{t=2}^{T} log\left(\Phi\left(\frac{0 - \lambda^w_{t-1}}{\sigma_{\lambda w}}\right) - \Phi\left(\frac{-1 - \lambda^w_{t-1}}{\sigma_{\lambda w}}\right)\right)$$

Combining the above two conditional densities we obtain,

$log\ \ p(\lambda^w|Y, \bullet) \propto -\frac{1}{2}(\lambda^w - \hat{\lambda}^w)' D^{-1}_{\lambda^w}(\lambda^w - \hat{\lambda}^w) + g_{\lambda w}(\lambda^w, \sigma^2_{\lambda w})$

where,

$D_{\lambda^w} = (H' \Omega^{-1}_{\lambda w} H + X'_u \Omega^{-1}_w X_u)^{-1}$

$\hat{\lambda}^w = D_{\lambda^w}(X'_u \Omega^{-1}_w B^w)$

The addition of the term $g_{\lambda w}(\lambda^w, \sigma^2_{\lambda w})$ leads to a non-standard density. Accordingly, we sample λ^w using an independence-chain Metropolis-Hastings (MH) procedure. This involves first generating candidate draws from $N(\hat{\lambda}^w, D_{\lambda^w})$ using the precision-based algorithm that are then accepted or rejected based on the accept-reject Metropolis-Hastings (ARMH) algorithm (discussed in Chan and Strachan, 2012).

Step 12. Derive the conditional distribution $p(\kappa^w|Y, \bullet)$

The information about κ^w comes from two sources. Below, we derive an expression for each of these two sources.

First, we define some notation,

$\tilde{w}_t = w_t - w^*_t$
$\tilde{u}_t = U_t - U^*_t$
$\tilde{\pi}_t = \pi_t - \pi^*_t$
$B^{\kappa w} = (\tilde{w}_1 - \rho^w_1 \tilde{w}_0 - \lambda^w_1 \tilde{u}_1, ..., \tilde{w}_T - \rho^w_T \tilde{w}_{T-1} - \lambda^w_{T-1} \tilde{u}_T)'$

The first source is the wage inflation measurement equation. Rewrite it in a matrix notation,

$$B^{\kappa w} = X_\pi \kappa^w + \varepsilon^w \qquad \varepsilon^w \sim N(0, \Omega_w) \tag{75}$$

where,

$X_\pi = diag(\tilde{\pi}_1, ..., \tilde{\pi}_T)$

Ignoring any terms not involving κ^w, we have the likelihood

$$log\ \ p(w|\kappa^w, \bullet) \propto -\tfrac{1}{2}(B^{\kappa w} - X_\pi \kappa^w)' \Omega_w^{-1} (B^{\kappa w} - X_\pi \kappa^w)$$

The second source comes from the state equation for κ^w. We rewrite it in a matrix notation as follows,

$$H\kappa^w = \varepsilon^{\kappa w} \qquad \varepsilon^{\kappa w} \sim N(0, \Omega_{\kappa w}), \quad where\ \ \Omega_{\kappa w} = diag(\omega_{\kappa w}^2, \sigma_{\kappa w}^2, ..., \sigma_{\kappa w}^2) \tag{76}$$

Ignoring any terms not involving κ^w, the prior density for κ^w is given by

$$log\ \ p(\kappa^w | \sigma_{\kappa w}^2, \Omega_{\kappa w}) \propto -\tfrac{1}{2}(\kappa^w)' H' \Omega_{\kappa w}^{-1} H(\kappa^w)$$

Combining the above two conditional densities we obtain,

$$log\ \ p(\kappa^w | Y, \bullet) \propto -\tfrac{1}{2}(\kappa^w - \hat{\kappa}^w)' D_{\kappa^w}^{-1} (\kappa^w - \hat{\kappa}^w)$$

where,

$D_{\kappa^w} = (H' \Omega_{\kappa w}^{-1} H + X_\pi' \Omega_w^{-1} X_\pi)^{-1}$

$\hat{\kappa}^w = D_{\kappa^w} (X_\pi' \Omega_w^{-1} B^{\kappa w})$

The candidate draws are sampled from $N(\hat{\kappa}^w, D_{\kappa^w})$ using the precision-based algorithm.

Step 13. Derive the conditional distribution $p(h^u, h^o, h^p, h^\pi, h^w, h^i | Y, \bullet)$

Given parameters and other latent states, $h^u, h^o, h^p, h^\pi, h^w, h^i$ are conditionally independent and so can be drawn separately. Following, Chan, Koop, and Potter (2013; 2016), we draw $h^u, h^o, h^p, h^\pi, h^w, h^i$ using the accept-reject independence-chain Metropolis Hastings (ARMH) algorithm of Chan and Strachan (2012; pages 32-34).

Step 14. Derive the conditional distribution $p(C^u, C^g, C^\pi, C^r, Wedge | Y, \bullet)$

Given parameters and other latent states, C^u, C^g, C^π, C^r are conditionally independent and so can be drawn separately.

Beginning with C^u, the information about it comes from two sources. Below, we derive an expression for each of these two sources.

The first source is the measurement equation linking survey expectations to U^*. Rewrite it in a matrix notation,

$$N^{zu} = C^u + \varepsilon^{zu} \qquad \varepsilon^{zu} \sim N(0, \Omega_{zu}) \tag{77}$$

where,

$n_t^{zu} = Z_t^u - \beta^u U^*$
$N^{zu} = (n_1^{zu}, n_2^{zu}, ..., n_T^{zu})'$
$\Omega_{zu} = diag(\sigma_{zu}^2, ..., \sigma_{zu}^2)$

Ignoring any terms not involving C^u, we have the likelihood

$$log\ \ p(Z^u|C^u, \bullet) \propto -\tfrac{1}{2}(N^{zu} - C^u)'\Omega_{zu}^{-1}(N^{zu} - C^u)$$

The second source comes from the state equation for C^u. We rewrite it in a matrix notation as follows,

$$HC^u = \alpha_{cu} + \varepsilon^{cu} \qquad \varepsilon^{cu} \sim N(0, \Omega_{cu}), \quad where\ \ \Omega_{cu} = diag(\omega_{cu}^2, \sigma_{cu}^2, ..., \sigma_{cu}^2) \tag{78}$$

where,

$$\alpha_{cu} = \begin{pmatrix} C_0^u \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Ignoring any terms not involving C^u, the prior density for C^u is given by

$$log\ \ p(C^u|\sigma_{cu}^2, \Omega_{cu}) \propto -\tfrac{1}{2}(C^u - H^{-1}\alpha_{cu})'H'\Omega_{cu}^{-1}H(C^u - H^{-1}\alpha_{cu})$$

Combining the above two conditional densities we obtain,

$$log\ \ p(C^u|Y, \bullet) \propto -\tfrac{1}{2}(C^u - \hat{C}^u)'D_{C^u}^{-1}(C^u - \hat{C}^u)$$

where,

$D_{C^u} = (H'\Omega_{cu}^{-1}H + \Omega_{zu}^{-1})^{-1}$

$\hat{C}^u = D_{C^u}(H'\Omega_{cu}^{-1}\alpha_{cu} + \Omega_{zu}^{-1}N^{zu})$

The candidate draws are sampled from $N(\hat{C}^u, D_{C^u})$ using the precision-based algorithm.

Following similar logic,

$N(\hat{C}^r, D_{C^r})$

$D_{C^r} = (H'\Omega_{cr}^{-1}H + \Omega_{zr}^{-1})^{-1}$

$\hat{C}^r = D_{C^r}(H'\Omega_{cr}^{-1}\alpha_{cr} + \Omega_{zr}^{-1}N^{zr})$

where,

$n_t^{zr} = Z_t^r - \beta^r r^*$
$N^{zr} = (n_1^{zr}, n_2^{zr}, ..., n_T^{zr})'$
$\Omega_{zr} = diag(\sigma_{zr}^2, ..., \sigma_{zr}^2)$

$N(\hat{C}^\pi, D_{C^\pi})$

$D_{C^\pi} = (H'\Omega_{c\pi}^{-1}H + \Omega_{z\pi}^{-1})^{-1}$

$\hat{C}^\pi = D_{C^\pi}(H'\Omega_{c\pi}^{-1}\alpha_{c\pi} + \Omega_{z\pi}^{-1}N^{z\pi})$

where,

$n_t^{z\pi} = Z_t^\pi - \beta^\pi \pi^*$
$N^{z\pi} = (n_1^{z\pi}, n_2^{z\pi}, ..., n_T^{z\pi})'$
$\Omega_{z\pi} = diag(\sigma_{z\pi}^2, ..., \sigma_{z\pi}^2)$

$N(\hat{C}^g, D_{C^g})$

$D_{C^g} = (H'\Omega_{cg}^{-1}H + \Omega_{zg}^{-1})^{-1}$

$\hat{C}^g = D_{C^g}(H'\Omega_{cg}^{-1}\alpha_{cg} + \Omega_{zg}^{-1}N^{zg})$

where,

$n_t^{zg} = Z_t^g + \beta^g \alpha_g - \beta^g gdp^*$
$N^{zg} = (n_1^{zg}, n_2^{zg}, ..., n_T^{zg})'$
$\Omega_{zg} = diag(\sigma_{zg}^2, ..., \sigma_{zg}^2)$
$\alpha_g = (gdp_0^*, 0, 0, ..., 0)'$

$N(\hat{Wedge}, D_{Wedge})$

$D_{Wedge} = (H'\Omega_{wlr}^{-1}H + \Omega_{w*}^{-1})^{-1}$

$\hat{Wedge} = D_{Wedge}(H'\Omega_{wlr}^{-1}\alpha_{wedge} + \Omega_{w*}^{-1}N^{wedge})$

where,

$n_t^{wedge} = W_t^* - P_t^* - \pi *_t$
$N^{wedge} = (n_1^{wedge}, n_2^{wedge}, ..., n_T^{wedge})'$
$\Omega_{w*} = diag(\sigma_{w*}^2, ..., \sigma_{w*}^2)$
$\Omega_{wlr} = diag(\omega_{wlr}^2, \sigma_{wlr}^2, ..., \sigma_{wlr}^2)$
$\alpha_{wedge} = (Wedge_0, 0, 0, ..., 0)'$

Step 15. Derive the conditional distribution $p(D|Y, \bullet)$

Given the posterior draws of r^*, ζ, and g^*, the posterior draw for D is constructed as,

$$D = r^* - \zeta g^* \tag{79}$$

Step 16. Derive the conditional distribution $p(\theta|Y, \bullet)$

There are 41 parameters in the vector θ. These parameters are drawn in 39 separate blocks using standard regression procedures. Following notation similar to that in Chan, Koop, and Potter (2016), we denote θ_{-x} to refer to all parameters in θ except the parameter x.

Substep 16.1 Derive the conditional distribution $p(\rho^u|Y, \bullet)$

Given the stationary constraints, $\rho_1^u + \rho_2^u < 1$, $\rho_2^u - \rho_1^u < 1$, and $|\rho_2^u| < 1$

$\rho^u = (\rho_1^u, \rho_2^u)'$ is a bivariate truncated normal. To obtain draws from this truncated normal distribution, an ARMH sampling algorithm is applied to the candidate draws from the proposal density, $N(\hat{\rho}^u, D_{\rho u})$.

$D_{\rho u} = (V_{\rho u}^{-1} + X_u' \Omega_u^{-1} X_u)^{-1}$

$\hat{\rho}^u = D_{\rho u}(V_{\rho u}^{-1} \rho_0^u + X_u' \Omega_{wlr}^{-1}(\tilde{u} - \phi^u ogap))$

where,

$V_{\rho u}^{-1}$ is the prior variance and ρ_0^u is the prior mean,

$$X_u = \begin{pmatrix} \tilde{u}_0 & \tilde{u}_{-1} \\ \tilde{u}_1 & \tilde{u}_0 \\ \vdots & \\ \tilde{u}_{T-1} & \tilde{u}_{T-2} \end{pmatrix}$$

Substep 16.2 Derive the conditional distribution $p(\sigma_{hu}^2|Y, \bullet)$

$p(\sigma^2_{hu}|Y,\bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma^2_{hu}|Y,\bullet) \sim IG(\nu_{hu0} + \tfrac{T-1}{2}, S_{hu0} + \tfrac{1}{2}\sum_{t=2}^{T}(h^u_t - h^u_{t-1})^2)$$

Substep 16.3 Derive the conditional distribution $p(\phi^u|Y,\bullet)$

Given the constraint $\phi^u < 0$, the conditional distribution $p(\phi^u|Y,\bullet)$ is a truncated normal density. The candidate draws are sampled from the proposal distribution $N(\hat{\phi}^u, D_{\phi u})$ using the precision-based algorithm, and a simple accept-reject step is applied to the candidate draws.

Rewrite the unemployment rate (gap) measurement equation in matrix notation as

$$Y^{\phi} = \phi^u ogap + \varepsilon^u \qquad \varepsilon^u \sim N(0, \Omega_u) \tag{80}$$

where,

$y^{\phi}_t = \tilde{u}_t - \rho^u_1 \tilde{u}_{t-1} - \rho^u_2 \tilde{u}_{t-2}$
$Y^{\phi} = (y^{\phi}_1, ..., y^{\phi}_T)'$

$D_{\phi u} = (V^{-1}_{\phi u} + ogap^{'}\Omega^{-1}_u ogap)^{-1}$

$\hat{\phi}^u = D_{\phi u}(V^{-1}_{\phi u}\phi^u_0 + ogap^{'}\Omega^{-1}_u Y^{\phi})$

where,

$V^{-1}_{\phi u}$ is the prior variance and ϕ^u_0 is the prior mean,

Substep 16.4 Derive the conditional distribution $p(\sigma^2_{u*}|Y,\bullet)$

$p(\sigma^2_{u*}|Y,\bullet)$ is a non-standard density because U^* is a bounded random walk,

$$log\ \ p(\sigma^2_{u*}|Y,\bullet) \propto -(\nu_{u*0}+1) log\ \ \sigma^2_{u*} - \tfrac{S_{u*0}}{\sigma^2_{u*}} - \tfrac{T-1}{2} log\ \ \sigma^2_{u*} - \tfrac{1}{2\sigma^2_{u*}}\textstyle\sum_{t=2}^{T}(U^*_t - U^*_{t-1})^2 + g_{u*}(U^*, \sigma^2_{u*})$$

The candidate draws from $p(\sigma^2_{u*}|Y,\bullet)$ are obtained via the MH step with the proposal density

$$IG(\nu_{u*0} + \tfrac{T-1}{2}, S_{u*0} + \tfrac{1}{2}\sum_{t=2}^{T}(U^*_t - U^*_{t-1})^2)$$

Substep 16.5 Derive the conditional distribution $p(\beta^u|Y,\bullet)$

Candidate draws are sampled from $N(\hat{\beta}^u, D_{\beta u})$ using the precision-based algorithm.

where,

$D_{\beta u} = (V_{\beta u}^{-1} + U^{*'}\Omega_{zu}^{-1}U^*)^{-1}$

$\hat{\beta}^u = D_{\beta u}(V_{\beta u}^{-1}\beta_0^u + U^{*'}\Omega_{zu}^{-1}J^{zu})$

$j_t^{zu} = Z_t^u - C_t^u$
$J^{zu} = (j_1^{zu}, ..., j_T^{zu})'$

$V_{\beta u}^{-1}$ is the prior variance and β_0^u is the prior mean for β^u

Substep 16.6 Derive the conditional distribution $p(\sigma_{zu}^2|Y, \bullet)$

$p(\sigma_{zu}^2|Y, \bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma_{zu}^2|Y, \bullet) \sim IG(\nu_{zu0} + \tfrac{T}{2}, S_{zu0} + \tfrac{1}{2}\sum_{t=1}^{T}(Z_t^u - C_t^u - \beta^u U^*)^2)$$

Substep 16.7 Derive the conditional distribution $p(\sigma_{cu}^2|Y, \bullet)$

$p(\sigma_{cu}^2|Y, \bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma_{cu}^2|Y, \bullet) \sim IG(\nu_{cu0} + \tfrac{T-1}{2}, S_{cu0} + \tfrac{1}{2}\sum_{t=2}^{T}(C_t^u - C_{t-1}^u)^2)$$

Substep 16.8 Derive the conditional distribution $p(\sigma_{gdp*}^2|Y, \bullet)$

$p(\sigma_{gdp*}^2|Y, \bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma_{gdp*}^2|Y, \bullet) \sim IG(\nu_{gdp*0} + \tfrac{T-1}{2}, S_{gdp*0} + (gdp^* - \alpha_{gdp*})' * H_2 H_2 * (gdp^* - \alpha_{gdp*})/2)$$

where (although they are defined above but for convenience we redefine them),

$$\alpha_{gdp*} = \begin{pmatrix} gdp_0^* + \triangle gdp_0^* \\ -gdp_0^* \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad H_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -2 & 1 & 0 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -2 & 1 \end{pmatrix}$$

H_2 is a band matrix with unit determinant and hence is invertible.

Substep 16.9 Derive the conditional distribution $p(\rho^g|Y,\bullet)$

Given the stationary constraints, $\rho_1^g + \rho_2^g < 1$, $\rho_2^g - \rho_1^g < 1$, and $|\rho_2^g| < 1$

$\rho^g = (\rho_1^g, \rho_2^g)'$ is a bivariate truncated normal. To obtain draws from this truncated normal distribution, an ARMH sampling algorithm is applied to the candidate draws from the proposal density, $N(\hat{\rho}^g, D_{\rho g})$.

$D_{\rho g} = (V_{\rho g}^{-1} + X_{\rho g}^{'}\Omega_{ogap}^{-1}X_{\rho g})^{-1}$

$\hat{\rho}^g = D_{\rho g}(V_{\rho g}^{-1}\rho_0^g + X_{\rho g}^{'}\Omega_{ogap}^{-1}Y_{ogap})$

where,

$V_{\rho g}^{-1}$ is the prior variance and ρ_0^g is the prior mean,

$$X_{\rho g} = \begin{pmatrix} 0 & 0 \\ ogap_1 & 0 \\ ogap_2 & ogap_1 \\ \vdots & \\ ogap_{T-1} & ogap_{T-2} \end{pmatrix}$$

$y_t^{ogap} = ogap_t - a^r(r_t - r_{t-1}) - \lambda^g \tilde{u}_t)$
$Y_{ogap} = (y_1^{ogap}, ..., y_T^{ogap})'$

Substep 16.10 Derive the conditional distribution $p(a^r|Y,\bullet)$

Candidate draws are sampled from $N(\hat{a}^r, D_{ar})$ using the precision-based algorithm.

where,

$D_{ar} = (V_{ar}^{-1} + X_{ar}^{'}\Omega_{ogap}^{-1}X_{ar})^{-1}$

$\hat{a}^r = D_{ar}(V_{ar}^{-1}a_0^r + X_{ar}^{'}\Omega_{ogap}^{-1}J^{ar})$

$j_t^{ar} = ogap_t - \rho_1^g ogap_{t-1} - \rho_2^g ogap_{t-2} - \lambda^g \tilde{u}_t$
$J^{ar} = (j_1^{ar}, ..., j_T^{ar})'$
$X_{ar} = (\tilde{r}_1, ..., \tilde{r}_T)'$
$\tilde{r}_t = r_t - r_t^*$

V_{ar}^{-1} is the prior variance and a_0^r is the prior mean for a^r

Substep 16.11 Derive the conditional distribution $p(\lambda^g|Y,\bullet)$

Given the constraint $\lambda^g < 0$, the conditional distribution $p(\lambda^g|Y, \bullet)$ is a truncated normal density. The candidate draws are sampled from the proposal distribution $N(\hat{\lambda}^g, D_{\lambda g})$ using the precision-based algorithm, and a simple accept-reject step is applied to the candidate draws.

where,

$D_{\lambda g} = (V_{\lambda g}^{-1} + X_u^{'}\Omega_{ogap}^{-1}X_u)^{-1}$

$\hat{\lambda}^g = D_{\lambda g}(V_{\lambda g}^{-1}\lambda_0^g + X_u^{'}\Omega_{ogap}^{-1}B^g)$

$b_t^g = ogap_t - \rho_1^g ogap_{t-1} - \rho_2^g ogap_{t-2} - a^r\tilde{r}_t$
$B^g = (b_1^g, ..., b_T^g)'$
$X_u = diag(\tilde{u}_1, ..., \tilde{u}_T)'$
$\tilde{r}_t = r_t - r_t^*$

$V_{\lambda g}^{-1}$ is the prior variance and λ_0^g is the prior mean for λ^g

Substep 16.12 Derive the conditional distribution $p(\sigma_{ho}^2|Y, \bullet)$

$p(\sigma_{ho}^2|Y, \bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma_{ho}^2|Y, \bullet) \sim IG(\nu_{ho0} + \tfrac{T-1}{2}, S_{ho0} + \tfrac{1}{2}\sum_{t=2}^{T}(h_t^o - h_{t-1}^o)^2)$$

Substep 16.13 Derive the conditional distribution $p(\sigma_{zg}^2|Y, \bullet)$

$p(\sigma_{zg}^2|Y, \bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma_{zg}^2|Y, \bullet) \sim IG(\nu_{zg0} + \tfrac{T}{2}, S_{zg0} + \tfrac{1}{2}\sum_{t=1}^{T}(Z_t^g - C_t^g - \beta^g gdp_{t-1}^* + \beta^g gdp_t^*)^2)$$

Substep 16.14 Derive the conditional distribution $p(\sigma_{cg}^2|Y, \bullet)$

$p(\sigma_{cg}^2|Y, \bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma_{cg}^2|Y, \bullet) \sim IG(\nu_{cg0} + \tfrac{T-1}{2}, S_{cg0} + \tfrac{1}{2}\sum_{t=2}^{T}(C_t^g - C_{t-1}^g)^2)$$

Substep 16.15 Derive the conditional distribution $p(\beta^g|Y, \bullet)$

Candidate draws are sampled from $N(\hat{\beta}^g, D_{\beta g})$ using the precision-based algorithm.

where,

$D_{\beta g} = (V_{\beta g}^{-1} + (Hgdp^* - \alpha_g)' \Omega_{zg}^{-1} (Hgdp^* - \alpha_g))^{-1}$

$\hat{\beta}^g = D_{\beta g}(V_{\beta g}^{-1} \beta_0^g + (Hgdp^* - \alpha_g)\Omega_{zg}^{-1} J^{zg})$

$j_t^{zg} = Z_t^g - C_t^g$
$J^{zg} = (j_1^{zg}, ..., j_T^{zg})'$
$\alpha_g = (gdp_0^*, 0, 0, ..., 0)'$

$V_{\beta g}^{-1}$ is the prior variance and β_0^g is the prior mean for β^g.

Substep 16.16 Derive the conditional distribution $p(\rho^p|Y, \bullet)$

Given the stationary constraint, $|\rho^p| < 1$

ρ^p is a truncated normal. To obtain draws from this truncated normal distribution, an AR sampling step is applied to the candidate draws from the proposal density, $N(\hat{\rho}^p, D_{\rho p})$.

$D_{\rho p} = (V_{\rho p}^{-1} + X_{prod}^{'} \Omega_P^{-1} X_{prod})^{-1}$

$\hat{\rho}^p = D_{\rho p}(V_{\rho p}^{-1} \rho_0^p + X_{prod}^{'} \Omega_P^{-1} Y^{prod})$

where,

$V_{\rho p}^{-1}$ is the prior variance and ρ_0^p is the prior mean,

$\tilde{p}_t = P_t - P_t^*$

$X_{prod} = (\tilde{p}_0, ..., \tilde{p}_{T-1})'$

$y_t^{prod} = \tilde{p}_t - \lambda_t^p \tilde{u}_t$

$Y^{prod} = (y_1^{prod}, ..., y_T^{prod})'$

Substep 16.17 Derive the conditional distribution $p(\sigma_{hp}^2|Y, \bullet)$

$p(\sigma_{hp}^2|Y, \bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma^2_{hp}|Y,\bullet) \sim IG(\nu_{hp0} + \tfrac{T-1}{2}, S_{hp0} + \tfrac{1}{2}\sum_{t=2}^{T}(h^p_t - h^p_{t-1})^2)$$

Substep 16.18 Derive the conditional distribution $p(\sigma^2_{p*}|Y,\bullet)$

$p(\sigma^2_{p*}|Y,\bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma^2_{p*}|Y,\bullet) \sim IG(\nu_{p*0} + \tfrac{T-1}{2}, S_{p*0} + \tfrac{1}{2}\sum_{t=2}^{T}(P^*_t - P^*_{t-1})^2)$$

Substep 16.19 Derive the conditional distribution $p(\sigma^2_{\lambda\pi}|Y,\bullet)$

$p(\sigma^2_{\lambda\pi}|Y,\bullet)$ is a non-standard density because of the constraints on λ^π,

$$log\ \ p(\sigma^2_{\lambda\pi}|Y,\bullet) \propto -(\nu_{\lambda\pi0}+1)log\ \ \sigma^2_{\lambda\pi} - \tfrac{S_{\lambda\pi0}}{\sigma^2_{\lambda\pi}} - \tfrac{T-1}{2} log\ \ \sigma^2_{\lambda\pi} - \tfrac{1}{2\sigma^2_{\lambda\pi}}\textstyle\sum_{t=2}^{T}(\lambda^\pi_t - \lambda^\pi_{t-1})^2 + g_{\lambda\pi}(\lambda^\pi, \sigma^2_{\lambda\pi})$$

The candidate draws from $p(\sigma^2_{\lambda\pi}|Y,\bullet)$ are obtained via the MH step with the proposal density

$$IG(\nu_{\lambda\pi0} + \tfrac{T-1}{2}, S_{\lambda\pi0} + \tfrac{1}{2}\sum_{t=2}^{T}(\lambda^\pi_t - \lambda^\pi_{t-1})^2)$$

Substep 16.20 Derive the conditional distribution $p(\sigma^2_{\rho\pi}|Y,\bullet)$

$p(\sigma^2_{\rho\pi}|Y,\bullet)$ is a non-standard density because of the constraints on ρ^π,

$$log\ \ p(\sigma^2_{\rho\pi}|Y,\bullet) \propto -(\nu_{\rho\pi0}+1)log\ \ \sigma^2_{\rho\pi} - \tfrac{S_{\rho\pi0}}{\sigma^2_{\rho\pi}} - \tfrac{T-1}{2} log\ \ \sigma^2_{\rho\pi} - \tfrac{1}{2\sigma^2_{\rho\pi}}\textstyle\sum_{t=2}^{T}(\rho^\pi_t - \rho^\pi_{t-1})^2 + g_{\rho\pi}(\rho^\pi, \sigma^2_{\rho\pi})$$

The candidate draws from $p(\sigma^2_{\rho\pi}|Y,\bullet)$ are obtained via the MH step with the proposal density

$$IG(\nu_{\rho\pi0} + \tfrac{T-1}{2}, S_{\rho\pi0} + \tfrac{1}{2}\sum_{t=2}^{T}(\rho^\pi_t - \rho^\pi_{t-1})^2)$$

Substep 16.21 Derive the conditional distribution $p(\sigma^2_{h\pi}|Y,\bullet)$

$p(\sigma^2_{h\pi}|Y,\bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma^2_{h\pi}|Y,\bullet) \sim IG(\nu_{h\pi0} + \tfrac{T-1}{2}, S_{h\pi0} + \tfrac{1}{2}\sum_{t=2}^{T}(h^\pi_t - h^\pi_{t-1})^2)$$

Substep 16.22 Derive the conditional distribution $p(\sigma^2_{\pi*}|Y,\bullet)$

$p(\sigma^2_{\pi*}|Y,\bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma^2_{\pi*}|Y,\bullet) \sim IG(\nu_{\pi*0} + \tfrac{T-1}{2}, S_{\pi*0} + \tfrac{1}{2}\sum_{t=2}^{T}(\pi^*_t - \pi^*_{t-1})^2)$$

Substep 16.23 Derive the conditional distribution $p(\sigma^2_{z\pi}|Y,\bullet)$

$p(\sigma^2_{z\pi}|Y,\bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma^2_{z\pi}|Y,\bullet) \sim IG(\nu_{z\pi0} + \tfrac{T}{2}, S_{z\pi0} + \tfrac{1}{2}\sum_{t=1}^{T}(Z^\pi_t - C^\pi_t - \beta^\pi\pi^*)^2)$$

Substep 16.24 Derive the conditional distribution $p(\sigma^2_{c\pi}|Y,\bullet)$

$p(\sigma^2_{c\pi}|Y,\bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma^2_{c\pi}|Y,\bullet) \sim IG(\nu_{c\pi0} + \tfrac{T-1}{2}, S_{c\pi0} + \tfrac{1}{2}\sum_{t=2}^{T}(C^\pi_t - C^\pi_{t-1})^2)$$

Substep 16.25 Derive the conditional distribution $p(\beta^\pi|Y,\bullet)$

Candidate draws are sampled from $N(\hat{\beta}^\pi, D_{\beta\pi})$ using the precision-based algorithm.

where,

$$D_{\beta\pi} = (V^{-1}_{\beta\pi} + \pi^{*\prime}\Omega^{-1}_{z\pi}\pi^*)^{-1}$$

$$\hat{\beta}^\pi = D_{\beta\pi}(V^{-1}_{\beta\pi}\beta^\pi_0 + \pi^{*\prime}\Omega^{-1}_{z\pi}J^{z\pi})$$

$$j^{z\pi}_t = Z^\pi_t - C^\pi_t$$
$$J^{z\pi} = (j^{z\pi}_1, ..., j^{z\pi}_T)'$$

$V^{-1}_{\beta\pi}$ is the prior variance and β^π_0 is the prior mean for β^π

Substep 16.26 Derive the conditional distribution $p(\sigma^2_{w*}|Y,\bullet)$

$p(\sigma^2_{w*}|Y,\bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma^2_{w*}|Y,\bullet) \sim IG(\nu_{w*0} + \tfrac{T-1}{2}, S_{w*0} + \tfrac{1}{2}\sum_{t=2}^{T}(w_t^* - \pi_t^* - P_t^*)^2)$$

Substep 16.27 Derive the conditional distribution $p(\sigma^2_{hw}|Y,\bullet)$

$p(\sigma^2_{hw}|Y,\bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma^2_{hw}|Y,\bullet) \sim IG(\nu_{hw0} + \tfrac{T-1}{2}, S_{hw0} + \tfrac{1}{2}\sum_{t=2}^{T}(h_t^w - h_{t-1}^w)^2)$$

Substep 16.28 Derive the conditional distribution $p(\sigma^2_{\rho w}|Y,\bullet)$

$p(\sigma^2_{\rho w}|Y,\bullet)$ is a non-standard density because of the constraints on ρ^w,

$$log\ \ p(\sigma^2_{\rho w}|Y,\bullet) \propto -(\nu_{\rho w0}+1)log\ \ \sigma^2_{\rho w} - \tfrac{S_{\rho w0}}{\sigma^2_{\rho w}} - \tfrac{T-1}{2} log\ \ \sigma^2_{\rho w} - \tfrac{1}{2\sigma^2_{\rho w}}\textstyle\sum_{t=2}^{T}(\rho_t^w - \rho_{t-1}^w)^2 + g_{\rho w}(\rho^w, \sigma^2_{\rho w})$$

The candidate draws from $p(\sigma^2_{\rho w}|Y,\bullet)$ are obtained via the MH step with the proposal density

$$IG(\nu_{\rho w0} + \tfrac{T-1}{2}, S_{\rho w0} + \tfrac{1}{2}\sum_{t=2}^{T}(\rho_t^w - \rho_{t-1}^w)^2)$$

Substep 16.29 Derive the conditional distribution $p(\sigma^2_{\lambda w}|Y,\bullet)$

$p(\sigma^2_{\lambda w}|Y,\bullet)$ is a non-standard density because of the constraints on λ^w,

$$log\ \ p(\sigma^2_{\lambda w}|Y,\bullet) \propto -(\nu_{\lambda w0}+1)log\ \ \sigma^2_{\lambda w} - \tfrac{S_{\lambda w0}}{\sigma^2_{\lambda w}} - \tfrac{T-1}{2} log\ \ \sigma^2_{\lambda w} - \tfrac{1}{2\sigma^2_{\lambda w}}\textstyle\sum_{t=2}^{T}(\lambda_t^w - \lambda_{t-1}^w)^2 + g_{\lambda w}(\lambda^w, \sigma^2_{\lambda w})$$

The candidate draws from $p(\sigma^2_{\lambda w}|Y,\bullet)$ are obtained via the MH step with the proposal density

$$IG(\nu_{\lambda w0} + \tfrac{T-1}{2}, S_{\lambda w0} + \tfrac{1}{2}\sum_{t=2}^{T}(\lambda_t^w - \lambda_{t-1}^w)^2)$$

Substep 16.30 Derive the conditional distribution $p(\sigma^2_{\kappa w}|Y,\bullet)$

The candidate draws are obtained from

$$IG(\nu_{\kappa w0} + \tfrac{T-1}{2}, S_{\kappa w0} + \tfrac{1}{2}\sum_{t=2}^{T}(\kappa_t^w - \kappa_{t-1}^w)^2)$$

Substep 16.31 Derive the conditional distribution $p(\rho^i|Y,\bullet)$

Given the constraint $|\rho^i| < 1$, the conditional distribution $p(\rho^i|Y,\bullet)$ is a truncated normal density. The candidate draws are sampled from the proposal distribution $N(\hat{\rho}^i, D_{\rho i})$ using the precision-based algorithm, and a simple accept-reject step is applied to the candidate draws.

where,

$$D_{\rho i} = (V_{\rho i}^{-1} + X_{\rho i}^{'}\Omega_i^{-1}X_{\rho i})^{-1}$$

$$\hat{\rho}^i = D_{\rho i}(V_{\rho i}^{-1}\rho_0^i + X_{\rho i}^{'}\Omega_i^{-1}M^{\rho i})$$

$m_t^{\rho i} = i_t - \pi_t^* - r_t^* - \lambda^i\tilde{u}_t - \kappa^i\tilde{\pi}_t$
$M^{\rho i} = (m_1^{\rho i}, ..., m_T^{\rho i})'$
$X_{\rho i} = (i_0 - \pi_0^* - r_0^*, ..., i_{T-1} - \pi_{T-1}^* - r_{T-1}^*)'$

$V_{\rho i}^{-1}$ is the prior variance and ρ_0^i is the prior mean for ρ^i

Substep 16.32 Derive the conditional distribution $p(\lambda^i|Y,\bullet)$

The candidate draws are sampled from the proposal distribution $N(\hat{\lambda}^i, D_{\lambda i})$ using the precision-based algorithm.

where,

$$D_{\lambda i} = (V_{\lambda i}^{-1} + X_{\lambda i}^{'}\Omega_i^{-1}X_{\lambda i})^{-1}$$

$$\hat{\lambda}^i = D_{\lambda i}(V_{\lambda i}^{-1}\lambda_0^i + X_{\lambda i}^{'}\Omega_i^{-1}M^{\lambda i})$$

$m_t^{\lambda i} = i_t - \pi_t^* - r_t^* - \rho^i(i_{t-1} - \pi_{t-1}^* - r_{t-1}^*) - \kappa^i\tilde{\pi}_t$
$M^{\lambda i} = (m_1^{\lambda i}, ..., m_T^{\lambda i})'$
$X_{\lambda i} = (\tilde{u}_1, ..., \tilde{u}_T)'$

$V_{\lambda i}^{-1}$ is the prior variance and λ_0^i is the prior mean for λ^i

Substep 16.33 Derive the conditional distribution $p(\kappa^i|Y,\bullet)$

The candidate draws are sampled from the proposal distribution $N(\hat{\kappa}^i, D_{\kappa i})$ using the precision-based algorithm.

where,

$D_{\kappa i} = (V_{\kappa i}^{-1} + X_{\kappa i}^{'}\Omega_i^{-1}X_{\kappa i})^{-1}$

$\hat{\kappa}^i = D_{\kappa i}(V_{\kappa i}^{-1}\kappa_0^i + X_{\kappa i}^{'}\Omega_i^{-1}M^{\kappa i})$

$m_t^{\kappa i} = i_t - \pi_t^* - r_t^* - \rho^i(i_{t-1} - \pi_{t-1}^* - r_{t-1}^*) - \lambda^i\tilde{u}_t$
$M^{\kappa i} = (m_1^{\kappa i}, ..., m_T^{\kappa i})'$
$X_{\kappa i} = (\tilde{\pi}_1, ..., \tilde{\pi}_T)'$

$V_{\kappa i}^{-1}$ is the prior variance and κ_0^i is the prior mean for κ^i

Substep 16.34 Derive the conditional distribution $p(\sigma_{hi}^2|Y, \bullet)$

$p(\sigma_{hi}^2|Y, \bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$p(\sigma_{hi}^2|Y, \bullet) \sim IG(\nu_{hi0} + \frac{T-1}{2}, S_{hi0} + \frac{1}{2}\sum_{t=2}^{T}(h_t^i - h_{t-1}^i)^2)$

Substep 16.35 Derive the conditional distribution $p(\sigma_{zr}^2|Y, \bullet)$

$p(\sigma_{zr}^2|Y, \bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$p(\sigma_{zr}^2|Y, \bullet) \sim IG(\nu_{zr0} + \frac{T}{2}, S_{zr0} + \frac{1}{2}\sum_{t=1}^{T}(Z_t^r - C_t^r - \beta^r r_t^*)^2)$

Substep 16.36 Derive the conditional distribution $p(\sigma_{cr}^2|Y, \bullet)$

$p(\sigma_{cr}^2|Y, \bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$p(\sigma_{cr}^2|Y, \bullet) \sim IG(\nu_{cr0} + \frac{T-1}{2}, S_{cr0} + \frac{1}{2}\sum_{t=2}^{T}(C_t^r - C_{t-1}^r)^2)$

Substep 16.37 Derive the conditional distribution $p(\beta^r|Y, \bullet)$

Candidate draws are sampled from $N(\hat{\beta}^r, D_{\beta r})$ using the precision-based algorithm.

where,

$D_{\beta r} = (V_{\beta r}^{-1} + r^{*'}\Omega_{zr}^{-1}r^*)^{-1}$

$$\hat{\beta}^r = D_{\beta r}(V_{\beta r}^{-1}\beta_0^r + r^{*\prime}\Omega_{zr}^{-1}J^{zr})$$

$j_t^{zr} = Z_t^r - C_t^r$
$J^{zr} = (j_1^{zr}, ..., j_T^{zr})'$

$V_{\beta r}^{-1}$ is the prior variance and β_0^r is the prior mean for β^r

Substep 16.38 Derive the conditional distribution $p(\sigma_d^2|Y, \bullet)$

$p(\sigma_d^2|Y, \bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma_d^2|Y, \bullet) \sim IG(\nu_{d0} + \tfrac{T-1}{2}, S_{d0} + \tfrac{1}{2}\sum_{t=2}^{T}(D_t - D_{t-1})^2)$$

Substep 16.39 Derive the conditional distribution $p(\sigma_{wlr}^2|Y, \bullet)$

$p(\sigma_{wlr}^2|Y, \bullet)$ is a standard inverse-Gamma density,

Candidate draws are sampled from

$$p(\sigma_{wlr}^2|Y, \bullet) \sim IG(\nu_{wlr0} + \tfrac{T-1}{2}, S_{wlr0} + \tfrac{1}{2}\sum_{t=2}^{T}(Wedge_t - Wedge_{t-1})^2)$$

A1.d Marginal likelihood computation

Bayesian model comparison is based on the marginal likelihood (marginal data density) metric. In computing marginal likelihood for various models, we use the approach proposed by CCK, which decomposes the marginal density of the data (e.g., inflation) into the product of predictive likelihoods. This approach allows us to separately compute marginal data density for each variable of interest: inflation, nominal wages, interest rate, real GDP, the unemployment rate, and labor productivity. The variable-specific marginal densities prove quite useful because they allow for deeper insights about the source of the deficiencies, which in turn helps differentiate models at a more disaggregated level.

Specifically, the marginal data density of the variables of interest is computed as follows,

$$p(y^j|X_i^j, M_i) = \prod_{t=3}^{T} p(y_t^j|y_{1:t-1}^j, X_{1:t,i}^j, M_i) \tag{81}$$

where, $j =$ PCE inflation (π), unemployment rate(ur), real GDP(gdp), labor productivity($prod$), nominal wage inflation($wage$), nominal short-term interest rate(int);
$p(y_t^j|y_{1:t-1}^j, X_{1:t,i}^j, M_i)$ is the predictive likelihood for variable j, and X_i^j is a set of covariates that influences variable j in model M_i. For example, in the case of the short-term interest rate, the covariates in the Base model include ur, π, gdp, and survey data, whereas in the Base-NoSurv model, the covariates will not include the survey data.

To compute the overall marginal data density of $Y = (y^{\pi}, y^{ur}, y^{gdp}, y^{prod}, y^{wage}, y^{int})'$ for model M_i,

$$\begin{aligned} p(Y|X_i, M_i) = p(y^{\pi}|X_i^{\pi}, M_i) \times p(y^{ur}|X_i^{ur}, M_i) \times p(y^{gdp}|X_i^{gdp}, M_i)... \\ \times p(y^{prod}|X_i^{prod}, M_i) \times p(y^{wage}|X_i^{wage}, M_i) \times p(y^{int}|X_i^{int}, M_i) \end{aligned} \tag{82}$$

A2. Prior Sensitivity Analysis

In the paper, we noted that our prior settings are similar to those used in CKP, CCK, and Gonzalez-Astudillo and Laforte (2020). As discussed in CCK, UC models with several unobserved variables, such as the one developed in this paper, require informative priors. That said, our priors settings for most variables are only slightly informative. The use of inequality restrictions on some parameters such as the Phillips curve, persistence, and bounds on u-star could be viewed as additional sources of information that eliminate the need for tight priors, something also noted by CKP. For the parameters for which there is a strong agreement in the empirical literature on their values, such as the Taylor-rule equation parameters, we use relatively tight priors, such that prior distributions are centered on prior means with small variance. So besides the prior on the Taylor-rule equation parameters, all other priors settings are taken from related papers.

Here, we examine the sensitivity of loosening the priors on the variances of the shocks to pi-star, p-star, u-star, and r-star (i.e., for the process D). Specifically, we double the prior mean of the variances from 0.01 to 0.03. Table A2 reports the posterior estimates. The top panel reports the posterior estimates from the main text to facilitate easy comparison, and panel (B)

reports the posterior estimates of re-running the Base model with these new prior values. It is worth noting that these new prior values are too loose to estimate Base-NoSurv feasibly; panel C reports results for prior values that are less loose. The results for p-star are as expected. In the paper, we noted that the prior view primarily shapes p-star, and we see that manifested here too; prior ($E(\sigma^2_{p*})$) and posterior ($E(\sigma^2_{p*}|Data)$) are fairly identical. Similar evidence is seen in the case of r-star (i.e., D) for the Base-NoSurv model. For pi-star, u-star, and r-star (in the case of Base), the differences in the posterior mean estimates' between the two panels are small. In fact, the posterior mean estimates from the runs with looser priors are pushed toward values that are closer to the prior mean estimates used in the main paper, lending credibility to our default priors settings used in the main paper.

Table A2: Parameter Estimates

Panel A: From the main paper, where prior $E(\sigma^2_{\pi^*}) = E(\sigma^2_{u^*}) = E(\sigma^2_d) = 0.1^2$ and $E(\sigma^2_{p^*}) = 0.14^2$

Parameter	Parameter description	Posterior estimates					
		Base			Base-NoSurv		
		Mean	5%	95%	Mean	5%	95%
$\sigma^2_{\pi^*}$	Variance of the shocks to π^*	0.119^2	0.100^2	0.140^2	0.118^2	0.082^2	0.162^2
$\sigma^2_{p^*}$	Variance of the shocks to p^*	0.142^2	0.110^2	0.180^2	0.140^2	0.108^2	0.178^2
$\sigma^2_{u^*}$	Variance of the shocks to u^*	0.093^2	0.079^2	0.106^2	0.121^2	0.103^2	0.139^2
σ^2_d	Variance of the shocks to d	0.094^2	0.078^2	0.112^2	0.100^2	0.076^2	0.128^2

Panel B: Prior sensitivity, where prior $E(\sigma^2_{\pi^*}) = E(\sigma^2_{u^*}) = E(\sigma^2_d) = E(\sigma^2_{p^*}) = 0.173^2$

Parameter	Parameter description	Posterior estimates					
		Base			Base-NoSurv		
		Mean	5%	95%	Mean	5%	95%
$\sigma^2_{\pi^*}$	Variance of the shocks to π^*	0.143^2	0.124^2	0.163^2	—	—	—
$\sigma^2_{p^*}$	Variance of the shocks to p^*	0.172^2	0.134^2	0.214^2	—	—	—
$\sigma^2_{u^*}$	Variance of the shocks to u^*	0.102^2	0.090^2	0.115^2	—	—	—
σ^2_d	Variance of the shocks to d	0.122^2	0.106^2	0.140^2	—	—	—

Panel C: Prior sensitivity, where prior $E(\sigma^2_{\pi^*}) = E(\sigma^2_{u^*}) = E(\sigma^2_d) = E(\sigma^2_{p^*}) = 0.14^2$

Parameter	Parameter description	Posterior estimates					
		Base			Base-NoSurv		
		Mean	5%	95%	Mean	5%	95%
$\sigma^2_{\pi^*}$	Variance of the shocks to π^*	0.132^2	0.114^2	0.150^2	0.152^2	0.116^2	0.196^2
$\sigma^2_{p^*}$	Variance of the shocks to p^*	0.144^2	0.112^2	0.181^2	0.138^2	0.107^2	0.173^2
$\sigma^2_{u^*}$	Variance of the shocks to u^*	0.105^2	0.092^2	0.117^2	0.133^2	0.114^2	0.152^2
σ^2_d	Variance of the shocks to d	0.112^2	0.095^2	0.129^2	0.139^2	0.107^2	0.175^2

A3. MCMC Convergence Diagnostics

In this section, we document the diagnostic properties of our MCMC algorithm in the Base model. Following Primiceri (2005), Koop, Leon-Gonzalez, and Strachan (2010), and Korobilis (2017), we report the autocorrelation functions of the posterior draws (10th and 50th order sample autocorrelation), inefficiency factors (IFs), and convergence diagnostics (CD) of Geweke (1992).[1]

One of the most common metrics examined to assess the efficiency of the MCMC sampler is to look at the autocorrelation function of the draws, which indicates how well the chain is mixing. Low autocorrelations are preferred to high because the lower the autocorrelation, the closer the draws are to being independent and the higher the efficiency of the algorithm. The plots shown in the top panel of Figure A1 correspond to 10th and 50th order autocorrelations in the draws, and as can be seen, they indicate very low autocorrelation. In the case of 50th order autocorrelation, except for a couple of them, most indicate correlation close to zero, and in the case of the 10th order except for a few, most indicate correlation below 0.2.

The inefficiency factor related to the autocorrelation functions is the inverse of Geweke's (1992) relative numerical efficiency measure (RNE). It is computed using the following formula, $(1 + 2\sum_{i=1}^{\infty} \rho_i)$, where ρ_i refers to the $k - th$ order autocorrelation of the chain. The middle panel in Figure A1 plots the IF for the model parameters. The values lower than or close to 20 are considered desirable. As shown, in the case of the Base model, most of the IFs are below 20. (Note: IFs are computed using the default setting in LeSage's toolbox: estimation of spectral density at frequency zero uses a tapered window of 4%.)

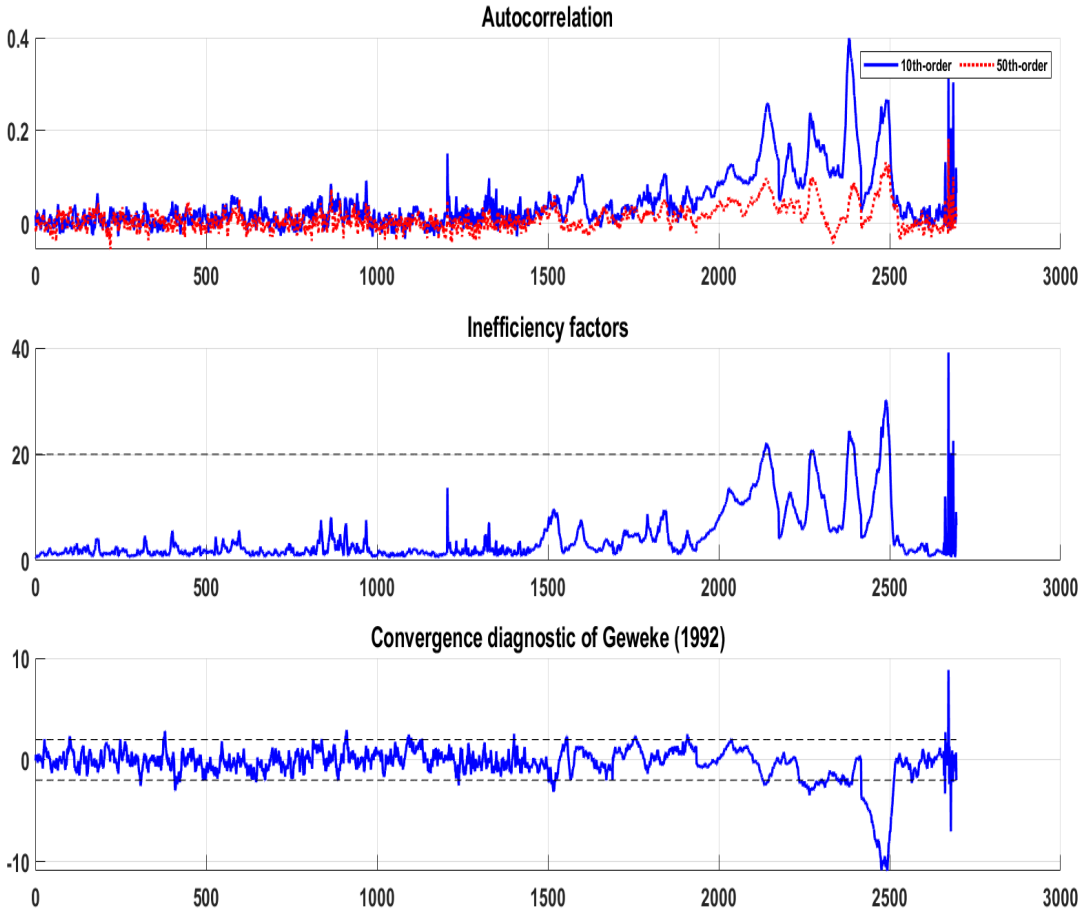
As discussed in Koop, Leon-Gonzalez, and Strachan (2010), to assess whether the MCMC sampler has converged, a rough rule of thumb is to look at the CDs and see whether 95% of them are less than 2. If they are, then convergence is likely achieved. Based on the plots in Figure A1 (third panel), most CDs are within +/- 2.

We also note that the results are fairly identical to the different initial conditions of the chain (picked randomly) and to a significantly lower number of simulations (and burn-in). For example, a run using 320K posterior draws with a burn-in of the first 300K and retaining all the remaining draws provides similar inference. However, the MCMC convergence properties favor higher simulations because it allows for a greater degree of thinning.

Overall, these diagnostic measures give us confidence in the good convergence properties of our MCMC algorithm in our Base model.

[1] In computing some of these metrics, we have benefitted from the Matlab toolbox developed by James P. LeSage. A detailed explanation, including intuition for these convergence diagnostics, is provided in Koop (2003; pages 67-68) and Chan et al. (2019; page 209).

Figure A1: MCMC Diagnostics of Base Model



A4. Forecasting Results I: Base vs. Base-NoSurv and Base vs. Benchmarks

Base vs. Base-NoSurv

Table A3 presents the results from comparing the out-of-sample forecasting performance (both point and density) between the two models over the forecast evaluation sample spanning 1999Q1 through 2019Q4. The forecast evaluation is based on real-time data vintages and uses a recursively expanding estimation window, where each recursive run uses an additional quarterly data point in the estimation sample.[2] Accordingly, we devise a systematic approach to adjusting the prior on the scale parameters of the inverse gamma distributions defining the variances of the stars. We multiply the scale parameter with the $factor = (\frac{2T}{N} - 1) * (\frac{T}{N+5(N-T)})$, where N is the total sample size from 1959Q4 through 2019Q4, and T refers to the number of data points in a given data vintage. At the end of the sample, the $factor = 1$ because $T = N$. The forecast accuracy (point and density) is computed from one quarter ahead to 20 quarters out. Partly due to our focus on the medium-term horizon and partly in the interest of space, we report accuracy metrics for 4, 8, 12, 16, and 20 quarters ahead. We evaluate the forecast accuracy using real-time data; specifically, we treat the "actual" as the *third* release of a given quarterly estimate.[3] For instance, in the case of real GDP, the third estimate for 2018Q4 corresponds to the GDP data available in late 2019Q1. The point forecast accuracy is assessed using the root mean squared error (RMSE) metric, and the density forecast accuracy is evaluated using the log predictive score (LPS). The statistical significance of the point forecast accuracy is gauged using the Diebold-Mariano and West test and, in the case of the density forecast, accuracy is based on the likelihood-ratio test of Amisano and Giacomini (2007).

The top panel of the table reports the results corresponding to the point forecast accuracy, while the lower panel reports results for the density forecast accuracy. The numbers reported in the table correspond to relative RMSE –RMSE Base relative to RMSE of Base-NoSurv – in the point forecast comparison, and relative mean LPS – LPS Base minus LPS Base-NoSurv – in the case of the density forecast comparison. Hence, numbers less than one in the top panel suggest that the point forecast accuracy of the Base forecast is more accurate on average, and positive numbers in the bottom panel suggest that the density forecast accuracy of the Base forecast is more accurate than that of the Base-NoSurv forecast.

As is evident by the numbers reported in the table, except for the point forecast accuracy of the shadow federal funds rate, the evidence generally favors the Base model as more accurate than Base-NoSurv. The evidence in support of the Base model is strongest for PCE inflation and nominal wage inflation. In the case of the unemployment rate and real GDP growth, both models perform comparably. In the case of the shadow federal funds rate, the Base-NoSurv model outperforms the Base model in point forecast accuracy. However, this improved point forecast accuracy of the fed funds rate does not translate into improved density forecast accuracy. In the case of real GDP growth, even though the point forecast accuracy between the two models is similar on average, the density forecasts from the Base-NoSurv model are slightly more accurate than those from the Base model for forecast horizons up to h=8Q.

[2]Going back in time means that we are using relatively fewer observations to estimate model(s). As is commonly done when performing real-time forecasting using multivariate UC models, we need to tighten priors on the shocks' variances driving the latent components (see, for instance, Barbarino et. al., 2020).

[3]Results are qualitatively similar if we instead use the revised data (2020Q1 vintage data) as the actual values in the forecast evaluation exercises. The results are available from the author on request.

Table A3: Real-Time Forecasting Accuracy: Base vs. Base-NoSurv

Panel A: Point Forecast Accuracy (Recursive evaluation: 1999.Q1-2019.Q4)

Relative RMSE: RMSE **Base** / RMSE **BaseNoSurv**

	h=4Q	h=8Q	h=12Q	h=16Q	h=20Q
Real GDP	1.02	0.99	1.00	0.99	1.00
PCE Inflation	0.96	0.93*	0.90*	0.93*	0.97
Productivity	0.99	1.01	0.98*	0.99	1.00
Nominal Wage (AHE)	0.98	0.90*	0.93*	0.94	0.98
Unemployment Rate	1.02	1.01	0.98	0.97	0.97
Shadow FFR	1.05	1.11*	1.14*	1.16*	1.21*

Panel B: Density Forecast Accuracy (Recursive evaluation: 1999.Q1-2019.Q4)

Relative Log Predictive Score (LPS): LPS Base - LPS BaseNoSurv

	h=4Q	h=8Q	h=12Q	h=16Q	h=20Q
Real GDP	-0.004*	-0.002*	-0.001	0.000	0.000
PCE Inflation	0.014*	0.013*	0.016*	0.016*	0.015*
Productivity	0.001	0.000	0.000	0.001	0.002
Nominal Wage (AHE)	0.003	0.011*	0.008*	0.004	0.001
Unemployment Rate	0.000	0.012	0.004	0.003	0.007*
Shadow FFR	0.021*	0.004	0.000	-0.002	0.000

Notes: The top panel compares the point forecast accuracy of the Base model with the Base-NoSurv model. Numbers less than 1 indicate that the Base model is more accurate on average. The bottom panel reports the corresponding density forecast accuracy performance. A positive value (for the relative mean predictive log score) suggests that the Base model is on average more accurate. The log predictive scores are computed using parametric normal approximation. The table reports statistical significance based on the likelihood-ratio test of Amisano and Giacomini (2007) for the density forecast accuracy, and based on the Diebold-Mariano and West test (with the lag $h-1$ truncation parameter of the HAC variance estimator) for the point forecast accuracy. The test statistics for the likelihood-ratio test use a two-sided t-test. In the case of the Diebold-Mariano and West test, the test statistics use two-sided standard normal critical values for horizons less than or equal to 8 quarters, and two-sided t-statistics for horizons greater than 8 quarters. *up to 10% significance level.

Base vs. Benchmarks

In this section we compare the real-time forecasting performance of our Base model to the outside benchmark models, which the forecasting literature has shown to be useful forecasting devices. Specifically, we compare the accuracy of the inflation forecasts from our Base model to the following three models: UCSV of Stock and Watson (2007) [UCSV], Chan, Koop, and Potter (2016) [CKP], and Chan, Clark, and Koop (2018) [CCK]. We compare the accuracy of the unemployment rate forecasts from our Base model to the CKP model, and the accuracy of the nominal wage inflation from the Base model to the UCSV model applied to the nominal wage inflation, motivated by Knotek (2015).

Table A4 presents the forecast evaluation results for headline PCE inflation, nominal wage inflation, and the unemployment rate. These results indicate the following three observations. First, in terms of point forecast accuracy, inflation forecasts from all four models considered are competitive with each other. There is some statistically significant evidence that the Base

model is more accurate than UCSV at h=12Q. Regarding the density forecast accuracy, the Base model is more accurate than the UCSV but inferior to CCK, as the latter produces more precise intervals than the Base model. Second, in the case of nominal wage inflation, the Base model generates more accurate forecasts (both point and density) than UCSV, and the gains are statistically significant for the most part.

Third, the accuracy of the unemployment forecasts from the Base model is competitive with the CKP model statistically speaking, even though the relative numbers favor CKP. A closer inspection of the forecast errors reveals that the Base model, which incorporates survey forecasts of the unemployment rate, experienced significantly bigger misses than the CKP model around the Great Recession period. Outside of this period, the Base model is slightly more accurate than the CKP, and when combined with the Great Recession period, on net, the much bigger misses of the Base model result in overall slightly higher RMSE.

As illustrated in Tallman and Zaman (2020), just before and at the onset of the Great Recession, survey participants projected relatively upbeat long-run forecasts of unemployment, which indicated a declining natural rate of unemployment. It was not until a few months into the recession that survey participants recognized the extent of the labor market damage and began to revise their estimates of the long-run unemployment rate higher. Hence, models such as the Base model that take signals from the survey forecasts experienced big misses.

To sum up, we view these forecasting results as providing evidence supporting our Base model's competitive forecasting properties.

Table A4: Out-of-Sample Forecasting Performance: **Base vs. Benchmarks**

Full Sample (Recursive evaluation: 1999.Q1-2019.Q4)

	Point forecasting					Density forecasting			
	4Q	8Q	12Q	20Q		4Q	8Q	12Q	20Q
PCE Inflation									
Relative RMSE					Relative Log Score				
Base/UCSV	0.96	0.96	0.93*	0.95	Base - UCSV	0.012*	0.022*	0.027*	0.025*
Base/CCK	1.02	1.04*	1.02	1.02	Base - CCK	-0.018*	-0.032*	-0.048*	-0.076*
Base/CKP	1.00	0.98	0.98	1.00	Base - CKP	0.001	-0.001	-0.004*	-0.024*
Nominal Wage									
Relative RMSE					Relative Log Score				
Base/UCSV	0.88*	0.78*	0.80*	0.50	Base - UCSV	0.017*	0.035*	0.031*	0.010
Unemployment Rate									
Relative MSE					Relative Log Score				
Base/CKP	1.01	1.03	1.05	1.06	Base - CKP	0.115*	0.025	-0.019	-0.059*

Notes: For variables PCE inflation and nominal wage (i.e., average hourly earnings), the forecasts and associated accuracy correspond to the quarterly annualized rate. Base forecast is defined as the Steady-State (SS) VAR forecast in which the steady states are assumed to be the estimates of the stars from the Base model. UCSV forecast corresponds to the forecast from the univariate unobserved component stochastic volatility model similar to Stock and Watson (2007). The model is used to construct forecasts of PCE inflation and nominal wage inflation. CCK forecast corresponds to the forecast from the bivariate unobserved component stochastic volatility model of Chan, Clark and Koop (2018). CKP forecast corresponds to the forecast from the bivariate unobserved component stochastic volatility model of Chan, Koop and Potter (2016), with the bounds on u-star fixed to values identical to the Base model. The left panel reports results for the point forecast accuracy (relative root mean squared errors) and the right panel reports the corresponding density forecast accuracy (mean of the relative log predictive score). The table reports statistical significance based on the Diebold-Mariano and West test with the lag $h-1$ truncation parameter of the HAC variance estimator and adjusts the test statistic for the finite sample correction proposed by Harvey, Leybourne, and Newbold (1997); *up to 10% significance level. The test statistics use two-sided standard normal critical values for horizons less than or equal to 8 quarters, and two-sided t-statistics for horizons greater than 8 quarters.

A5. Forecasting Results II: SSBVAR, Base stars vs. Survey

In macroeconomic forecasting, research by Wright (2013) and Tallman and Zaman (2020), among others, using workhorse Bayesian VAR models shows that the predictive performance boils down to good starting conditions (i.e., nowcasts) and terminal conditions (i.e., steady states proxied by stars). Survey forecasts provide both nowcasts and long-run projections, whose accuracy has been shown by past research to be quite good. Wright (2019) emphasizes the desirable forecasting properties of the survey forecasts and highlights that econometric approaches utilizing survey projections are at the forecasting frontier, especially in inflation forecasting. Most empirical research on forecasting has focused on proposing methods to improve the accuracy of the nowcast estimates relative to survey nowcasts' accuracy, but only little effort has been dedicated to improving estimates of long-run projections. Hence, this paper raises a natural curiosity about the usefulness of the stars' estimates from our modeling framework for macroeconomic forecasting using Bayesian VARs (via the imposition of steady states).

To assess the efficacy of our stars' estimates for the external VAR models, we perform a real-time out-of-sample forecasting evaluation similar to Wright (2013) and Tallman and Zaman (2020). These studies informed the time-varying steady states for the steady-state (SS) BVAR using long-run survey projections and found that doing so leads to significant gains in accuracy. Accordingly, the design of our forecasting examination is as follows. We take the SSBVAR from Tallman and Zaman (2020) and perform three sets of recursive real-time out-of-sample forecasting runs. In the first run, we inform the steady states for real GDP growth, PCE inflation, core PCE inflation, the unemployment rate, nominal wage inflation, and labor productivity growth using long-run survey projections. For the latter two variables, we use the survey expectations from the SPF.[4] The forecasts from this run are denoted 'Survey' in Table A5. In the second run, we repeat the exercise, but this time inform the steady states using the real-time estimates of the stars from the Base-NoSurv model, denoted 'BaseNoSurv.' In the third run, we inform the steady states using the real-time estimates of stars from the Base model, denoted 'Base.'

Each of the three forecasting runs is based on estimating the SSBVAR with a recursively expanding sample, i.e., the recursive execution uses an additional quarterly data point in the estimation. The SSBVAR is estimated with quarterly data beginning 1959Q2. The model consists of ten variables: (1) real GDP growth; (2) real consumption expenditures; (3) headline PCE inflation; (4) core PCE inflation; (5) labor productivity growth; (6) growth in average hourly earnings; (7) growth in payroll employment; (8) the unemployment rate; (9) the shadow federal funds rate; and (10) the risk spread, defined as the difference between the yield on the 10-year Treasury bond and the yield on a BAA-rated bond. The out-of-sample forecasting period spans 1999Q1 through 2019Q4. The forecast accuracy (point and density) is computed from one quarter ahead to 20 quarters out. Partly due to our focus on the medium-term horizon and partly in the interest of space, we report accuracy metrics for 4, 8, 12, and 20 quarters ahead.

We evaluate the forecast accuracy using real-time data; specifically, we treat the "actual" as the third quarterly estimate. For instance, in the case of real GDP, the third estimate for 2018Q4 corresponds to the GDP data available in late 2019Q1. The point forecast accuracy is assessed using the root mean squared error (RMSE) metric, and the density forecast accuracy is assessed using the continuous ranked probability score (CRPS). Forecasts with lower RMSE

[4] In the case of nominal wage inflation, we construct an implied survey projection by adding the survey expectations of PCE inflation and productivity, both of which are obtained from the SPF.

and CRPS are preferred. The statistical significance of the point and density forecast accuracy is gauged using the Diebold-Mariano and West test. The description of these tests is listed in the notes accompanying the tables reporting forecast accuracy.

Table A5 reports forecast evaluation results corresponding to this exercise. The left panel reports the point forecast accuracy results, while the right panel reports results for the density forecast accuracy. We evaluate and compare the point and density forecast accuracies among the Base, BaseNoSurv, and Survey forecasts in a pairwise fashion. For each variable, the three rows report the relative RMSE (for point forecast accuracy) and the relative CRPS (for density forecast accuracy). The first row reports the RMSE of the Base relative to Survey, the second row reports the RMSE of BaseNoSurv relative to Survey, and the third row reports the RMSE of BaseNoSurv relative to Base. A model with lower values of RMSE and CRPS is preferred to a model with higher values. These relative metrics indicate the competitive accuracy of the stars' estimates from the Base model compared to Survey and Base-NoSurv. The exception is the shadow federal funds rate, where Base-NoSurv outperforms both Base and Survey. For real GDP growth, statistically speaking, Base outperforms Base-NoSurv and is competitive with Survey.

For headline PCE inflation and labor productivity, all three are competitive with each other. In the case of nominal wage inflation, both Base and Base-NoSurv generate forecasts that are substantially more accurate than Survey on average. The gains are statistically significant for the most part. The SSBVAR with steady states informed by the Base model generates more accurate unemployment forecasts than Base-NoSurv and Survey, but the forecast gains are statistically significant only when compared to Survey.

Overall, these forecasting results lend credibility to our stars' estimates in their use to inform steady states for VAR forecasting models. We also note that the results in this section lend support to the survey projections in their use as proxies for stars, something also documented by Tallman and Zaman (2020), among others.

We believe the fact that the estimates of stars from our models are generally competitive with survey long-run projections is a good outcome. It has been well-established that survey expectations are at the frontier of forecasting (e.g., Wright, 2019). However, the preference is for forecasts (or estimates of stars) obtained using a single multivariate model because the resulting forecasts will be coherent and allow for a credible narrative in a systematic manner.

Table A5: Out-of-Sample Forecasting Performance: **Steady-State BVAR**

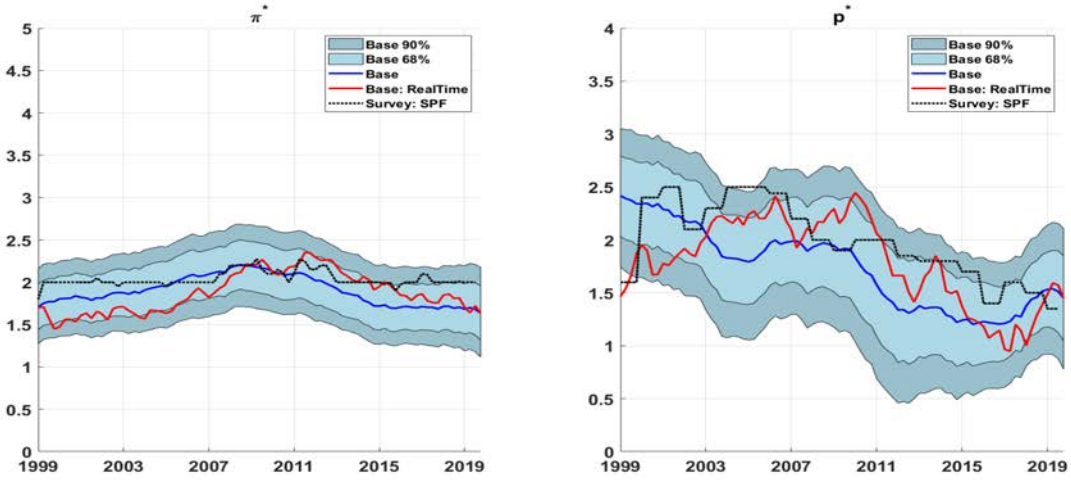
Full Sample (Recursive evaluation: 1999.Q1-2019.Q4)

	Point forecasting					Density forecasting			
	4Q	8Q	12Q	20Q		4Q	8Q	12Q	20Q
Real GDP									
Relative RMSE					Relative CRPS				
Base/Survey	0.96	1.00	1.03	1.01	Base - Survey	-0.05	0.02	0.04	0.02
BaseNoSurv/Survey	1.01	1.04	1.04	1.02	BaseNoSurv - Survey	0.02	0.06	0.05	0.02
BaseNoSurv/Base	1.05*	1.03*	1.01	1.01	BaseNoSurv - Base	0.07*	0.04	0.01	0.01
PCE Inflation									
Relative RMSE					Relative CRPS				
Base/Survey	0.98	1.00	1.02	1.01	Base - Survey	-0.02*	0.01	0.03	0.02
BaseNoSurv/Survey	0.99	1.02	1.09	1.05	BaseNoSurv - Survey	-0.01	0.02	0.08	0.06
BaseNoSurv/Base	1.00	1.01	1.07*	1.03	BaseNoSurv - Base	0.00	0.01	0.06*	0.04
Productivity									
Relative RMSE					Relative CRPS				
Base/Survey	1.03	1.04	1.04	1.01*	Base - Survey	0.02	0.04	0.04	0.01*
BaseNoSurv/Survey	1.04	1.04	1.04	1.01	BaseNoSurv - Survey	0.03*	0.03	0.05	0.01
BaseNoSurv/Base	1.01	0.99	1.00	1.00	BaseNoSurv - Base	0.01	-0.01	0.01	0.00
Nominal Wage									
Relative RMSE					Relative CRPS				
Base/Survey	0.71*	0.68*	0.67*	0.71*	Base - Survey	-0.09*	-0.13*	-0.18*	-0.26*
BaseNoSurv/Survey	0.68*	0.67*	0.74*	0.80*	BaseNoSurv - Survey	-0.09*	-0.12*	-0.14*	-0.19*
BaseNoSurv/Base	0.95	0.98	1.11*	1.12*	BaseNoSurv - Base	0.00	0.00	0.04*	0.07*
Unemployment Rate									
Relative RMSE					Relative CRPS				
Base/Survey	0.93*	0.91*	0.90	0.92	Base - Survey	-0.04	-0.12	-0.16	-0.12
BaseNoSurv/Survey	0.95*	0.97	0.99	0.99	BaseNoSurv - Survey	-0.05	-0.07	-0.06	0.00
BaseNoSurv/Base	1.03	1.07	1.09	1.08	BaseNoSurv - Base	0.00	0.05	0.10	0.12
Shadow FFR									
Relative RMSE					Relative CRPS				
Base/Survey	0.99	0.97	0.96	0.95*	Base - Survey	-0.01	-0.07	-0.13	0.16*
BaseNoSurv/Survey	0.93*	0.91*	0.93	0.95*	BaseNoSurv - Survey	-0.07*	-0.17*	-0.19	-0.14*
BaseNoSurv/Base	0.93*	0.94*	0.97	1.00	BaseNoSurv - Base	-0.06*	-0.10*	-0.07	0.03*

Notes: For the variables real GDP, PCE inflation, productivity, nominal wage (i.e., average hourly earnings), the forecasts and the associated accuracy correspond to the quarterly annualized rate. Base forecast is defined as the Steady-State (SS) VAR forecast in which the steady states are assumed to be the estimates of the stars from the Base model. BaseNoSurv forecast is defined as the SS-VAR forecast in which the steady states are taken from the Base-NoSurv model. The left panel reports results for the point forecast accuracy (relative root mean squared errors) and the right panel reports the corresponding density forecast accuracy (mean of the relative continuous ranked probability score). The table reports statistical significance based on the Diebold-Mariano and West test with the lag $h-1$ truncation parameter of the HAC variance estimator and adjusts the test statistic for the finite sample correction proposed by Harvey, Leybourne, and Newbold (1997); *up to 10% significance level. The test statistics use two-sided standard normal critical values for horizons less than or equal to 8 quarters, and two-sided t-statistics for horizons greater than 8 quarters.

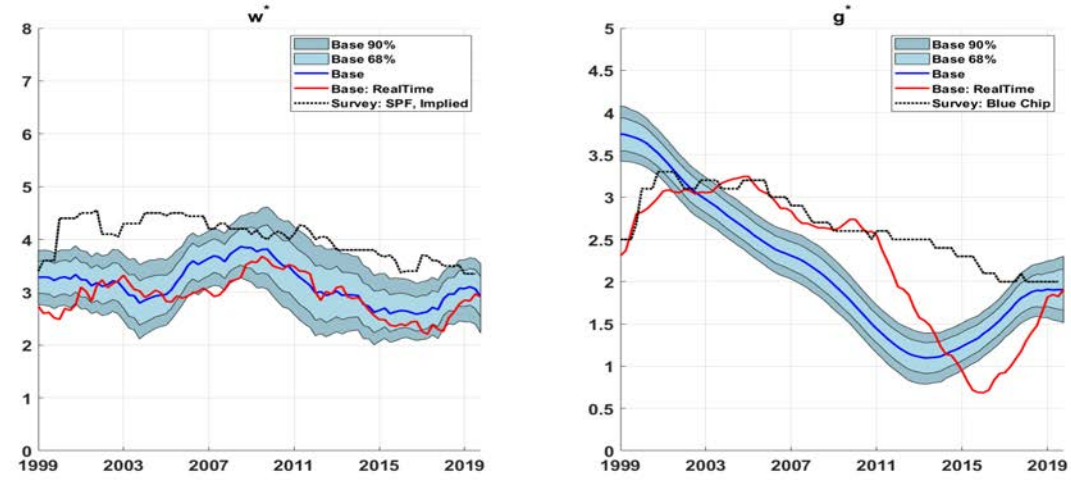
A6. Additional Real-time Estimates of Stars

Figure A2: Real-time Recursive Estimates of pi-star and p-star: Base model



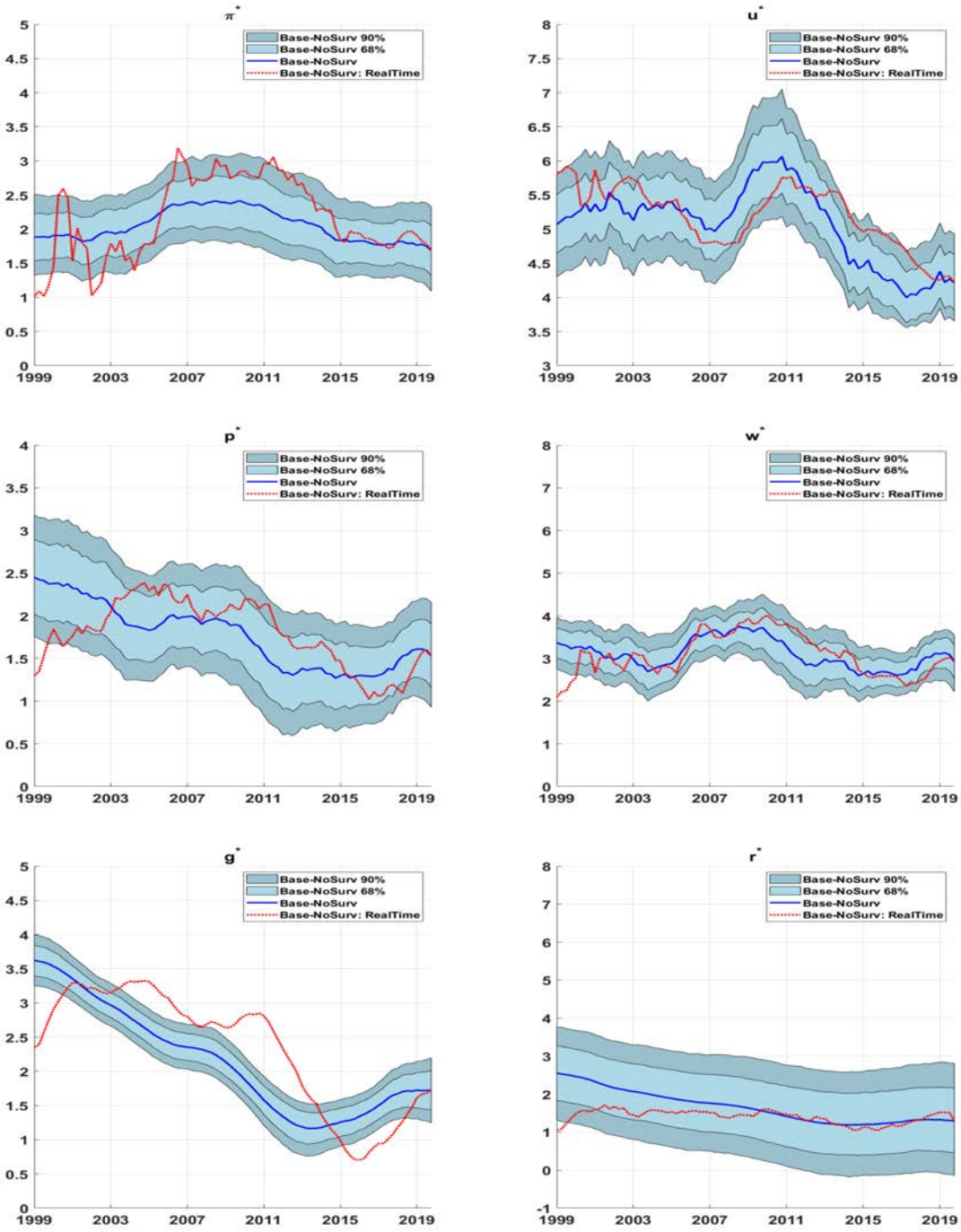
Notes: The plot denoted Base corresponds to smoothed (posterior mean) estimates based on the full sample information, i.e., 1959.Q4 through 2019.Q4. The plot denoted Base: RealTime corresponds to real-time recursive (posterior mean) estimates generated by estimating the Base model at different points in time, specifically 1999.Q1 through 2019.Q4. The credible intervals reflect the uncertainty around the posterior mean smoothed estimates.

Figure A3: Real-time Recursive Estimates of w-star and g-star: Base model



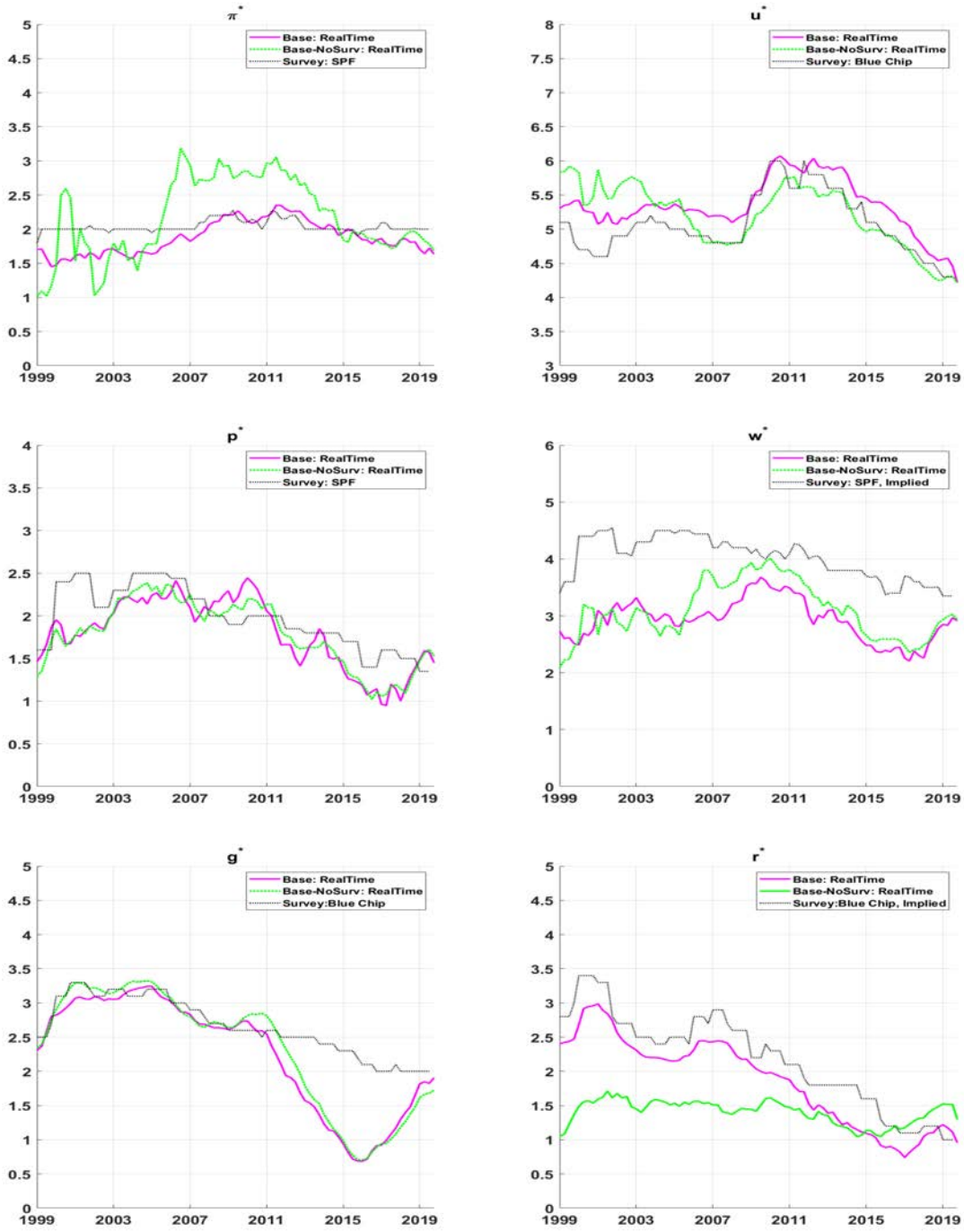
Notes: The plot denoted Base corresponds to smoothed (posterior mean) estimates based on the full sample information, i.e., 1959.Q4 through 2019.Q4. The plot denoted Base: RealTime corresponds to real-time recursive (posterior mean) estimates generated by estimating the Base model at different points in time, specifically 1999.Q1 through 2019.Q4. The credible intervals reflect the uncertainty around the posterior mean smoothed estimates. For w-star, plotted are the implied survey estimates of w-star, constructed by adding the survey estimates of p-star and pi-star.

Figure A4: Real-time Recursive Estimates of Stars: Base-NoSurv model



Notes: The plots denoted Base-NoSurv correspond to smoothed estimates based on the full sample information, i.e., 1959.Q4 through 2019.Q4. The plots denoted Base-NoSurv: RealTime correspond to real-time recursive estimates generated by estimating the Base-NoSurv model at different points in time, specifically 1999.Q1 through 2019.Q4.

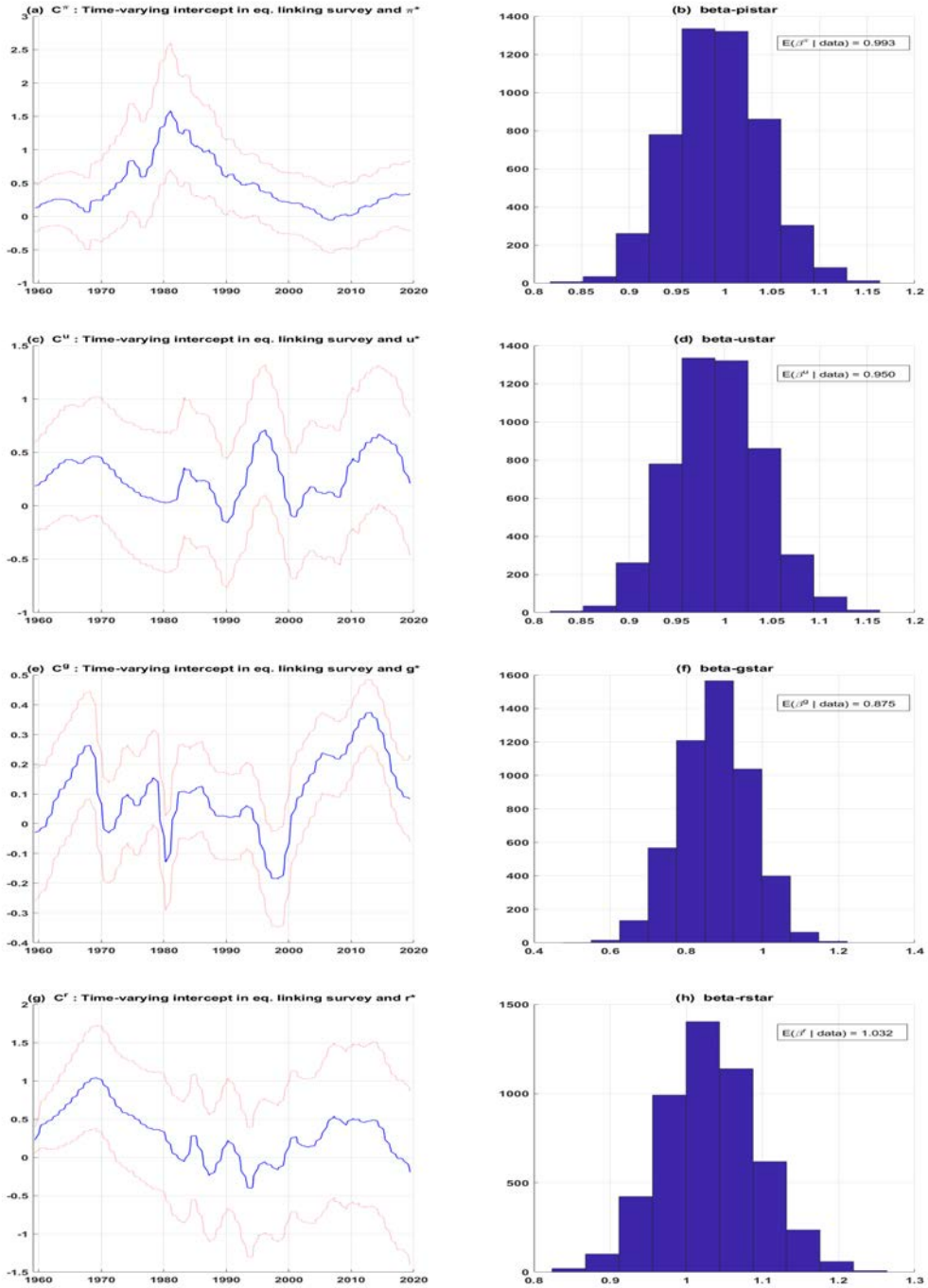
Figure A5: Real-time Recursive Estimates of Stars: Base model vs. Base-NoSurv model



Notes: The plots correspond to real-time recursive estimates generated by estimating Base and Base-NoSurv models at different points in time, specifically 1999.Q1 through 2019.Q4. To facilitate comparison, real-time estimates from either the Blue Chip or the Survey of Professional Forecasters (SPF) are also plotted.

A7. Estimated Relationship between Surveys and Stars

Figure A6: Estimated Link Between Survey Forecasts and Stars



Notes: The posterior estimates are based on the full sample (from 1959Q4 through 2019Q4).

A8. Additional COVID-19 Pandemic Results

Figure A7 presents posterior estimates of u-star, g-star, and r-star from the Base and Base-NoSurv models based on estimating data through 2020Q3. Also plotted to facilitate comparison are the corresponding posterior estimates based on estimation through 2019Q4. Figure A8, similarly, provides estimates of pi-star, p-star, and w-star. A visual inspection of the plots suggests the following four observations. First, estimates appear reasonable, indicating the model isn't blowing up. Second, adding pandemic data to the estimation sample has small effects on the historical (posterior mean) estimates of stars in the Base model and, for the most part (u-star being exception), also applies to the Base-NoSurv model. For u-star, there are some notable revisions in the estimates obtained from the Base-NoSurv model comparing between estimations pre- and post-pandemic. This considerable revision in the posterior mean of u-star is associated with significantly increased uncertainty, as evidenced by the larger width of the 90% credible intervals; however, in the Base model, the estimation with pandemic data leads to relatively small increase in the uncertainty about u-star.

Third, as would be expected (see Carriero et al., 2021), the precision plots indicate an uptick in uncertainty toward the end of the sample period associated with the pandemic data. Relatedly, the revisions to the historical estimates of precision are small in the case of the Base model but quite large for the Base-NoSurv model.

Overall, the Base model generally held up better because the survey forecasts help anchor the econometric estimates of stars to a reasonable range. Without it, extreme data movements in the unemployment rate profoundly influenced the econometric estimates of u-star in the Base-NoSurv model.

We believe that the rich features of our models, which include: (1) modeling the changing economic relationships via the implementation of time-varying parameters; (2) allowing for the changing variance of the innovations to various equations (i.e., SV); (3) imposing bounds on some of the random walk processes; (4) joint modeling of the output gap and unemployment gap in particular; and (5) using survey forecasts, helped position our models to handle the pandemic data better.

Carriero et al. (2021) using monthly Bayesian VARs show that models that allow for SV better handle pandemic observations than those without SV. But even models with SV have a drawback in the context of the pandemic data. This drawback arises from the standard approach to modeling SV, which assumes a random walk process or a very persistent AR process. So in the face of a temporary spike in volatility, the model will attribute this spike incorrectly to a persistent increase in volatility. Inspired by the outlier treatment method of Stock and Watson (2016) for UCSV models, Carriero et al. (2021) propose an outlier-adjusted SV method that models the VAR residuals as a combination of persistent and transitory changes in volatility.

We believe that this drawback of standard SV applies more to monthly VARs and to a lesser extent in quarterly models, as is the case here. However, we stress that Stock and Watson's treatment method for outliers can be conveniently implemented in our modeling framework. To keep the length of the paper manageable, we leave this extension for future research.

The COVID-19 pandemic provides an excellent real-time illustration of the importance of using survey expectations data in the econometric estimation of the stars. The unprecedented nature of the pandemic crisis and the extreme movements in the data induced by the pandemic are too volatile to provide a timely and credible signal about the long-run macroeconomic consequences. Complicating problem of extracting the signal from the data during the pandemic period is that consensus has been developing (perhaps rightly so) to treat macroeconomic data

for the periods 2020Q2 and 2020Q3 as outliers in estimating the macroeconometric models; see Schorfhedie and Song (2020), Carriero et al. (2021), among others.

On the other hand, judgment assessment informed by past event studies and an understanding of many decades of economic research indicate that the COVID pandemic is likely going to have implications for long-run productivity growth (p-star), the growth rate of potential output (g-star), the natural rate of unemployment (u-star), and the long-run real rate of interest (r-star); see Jorda, Singh, and Taylor (2020). As time rolls forward, and more is revealed about the possible long-term macroeconomic impact of the pandemic on the underlying trends, survey participants would judgmentally adjust their estimates of long-run projections in a more timely manner, and, by extension, our Base model, which incorporates the long-run survey projections.

Base model vs. external sources: Post-pandemic recession

We next compare our Base model estimates with those produced by external sources (and/or models) to assess further the reliability of our Base model estimates post-pandemic recession. Figure A9 compares the estimates of the output gap (panel a), r-star (panel b), u-star (panel c), and pi-star (panel d) from the Base model to the outside estimates.[5] The estimates are based on data through 2020Q3 (specifically the data vintage corresponding to late November 2020). In the case of the CBO, the projections correspond to an update as of late July 2020.

The plots in panel (a) indicate remarkable similarity between the posterior mean estimate of the Base model's output gap and the CBO output gap. Compared with Morley and Wong (2020), even though before the pandemic, the base model's output gap estimates indicated less tight resource utilization, for 2020, all three imply broadly similar inference. Morley and Wong (2020), based on a BVAR approach, could be viewed more flexibly than ours because it explicitly considers the possible error correlation across model equations. However, at the same time, their approach could be deemed less flexible than ours because it does not explicitly model time variation in parameters and stochastic volatility – i.e., it abstracts from the issue of "a changing economic environment." Both the CBO and Morley and Wong (2020) estimate the output gap at -3.6% for 2020Q3, with the Base model at -2.5%. (Interestingly, the output gap estimate from the Base model variant without SV in output gap and the unemployment gap is at -3.5%.)

Panel (b) plots the estimates of r-star from various sources. Except for Laubach and Williams (2003) [LW], all others are based on information available as of late November 2020. LW's estimate reflects information through August 2020. Comparing between 2019Q4 and 2020Q3, the Base model, Johannsen and Mertens (2019), and Del Negro et al. (2017), all estimate r-star to have changed only a little: Base model: from 0.87% to 0.76%; Del Negro et al. (2017): from 1.11% to 1.08%; and Johannsen and Mertens (2019): from 1.48% to 1.47%. In contrast, Lubik and Matthes (2015) have r-star increasing from 0.64% to 1.00%. However, in their estimate, r-star first falls from 0.64% to -0.68% and then bounces back to 1.0% in 2020Q3. Their estimate of r-star displays considerable volatility compared to others' estimates.

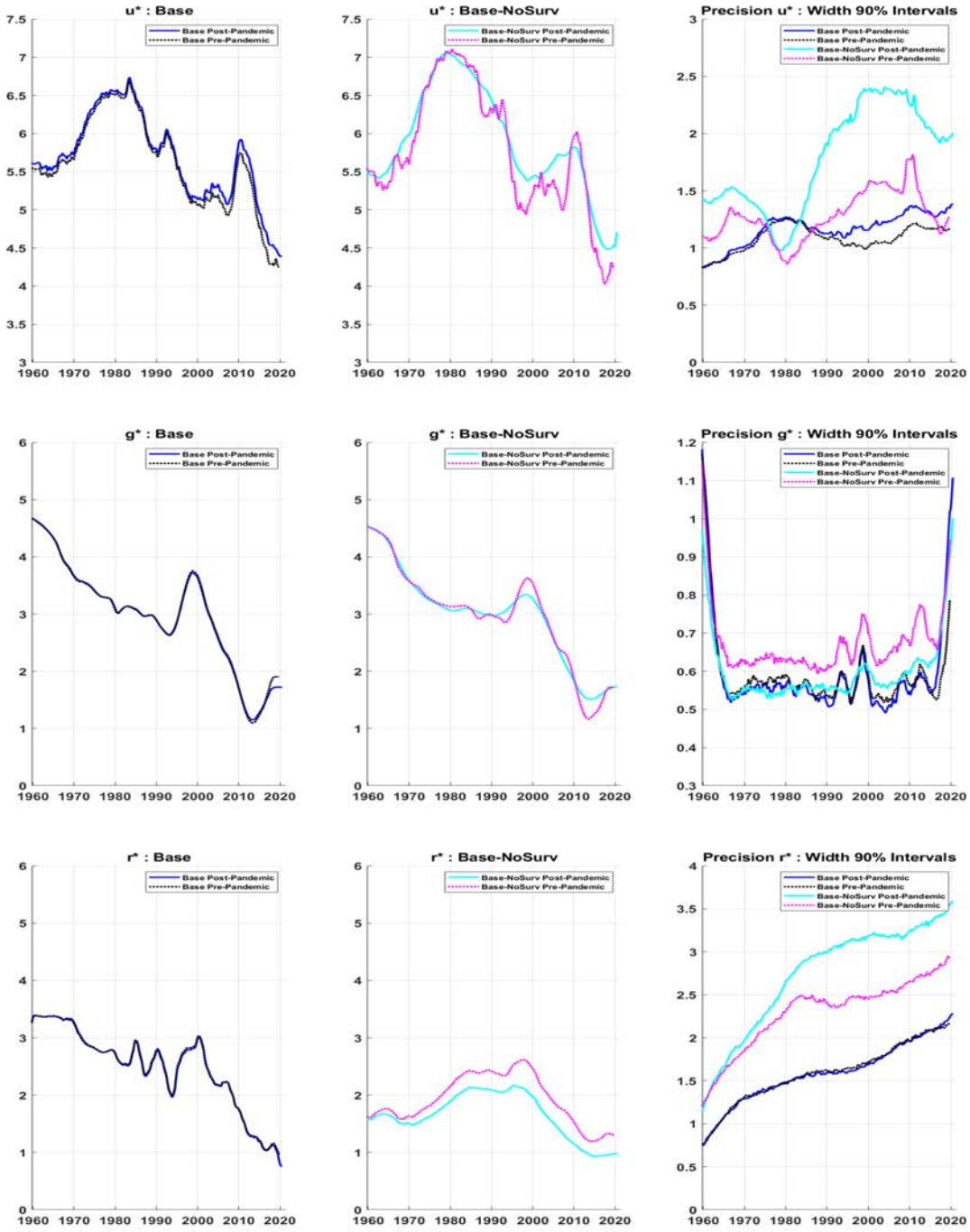
[5]Morley and Wong (2020) estimates are based on their updated work (Berger, Morley, and Wong (forthcoming)) and are available to download from outputgapnow.com. The estimates were downloaded in the last week of November, which included the nowcast estimate for 2020Q4 that we do not plot. We thank Murat Tasci for providing the estimates of u-star from the Tasci (2012) model. We also thank Benjamin Johannsen for providing the r-star estimates from Johannsen and Mertens (2019). The LW estimates of r-star were downloaded from the New York Fed's website. Del Negro et al. (2017) estimates of r-star were downloaded from github.com/FRBNY-DSGE/rstarBrookings2017. Lubik and Matthes' estimates were downloaded from the Richmond Fed's website in late November 2020.

Panel (c) plots the estimates of u-star from four sources: Base model, the CBO, Tasci (2012), and Chan, Koop, and Potter (2016). Comparing between the Base model and the CBO, the contours of the u-star plots are similar. But the levels through 2010 are notably different, with the CBO lower than the Base model. In 2020Q3, both CBO and Base indicate u-star at 4.4%. Interestingly, both the CBO and the Base model have u-star remaining mostly stable between 2019Q4 and 2020Q3, suggesting that they attribute most of the increase in the pandemic's unemployment rate to the cyclical component. It is worth highlighting that the (median) estimate of u-star reported in the September 2020 Summary of Economic Projections, which the Federal Reserve compiles, also indicated a stable u-star (at 4.1%) between 2019Q4 and 2020Q3.

Broadly speaking, the contour of the u-star implied by the CKP (bivariate Phillips curve) is similar to the Base model and the CBO. But the estimated level of u-star is significantly higher. According to the CKP model, the estimated u-star in 2020Q3 is 5.7%, just a tenth higher than in 2019Q4. The Tasci (2012) model, which is based on the flow rates in-and-out of unemployment, is significantly impacted by the pandemic data, as u-star is estimated to have increased from 4.7% in late 2019 to 5.2% in 2020Q3. Part of the explanation of more significant movements in u-star seen in the Tasci model in response to pandemic data is that the model is estimated using maximum likelihood methods, which are known to have done a relatively inferior job in handling extreme pandemic-induced movements in variables. More generally, Tasci (2019) documents the challenges of estimating u-star in real time with these models during crisis periods.

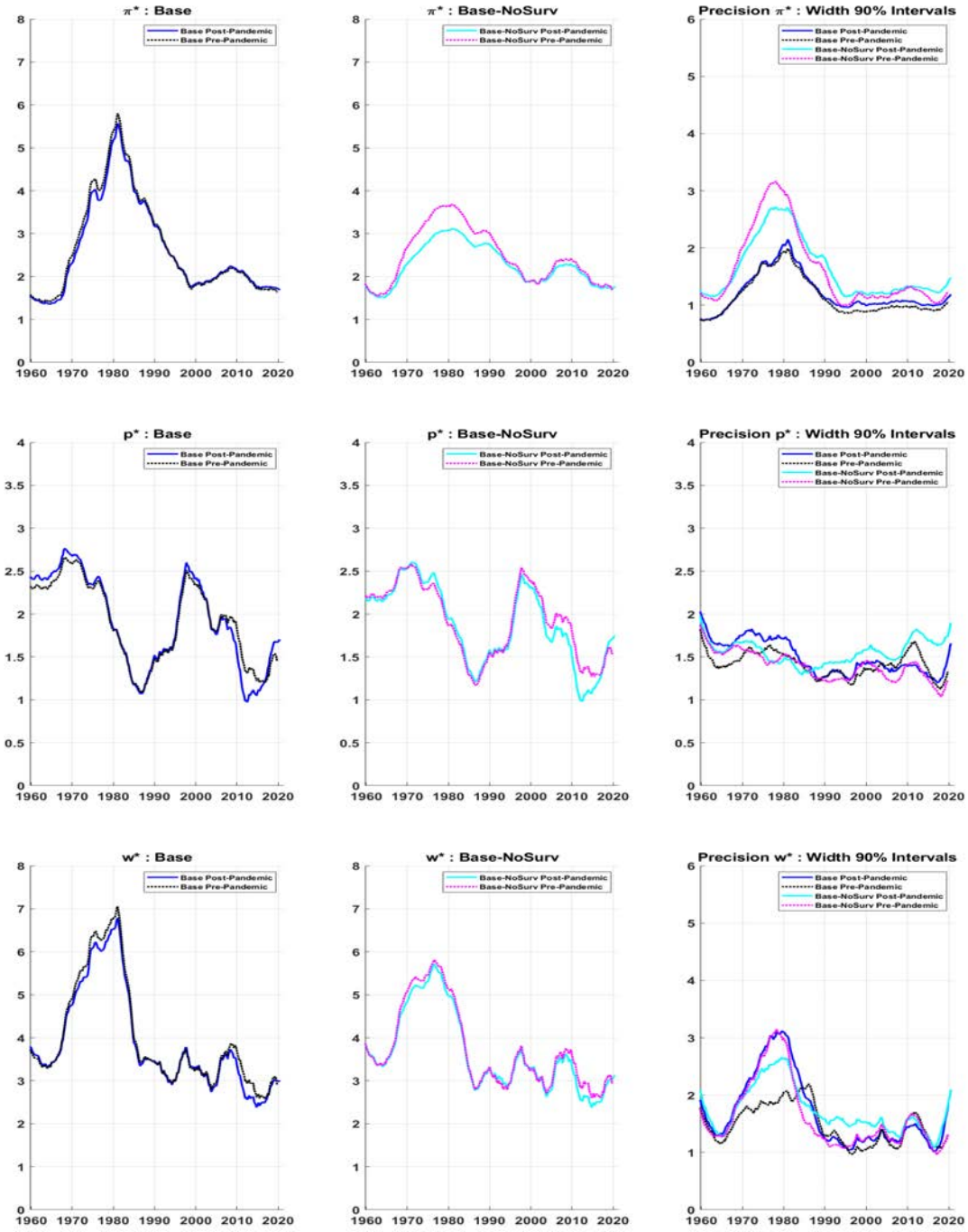
Panel (d) presents pi-star estimates from three sources: the Base model, CCK model, and CKP model. All three models indicate that pi-star remained stable between 2019Q4 and 2020Q3. However, the pi-star estimates differ slightly across models, with the Base model at 1.70%, CCK at 1.50%, and CKP at 1.44% (in 2020Q3).

Figure A7: Estimates of Stars pre- vs. post-COVID Recession



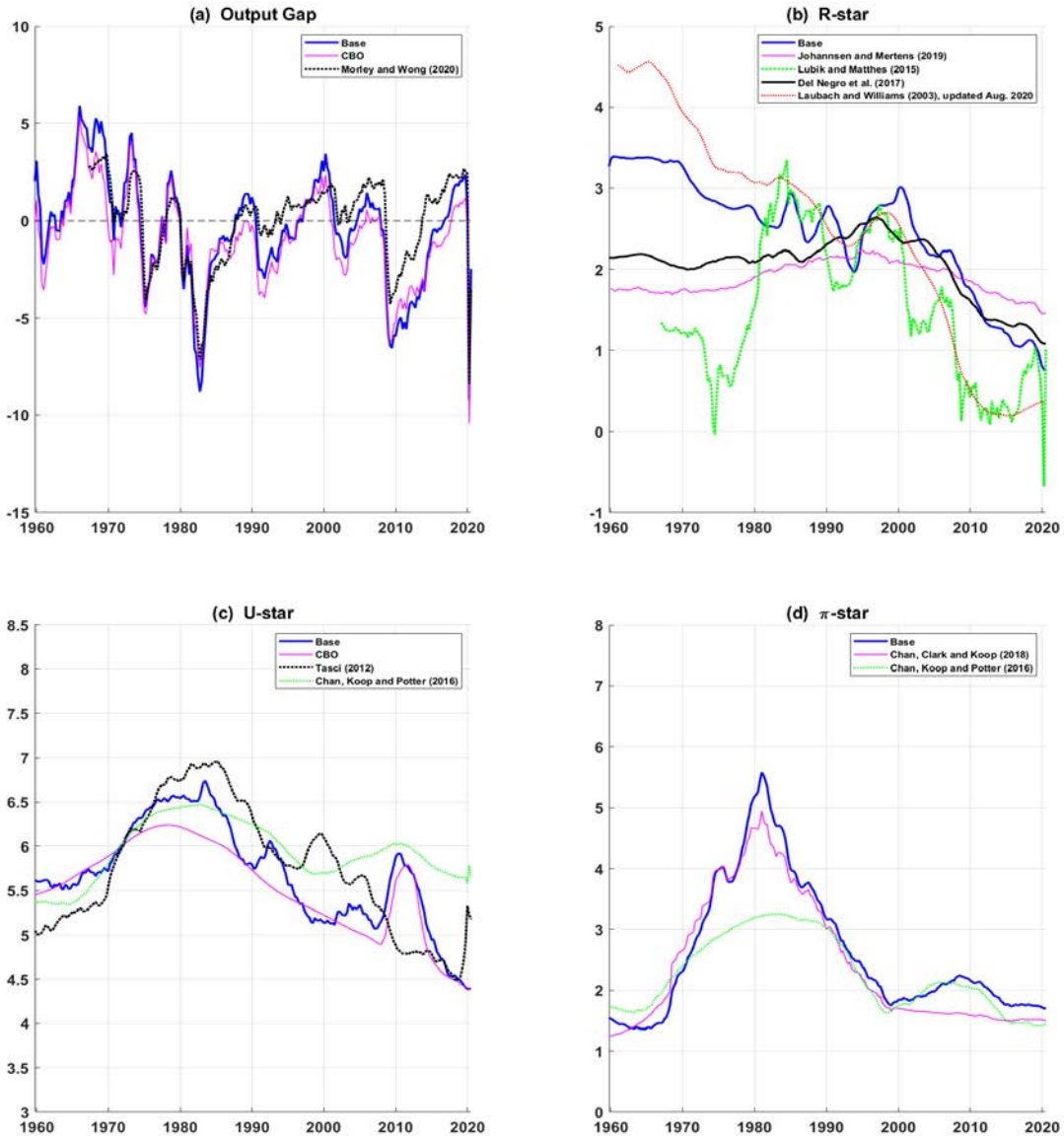
Notes: The plots labeled Pre-Pandemic reflect posterior estimates based on information in the sample 1959Q4 through 2019Q4, and plots labeled Post-Pandemic reflect posterior estimates based on the sample 1959Q4 through 2020Q3.

Figure A8: Estimates of Stars pre- vs. post-COVID Recession



Notes: The plots labeled Pre-Pandemic reflect posterior estimates based on information in the sample 1959Q4 through 2019Q4, and plots labeled Post-Pandemic reflect posterior estimates based on the sample 1959Q4 through 2020Q3.

Figure A9: Estimates of Stars post-COVID Recession: Base vs. Outside



Notes: In the case of Johannsen and Mertens (2021), Del Negro et al. (2017), and Lubik and Matthes (2015), the estimates plotted are the posterior median; for all others it is the (posterior) mean estimate.

A9. R*: Backcast Survey R* from 1959-1982

The survey estimates of g-star, u-star, and pi-star are direct reads from the survey. In contrast, the r-star survey estimate is not a direct estimate. Instead, it is inferred from the Blue Chip survey's long-run estimates of the GDP deflator and short-term interest rate (3-month Treasury bill) using the long-run Fisher equation, specifically, the long-run forecast of 3-month Treasury bills less the long-run forecast of the GDP deflator. To this differential, we add +0.3 to reflect the average differential between the federal funds rate and the 3-month Treasury bill (r-star refers to the long-run equilibrium federal funds rate).

Survey projections are not available before 1983Q1. To fill in estimates for the survey variables between 1959Q4 and 1982Q4, we use the CBO's long-run projections in the case of real GDP growth and the unemployment rate. In the case of inflation, we use the PTR series available from the Federal Reserve Board's website; this series is used in many studies employing long-run expectations of inflation (e.g., CCK, Tallman and Zaman, 2020). We do not have a readily available historical source for long-run forecasts for interest rates (and r-star). So we backcast a particular time series of implied r-star using the CBO's long-run projections of g-star. Specifically, we first fit a simple linear regression model over the post-1983 period that regresses survey r-star on a constant, its lags (2 lags), and a one-period lag "gap," defined as the difference between survey r-star and survey g-star. We use the estimated model and the CBO's long-run projections of g-star over the sample 1959Q4 through 1982Q4 to backcast the implied survey r-star estimates. (When backcasting, the initial values of r-star for 1959Q2 and 1959Q3 are assumed to be identical to the CBO's g-star.)

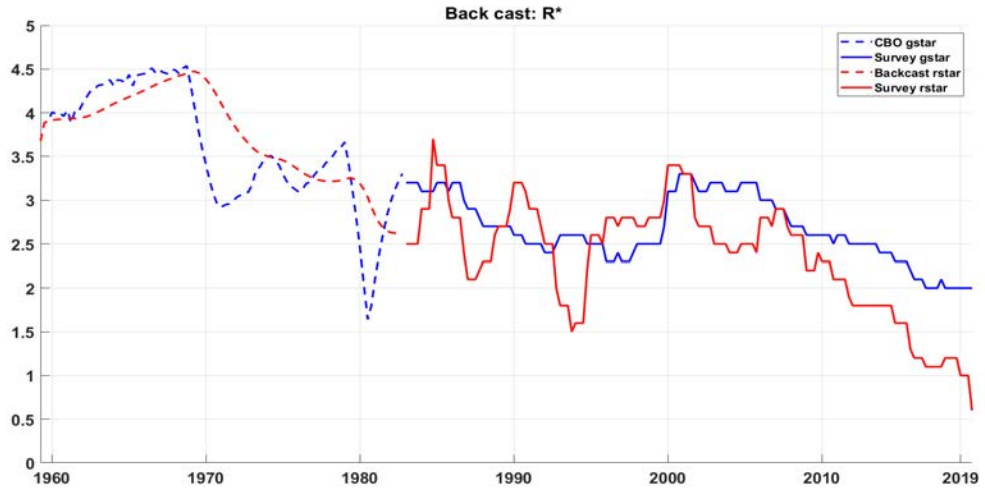
$$r_t^{*,Surv} = c + \beta_1 gap_{t-1}^{r*,g*,Surv} + \beta_2 r_{t-1}^{*,Surv} + \beta_3 r_{t-2}^{*,Surv} + \varepsilon_t^{*,Surv}, \;\; \varepsilon_t^{*,Surv} \sim N(0, \sigma^2_{*,Surv}) \quad (83)$$

where, $gap_t^{r*,g*,Surv} = g_t^{*,Surv} - r_t^{*,Surv}$

The OLS estimation yields $c = -0.0745$; $\beta_1 = 0.06$; $\beta_2 = 1.167$; $\beta_3 = -0.148$.

Figure A10 plots the survey g-star and r-star estimates in solid lines, and the CBO's g-star and the backcast r-star in dashed lines.

Figure A10: Survey r* and g*



A10. R* : Additional Full Sample Results

A10.a. Role of data vs. prior in shaping r-star

Kiley (2020), using a model in which r-star follows an RW process, documents an essential finding that data provide very little information in shaping the r-star process. Hence, the model-based r-star estimate is mainly driven by the modeler's prior views. Our results generally confirm Kiley's findings. However, in our Base specification, where the variance of the g-star process influences both the prior and the posterior for the r-star process, the data do influence the r-star estimate because we find evidence that the data provide information about the g-star process. This latter evidence of the data's influence on the identification of g-star is also noted by Kiley (2020).

We begin by comparing the prior and posterior estimates of the parameter σ^2_{r*}, which governs the shock variance of the r-star process, in the Base-NoSurv-R*RW model specification, which is a variant of the Base model that excludes survey information and assumes RW for r-star similar to Kiley (2020). The default prior for $E(\sigma^2_{r*}) = 0.1^2$), which is the same as in Gonzalez-Astudillo and Laforte (2020) but tighter than the 0.25^2 used by Kiley.[6] (Our choice of a tighter prior than Kiley is due to a significantly more complex model.) Our model estimation yields posterior estimates of 0.09^2 (with 90% credible intervals 0.07^2 to 0.12^2) for the parameter σ^2_{r*}, which is slightly different from the prior mean, suggesting that the data do play a role in shaping r-star.

We next confirm our finding by re-doing our exercise setting a looser prior for $E(\sigma^2_{r*}) = 0.173^2$), which equates to doubling the prior mean of the variance of the process governing the evolution of r-star. The updated model estimation yields posterior estimates of $E(\sigma^2_{r*}) = 0.14^2$ (with 90% credible intervals 0.11^2 to 0.17^2). The fit of this model (looser prior) to the interest rate data (and other model data) is significantly worse compared to the model specification with the default prior.

We explored the impact on the r-star estimates of even looser priors on the shock process governing r-star. We find that as the prior on the r-star process loosens, the data become more informative in shaping the r-star estimate (echoing Lewis and Vazquez-Grande, 2019). But it comes at the cost of a worsening model fit, higher volatility in the r-star estimate, and a worsening precision of r-star.

A10.b. Base vs. external models

In Figure A11, the left panel plots r-star from the Base, Base-NoSurv, and two external models: the seminal model of LW (dashed line) and the more recent model developed in Del Negro et al. (2017) (dotted line). As is the case with most r-star estimates presented in the literature, the LW estimate shows a marked decline in r-star from 2000 and beyond. As shown in the figure, compared to the r-star estimate from the Base, the LW estimate is lower over this period. Part of the explanation of this difference in the estimates comes from the different estimates of g-star (not shown).

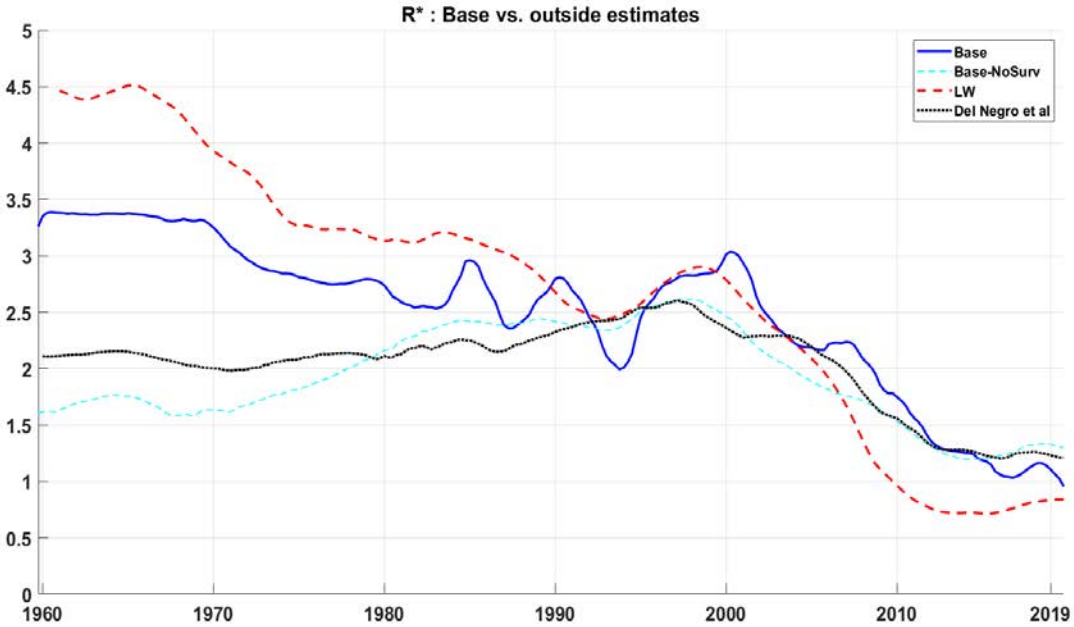
In the LW model, the mechanical reason for this steadily declining trajectory of r-star is coming from the fact that their model estimate of g-star has been steadily declining over the same period. Over this period, GDP grew just slightly above their estimate of g-star, even

[6] We also explore a model specification in which prior variance is set at 0.25^2. The fit of this specification was significantly inferior, and the r-star estimate was quite volatile.

though the real short-term interest rate is significantly below zero over this period. The model explains the combination of moderate growth in GDP (suggesting a small positive output gap) and a negative real short-term interest rate through a low level of the r-star estimate so as to obtain a negative real interest rate gap (see LW). In our Base (and Base-NoSurv) model, because the estimate of g-star is even lower than LW's, which implies a more positive output gap (than LW), a less negative real interest rate gap (than LW) is required to explain the output gap. The less negative real interest rate gap (i.e., a smaller interest rate gap) implies a higher level of r-star than LW.

The r-star estimate from Del Negro et al. is stable around 2% from 1960 through early 1980 and then slowly moves up, reaching 2.5% by late 1990. From there on, it begins a gradual decline, ending 2019 at 1.2%, identical to the Base-NoSurv, and two-tenths lower than Base. Broadly speaking, the contours of r-star from Del Negro et al. model closely resemble Base-NoSurv model. It is worth noting that Del Negro et al. also utilize survey expectations on r-star to estimate r-star but their approach to how they model the link between the two is very different from ours.[7] They also assume a relationship between g-star (in their case, long-run productivity growth) and r-star. However, their model structure is different compared to ours.

Figure A11: R* estimates



[7]Del Negro et al. use survey expectations from the Survey of Professional Forecasters, which start from 1992 onward. In addition, in their framework survey expectations are one of the several financial indicators they use to extract a common trend. So arguably, in their approach, the survey expectations of r-star will be less influential in driving r-star than in our approach, in which a direct connection between r-star and the survey data is assumed.

A10.c. Assessment of policy stance

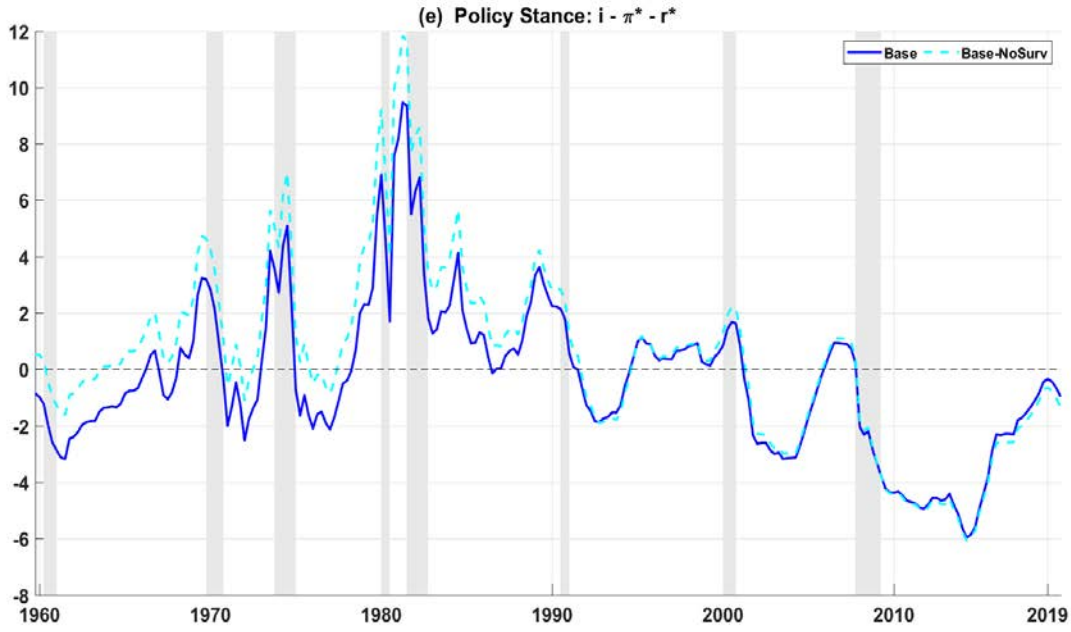
Figure A12 provides an assessment of the stance of monetary policy. Following Pescatori and Turunen (2016), we gauge monetary policy's stance as the deviation of the short-term nominal interest rate from the long-run nominal neutral rate of interest (defined as the sum of r* and pi*) – this is the interest rate gap from the Taylor-rule equation. A positive interest rate gap characterizes a restrictive monetary policy stance, and a negative interest rate gap implies a stimulative stance. The solid line corresponds to the posterior mean estimate of the policy stance inferred from the Base model and the dashed line to that inferred from the Base-NoSurv model. Even with notable differences in the estimates of r-star across the two models, the assessment of the policy stance is remarkably similar throughout the sample.[8]

According to our model(s) estimates, the policy stance appeared to be slightly restrictive before the Great Recession, but at the onset of the Great Recession, the policy stance immediately turned accommodative. Since then, it has remained very accommodative (reflecting the effects of unconventional monetary policy). After peaking in late 2015, the degree of accommodation has gradually declined (i.e., the interest rate gap has become less negative), such that, by the end of 2019, it has edged closer to the neutral threshold.

A closer inspection of the figure reveals an interesting insight. Since 1990, both the degree and duration of policy accommodation in response to the recession have been more significant than in the previous recession. For instance, the monetary policy stance was more accommodative in terms of both level and duration following the 2001 recession than following the 1990-1991 recession. Similarly, during and following the Great Recession, the policy stance, in terms of level and duration, was more significant than following the 2001 recession.

[8]So why are they so similar? The answer lies in the differences in pi-star estimates across two model specifications. In other words, the differences between r-star estimates across the two models are compensated (i.e., offset) by the differences between pi-star, such that the assessment about the stance of monetary policy across the two models is strikingly similar. For instance, in the 1960s, the r-star estimate from the Base is on average 1.34 percentage points (ppts) higher than that of Base-NoSurv. However, over the same period, the pi-star estimate from the Base model is 0.54 ppt lower than that of the Base-NoSurv model, which reduces the difference between their associated assessments of policy stance.

Figure A12: Policy Stance



Notes: Plotted are the posterior mean estimates based on estimation using the full sample (from 1959Q4 through 2019Q4).

A11. Pi* : Additional Full Sample Results

A11.a. Pi-star comparison Base vs. outside models

In Figure A13, panel (a) plots posterior mean estimates of pi-star from some related (smaller) models from the literature alongside Base to facilitate comparison. In particular, estimates are shown for CKP, CCK, and the celebrated UCSV model of Stock and Watson (2007). (These three models could be viewed as restricted variants of our Base model.)[9] Panel (b) plots the corresponding precision estimates of pi-star.

There are some interesting similarities and differences across the pi-star estimates. Whereas UCSV displays very volatile and erratic estimates of pi-star, other models show a smoother evolution of pi-star. CKP indicate a lower estimate of pi-star than others from the early 1970s through the late 1980s. The primary factor contributing to lower pi-star in CKP is the model assumption of a bounded random walk for pi-star. As discussed in CKP, the addition of bounds on pi-star leads the model to attribute a substantial share of the observed high inflation of the 1970s to the increased persistence of the inflation gap and only a small increase in pi-star. Hence, pi-star is estimated to have risen less than implied by other models. For instance, the CCK model had pi-star peaking at 4.9%, Base at 5.8% and CKP at 3.2%. As alluded to in CKP, this small rise in pi-star is consistent with a specific narrative that during the Great Inflation period, the Fed had a low implicit target for inflation but was either unable to or unwilling to correct large deviations of inflation from the target.

The contours of pi-star from Base are similar to CCK through 2000, but from 2000 to 2012, Base is identical to CKP, with CCK a touch lower. It is interesting to note that from the early 2000s through 2010, both Base and CKP indicate pi-star at 2%. From 2012 through 2019, both Base and CKP gradually drift lower to 1.6% and 1.3%, respectively; CCK is at 1.5%.

Panel (b), which plots the corresponding precision of pi-star, reveals some interesting patterns. First, the precision of pi-star evolved generally with the level of pi-star. As pi-star increased during the Great Inflation, pi-star became more uncertain, i.e., more imprecise. Subsequently, as pi-star trended lower during the Volcker disinflation, so did the uncertainty about it (i.e., precision increased). Second, comparing across models, there is significant heterogeneity in the precision of pi-star. From 1960 through the mid-1970s, the Base model indicates the most precise pi-star, followed by CCK and CKP. The UCSV model shows volatile estimates of precision, sharply fluctuating between the most precise to the least precise. From the mid-1970s through 2019, the CCK model indicates the most precise (least uncertain) pi-star, followed by Base, CKP, and UCSV. CCK had the uncertainty of pi-star gradually trending down starting in the mid-1970s. In contrast, in others, the uncertainty continued to trend higher until peaking in the early 1980s.

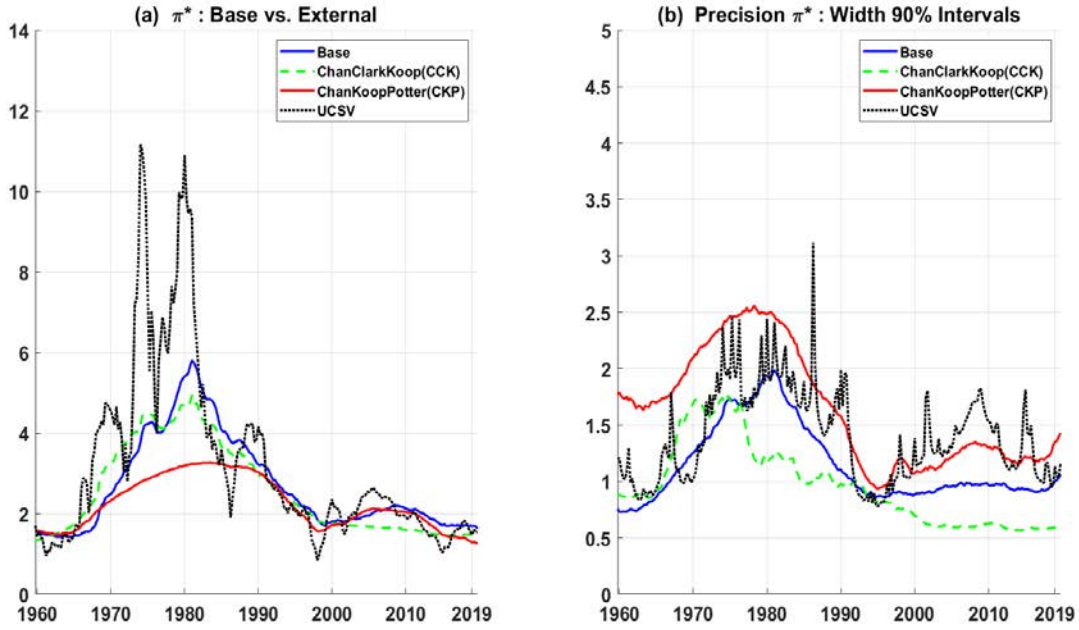
Third, between 2000 and 2019, the uncertainty around pi-star implied by CCK and Base has been reasonably stable, an artifact of the use of survey data. During this period, the precision of pi-star implied by CCK is on average 40 basis points higher (i.e., uncertainty is lower) compared to Base. This improved precision of CCK is interesting because both CCK and Base utilize information from survey expectations of inflation. However, at the same time, compared to Base, which has a rich structure (hence more parameters), the CCK model is parsimonious, as it uses information from survey expectations only (in addition to inflation's own history) to estimate pi-star.

An additional factor that could contribute to the differential in precision is that, unlike Base,

[9] Whereas in estimating the UCSV model, Stock and Watson (2007) fix the parameters governing the smoothness of the SV processes, we estimate them.

CCK allows SV in the pi-star equation. A more in-depth inspection of the estimation results reveals that the primary factor driving the superior precision of the CCK estimate of pi-star compared to Base is tighter priors on the assumed relationship between survey expectations and pi-star. And that translates into a posterior estimate implying a stronger connection between survey expectations and pi-star in CCK than Base.

Figure A13: Pi* estimates: Base vs. External models



Notes: The posterior estimates are based on the full sample (from 1959Q4 through 2019Q4). In all cases, the inflation measure is PCE inflation.

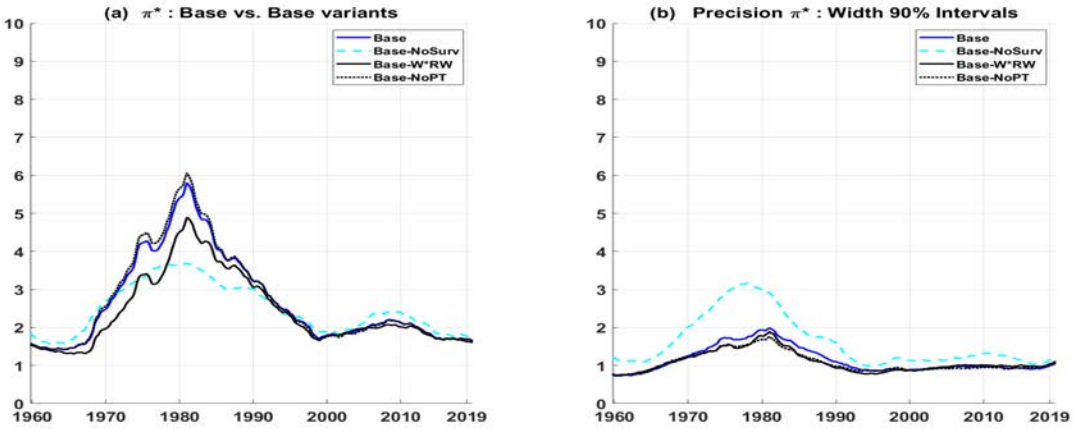
A11.b. Sensitivity of pi-star to modeling assumptions

Figure A14, panel (a) indicates the sensitivity of the pi-star estimates to modeling assumptions. The plot labeled Base-W*RW is the variant of the Base model that removes the theoretical restriction imposed by equation (26) and instead assumes a random walk assumption for w-star. Comparing the Base and Base-W*RW plots indicates the effects of the theoretical restriction on pi-star. As shown, the posterior mean estimate of pi-star from Base-W*RW is marginally lower than Base in the period 1970 through the early 1980s (Great Inflation period). However, from there on, estimates of pi-star are identical. During the high-inflation period, compared to the Base model, the Base-W*RW allocates a higher share of the increase in inflation to the persistence component than pi-star (i.e., the random walk component); see figure A15. Hence, the lower level of pi-star in Base-W*RW than Base.

The plot labeled Base-NoPT is the variant of the Base model that eliminates the pass-through from prices to wages, modeled via equation (29b)—doing so results in a slightly higher pi-star (Base-NoPT) from 1970 through the early 1980s. However, thereafter, estimates of pi-star are identical between Base and Base-NoPT. During the high-inflation period, compared to the Base model, the Base-NoPT allocates a lower share of the increase in inflation to the persistence component than pi-star; hence, the higher level of pi-star in Base-NoPT than Base. Based on the model comparison (shown in table A6), the Base-W*RW model's fit to the inflation data is comparable to the Base but the fit to other data is significantly inferior compared to Base. In the case of Base-NoPT, the fit to the inflation data is marginally better than Base. However, the overall fit of the Base-NoPT is significantly worse than Base. The Base-NoPT model's reduced fit is the net effect of its reduced ability to fit wages and its improved ability to fit prices.

We also explored a variant of the Base model that allowed the pass-through from wages to prices in the price inflation equation, denoted Base-PT-Wage-to-Prices. The estimates of pi-star (and of other parameters) are identical to those of the Base; hence, they are not shown. Therefore, not surprisingly, as reported in Table A6, both models' ability to fit inflation data is very similar. We also highlight that allowing SV in the inflation equation is very important, as evidenced by a significantly reduced fit of the Base-NoSV model, which is the Base model variant that does not feature SV in any model equations.

Figure A14: More Estimates for Price Inflation Block



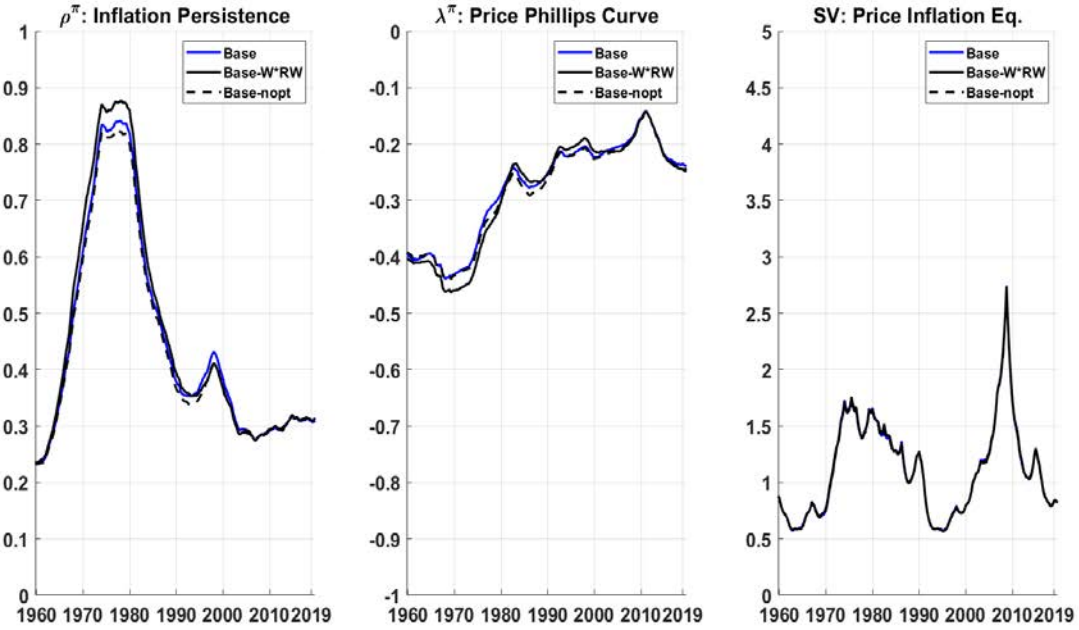
Notes: The posterior estimates are based on the full sample (from 1959Q4 through 2019Q4).

Table A6: Model Variants for Price Inflation

Model variant	MDD of Price inflation	MDD Model
Base	-365.2	-1587.3
Base-NoSurv	-365.6	-1587.1
Base-W*RW	-365.2	-1594.3
Base-PT-Wages-to-Prices	-365.0	-1586.8
Base-NoPT	-364.4	-1592.3
Base-NoSV	-412.5	-1939.7

A11.c. Pi-star estimates for some variants of the Base model

Figure A15: Pi* estimates: Base vs. Base model variants



Notes: The posterior estimates are based on the full sample (from 1959Q4 through 2019Q4).

A12. P*: Base Comparison with Kahn and Rich (2007)

In this section, we compare our model-based estimates of p-star with the narrative about p-star implied from the two-regime Markov switching model of Kahn and Rich (2007).[10] A regime-switching framework (as in Kahn and Rich) allows for deterministic values of p-star, where the number of deterministic values equals the number of possible regimes. Accordingly, in a 2-regime setup, the estimated p-star would periodically alternate from one regime (e.g., a low productivity regime) to the other regime (e.g., a high productivity regime). In contrast, the random walk assumption for p-star adopted in this paper (and in others such as Roberts, 2001; Edge et al., 2007; Benati, 2007) allows for the possibility that p-star may be (slowly) changing in every period. This latter assumption implies that the possible values of p-star could equal the number of periods in the estimation sample. The differences in the stochastic conception between the two frameworks complicate direct comparison in p-star.

One possible, albeit imperfect, approach to comparing the implied p-star from two frameworks is to use the regime-switching model's identified regimes to assess how well those corroborate p-star estimates implied from the RW assumption model. Specifically, for the RW model, compute the "average" p-star over the specific periods (identified regimes). Then assess the following: (1) whether the "average" rates imply a characterization of regimes that corroborate the identified regimes; and (2) how close the "average" rates of p-star are to the deterministic values of p-star estimated in the regime-switching model. We use this approach to compare the estimates of p-star from our models to the p-star estimated by the Kahn and Rich model.

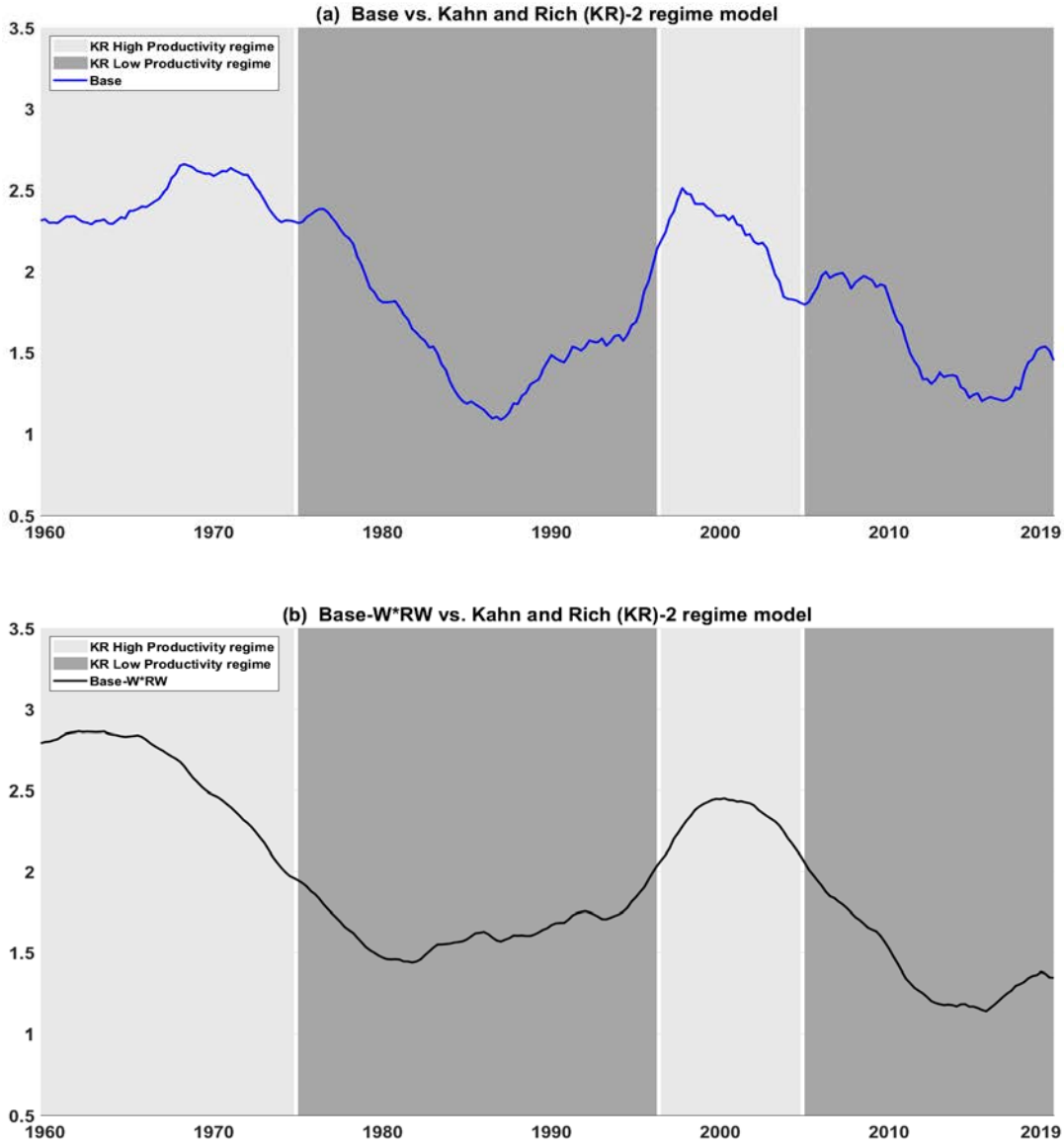
Figure A16 presents the comparison of p-star. Panel (a) compares the Base model with the Kahn and Rich model, and panel (b) compares the Base-W*RW model with the Kahn and Rich model. In the panels, the shaded areas refer to the two regimes identified by the Kahn and Rich model using the same vintage of data as our models. The lighter shaded area corresponds to the "high productivity regime," and the darker shaded area corresponds to the "low productivity regime." Their model identifies two subperiods of high productivity regimes: the beginning of our sample through 1974Q4 and 1996Q3 through 2004Q4. Similarly, their model identifies two subperiods of low productivity regimes: 1975Q1 through 1996Q2 and 2005Q1 through the end of the sample, 2019Q4.

Next, we compare the "average" rates for the two regimes implied by our models to the Kahn and Rich model. The Base model implies for a low productivity regime an "average" rate of 1.6% (for both subperiods) and for a high productivity regime an "average" rate of 2.4% (subperiod beginning of our sample through 1974Q4) and 2.2% (subperiod 1996Q3 through 2004Q4). The Base-W*RW model implies for a low productivity regime "average" rate of 1.7% (in the first subperiod) and 1.4% (in the second subperiod) and for a high productivity regime "average" rates of 2.6% and 2.3%, respectively. In comparison, Kahn and Rich's model implies a p-star of 1.33% for a low productivity regime for both subperiods (p-star is equal across subperiods by construction) and 2.96% p-star for a high productivity regime. For the low productivity regime, the implied p-star from the Base is slightly higher than Kahn and Rich's, but for the high productivity regime, Kahn and Rich's model is slightly higher.

Overall, this illustration suggests that the two approaches provide generally similar inferences about developments in p-star, and we view this as a useful result for macroeconomists tasked with modeling and tracking productivity developments.

[10] The estimates of p-star implied by the Kahn and Rich (2007) model are routinely updated and made available for download at James A. Kahn's website: http://sites.google.com/view/james-a-kahn-economics/home/trend-productivity-update .

Figure A16: P* Consistent with Narrative from 2-Regime Markov-Switching Model



Notes: The shaded areas refer to the two regimes identified by the Kahn and Rich model using the same vintage of data as our models. The lighter shaded area corresponds to the "high productivity regime," and the darker shaded area the "low productivity regime." The plots labeled Base and Base-W*RW are the posterior mean estimates based on the full sample (from 1959Q4 through 2019Q4).

A13. U* : Additional Full Sample Results

A13.a. Sensitivity of u-star to modeling assumptions

Figure A17, panel (a) plots estimates of u-star (posterior mean) from the variants of the Base model to highlight the sensitivity of u-star to modeling assumptions and the informational aspect of joint modeling. The plot denoted Base-NoBoundU* represents the Base model variant that eliminates the bound on the random walk process describing u-star. Doing so has a trivial effect on the estimates of u-star, the precision of u-star, and model fit. Comparing between panels (a) and (b), the posterior mean estimate of u-star is quite similar across Base and Base-NoBoundU*. Similarly, there is little change in the u-star estimate's precision across the two models, with Base only marginally better in the latter part of the sample (as shown in panel e). Not surprisingly, the Bayesian model comparison suggests equal support for both Base and Base-NoBoundU*.

We highlight two noteworthy comments in regard to the implementation of bounds on u-star. First, the trivial difference in the estimates between Base and Base-NoBoundU* arises because the bounds defined on u-star are wide. Put differently, the values of the bound we have set are not binding on the Base model. Second, we find that using bounds on u-star is extremely important in the Base-NoSurv, as it helps keep the estimation tractable. In other words, the advantages of using bounds on the random walk processes that were stressed in CKP were in full display in the estimation of Base-NoSurv. Hence, we prefer to keep bounds on u-star in our main models.

The other model variants plotted in panel (a) are all nested specifications of the Base model: the Bivariate model of GDP and the unemployment rate (a Base model that excludes survey information and everything else except the equations describing the dynamics of GDP and the unemployment rate); Bivariate+Surv, which is bivariate but adds survey data for GDP and unemployment; and CKP Adjusted, which is a bivariate model of inflation and the unemployment rate as in CKP but with no bounds on pi-star and augmented with SV. For visual reasons (to limit the number of plots), u-star from Base in panel (a) is not shown. However, for the sake of discussion, we could treat the plot representing Base-NoBoundU* as the estimate for the Base model, since they are identical to each other (as discussed in the preceding paragraph).

These plots show that different model specifications could provide very different signals about the level of long-run unemployment, indicating the sensitivity of u-star to modeling assumptions. A model specification that infers the estimate of u-star from inflation and unemployment data only, i.e., the price Phillips curve (CKP Adj. model), has a lower trajectory of u-star compared to Base. The model specification that infers the estimate of u-star from GDP and the unemployment data only, i.e., the Okun's law relationship (Bivariate model), displays larger fluctuations than the CKP Adj. model. Once the Bivariate model is augmented with survey data for GDP and unemployment, its trajectory resembles that of the Base model.

Panel (b) compares the u-star estimates from the main model specifications with the CBO estimate of the long-run unemployment rate. Interestingly, for the most part, the CBO u-star's contour is similar to our model-based estimates, though the level of the CBO u-star estimate is lower from 1960 through the mid-1990s. At the onset of the Great Recession and through the early phase of the economic recovery, all three have u-star continuing to move higher. Whereas Base and Base-NoSurv peak in late 2010 at 5.7% and 6.0%, respectively, the CBO has u-star peaking in late 2011 at 5.8%. Since then, u-star has steadily moved lower, with the pace of decline quite similar across CBO and Base. The CBO has u-star at the end of 2019 at 4.4%, about two-tenths higher than the Base model.

Precision of u-star

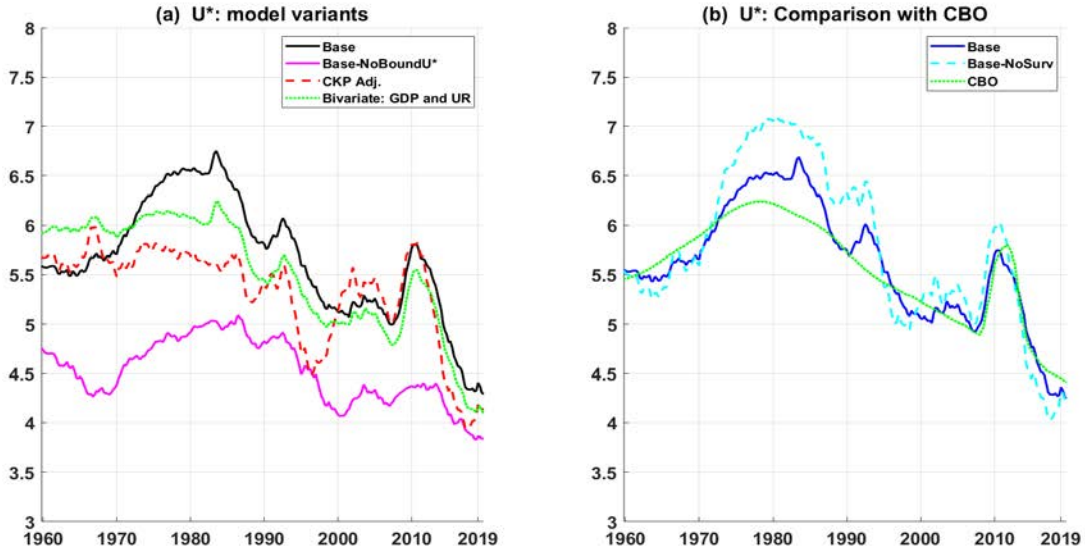
Panel (c) plots the precision of u-star estimates for Base, Base-NoSurv, Base-NoBoundU*, and Bivariate+Surv. We make several observations. First, comparing between Base and Bivariate+Surv, both of these model specifications use information from the surveys and Okun's relationship. However, Base relies on greater information and additional structure (e.g., wage Phillips curve, price Phillips curve, cyclical productivity, monetary policy stance via the Taylor-type policy rule) to infer u-star compared to Bivariate+Surv; hence, the more improved precision of the resulting u-star. And the model comparison indicates a significantly higher fit of the Base model to the unemployment data compared to the Bivariate+Surv.

Second, comparing with Base and Base-NoSurv, additional information from survey forecasts improves the precision of u-star, but this improved precision does not translate into improved model fit, which instead worsens somewhat (as shown in Table A7).

Panel (d) shows the precision of u-star for Base-NoSurv, Bivariate (GDP and unemployment), and CKP-Adj (which is a bivariate model of price inflation and the unemployment rate). As indicated earlier, u-star from Base-NoSurv is inferred from a broader information set and structure than the other two small-scale models. Accordingly, the Base-NoSurv estimate of u-star is, for the most part, more precise and the model comparison indicates a substantially higher fit to the unemployment data than the other two. The plots also show that u-star inferred from the Okun's law relationship (i.e., Bivariate model) is less precise than u-star inferred from the price Phillips curve (i.e., CKP-Adj model). In contrast, the Bayesian model comparison lends support to the Bivariate model over the CKP-Adj model.

The results also provide evidence that adding survey data to the Bivariate model (Bivariate+Surv) further improves both the precision of u-star (comparing Bivariate and Bivariate-Surv in panels c and d) and the fit to the unemployment data (Bivariate: 88.5 vs. Bivariate+Surv: 85.2, as shown in Table A7). This latter finding of improved fit from adding survey data is interesting because in the case of Base, adding survey data worsens the model fit (Base: 89.4 vs. Base-NoSurv: 93.1). Importantly, it suggests that survey forecasts of u-star are likely useful in the case of parsimonious models but of limited use for models that are already utilizing various sources of information to infer u-star.

Figure A17: U*



Notes: The posterior estimates are based on the full sample (from 1959Q4 through 2019Q4). In panel (a), CKP Adj. refers to the bivariate model of inflation and the unemployment rate as in CKP but with no bounds on pi-star.

Table A7: Model Comparison: Variants Focused on Unemployment Rate (UR)

	Base	Base-NoSurv	Base-NoBoundU*	Bivariate	Bivariate+Surv	CKP-Adj
MDD of UR	89.4	93.1	88.8	88.5	85.2	47.7

References

[1] Amisano, Gianni, & Raffaella Giacomini (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business and Economic Statistics*, 25(2): 177-90

[2] Barbarino, Alessandro, Berge, Travis J., Chen, Han, & Andrea Stella (2020). Which output gap estimates are stable in real time and why? Technical Report 2020-101, Board of Governors of the Federal Reserve System.

[3] Benati, Luca (2007). Drift and breaks in labor productivity. *Journal of Economic Dynamics & Control*, 31(8): 2847-2877.

[4] Berger, Tino, Morley, James, & Benjamin Wong (forthcoming). Nowcasting the output gap. *Journal of Econometrics*

[5] Carriero, Andrea, Clark, Todd E., Marcellino, Massimiliano, & Elmar Mertens (2021). Addressing COVID-19 outliers in BVARs with stochastic volatility. Federal Reserve Bank of Cleveland, Working Paper No.21-02. https://doi.org/10.26509/frbc-wp-202102

[6] Chan, Joshua, Clark, Todd E., & Gary Koop (2018). A new model of inflation, trend inflation, and long-run inflation expectations. *Journal of Money, Credit, and Banking*, 50(1): 5-53

[7] Chan, Joshua, & Ivan Jeliazkov (2009). Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modeling and Numerical Optimisation* 1: 101-120

[8] Chan, Joshua, Koop, Gary, Poirier, Dale J., & Justin Tobias L. (2019). Bayesian Econometric Methods 2nd edition. *Cambridge University Press*

[9] Chan, Joshua, Koop, Gary, & Simon M. Potter (2013). A new model of trend inflation. *Journal of Business and Economic Statistics* 31: 94-106

[10] Chan, Joshua, Koop, Gary, & Simon M. Potter (2016). A bounded model of time variation in trend inflation, NAIRU and the Phillips curve. *Journal of Applied Econometrics* 31: 551-565

[11] Chan, Joshua, & Rodney Strachan (2012). Estimation in non-linear non-Gaussian state-space models with precision-based methods. CAMA Working Paper series 2012-13.

[12] Del Negro, Marco, Giannone, Domenico, Giannoni, Marc, & Andrea Tambalotti (2017). Safety, liquidity, and the natural rate of interest. *Brookings Papers on Economic Activity*

[13] Edge, Rochelle M., Laubach, Thomas & John C. Williams (2007). Learning and shifts in long-run productivity growth. *Journal of Monetary Economics* 54: 2421-2438

[14] Geweke, John (1992). Evaluating the accuracy of sample-based approaches to the calculation of posterior moments. In *Bayesian Statistics 4* (eds. J.M. Bernardo, J.O. Berger, A.P. Dawid, A.F.M.Smith), 641-649. Oxford: Clarendon Press.

[15] Gonzalez-Astudillo, Manuel, & Jean-Philippe Laforte (2020). Estimates of r* consistent with a supply-side structure and a monetary policy rule for the U.S. economy. Finance and Economics Discussion Series Paper 2020-085

[16] Harvey, David, Leybourne, Stephen, & Paul Newbold (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13: 281-291.

[17] Johannsen, Benjamin K., & Elmar Mertens (2019). A time series model of interest rates with the effective lower bound. *Journal of Money, Credit, and Banking*

[18] Jorda, Oscar, Singh, Sanjay R., & Alan M. Taylor (2020). Longer-Run economic consequences of pandemics. San Francisco Fed Working Paper 2020-09

[19] Kahn, James A., & Robert W. Rich (2007). Tracking the new economy: Using growth theory to detect changes in trend productivity. *Journal of Monetary Economics* 54: 1670-1701

[20] Kiley, Michael T. (2020). What can data tell us about the equilibrium real interest rate? *International Journal of Central Banking*, June 2020

[21] Knotek, Edward S. (2015). Difficulties forecasting wage growth. Federal Reserve Bank of Cleveland *Economic Trends*

[22] Koop, Gary (2003). Bayesian Econometrics. *Chichester: Wiley.*

[23] Koop, Gary, Leon-Gonzalez, Roberto & Rodney W. Strachan (2010). Dynamic probabilities of restrictions in state space models: An application to the Phillips curve. *Journal of Business and Economic Statistics* 28(3): 370-379

[24] Korobilis, Dimitris (2017). Quantile regression forecasts of inflation under model uncertainty. *International Journal of Forecasting* 33(1): 11-20

[25] Laubach, Thomas & John C. Williams (2003). Measuring the natural rate of interest. *The Review of Economics and Statistics*, 85(4): 1063-1070

[26] Lewis, Kurt F. & Francisco Vazquez-Grande (2019). Measuring the natural rate of interest: A note on transitory shocks. *Journal of Applied Econometrics* 34: 425-436

[27] Lubik, Thomas A. & Christian Matthes (2015). Calculating the natural rate of interest: A comparison of two alternative approaches. *Richmond Fed Economic Brief*

[28] Morley, James, & Benjamin Wong (2020). Estimating and accounting for the output gap with large Bayesian vector autoregressions. *Journal of Applied Econometrics*, 35: 1-18

[29] Pescatori, Andrea, & Jarkko Turunen (2016). Lower for longer: Neutral rate in the US. *IMF Economic Review*, 64(4): 708-731

[30] Primiceri, Giorgio E. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, 72, 821-852 doi:10.1111/j.1467-937x.2005.00353.x

[31] Roberts, John. (2001). Estimates of the productivity trend using time-varying parameter techniques. *Contributions to Macroeconomics*, 1(1):3

[32] Schorfheide, Frank, & Dongho Song. (2020). Real-Time Forecasting with a (Standard) Mixed-Frequency VAR During a Pandemic. Philadelphia Fed Working Paper 20-26.

[33] Stock, James H., & Mark W. Watson (2007). Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking*, 39: 3-33

[34] Stock, James H., & Mark W. Watson (2016). Core inflation and trend inflation. *Review of Economics and Statistics*, 98(4): 770-784

[35] Tallman, Ellis W., & Saeed Zaman (2020). Combining survey long-run forecasts and nowcasts with BVAR forecasts using relative entropy. *International Journal of Forecasting*, 36(2): 373-398

[36] Tasci, Murat (2012).The ins and outs of unemployment in the long run: Unemployment flows and the natural rate. Federal Reserve Bank of Cleveland Working Paper

[37] Tasci, Murat (2019). Challenges with estimating u star in real time. Federal Reserve Bank of Cleveland *Economic Commentary*, 2019-18

[38] Wright, Jonathan (2013). Evaluating real-time VAR forecasts with an informative democratic prior. *Journal of Applied Econometrics*, 28: 762-776 doi:10.1002/jae.2268

[39] Wright, Jonathan (2019). Some observations on forecasting and policy. *International Journal of Forecasting*, 35(3): 1186-1192