Unpacking trend inflation: Evidence from a factor correlated unobserved components model of sticky and flexible prices — Online appendix

Mengheng Li*

Ivan Mendieta-Muñoz†

May 11, 2026

This online appendix accompanies: Li, Mengheng, and Mendieta-Muñoz, Ivan. (2026). "*Unpacking trend inflation: Evidence from a factor correlated unobserved components model of sticky and flexible prices*". This appendix provides: (i) formal proofs of the identifiability of the factor correlated unobserved components (FCUC) model; (ii) relevant details of the Bayesian estimation procedure for the FCUC model, introduced in section 2 of the main text; (iii) further details on the relevance of the FCUC model by highlighting its flexibility with respect to other relevant models, extending section 3 of the main text; (iv) additional robustness checks that complement the discussions in section 5 of the main text; and (v) the relevant mixing statistics of the Markov chain that fits the FCUC model to our data set, which consists of sticky and flexible inflation rates and a measure of long-run inflation expectations. This appendix provides supporting materials and it is not meant for publication. Notations follow the main text, unless stated otherwise.

Keywords: trend inflation; sticky inflation; flexible inflation; stochastic volatility; dynamic factor model.

JEL Classification: C32; C53; E37.

*Senior Lecturer (Assistant Professor). University of Technology Sydney (UTS) Business School and Centre for Applied Macroeconomic Analysis (CAMA). Email: mengheng.li@uts.edu.au

†Associate Professor. Department of Economics, University of Utah. Email: ivan.mendietamunoz@utah.edu. ORCID: https://orcid.org/0000-0003-0018-5642

1 Identification

As discussed by Trenkler and Weber (2016) and Li and Mendieta-Muñoz (2022), there exists a set of non-trivial conditions for the identification of multivariate CUC models. In order to understand how the FCUC model is identified, we first highlight that a local level model (Shephard, 2015; Li and Koopman, 2021) leads to an integrated moving average of order 1, IMA(1), reduced-form representation.[1] The latter includes two parameters: the MA coefficient and the innovation variance. Therefore, a bivariate local level model admits a 2-dimensional vector IMA(1) representation with an MA coefficient matrix and a covariance matrix, yielding 7 reduced-form parameters. A constant bivariate FCUC model contains 6 parameters: a factor loading, two correlation coefficients, and three innovation variances, thus satisfying the order condition with an overidentified system.

We proceed to show the identification of the FCUC model more formally. It suffices to consider $\boldsymbol{y}_t = (\pi_t^S, \pi_t^F)'$ and a constant-parameter model in which we drop the subscript t for the factor loading and the correlation terms, and replace the variances of η_t^S, η_t^F and η_t^C by σ_S^2, σ_F^2 and σ_C^2, respectively. Let $\boldsymbol{\Gamma}_k = \mathrm{Cov}(\Delta\boldsymbol{y}_t, \Delta\boldsymbol{y}_{t-k})$ define the k-th order autocovariance. We introduce two assumptions.

Assumption 1. *(Non-degeneracy.) The innovation variance of the factor inflation* $\mathrm{Var}(\Delta\tau_t^C) = \sigma_\Delta^2$ *is strictly positive, or*

$$\sigma_\Delta^2 = \gamma^2\sigma_S^2 + \delta^2\sigma_F^2 + \sigma_C^2 > 0,$$

and the first-order autocovariance $\boldsymbol{\Gamma}_1$ *has full rank.*

Assumption 2. *(Spectral identifiability at low frequency.) The constant FCUC model does not reduce to a local level model, or*

$$a\gamma + b\delta \neq 0, \text{ for all } a, b \in \mathbb{R} \text{ satisfying } ab \neq 0.$$

Identification is given in the following proposition:

Proposition 1. *(i) Under Assumptions 1 and 2, the differenced process* $\Delta\boldsymbol{y}_t = (\Delta\pi_t^S, \Delta\pi_t^F)'$ *admits a 2-dimensional IMA(1) representation with positive-definite innovation covariance. (ii) The vector of structural parameters* $(\lambda, \gamma, \delta, \sigma_S^2, \sigma_F^2, \sigma_C^2)$ *is locally identified from the reduced-form*

[1] A local level model is an unobserved components model with a random walk trend and a white noise cycle. In other words, it is the UCSV model in (13) but with constant innovation variances.

second moments $(\boldsymbol{\Gamma}_0, \boldsymbol{\Gamma}_1)$.

Proof of Proposition 1 (i). Differencing the constant FCUC model gives

$$\Delta\pi_t^S = (\gamma+1)\eta_t^S + \delta\eta_t^F + \eta_t^C - \eta_{t-1}^S,$$
$$\Delta\pi_t^F = \lambda\gamma\eta_t^S + (\lambda\delta+1)\eta_t^F + \lambda\eta_t^C - \eta_{t-1}^F.$$

We define $\boldsymbol{\xi}_t = (\eta_t^S, \eta_t^F, \eta_t^C)'$. Then

$$\Delta\boldsymbol{y}_t = \boldsymbol{B}\boldsymbol{\xi}_t + \boldsymbol{C}\boldsymbol{\xi}_{t-1}, \quad \boldsymbol{B} = \begin{bmatrix} \gamma+1 & \delta & 1 \\ \lambda\gamma & \lambda\delta+1 & \lambda \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}.$$

The shocks satisfy the condition that $\boldsymbol{\xi}_t \overset{\text{i.i.d}}{\sim} (\mathbf{0}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = \text{diag}(\sigma_S^2, \sigma_F^2, \sigma_C^2)$. It follows that

$$\boldsymbol{\Gamma}_0 = \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}' + \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}',$$
$$\boldsymbol{\Gamma}_1 = \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{B}'.$$

Let us define the innovations $\boldsymbol{\varepsilon}_t = \Delta\boldsymbol{y}_t - \boldsymbol{C}\boldsymbol{\xi}_{t-1} = \boldsymbol{B}\boldsymbol{\xi}_t$. Since $\boldsymbol{B}$ has full row rank by Assumption 1, $\boldsymbol{S} = \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}' \succ 0$, where $\boldsymbol{X} \succ 0$ means that the smallest eigenvalue of $\boldsymbol{X}$ is larger than 0, and $\boldsymbol{\varepsilon}_t$ is non-degenerate. Rewriting $\boldsymbol{\xi}_{t-1} = \boldsymbol{\Sigma}\boldsymbol{B}'\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_{t-1}$ gives

$$\Delta\boldsymbol{y}_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\Theta}\varepsilon_{t-1}, \quad \Theta = \boldsymbol{\Gamma}_1(\boldsymbol{\Gamma}_0 - \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}')^{-1}, \quad \boldsymbol{\varepsilon}_t \overset{\text{i.i.d}}{\sim} N(\mathbf{0}, \boldsymbol{S}).$$

Hence $\Delta\boldsymbol{y}_t$ is a vector MA(1) with MA coefficient matrix $\boldsymbol{\Theta}$ and innovation covariance $\boldsymbol{S} \succ 0$. □

Proof of Proposition 1 (ii). Denote the parameter vector $\boldsymbol{\theta} = (\lambda, \gamma, \delta, \sigma_S^2, \sigma_F^2, \sigma_C^2)'$. Notice that the autocovariances are functions of $\boldsymbol{\theta}$, that is, we have moment mappings $\boldsymbol{\Gamma}_0(\boldsymbol{\theta})$ and $\boldsymbol{\Gamma}_1(\boldsymbol{\theta})$. Let $\boldsymbol{X}(i,j)$ denote the (i,j)-th element of $\boldsymbol{X}$. It can be shown that

$$\boldsymbol{\Gamma}_1(1,1) = -\sigma_S^2(\gamma+1), \qquad \boldsymbol{\Gamma}_1(1,2) = -\sigma_S^2\lambda\gamma,$$
$$\boldsymbol{\Gamma}_1(2,1) = -\sigma_F^2\delta, \qquad \boldsymbol{\Gamma}_1(2,2) = -\sigma_F^2(\lambda\delta+1).$$

To show the exact mapping, and thus local identification, we proceed in three steps. First, we

see that the row-wise ratios given by

$$r_S = \frac{\boldsymbol{\Gamma}_1(1,2)}{\boldsymbol{\Gamma}_1(1,1)} = \frac{\lambda\gamma}{\gamma+1}, \qquad r_F = \frac{\boldsymbol{\Gamma}_1(2,2)}{\boldsymbol{\Gamma}_1(2,1)} = \frac{\lambda\delta+1}{\delta},$$

identify $(\lambda, \gamma, \delta)$ provided that neither γ nor δ is zero (Assumption 2), while σ_S^2 and σ_F^2 follow directly once $(\lambda, \gamma, \delta)$ are known.

Second, with $(\lambda, \gamma, \delta, \sigma_S^2, \sigma_F^2)$ fixed, it is instrumental to notice that $\boldsymbol{\Gamma}_0(\boldsymbol{\theta})$ is affine in σ_C^2:

$$\boldsymbol{\Gamma}_0(\theta) = \boldsymbol{M}(\lambda, \gamma, \delta, \sigma_S^2, \sigma_F^2) + \begin{bmatrix} 1 & \lambda \\ \lambda & \lambda^2 \end{bmatrix} \sigma_C^2,$$

where $\boldsymbol{M}(\cdot)$ collects all other terms. Since the matrix multiplying σ_C^2 is known and nonzero, σ_C^2 is identified from any element of $\boldsymbol{\Gamma}_0(\boldsymbol{\theta})$.

Third, given that the mapping $\boldsymbol{\theta} \mapsto (\operatorname{vec} \boldsymbol{\Gamma}_0, \operatorname{vec} \boldsymbol{\Gamma}_1)$ is continuously differentiable, we explicitly compute the Jacobian and show it has full column rank at any generic parameter vector $\boldsymbol{\theta}$ satisfying Assumption (2). Define the six scalar moments used for identification:

$$\boldsymbol{\Psi}(\boldsymbol{\theta}) = (\boldsymbol{\Gamma}_1(1,1), \boldsymbol{\Gamma}_1(1,2), \boldsymbol{\Gamma}_1(2,1), \boldsymbol{\Gamma}_1(2,2), \boldsymbol{\Gamma}_0(1,2), \boldsymbol{u}'\boldsymbol{\Gamma}_0\boldsymbol{u})',$$

where $\boldsymbol{u} = (1, \lambda)$ isolates the linear term in σ_C^2. The partial derivatives, or the 6×6 Jacobian $\boldsymbol{J}(\boldsymbol{\theta}) = \partial\boldsymbol{\Psi}(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$, can be written in block-triangular form as

$$\boldsymbol{J}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{A}(\boldsymbol{\theta}) & 0 \\ \star & (1+\lambda^2)^2 \end{bmatrix}, \quad \boldsymbol{A}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{S}(\boldsymbol{\theta}) & \mathbf{0} & 0 \\ \mathbf{0} & \boldsymbol{F}(\boldsymbol{\theta}) & 0 \\ \star & \star & \frac{\partial \boldsymbol{u}'\boldsymbol{\Gamma}_0\boldsymbol{u}}{\partial\lambda} \end{bmatrix}.$$

The top-left 5×5 block $\boldsymbol{A}(\boldsymbol{\theta})$ in the left has rows ordered as $(\boldsymbol{\Gamma}_1(1,1), \boldsymbol{\Gamma}_1(1,1), \boldsymbol{\Gamma}_1(2,1), \boldsymbol{\Gamma}_1(2,2), \boldsymbol{\Gamma}_0(1,2))$ and columns as $(\gamma, \sigma_S^2, \delta, \sigma_F^2, \lambda)$ with

$$\boldsymbol{S}(\boldsymbol{\theta}) = \begin{bmatrix} -\sigma_S^2 & -(\gamma+1) \\ -\sigma_S^2\lambda & -\lambda\gamma \end{bmatrix}, \quad \boldsymbol{F}(\boldsymbol{\theta}) = \begin{bmatrix} -\sigma_F^2 & -\delta \\ -\sigma_F^2\lambda & -(\lambda\delta+1) \end{bmatrix}.$$

It is easy to show that

$$\det \boldsymbol{S}(\boldsymbol{\theta}) = -\sigma_S^2 \lambda, \quad \det \boldsymbol{F}(\boldsymbol{\theta}) = \sigma_F^2, \quad \frac{\partial \boldsymbol{u}' \boldsymbol{\Gamma}_0 \boldsymbol{u}}{\partial \lambda} = (\lambda+1)\lambda + \delta^2 \sigma_F^2 + \sigma_C^2.$$

Thus, the top-left 4×4 block $\mathrm{diag}(\boldsymbol{S}, \boldsymbol{F})$ is invertible whenever $\sigma_S^2 > 0$, $\sigma_F^2 > 0$, and $\lambda \neq 0$. Therefore, we have invertibility of the triangular Jacobian, namely,

$$\det \boldsymbol{J}(\boldsymbol{\theta}) = -\lambda \sigma_S^2 \sigma_F^2 \left[(\lambda+1)\lambda + \delta^2 \sigma_F^2 + \sigma_C^2\right] (1+\lambda^2)^2 \neq 0$$

and the Jacobian $\boldsymbol{J}(\boldsymbol{\theta})$ has full rank. □

Proposition 1 establishes identification of the constant-parameter FCUC model in the time domain, relying on the mapping from finite-lag autocovariances $(\boldsymbol{\Gamma}_0, \boldsymbol{\Gamma}_1)$ to the structural parameters. This ensures that the model is locally identified using observable second moments. In what follows, we provide a complementing lemma from a frequency-domain perspective.

Lemma 1. *Let $f_{SS}(\omega)$ and $f_{FF}(\omega)$ denote the spectral densities of (π_t^S, π_t^F) and $f_{SF}(\omega)$ their cross-spectrum. Under Assumption 1 and 2, we have*

$$\lim_{\omega \to 0} \omega^2 f_{SS}(\omega) = \sigma_\Delta^2,$$
$$\lim_{\omega \to 0} \omega^2 f_{FF}(\omega) = \lambda^2 \sigma_\Delta^2,$$
$$\lim_{\omega \to 0} \omega^2 f_{SF}(\omega) = \lambda \sigma_\Delta^2.$$

Thus, λ and σ_Δ^2 are jointly identified from the low-frequency behavior of (f_{SS}, f_{FF}, f_{SF}).

Proof of Lemma 1. $\Delta\tau_t^C = \gamma\eta_t^S + \delta\eta_t^F + \eta_t^C$ has variance σ_Δ^2 and spectrum

$$f_{CC}(\omega) = \frac{\sigma_\Delta^2}{|1 - e^{-i\omega}|^2} = \frac{\sigma_\Delta^2}{2(1-\cos\omega)} \Longrightarrow \lim_{\omega \to 0} f_{CC}(\omega) = \frac{\sigma_\Delta^2}{\omega^2}.$$

Since $\pi_t^S = \tau_t^C + \eta_t^S$ and $\pi_t^F = \lambda\tau_t^C + \eta_t^F$ with η_t^S, η_t^F orthogonal to τ_t^C, then

$$f_{SS}(\omega) = f_{CC}(\omega) + \sigma_S^2,$$
$$f_{FF}(\omega) = \lambda^2 f_{CC}(\omega) + \sigma_F^2,$$
$$f_{SF}(\omega) = \lambda f_{CC}(\omega).$$

Multiplying by ω^2 and taking limits gives the stated results, and constructing the relevant ratios we have

$$\lim_{\omega\to 0} \frac{f_{FF}(\omega)}{f_{SS}(\omega)} = \lambda^2, \quad \lim_{\omega\to 0} \frac{f_{SF}(\omega)}{f_{SS}(\omega)} = \lambda,$$

which identify λ from the slope of the low-frequency cross-spectrum and σ_Δ^2 from the common scaling constant. □

Lemma 1 shows that the common-trend loading λ and the innovation variance σ_Δ^2 are identified from the behavior of spectral density near zero frequency. This spectral argument demonstrates that identification is not merely a coincidence derived from the chosen IMA(1) representation, but a structural property derived from the τ_t^C's $1/\omega^2$ pole at low frequencies. We believe that this dual time- and frequency-domain identification provides robustness and clarifies that the long-run component of inflation dominates the low-frequency spectrum, linking our FCUC structure to both the univariate Beveridge–Nelson (Morley et al., 2003; Li and Mendieta-Muñoz, 2022) and Stock–Watson decomposition (Stock and Watson, 2007, 2020; Li and Koopman, 2021) frameworks in the unobserved components literature.

Finally, we point out that, while the above propositions define a constant-parameter model, the proposed time-varying FCUC model with random walk states can be treated as locally constant over short windows, so the same local identification applies. Specifically, a time-varying parameter process with random walk dynamics, such as $\theta_t = \theta_{t-1} + \nu_t$ with innovation variance σ_ν^2, can be treated as *locally constant* over short horizons due to $E(\theta_{t+k} - \theta_t)^2 = k\sigma_\nu^2 \sim \mathcal{O}(k)$, or $\theta_{t+k} - \theta_t \sim \mathcal{O}_p(\sqrt{k})$. As $k/T \to 0$, the scaled path $\theta(u) = \theta_{\lfloor Tu \rfloor}$, where $u = t/T \in [0, 1]$, satisfies the local stationarity condition of Dahlhaus (1997):

$$\sup_{|h|<H/T} \left|\theta(u+h) - \theta(u)\right| \xrightarrow{p.} 0,$$

for any fixed H. Therefore, over any short horizon with at least $k = o(T)$ observations, θ_t can be treated as constant to first order, rendering the process a constant-parameter system in local windows with sufficient moments to identify model parameters.

2 Sampling procedure

The MCMC sampler iterates over the following three major blocks:

- Sample model parameters $\boldsymbol{\theta} = (\sigma_g, \phi)$, $g \in \{h^S, h^F, h^C, E, \alpha, \beta, \lambda, \gamma, \delta\}$, conditional on latent states and initializations, or $(\boldsymbol{\theta}|\boldsymbol{\pi}, \boldsymbol{f}, f_0)$, where $f = \{h^S, h^F, h^C, \tau^C, \alpha, \beta, \lambda, \gamma, \delta\}$;
- Sample latent states conditional on model parameters and initializations, or $(\boldsymbol{f}|\boldsymbol{\pi}, f_0, \boldsymbol{\theta})$;
- Sample initializations conditional on model parameters and latent states, or $(f_0|\boldsymbol{\pi}, \boldsymbol{f}, \boldsymbol{\theta})$.

In the description of the sampling procedure above, $\boldsymbol{\pi} = \left((\boldsymbol{\pi}^S)', (\boldsymbol{\pi}^F)', (\boldsymbol{\pi}^E)'\right)'$ collects the data, where bold fonts indicate the vector of a time series variable, *e.g.*, $\boldsymbol{\pi}^S = (\pi_1^S, ..., \pi_T^S)'$.

2.1 Sample model parameters

Sample variances

Let the gamma prior in section 2.3 of the main text be denoted by $\sigma_g^2 \sim \Gamma(a, b)$ with prior

$$p_0(\sigma_g^2) \propto (\sigma_g^2)^{a-1} \exp(-b\sigma_g^2). \tag{1}$$

For σ_f^2, we have

$$\boldsymbol{H}\boldsymbol{f} = \boldsymbol{e}_1 f_0 + \boldsymbol{\epsilon}^f, \quad \boldsymbol{\epsilon}^f \sim N(\boldsymbol{0}, \sigma_f^2 \boldsymbol{I}_T),$$

where $\boldsymbol{I}_T$ is a $T \times T$ identity matrix, $\boldsymbol{e}_s$ is the s-th unit vector in $\mathbb{R}^T$, and $\boldsymbol{H}$ is a $T \times T$ matrix that first differences $\boldsymbol{f}$, or

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

This yields the conditional likelihood

$$p(\boldsymbol{f}|\sigma_f^2, f_0) \propto (\sigma_f^2)^{-\frac{T}{2}} \exp\left(-\frac{1}{2\sigma_f^2}\tilde{\boldsymbol{f}}'\boldsymbol{H}'\boldsymbol{H}\tilde{\boldsymbol{f}}\right), \tag{2}$$

where $\tilde{\boldsymbol{f}} = \boldsymbol{f} - \boldsymbol{H}^{-1}\boldsymbol{e}_1 f_0 = \boldsymbol{f} - \mathbf{1}f_0$, and $\mathbf{1}$ is a vector of ones. Combining the likelihood (2) with the prior (1), we can derive

$$p(\sigma_f^2|\boldsymbol{f}, f_0) \propto (\sigma_f^2)^{-\frac{T}{2}-a} \exp\left(-\frac{\tilde{\boldsymbol{f}}'\boldsymbol{H}'\boldsymbol{H}\tilde{\boldsymbol{f}}}{2\sigma_f^2} + 2b\sigma_f^2\right).$$

This is the kernel of the generalized inverse Gaussian (GIG) distribution. We follow Eisenstat et al. (2016) and sample the variance parameters as follows:

$$\sigma_f^2|\boldsymbol{f}, f_0 \sim GIG\left(a - \frac{T}{2}, 2b, \tilde{\boldsymbol{f}}'\boldsymbol{H}'\boldsymbol{H}\tilde{\boldsymbol{f}}\right),$$

for $f = \{h^S, h^F, h^C, \alpha, \beta, \lambda, \gamma, \delta\}$.

To sample σ_E^2, we use equation (4) in the main text. We write the latter in matrix form:

$$\begin{aligned} \tilde{\boldsymbol{\pi}}^E &= \boldsymbol{H}_\phi \boldsymbol{\epsilon}^E, \quad \boldsymbol{\epsilon}^E \sim N(\mathbf{0}, \sigma_E^2 \boldsymbol{I}_T), \\ \tilde{\boldsymbol{\pi}}^E &= \boldsymbol{\pi}^E - \boldsymbol{\alpha} - \boldsymbol{\beta} \odot \boldsymbol{\pi}^C, \end{aligned} \tag{3}$$

where $\odot$ indicates element-by-element product, and

$$\boldsymbol{H}_\phi = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\phi & 1 & 0 & \cdots & 0 \\ 0 & -\phi & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -\phi & 1 \end{bmatrix}.$$

Following (2), the conditional likelihood in this case is

$$p(\boldsymbol{\pi}^E|\sigma_E^2, \boldsymbol{\alpha}, \boldsymbol{\beta}, \phi) \propto (\sigma_E^2)^{-\frac{T}{2}} \exp\left(-\frac{1}{2\sigma_E^2}(\tilde{\boldsymbol{\pi}}^E)'(\boldsymbol{H}_\phi \boldsymbol{H}_\phi')^{-1}\tilde{\boldsymbol{\pi}}^E\right).$$

Combing it with the prior (1) gives the conditional posterior

$$p(\sigma_E^2|\boldsymbol{\pi}^E, \boldsymbol{\alpha}, \boldsymbol{\beta}, \phi) \propto (\sigma_E^2)^{-\frac{T}{2}-a} \exp\left(-\frac{(\tilde{\boldsymbol{\pi}}^E)'(\boldsymbol{H}_\phi \boldsymbol{H}_\phi')^{-1}\tilde{\boldsymbol{\pi}}^E}{2\sigma_E^2} + 2b\sigma_E^2\right).$$

Thus, we sample from the following GIG distribution:

$$\sigma_E^2|\boldsymbol{\pi}^E, \boldsymbol{\alpha}, \boldsymbol{\beta}, \phi \sim GIG\left(a - \frac{T}{2}, 2b, (\tilde{\boldsymbol{\pi}}^E)'(\boldsymbol{H}_\phi \boldsymbol{H}_\phi')^{-1}\tilde{\boldsymbol{\pi}}^E\right).$$

Sample MA coefficient

Using (3), we have the conditional likelihood

$$p(\boldsymbol{\pi}^E|\sigma_E^2, \boldsymbol{\alpha}, \boldsymbol{\beta}, \phi) \propto \exp\left(-\frac{1}{2\sigma_E^2}(\tilde{\boldsymbol{\pi}}^E)'(\boldsymbol{H}_\phi \boldsymbol{H}_\phi')^{-1}\tilde{\boldsymbol{\pi}}^E\right),$$

such that the gamma prior (1) is not conjugate. We implement a standard Metropolis-Hastings step. Let $\hat{\phi}$ and $\widehat{\text{Var}}(\hat{\phi})$ denote the mean and variance, respectively, of the Laplace approximation (denoted by $\sim^*$, meaning approximate distribution up to location and scale),

$$\phi|\boldsymbol{\pi}^E, \sigma_E^2, \boldsymbol{\alpha}, \boldsymbol{\beta} \sim^* N\big(\hat{\phi}, \widehat{\text{Var}}(\hat{\phi})\big), \tag{4}$$

where the parameters are defined by:

$$\hat{\phi} = \arg\max_\phi \left[\log p(\boldsymbol{\pi}^E|\sigma_E^2, \boldsymbol{\alpha}, \boldsymbol{\beta}, \phi) + \log p_0(\phi)\right],$$
$$\widehat{\text{Var}}(\hat{\phi}) = -\left(\frac{\partial^2}{(\partial\phi)^2}\left[\log p(\boldsymbol{\pi}^E|\sigma_E^2, \boldsymbol{\alpha}, \boldsymbol{\beta}, \phi) + \log p_0(\phi)\right]\Big|_{\phi=\hat{\phi}}\right)^{-1},$$

where $p_0(\phi)$ is the truncated normal prior for the MA coefficient defined in section 2.3 of the main text.

We generate a draw ϕ^{new} from the approximate distribution (4). Let ϕ^{old} denote the current draw in the Markov chain and let $N(x; u, v)$ denote the density ordinate of a normal distribution with mean u and variance v evaluated at x. The draw is accepted with probability

$$\min\left[\frac{p(\boldsymbol{\pi}^E|\sigma_E^2, \boldsymbol{\alpha}, \boldsymbol{\beta}, \phi^{\text{new}})p_0(\phi^{\text{new}})N(\phi^{\text{old}}; \hat{\phi}, \widehat{\text{Var}}(\hat{\phi}))}{p(\boldsymbol{\pi}^E|\sigma_E^2, \boldsymbol{\alpha}, \boldsymbol{\beta}, \phi^{\text{old}})p_0(\phi^{\text{old}})N(\phi^{\text{new}}; \hat{\phi}, \widehat{\text{Var}}(\hat{\phi}))}, 1\right],$$

otherwise $\phi^{\text{new}} = \phi^{\text{old}}$.

2.2 Sample latent states

Sample common trend $\boldsymbol{\tau}^C$

In brief, to sample $\boldsymbol{\tau}^C$ is to sample $\boldsymbol{\epsilon}^S$. First, we notice that

$$\pi_t^f = \lambda_t(\pi_t^S - \epsilon_t^S) + \epsilon_t^F. \tag{5}$$

Similarly, we rewrite the equation for the inflation expectations, or equation (4) in the main text:

$$\pi_t^E - \alpha_t = \beta_t(\pi_t^S - \epsilon_t^S) + \epsilon_t^E + \theta\epsilon_t^E,$$

which yields

$$\underbrace{\hat{\pi}_t^E}_{\pi_t - \alpha_t - \beta_t\pi_t^S} = -\beta_t\epsilon_t^S + \epsilon_t^E + \theta\epsilon_t^E. \tag{6}$$

Using (5) to replace ϵ_t^F in the transition dynamics of the common trend, or equation (5) in the main text, we have

$$\begin{aligned}\tau^C &= \tau_{t-1}^C + \gamma_t\epsilon_t^S + \delta_t(\pi_t^F - \lambda_t\pi_t^S + \lambda_t\epsilon_t^S) + \eta_t^C \\ &= \tau_{t-1}^C + \underbrace{\delta_t(\pi_t^F - \lambda_t\pi_t^S)}_{z_t} + \underbrace{(\gamma_t + \delta_t\lambda_t)\epsilon_t^S}_{x_t} + \eta_t^C.\end{aligned} \tag{7}$$

In matrix form, equation (7) becomes

$$\boldsymbol{H}\boldsymbol{\tau}^C = \boldsymbol{e}_1\tau_0^C + \boldsymbol{z} + \boldsymbol{x} \odot \boldsymbol{\epsilon}^S + \boldsymbol{\eta}^C.$$

Combining the equation above with $\boldsymbol{H}\boldsymbol{\pi}^S = \boldsymbol{H}\boldsymbol{\tau}^S + \boldsymbol{H}\boldsymbol{\epsilon}^S$ (that is, pre-multiplying $\boldsymbol{H}$ to equation (2) in the main text) yields

$$\boldsymbol{H}\boldsymbol{\tau}^S = \boldsymbol{e}_1\tau_0^C + \boldsymbol{z} + (\text{diag}(\boldsymbol{x}) + \boldsymbol{H})\boldsymbol{\epsilon}^S + \boldsymbol{\eta}^C, \tag{8}$$

where $\text{diag}(\boldsymbol{y})$ denotes a diagonal matrix of the same dimension as the vector $\boldsymbol{y}$ whose elements

also form the diagonal elements of the matrix. Casting equation (3) in matrix form yields

$$\hat{\boldsymbol{\pi}}^E = \text{diag}(-\boldsymbol{\beta})\boldsymbol{\epsilon}^S + \boldsymbol{H}_\phi \boldsymbol{\epsilon}^E. \tag{9}$$

Putting (8) and (9) together yields

$$\underbrace{\begin{pmatrix} \boldsymbol{H}\boldsymbol{\pi}^S - \boldsymbol{e}_1\tau_0^C - \boldsymbol{z} \\ \boldsymbol{H}_\phi^{-1}\hat{\boldsymbol{\pi}}^E \end{pmatrix}}_{\boldsymbol{\pi}^*} = \underbrace{\begin{bmatrix} \text{diag}(\boldsymbol{x}) + \boldsymbol{H} \\ \boldsymbol{H}_\phi^{-1}\text{diag}(-\boldsymbol{\beta}) \end{bmatrix}}_{\boldsymbol{X}} \boldsymbol{\epsilon}^S + \underbrace{\begin{pmatrix} \boldsymbol{\eta}^C \\ \boldsymbol{\epsilon}^E \end{pmatrix}}_{\boldsymbol{\epsilon}^*}, \quad \boldsymbol{\epsilon}^* | \boldsymbol{h}^C, \sigma_E^2 \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}),$$

where

$$\boldsymbol{\Sigma} = \begin{bmatrix} \text{diag}(\exp^\bullet(\boldsymbol{h}^C)) & \boldsymbol{O} \\ \boldsymbol{O} & \sigma_E^2 \boldsymbol{I}_T \end{bmatrix},$$

$\exp^\bullet(.)$ denotes element-wise exponential and $\boldsymbol{O}$ is a $T \times T$ matrix of zeros. Using the Bayesian regression lemma (where $\boldsymbol{\pi}^*$ is the vector of responses and $\boldsymbol{X}$ is the matrix of regressors), we have the conditional posterior

$$\boldsymbol{\epsilon}^S | \boldsymbol{\pi}, \boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\gamma}, \boldsymbol{h}^C, \boldsymbol{h}^S, \phi, \sigma_E^2 \sim N(\boldsymbol{\mu}_C, \boldsymbol{P}_C^{-1}),$$

where

$$\begin{aligned} \boldsymbol{P}_C &= \exp^\bullet(-\boldsymbol{h}^S) + \boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}, \\ \boldsymbol{\mu}_C &= \boldsymbol{P}_C^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\pi}^*. \end{aligned}$$

Since all the matrices involved in the derivation above are tri-diagonal, we use the fast precision sampler of Chan and Jeliazkov (2009) to sample from the above multivariate Gaussian conditional posterior, utilizing efficient routines of sparse matrix inversion and Cholesky decomposition that are available in Matlab® 2022b.

Sample loading $\boldsymbol{\lambda}$

It is instrumental to notice that we cannot sample λ_t directly from the flexible inflation equation (3) in the main text, conditional on τ_t^C. This is so because τ_t^C and the innovation term (flexible

cycle) η_t^F are correlated, or $E(\boldsymbol{\tau}^C(\boldsymbol{\eta}^F)'|\boldsymbol{\gamma},\boldsymbol{\delta}) \neq \boldsymbol{O}$. In other words, there exists an endogeneity issue that needs to be addressed for valid posterior inference.

We write equation (3) in the main text in matrix form:

$$\begin{aligned}\boldsymbol{\pi}^F &= \boldsymbol{\tau}^C \odot \boldsymbol{\lambda} + \boldsymbol{\epsilon}^F \\ &= \boldsymbol{\nu} + \text{diag}(\boldsymbol{\tau}^C)\boldsymbol{\lambda} + \tilde{\boldsymbol{\eta}}^F, \quad \tilde{\boldsymbol{\eta}}^F|\boldsymbol{h}^F, \boldsymbol{h}^C m\boldsymbol{\delta} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}^F),\end{aligned} \tag{10}$$

where $\boldsymbol{\nu}$ and $\boldsymbol{\Sigma}^F$ is the conditional mean and covariance of $\boldsymbol{\epsilon}^F$ (conditional on $\boldsymbol{\tau}^C$). To derive this, we write the transition equation of the common trend in matrix form:

$$\boldsymbol{H}\boldsymbol{\tau}^C = \boldsymbol{e}_1\tau_0^C + \boldsymbol{\gamma} \odot \boldsymbol{\epsilon}^S + \text{diag}(\boldsymbol{\delta})\boldsymbol{\epsilon}^F + \boldsymbol{\eta}^C, \quad \boldsymbol{\eta}^C|\boldsymbol{h}^C \sim N(\boldsymbol{0}, \text{diag}(\exp^\bullet(\boldsymbol{h}^C))), \tag{11}$$

which is a regression on $\boldsymbol{\eta}^F$. Given the conditional prior $\boldsymbol{\eta}^F|\boldsymbol{h}^F \sim N(\boldsymbol{0}, \text{diag}(\exp^\bullet(\boldsymbol{h}^F)))$, we can compute the conditional mean and covariance matrix using the Bayesian regression lemma as follows:

$$\begin{aligned}\left(\boldsymbol{\Sigma}^F\right)^{-1} &= \text{diag}(\exp^\bullet(-\boldsymbol{h}^F)) + \text{diag}(\boldsymbol{\delta})'\text{diag}(\exp^\bullet(-\boldsymbol{h}^C))\text{diag}(\boldsymbol{\delta}) \\ \boldsymbol{\nu} &= \boldsymbol{\Sigma}^F\text{diag}(\boldsymbol{\delta})'\text{diag}(\exp^\bullet(-\boldsymbol{h}^C))(\boldsymbol{H}\boldsymbol{\tau}^C - \boldsymbol{e}_1\tau_0^C - \boldsymbol{\gamma} \odot \boldsymbol{\epsilon}^S).\end{aligned}$$

With the equation above determined, we go back to equation (10), which is a regression on $\boldsymbol{\lambda}$. Based on the conditional prior $\boldsymbol{\lambda}|\lambda_0, \sigma_\lambda^2 \sim N(\boldsymbol{1}\lambda_0, \sigma_\lambda^2(\boldsymbol{H}'\boldsymbol{H})^{-1})$ derived from $\boldsymbol{H}\boldsymbol{\lambda} = \boldsymbol{e}_1\lambda_0 + \boldsymbol{\epsilon}^\lambda$ with $\boldsymbol{\epsilon}^\lambda|\sigma_\lambda^2 \sim N(\boldsymbol{0}, \sigma_\lambda^2\boldsymbol{I}_T)$, using the Bayesian regression lemma we obtain

$$\boldsymbol{\lambda}|\boldsymbol{\pi}^F, \boldsymbol{\pi}^S, \boldsymbol{\tau}^C, \boldsymbol{\gamma}, \boldsymbol{\delta}, \boldsymbol{h}^C, \boldsymbol{h}^F, \sigma_\lambda^2, \lambda_0 \sim N(\boldsymbol{\mu}_\lambda, \boldsymbol{P}_\lambda^{-1}),$$

where

$$\begin{aligned}\boldsymbol{P}_\lambda &= \frac{1}{\sigma_\lambda^2}(\boldsymbol{H}'\boldsymbol{H}) + \text{diag}(\boldsymbol{\tau}^C)'(\boldsymbol{\Sigma}^F)^{-1}\text{diag}(\boldsymbol{\tau}^C), \\ \boldsymbol{\mu}_\lambda &= \boldsymbol{P}_\lambda^{-1}\left(\frac{\lambda_0}{\sigma_\lambda^2}\boldsymbol{H}'\boldsymbol{e}_1 + \text{diag}(\boldsymbol{\tau}^C)'(\boldsymbol{\Sigma}^F)^{-1}(\boldsymbol{\tau}^F - \boldsymbol{v})\right).\end{aligned}$$

Sample regression coefficients γ and δ

We rewrite (11) as a regression on $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$. It follows that

$$\underbrace{\boldsymbol{H}\boldsymbol{\tau}^C - \boldsymbol{e}_1\tau_0^C}_{\hat{\boldsymbol{\pi}}^C} = \underbrace{\text{diag}(((\boldsymbol{\epsilon}^S)', (\boldsymbol{\epsilon}^F)')')}_{\boldsymbol{X}_\epsilon} \begin{pmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\delta} \end{pmatrix} + \boldsymbol{\eta}^C, \quad \boldsymbol{\eta}^C | \boldsymbol{h}^C \sim N(\boldsymbol{0}, \text{diag}(\exp^\bullet(\boldsymbol{h}^C))).$$

Given the conditional prior

$$\begin{pmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\delta} \end{pmatrix} \Big| \gamma_0, \delta_0, \sigma_\gamma^2, \sigma_\delta^2 \sim N\left((\gamma_0, \delta_0)' \otimes \boldsymbol{1}, (\sigma_\gamma^2, \sigma_\delta^2)' \otimes (\boldsymbol{H}'\boldsymbol{H})^{-1}\right),$$

using the Bayesian regression lemma yields the following conditional posterior:

$$\begin{pmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\delta} \end{pmatrix} \Big| \boldsymbol{\pi}^S, \boldsymbol{\pi}^F, \boldsymbol{\tau}^C, \boldsymbol{\lambda}, \boldsymbol{h}^C, \sigma_\gamma^2, \sigma_\delta^2, \gamma_0, \delta_0 \sim N(\boldsymbol{\mu}_{\gamma,\delta}, \boldsymbol{P}_{\gamma,\delta}^{-1}),$$

where

$$\begin{aligned} \boldsymbol{P}_{\gamma,\delta} &= \left(\frac{1}{\sigma_\gamma^2}, \frac{1}{\sigma_\delta^2}\right)' \otimes (\boldsymbol{H}'\boldsymbol{H}) + \boldsymbol{X}_\epsilon' \text{diag}(\exp^\bullet(-\boldsymbol{h}^C))\boldsymbol{X}_\epsilon, \\ \boldsymbol{\mu}_{\gamma,\delta} &= \boldsymbol{P}_{\gamma,\delta}^{-1}\left(\left(\frac{\gamma_0}{\sigma_\gamma^2}, \frac{\delta_0}{\sigma_\delta^2}\right)' \otimes \boldsymbol{H}'\boldsymbol{e}_1 + \boldsymbol{X}_\epsilon' \text{diag}(\exp^\bullet(-\boldsymbol{h}^C))\hat{\boldsymbol{\pi}}^C\right). \end{aligned}$$

Sample regression coefficients α and β

We rewrite (3) as a regression on $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. It follows that

$$\boldsymbol{\pi}^E = \underbrace{\text{diag}((\boldsymbol{1}', (\boldsymbol{\tau}^C)')')}_{\boldsymbol{X}_E} \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} + \tilde{\epsilon}^E, \quad \tilde{\epsilon}^E | \phi, \sigma_E^2 \sim N(\boldsymbol{0}, \sigma_E^2(\boldsymbol{H}_\phi \boldsymbol{H}_\phi')).$$

Given the conditional prior

$$\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} \Big| \sigma_\alpha^2, \sigma_\beta^2, \alpha_0, \beta_0 \sim N\left((\alpha_0, \beta_0)' \otimes \boldsymbol{1}, (\sigma_\alpha^2, \sigma_\beta^2)' \otimes (\boldsymbol{H}'\boldsymbol{H})^{-1}\right),$$

we use the Bayesian regression lemma to construct the following conditional posterior:

$$\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} \Big| \boldsymbol{\pi}^E, \boldsymbol{\tau}^C, \phi, \boldsymbol{h}^C, \sigma_E^2, \sigma_\alpha^2, \sigma_\beta^2, \alpha_0, \beta_0 \sim N(\boldsymbol{\mu}_{\alpha,\beta}, \boldsymbol{P}_{\alpha,\beta}^{-1}),$$

where

$$\begin{aligned} \boldsymbol{P}_{\alpha,\beta} &= \left(\frac{1}{\sigma_\alpha^2}, \frac{1}{\sigma_\beta^2}\right)' \otimes (\boldsymbol{H}'\boldsymbol{H}) + \frac{1}{\sigma_E^2} \boldsymbol{X}'_E (\boldsymbol{H}_\phi \boldsymbol{H}'_\phi)^{-1} \boldsymbol{X}_E, \\ \boldsymbol{\mu}_{\alpha,\beta} &= \boldsymbol{P}_{\alpha,\beta}^{-1} \left(\left(\frac{\alpha_0}{\sigma_\alpha^2}, \frac{\beta_0}{\sigma_\beta^2}\right)' \otimes \boldsymbol{H}'\boldsymbol{e}_1 + \frac{1}{\sigma_E^2} \boldsymbol{X}'_E (\boldsymbol{H}_\phi \boldsymbol{H}'_\phi)^{-1} \hat{\boldsymbol{\pi}}^E \right). \end{aligned}$$

Sample stochastic volatility $\boldsymbol{h}^m$, $m \in \{S, F, C\}$

Given the aforementioned latent states sampled, we obtain $\boldsymbol{\eta}^m$, $m \in \{S, F, C\}$, which are sufficient statistics for the log volatility $\boldsymbol{h}^m$. Let $\boldsymbol{\kappa}^m = (\log[(\eta_1^m)^2], ..., \log[(\eta_T^m)^2])'$. We suppress the superscript for presentation. Then, equation (6) in the main text becomes

$$\boldsymbol{\kappa} = \boldsymbol{I}_T \boldsymbol{h} + \boldsymbol{\zeta}, \tag{12}$$

where $\boldsymbol{\zeta} = (\zeta_1, ..., \zeta_T)'$ with ζ_t being the logarithm of the square of standard Gaussian innovation terms. Specifically, ζ_t is log-χ^2 distributed with one degree of freedom. We follow Kim et al. (1998) and approximate the nonstandard distribution by using a 7-component Gaussian mixture via data augmentation:

$$\zeta_t \approx \sum_{i=1}^{7} \mathbb{1}_{\{q_t = i\}} N(\nu_i, v_i),$$

where ν_i and v_i are tabulated by Kim et al. (1998), and the prior probabilities of the mixing component $q_t \in \{1, ..., 7\}$ are also tabulated and given by

$$P(q_t = i) = q_i, \quad i = 1, ..., 7,$$

with $q_i > 0$ and $\sum_{i=1}^7 q_i = 1$. It can be verified that the first several moments of ζ_t are well approximated by the auxiliary representation. This means that we can modify (12) as follows:

$$\boldsymbol{\kappa} = \boldsymbol{I}_T \boldsymbol{h} + \boldsymbol{\epsilon}^\star, \quad \boldsymbol{\epsilon}^\star | \boldsymbol{q} \sim N(\boldsymbol{\nu}, \text{diag}(\boldsymbol{v})).$$

Given the conditional prior $\boldsymbol{h}|h_0, \sigma_h^2 \sim N(\mathbf{1}h_0, \sigma_h^2(\boldsymbol{H}'\boldsymbol{H})^{-1})$, we use the Bayesian regression lemma and derive the following conditional posterior:

$$\boldsymbol{h}|\boldsymbol{\eta}, h_0, \sigma_h^2, \boldsymbol{q} \sim N(\boldsymbol{\mu}_h, \boldsymbol{P}_h^{-1}),$$

where

$$\begin{aligned}\boldsymbol{P}_h &= \frac{1}{\sigma_h^2}\boldsymbol{H}'\boldsymbol{H} + \text{diag}(-\boldsymbol{v}),\\ \boldsymbol{\mu}_h &= \boldsymbol{P}_h^{-1}\left(\frac{h_0}{\sigma_h^2}\boldsymbol{H}'\boldsymbol{e}_1 + \text{diag}(-\boldsymbol{v})(\boldsymbol{\kappa} - \boldsymbol{v})\right).\end{aligned}$$

We then update the mixing components via a categorical distribution with the following conditional posterior:

$$P(q_t = i)|\kappa_t, h_t \propto \frac{q_i}{\sqrt{v_i}}\exp\left(-\frac{1}{2v_i}(\kappa_t - \nu_i - h_t)^2\right), \quad t = 1, ..., T.$$

2.3 Sample initializations

For $f \in \{h^S, h^F, h^C, \alpha, \beta, \lambda, \gamma, \delta\}$, we can write

$$\boldsymbol{H}\boldsymbol{f} = \mathbf{1}f_0 + \boldsymbol{\epsilon}^f, \quad \boldsymbol{\epsilon}^f|\sigma_f^2 \sim N(\mathbf{0}, \sigma_f^2\boldsymbol{I}_T).$$

Provided the diffuse prior shown in section 2.3 of the main text, or $f_0 \sim N(0, M)$ with $M = 10^6$, we have the conditional posterior given by

$$f_0|\boldsymbol{f}, \sigma_f^2 \sim N(\mu_{f_0}, \sigma_{f_0}^2),$$

where

$$\begin{aligned}\frac{1}{\sigma_{f_0}^2} &= \frac{1}{M} + \frac{1}{\sigma_f^2}\boldsymbol{e}_1'\boldsymbol{e}_1,\\ \mu_{f_0} &= \frac{\sigma_{f_0}^2}{\sigma_f^2}\boldsymbol{e}_1'\boldsymbol{H}\boldsymbol{f}.\end{aligned}$$

We can further simplify the expression above as follows:

$$f_0|f_1,\sigma_f^2 \sim N\left(\frac{f_1/\sigma_f^2}{1/M+1/\sigma_f^2},\left(\frac{1}{M}+\frac{1}{\sigma_f^2}\right)^{-1}\right).$$

A straightforward modification is made for the initialization of the common inflation trend, given by

$$\tau_0^C|\tau_1^C,\gamma_1,\delta_1,\epsilon_1^S,\epsilon_1^F,h_1^C \sim N\left(\frac{(\tau_1^C-\gamma_1\epsilon_1^S-\delta_1\epsilon_1^F)\exp(-h_1^C)}{1/M+\exp(-h_1^C)},\left(\frac{1}{M}+\exp(h_1^C)\right)^{-1}\right).$$

3 Further details on key stylized facts of U.S. inflation: sticky and flexible components

3.1 Are sticky, flexible, and aggregate inflation similar?

The seminal unobserved components model with stochastic volatility model (UCSV model) of Stock and Watson (2007) is defined by:

$$\begin{aligned}
\pi_t &= \tau_t+\eta_t, \quad \eta_t=\exp\left(\frac{h_t}{2}\right)\epsilon_t,\\
\tau_t &= \tau_{t-1}+\eta_t^\tau, \quad \eta_t^\tau=\exp\left(\frac{h_t^\tau}{2}\right)\epsilon_t^\tau,\\
f_t &= f_{t-1}+\sigma_f\epsilon_t^f, \quad f\in\{h,h^\tau\},
\end{aligned} \tag{13}$$

where τ_t and η_t represent the trend and cycle, respectively, while ϵ_t, ϵ_t^τ, ϵ_t^h, and $\epsilon_t^{h^\tau}$ are independent standard Gaussian terms.

Shephard (2015) and Li and Koopman (2021) showed that the UCSV model implies an MA forecasting function that dynamically discounts past observations, depending on the permanent (or trend) volatility $\exp(h_t^\tau/2)$ and the transitory (or cycle) volatility $\exp(h_t/2)$. This feature allows the model to attribute inflation swings during the 1970s-80s oil crises and the 2007-9 Global Financial Crisis (GFC) to permanent and transitory changes, respectively, while tracking the Great Moderation amid the Volcker-Greenspan-Bernanke regimes. In this sense, as documented by Shephard (2015), Lansing (2022) and Jørgensen and Lansing (2024), the UCSV model allows for a dynamic signal-to-noise ratio (SNR) that captures the

time-varying persistence of trend inflation.

We fit the UCSV model to π_t^S and π_t^F individually, with the respective superscripts added to τ_t and ϵ_t and their volatilities. The weighted average of the estimated sticky and flexible trends (and cycles) corresponds to the trend (and cycle) for aggregate inflation. The estimation results are shown in figure 1.

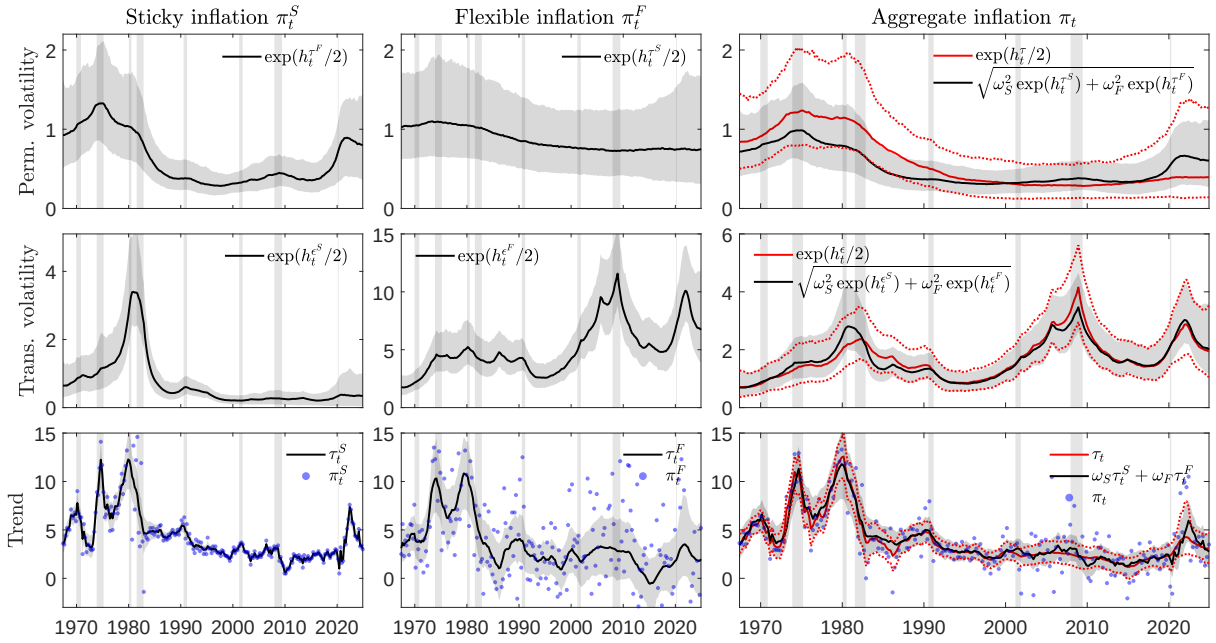

Figure 1: Permanent and transitory volatilities and trend inflation rates obtained from the UCSV model. Top to bottom: estimated permanent volatility, transitory volatility, and trend inflation. Left to right: estimation results for sticky inflation, flexible inflation, and aggregate inflation. Rightmost: red indicates aggregate inflation; whereas black shows volatilities and trend implied by the sticky-flexible inflation decomposition in equation (1) in the main text, assuming independent components. Solid lines show the posterior medians. Shaded areas and dotted lines show the $(0.05, 0.95)$ posterior quantiles. Bottom figures were cropped for readability, with some extreme inflation values (blue dots) outside the range.

3.2 Is time-varying cross-frequency correlation present?

Standard UCSV models assume uncorrelated trends and cycles, so permanent and transitory shocks affect the system dynamics separately. There are two types of correlations that are relevant in the context of multivariate UCSV models: (i) within-frequency correlation; and (ii) cross-frequency correlation.

Within-frequency correlation captures either the correlation between trends or the correlation between cycles, since trends and cycles capture the low- and high-frequency dynamics of the

time series, respectively. Given that a trend-cycle decomposition is applied to each component of inflation, it is reasonable to assume within-frequency correlation due to overlapping item categories and the potential time variation over the business cycle. Recently, Eo et al. (2023) considered a bivariate UCSV (BUCSV) model for the inflation of manufacturing goods (MI) and services (SI) along these lines. This model considers an UCSV model for both MI and SI, as well as stochastic correlations between the trends and between the cycles. In doing so, it is possible to attribute aggregate trend inflation variations to those of MI and SI.

We fit the bivariate UCSV (BUCSV) model of Eo et al. (2023) to sticky and flexible inflation, with the estimates of within-frequency correlations shown in figure 2.

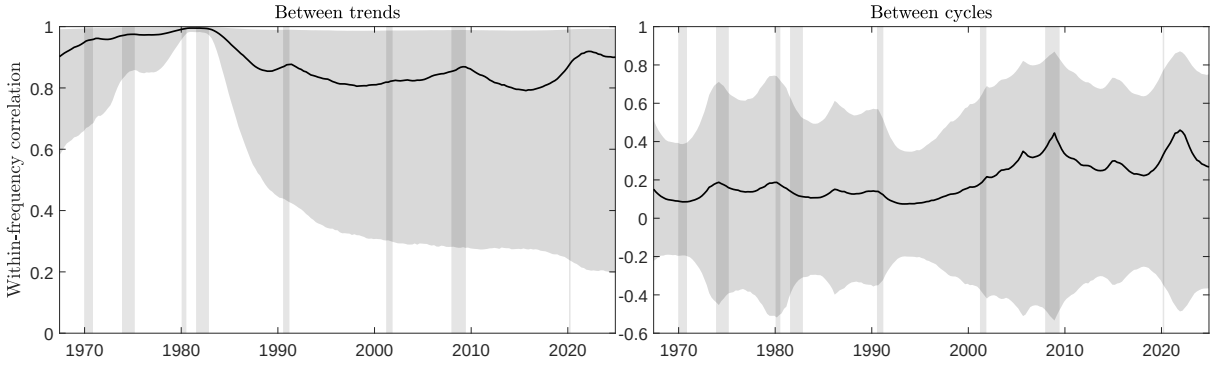

Figure 2: Within-frequency stochastic correlations obtained from the BUCSV model. We estimated the BUCSV model of Eo et al. (2023) for the sticky and flexible inflation rates. We report the stochastic correlation coefficients between trends and between cycles of sticky and flexible inflation. Lines and shaded areas indicate the posterior medians and $(0.05, 0.95)$ posterior quantiles, respectively.

We point out that, as discussed in section 1, our proposed bivariate FCUC model satisfies the order condition with an overidentified system. In this sense, we could add one extra parameter to incorporate the within-frequency correlation for cycles and obtain exact identification. However, as shown in figure 2, the results obtained from the BUCSV model do not suggest that this is a relevant feature. We also mention this in section 3.2 of the main text. Hence, give these results, we prefer to estimate a more parsimonious FCUC model by only incorporating within-frequency correlation for trends.

3.3 Can long-run inflation expectations anchor trend inflation?

It is known that CUC models tend to exacerbate the cross-frequency correlation in finite samples. For example, Morley et al. (2003), Trenkler and Weber (2016), Grant and Chan (2017), and Li

and Mendieta-Muñoz (2022) found that the trend-cycle correlation in real GDP is approximately −0.9 with trend (potential) output subject to excess volatility and an uneven path. Moreover, CUC models with a near-perfect (negative) correlation are also counter-intuitive to standard New Keynesian models, as shown by Canova and Ferroni (2022), and can provide poor policy guidance in practice.

One remedy is to connect trend inflation with long-run inflation expectations. Based on the Beveridge-Nelson decomposition (Beveridge and Nelson, 1981), there is an equivalence between trend and long-run rational expectations: $\tau_t = \lim_{s\to\infty} E_t(\pi_{t+s})$. Figure 3 shows our measure of inflation expectations together with sticky, flexible, and aggregate inflation.

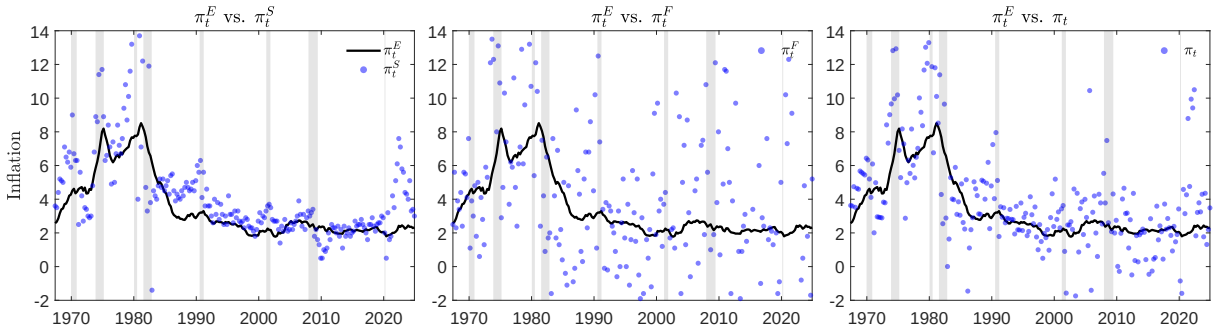

Figure 3: Long-run inflation expectations, sticky inflation, flexible inflation, and aggregate inflation. Long-run inflation expectations (black line) is plotted with sticky inflation (left figure), flexible inflation (middle figure), and aggregate inflation (right figure). Blue dots show the actual values for the different measures of inflation. Our measure of inflation expectation corresponds to the 10-year inflation expectations in Haubrich et al. (2012). Section 4.1 in the main text provides further details about the latter. Figures were cropped for readability, with some extreme inflation values outside the range.

4 Additional robustness checks

4.1 Pre-COVID-19 sample analysis and an expanding window exercise

Figure 4 shows the trend inflation rate derived from the FCUC model, together with the UCSV estimate, considering a sample ending in 2019:Q4, *i.e.*, before the COVID-19 shock. On the other hand, figures 5 and 6 below correspond to figures 4 and 5 in the main text, respectively, but these also consider only pre-COVID-19 data (period 1967:Q3-2019:Q4).

First, these figures show that the results obtained from the FCUC model for the period 1967:Q3-2019:Q4 are highly stable, robust, and can quickly adapt to the sudden fluctuations

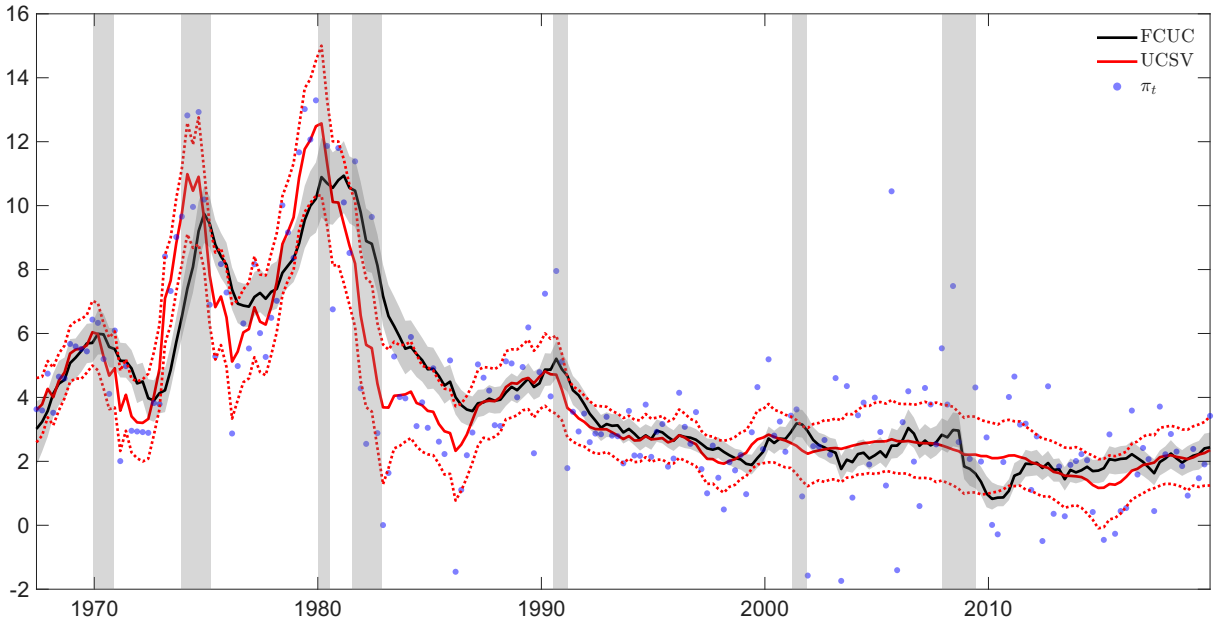

Figure 4: Posterior estimates of the trend inflation obtained from the FCUC and UCSV models using only pre-COVID-19 data (period 1967:Q3-2019:Q4). Trend inflation estimates obtained from the FCUC model are constructed following equation (7) in the main text. Lines show the posterior medians; whereas shaded areas and dotted lines show the $(0.05, 0.95)$ posterior quantiles for each model. Figure is cropped for readability, with some extreme inflation values (blue dots) outside the range.

derived from the COVID-19 shock.

Second, the results shown in figure 7 below indicate that, with the sample ending in 2019:Q4, there seems to be very limited cross-frequency correlation (or spillover) from the flexible price inflation to the trend inflation rate. In other words, these results imply that the increasing influence of the flexible inflation cycle on trend inflation observed in figure 6 in the main text is largely driven by the twofold increase in the cross-frequency coefficient ρ_t^{CF}. Therefore, the COVID-19 shock has changed the structural relationship between the permanent and transitory inflation components.

Lastly, central banks revise published inflation data as new data becomes available. To assess the sensitivity of our historical trend estimate to new data, we conduct an expanding window exercise that starts from 2019Q4. With each additional quarter, we re-estimate the model and store the trend estimate. Results shown in figure 8 suggests that our dynamic time series model is sufficiently flexible, able to pick up parameter shifts, and maintains stability even during noisy periods.

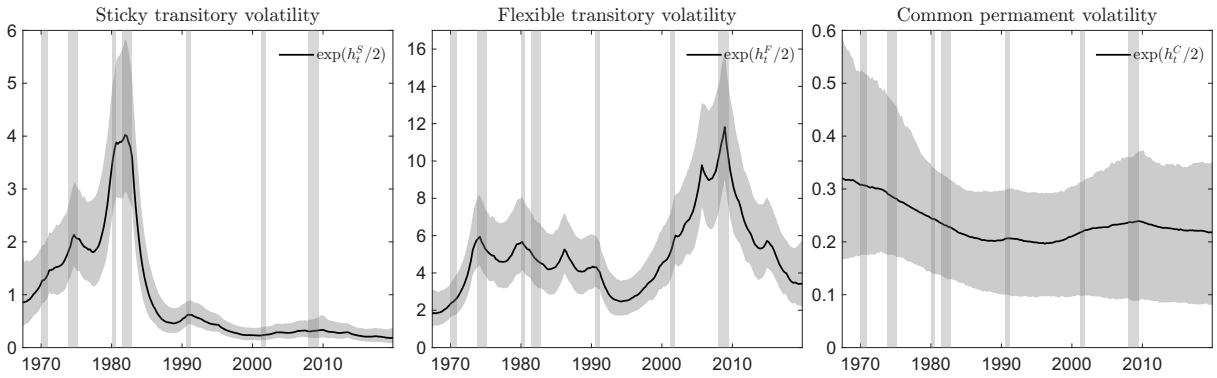

Figure 5: Posterior estimates of the stochastic volatility components obtained from the FCUC model using only pre-COVID-19 data (period 1967:Q3-2019:Q4). Left to right: sticky transitory volatility, flexible transitory volatility, and common permanent volatility. Lines show the posterior medians; whereas shaded areas show the $(0.05, 0.95)$ posterior quantiles.

4.2 Different measures of inflation expectations

Figures 9 and 10 show the variance decomposition analyses of the trend volatility estimates obtained from the FCUC model using 10-year-ahead inflation expectations data extracted from the University of Michigan Surveys of Consumers (UoM) and the Survey of Professional Forecasters (SPF), respectively. It is clear that, although the sample period is significantly shortened, the findings in the main text (figure 6 there) remain highly robust. This means that our findings are not sensitive to the structural assumptions used in the no-arbitrage model of Haubrich et al. (2012) and Federal Reserve Bank of Cleveland (2022).

5 Mixing properties and posterior statistics of model parameters

All results obtained from the FCUC model presented in the main text, except for the expanding-window forecasting exercise, are based on a post-burn in sample of 10,000 MCMC runs (with 5,000 initial runs). We further thin the chain, so that the posterior sample of size $N = 2,500$ consists of every fourth draw obtained from the MCMC runs. Estimation takes approximately 6 minutes on a low-end laptop bought in 2015. Results based on 150,000 runs are virtually identical.

Figure 11 shows the posterior traces obtained from the post-burn in sample of the FCUC model parameters. We also indicate posterior quantiles, including the 5th, 50th, and 95th percentile of each parameter's posterior distribution. Visual inspection suggests satisfactory mixing of the

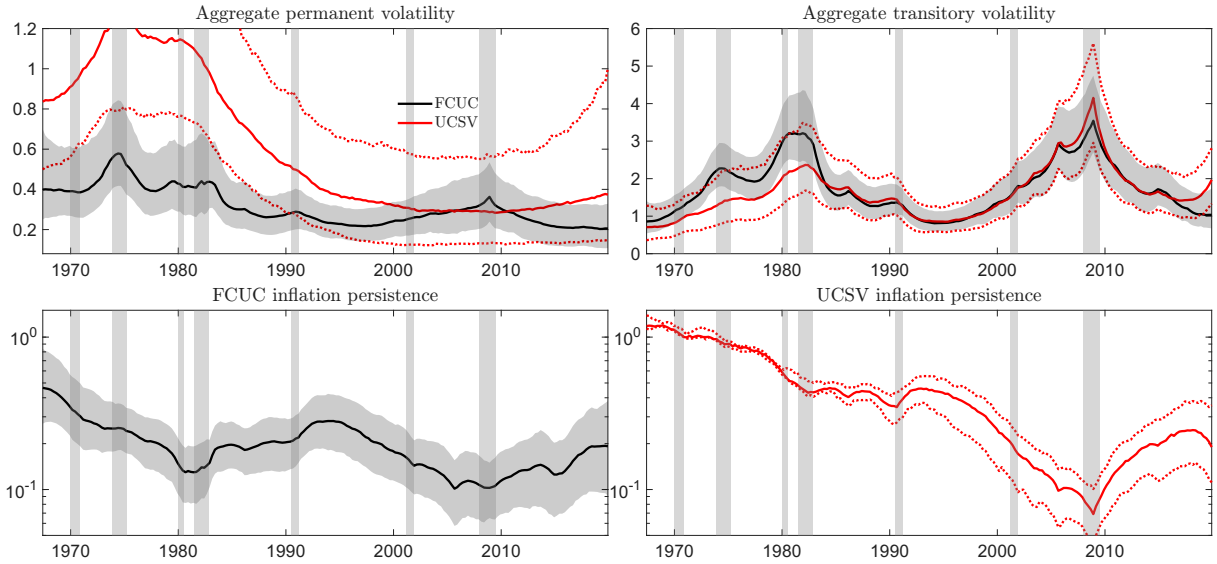

Figure 6: Aggregate permanent volatility, aggregate transitory volatility, and inflation persistence obtained from the FCUC model using only pre-COVID-19 data (period 1967:Q3-2019:Q4). Top row: Aggregate permanent and transitory volatility components obtained from the FCUC and UCSV models. Bottom row (y-axis in log scale): implied inflation persistence measured by the SNR. Lines show the posterior medians; whereas shaded areas and dotted lines show the $(0.05, 0.95)$ posterior quantiles for each model. The top left figure is cropped for readability.

Markov chain.

Figure 12 shows the autocorrelation function (ACF) of model parameters. For all plots, we observe that the ACFs converge to zero, which suggests a well-mixed Markov chain. The least efficient parameters are β_0 and α_0, the initialization parameters of the regression coefficients in the measurement equation of the long-run inflation expectations, *i.e.*, equation (4) in the main text. This may have to do with the fact that the majority of the information used for identifying trend inflation comes from the sticky and flexible inflation rate time series, with long-run inflation expectations offering only a marginal contribution. This result also resonates with our finding in the main text that long-run inflation expectations show persistent deviations from the assumption of rational expectations and their anchoring effect for the model is mainly effective during the two oil crises of the 1970s-80s.

We summarize the convergence diagnostics via the inefficiency factor (IEF) and the effective sample size (ESS), also reported in figure 12. The IEF measures the number of posterior draws needed to achieve the same inferential accuracy as one independent draw from the posterior. Thus, we prefer smaller IEFs (closer to one). The ESS indicates the number of posterior draws that can be regarded as independent and, therefore, we prefer a larger ESS (closer to N, the

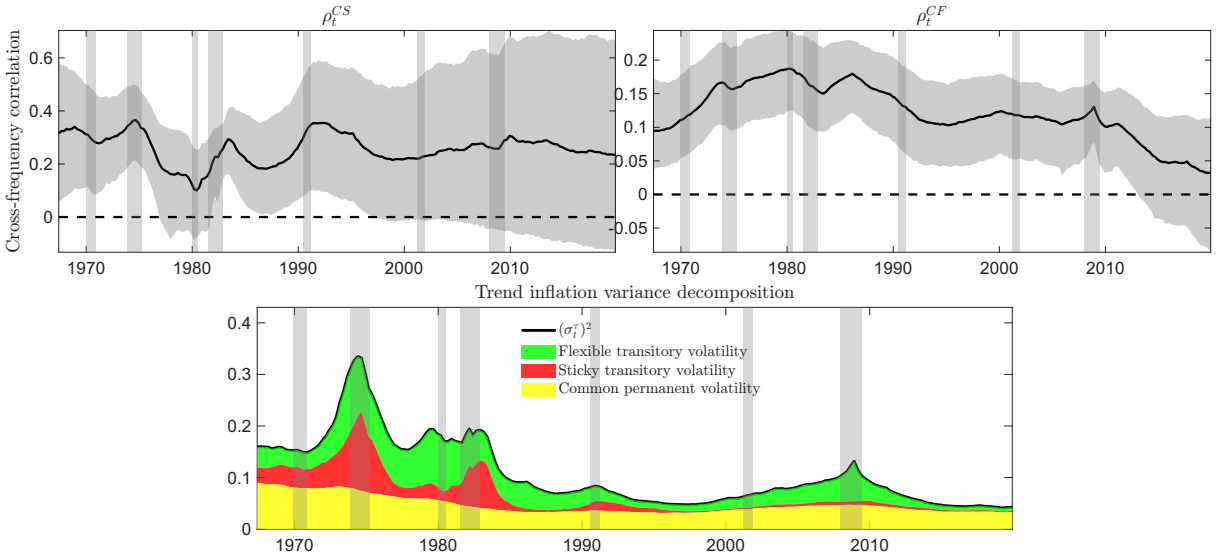

Figure 7: Cross-frequency correlations and variance decomposition of the aggregate trend inflation obtained from the FCUC model using only pre-COVID-19 data (period 1967:Q3-2019:Q4). Top row: ρ_t^{CS} (left figure) and ρ_t^{CF} (right figure) as depicted by equation (9) in the main text. Lines show the posterior medians; whereas shaded areas show the $(0.05, 0.95)$ posterior quantiles. Bottom row: variance decomposition based on equation (8) in the main text.

posterior sample size). Following Robert et al. (1999), these are defined by

$$\text{IEF}_n = 1 + 2\sum_{i=1}^{n}\left(1 - \frac{i}{n}\right)\rho_i,$$
$$\text{ESS}_n = \frac{N}{1 + 2\sum_{i=1}^{n}\rho_i},$$

where ρ_i denotes the i-th ACF of the posterior trace of a model parameter, $1 - i/n$ is the sequence of Bartlett weights, and n is a chosen cutoff that truncates the ACF. We choose n to be 120.

It matters little if we further increase the cutoff, as the Bartlett weights effectively end ρ_i at a

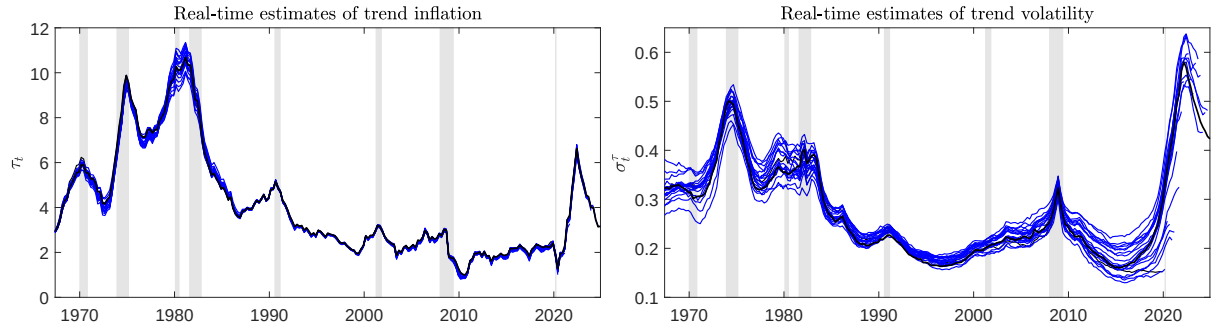

Figure 8: Real-time estimates of trend inflation and volatility. Reported are the posterior means of τ_t and σ_t^τ, re-estimated in an expanding window exercise starting from 2019:Q4. Black lines indicate the full-sample, or final, estimates.

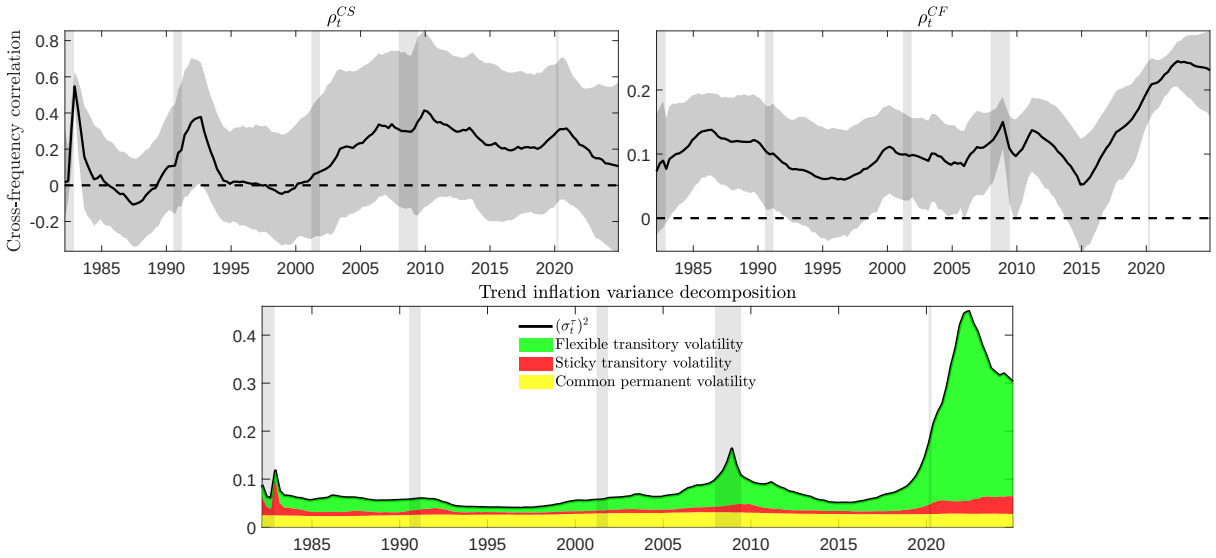

Figure 9: Cross-frequency correlations and variance decomposition of the aggregate trend inflation obtained from the FCUC model using UoM inflation expectations data (period 1982:Q1-2024:Q4). Top row: ρ_t^{CS} (left figure) and ρ_t^{CF} (right figure) as depicted by equation (9) in the main text. Lines show the posterior medians; whereas shaded areas show the $(0.05, 0.95)$ posterior quantiles. Bottom row: variance decomposition based on equation (8) in the main text.

linear rate, which is sufficient for a well-mixed chain. It can be shown that the IEF_n converges to the integrated autocorrelation time (long-run variance of the Markov chain, scaled by the variance of the chain) as $n \to \infty$. If the chain is not convergent, then the IEF explodes.

A rule-of-thumb for valid posterior inference is that the ESS must be larger than $N/10$, which is satisfied by all the parameters of the FCUC model, even the least efficient ones, β_0 and α_0. For σ_{h^S}, the chain is highly well-mixed. Note that its ESS is larger than N. This, however, does not suggest a numerical issue, but simply shows the fact that we were too conservative and thinned the chain more than necessary. Therefore, the MCMC algorithm for the FCUC model behaves satisfactorily, yielding a well-mixed chain and enabling valid posterior inference.

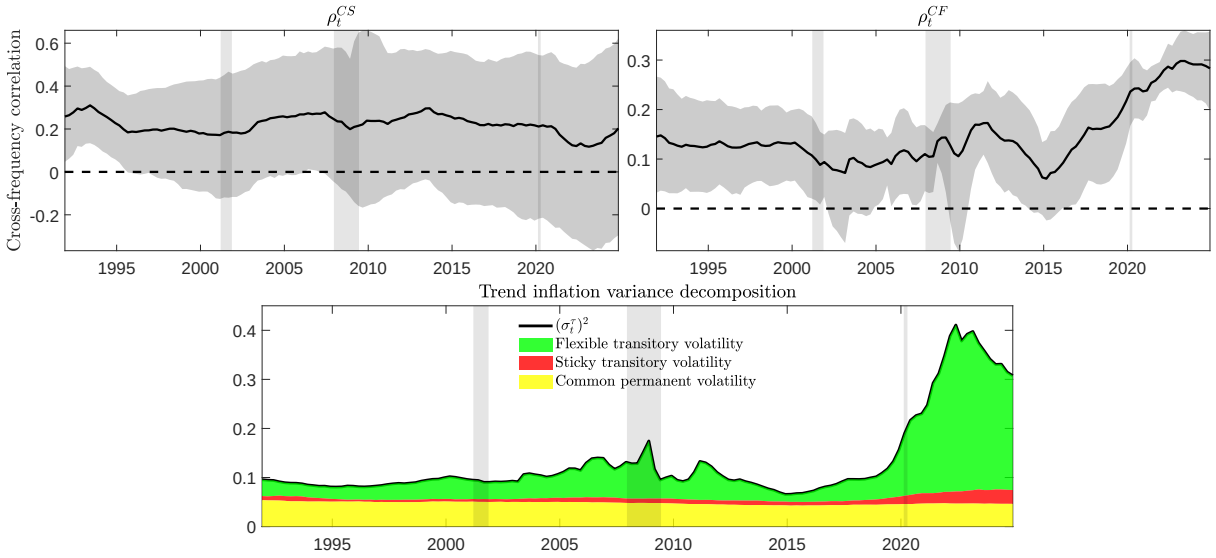

Figure 10: Cross-frequency correlations and variance decomposition of the aggregate trend inflation obtained from the FCUC model using SPF inflation expectations data (period 1991:Q1-2024:Q4). Top row: ρ_t^{CS} (left figure) and ρ_t^{CF} (right figure) as depicted by equation (9) in the main text. Lines show the posterior medians; whereas shaded areas show the $(0.05, 0.95)$ posterior quantiles. Bottom row: variance decomposition based on equation (8) in the main text.

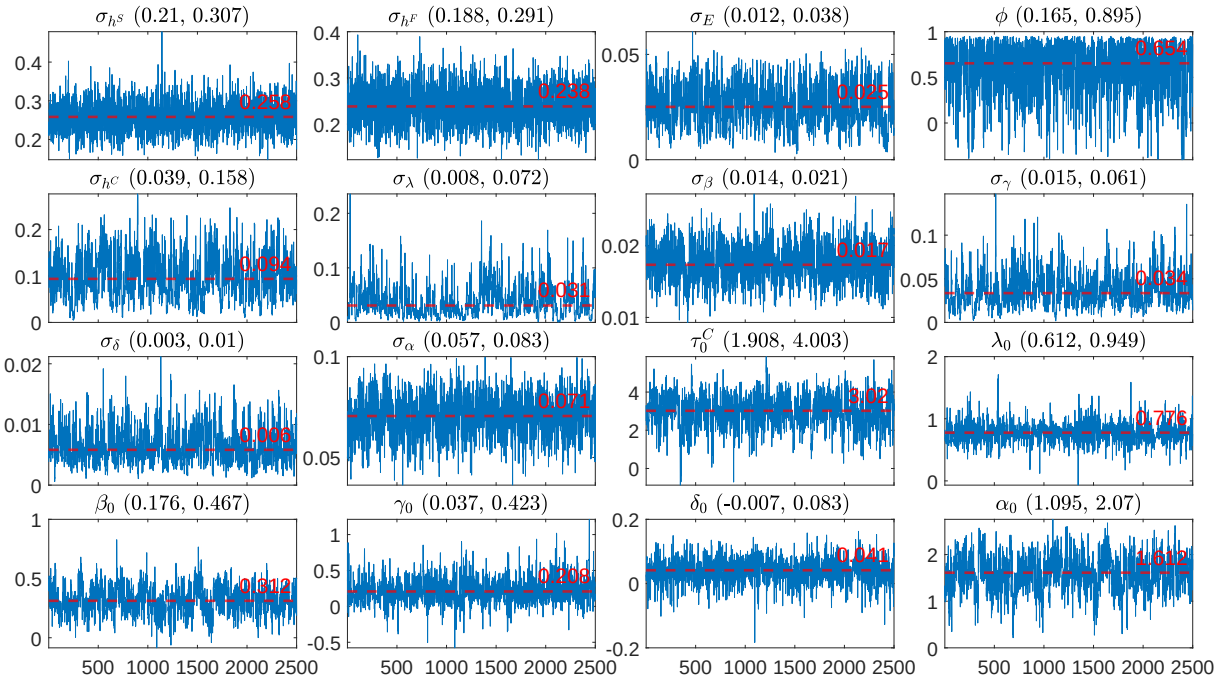

Figure 11: Posterior traces of parameters obtained from the FCUC model. Titles indicate parameters, following the same notations of the main text, and the 0.05 and 0.95 quantiles of their posterior distributions. Red lines and values show the posterior medians.

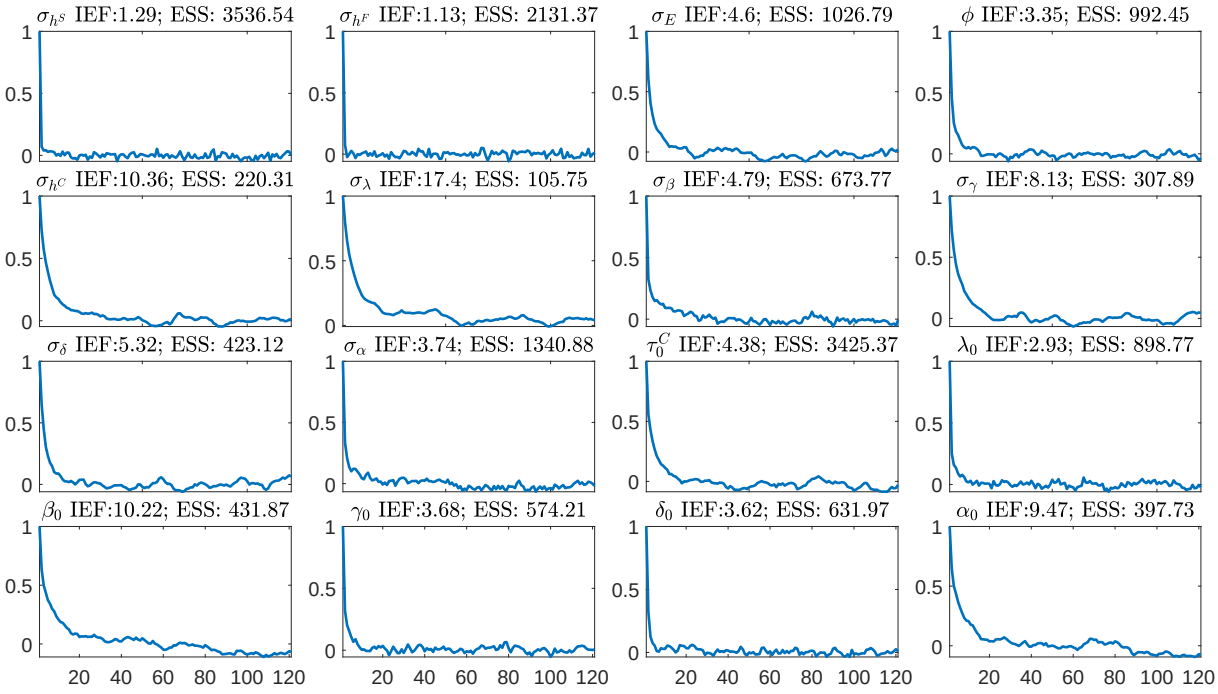

Figure 12: Autocorrelation functions of parameters obtained from the FCUC model. Titles indicate parameters, following the same notations of the main text, inefficiency factor (IEF) and effective sample size (ESS).

References

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

Canova, F. and F. Ferroni (2022). Mind the gap! Stylized dynamic facts and structural models. *American Economic Journal: Macroeconomics 14*(4), 104–135.

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

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *The annals of Statistics 25*(1), 1–37.

Eisenstat, E., J. C. Chan, and R. W. Strachan (2016). Stochastic model specification search for time-varying parameter VARs. *Econometric Reviews 35*(8-10), 1638–1665.

Eo, Y., L. Uzeda, and B. Wong (2023). Understanding trend inflation through the lens of the goods and services sectors. *Journal of Applied Econometrics 38*(5), 751–766.

Federal Reserve Bank of Cleveland (2022, August). Inflation expectations. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-infexp. Accessed May 2025.

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

Haubrich, J., G. Pennacchi, and P. Ritchken (2012). Inflation expectations, real rates, and risk premia: Evidence from inflation swaps. *Review of Financial Studies 25*(5), 1588–1629.

Jørgensen, P. L. and K. J. Lansing (2024). Anchored inflation expectations and the slope of the Phillips curve. *Federal Reserve Bank of San Francisco Working Paper 27*.

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

Lansing, K. J. (2022). Untangling persistent versus transitory shocks to inflation. *FRBSF Economic Letter 13*.

Li, M. and S. J. Koopman (2021). Unobserved components with stochastic volatility: Simulation-based estimation and signal extraction. *Journal of Applied Econometrics 36*(5), 614–627.

Li, M. and I. Mendieta-Muñoz (2022). Bayesian analysis of structural correlated unobserved components and identification via heteroskedasticity. *Studies in Nonlinear Dynamics & Econometrics 26*(3), 337–359.

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

Robert, C. P., G. Casella, and G. Casella (1999). *Monte Carlo statistical methods*, Volume 2. Springer.

Shephard, N. (2015). Martingale unobserved component models. In S. J. Koopman and N. Shephard (Eds.), *Unobserved Components and Time Series Econometrics*, pp. Chapter 10. Oxford University Press.

Stock, J. H. and M. W. Watson (2007). Why has US inflation become harder to forecast?

Journal of Money, Credit and Banking 39, 3–33.

Stock, J. H. and M. W. Watson (2020). Slack and cyclically sensitive inflation. *Journal of Money, Credit and Banking 52*(S2), 393–428.

Trenkler, C. and E. Weber (2016). On the identification of multivariate correlated unobserved components models. *Economics Letters 138*, 15–18.