Federal Reserve Bank of Cleveland Working Paper Series

Sticky Information Versus Sticky Prices Revisited: A Bayesian VAR-GMM Approach

Takushi Kurozumi, Ryohei Oishi, and Willem Van Zandweghe

Working Paper No. 22-34

November 2022

Suggested citation: Kurozumi, Takushi, Ryohei Oishi, and Willem Van Zandweghe. 2022. "Sticky Information Versus Sticky Prices Revisited: A Bayesian VAR-GMM Approach." Working Paper No. 22-34. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202234.

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

Sticky Information Versus Sticky Prices Revisited: A Bayesian VAR-GMM Approach*

Takushi Kurozumi† Ryohei Oishi‡ Willem Van Zandweghe§

November 2022

Abstract

Several Phillips curves based on sticky information and sticky prices are estimated and compared using Bayesian VAR-GMM. This method derives expectations in each Phillips curve from a VAR and estimates the Phillips curve parameters and the VAR coefficients simultaneously. Quasi-marginal likelihood-based model comparison selects a dual stickiness Phillips curve in which, each period, some prices remain unchanged, consistent with micro evidence. Moreover, sticky information is a more plausible source of inflation inertia in the Phillips curve than other sources proposed in previous studies. Sticky information, sticky prices, and unchanged prices in each period are all needed to better describe inflation dynamics.

JEL Classification: C11, C26, C52, E31

Keywords: Sticky information, Sticky price, Steady-state inflation, Inflation inertia, Bayesian VAR-GMM

---

*The authors are grateful to Yoosoon Chang, Todd Clark, Paolo Gelain, Ina Hajdini, Pierlauro Lopez, Kurt Lunsford, James Mitchell, Jouchi Nakajima, Hikaru Saijo, Mototsugu Shintani, Ellis Tallman, Takayuki Tsuruga, Randy Verbrugge, Toshiaki Watanabe, and participants at the ISER Macro/Trade Workshop, SETA 2022, 2022 AMES East and South East Asia, and a seminar at the Federal Reserve Bank of Cleveland for comments and discussions. The views expressed in the paper are those of the authors and do not necessarily reflect those of the Bank of Japan, the Federal Reserve Bank of Cleveland, or the Federal Reserve System.

†Monetary Affairs Department, Bank of Japan, 2-1-1 Nihonbashi-Hongokucho, Chuo-ku, Tokyo 103-8660, Japan (e-mail: takushi.kurozumi@boj.or.jp)

‡University College London, Gower Street, London, WC1E 6BT, UK and Bank of Japan, 2-1-1 Nihonbashi-Hongokucho, Chuo-ku, Tokyo 103-8660, Japan (e-mail: ryohei.oishi.22@ucl.ac.uk)

§Research Department, Federal Reserve Bank of Cleveland, P.O Box 6387, Cleveland, OH 44101, USA (e-mail: willem.vanzandweghe@clev.frb.org)

1 Introduction

Since the seminal work of Mankiw and Reis (2002), sticky information—slow dissemination of information throughout the population—has received much attention in macroeconomics. They propose the sticky information Phillips curve (SIPC) as a replacement for the New Keynesian Phillips curve (NKPC), which is based on sticky prices, and indicate that sticky information helps Phillips curves generate a plausible response of the economy to monetary policy shocks. Subsequent studies, such as Kiley (2007) and Coibion (2010), estimate and compare the SIPC and the NKPC. Despite its theoretical appeal for Mankiw and Reis (2002), the SIPC receives little support from these empirical studies, but they point to complementarity between sticky information and sticky prices within Phillips curves.[1] Thus, Dupor et al. (2010) and Knotek (2010) develop and estimate a Phillips curve based on both sticky information and sticky prices, and emphasize that its empirical performance is better than that of the SIPC and the NKPC. Dupor et al. (2010) call such a Phillips curve the dual stickiness Phillips curve (DSPC) and show that sticky information gives rise to inflation inertia. These studies, however, conduct no formal model comparison among the Phillips curves and take little account of the role of steady-state inflation in inflation dynamics, which has received increasing attention in recent studies reviewed by, for example, Ascari and Sbordone (2014).

This paper estimates and compares several Phillips curves based on sticky information and sticky prices. The benchmark Phillips curve is a DSPC and is derived from a staggered price model of Calvo (1983) in which, each period, some prices remain unchanged, consistent with micro evidence, while the other prices are set subject to sticky information as in Mankiw and Reis (2002). In the presence of the unchanged prices, the DSPC's inflation inertia and slope coefficients are influenced by the level of steady-state inflation. This feature contrasts sharply with the DSPC of Dupor et al. (2010), where steady-state inflation plays little role because price indexation to steady-state inflation is (implicitly) assumed. Their DSPC is thus referred to as the DSPC with indexation in this paper. Our DSPC is reduced to the SIPC in the absence of nominal price rigidity, in line with Dupor et al. (2010), whereas it

[1] In this context, Carroll (2003) points out that "the real world presumably combines some degree of price stickiness and a degree of expectational stickiness" (p. 295).

is reduced to the NKPC of Ascari (2004) in the absence of information rigidity, in contrast with Dupor et al. (2010)'s paper. These Phillips curves are estimated and compared during two periods: the Great Inflation period and thereafter.

To estimate the Phillips curves, this paper adopts Bayesian VAR-GMM. This limited-information method is particularly suitable for estimating Phillips curves with lagged expectations, such as the SIPC and the DSPC.[2] The method differs substantially from those used in the empirical literature on Phillips curves mainly in the following three respects.

First, in the GMM estimation, expectational variables in each Phillips curve are derived from a VAR, instead of being replaced with their realizations as in previous studies, including Galí and Gertler (1999). This method is referred to as VAR-GMM by Mavroeidis et al. (2014) and is adopted by Guerrieri et al. (2010) to estimate an open-economy version of the NKPC. As in Guerrieri et al. (2010), VAR-GMM can estimate Phillips curve parameters and VAR coefficients in one step, which contrasts with the two-step procedure employed in previous studies, such as Sbordone (2002), Cogley and Sbordone (2008), and Dupor et al. (2010), that first estimates a VAR and then infers Phillips curve parameters, given expectations derived from the separately estimated VAR.

Second, Bayesian methods are applied to the VAR-GMM estimation. This application is similar to that of Inoue and Shintani (2018) and Gemma et al. (2017), who adopt Bayesian methods in the classical GMM estimation. The issue of weak identification has been extensively discussed in the estimation of NKPCs, as in a review of the empirical literature by Mavroeidis et al. (2014). Kleibergen and Mavroeidis (2014) point out that this issue can be mitigated using Bayesian methods. Moreover, our paper utilizes a Block Metropolis-Hastings algorithm (see, e.g., Herbst and Schorfheide, 2015) to estimate Phillips curve parameters and VAR coefficients simultaneously.

Third, quasi-marginal likelihood (QML) is used for model selection, that is, not only selection from several Phillips curves but also selection of the lag length in the VAR for each Phillips curve. As shown by Inoue and Shintani (2018), a model with higher QML can be regarded as a better one, and this model selection procedure is valid even when some model parameters are weakly identified.

[2] The limited-information method is adopted because it leaves unspecified other equations in models and it is therefore much less subject to misspecification issues than full-information methods.

The main findings of the paper are twofold.

First, the DSPC is the best Phillips curve both during and after the Great Inflation period among all those considered. Model comparison based on QML shows that the DSPC is superior to the SIPC, the NKPC of Ascari (2004), and the DSPC with indexation of Dupor et al. (2010), as well as the extensively used NKPC of Smets and Wouters (2007). The estimated DSPC indicates that when steady-state inflation fell after the Great Inflation period, the probability of price change decreased, thereby generating a flattening of the Phillips curve. This empirical finding is consistent with the theoretical prediction in the literature on endogenous price stickiness, initiated by Ball et al. (1988) and subsequently developed by Levin and Yun (2007) and Kurozumi (2016). The decrease in the estimated probability of price change leads to an increase in its implied average duration, which is in line with the rise in the duration implied by the frequency of regular price change reported in Nakamura et al. (2018).

Second, sticky information is a more plausible source of inflation inertia in the DSPC both during and after the Great Inflation period than other sources proposed in previous studies. As a source of inflation inertia in Phillips curves, Galí and Gertler (1999) introduce rule-of-thumb price setters, while Kurozumi and Van Zandweghe (2019) suggest variable elasticity of demand generated by a non-CES goods aggregator of the sort proposed by Kimball (1995).[3] Replacing sticky information with each of the two sources in the DSPC reduces QML, thus deteriorating the empirical performance of the resulting Phillips curve. As pointed out by Mankiw and Reis (2002), sticky information is due to the presence of firms' costs of information acquisition and reoptimization. The importance of these costs is demonstrated by Zbaracki et al. (2004), who indicate that managerial costs, including information gathering and decision-making costs, are much larger than menu costs, using data from a large industrial manufacturer.

These two main findings suggest that sticky information, sticky prices, and unchanged prices in each quarter are all needed to better describe inflation dynamics both during and after the Great Inflation period. This extends the result of Dupor et al. (2010) in the direction of more consistency with micro evidence on price setting.

[3]In the literature there is another source of inflation inertia, the upward-sloping hazard function proposed by Sheedy (2010). This source is not employed due to the lack of tractability.

This paper contributes to two strands of the literature. Since it was proposed by Mankiw and Reis (2002), the SIPC has been estimated by Khan and Zhu (2006), Kiley (2007), Coibion (2010), and Coibion and Gorodnichenko (2011). Subsequent studies of Klenow and Willis (2007), Dupor et al. (2010), and Knotek (2010) develop DSPCs and demonstrate their empirical relevance relative to the SIPC (and the NKPC). Our paper extends the DSPC further along the lines of the literature on NKPCs reviewed by Ascari and Sbordone (2014) and shows that such an extension improves the empirical performance of the DSPC.

Many of the previous studies have estimated Phillips curves using limited-information methods.[4] Galí and Gertler (1999) and Galí et al. (2005) employ GMM in estimating NKPCs. Guerrieri et al. (2010) utilize VAR-GMM to estimate an open-economy version of the NKPC. Dupor et al. (2010) exploit the aforementioned two-step procedure to estimate the DSPC with indexation. To circumvent the weak-identification problem with estimated NKPCs, Inoue and Shintani (2018) and Gemma et al. (2017) adopt Bayesian GMM and compare estimated NKPCs using QML. Our paper utilizes Bayesian VAR-GMM and a Block Metropolis-Hastings algorithm to estimate parameters of our DSPC and its associated VAR coefficients simultaneously, and conducts model selection using QML to show that the DSPC is the best Phillips curve among all those considered, including the DSPC with indexation.

The remainder of the paper proceeds as follows. Section 2 presents several Phillips curves. Section 3 explains our method and data for estimating them. Section 4 shows our empirical results. Section 5 concludes.

2 Phillips Curves

This section presents several Phillips curves based on sticky information and sticky prices.

2.1 Dual stickiness Phillips curve

Our benchmark Phillips curve is a dual stickiness Phillips curve (DSPC). This Phillips curve is derived from a staggered price model of Calvo (1983) in which, each period, a fraction $\lambda \in [0, 1)$ of prices remains unchanged, while the other prices are set subject to sticky

[4] Full-information methods are employed in, for example, Smets and Wouters (2007) and Hirose et al. (2021).

information as in Mankiw and Reis (2002): a fraction $\omega \in [0, 1)$ of the prices is chosen optimally without information update, whereas the remaining fraction is optimized with it.

As shown in Appendix A, under the assumption

$$\lambda \max\left(\pi^{\theta}, \pi^{\theta-1}\right) < 1, \tag{1}$$

the DSPC can be obtained as

$$\begin{aligned}
\hat{\pi}_t = \kappa_b \hat{\pi}_{t-1} + \kappa_f E_t \hat{\pi}_{t+1} + \kappa\, \hat{mc}_t + \kappa_\phi \sum_{j=1}^{\infty} (\beta\lambda\pi^{\theta-1})^j E_t \phi_{1,t+j} - \kappa_\omega \sum_{j=1}^{\infty} (\beta\lambda\pi^{\theta})^j E_t \phi_{2,t+j} \\
+ \kappa_\omega \sum_{j=1}^{\infty} \omega^{j-1} \Bigg[\omega\beta\lambda\pi^{\theta} E_{t-j} \hat{mc}_t - E_{t-j} \hat{mc}_{t-1} - \sum_{k=1}^{\infty} (\beta\lambda\pi^{\theta})^{k-1} \left(\omega\beta\lambda\pi^{\theta} E_{t-j} \phi_{2,t+k} - E_{t-j} \phi_{2,t+k-1} \right) \\
+ \frac{1 - \beta\lambda\pi^{\theta-1}}{1 - \beta\lambda\pi^{\theta}} \sum_{k=1}^{\infty} (\beta\lambda\pi^{\theta-1})^{k-1} \left(\omega\beta\lambda\pi^{\theta} E_{t-j} \phi_{1,t+k} - E_{t-j} \phi_{1,t+k-1} \right) \Bigg],
\end{aligned} \tag{2}$$

where hatted variables denote log-deviations from steady-state values (e.g., $\hat{\pi}_t \equiv \log \pi_t - \log \pi$), E_t is the expectation operator conditional on information available in period t, π_t is the inflation rate, mc_t is the real marginal cost, $\phi_{1,t}$ and $\phi_{2,t}$ are auxiliary variables defined as

$$\phi_{1,t} \equiv \hat{gy}_t + \theta\, \hat{\pi}_t - \hat{r}_{t-1}, \quad \phi_{2,t} \equiv \phi_{1,t} + \hat{\pi}_t + (1 - \beta\lambda\pi^{\theta})\, \hat{mc}_t,$$

gy_t is the output growth rate, r_t is the nominal interest rate, π is the steady-state inflation rate, $\beta \in (0, 1)$ is the subjective discount factor, $\theta > 1$ is the elasticity of substitution between individual goods, and the coefficients κ_b, κ_f, κ, κ_ϕ, and κ_ω are presented in Appendix A.

In the DSPC (2), two points are worth noting. First, the real marginal cost $\hat{mc}_t$ consists not only of the real unit labor cost ulc_t but also of the relative price distortion Δ_t:

$$\hat{mc}_t = \hat{ulc}_t - \hat{\Delta}_t, \tag{3}$$

since the distortion has a first-order effect in the DSPC under nonzero steady-state inflation. The law of motion of the distortion is then given by

$$\hat{\Delta}_t = \rho_\Delta \hat{\Delta}_{t-1} + \kappa_\Delta \hat{\pi}_t, \tag{4}$$

where the coefficients ρ_Δ and κ_Δ are presented in Appendix A.

Second, all the coefficients in the DSPC depend not only on the probability of no price change λ and the probability of no information update ω but also on steady-state inflation π. The inflation inertia coefficient κ_b is present under sticky information (i.e., $\omega > 0$), as emphasized by Dupor et al. (2010), and it declines with a lower probability ω of no information update, a lower probability λ of no price change, and a lower steady-state inflation rate π. In contrast, the slope coefficient κ decreases with a higher ω, a higher λ, and a higher π.

In the absence of nominal price rigidity (i.e., $\lambda = 0$), the DSPC is reduced to the SIPC

$$\hat{\pi}_t = \frac{1-\omega}{\omega}\,\hat{ulc}_t + (1-\omega)\sum_{j=1}^{\infty}\omega^{j-1}\Big(E_{t-j}\hat{\pi}_t + E_{t-j}\hat{ulc}_t - E_{t-j}\hat{ulc}_{t-1}\Big), \tag{5}$$

where $\hat{mc}_t = \hat{ulc}_t$ because no nominal price rigidity implies no first-order effect of the relative price distortion, i.e., $\hat{\Delta}_t = 0$.[5] On the other hand, in the absence of information rigidity (i.e., $\omega = 0$), the DSPC is reduced, under assumption (1), to the NKPC of Ascari (2004)

$$\hat{\pi}_t = \beta\pi E_t\hat{\pi}_{t+1} + \kappa_0\Big(\hat{ulc}_t - \hat{\Delta}_t\Big) + \kappa_{\phi 0}\sum_{j=1}^{\infty}(\beta\lambda\pi^{\theta-1})^j\big(E_t\hat{gy}_{t+j} + \theta E_t\hat{\pi}_{t+j} - E_t\hat{r}_{t+j-1}\big), \tag{6}$$

where the coefficients κ_0 and $\kappa_{\phi 0}$ correspond to κ and κ_ϕ at $\omega = 0$, respectively. The law of motion of the relative price distortion $\hat{\Delta}_t$ remains the same as (4).

2.2 Phillips curves with price indexation to steady-state inflation

The DSPC with indexation of Dupor et al. (2010) is a simple variant of the DSPC (2). It can be obtained only by altering the model so that the prices that remain unchanged in the aforementioned setting are instead adjusted using indexation to steady-state inflation. Thus the parameter λ represents the probability of steady-state inflation-indexed price setting. This implies that all prices change in every period, which contradicts micro evidence. Moreover, the level of steady-state inflation has no influence on the Phillips curve coefficients, and the relative price distortion has no first-order effect, i.e., $\hat{\Delta}_t = 0$. Indeed, the DSPC with indexation coincides with the DSPC (2) in which its coefficients are set at their values under

[5] In the absence of nominal price rigidity (i.e., $\lambda = 0$), assumption (1) is always met.

zero steady-state inflation (i.e., $\pi = 1$) and there is no distortion term:

$$\hat{\pi}_t = \kappa_{b1}\hat{\pi}_{t-1} + \kappa_{f1}E_t\hat{\pi}_{t+1} + \kappa_1\hat{ulc}_t - \kappa_{\omega 1}\sum_{j=1}^{\infty}(\beta\lambda)^j\Big(E_t\hat{\pi}_{t+j} + (1-\beta\lambda)E_t\hat{ulc}_{t+j}\Big)$$
$$+ \kappa_{\omega 1}\sum_{j=1}^{\infty}\omega^{j-1}\Bigg\{\omega\beta\lambda E_{t-j}\hat{ulc}_t - E_{t-j}\hat{ulc}_{t-1} - \sum_{k=1}^{\infty}(\beta\lambda)^{k-1}\Big[\omega\beta\lambda E_{t-j}\hat{\pi}_{t+k} - E_{t-j}\hat{\pi}_{t+k-1}$$
$$+ (1-\beta\lambda)\Big(\omega\beta\lambda E_{t-j}\hat{ulc}_{t+k} - E_{t-j}\hat{ulc}_{t+k-1}\Big)\Big]\Bigg\}, \tag{7}$$

where the coefficients κ_{b1}, κ_{f1}, κ_1, and $\kappa_{\omega 1}$ correspond to κ_b, κ_f, κ, and κ_ω at $\pi = 1$, respectively. Note that the DSPC with indexation (7) is reduced to the SIPC (5) in the absence of nominal price rigidity (i.e., $\lambda = 0$), while it is reduced to the textbook NKPC

$$\hat{\pi}_t = \beta E_t\hat{\pi}_{t+1} + \kappa_1\hat{ulc}_t \tag{8}$$

in the absence of information rigidity (i.e., $\omega = 0$).

3 Estimation Method and Data

This section explains our method and data for estimating the Phillips curves presented in the preceding section.

3.1 Bayesian VAR-GMM

This paper utilizes Bayesian VAR-GMM to estimate the Phillips curves. This limited-information method is particularly suitable for estimating Phillips curves with lagged expectations, such as the SIPC and the DSPC. The method differs substantially from those used in the empirical literature on Phillips curves mainly in the following three respects.

First, in the GMM estimation, expectational variables in each Phillips curve are derived from a VAR, instead of being replaced with their realizations as in previous studies, such as Galí and Gertler (1999) and Galí et al. (2005). This method is referred to as VAR-GMM by Mavroeidis et al. (2014) and is adopted by Guerrieri et al. (2010) to estimate an open-economy version of the NKPC. With VAR-GMM, the Phillips curve parameters and the

VAR coefficients can be estimated simultaneously, as in Guerrieri et al. (2010). This one-step procedure contrasts with the two-step one employed in previous studies (e.g., Sbordone, 2002; Cogley and Sbordone, 2008; Dupor et al., 2010), which first estimate a VAR and then infer Phillips curve parameters, given expectations derived from the separately estimated VAR. In our model, the steady-state inflation rate π is among not only the Phillips curve parameters but also the VAR coefficients, and thus our one-step procedure can estimate the rate π so as to meet restrictions imposed by both the Phillips curve and the VAR.[6]

Second, Bayesian methods are applied to the VAR-GMM estimation. This application is similar to that of Inoue and Shintani (2018) and Gemma et al. (2017), who adopt Bayesian methods in the classical GMM estimation. As reviewed by Mavroeidis et al. (2014), the literature has extensively discussed the weak identification of parameters, which arises when instruments are weakly correlated with endogenous regressors. It is an issue in the estimation of NKPCs that reflects the difficulty of forecasting inflation. Kleibergen and Mavroeidis (2014) point out that Bayesian methods can mitigate this issue and be considered as an alternative to the weak-identification robust GMM statistic of Kleibergen and Mavroeidis (2009) for inference of weakly identified parameters of NKPCs. Moreover, our paper utilizes a Block Metropolis-Hastings algorithm (see, e.g., Herbst and Schorfheide, 2015) to estimate the Phillips curve parameters and the VAR coefficients jointly.

Third, quasi-marginal likelihood (QML) is used for model selection, that is, not only selection from several Phillips curves but also selection of the lag length in the VAR for each Phillips curve. As shown by Inoue and Shintani (2018), a model with higher QML can be regarded as a better one, and this procedure of model selection is valid even when some model parameters are weakly identified or set identified. Using QML, Inoue and Shintani (2018) compare the NKPCs of Galí and Gertler (1999) and Smets and Wouters (2007), while Gemma et al. (2017) compare NKPCs with price indexation to steady-state inflation and without it.

[6] Another advantage of VAR-GMM is that it allows estimation up to the latest period. For example, when the Ascari NKPC (6) is estimated using GMM as in previous studies, such as Galí and Gertler (1999) and Galí et al. (2005), the NKPC contains the infinite-horizon expectations $\sum_{j=1}^{\infty}(\beta\lambda\pi^{\theta-1})^j\left(E_t\hat{gy}_{t+j}+\theta E_t\hat{\pi}_{t+j}-E_t\hat{r}_{t+j-1}\right)$ and thus such expectations need to be truncated at a finite horizon to replace them with their realizations, so the NKPC cannot be estimated up to the latest period. VAR-GMM replaces the infinite-horizon expectations with their corresponding VAR forecasts, and thus it can estimate the NKPC up to the latest period.

Next, the Bayesian VAR-GMM estimation of the Phillips curves is described. The DSPC (2) includes four variables: the inflation rate $\hat{\pi}_t$, the unit labor cost $\hat{ulc}_t$, the output growth rate $\hat{gy}_t$, and the nominal interest rate $\hat{r}_t$. Thus, it is assumed that expectational variables in each Phillips curve are derived from a finite-order VAR with the four variables,

$$X_t \equiv [\hat{\pi}_t, \hat{ulc}_t, \hat{gy}_t, \hat{r}_t, \dots, \hat{\pi}_{t-n+1}, \hat{ulc}_{t-n+1}, \hat{gy}_{t-n+1}, \hat{r}_{t-n+1}]' = A\, X_{t-1} + \varepsilon_t,$$

where n denotes the lag length in the VAR. Under this assumption, we derive, for example, the j-period-ahead inflation forecast as $E_t\hat{\pi}_{t+j} = e_\pi' E_t X_{t+j} = e_\pi' A^j X_t$ and the j-period-ago forecast of current inflation as $E_{t-j}\hat{\pi}_t = e_\pi' A^j X_{t-j}$, where e_π is the selection vector for inflation. Replacing expectational variables with their corresponding VAR forecasts leads to the representation of each Phillips curve for estimation. Specifically, for the DSPC, such a representation can be obtained by combining the DSPC (2), the marginal cost equation (3), and the law of motion of the relative price distortion (4) and then replacing expectational variables in the resulting equation with their corresponding VAR forecasts. The representation of the other Phillips curves can be derived likewise. Then, the DSPC (2), the one with indexation (7), and the SIPC (5) include infinite backward summation, which is approximated with the truncated sum of 16 lags, following Dupor et al. (2010).[7]

Let $\varphi \equiv [\vartheta', vec(A)']'$ denote a vector that combines Phillips curve parameters ϑ and its associated VAR coefficients $vec(A)$, and let $g_t(\varphi)$ be a vector of moment functions that satisfies $E(g_t(\varphi)) = 0$ at a true value of $\varphi = \varphi_0$, where E is the unconditional expectation operator. Define $g_t(\varphi)$ as

$$g_t(\varphi) = \begin{bmatrix} h_t(\varphi) Z_t \\ (X_t - A\, X_{t-1}) X_{t-1} \\ \log \pi_t - \log \pi \end{bmatrix}, \tag{9}$$

where $h_t(\varphi)$ is each Phillips curve's residual function and Z_t is the vector of instruments including a constant of unity. The top and middle parts of $g_t(\varphi)$ imply the orthogonality conditions for each Phillips curve and the VAR, respectively. The bottom part presents our estimation's assumption that $\log \pi = E(\log \pi_t)$, i.e., $E\hat{\pi}_t = 0$.

The efficient two-step GMM estimator is employed, as in previous studies, such as Galí

[7] We also experimented with 12 or 20 lags and found that the results were qualitatively unaffected.

and Gertler (1999), Galí et al. (2005), and Guerrieri et al. (2010). The estimator chooses $\varphi \in \Phi$ so as to maximize the objective function $q(\varphi) = -(1/2)\, g(\varphi)'\, W\, g(\varphi)$, where $g(\varphi) = (1/\sqrt{T}) \sum_{t=1}^{T} g_t(\varphi)$ and W is the optimal weighting matrix based on the HAC estimator of Newey and West (1987).[8] The matrix W is calculated as $W = [\Gamma_0(\tilde{\varphi}) + \sum_{j=1}^{J}((J - j)/J)(\Gamma_j(\tilde{\varphi}) + \Gamma_j(\tilde{\varphi})')]^{-1}$, where $\Gamma_j(\tilde{\varphi}) = [1/(T-j)] \sum_{t=j+1}^{T} g_t(\tilde{\varphi}) g_{t-j}(\tilde{\varphi})'$, $\tilde{\varphi}$ is an arbitrary consistent estimator, and the lag length J is selected by the automatic bandwidth selection method of Andrews (1991).

Following Chernozhukov and Hong (2003), Bayesian methods are applied to the VAR-GMM estimation. The quasi-posterior distribution of φ is defined as

$$\frac{\exp(q(\varphi))p(\varphi)}{\int_{\Phi} \exp(q(\varphi))p(\varphi)d\varphi},$$

where $p(\varphi)$ is the prior distribution for φ. To obtain the quasi-posterior distribution, the Markov Chain Monte Carlo (MCMC) method is used. Specifically, our paper utilizes the Block Metropolis-Hastings algorithm with two blocks for Phillips curve parameters and VAR coefficients to estimate them simultaneously without time-consuming computation of a large Hessian matrix of the quasi-posterior probability density for the moment functions $g_t(\varphi)$ in (9) and its computational error.[9] To obtain initial values for MCMC draws, $10,000$ draws are generated from the prior distribution of Phillips curve parameters, and the quasi-posterior probability density is numerically maximized to find the mode, given values of coefficients in the separately estimated VAR. The Block Metropolis-Hastings algorithm is then used to generate $210,000$ MCMC draws, and the first $10,000$ draws are discarded as a burn-in to obtain the quasi-posterior distribution.[10]

In conducting model selection, this paper follows Inoue and Shintani (2018) and uses the

[8] The HAC covariance matrix estimator of the moment functions $g_t(\varphi)$ in (9) is employed for two reasons. First, the inflation gap $\log \pi_t - \log \pi$ in (9) is possibly serially correlated. Second, the use of the HAC estimator makes the resulting estimation valid not only for the exact representation of each Phillips curve but also for the case in which a disturbance (e.g., cost-push shock) is incorporated in it.

[9] See Appendix C for details of the Block Metropolis-Hastings algorithm in the Bayesian VAR-GMM estimation of Phillips curves.

[10] In estimating the DSPC (2) and the NKPC (6), $210,000$ MCMC draws that meet assumption (1) are generated.

QML defined as

$$\int_{\Phi} \exp(q(\varphi))p(\varphi)d\varphi\,.$$

With the modified harmonic mean method of Geweke (1999), the QML is calculated for each Phillips curve with each lag length of the VAR.[11] Then, a Phillips curve with a lag length of the VAR that has higher QML is selected as a better model.[12]

3.2 Data

The primary data used in the estimation are four US quarterly time series: the inflation rate π_t, the real unit labor cost ulc_t, the output growth rate gy_t, and the nominal interest rate r_t. Following Galí et al. (2005), we employ four lags of the inflation rate and two lags of the nominal wage growth rate, the real unit labor cost, the output growth rate, and the nominal interest rate as instruments in Z_t.

As in Galí and Gertler (1999) and Galí et al. (2005), the data on π_t is the inflation rate of the GDP implicit price deflator, and that on ulc_t is the labor income share in the nonfarm business sector. Those on gy_t and r_t are the per capita real GDP growth rate and the three-month Treasury bill rate, respectively. The nominal wage growth data is based on hourly compensation in the nonfarm business sector. To take into account a shift in steady-state inflation in post-WWII US inflation history, we perform each estimation separately for the Great Inflation period (1966:Q1–1982:Q3) and the period thereafter (1982:Q4–2019:Q4). Recent samples (since 2020) are not used due to the COVID-19 pandemic.

For the real unit labor cost, the output growth rate, and the nominal interest rate, the time series of their log-deviations from steady-state values $\{\hat{ulc}_t, \hat{gy}_t, \hat{r}_t\}$ are constructed separately for each sample period; they are all demeaned using their respective sample period averages.

[11] This paper reports the Geweke (1999) modified harmonic mean estimator of QML using the value of its truncation parameter of $\tau = 0.5$. We confirmed the robustness of the model selection results with respect to the modified harmonic mean method, by using an alternative value of the truncation parameter of $\tau = 0.9$ in the Geweke (1999) estimator and employing an alternative estimator of Sims et al. (2008) with the values of the truncation parameter of $q = 0.5, 0.9$, as in Herbst and Schorfheide (2015).

[12] The difference in the QML between two Phillips curves or models can be evaluated using, for example, the Jeffreys (1961) criterion, which is similar to the marginal likelihood-based model comparison in full-information Bayesian estimation.

3.3 Fixed parameters and prior distributions

In each estimation, values of two parameters are fixed to avoid identification issues. The subjective discount factor is set at $\beta = 0.9975$, and the elasticity of substitution between goods is chosen at $\theta = 9.32$, which is the estimate of Ascari and Sbordone (2014). All the remaining parameters in each Phillips curve and coefficients in its associated VAR are estimated.

Table 1 presents the prior distributions for the parameters in each Phillips curve. The prior for the annualized steady-state inflation rate $\bar{\pi}$ ($\equiv 400 \log \pi$) is centered around 3.4, which is an average of the inflation rate over the period 1966:Q1–2019:Q4, with standard deviation 1.5. The prior distributions for the probability of no price change λ and the probability of no information update ω are set to be the beta distributions with mean 0.5 and standard deviation 0.1.[13]

Table 1: Prior distributions for Phillips curve parameters

Phillips curve parameter		Distribution	Mean	Std. dev.
$\bar{\pi}$	annualized steady-state inflation rate	normal	3.4	1.5
λ	probability of no price change	beta	0.5	0.1
ω	probability of no information update	beta	0.5	0.1
ι	degree of price indexation to lagged inflation	beta	0.5	0.1
ω_r	fraction of rule-of-thumb price setters	beta	0.5	0.1
$-\epsilon$	parameter governing the curvature of demand curves	gamma	3	1

Notes: $\bar{\pi} \equiv 400 \log \pi$. In the DSPC with indexation (7), λ represents the probability of steady-state inflation-indexed price setting.

For the VAR coefficients, this paper employs the Minnesota prior in which the covariance matrix of the VAR error term is replaced with the OLS estimate, following Canova (2007). The hyper-parameters of the prior are also set in the same manner as in Canova (2007).

4 Empirical Results

This section presents the results of model selection from the Phillips curves presented above and then analyzes the selected Phillips curve.

[13] Phillips curve parameters reported in the lower panel of Table 1 are explained later.

4.1 Model comparison of Phillips curves

This paper utilizes QML to conduct model selection from the Phillips curves. To evaluate their empirical performance, the extensively used NKPC of Smets and Wouters (2007) is also considered:

$$\hat{\pi}_t = \frac{\iota}{1+\beta\iota}\,\hat{\pi}_{t-1} + \frac{\beta}{1+\beta\iota}\,E_t\hat{\pi}_{t+1} + \frac{(1-\lambda)(1-\beta\lambda)}{\lambda(1+\beta\iota)}\,\hat{ulc}_t, \tag{10}$$

where $\iota \in [0,1]$ denotes the degree of price indexation to lagged inflation. To estimate this NKPC with Bayesian VAR-GMM, we set the prior distribution for ι to be the beta distribution with mean 0.5 and standard deviation 0.1, as presented in Table 1.

Table 2: Model comparison of Phillips curves based on sticky information and sticky prices

		VAR lag length		
Period	Phillips curve	$n = 1$	$n = 2$	$n = 3$
	DSPC (2)	−40.72	−80.82	−94.72
	SIPC (5)	−55.54	−79.32	−89.58
Great Inflation	Ascari NKPC (6)	−55.15	−64.66	−83.78
(1966:Q1–1982:Q3)	DSPC with indexation (7)	−56.18	−84.20	−95.81
	Textbook NKPC (8)	−59.96	−86.03	−98.73
	Smets-Wouters NKPC (10)	−57.96	−71.75	−106.56
	DSPC (2)	−50.28	−93.28	−108.99
	SIPC (5)	−87.43	−119.51	−144.54
Post-Great Inflation	Ascari NKPC (6)	−64.70	−91.37	−143.14
(1982:Q4–2019:Q4)	DSPC with indexation (7)	−63.43	−106.83	−122.28
	Textbook NKPC (8)	−81.79	−117.82	−144.75
	Smets-Wouters NKPC (10)	−85.90	−117.94	−143.53

Note: The table reports the value of log QML for each Phillips curve with the VAR lag length of $n = 1, 2, 3$ during the Great Inflation period (1966:Q1–1982:Q3) and the period thereafter (1982:Q4–2019:Q4).

Table 2 reports the value of log QML for each Phillips curve with the VAR lag length of $n = 1, 2, 3$. In this table, two findings are detected. First, for each Phillips curve, a VAR lag length of $n = 1$ is selected both during and after the Great Inflation period. On each line of the table, which corresponds to each Phillips curve for each estimation period, the value of log QML is the largest in the case of a VAR lag length of $n = 1$. Second, the DSPC (2) with a VAR lag length of $n = 1$ is the best Phillips curve among all those considered both during and after the Great Inflation period. It has the largest value of log QML among them during both periods. In addition, the Ascari NKPC (6) and the DSPC with indexation (7) have relatively good empirical performance. In contrast, the textbook NKPC (8) and the Smets-Wouters NKPC (10) have relatively poor performance, and so does the SIPC (5) during the post-Great Inflation period. These results suggest that three factors are all needed to better describe inflation dynamics both during and after the Great Inflation period: sticky information, sticky prices, and unchanged prices in each quarter.[14]

4.2 Model comparison of alternative sources of inflation inertia

We have shown that the DSPC (2) with the VAR lag length of one is the best Phillips curve among all those considered, and that the three factors—sticky information, sticky prices, and unchanged prices in each quarter—are needed to better describe inflation dynamics. Dupor et al. (2010) stress that sticky information generates inflation inertia in the DSPC with indexation. This raises the question as to what if an alternative source of inflation inertia is incorporated in the DSPC instead of sticky information.

As the alternative source, Galí and Gertler (1999) propose rule-of-thumb (ROT) price setters, while Kurozumi and Van Zandweghe (2019) suggest variable elasticity of demand (VED) that is based on a non-CES goods aggregator of the sort developed by Kimball (1995).

First, the ROT price setters are considered. Let $\omega_r \in [0, 1)$ denote the fraction of ROT price setters among firms. Under assumption (1), replacing sticky information with ROT price setters in the Calvo staggered price model leads to the NKPC

$$\hat{\pi}_t = \kappa_{b,r}\hat{\pi}_{t-1} + \kappa_{f,r} E_t\hat{\pi}_{t+1} + \kappa_r\left(\hat{ulc}_t - \hat{\Delta}_t\right) + \kappa_{\phi,r}\sum_{j=1}^{\infty}(\beta\lambda\pi^{\theta-1})^j\left(E_t\hat{gy}_{t+j} + \theta E_t\hat{\pi}_{t+j} - E_t\hat{r}_{t+j-1}\right), \tag{11}$$

where the coefficients $\kappa_{b,r}$, $\kappa_{f,r}$, κ_r, and $\kappa_{\phi,r}$ are presented in Appendix B. The law of motion of the relative price distortion $\hat{\Delta}_t$ is the same as (4). The NKPC (11) is called the ROT-NKPC in this paper and can be reduced to the Ascari NKPC (6) in the absence of ROT price setters, i.e., $\omega_r = 0$.

[14] By employing alternative priors (e.g., alternative values of the prior mean and variance of the probability of no price change λ and the probability of no information update ω), we confirmed the robustness of the result that the DSPC (2) with a VAR lag length of $n = 1$ is the best Phillips curve among all those considered both during and after the Great Inflation period.

Next, we turn to the VED. Under assumption

$$\lambda \max\left(\pi^{\gamma}, \pi^{\gamma-1}, \pi^{-1}, \rho_d \beta \pi^{\gamma-1}\right) < 1, \tag{12}$$

introducing the non-CES aggregator instead of sticky information in the Calvo model along the lines of Kurozumi and Van Zandweghe (2019) yields the NKPC

$$\begin{aligned}
\hat{\pi}_t = {} & \frac{\kappa_{\epsilon d}(1+\rho_d \tilde{\kappa}_{d0})}{1-\kappa_{\epsilon d}\tilde{\kappa}_{d0}} \sum_{j=1}^{\infty} \rho_d^{j-1} \hat{\pi}_{t-j} + \frac{\beta\pi(1+\kappa_{\epsilon d})}{1-\kappa_{\epsilon d}\tilde{\kappa}_{d0}} E_t \hat{\pi}_{t+1} + \frac{\kappa_v}{1-\kappa_{\epsilon d}\tilde{\kappa}_{d0}} \left(\hat{ulc}_t - \hat{\Delta}_t\right) \\
& + \frac{\kappa_{\phi,v}}{1-\kappa_{\epsilon d}\tilde{\kappa}_{d0}} \sum_{j=1}^{\infty} (\beta\lambda\pi^{\gamma-1})^j \left[E_t \hat{gy}_{t+j} + \gamma\left(1 + \frac{\kappa_{\varepsilon d}(1-\beta\lambda\pi^{\gamma-1})}{1-\rho_d\beta\lambda\pi^{\gamma-1}}\right) E_t \hat{\pi}_{t+j} - E_t \hat{r}_{t+j-1} \right] \\
& + \frac{\kappa_{\epsilon\psi}}{1-\kappa_{\epsilon d}\tilde{\kappa}_{d0}} \sum_{j=1}^{\infty} (\beta\lambda\pi^{-1})^j \left(E_t \hat{gy}_{t+j} - E_t \hat{r}_{t+j-1} \right)
\end{aligned} \tag{13}$$

and the law of motion of the relative price distortion

$$\hat{\Delta}_t = \rho_s \hat{\Delta}_{t-1} + \kappa_{\Delta}\kappa_s \left[\hat{\pi}_t + \kappa_{\epsilon d} \left(\sum_{j=0}^{\infty} \rho_d^j \hat{\pi}_{t-j} - \sum_{j=0}^{\infty} \rho_d^j \hat{\pi}_{t-j-1} \right) \right],$$

where $\gamma \equiv \theta(1+\epsilon)$, ϵ is the parameter that governs the curvature of demand curves, and the coefficients $\kappa_{\epsilon d}$, ρ_d, $\tilde{\kappa}_{d0}$, κ_v, $\kappa_{\phi,v}$, $\kappa_{\epsilon\psi}$, ρ_s, κ_{Δ}, and κ_s are presented in Appendix B. The NKPC (13) is called the VED-NKPC in this paper and can be reduced to the Ascari NKPC (6) in the case of constant elasticity of demand, i.e., $\epsilon = 0$.

To estimate these NKPCs with Bayesian VAR-GMM, we set the prior distribution for the fraction of ROT price setters ω_r in the ROT-NKPC (11) to be the beta distribution with mean 0.5 and standard deviation 0.1, as presented in Table 1. As for the parameter governing the curvature of demand curves ϵ in the VED-NKPC (13), we set the prior for $-\epsilon$ to be the gamma distribution with mean 3 and standard deviation 1, following Hirose et al. (2021).[15]

Table 3 reports the values of log QML for the DSPC (2), the ROT-NKPC (11), and the VED-NKPC (13) with the VAR lag length of $n = 1, 2, 3$. For each Phillips curve, a VAR

[15] This paper focuses on the case of $\epsilon < 0$ in which relative demand for each good is more price-elastic for an increase in the relative price of the good and less price-elastic for a decrease in the relative price, so that the demand curve has a smoothed-off kink as in Dotsey and King (2005) and Levin et al. (2008). In estimating the VED-NKPC (13), 210,000 MCMC draws that meet assumption (12) are generated.

lag length of $n = 1$ is selected during both estimation periods. More importantly, the DSPC empirically outperforms the ROT-NKPC and the VED-NKPC during both periods. This result suggests that sticky information is a more plausible source of inflation inertia in the Phillips curve than the ROT price setters and VED, and confirms that the three factors of sticky information, sticky prices, and unchanged prices in each quarter are all needed to better describe inflation dynamics both during and after the Great Inflation period. This extends the result of Dupor et al. (2010) in the direction of more consistency with micro evidence on price setting.

Table 3: Model comparison of Phillips curves with alternative sources of inflation inertia

		VAR lag length		
Period	Phillips curve	$n = 1$	$n = 2$	$n = 3$
Great Inflation	DSPC (2)	−40.72	−80.82	−94.72
(1966:Q1–1982:Q3)	ROT-NKPC (11)	−56.07	−79.50	−93.11
	VED-NKPC (13)	−45.62	−61.48	−96.43
Post-Great Inflation	DSPC (2)	−50.28	−93.28	−108.99
(1982:Q4–2019:Q4)	ROT-NKPC (11)	−56.20	−103.10	−135.60
	VED-NKPC (13)	−52.44	−92.69	−122.22

Note: The table reports the values of log QML for each Phillips curve with the VAR lag length of $n = 1, 2, 3$ during the Great Inflation period (1966:Q1–1982:Q3) and the period thereafter (1982:Q4–2019:Q4).

4.3 Posterior estimates of the dual stickiness Phillips curve

We have shown that the DSPC (2) with the VAR lag length of $n = 1$ is the best Phillips curve both during and after the Great Inflation period among all those considered. This subsection analyzes the Phillips curve in detail.

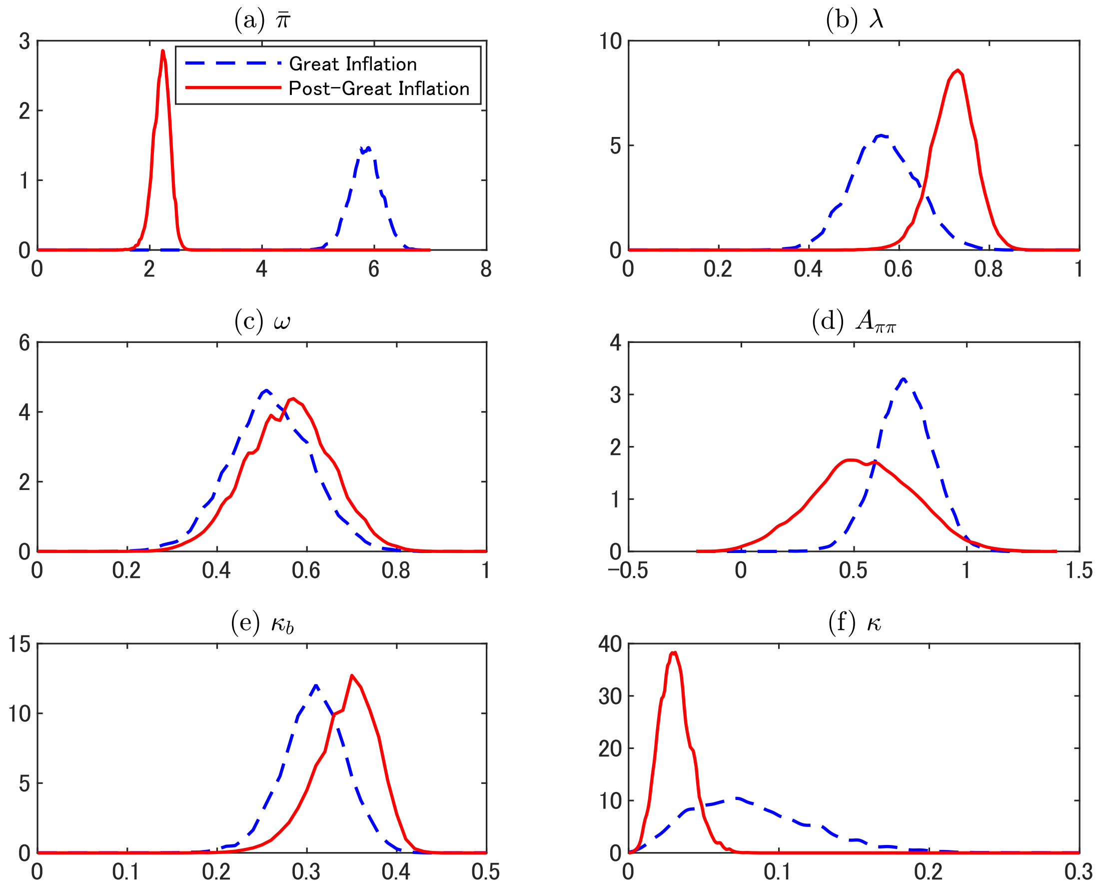
For each of the DSPC parameters and VAR coefficients, its quasi-posterior mean and 90 percent highest quasi-posterior density interval are reported in Table 4. The quasi-posterior mean estimates show that when the annualized steady-state inflation rate $\bar{\pi}$ fell from 5.84 percent during the Great Inflation period to 2.23 percent thereafter, the probability of no price change λ increased from 0.57 to 0.72 and the autoregressive coefficient $A_{\pi,\pi}$ in the VAR's inflation equation—which captures the persistence in the formation of inflation expectations—decreased from 0.72 to 0.52, while the probability of no information update ω rose slightly from 0.52 to 0.56. Then, the DSPC's slope coefficient κ diminished from 0.08

Table 4: Quasi-posterior estimates of the DSPC parameters and VAR coefficients

	Great Inflation (1966:Q1–1982:Q3)		Post-Great Inflation (1982:Q4–2019:Q4)	
	Mean	90% interval	Mean	90% interval
$\bar{\pi}$	5.84	$[5.40, 6.30]$	2.21	$[1.98, 2.46]$
λ	0.57	$[0.44, 0.68]$	0.72	$[0.64, 0.80]$
ω	0.52	$[0.36, 0.66]$	0.56	$[0.41, 0.72]$
$A_{\pi,\pi}$	0.72	$[0.52, 0.93]$	0.54	$[0.17, 0.90]$
$A_{\pi,ulc}$	-0.01	$[-0.07, 0.05]$	-0.02	$[-0.03, 0.00]$
$A_{\pi,gy}$	-0.01	$[-0.09, 0.06]$	0.04	$[-0.04, 0.11]$
$A_{\pi,r}$	0.07	$[-0.09, 0.23]$	0.12	$[-0.04, 0.26]$
$A_{ulc,\pi}$	-0.34	$[-0.72, 0.05]$	0.01	$[-0.94, 1.01]$
$A_{ulc,ulc}$	0.79	$[0.68, 0.91]$	0.94	$[0.91, 0.96]$
$A_{ulc,gy}$	0.00	$[-0.14, 0.14]$	0.41	$[0.18, 0.63]$
$A_{ulc,r}$	0.33	$[0.13, 0.52]$	0.28	$[0.01, 0.55]$
$A_{gy,\pi}$	-0.01	$[-0.41, 0.35]$	-0.24	$[-0.96, 0.48]$
$A_{gy,ulc}$	-0.10	$[-0.21, 0.00]$	0.02	$[-0.00, 0.05]$
$A_{gy,gy}$	-0.07	$[-0.23, 0.08]$	0.35	$[0.16, 0.54]$
$A_{gy,r}$	-0.54	$[-0.80, -0.27]$	0.08	$[-0.17, 0.32]$
$A_{r,\pi}$	0.18	$[0.07, 0.30]$	0.04	$[-0.12, 0.21]$
$A_{r,ulc}$	-0.02	$[-0.06, 0.02]$	0.00	$[-0.01, 0.00]$
$A_{r,gy}$	0.03	$[-0.03, 0.08]$	0.08	$[0.05, 0.10]$
$A_{r,r}$	0.86	$[0.76, 0.95]$	0.97	$[0.91, 1.03]$
κ_b	0.31	$[0.25, 0.36]$	0.34	$[0.29, 0.40]$
κ_f	0.61	$[0.53, 0.70]$	0.62	$[0.55, 0.69]$
κ	0.08	$[0.02, 0.14]$	0.03	$[0.01, 0.05]$
κ_ϕ	0.04	$[0.01, 0.06]$	0.01	$[0.00, 0.02]$
κ_ω	0.03	$[0.01, 0.06]$	0.01	$[0.00, 0.02]$

Note: The table reports the quasi-posterior mean and 90 percent highest quasi-posterior density interval for each of the DSPC parameters and VAR coefficients during the Great Inflation period (1966:Q1–1982:Q3) and the period thereafter (1982:Q4–2019:Q4).

Figure 1: Quasi-posterior distribution of the DSPC parameters and key coefficients



Note: In the figure, panels (a)–(f) display the quasi-posterior distribution during the Great Inflation period 1966:Q1–1982:Q3 (dashed blue line) and the period thereafter 1982:Q4–2019:Q4 (solid red line) of the annualized steady-state inflation rate $\bar{\pi}$, the probability of no price change λ, the probability of no information update ω, the autoregressive coefficient $A_{\pi,\pi}$ in the VAR's inflation equation, the DSPC's inflation inertia coefficient κ_b, and its slope coefficient κ, respectively.

to 0.03, whereas its inflation inertia coefficient κ_b increased slightly from 0.31 to 0.34. These shifts are also detected in the quasi-posterior distribution of the DSPC parameters and key coefficients illustrated in Figure 1.

In the estimated DSPC, three points are worth noting. First, the increase in the mean estimate of the probability of no price change from $\lambda = 0.57$ during 1966:Q1–1982:Q3 to $\lambda = 0.72$ during 1982:Q4–2019:Q4 implies an increase in the average duration from 7.0 months to 10.7 months, which is in line with the micro evidence reported by Nakamura et al. (2018) that the duration implied by the frequency of regular price change increased from 7.6 months during 1978–1987 to 9.4 months during 1988–2014.

Second, the estimated DSPC indicates that when steady-state inflation fell after the Great Inflation period, the probability of price change decreased, thereby generating a flattening of the Phillips curve. This empirical finding is in line with the theoretical prediction in the literature on endogenous price stickiness, initially started by Ball et al. (1988) and subsequently developed by Levin and Yun (2007) and Kurozumi (2016). Moreover, the flattening of the DSPC is consistent with the empirical result of Benati (2007), Ball and Mazumder (2011), and the International Monetary Fund (2013) that Phillips curves flattened after the Volcker disinflation.

Third, the decrease in the autoregressive coefficient $A_{\pi,\pi}$ in the VAR's inflation (gap) equation after the Great Inflation period is consistent with the empirical evidence of Cogley et al. (2010) that the persistence of the inflation gap declined after the Volcker disinflation. Thus, the decline in the inflation gap persistence is captured by the decrease in the autoregressive coefficient $A_{\pi,\pi}$ but not by the slight increase in the inflation inertia coefficient κ_b in the DSPC. The decrease in $A_{\pi,\pi}$ may also imply that inflation expectations have become more anchored after the Great Inflation period.

5 Concluding Remarks

This paper has estimated and compared several Phillips curves based on sticky information and sticky prices using Bayesian VAR-GMM. This method derives expectational variables in each Phillips curve from a VAR and estimates the Phillips curve parameters and the VAR coefficients simultaneously. Model selection based on QML has shown that the DSPC, where each period some prices remain unchanged in line with micro evidence, is the best Phillips curve among all those considered, including the DSPC with indexation of Dupor et al. (2010). Moreover, the model comparison has demonstrated that sticky information is a more plausible source of inflation inertia in the Phillips curve than other sources considered in previous studies, such as the ROT price setters and VED. These results suggest that sticky information, sticky prices, and unchanged prices in each quarter are all needed to better describe inflation dynamics both during and after the Great Inflation period.

Our empirical findings suggest that informational frictions are a crucial source of inflation inertia in Phillips curves. As an alternative specification of informational frictions, previous studies propose noisy information (Phelps, 1970; Lucas, 1972; Woodford, 2003) and rational inattention (Sims, 2003; Maćkowiak and Wiederholt, 2009). Using these specifications to extend the DSPC further and estimating and comparing the DSPCs with alternative specifications of informational frictions could be a fruitful agenda for future research.

Appendix

A Derivation of the dual stickiness Phillips curve

The DSPC (2) is derived from a staggered price model of Calvo (1983) in which, each period, some prices remain unchanged, while the other prices are set subject to sticky information.

In the economy there is a representative composite-good producer that combines the output of a continuum of firms $f \in [0,1]$ using the CES production technology $Y_t = \left[\int_0^1 (Y_t(f))^{(\theta-1)/\theta} df\right]^{\theta/(\theta-1)}$, where Y_t is the output of the composite good and $Y_t(f)$ is firm f's output of an individual differentiated good. Given the composite good's price P_t and individual goods' prices $\{P_t(f)\}$, the composite-good producer maximizes its profit $P_t Y_t - \int_0^1 P_t(f) Y_t(f)\, df$ subject to the CES production technology. The first-order condition for profit maximization yields the demand curve for each individual good

$$Y_t(f) = Y_t \left(\frac{P_t(f)}{P_t}\right)^{-\theta}, \tag{A1}$$

and thus the CES production technology leads to the composite good's price equation

$$P_t = \left[\int_0^1 (P_t(f))^{1-\theta} df\right]^{\frac{1}{1-\theta}}. \tag{A2}$$

Each firm f produces one kind of differentiated good $Y_t(f)$ using the Cobb–Douglas production technology $Y_t(f) = A_t (K_t(f))^{\alpha} (l_t(f))^{1-\alpha}$, where A_t is total factor productivity (TFP) and is assumed to be identical across firms, $K_t(f)$ and $l_t(f)$ are firm f's capital and labor inputs, and $\alpha \in (0,1)$ is the capital elasticity of output. The TFP is assumed to follow the nonstationary stochastic process

$$\log (A_t)^{\frac{1}{1-\alpha}} = \log gy + \log (A_{t-1})^{\frac{1}{1-\alpha}} + \varepsilon_t, \tag{A3}$$

where gy is the steady-state rate of technological change $(A_t/A_{t-1})^{1/(1-\alpha)}$, which coincides with the steady-state rate of output growth $gy_t = Y_t/Y_{t-1}$, and ε_t is an i.i.d. technology shock. In the presence of the economy-wide factor markets with the capital rental rate $P_t r_{k,t}$ and the wage rate $P_t W_t$, the firm minimizes its cost $P_t r_{k,t} K_t(f) + P_t W_t l_t(f)$ subject to the Cobb–Douglas production technology. Combining the first-order conditions for cost minimization shows that all firms face the same real marginal cost mc_t, so that aggregating the labor input condition $W_t l_t(f) = (1-\alpha) mc_t Y_t(f)$ over firms $f \in [0,1]$ and using the demand curve (A1) leads to

$$mc_t = \frac{\int_0^1 W_t l_t(f)\, df}{(1-\alpha)\int_0^1 Y_t(f)\, df} = \frac{W_t l_t}{(1-\alpha) Y_t \Delta_t} = \frac{ulc_t}{(1-\alpha)\Delta_t}, \tag{A4}$$

where $l_t \equiv \int_0^1 l_t(f)\, df$ is aggregate labor, $ulc_t \equiv W_t l_t / Y_t$ is the *composite-good-based* real unit labor cost, and

$$\Delta_t \equiv \int_0^1 \left(\frac{P_t(f)}{P_t}\right)^{-\theta} df \tag{A5}$$

is the relative price distortion. It is worth noting that the marginal cost equation (A4) shows that each firm's real marginal cost mc_t is equal to the *individual-good-based* real unit labor cost $\int_0^1 W_t l_t(f)\, df \Big/ \int_0^1 Y_t(f)\, df$ divided by the labor elasticity of output $1-\alpha$. Therefore, the role of the distortion Δ_t in (A4) is to merely shift the basis of the unit labor cost from composite-good to individual-good production.

Firms set their product prices subject to the demand curve (A1) and the real marginal cost (A4) on a staggered basis as in Calvo (1983). In each period, a fraction $\lambda \in [0,1)$ of firms keeps prices unchanged, while the other firms set prices subject to sticky information along the lines of Mankiw and Reis (2002); that is, with probability $1-\omega \in (0,1]$, each of the price-setting firms updates its information set and maximizes the relevant profit $E_t \sum_{j=0}^{\infty} \lambda^j M_{t,t+j} \left(P_t(f) - P_{t+j} mc_{t+j}\right) Y_{t+j} \left(P_t(f)/P_{t+j}\right)^{-\theta}$, where $M_{t,t+j}$ is the nominal stochastic discount factor between period t and period $t+j$, which meets $M_{t,t+j} = \prod_{k=1}^{j} M_{t+k-1,t+k}$. The first-order condition for profit maximization can be written as

$$E_t \sum_{j=0}^{\infty} \lambda^j \prod_{k=1}^{j} M_{t+k-1,t+k}\, \pi_{t+k}\, g y_{t+k}\, (p_t^*)^{-\theta}\, \pi_{t+k}^{\theta} \left(p_t^* \prod_{k=1}^{j} \pi_{t+k}^{-1} - \frac{\theta}{\theta-1} mc_{t+j} \right) = 0, \tag{A6}$$

where $p_t^* \equiv P_t^*/P_t$, P_t^* is the price set by firms that adjust prices conditional on the contemporaneous information set in period t, and $\pi_t \equiv P_t/P_{t-1}$. Under the aforementioned price setting, the composite good's price equation (A2) and the relative price distortion equation

(A5) can be reduced to, respectively,

$$1 = \lambda \pi_t^{\theta-1} + (1-\lambda)(1-\omega) \sum_{j=0}^{\infty} \omega^j \left[E_{t-j} \left(p_t^* \prod_{k=0}^{j-1} \pi_{t-k} \right) \prod_{k=0}^{j-1} \pi_{t-k}^{-1} \right]^{1-\theta}, \tag{A7}$$

$$\Delta_t = \lambda \pi_t^{\theta} \Delta_{t-1} + (1-\lambda)(1-\omega) \sum_{j=0}^{\infty} \omega^j \left[E_{t-j} \left(p_t^* \prod_{k=0}^{j-1} \pi_{t-k} \right) \prod_{k=0}^{j-1} \pi_{t-k}^{-1} \right]^{\theta}. \tag{A8}$$

In the presence of one-period nominal bonds, the nominal interest rate r_t satisfies

$$1 = E_t(M_{t,t+1}\, r_t)\,. \tag{A9}$$

Let $\beta_{t,t+1} \equiv M_{t,t+1}\, \pi_{t+1}\, (A_{t+1}/A_t)^{1/(1-\alpha)}$ and $\tilde{\omega} \equiv \omega[1 - \lambda\pi^{\theta-1}(1-\beta\lambda\pi^{\theta})]$. Then, log-linearizing (A3), (A6), (A7), and (A9) under assumption (1) and combining the resulting equations give rise to the DSPC (2), where the coefficients κ_b, κ_f, κ, κ_ϕ, and κ_ω are given by

$$\kappa_b \equiv \frac{\omega\lambda\pi^{\theta-1}}{\lambda\pi^{\theta-1} + \tilde{\omega}}, \quad \kappa_f \equiv \frac{\beta\lambda\pi^{\theta}}{\lambda\pi^{\theta-1} + \tilde{\omega}}, \quad \kappa \equiv \frac{(1-\lambda\pi^{\theta-1})(1-\beta\lambda\pi^{\theta})(1-\omega)(1+\omega\beta\lambda\pi^{\theta})}{\lambda\pi^{\theta-1} + \tilde{\omega}},$$
$$\kappa_\phi \equiv \frac{\{\pi[1+\omega(1-\beta\lambda\pi^{\theta-1})] - 1\}(1-\omega)(1-\lambda\pi^{\theta-1})}{\lambda\pi^{\theta-1} + \tilde{\omega}}, \quad \kappa_\omega \equiv \frac{\omega(1-\omega)(1-\lambda\pi^{\theta-1})(1-\beta\lambda\pi^{\theta})}{\lambda\pi^{\theta-1} + \tilde{\omega}}.$$

In addition, the marginal cost equation (3) can be obtained from (A4), while the law of motion of the relative price distortion (4) can be derived by combining (A7) and (A8), and its coefficients ρ_Δ and κ_Δ are given by

$$\rho_\Delta \equiv \lambda\pi^{\theta}, \quad \kappa_\Delta \equiv \frac{\theta\lambda\pi^{\theta-1}(\pi-1)}{1-\lambda\pi^{\theta-1}}.$$

B Coefficients of Phillips curves with inflation inertia

This appendix presents coefficients of the ROT-NKPC (11) and the VED-NKPC (13). In the ROT-NKPC, the coefficients $\kappa_{b,r}$, $\kappa_{f,r}$, κ_r, and $\kappa_{\phi,r}$ are given by

$$\kappa_{b,r} \equiv \frac{\omega_r}{\lambda\pi^{\theta-1} + \tilde{\omega}_r}, \quad \kappa_{f,r} \equiv \frac{\beta\lambda\pi^{\theta}}{\lambda\pi^{\theta-1} + \tilde{\omega}_r}, \quad \kappa_r \equiv \frac{(1-\lambda\pi^{\theta-1})(1-\beta\lambda\pi^{\theta})(1-\omega_r)}{\lambda\pi^{\theta-1} + \tilde{\omega}_r},$$
$$\kappa_{\phi,r} \equiv \frac{(\pi-1)(1-\lambda\pi^{\theta-1})(1-\omega_r)}{\lambda\pi^{\theta-1} + \tilde{\omega}_r}, \quad \tilde{\omega}_r \equiv \omega_r[1 - \lambda\pi^{\theta-1}(1-\beta\pi)].$$

In the VED-NKPC, the coefficients $\kappa_{\epsilon d}$, ρ_d, $\tilde{\kappa}_{d0}$, κ_v, $\kappa_{\phi,v}$, $\kappa_{\epsilon\psi}$, ρ_s, κ_Δ, and κ_s are given by

$$\kappa_{\epsilon d} \equiv -\frac{\epsilon_1 \lambda \pi^{-1}(\pi^\gamma - 1)}{(1+\epsilon_1)(1-\lambda\pi^{-1})}, \quad \rho_d \equiv \frac{\lambda\pi^{-1}(1+\epsilon_1\pi^\gamma)}{1+\epsilon_1},$$
$$\tilde{\kappa}_{d0} \equiv \kappa_{d0} + \rho_d \beta \left[\pi + \frac{\lambda\pi^{\gamma-1}\kappa_{\phi,v}\gamma(1-\beta\lambda\pi^{\gamma-1})}{1-\rho_d\beta\lambda\pi^{\gamma-1}}\right], \quad \kappa_v \equiv \frac{(1-\lambda\pi^{\gamma-1})(1-\beta\lambda\pi^\gamma)}{\lambda\pi^{\gamma-1}[1-\epsilon_2\gamma/(\gamma-1-\epsilon_2)]},$$
$$\kappa_{\phi,v} \equiv \frac{(\pi-1)(1-\lambda\pi^{\gamma-1})}{\lambda\pi^{\gamma-1}[1-\epsilon_2(1+\gamma)/(\gamma-1)]}, \quad \kappa_{\epsilon\psi} \equiv -\frac{\epsilon_2(\pi^{1+\gamma}-1)(1-\lambda\pi^{\gamma-1})}{\lambda\pi^{\gamma-1}[\gamma-1-\epsilon_2(1+\gamma)]},$$
$$\rho_s \equiv \lambda\pi^\gamma, \quad \kappa_\Delta \equiv \left(1+\epsilon_1\frac{1-\lambda\pi^\gamma}{1-\lambda}\right)^{-1}, \quad \kappa_s \equiv \frac{\gamma\lambda\pi^{\gamma-1}(\pi-1)}{1-\lambda\pi^{\gamma-1}},$$

where ϵ_1, ϵ_2, and κ_{d0} are given by

$$\epsilon_1 \equiv \epsilon\left(\frac{1-\lambda\pi^{\gamma-1}}{1-\lambda}\right)^{\frac{\gamma}{1-\gamma}}, \quad \epsilon_2 \equiv \epsilon_1\frac{1-\beta\lambda\pi^{\gamma-1}}{1-\beta\lambda\pi^{-1}}, \quad \kappa_{d0} \equiv \gamma(\kappa_v - \tilde{\kappa}) - \frac{1}{\lambda\pi^{\gamma-1}} - \beta\lambda\pi^\gamma,$$
$$\tilde{\kappa} \equiv \frac{(1-\lambda\pi^{\gamma-1})(1-\beta\lambda\pi^{\gamma-1})}{\lambda\pi^{\gamma-1}[1-\epsilon_2(1+\gamma)/(\gamma-1)]}.$$

C Block Metropolis-Hastings algorithm in Bayesian VAR-GMM estimation of Phillips curves

This appendix explains the procedure of the Block Metropolis-Hastings algorithm in the Bayesian VAR-GMM estimation of Phillips curves. Specifically, this algorithm is applied to obtain the quasi-posterior distribution of the Phillips curve parameters ϑ and the VAR coefficients $vec(A)$. The algorithm is a natural application of the standard Block Metropolis-Hastings algorithm described in Herbst and Schorfheide (2015) but with clear-cut parameter blocks in light of the fundamental property of VAR-GMM. Two blocks are set: one for the VAR coefficients and the other for the Phillips curve parameters. The algorithm consists of the following steps:

1. Estimate a VAR solely to obtain the quasi-posterior mean A_{-1} and the quasi-posterior variance Σ_{-1}. Initialize $vec(A_0)$ at $vec(A_{-1})$.

2. Initialize ϑ_0 at their quasi-posterior mode $\hat{\vartheta}$, fixing $vec(A)$ at $vec(A_{-1})$. This requires numerical maximization of their log quasi-posterior probability density.

3. Apply the Block Metropolis-Hastings algorithm. First, draw candidate values $vec(\tilde{A})$ of the VAR coefficients from a Gaussian proposal distribution with mean $vec(A_{j-1})$ and variance $c_1^2\Sigma_{-1}$, where $vec(A_{j-1})$ is the previous draw of $vec(A)$ and c_1 is the scaling parameter chosen to obtain an acceptance rate of approximately 25 percent.

   Set
$$A_j = \begin{cases} \tilde{A} & \text{with probability } \alpha_{1,j} \\ A_{j-1} & \text{with probability } 1-\alpha_{1,j}, \end{cases}$$
   where
$$\alpha_{1,j} = \min\left\{1, \frac{p(\vartheta_{j-1}, vec(\tilde{A})|Y)}{p(\vartheta_{j-1}, vec(A_{j-1})|Y)}\right\}.$$

4. Then, draw candidate values $\tilde{\vartheta}$ of the Phillips curve parameters from a Gaussian proposal distribution with mean ϑ_{j-1} and variance $c_2^2\hat{\Sigma}$, where ϑ_{j-1} is the previous draw of ϑ, $\hat{\Sigma}$ is the negative of the inverse Hessian of the log quasi-posterior probability density of ϑ evaluated at $\hat{\vartheta}$, calculated as
$$\hat{\Sigma} = -\left(\left.\frac{\partial^2 \log(p(\vartheta, vec(A_{-1})|Y))}{\partial\vartheta\partial\vartheta'}\right|_{\vartheta=\hat{\vartheta}}\right)^{-1}$$
   and c_2 is the scaling parameter chosen to obtain an acceptance rate of approximately 25 percent.

   Set
$$\vartheta_j = \begin{cases} \tilde{\vartheta} & \text{with probability } \alpha_{2,j} \\ \vartheta_{j-1} & \text{with probability } 1-\alpha_{2,j}, \end{cases}$$
   where
$$\alpha_{2,j} = \min\left\{1, \frac{p(\tilde{\vartheta}, vec(A_j)|Y)}{p(\vartheta_{j-1}, vec(A_j)|Y)}\right\}.$$

5. Increment j to $j+1$ and return to step 3.

6. Repeat from step 3 to step 5. Discard a certain number of first draws as a burn-in and use the remaining draws to obtain the quasi-posterior distribution of the Phillips curve parameters and the VAR coefficients.

References

Andrews, Donald W. K. (1991). "Heteroskedasticity and autocorrelation consistent covariance matrix estimation." *Econometrica*, 59(3), pp. 817–858. doi:/10.2307/2938229.

Ascari, Guido (2004). "Staggered prices and trend inflation: Some nuisances." *Review of Economic Dynamics*, 7(3), pp. 642–667. doi:10.1016/j.red.2004.01.002.

Ascari, Guido and Argia M. Sbordone (2014). "The macroeconomics of trend inflation." *Journal of Economic Literature*, 52(3), pp. 679–739. doi:10.1257/jel.52.3.679.

Ball, Laurence, N. Gregory Mankiw, and David Romer (1988). "The New Keynesian economics and the output-inflation trade-off." *Brookings Papers on Economic Activity*, 19(1), pp. 1–65. doi:10.2307/2534424.

Ball, Laurence and Sandeep Mazumder (2011). "Inflation dynamics and the Great Recession." *Brookings Papers on Economic Activity*, 42(1), pp. 337–381. doi:10.1353/eca.2011.0005.

Benati, Luca (2007). "The time-varying Phillips correlation." *Journal of Money, Credit and Banking*, 39(5), pp. 1275–1283. doi:10.1111/j.1538-4616.2007.00066.x.

Calvo, Guillermo A. (1983). "Staggered prices in a utility-maximizing framework." *Journal of Monetary Economics*, 12(3), pp. 383–398. doi:10.1016/0304-3932(83)90060-0.

Canova, Fabio (2007). *Methods for Applied Macroeconomic Research*. Princeton University Press. doi:10.1515/9781400841028.

Carroll, Christopher D. (2003). "Macroeconomic expectations of households and professional forecasters." *Quarterly Journal of Economics*, 118(1), pp. 269–298. doi:10.1162/00335530360535207.

Chernozhukov, Victor and Han Hong (2003). "An MCMC approach to classical estimation." *Journal of Econometrics*, 115(2), pp. 293–346. doi:10.1016/S0304-4076(03)00100-3.

Cogley, Timothy, Giorgio E. Primiceri, and Thomas J. Sargent (2010). "Inflation-gap persistence in the US." *American Economic Journal: Macroeconomics*, 2(1), pp. 43–69. doi:10.1257/mac.2.1.43.

Cogley, Timothy and Argia M. Sbordone (2008). "Trend inflation, indexation, and inflation persistence in the New Keynesian Phillips curve." *American Economic Review*, 98(5), pp. 2101–2126. doi:10.1257/aer.98.5.2101.

Coibion, Olivier (2010). “Testing the sticky information Phillips curve.” *Review of Economics and Statistics*, 92(1), pp. 87–101. doi:10.1162/rest.2009.11148.

Coibion, Olivier and Yuriy Gorodnichenko (2011). “Strategic interaction among heterogeneous price-setting in an estimated DSGE model.” *Review of Economics and Statistics*, 93(3), pp. 920–940. doi:10.1162/REST_a_00102.

Dotsey, Michael and Robert G. King (2005). “Implications of state-dependent pricing for dynamic macroeconomic models.” *Journal of Monetary Economics*, 52(1), pp. 213–242. doi:10.1016/j.jmoneco.2004.10.004.

Dupor, Bill, Tomiyuki Kitamura, and Takayuki Tsuruga (2010). “Integrating sticky prices and sticky information.” *Review of Economics and Statistics*, 92(3), pp. 657–669. doi:10.1162/rest_a_00017.

Galí, Jordi and Mark Gertler (1999). “Inflation dynamics: A structural econometric analysis.” *Journal of Monetary Economics*, 44(2), pp. 195–222. doi:10.1016/s0304-3932(99)00023-9.

Galí, Jordi, Mark Gertler, and J. David López-Salido (2005). “Robustness of the estimates of the hybrid New Keynesian Phillips curve.” *Journal of Monetary Economics*, 52(6), pp. 1107–1118. doi:10.1016/j.jmoneco.2005.08.005.

Gemma, Yasufumi, Takushi Kurozumi, and Mototsugu Shintani (2017). “Trend inflation and evolving inflation dynamics: A Bayesian GMM analysis of the generalized New Keynesian Phillips curve.” Discussion Paper 2017-E-10, Institute for Monetary and Economic Studies, Bank of Japan. URL https://www.imes.boj.or.jp/research/papers/english/17-E-10.pdf.

Geweke, John (1999). “Using simulation methods for Bayesian econometric models: Inference, development, and communication.” *Econometric Reviews*, 18(1), pp. 1–73. doi:10.1080/07474939908800428.

Guerrieri, Luca, Christopher Gust, and J. David López-Salido (2010). “International competition and inflation: A New Keynesian perspective.” *American Economic Journal: Macroeconomics*, 2(4), pp. 247–280. doi:10.1257/mac.2.4.247.

Herbst, Edward P. and Frank Schorfheide (2015). *Bayesian Estimation of DSGE Models*. Princeton University Press. doi:10.23943/princeton/9780691161082.001.0001.

Hirose, Yasuo, Takushi Kurozumi, and Willem Van Zandweghe (2021). “Inflation gap persistence, indeterminacy, and monetary policy.” Working Paper 21-05, Federal Reserve Bank of Cleveland. doi:10.26509/frbc-wp-202105.

Inoue, Atsushi and Mototsugu Shintani (2018). "Quasi-Bayesian model selection." *Quantitative Economics*, 9(3), pp. 1265–1297. doi:10.3982/QE587.

International Monetary Fund (2013). "The dog that didn't bark: Has inflation been muzzled or was it just sleeping?" In *World Economic Outlook*, chapter 3. International Monetary Fund. doi:10.5089/9781616355555.081.

Jeffreys, Harold (1961). *Theory of Probability*. Clarendon Press, Oxford, third edition. doi:10.1515/9781400841028.

Khan, Hashmat and Zhenhua Zhu (2006). "Estimates of the sticky-information Phillips curve for the United States." *Journal of Money, Credit and Banking*, 38(1), pp. 195–207. URL https://www.jstor.org/stable/3839074.

Kiley, Michael T. (2007). "A quantitative comparison of sticky-price and sticky-information models of price setting." *Journal of Money, Credit and Banking*, 39(Supplement 1), pp. 101–125. URL https://www.jstor.org/stable/4123057.

Kimball, Miles S. (1995). "The quantitative analytics of the basic neomonetarist model." *Journal of Money, Credit and Banking*, 27(4), pp. 1241–1277. doi:10.2307/2078048.

Kleibergen, Frank and Sophocles Mavroeidis (2009). "Weak instrument robust tests in GMM and the New Keynesian Phillips curve." *Journal of Business and Economic Statistics*, 27(3), pp. 293–311. doi:10.1198/jbes.2009.08280.

Kleibergen, Frank and Sophocles Mavroeidis (2014). "Identification issues in limited-information Bayesian analysis of structural macroeconomic models." *Journal of Applied Econometrics*, 29, pp. 1183–1209. doi:10.1002/jae.2398.

Klenow, Peter J. and Jonathan L. Willis (2007). "Sticky information and sticky prices." *Journal of Monetary Economics*, 54(Supplement), pp. 79–99. doi:10.1016/j.jmoneco.2007.06.003.

Knotek, Edward S., II (2010). "A tale of two rigidities: Sticky prices in a sticky-information environment." *Journal of Money, Credit and Banking*, 42(8), pp. 1543–1564. URL https://www.jstor.org/stable/40925702.

Kurozumi, Takushi (2016). "Endogenous price stickiness, trend inflation, and macroeconomic stability." *Journal of Money, Credit and Banking*, 48(6), pp. 1267–1291. doi:10.1111/jmcb.12333.

Kurozumi, Takushi and Willem Van Zandweghe (2019). "A theory of intrinsic inflation persistence." Working Paper 19-16, Federal Reserve Bank of Cleveland. doi:10.26509/frbc-wp-201916.

Levin, Andrew T., J. David López-Salido, Edward Nelson, and Tack Yun (2008). "Macroeconomic equivalence, microeconomic dissonance, and the design of monetary policy." *Journal of Monetary Economics*, 55(Supplement), pp. S48–S62. doi:10.1016/j.jmoneco.2008.07.013.

Levin, Andrew T. and Tack Yun (2007). "Reconsidering the natural rate hypothesis in a New Keynesian framework." *Journal of Monetary Economics*, 54(5), pp. 1344–1365. doi:10.1016/j.jmoneco.2007.06.006.

Lucas, R. E., Jr. (1972). "Expectations and the neutrality of money." *Journal of Economic Theory*, 4(2), pp. 103–124.

Maćkowiak, Bartosz and Mirko Wiederholt (2009). "Optimal sticky prices under rational inattention." *American Economic Review*, 99(3), pp. 769–803. doi:10.1257/aer.99.3.769.

Mankiw, N. Gregory and Ricard Reis (2002). "Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve." *Quarterly Journal of Economics*, 117(4), pp. 1295–1328. doi:10.1162/003355302320935034.

Mavroeidis, Sophocles, Mikkel Plagborg-Møller, and James H. Stock (2014). "Empirical evidence on inflation expectations in the New Keynesian Phillips curve." *Journal of Economic Literature*, 52(1), pp. 124–188. doi:10.1257/jel.52.1.124.

Nakamura, Emi, Jón Steinsson, Patrick Sun, and Daniel Villar (2018). "The elusive costs of inflation: Price dispersion during the U.S. Great Inflation." *Quarterly Journal of Economics*, 133(4), pp. 1933–1980. doi:10.1093/qje/qjy017.

Newey, Whitney K. and Kenneth D. West (1987). "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix." *Econometrica*, 55(3), pp. 703–708. doi:10.2307/1913610.

Phelps, Edmund S. (1970). "Introduction: the new microeconomics in employment and inflation theory." In Edmund S. Phelps et al., editors, *Microeconomic Foundations of Employment and Inflation Theory*, pp. 1–23. Norton.

Sbordone, Argia M. (2002). "Prices and unit labor costs: A new test of price stickiness." *Journal of Monetary Economics*, 49(2), pp. 265–292. doi:10.1016/S0304-3932(01)00111-8.

Sheedy, Kevin D. (2010). "Intrinsic inflation persistence." *Journal of Monetary Economics*, 57(8), pp. 1049–1061. doi:10.1016/j.jmoneco.2010.10.002.

Sims, Christopher A. (2003). "Implications of rational inattention." *Journal of Monetary Economics*, 50(3), pp. 665–690. doi:10.1016/S0304-3932(03)00029-1.

Sims, Christopher A., Daniel F. Waggoner, and Tao Zha (2008). "Methods for inference in large multiple-equation Markov-switching models." *Journal of Econometrics*, 146(2), pp. 255–274. doi:10.1016/j.jeconom.2008.08.023.

Smets, Frank and Rafael Wouters (2007). "Shocks and frictions in US business cycles: A Bayesian DSGE approach." *American Economic Review*, 97(3), pp. 586–606. doi:10.1257/aer.97.3.586.

Woodford, Michael (2003). "Imperfect common knowledge and the effects of monetary policy." In Philippe Aghion, Roman Frydman, Joseph Stiglitz, and Michael Woodford, editors, *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, pp. 25–58. Princeton University Press. doi:10.2307/j.ctv18zhdvn.6.

Zbaracki, Mark J., Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen (2004). "Managerial and customer costs of price adjustment: Direct evidence from industrial markets." *Review of Economics and Statistics*, 86(2), pp. 514–533. doi:10.1162/003465304323031085.