Economic Modelling 105 (2021) 105644

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.journals.elsevier.com/economic-modelling

# On the role of dependence in sticky price and sticky information Phillips curve: Modelling and forecasting☆

Roberto Casarin [a], Mauro Costantini [b,*], Antonio Paradiso [a]

[a] Department of Economics, Ca' Foscari University of Venice, Italy
[b] Department of Industrial and Information Engineering and Economics, University of L'Aquila, Italy

ARTICLE INFO

*JEL classification:*
E31
C11
C32
C53

*Keywords:*
Phillips curve
Dual stickiness
Dependence
Bayesian analysis
Forecasting

ABSTRACT

Understanding the role of sticky price and sticky information for inflation dynamics is a key issue in economics. The literature has treated the two forms of stickiness as independent. This paper proposes a new dual stickiness Phillips curve based on dependence among the events of setting prices and updating information. Using US data over the period 1947Q1–2020Q1, the new model is scrutinized against a dual stickiness model without dependence, a pure sticky price model, and a pure sticky information model, through in- and out-of-sample analyses. The results show: (*i*) the new model outperforms the model without dependence in-sample; (*ii*) the dual stickiness models perform similarly out-of-sample; and (*iii*) the pure sticky models yield the worst forecasts. The results have some implications for policy makers and practitioners. A policy maker may consider the new model given its performance in- and out-of-sample, while a practitioner may prefer the model without dependence, given its lesser complexity and its competitive forecasting performance.

## 1. Introduction

Understanding the inflation dynamics is one of the main issues in economics. Over the last 20 years, models based on price or information rigidities (see, for example, Galí and Gertler, 1999; Mankiw and Reis, 2002; Coibion, 2006; Carrera and Ramírez-Rondán, 2019; Bilbiie, 2021), on both rigidities but taken separately (see Keen, 2007; Kiley, 2007; Coibion, 2010; Carrillo, 2012), and on the combination of the two form of stickiness (Klenow and Willis, 2007; Knotek, 2010; Dupor et al., 2010; Arslan, 2010; Coibion and Gorodnichenko, 2011; Kim and Kim, 2019), have been proposed to study inflation dynamics.

The literature on dual stickiness Phillips curve has offered a fundamental contribution to the interpretation of inflation dynamics, both theoretically and empirically. In particular, Klenow and Willis (2007) simulated a general equilibrium model featuring state-dependent sticky price and sticky information. To keep the model tractable, they assumed that information regarding macro state variables arrives exogenously in a staggered fashion. Their results show that price changes reflect old inflation innovations. Knotek (2010) proposed a micro-founded Phillips curve that relies on both form of stickiness and is on based on state-dependent pricing decisions (as in Klenow and Willis, 2007).[1] In the empirical application, Knotek (2010) shows that the sticky information plays a role in explaining the dynamics of inflation at both micro and macro level for the US over the period 1983–2005.

Dupor et al. (2010) proposed a time-dependent approach to model a dual stickiness Phillips curve (DS-PC) and show that both rigidities are present in US data for the period 1960–2007. Similarly, Arslan (2010) derived a DS-PC using a time-dependent approach. The estimates of the structural model show that both kind of stickiness are statistically and quantitatively important for price setting in the US over the period 1960–2007. Coibion and Gorodnichenko (2011) estimated a DSGE

☆ The authors thank the Editor and two anonymous Reviewers for their valuable comments and suggestions. The authors also thank Monica Billio, Josep Lluís Carrion-i-Silvestre, Michele Costola, Saten Kumar, and Francesco Ravazzolo for helpful comments. R. Casarin and A. Paradiso thank VERA center at Ca' Foscari University of Venice for supporting open access publication.

* Corresponding author. Department of Industrial and Information Engineering and Economics, University of L'Aquila, Via Giuseppe Mezzanotte, 67100, L'Aquila, Italy.

*E-mail addresses:* r.casarin@unive.it (R. Casarin), mauro.costantini@univaq.it (M. Costantini), antonio.paradiso@unive.it (A. Paradiso).

[1] Time-dependent pricing models assume an exogenously given probability of price adjustment and/or information upgrading. In state-dependent pricing models, firms choose when to change price or update information subject to menu or information costs.

https://doi.org/10.1016/j.econmod.2021.105644
Received 15 April 2021; Received in revised form 2 September 2021; Accepted 7 September 2021
Available online 16 September 2021
0264-9993/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

model with different forms of time-dependent price setting regimes. They find that the model embodying sticky price and sticky information outperforms a pure sticky price model and a pure sticky information model for the US over the period 1984–2008. Finally, Kim and Kim (2019) studied the dynamics of inflation using a Bayesian DSGE model with (time-dependent) dual stickiness. The empirical results based on an historical decomposition analysis reveal that supply and inflation target shocks are found to be dominant forces driving long-run changes in inflation rate for the US over the period 1968–2008.

In the dual stickiness models, firms may change prices and update information. However, the aforementioned studies have ruled out any dependence among the events of setting price and updating information. This is primarily due to the mathematical tractability of the solution, as pointed out by Arslan (2010) and Dupor et al. (2010). This paper fills the gap and proposes a new DS-PC model (with time-dependent pricing rule) for inflation that assumes dependence among the events of changing price and updating information through a bivariate Bernoulli distribution. This is the first contribution of the paper.

The assumption of dependence among the events of setting price and updating information relies on rational arguments. Under conditions of uncertainty, that is absence of adequate information, economic agents tend to maintain the *status quo* when making decisions concerning target variables (see Samuelson and Zeckhauser, 1988; Sautua, 2017; Foellmi et al., 2019). When the target variable is price, then price stickiness and information stickiness may be correlated, and firms modify the target value of price (expected optimal price) only when the available information is adequately updated, given the fact that firms do not know the probability distribution of all the potential outcomes (Bewley, 2002). Therefore, firms anchor the price to the value previously set to not incur in potential losses, if they are not fully and adequately informed, and the events of (not) updating price and of (not) updating information turn out to be positively correlated. This paper assumes positive correlation among the events of adjusting price and updating information. In this respect, Ball and Mankiw (1994) argue that firms adjust prices only after updating the information, and the stickiness of price is mainly due to the costs of gathering information.[2] Further, Zbaracki et al. (2004) and Harris et al. (2020) provide some empirical evidence on the correlation among the events of chancing price and updating information using firms data for the US and France, respectively.

Using US data over the period 1947Q1–2020Q1, this paper focuses on in- and out-of-sample analyses for inflation by considering the GDP deflator as a measure of inflation and the labour share (a proxy of real marginal costs) as forcing variable.[3] This is the second contribution of the paper.

The in-sample analysis is carried out for the new DS-PC model with dependence and the DS-PC model without dependence proposed by Dupor et al. (2010). For the out-of-sample forecasts, two additional models, such as a pure sticky price model and a pure sticky information model, are also taken into account.

For the estimation of the models, a two-step procedure based on a rolling estimation scheme is adopted. In the first step, expectations of inflation and real marginal costs are estimated by a rolling VAR model. In the second step, the parameters of the different Phillips curves are estimated by Bayesian inference.

The in-sample analysis proceeds as follows. A rolling window estimation of the key parameters of the two DS-PC models is first carried out to ascertain the significance of these parameters. Then, the performance of the two DS-PC models is compared by the Bayes factor.

For the out-of-sample analysis, we use the four different Phillips curves and two different measures of forecast accuracy, namely the mean square error (MSE) and the directional accuracy (DA) by Blaskowitz and Herwartz (2009, 2011). To the best of our knowledge, this is the first paper to use the DA measure to evaluate the forecasts of sticky price and sticky information models. The relevance of the direction of the forecasts is widely documented in forecasting literature. Leitch and Tanner (1995) pointed out that the direction is what mostly concerns entrepreneurs when making a decision on how to invest. Likewise the direction of the forecasts is at core of the decisions of the policy makers (see, for example, Öller and Barot, 2000; Sinclair et al., 2010; Bergmeir et al., 2014; Chen et al., 2016; Costantini et al., 2016).[4]

The main empirical results are as follows. First, the in-sample analysis shows that: (*i*) the parameter that rules the dependence is highly significant, while it tends to reduce after the crisis of 2008; and (*ii*) the new DS-PC model with dependence outperforms the model without dependence. Second, the out-of-sample analysis shows that the DS-PC models perform similarly in terms of both MSE and DA, and the two DS-PC models outperform the pure sticky price and pure sticky information models, with the latter model being the worst performer.

For both in- and out-of-sample analysis, a set of robustness checks is performed. First, diverse combinations of the rolling window for the expectations and the estimation of the models are considered. Second, a different measure of inflation (CPI) and a forcing variable (output gap) are used. The results show that the parameter of dependence is always statistically significant and the DS-PC models are equally competitive in terms of forecasts and outperform the pure sticky price and pure sticky information models. These findings confirm the main conclusions.

The empirical findings may have some implications for policy makers and practitioners. A policy maker, who is interested in the effectiveness of policy interventions and in the predictions of inflation, may favour the use of the dual stickiness model with dependence given its in-an out-of-sample performance. Further, given the relevance of the parameter of dependence for firms, policy interventions should look at the dynamics of dependence parameter. On the other hand, a practitioner may advocate the use of the dual stickiness model without dependence given its lesser complexity and its competitive forecasting performance.

The rest of the paper is organized as follows. Section 2 presents the new DS-PC model with dependence. Section 3 describes the data and methodology. Section 4 presents and discusses the empirical results. Section 5 concludes.

## 2. Dependence in sticky price and sticky information Phillips curve

Consider a continuum of firms engaged in monopolistic competition. Suppose that each firm updates information and adjusts prices infrequently (dual stickiness). Let $Z_1$ and $Z_2$ be the price adjustment and information updating events, respectively. If $(Z_1, Z_2)$ has a bivariate Bernoulli distribution, then there are four possible events, (1,1), (1,0), (0,1), (0,0), with the following probabilities

$$\begin{aligned} P(Z_1 = 1, Z_2 = 1) &= p_{11} = (1-\gamma)(1-\phi) + \delta \\ P(Z_1 = 1, Z_2 = 0) &= p_{10} = (1-\gamma)\,\phi - \delta \\ P(Z_1 = 0, Z_2 = 1) &= p_{01} = \gamma\,(1-\phi) - \delta \\ P(Z_1 = 0, Z_2 = 0) &= p_{00} = \gamma\phi + \delta, \end{aligned} \tag{1}$$

where $p_{11} + p_{10} + p_{01} + p_{00} = 1$. In our setting, $p_{11}$ indicates the probability that a firm updates both price and information, $p_{00}$ is

[2] "The most important cost of price adjustment are the time and attention required of managers to gather the relevant information and make and implement decisions" (Ball and Mankiw, 1994, p.24-25).

[3] The forcing variable represents the excess of demand. In the empirical literature, it is measured as output gap (difference between real and potential output) or labour share (a proxy of real marginal costs).

[4] For example, monetary policy interventions change as a consequence of unexpected inflation shocks (see, for example, Mallick and Sousa, 2013).

the probability that neither price nor information are updated, and $p_{10}$ and $p_{01}$ are probabilities of updating only price and only information, respectively.

The parameter $\delta$ captures the dependence among the probability related to the events of (not) changing prices $1-\gamma$ ($\gamma$) and (not) updating information $1-\phi$ ($\phi$).[5] Firms usually operate in a context where both the aforementioned events may be correlated: firms may decide to not changing the (target) price if not adequately informed, whereas they are inclined to revise prices if they are fully informed (Ball and Mankiw, 1994). It is easy to show that when $\delta = 0$ (independence among events), our parametrization collapses into that of Dupor et al. (2010).

In detail, the marginal distributions are

$$P(Z_1 = 1) = p_{1+} = p_{10} + p_{11} = 1-\gamma$$

$$P(Z_1 = 0) = p_{0+} = p_{00} + p_{01} = \gamma$$

$$P(Z_2 = 1) = p_{+1} = p_{11} + p_{01} = 1-\phi$$

$$P(Z_2 = 0) = p_{+0} = p_{10} + p_{00} = \phi.$$

In our model, the dependence between price and information stickiness is ruled by (see Marshall and Olkin, 1985)

$$\begin{aligned} Cov(Z_1, Z_2) &= E(Z_1 \cdot Z_2) - E(Z_1)\cdot E(Z_2) \\ &= p_{11} - (p_{10}+p_{11})(p_{01}+p_{11}) \\ &= p_{11}p_{00} - p_{10}p_{01} = \delta. \end{aligned} \tag{2}$$

Using Fréchet bound inequalities (Joe, 2014, p. 48–49), and the fact that covariance is assumed to be positive (see section 1), we can define the lower and upper bound of $\delta$[6]

$$0 \leqslant \delta \leqslant \min\{(1-\gamma)\phi, (1-\phi)\gamma\}. \tag{3}$$

Using the marginal probabilities and equation (2), the dynamics of the log aggregate price level $p_t$ can be described by

$$\begin{aligned} p_t &= (p_{00}+p_{01})p_{t-1} + p_{11}p_t^f + p_{10}p_t^b \\ &= [(\gamma\phi+\delta) + [\gamma(1-\phi)-\delta]]\, p_{t-1} + \\ &\quad [(1-\gamma)(1-\phi)+\delta]\, p_t^f + [(1-\gamma)\phi - \delta]\, p_t^b \\ &= \gamma p_{t-1} + (1-\gamma) q_t, \end{aligned} \tag{4}$$

where $q_t = (1-\phi)p_t^f + \phi p_t^b + \frac{\delta}{1-\gamma}\left(p_t^f - p_t^b\right)$. Equation (4) says that the log aggregate price can be expressed as a weighted average of the price observed in the previous period, $p_{t-1}$, and the price index for all newly set prices at time $t$, $q_t$. In particular, $q_t$ is the sum of $p_t^f$ (the price set on the basis of new information), $p_t^b$ (the price set on the old information) and $\delta/(1-\gamma)$, which captures the dependence among events for new and old informed firms.

The prices $p_t^f$ and $p_t^b$ are given by

$$\begin{aligned} p_t^f &= (1-\gamma)E_t p_t^* + (1-\gamma)\gamma E_t p_{t+1}^* + (1-\gamma)\gamma^2 E_t p_{t+2}^* + \dots \\ &= (1-\gamma)\sum_{j=0}^{\infty} \gamma^j E_t p_{t+j}^* \end{aligned} \tag{5}$$

$$\begin{aligned} p_t^b &= (1-\phi)E_{t-1}p_t^f + (1-\phi)\phi E_{t-2}p_t^f + (1-\phi)\phi^2 E_{t-3}p_t^f + \dots \\ &= (1-\phi)\sum_{k=0}^{\infty} \phi^k E_{t-k-1} p_t^f, \end{aligned} \tag{6}$$

where $E_t(\cdot) = E(\cdot|\mathcal{F}_t)$ indicates the conditional expectation given the information set $\mathcal{F}_t$ available at time $t$ and $p_t^*$ represents the desired price at time $t$. Equation (5) states that firms with zero period old information account for the expected future path of desired prices given the likelihood that the price may remain fixed for multiple periods ($\gamma$). Instead, equation (6) collects the individual optimal prices conditional on old information set; it consists of a weighted average of each individual price for inattentive firms with old information sets $\mathcal{F}_{t-k}$, $k \geq 1$.

After some algebra (see A.1 in Appendix A), one can obtain the following equation for inflation

$$\begin{aligned} \frac{\gamma+\phi(1-\gamma)}{1-\gamma}\pi_t &= (1-\phi)\left(p_t^f - p_t\right) + \frac{\gamma}{1-\gamma}\phi\pi_{t-1} + (1-\phi)\left[\phi - \frac{\delta}{1-\gamma}\right] \\ &\times \sum_{k=0}^{\infty}\phi^k E_{t-k-1}\Delta p_t^f \ + \delta\Delta p_t^* + \delta\gamma\sum_{j=0}^{\infty}\gamma^j E_t \Delta p_{t-j+1}^* + \delta\gamma\tau_t, \end{aligned} \tag{7}$$

where $\tau_t$ is a (random) variable capturing the revision of expectations (see A.1 in Appendix A for more details).

The Phillips curve (PC) can be derived by complementing equation (7) with the optimal price equation that is given by the marginal cost function $p_t^* = mc_t^n = mc_t + p_t$, where $mc_t^n$ denotes the nominal marginal cost. Therefore, equation (7) can be re-written as (see A.2 in Appendix A)

$$\begin{aligned} \pi_t &= \rho^D\pi_{t-1} + \zeta_1^D(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_t\left(mc_{t+j} + \sum_{k=1}^{j}\pi_{t+k}\right) \\ &+ \zeta_2^D(1-\phi)\sum_{k=0}^{\infty}\phi^k\left[(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_{t-k-1}\left(\Delta mc_{t+j} + \pi_{t+j}\right)\right] + \zeta_1^N \Delta mc_t \\ &+ \zeta_2^N(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_t\left(\Delta mc_{t+j+1} + \pi_{t+j+1}\right) + \zeta_2^N\tau_t, \end{aligned} \tag{8}$$

where $\rho^D \equiv \frac{p_{00}-\delta}{1-p_{11}+(p_{00}+p_{01})\delta} \equiv \frac{\phi\gamma}{\gamma+(1-\gamma)(\phi-\delta)}$, $\zeta_1^D \equiv \frac{p_{11}-\delta}{1-p_{11}+(p_{00}+p_{01})\delta} \equiv \frac{(1-\gamma)(1-\phi)}{\gamma+(1-\gamma)(\phi-\delta)}$, $\zeta_2^D \equiv \frac{p_{10}}{1-p_{11}+(p_{00}+p_{01})\delta} \equiv \frac{\phi(1-\gamma)-\delta}{\gamma+(1-\gamma)(\phi-\delta)}$, $\zeta_1^N \equiv \frac{\delta(p_{10}+p_{11})}{1-p_{11}+(p_{00}+p_{01})\delta} \equiv \frac{\delta(1-\gamma)}{\gamma+(1-\gamma)(\phi-\delta)}$, $\zeta_2^N \equiv \frac{\delta(p_{00}+p_{01})}{1-p_{11}+(p_{00}+p_{01})\delta} \equiv \frac{\delta\gamma}{\gamma+(1-\gamma)(\phi-\delta)}$. Equation (8) represents the new DS-PC model with dependence.

When $\delta = 0$, $\zeta_1^N = \zeta_2^N = 0$, equation (8) collapses into the DS-PC model without dependence (i.e., Dupor et al. (2010)'s specification)

$$\begin{aligned} \pi_t &= \rho^D\pi_{t-1} + \zeta_1^D(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_t\left(mc_{t+j} + \sum_{k=1}^{j}\pi_{t+k}\right) \\ &+ \zeta_2^D(1-\phi)\sum_{k=0}^{\infty}\phi^k\left[(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_{t-k-1}\left(\Delta mc_{t+j} + \pi_{t+j}\right)\right]. \end{aligned} \tag{9}$$

Pure sticky price PC is obtained from (8) setting $\delta = 0$ and $\phi = 0$

$$\pi_t = \kappa mc_t + E_t\pi_{t+1}, \tag{10}$$

where $\kappa \equiv \frac{(1-\gamma)^2}{\gamma}$. Since the derivation of (10) is not straightforward, more details are reported in A.3, Appendix A.

Finally, pure sticky information PC is easily obtained from (8) setting $\delta = 0$ and $\gamma = 0$

$$\pi_t = \nu mc_t + (1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k-1}(\Delta mc_t + \pi_t), \tag{11}$$

where $\nu \equiv \frac{(1-\phi)}{\phi}$.

---

5 One may argue that firms are reluctant to change prices due to the presence of imperfect competition or menu costs (see, for example Akerlof and Yellen, 1985; Ball and Mankiw, 1994), and they update the information very rarely because of the presence of costs related to the acquisition of information or the limited capacity of elaborating the information (see, for example Sims, 2003; Reis, 2006).

6 The correlation is expressed as $\rho = \frac{\delta}{\sqrt{(1-\gamma)\gamma(1-\phi)\phi}}$, and the admissible range of correlation is $0 \leqslant \rho \leqslant \min\left\{\frac{\sqrt{(1-\gamma)\phi}}{\sqrt{(1-\phi)\gamma}}, \frac{\sqrt{(1-\phi)\gamma}}{\sqrt{(1-\gamma)\phi}}\right\}$.

## 3. Data and methodology

This section describes the data and the methodology used for the estimation of models (8)–(11) and the comparison among them. More specifically, in section 3.1 the data and the empirical strategy for the in-sample and out-of-sample analyses are presented. Section 3.2 describes the Bayesian inference approach for the estimation of equations 8–11, while sections 3.3 and 3.4 detail the Bayesian approach for the in-sample evaluation and the forecast accuracy measures for the out-of-sample evaluation, respectively.

### 3.1. Data and empirical strategy

All versions of Phillips curves are estimated using the following variables for the US: inflation ($\pi_t$), labour share ($L_t$) – the proxy of real marginal costs (see, for example, Rotemberg and Woodford, 1999; Lawless and Whelan, 2011) – and their expectations when needed (see equations 8–11). For the reconstruction of expectations we consider a VAR model with the following variables: $\pi_t$, $L_t$, output gap ($y_t$), and interest rate spread ($S_t$).[7] US data are at quarterly frequency and cover the period 1947Q1–2020Q1. Inflation is measured as the percentage change from preceding period of GDP implicit price deflator (source: U.S. Bureau of Economic Analysis). Labour share is defined as the amount of income paid out in wages, salaries, and benefits in the US private business sector (source: Giandrea and Sprague (2017) (1947Q1–2013Q4) and U.S. Bureau of Labor Statistics (2014Q1–2020Q1)). Output gap is defined as the different between real GDP (source: U.S. Bureau of Economic Analysis) and a quadratic time trend. Interest rate term spread is defined as the difference between long (10y) and short (3m) term interest rates (source: Federal Reserve Bank of St. Louis). All variables (except for interest rates) are seasonally adjusted.

The in-sample analysis is conducted through a comparison among the new DS-PC model with dependence (equation (8)) and the DS-PC model without dependence (equation (9)). The analysis proceeds as follows. First, we generate the series of expectations of inflation and real marginal costs by estimating a rolling VAR model with $\pi_t$, $L_t$, $y_t$, and $S_t$. The rolling window used to generate the expectations ($w_{ex}$) is equal to 60 observations. The lag order of the VAR is set to two. In the second step, we estimate (8) and (9) using a rolling Bayesian procedure according to the prior specification and the Markov Chain Monte Carlo (MCMC) procedure for posterior approximation given in section 3.2 to check the statistically significance of key parameters.[8] The rolling window used for the estimations of the models ($w_{es}$) is of 40 observations. It is well known (see, for example, Schorfheide, 2010; Bauwens and Korobilis, 2013) that the Bayesian procedure is suitable for small sample. Third, the Bayes factor measure to compare the in-sample performance of the two models is applied (a full description of the Bayes factor is given in section 3.3).

In the out-of-sample analysis, we use the four different Phillips curve models described in section 2, and MSE and DA by Blaskowitz and Herwartz (2009, 2011) for the evaluation of the forecasts (both measures are described in section 3.4). The out-of-sample analysis considers three different out-of-sample periods, 1995Q1–2020Q1, 2000Q1–2020Q1, and 2008Q1–2020Q1, three forecast horizons, $h = 1, 4, 8$, and $w_{ex} = 60$ and $w_{es} = 40$.[9]

A robustness check for the in- and out-of-sample analyses is also carried out. For the in-sample analysis, two exercises are performed: (*i*) we consider two different values of $w_{es}$ equal to 50 and 60, respectively, along with $w_{ex} = 60$ for the estimation of the key parameters of the two dual stickiness models; and (*ii*) GDP deflator is first replaced with consumer price index (CPI), and then output gap is used instead of real marginal costs, both for $w_{ex} = 60$ and $w_{es} = 40$.[10] As for the forecasts, we also consider the following three combinations $w_{ex} = 60$ and $w_{es} = 50$, $w_{ex} = 80$ and $w_{es} = 40$, and $w_{ex} = 80$ and $w_{es} = 50$. Further, we first replace GDP deflator with CPI, and then real marginal costs with output gap, for $w_{ex} = 60$ and $w_{es} = 40$.

### 3.2. Posterior distribution and numerical approximation

Equation (8) can be rewritten as

$$\pi_t = \rho^D \pi_{t-1} + \zeta_1^D \mathbf{b}' X_t + \zeta_2^D \mathbf{c}' \sum_{k=0}^{\infty} \phi^k A^k X_{t-k-1} + \zeta_1^N \mathbf{e}_{mc}' \Delta X_t + \zeta_2^N (1-\gamma) \mathbf{d}' X_t + \zeta_2^N \tau_t, \tag{12}$$

where $\mathbf{b}'$, $\mathbf{c}'$, and $\mathbf{d}'$ corresponds to

$$\mathbf{b}' = \left[(1-\gamma)\, \mathbf{e}_{mc}' + \gamma \mathbf{e}_\pi' A\right] [I - \gamma A]^{-1}$$

$$\mathbf{c}' = (1-\gamma)(1-\phi) \left[\mathbf{e}_{mc}' (A - I) + \mathbf{e}_\pi' A\right] [I - \gamma A]^{-1}$$

$$\mathbf{d}' = \left(\mathbf{e}_{mc}' (A - I) + \mathbf{e}_\pi' A\right) (I - \gamma A)^{-1}$$

and $\mathbf{e}_{mc}$ and $\mathbf{e}_\pi$ are the selection vectors with $3p$ elements as defined in Dupor et al. (2010).

Let $\mathbf{y}_{1:T} = (\pi_1, \dots, \pi_T)'$ be the $T \times 1$ vector of observations, $\mathbf{x}_{1:T} = (X_1, \dots, X_T)$ the $T \times n$ matrix of covariates, and $\boldsymbol{\theta} = (\gamma, \phi, \delta)'$. The likelihood function is

$$L(\mathbf{y}_{1:T}|\boldsymbol{\theta}) = \prod_{t=1}^{T} (2\pi\sigma^2(\boldsymbol{\theta}))^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2\sigma^2(\boldsymbol{\theta})} \varepsilon_t(\boldsymbol{\theta})^2 \right\} \tag{13}$$

where

$$\varepsilon_t(\boldsymbol{\theta}) = \pi_t - \left( \boldsymbol{\beta}_1(\boldsymbol{\theta})' X_t + \boldsymbol{\beta}_2(\boldsymbol{\theta})' X_{t-1} + \sum_{k=1}^{K} \boldsymbol{\beta}_{2+k}(\boldsymbol{\theta})' X_{t-k-1} \right) \tag{14}$$

with $\boldsymbol{\beta}_1(\boldsymbol{\theta}) = \zeta_1^D \mathbf{b} + \zeta_1^N \mathbf{e}_{mc} + \zeta_2^N (1-\gamma)\mathbf{d}$, $\boldsymbol{\beta}_2(\boldsymbol{\theta}) = \rho^D \mathbf{e}_\pi + \zeta_2^D \mathbf{c} - \zeta_1^N \mathbf{e}_{mc}$, $\boldsymbol{\beta}_{2+k}(\boldsymbol{\theta}) = \zeta_2^D \phi^k (A^k)' \mathbf{c}$, $k = 1, \dots, K$, where $K$ indicates the truncation value, and $\sigma(\boldsymbol{\theta}) = \zeta_2^N$. In the empirical analysis $K$ is set equal to 5.[11]

To complete the Bayesian model, we elicit a prior distribution on the parameter vector $\boldsymbol{\theta} = (\gamma, \phi, \delta)'$. We assume the following uniform joint prior distribution $\pi(\boldsymbol{\theta}) \propto \mathbb{I}_\Theta(\boldsymbol{\theta})$, where $\Theta = \{\boldsymbol{\theta} s.t. \phi \in [0,1], \gamma \in [0,1], \delta \in [0, \min\{(1-\gamma)\phi, (1-\phi)\gamma\}], \forall\, i, j\}$.

Given these prior distributions, the goal of the Bayesian analysis is to know about the parameter $\boldsymbol{\theta}$ from the joint posterior distribution, $\pi(\boldsymbol{\theta}|\mathbf{y}_{1:T}) \propto \pi(\boldsymbol{\theta}) L(\mathbf{y}_{1:T}|\boldsymbol{\theta})$. The posterior distribution is not tractable, thus Markov Chain Monte Carlo (MCMC) procedure is applied to produce sample from this density and to approximate all posterior quantities of interest. Our MCMC algorithm is a Metropolis-Hastings (MH)

[7] The assumption that agents form expectations with a forecasting unrestricted VAR is common in the Phillips curve literature – see, for example, Rudd and Whelan (2005), Cogley and Sbordone (2008), Dupor et al. (2010). Interest rate spread is taken into account since is a good predictor of the economic cycle (see, for example, Wheelock and Wohar, 2009).

[8] All of our estimates are based on 50,000 MCMC samples, a burn-in sample of 10,000 iterations, and a thinning rate of 0.25 as to improve the efficiency of the MCMC estimator.

[9] The starting point of the three out-of-sample periods is selected on the basis of three relevant events: the beginning of world-wide diffusion of information and communication technology (ITC) (1995Q1), the peak of the ICT bubble (2000Q1), and the housing bubble burst and the beginning of global financial crisis (2008Q1).

[10] Data for CPI (seasonally adjusted) are taken from Federal Reserve of St. Louis (FRED) and cover the period 1947Q1–2020Q1.

[11] The results are robust to different values of $K$.

with the following target distribution:

$$f(\gamma,\phi,\delta|\mathbf{y}_{1:T}) \propto \left(\sigma^2(\theta)^{-\frac{T}{2}} \exp\left\{-\frac{1}{2\sigma^2(\theta)}\sum_{t=1}^{T}\varepsilon_t(\theta)^2\right\}\mathbb{I}_{\Theta}(\theta).\right. \quad (15)$$

In order to impose the restriction on the parameters the following re-parameterization is considered: $(\theta_1,\theta_2,\theta_3)' = \mathbf{g}(\xi)$, where $\xi = (\xi_1,\xi_2,\xi_3)$, $\theta_1 = \phi(\xi_1)$, $\theta_2 = \phi(\xi_2)$ and $\theta_3 = \phi(\xi_3)$, with $\phi(x) = 1/(1+\exp(-x))$ the logistic transform.

The random walk proposal distribution $\xi^* \sim \mathcal{N}_3(\xi^{(j-1)},\Lambda)$ is applied, with $\Lambda = \text{diag}\{(0.01, 0.01, 0.01)\}$ the scale matrix and $\xi^{(j-1)}$ the value of the MH chain at the previous iteration. The log-acceptance probability is

$$0 \wedge \left(\log f(\mathbf{g}(\xi^*)|\sigma^2,\mathbf{y}_{1:T}) + J(\widetilde{\xi}^*) - \log f(\mathbf{g}(\xi^{(j-1)})|\sigma^2,\mathbf{y}_{1:T}) - J(\widetilde{\xi}^{(j-1)})\right),$$

where $J(\xi) = \sum_{i=1}^{3}(\log\phi(\widetilde{\xi}_i) - \log(1-\phi(\widetilde{\xi}_i)))$ denotes the log-Jacobian of the parameter transform used.

### 3.3. Bayes Factor

This section describes the Bayes Factor (BF) used to compare the in-sample performance of the DS-PC model with dependence ($\mathcal{M}_2$, see equation (8)) with the DS-PC model without dependence ($\mathcal{M}_1$, see equation (9)). The BF is the ratio of the posterior normalizing constants of models under comparison

$$BF_{21} = \frac{f(\mathcal{M}_2|\mathbf{y}_{1:T})}{f(\mathcal{M}_1|\mathbf{y}_{1:T})}, \quad (16)$$

where

$$f(\mathcal{M}_j|\mathbf{y}_{1:T}) = m_j(\mathbf{y}_{1:T})\pi(\mathcal{M}_j) \quad (17)$$

is the model posterior, with $\pi(\mathcal{M}_j)$ the model prior and

$$m_j(\mathbf{y}_{1:T}) = \int L(\mathbf{y}_{1:T}|\theta,\mathcal{M}_j)\pi(\theta|\mathcal{M}_j)d\theta \quad (18)$$

the marginal likelihood. Following Geyer (1994) we evaluate the logarithmic BF $\kappa = K_2 - K_1$ with $K_j = \log f(\mathcal{M}_j|\mathbf{y}_{1:T})$ by maximizing the quasi-likelihood function of a reverse logistic regression

$$\ell_n(\kappa) = \sum_{i=1}^{n}\log p_1(x_{i1},K_1) + \sum_{i=1}^{n}\log p_2(x_{i2},K_2) \quad (19)$$

where

$$p(x|\mathcal{M}_j) = \frac{h_j(x)\exp(K_j)}{h_1(x)\exp(K_1)+h_2(x)\exp(K_2)},\ j=1,2 \quad (20)$$

is the probability assigned to the model $j$, $n$ is the number of MCMC draws for each model and $x_{ij} = \log f(\mathbf{y}_{1:T}|\theta^{(i)},\mathcal{M}_j)$ is the log-likelihood of the model $\mathcal{M}_j$ evaluated at the $i$-th MCMC sample for each model.

### 3.4. Forecast accuracy measures

For the evaluation of the forecasts of the models (8)–(11), we use a standard accuracy measure, MSE, and DA by Blaskowitz and Herwartz (2011, 2014).[12] As for the MSE, we have

$$MSE = \frac{1}{T}\sum_{t=1}^{T}(\pi_{t+h} - \widehat{\pi}_{t,t+h})^2, \quad (21)$$

where $\pi_{t+h}$ is the actual values of inflation, $\widehat{\pi}_{t+h}$ is the forecasts of inflation, and $h$ represents the forecast horizon.

Let $X_t^h$ and $Y_{t+h}$ be the h ahead forecast available at time $t$ and the realized value, respectively.[13] Let us consider $h = 1$. Using the indicator function $I(\bullet)$, the realized and predicted directions are given gained as $\widetilde{Y}_t = I(Y_{t+1} - Y_t > 0)$ and $\widetilde{X}_t = I(X_t^1 - Y_t > 0)$, respectively (see Blaskowitz and Herwartz, 2011).

The (In-)correct directional forecast can be defined by the binary variable $\widetilde{Z}_t = I(\widetilde{X}_t = \widetilde{Y}_t)$. A loss function for directional forecast is given by (see Blaskowitz and Herwartz, 2011):

$$L^{DA}(X_t^1, Y_{t+1}, Y_t) = \begin{cases} a & \text{if } \widetilde{Z}_t = 1 \\ b & \text{if } \widetilde{Z}_t = 0 \end{cases} \quad (22)$$

where $(a,b) \neq (0,0)$. A correct direction forecast will take the value of $a$, while an incorrectly predicted direction will take value of $b$. Usually $(a,b) = (1,0)$ or $(a,b) = (1,-1)$. In this paper, we use $(a,b) = (1,0)$ as in Swanson and White (1997a,b), Diebold (2007) and Costantini et al. (2016).

## 4. Empirical results

Fig. 1 reports in-sample estimates of $p_{ij}$ and $\delta$ for the new DS-PC model with dependence and of $p_{ij}$ for the DS-PC model without dependence in case of $w_{ex} = 60$ and $w_{es} = 40$, respectively. The results show that: (*i*)$\delta$ is highly statistically significant; (*ii*) the estimates of $p_{ij}$ show a similar dynamics over time, but different magnitude; and (*iii*) the new DS-PC model seems to give more weights to symmetric joint probabilities (i.e., $p_{ij}$ with $i = j$).

The parameter $\delta$, which measures the dependence among the events of adjusting price and updating information, displays a reduction in magnitude over time after some years of the occurrence of the financial crisis (see the vertical line in Fig. 1 for the financial crisis). This may be due to the fact that firms may have experienced difficulties in gathering information during the crisis, with a decreasing probability of updating information. However, even during this period of uncertainty, the parameter $\delta$ remains positive and statistically significant.

In order to establish the best in-sample performance among the two DS-PC models, the new model ($\mathcal{M}_2$) and the model with stickiness independence ($\mathcal{M}_1$), we use BF, which is the ratio $K_2/K_1$ of the posterior normalizing constants $K_j$ of the two models under comparison (see section 3.3). We evaluate the log-BF using the reversion logistic estimator proposed in Geyer (1994). In Bayesian inference, the Jeffrey's scale (Kass and Raftery, 1995) provides the thresholds to compare and assess the performance of two models. If the log-BF is larger than 5, there is a very strong evidence in favour of $\mathcal{M}_2$ against $\mathcal{M}_1$. The dynamics in Fig. 2 show that there is a very strong evidence in favour of the new DS-PC model with dependent stickiness.[14]

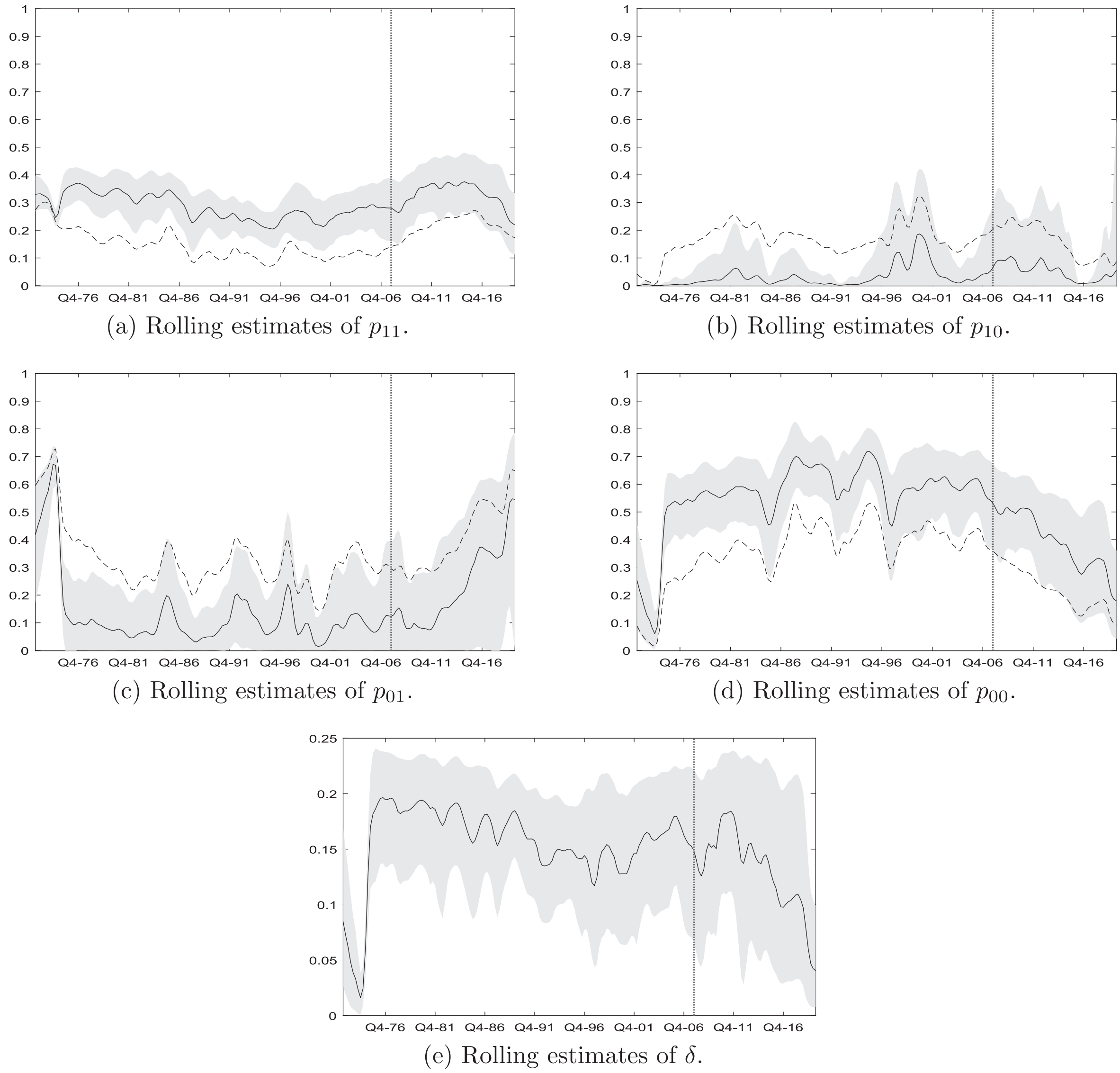
For a robustness check, two further exercises are carried out.[15] First, the estimates illustrated in Fig. 1 are also obtained for $w_{es} = 50$ and $w_{es} = 60$, and the results are reported in Figs. B1 and B2, respectively (see Appendix B). The symmetric joint probabilities ($p_{00}$ e $p_{11}$) of the DS-PC model with dependence are statistically significant and show larger values than those of the model without dependence (see Figs. B1a and B1d, and B2a and B2d, respectively). Further, the $\delta$ parameter is also statistical significant, and shows a reduction after the financial crisis as in Fig. 1. Second, GDP deflator is first replaced by CPI as inflation measure (see Fig. B3), and then the output gap is used instead of real marginal costs (see Fig. B4). When using CPI, the joint probabilities are statistically significant and larger in value for the new DS-PC model with dependence; the parameter $\delta$ is also statistical significant. Both the joint probabilities and the dependence parameter do not show a decreasing tendency as in case of GDP deflator. This may

[12] For the directional accuracy, see also Chen et al. (2016) and Costantini et al. (2016).

[13] In our out-of-sample analysis, this corresponds to the predicted and realized inflation, respectively.

[14] As pointed out in section 3.1, a battery of robustness checks is performed. For reasons of space, the results for the Bayes factor are not reported here. However, the results are qualitatively similar to those in Fig. 2.

[15] For reasons of space, we report in B all the results of robustness checks.



(a) Rolling estimates of $p_{11}$.
(b) Rolling estimates of $p_{10}$.
(c) Rolling estimates of $p_{01}$.
(d) Rolling estimates of $p_{00}$.
(e) Rolling estimates of $\delta$.

**Fig. 1.** Rolling estimates of key parameters for the model with dependent (solid lines) and independent (dashed lines) dual stickiness with $w_{ex} = 60$ and $w_{es} = 40$, GDP deflator as inflation measure, and real marginal costs as forcing variable. 90% confidence intervals (gray areas) refer to parameters estimates of the model with dependent dual stickiness. $p_{11}$, $p_{00}$, $p_{01}$, and $p_{10}$ indicates the probabilities of the four possible events of setting price and updating information (see equation (1)). $\delta$ measures the dependence among the two events. All parameters are estimated by using a two-step rolling scheme described in section 3.1. The effective sample is 1972Q3-2020Q1. The vertical dotted line corresponds to 2007Q4.

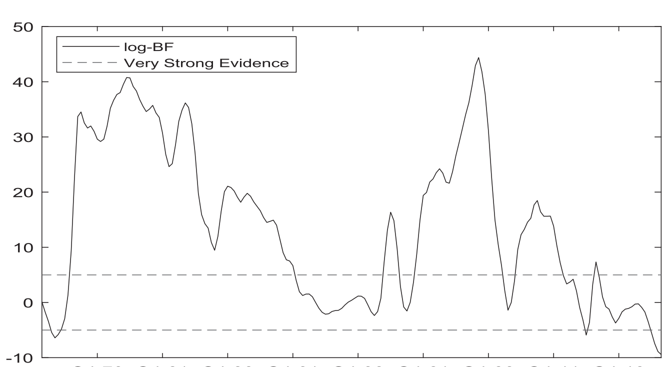



**Fig. 2.** Rolling estimate of the logarithmic Bayes Factor (solid line) and the very-strong evidence thresholds (dashed lines) following the Jeffrey's scale of evidence. *Log BF* $\overset{\text{def}}{=}\log(\text{BF})(\mathcal{M}_2, \mathcal{M}_1)$, where $\mathcal{M}_2$ indicates the new DS-PC model with dependence and $\mathcal{M}_1$ is the DS-PC model without dependence. When log-BF > 5, we have a very-strong evidence in favour of model $\mathcal{M}_2$; if log-BF < −5, there is a very-strong evidence in favour of model $\mathcal{M}_1$. Log-BF is calculated using the two-step rolling scheme procedure (see also notes in Fig. 1). The effective sample is 1972Q3–2020Q1.

**Table 1**
Mean square error (MSE) and directional accuracy (DA). $w_{ex} = 60$ and $w_{es} = 40$, GDP deflator as inflation measure, and real marginal costs as forcing variable.

| Out-of-sample period: 1995Q1–2020Q1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | MSE | | | | DA | | | |
| Horizons/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.47 | 0.48 | 1.21 | 1.37 | 0.62 | 0.63 | 0.59 | 0.54 |
| $h = 4$ | 0.48 | 0.50 | 1.24 | 1.42 | 0.63 | 0.63 | 0.60 | 0.54 |
| $h = 8$ | 0.54 | 0.55 | 1.32 | 1.53 | 0.62 | 0.62 | 0.58 | 0.50 |
| Out-of-sample period: 2000Q1–2020Q1 | | | | | | | | |
| Horizons/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.54 | 0.55 | 1.34 | 1.52 | 0.64 | 0.63 | 0.60 | 0.56 |
| $h = 4$ | 0.58 | 0.58 | 1.31 | 1.55 | 0.63 | 0.63 | 0.60 | 0.54 |
| $h = 8$ | 0.63 | 0.64 | 1.35 | 1.62 | 0.64 | 0.65 | 0.59 | 0.52 |
| Out-of-sample period: 2008Q1–2020Q1 | | | | | | | | |
| Horizon/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.68 | 0.69 | 1.25 | 1.44 | 0.69 | 0.67 | 0.67 | 0.59 |
| $h = 4$ | 0.75 | 0.75 | 1.46 | 1.63 | 0.65 | 0.65 | 0.65 | 0.57 |
| $h = 8$ | 0.82 | 0.83 | 1.62 | 1.81 | 0.69 | 0.69 | 0.67 | 0.59 |

Notes: $w_{ex} = 60$ and $w_{es} = 40$ denote the number of observations of the rolling window to generate the expectations and estimate the Phillips curve, respectively; $h = 1, 4, 8$ indicate the forecast horizons; $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$, and $\mathcal{M}_4$ denote the dual stickiness Phillips curve model without dependence (equation (9)), the dual stickiness Phillips curve model with dependence (equation (8)), the pure sticky price model (equation (10)) and the pure sticky information model (equation (11)), respectively; the numbers of DA denote the proportion of corrected predictions of the direction.

be attributed to the different features of CPI and GDP deflator.[16] Similar results for the joint probabilities and the dependence parameter are observed when output gap is used instead of real marginal costs as forcing variable (see Fig. B4). All this shows that our findings are robust to different rolling windows and measures of forcing variables, and the dependence parameter plays a role in influencing the probability of changing princes when firms update their information.

As far as the out-of-sample performance is concerned, the results for MSE and DA in case of $w_{ex} = 60$ and $w_{es} = 40$ are reported in Table 1. When looking at the MSE, the prediction errors vary over the different sub-sample periods and forecast horizons. More specifically, the forecasts tend to worsen in the last out-of-sample period (2008Q1–2020Q1) for all the models, likely due to the impact of financial crisis on the predictions, and for $h = 8$, likely due to the increasing uncertainty. Slight different findings are observed for DA over the three sub-sample periods, with larger values for this accuracy measure in the last sub-sample period, while DA values do not change much over the forecast horizons. The improvement of the forecasts in the last sub-sample period may be due to the relatively persistent and negative dynamics of inflation during the years of the crisis and afterwards (see, for example, Granville and Zeng, 2019), making the prediction of the direction less challenging. When looking into the performance of the single models, it can be noticed that the two DS-PC models perform similarly across the three forecast horizons ($h = 1, 4, 8$ quarters) and out-of-sample periods (1995Q1–2020Q1, 2000Q1–2020Q1, 2008Q1–2020Q1) in terms of both MSE and DA, while the pure sticky price and pure sticky information models do much worse, with the latter being the worst performer.

Tables B1–B3 illustrate the results of the first robustness check for MSE and DA in case of $w_{ex} = 60$ and $w_{es} = 50$, $w_{ex} = 80$ and $w_{es} = 40$, and $w_{ex} = 80$ and $w_{es} = 40$, respectively (see Appendix B). Some general findings emerge. First, the use of a lager rolling window to generate the expectations seems to produce better results in terms of MSE, while DA do not vary substantially. Second, the results in terms of MSE for all the models tend to worsen as the forecast horizon increases, and the worst performance is uncovered in the last sub-sample period. Third, the two dual stickiness models are equally competitive in terms of both MSE and DA, and they do better than the two pure sticky models, with a very few exceptions in case of DA for the pure sticky price model, which seems to gains some ground (see Table B3). All these findings seem to confirm those reported in Table 1.

Tables B4 and B5 display the results for MSE and DA for the second robustness check. In particular, Table B4 includes those findings when GDP deflator is replaced with CPI, whereas Table B5 reports the findings in case of output gap instead of real marginal costs. Some general evidence emerges. First, the two dual stickiness models are still equally competitive and outperform the pure sticky models, both in terms of MSE and DA. Second, when using CPI inflation instead of GDP deflator, the forecasts tend to deteriorate, especially in terms of MSE (the impact on DA is less pronounced). Yet, this result may depend on the features of CPI as a measure of inflation. Third, the use of the output gap instead of real marginal costs helps to improve the forecasts of all the models. This result is in line with that obtained in Rudd and Whelan (2005).

## 5. Conclusions

This paper investigates the role of dependence in the sticky price and sticky information Phillips curve. A new dual stickiness Phillips curve is derived by assuming dependence among the events of setting prices and updating information. The dependence is modelled through a bivariate Bernoulli distribution. This is the first contribution of the paper.

The performance of the new model with dependence is scrutinized against other models through in- and out-of-sample analyses for inflation using US data over the period 1947Q1–2020Q1. The in-sample analysis is carried out for two dual stickiness Phillips curves, with and without dependence. For the forecasts, two additional models, such as

[16] First, GDP deflator includes only domestic goods, while CPI includes all goods bought by consumers, included foreign goods. Second, GDP deflator is a measure of the prices of all goods and services, while the CPI is a measure of only goods bought by consumers. This implies that CPI captures the prices of goods not produced by domestic companies and reflects only the choices of consumers. More importantly, the two inflation series exhibit a different volatility (see, for example, McKnight et al., 2020).

a pure sticky price model and a pure sticky information model, are also taken into account. This is the second added value of the paper.

To estimate the models, a two-step procedure based on a rolling estimation scheme is adopted. Expectations of inflation and real marginal costs are first estimated by a rolling VAR model. Then, the parameters of the different Phillips curves are estimated by Bayesian inference.

For the in-sample analysis, we proceed as follows. A rolling window estimation of the key parameters of the two dual stickiness models is first carried out to ascertain the significance of these parameters. Then, the performance of these two models is compared by the Bayes factor. Two robustness checks are also carried out. First, diverse combinations of rolling windows for the expectations and the estimations of the models are considered. Second, a different measure of inflation, CPI, and a forcing variable, output gap, are used.

The out-of-sample analysis is conducted using four different Phillips curve models: the two dual stickiness models, a pure sticky price model and a pure sticky information model. For the evaluation of the forecasts, we use the mean square error and the directional accuracy by Blaskowitz and Herwartz (2009, 2011), three different out-of-sample periods, 1995Q1–2020Q1, 2000Q1–2020Q1, and 2008Q1–2020Q1, and three different forecast horizons, $h = 1, 4, 8$. Similarly to the in-sample analysis, a robustness check is also performed.

The main results in-sample show that the key parameters of the two dual stickiness Phillips curves are statistically significant: (*i*) the parameter that measures the dependence is positive; and (*ii*) the estimates of the two models reveal that the joint probabilities related to the events of setting price and updating information have a larger magnitude in case of the dual stickiness model with dependence. Further, the comparison among the two dual stickiness models through the Bayes factor point to strong evidence in favour of the new model with stickiness dependence with respect to the model without dependence. The robustness check results seem to confirm those in the main exercise.

The out-of-sample findings unveil similar performances among the two dual stickiness models over all the forecast horizons and across the sub-sample periods in terms of mean square error and directional accuracy. Further, the two dual stickiness models outperform the pure sticky models, both in terms of mean square error and directional accuracy. Moreover, among the two pure sticky models, the pure sticky information model records the worst performance, while the pure sticky price model seems to gain some ground in terms of directional accuracy. In general, these results indicate that the new dual stickiness model represents a good alternative to the model without dependence, and the robustness checks confirm these conclusions.

All in all, some implications for policy makers and practitioners can be drawn. A policy maker, who looks at both in and out-of-sample analyses for policy interventions and predictions respectively, may opt for the dual stickiness model with dependence given its performance in both analyses. This is because the dependence parameter plays a significant role for the events of setting price and updating information. On the other hand, a practitioner may be primarily interested in the predictions, and the dual stickiness model with no dependence may be preferable, given its lesser complexity and its competitive forecasting performance.

**Declaration of competing interest**

The authors declare that there is no conflict of interest.

## Appendix A. Derivations of the model

### A.1. Proof of equation (7)

By subtracting $\gamma p_t$ on both sides of equation (4) of the main text, and after some manipulations, one yields:

$$\pi_t = \frac{(1-\gamma)}{\gamma}(q_t - p_t). \tag{A.1}$$

According to equation (6), $q_t$ can be written as

$$q_t = (1-\phi)E_t p_t^f + \phi(1-\phi)E_{t-1}p_t^f + \phi^2(1-\phi)E_{t-2}p_t^f + \ldots + \frac{\delta}{1-\gamma}\left(p_t^f - p_t^b\right). \tag{A.2}$$

Equation (A.2) can be written in two different and equivalent ways:

$$q_t = (1-\phi)p_t^f + \phi(1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k-1}p_t^f + \frac{\delta}{1-\gamma}\left(p_t^f - p_t^b\right) \tag{A.3}$$

$$q_t = (1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k}p_t^f + \frac{\delta}{1-\gamma}\left(p_t^f - p_t^b\right). \tag{A.4}$$

Focusing on equation (A.3) and noting that $p_t^f = \Delta p_t^f + p_{t-1}^f$, we can write $q_t$ as follows:

$$q_t = \underbrace{\left[(1-\phi) + \frac{\delta}{1-\gamma}\right]}_{\equiv\psi} p_t^f + \phi(1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k-1}\left(\Delta p_t^f + p_{t-1}^f\right) - \frac{\delta}{1-\gamma}p_t^b. \tag{A.5}$$

Since

$$\begin{aligned}\frac{\delta}{1-\gamma}p_t^b &= \frac{\delta}{1-\gamma}(1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k-1}p_t^f \\ &= \frac{\delta}{1-\gamma}(1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k-1}\left(\Delta p_t^f + p_{t-1}^f\right),\end{aligned}$$

then equation (A.5) becomes

$$q_t = \psi p_t^f + \underbrace{\phi\,(1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k-1}p_{t-1}^f}_{\equiv A} + \phi\,(1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k-1}\Delta p_t^f - \frac{\delta}{1-\gamma}\,(1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k-1}\left(\Delta p_t^f + p_{t-1}^f\right). \tag{A.6}$$

Now we focus on the quantity $A$. According to equation (A.4)

$$\begin{aligned} A &= \phi(1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k-1}p_{t-1}^f \\ &= \phi q_{t-1} - \phi\frac{\delta}{1-\gamma}(p_{t-1}^f - p_{t-1}^b) \\ &= \phi q_{t-1} - \phi\frac{\delta}{1-\gamma}p_{t-1}^f + \phi\frac{\delta}{1-\gamma}(1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k-2}p_{t-1}^f. \end{aligned} \tag{A.7}$$

Inserting equation (A.7) into equation (A.6) and rearranging the quantities containing the term $\Delta p_t^f$ one obtains

$$\begin{aligned} q_t &= \psi p_t^f + \phi q_{t-1} - \phi\frac{\delta}{1-\gamma}p_{t-1}^f + \phi\frac{\delta}{1-\gamma}\,(1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k-2}p_{t-1}^f \\ &\quad + (1-\phi)\left[\phi - \frac{\delta}{1-\gamma}\right]\sum_{k=0}^{\infty}\phi^k E_{t-k-1}\Delta p_t^f - \underbrace{\frac{\delta}{1-\gamma}\,(1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k-1}p_{t-1}^f}_{\equiv B}. \end{aligned} \tag{A.8}$$

The last term can be written as:

$$\begin{aligned} B &= \frac{\delta}{1-\gamma}\,(1-\phi)\sum_{k=0}^{\infty}\phi^k E_{t-k-1}p_{t-1}^f \\ &= \frac{\delta}{1-\gamma}\,(1-\phi)\left[E_{t-1}p_{t-1}^f + \phi\sum_{k=0}^{\infty}\phi^k E_{t-k-2}p_{t-1}^f\right]. \end{aligned} \tag{A.9}$$

Substituting in equation (A.10) and rearranging terms, we obtain

$$\begin{aligned} q_t &= \psi p_t^f + \phi q_{t-1} \underbrace{-\phi\frac{\delta}{1-\gamma}p_{t-1}^f - \frac{(1-\phi)\,\delta}{1-\gamma}p_{t-1}^f}_{=-\frac{\delta}{1-\gamma}p_{t-1}^f} \\ &\quad + (1-\phi)\left[\phi - \frac{\delta}{1-\gamma}\right]\sum_{k=0}^{\infty}\phi^k E_{t-k-1}\Delta p_t^f \\ &\quad + \frac{(1-\phi)\,\delta}{1-\gamma}\sum_{k=0}^{\infty}\phi^k\left[\phi E_{t-k-2}p_{t-1}^f - \phi E_{t-k-2}p_{t-1}^f\right] \\ &= \psi p_t^f + \phi q_{t-1} - \frac{\delta}{1-\gamma}p_{t-1}^f \\ &\quad + (1-\phi)\left[\phi - \frac{\delta}{1-\gamma}\right]\sum_{k=0}^{\infty}\phi^k E_{t-k-1}\Delta p_t^f. \end{aligned} \tag{A.10}$$

Subtracting $p_t$ on both sides of (A.10), we have

$$\begin{aligned} q_t - p_t &= \psi p_t^f + \phi q_{t-1} - \frac{\delta}{1-\gamma}p_{t-1}^f + (1-\phi)\left[\phi - \frac{\delta}{1-\gamma}\right]\sum_{k=0}^{\infty}\phi^k E_{t-k-1}\Delta p_t^f - p_t \\ &= \psi p_t^f + \phi q_{t-1} - \frac{\delta}{1-\gamma}p_{t-1}^f + (1-\phi)\left[\phi - \frac{\delta}{1-\gamma}\right]\sum_{k=0}^{\infty}\phi^k E_{t-k-1}\Delta p_t^f - \phi\pi_t - \phi p_{t-1} - (1-\phi)\,p_t, \end{aligned}$$

which holds the expression

$$\begin{aligned} q_t - p_t &= \left[(1-\phi) + \frac{\delta}{1-\gamma}\right]p_t^f - \frac{\delta}{1-\gamma}p_{t-1}^f + \phi\,(q_{t-1} - p_{t-1}) - \phi\pi_t - (1-\phi)\,p_t + (1-\phi)\left[\phi - \frac{\delta}{1-\gamma}\right]\sum_{k=0}^{\infty}\phi^k E_{t-k-1}\Delta p_t^f \\ &= (1-\phi)\left(p_t^f - p_t\right) + \frac{\delta}{1-\gamma}\Delta p_t^f + \phi\,(q_{t-1} - p_{t-1}) - \phi\pi_t + (1-\phi)\left[\phi - \frac{\delta}{1-\gamma}\right]\sum_{k=0}^{\infty}\phi^k E_{t-k-1}\Delta p_t^f. \end{aligned} \tag{A.11}$$

Now we focus on the quantity $\Delta p_t^f$. This variable, according to equation (5) (see the main text), is equal to

$$\Delta p_t^f = (1-\gamma)\Delta p_t^* + \underbrace{\gamma(1-\gamma)\sum_{j=0}^{\infty}\gamma^j\left[E_t p_{t+j+1}^* - E_{t-1}p_{t+j}^*\right]}_{\equiv F}. \tag{A.12}$$

By using the rational expectation hypothesis, we can express the expectation at time $t-k$ of $p^*$ in time $t+j$ in the following way

$$E_{t-k}p_{t+j}^* = E_{t-k-1}p_{t+j}^* + \tau_{t-k}, \tag{A.13}$$

where $\tau_{t-k}$ is a (random) variable which captures the revision of expectations in correspondence of an information updating between time $t-k-1$ and $t-k$.[17]

We now consider the expression $F$ in equation (A.12). We can write it in the following way by using equation (A.13)

$$\begin{aligned}\Delta p_t^f &= (1-\gamma)\delta p_t^* + \gamma(1-\gamma)\left\{E_t p_{t+1}^* - E_t p_t^* + \tau_t + \gamma E_t p_{t+2}^* - \gamma E_t p_{t+1}^* + \gamma\tau_t \ + \gamma^2 E_t p_{t+3}^* - \gamma^2 E_t p_{t+2}^* + \gamma^2\tau_t + \dots\right\}\\ &= (1-\gamma)\Delta p_t^* + \gamma(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_t\Delta p_{t+j+1}^* + \gamma(1-\gamma)\sum_{j=0}^{\infty}\gamma^j\tau_t.\end{aligned} \tag{A.14}$$

As $j\to\infty$, the previous equation can be written as

$$\Delta p_t^f = (1-\gamma)\Delta p_t^* + \gamma(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_t\Delta p_{t+j+1}^* + \gamma\tau_t. \tag{A.15}$$

Multiplying both side of equation (A.15) by $\delta/(1-\gamma)$, we have

$$\frac{\delta}{1-\gamma}\Delta p_t^f = \frac{\delta}{1-\gamma}\left[(1-\gamma)\Delta p_t^* + \gamma(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_t\Delta p_{t+j+1}^* + \gamma\tau_t\right].$$

Now it is easy to show that equation (A.11) corresponds to equation (7) of the main text

$$\begin{aligned}&\left[\frac{\gamma+\phi(1-\gamma)}{1-\gamma}\right]\pi_t = (1-\phi)\left(p_t^f - p_t\right)\\ &+\frac{\gamma}{1-\gamma}\phi\pi_{t-1}\\ &+(1-\phi)\left[\phi-\frac{\delta}{1-\gamma}\right]\sum_{k=0}^{\infty}\phi^k E_{t-k-1}\Delta p_t^f\\ &+\delta\Delta p_t^* + \delta\gamma\sum_{j=0}^{\infty}\gamma^j E_t\Delta p_{t-j+1}^* + \delta\gamma\tau_t.\end{aligned} \tag{A.16}$$

### A.2. Proof of equation (8)

According to the optimal price $p_t^*$ expressed as a function of the marginal cost, we have

$$\begin{aligned}&\left[\frac{\gamma+\phi(1-\gamma)}{1-\gamma}\right]\pi_t = (1-\phi)\left[(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_t mc_{t+j}^n - p_t\right]\\ &+\frac{\gamma}{1-\gamma}\phi\pi_{t-1}\\ &+(1-\phi)\left[\phi-\frac{\delta}{1-\gamma}\right]\sum_{k=0}^{\infty}\phi^k\left[(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_{t-k-1}\left(\Delta mc_{t+j}+\pi_{t+j}\right)\right]\\ &+\delta\left(\pi_t+\Delta mc_t\right) + \delta\gamma\sum_{j=0}^{\infty}\gamma^j E_t\left(\Delta mc_{t+j+1}+\pi_{t+j+1}\right) + \delta\gamma\tau_t.\end{aligned} \tag{A.17}$$

Since $\gamma < 1$, it is possible to import $p_t$ inside the summation in the first term in the square brackets on the right-hand side of equation (A.17), thus obtaining

[17] In particular, note that $E\left[\tau_{t-k}\right] = 0$.

$$\left[\frac{\gamma+\phi(1-\gamma)}{1-\gamma}\right]\pi_t=(1-\phi)\left[(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_t\left(mc^n_{t+j}-p_t\right)\right]$$
$$+\frac{\gamma}{1-\gamma}\phi\pi_{t-1}$$
$$+(1-\phi)\left[\phi-\frac{\delta}{1-\gamma}\right]\sum_{k=0}^{\infty}\phi^k\left[(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_{t-k-1}\left(\Delta mc_{t+j}+\pi_{t+j}\right)\right]$$
$$+\delta\left(\pi_t+\Delta mc_t\right)+\delta\gamma\sum_{j=0}^{\infty}\gamma^j E_t\left(\Delta mc_{t+j+1}+\pi_{t+j+1}\right)+\delta\gamma\tau_t. \tag{A.18}$$

Noting that

$$mc^n_{t+j}-p_t=\begin{cases} mc_t, & \text{for } j=0\\ mc_{t+j}+\pi_{t+1}+\cdots+\pi_{t+j}, & \text{for } j\geqslant 1,\end{cases}$$

we can write equation (A.18) as:

$$\left[\frac{\gamma+\phi(1-\gamma)}{1-\gamma}\right]\pi_t=(1-\phi)\left[(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_t\left(mc_{t+j}+\sum_{k=1}^{j}\pi_{t+k}\right)\right]$$
$$+\frac{\gamma}{1-\gamma}\phi\pi_{t-1}$$
$$+(1-\phi)\left[\phi-\frac{\delta}{1-\gamma}\right]\sum_{k=0}^{\infty}\phi^k\left[(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_{t-k-1}\left(\Delta mc_{t+j}+\pi_{t+j}\right)\right]$$
$$+\delta\left(\pi_t+\Delta mc_t\right)+\delta\gamma\sum_{j=0}^{\infty}\gamma^j E_t\left(\Delta mc_{t+j+1}+\pi_{t+j+1}\right)+\delta\gamma\tau_t. \tag{A.19}$$

or equivalently as

$$\left[\frac{\gamma+(1-\gamma)(\phi-\delta)}{1-\gamma}\right]\pi_t=(1-\phi)\left[(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_t\left(mc_{t+j}+\sum_{k=1}^{j}\pi_{t+k}\right)\right]$$
$$+\frac{\gamma}{1-\gamma}\phi\pi_{t-1}$$
$$+(1-\phi)\left[\phi-\frac{\delta}{1-\gamma}\right]\sum_{k=0}^{\infty}\phi^k\left[(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_{t-k-1}\left(\Delta mc_{t+j}+\pi_{t+j}\right)\right]$$
$$+\delta\Delta mc_t+\delta\gamma\sum_{j=0}^{\infty}\gamma^j E_t\left(\Delta mc_{t+j+1}+\pi_{t+j+1}\right)+\delta\gamma\tau_t, \tag{A.20}$$

from which it is possible to obtain the result in equation (8) of the main text.

### *A.3. Proof of equation (10)*

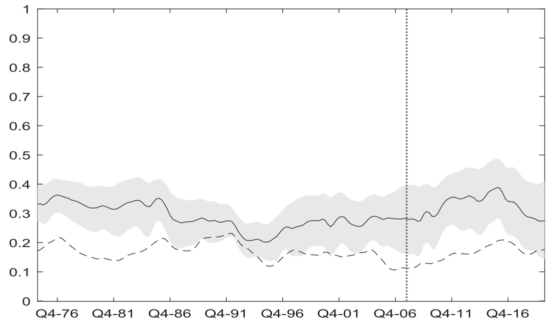
From (8), setting $\delta = 0$ and $\phi = 0$ we have

$$\pi_t=\frac{(1-\gamma)}{\gamma}(1-\gamma)\sum_{j=0}^{\infty}\gamma^j E_t\left(mc_{t+j}+\sum_{k=1}^{j}\pi_{t+k}\right)$$
$$=\frac{(1-\gamma)}{\gamma}\left[(1-\gamma)E_t mc_t+(1-\gamma)\sum_{j=1}^{\infty}\gamma^j E_t(mc^n_{t+j}-p_t)\right]$$
$$=\frac{(1-\gamma)^2}{\gamma}mc_t+(1-\gamma)E_t(p^f_{t+1}-p_t).$$

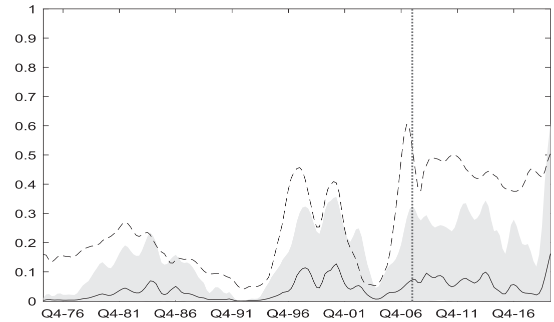
Since $p_t=\gamma p_{t-1}+(1-\gamma)p^f_t$, it is immediate to have $E_t\frac{\pi_{t+1}}{(1-\gamma)}=E_t(p^f_{t+1}-p_t)$. So that, we have

$$\pi_t=\kappa mc_t+E_t\pi_{t+1}. \tag{A.21}$$

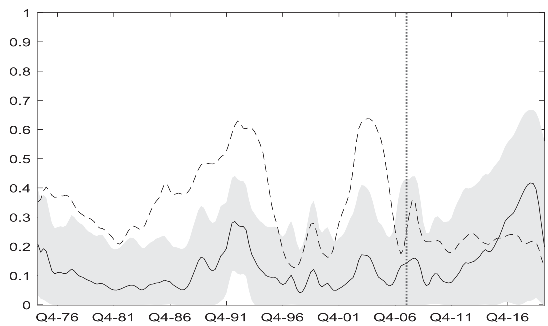
## Appendix B. Robustness checks



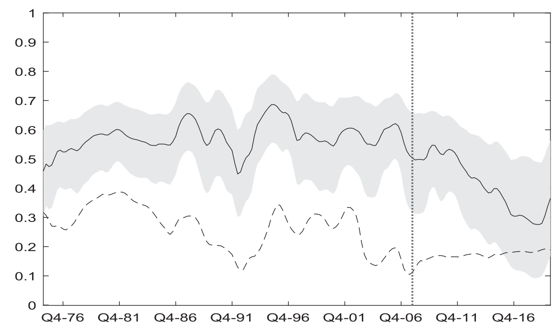

(a) Rolling estimates of $p_{11}$.



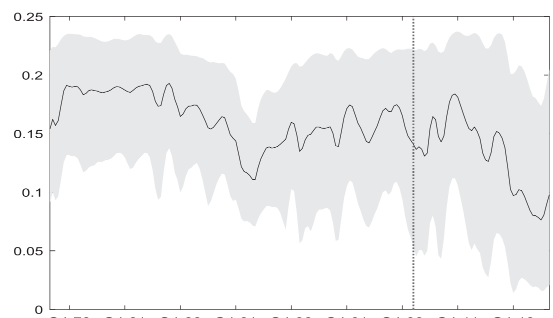

(b) Rolling estimates of $p_{10}$.




(c) Rolling estimates of $p_{01}$.




(d) Rolling estimates of $p_{00}$.




(e) Rolling estimates of $\delta$.

**Fig. B1** Rolling estimates of key parameters for the model with dependent (solid lines) and independent (dashed lines) dual stickiness with $w_{ex} = 60$ and $w_{es} = 50$, GDP deflator as inflation measure, and real marginal costs as forcing variable. 90% confidence intervals (gray areas) refer to parameters estimates of the model with dependent dual stickiness. $p_{11}$, $p_{00}$, $p_{01}$, and $p_{10}$ indicates the probabilities of the four possible events of setting price and updating information (see equation (1)). $\delta$ measures the dependence among the two events. All parameters are estimated by using a two-step rolling scheme described in section 3.1. The effective sample is 1975Q1-2020Q1. The vertical dotted line corresponds to 2007Q4.

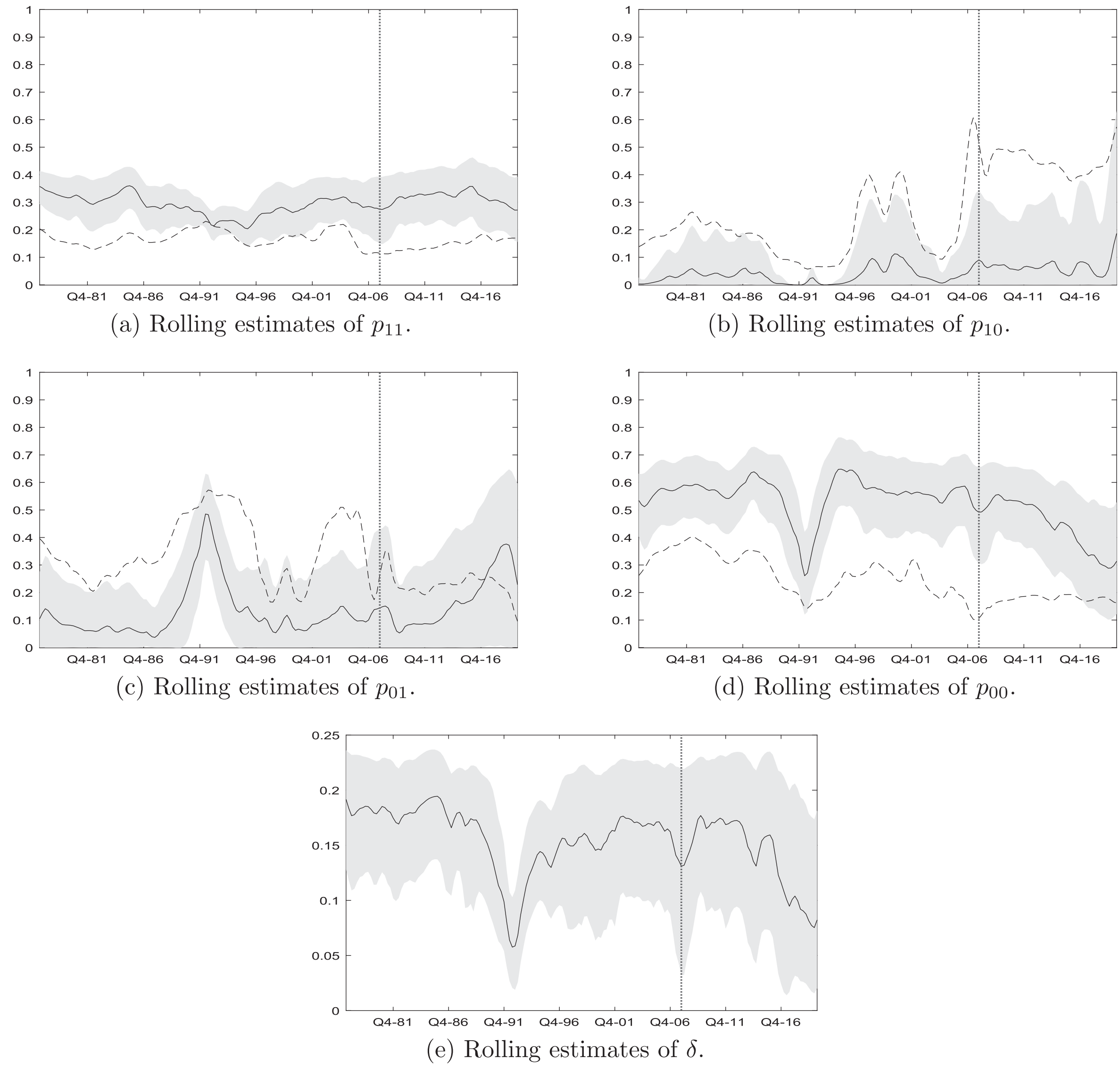
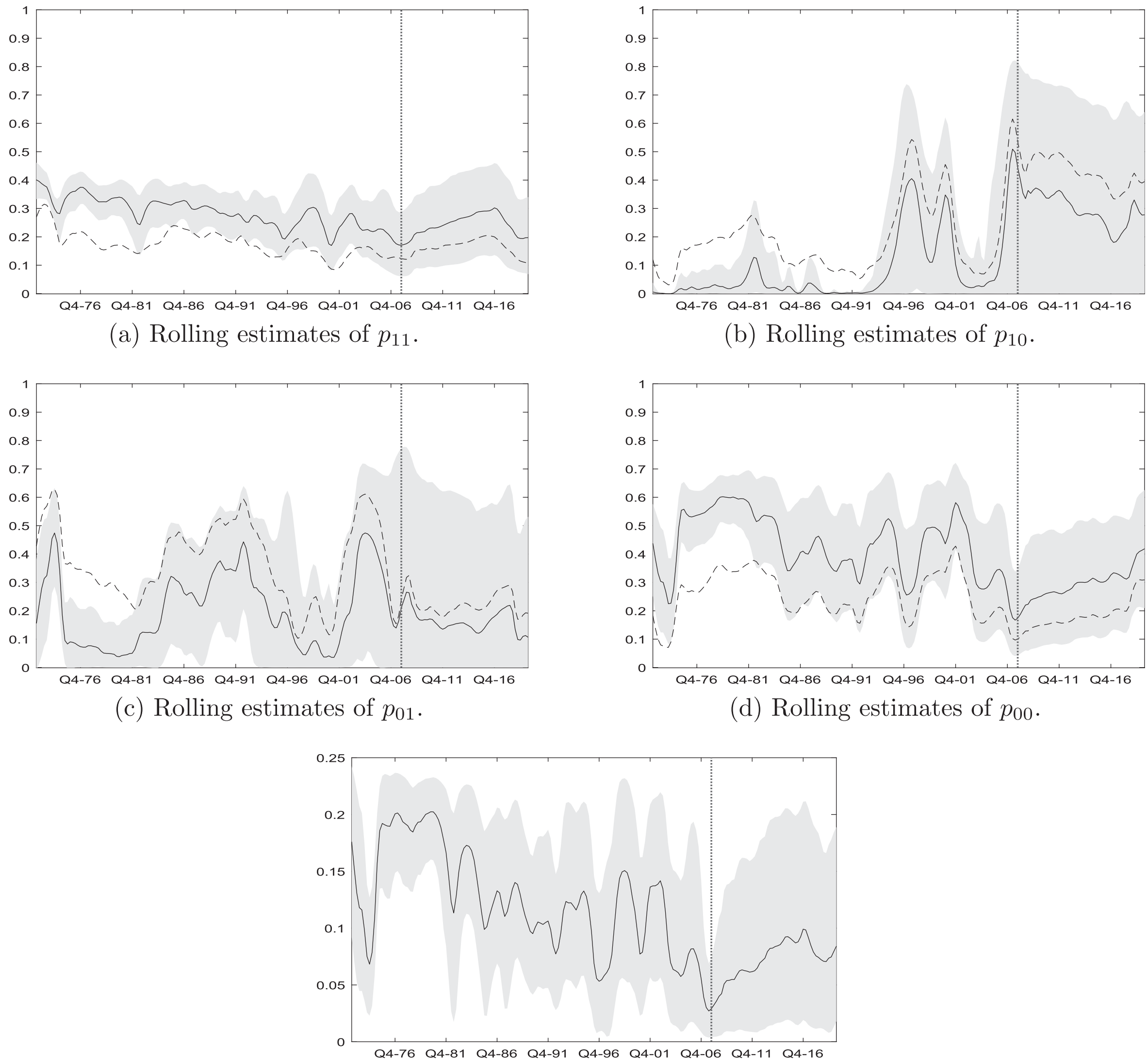



**Fig. B2** Rolling estimates of key parameters for the model with dependent (solid lines) and independent (dashed lines) dual stickiness with $w_{ex} = 60$ and $w_{es} = 60$, GDP deflator as inflation measure, and real marginal costs as forcing variable. 90% confidence intervals (gray areas) refer to parameters estimates of the model with dependent dual stickiness. $p_{11}$, $p_{00}$, $p_{01}$, and $p_{10}$ indicates the probabilities of the four possible events of setting price and updating information (see equation (1)). $\delta$ measures the dependence among the two events. All parameters are estimated by using a two-step rolling scheme described in section 3.1. The effective sample is 1977Q3-2020Q1. The vertical dotted line corresponds to 2007Q4.




(a) Rolling estimates of $p_{11}$.

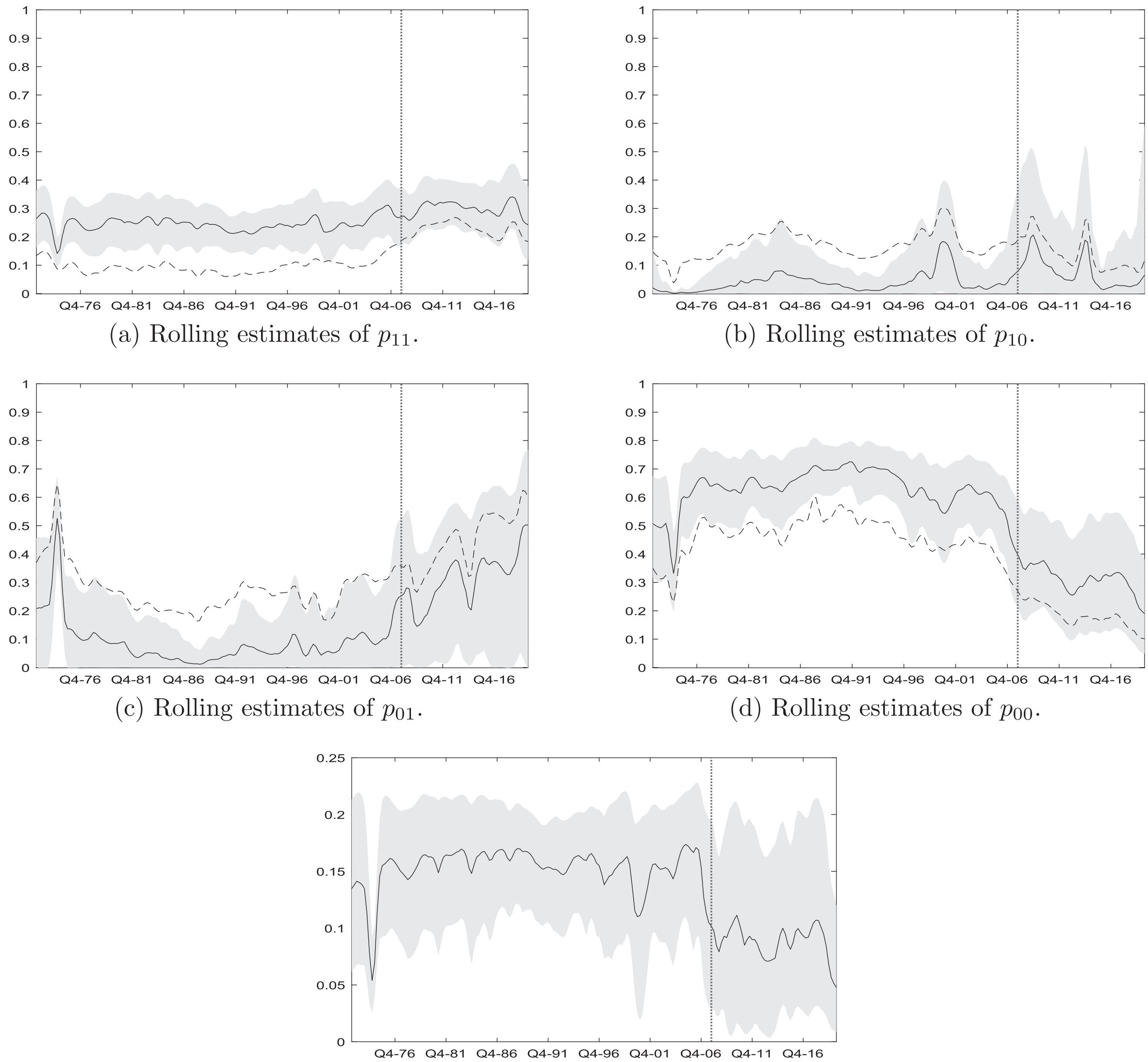
(b) Rolling estimates of $p_{10}$.

(c) Rolling estimates of $p_{01}$.

(d) Rolling estimates of $p_{00}$.

(e) Rolling estimates of $\delta$.

**Fig. B3** Rolling estimates of key parameters for the model with dependent (solid lines) and independent (dashed lines) dual stickiness with $w_{ex} = 60$ and $w_{es} = 40$, CPI inflation, and real marginal costs as forcing variable. 90% confidence intervals (gray areas) refer to parameters estimates of the model with dependent dual stickiness. $p_{11}$, $p_{00}$, $p_{01}$, and $p_{10}$ indicates the probabilities of the four possible events of setting price and updating information (see equation (1)). $\delta$ measures the dependence among the two events. All parameters are estimated by using a two-step rolling scheme described in section 3.1. The effective sample is 1972Q3-2020Q1. The vertical dotted line corresponds to 2007Q4.




(a) Rolling estimates of $p_{11}$.

(b) Rolling estimates of $p_{10}$.

(c) Rolling estimates of $p_{01}$.

(d) Rolling estimates of $p_{00}$.

(e) Rolling estimates of $\delta$.

**Fig. B4** Rolling estimates of key parameters for the model with dependent (solid lines) and independent (dashed lines) dual stickiness with $w_{ex} = 60$ and $w_{es} = 40$, GDP deflator as inflation measure, and output gap as forcing variable. 90% confidence intervals (gray areas) refer to parameters estimates of the model with dependent dual stickiness. $p_{11}$, $p_{00}$, $p_{01}$, and $p_{10}$ indicates the probabilities of the four possible events of setting price and updating information (see equation (1)). $\delta$ measures the dependence among the two events. All parameters are estimated by using a two-step rolling scheme described in section 3.1. The effective sample is 1972Q3-2020Q1. The vertical dotted line corresponds to 2007Q4.

**Table B1**
Mean square error (MSE) and directional accuracy (DA). $w_{ex} = 60$ and $w_{es} = 50$, GDP deflator as inflation measure, and real marginal costs as forcing variable

| Out-of-sample period: 1995Q1–2020Q1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | MSE | | | | DA | | | |
| Horizons/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.54 | 0.56 | 1.34 | 1.60 | 0.60 | 0.61 | 0.56 | 0.50 |
| $h = 4$ | 0.59 | 0.60 | 1.49 | 1.77 | 0.61 | 0.60 | 0.56 | 0.50 |
| $h = 8$ | 0.66 | 0.69 | 1.84 | 2.07 | 0.58 | 0.58 | 0.55 | 0.50 |
| Out-of-sample period: 2000Q1–2020Q1 | | | | | | | | |
| Horizons/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.63 | 0.63 | 1.36 | 1.67 | 0.63 | 0.65 | 0.59 | 0.51 |
| $h = 4$ | 0.67 | 0.69 | 1.54 | 1.86 | 0.60 | 0.60 | 0.58 | 0.49 |
| $h = 8$ | 0.74 | 0.78 | 1.92 | 2.21 | 0.60 | 0.60 | 0.57 | 0.51 |
| Out-of-sample period: 2008Q1–2020Q1 | | | | | | | | |
| Horizon/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.79 | 0.80 | 1.50 | 1.71 | 0.73 | 0.71 | 0.67 | 0.59 |
| $h = 4$ | 0.84 | 0.85 | 1.51 | 1.76 | 0.71 | 0.67 | 0.67 | 0.59 |
| $h = 8$ | 0.73 | 0.75 | 1.47 | 1.79 | 0.67 | 0.67 | 0.63 | 0.57 |

Notes: $w_{ex} = 60$ and $w_{es} = 50$ denote the number of observations of the rolling window to generate the expectations and estimate the Phillips curve, respectively; $h = 1, 4, 8$ indicate the forecast horizons; $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$, and $\mathcal{M}_4$ denote the dual stickiness Phillips curve model without dependence (equation (9)), the dual stickiness Phillips curve model with dependence (equation (8)), the pure sticky price model (equation (10)) and the pure sticky information model (equation (11)), respectively; the numbers of DA denote the proportion of corrected predictions of the direction.

**Table B2**
Mean square error (MSE) and directional accuracy (DA). $w_{ex} = 80$ and $w_{es} = 40$, GDP deflator as inflation measure, and real marginal costs as forcing variable

| Out-of-sample period: 1995Q1–2020Q1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | MSE | | | | DA | | | |
| Horizons/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.32 | 0.32 | 0.81 | 0.86 | 0.63 | 0.63 | 0.64 | 0.61 |
| $h = 4$ | 0.35 | 0.35 | 0.90 | 0.90 | 0.62 | 0.62 | 0.61 | 0.58 |
| $h = 8$ | 0.41 | 0.41 | 1.00 | 1.12 | 0.64 | 0.62 | 0.60 | 0.54 |
| Out-of-sample period: 2000Q1–2020Q1 | | | | | | | | |
| Horizons/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.38 | 0.37 | 0.82 | 0.97 | 0.62 | 0.62 | 0.60 | 0.57 |
| $h = 4$ | 0.40 | 0.40 | 0.94 | 1.03 | 0.63 | 0.62 | 0.59 | 0.55 |
| $h = 8$ | 0.48 | 0.47 | 1.09 | 1.23 | 0.63 | 0.62 | 0.60 | 0.53 |
| Out-of-sample period: 2008Q1–2020Q1 | | | | | | | | |
| Horizon/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.48 | 0.47 | 0.97 | 1.07 | 0.63 | 0.63 | 0.61 | 0.59 |
| $h = 4$ | 0.52 | 0.51 | 1.07 | 1.15 | 0.63 | 0.61 | 0.63 | 0.55 |
| $h = 8$ | 0.63 | 0.62 | 1.18 | 1.40 | 0.67 | 0.67 | 0.65 | 0.53 |

Notes: $w_{ex} = 80$ and $w_{es} = 40$ denote the number of observations of the rolling window to generate the expectations and estimate the Phillips curve, respectively; $h = 1, 4, 8$ indicate the forecast horizons; $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$, and $\mathcal{M}_4$ denote the dual stickiness Phillips curve model without dependence (equation (9)), the dual stickiness Phillips curve model with dependence (equation (8)), the pure sticky price model (equation (10)) and the pure sticky information model (equation (11)), respectively; the numbers of DA denote the proportion of corrected predictions of the direction.

**Table B3**

Mean square error (MSE) and directional accuracy (DA). $w_{ex} = 80$ and $w_{es} = 50$, GDP deflator as inflation measure, and real marginal costs as forcing variable

| Out-of-sample period: 1995Q1–2020Q1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | MSE | | | | DA | | | |
| Horizons/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.41 | 0.42 | 1.05 | 1.14 | 0.59 | 0.57 | 0.57 | 0.50 |
| $h = 4$ | 0.42 | 0.42 | 1.08 | 1.19 | 0.60 | 0.60 | 0.56 | 0.50 |
| $h = 8$ | 0.42 | 0.43 | 1.17 | 1.36 | 0.60 | 0.60 | 0.60 | 0.50 |
| Out-of-sample period: 2000Q1–2020Q1 | | | | | | | | |
| Horizons/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.49 | 0.49 | 1.18 | 1.28 | 0.59 | 0.58 | 0.59 | 0.51 |
| $h = 4$ | 0.48 | 0.50 | 1.26 | 1.37 | 0.60 | 0.60 | 0.57 | 0.52 |
| $h = 8$ | 0.50 | 0.51 | 1.37 | 1.55 | 0.69 | 0.67 | 0.69 | 0.57 |
| Out-of-sample period: 2008Q1–2020Q1 | | | | | | | | |
| Horizons/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.64 | 0.63 | 1.15 | 1.31 | 0.63 | 0.63 | 0.65 | 0.53 |
| $h = 4$ | 0.64 | 0.64 | 1.15 | 1.29 | 0.67 | 0.67 | 0.63 | 0.55 |
| $h = 8$ | 0.64 | 0.64 | 1.27 | 1.50 | 0.69 | 0.67 | 0.69 | 0.57 |

Notes: $w_{ex} = 80$ and $w_{es} = 50$ denote the number of observations of the rolling window to generate the expectations and estimate the Phillips curve, respectively; $h = 1, 4, 8$ indicate the forecast horizons; $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$, and $\mathcal{M}_4$ denote the dual stickiness Phillips curve model without dependence (equation (9)), the dual stickiness Phillips curve model with dependence (equation (8)), the pure sticky price model (equation (10)) and the pure sticky information model (equation (11)), respectively; the numbers of DA denote the proportion of corrected predictions of the direction.

**Table B4**

Mean square error (MSE) and directional accuracy (DA). $w_{ex} = 60$ and $w_{es} = 40$, CPI inflation, and real marginal costs as forcing variable

| Out-of-sample period: 1995Q1–2020Q1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | MSE | | | | DA | | | |
| Horizons/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.93 | 0.97 | 1.67 | 1.74 | 0.60 | 0.58 | 0.55 | 0.55 |
| $h = 4$ | 0.94 | 0.97 | 1.67 | 1.77 | 0.59 | 0.63 | 0.55 | 0.55 |
| $h = 8$ | 0.94 | 0.97 | 1.68 | 1.83 | 0.62 | 0.61 | 0.54 | 0.56 |
| Out-of-sample period: 2000Q1–2020Q1 | | | | | | | | |
| Horizons/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 1.10 | 1.14 | 1.93 | 1.99 | 0.60 | 0.59 | 0.56 | 0.56 |
| $h = 4$ | 1.11 | 1.15 | 1.85 | 1.96 | 0.60 | 0.62 | 0.57 | 0.56 |
| $h = 8$ | 1.10 | 1.14 | 1.77 | 1.93 | 0.64 | 0.64 | 0.56 | 0.59 |
| Out-of-sample period: 2008Q1–2020Q1 | | | | | | | | |
| Horizon/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 1.42 | 1.48 | 1.94 | 1.89 | 0.63 | 0.61 | 0.61 | 0.59 |
| $h = 4$ | 1.26 | 1.31 | 1.98 | 1.91 | 0.59 | 0.61 | 0.61 | 0.59 |
| $h = 8$ | 1.33 | 1.38 | 2.17 | 2.13 | 0.59 | 0.61 | 0.57 | 0.57 |

Notes: $w_{ex} = 60$ and $w_{es} = 40$ denote the number of observations of the rolling window to generate the expectations and estimate the Phillips curve, respectively; $h = 1, 4, 8$ indicate the forecast horizons; $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$, and $\mathcal{M}_4$ denote the dual stickiness Phillips curve model without dependence (equation (9)), the dual stickiness Phillips curve model with dependence (equation (8)), the pure sticky price model (equation (10)) and the pure sticky information model (equation (11)), respectively; the numbers of DA denote the proportion of corrected predictions of the direction.

**Table B5**
Mean square error (MSE) and directional accuracy (DA). $w_{ex} = 60$ and $w_{es} = 40$, GDP deflator as inflation measure, and output gap as forcing variable

| Out-of-sample period: 1995Q1–2020Q1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | MSE | | | | DA | | | |
| Horizons/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.41 | 0.43 | 1.23 | 1.35 | 0.65 | 0.67 | 0.59 | 0.52 |
| $h = 4$ | 0.44 | 0.45 | 1.15 | 1.31 | 0.64 | 0.65 | 0.59 | 0.53 |
| $h = 8$ | 0.46 | 0.46 | 0.91 | 1.07 | 0.67 | 0.66 | 0.66 | 0.58 |
| Out-of-sample period: 2000Q1–2020Q1 | | | | | | | | |
| Horizons/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.49 | 0.51 | 1.37 | 1.55 | 0.67 | 0.68 | 0.62 | 0.52 |
| $h = 4$ | 0.53 | 0.54 | 1.29 | 1.52 | 0.65 | 0.66 | 0.62 | 0.54 |
| $h = 8$ | 0.56 | 0.56 | 1.07 | 1.28 | 0.70 | 0.68 | 0.67 | 0.58 |
| Out-of-sample period: 2008Q1–2020Q1 | | | | | | | | |
| Horizon/*models* | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ | $\mathcal{M}_1$ | $\mathcal{M}_2$ | $\mathcal{M}_3$ | $\mathcal{M}_4$ |
| $h = 1$ | 0.57 | 0.58 | 1.11 | 1.39 | 0.71 | 0.73 | 0.67 | 0.55 |
| $h = 4$ | 0.64 | 0.65 | 1.18 | 1.51 | 0.69 | 0.69 | 0.65 | 0.55 |
| $h = 8$ | 0.71 | 0.72 | 1.04 | 1.35 | 0.73 | 0.69 | 0.71 | 0.59 |

Notes: $w_{ex} = 60$ and $w_{es} = 40$ denote the number of observations of the rolling window to generate the expectations and estimate the Phillips curve, respectively; $h = 1, 4, 8$ indicate the forecast horizons; $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$, and $\mathcal{M}_4$ denote the dual stickiness Phillips curve model without dependence (equation (9)), the dual stickiness Phillips curve model with dependence (equation (8)), the pure sticky price model (equation (10)) and the pure sticky information model (equation (11)), respectively; the numbers of DA denote the proportion of corrected predictions of the direction.

## References


Akerlof, G.A., Yellen, J.L., 1985. A near-rational model of the business cycle, with wage and price inertia. Q. J. Econ. 100 (Supplement), 823–838.

Arslan, M.M., 2010. Relative importance of sticky prices and sticky information in price setting. Econ. Modell. 27 (5), 1124–1135.

Ball, L., Mankiw, N.G., 1994. A sticky price manifesto. Carnegie-Rochester Conf. Ser. Public Policy 41, 127–151.

Bauwens, L., Korobilis, D., 2013. Bayesian methods. In: Hashimzade, N., Thornton, M.A. (Eds.), Handbook of Research Methods and Applications in Empirical Macroeconomics. Edward Elgar Publishing, Cheltenham, pp. 363–380.

Bergmeir, C., Costantini, M., Benítez, J., 2014. On the usefulness of cross-validation for directional forecast evaluation. Comput. Stat. Data Anal. 76, 132–143.

Bewley, T.F., 2002. Knightian decision theory. Part i. Decis. Econ. Finance 25 (2), 79–110.

Bilbiie, F.O., 2021. Monetary neutrality with sticky prices and free entry. Rev. Econ. Stat. 103 (3), 492–504.

Blaskowitz, O., Herwartz, H., 2009. Adaptive forecasting of the EURIBOR swap term structure. J. Forecast. 28, 575–594.

Blaskowitz, O., Herwartz, H., 2011. On economic evaluation of directional forecasts. Int. J. Forecast. 27, 1058–1065.

Blaskowitz, O., Herwartz, H., 2014. Testing the value of directional forecasts in the presence of serial correlation. Int. J. Forecast. 30, 30–42.

Carrera, C., Ramírez-Rondán, N., 2019. Inflation, information rigidity, and the sticky information phillips curve. Macroecon. Dyn. 23, 2597–2615.

Carrillo, J.A., 2012. How well does sticky information explain the dynamics of inflation, output, and real wages? J. Econ. Dynam. Contr. 36 (6), 830–850.

Chen, Q., Costantini, M., Deschamps, B., 2016. How accurate are professional forecasts in Asia? Evidence from ten countries. Int. J. Forecast. 32, 154–167.

Cogley, T., Sbordone, A.M., 2008. Trend inflation, indexation, and inflation persistence in the New Keynesian Phillips curve. Am. Econ. Rev. 98 (5), 2101–2126.

Coibion, O., 2006. Inflation inertia in sticky information models. B E J. Macroecon. 6 (1), 1–29.

Coibion, O., 2010. Testing the sticky information phillips curve. Rev. Econ. Stat. 92 (1), 87–101.

Coibion, O., Gorodnichenko, Y., 2011. Strategic interaction among heterogeneous price-setters in an estimated DSGE model. Rev. Econ. Stat. 93 (3), 920–940.

Costantini, M., Cuaresma, J., Hlouskova, J., 2016. Forecasting errors, directional accuracy and profitability of currency trading: the case of EUR/USD exchange rate. J. Forecast. 35, 652–668.

Diebold, F., 2007. Elements of Forecasting, fourth ed. Thomson, Ohio, US.

Dupor, B., Kitamura, T., Tsuruga, T., 2010. Integrating sticky prices and sticky information. Rev. Econ. Stat. 92 (3), 657–669.

Foellmi, R., Jaeggi, A., Rosenblatt-Wisch, R., 2019. Loss aversion at the aggregate level across countries and its relation to economic fundamentals. J. Macroecon. 61, 103136.

Galí, J., Gertler, M., 1999. Inflation dynamics: a structural econometric analysis. J. Monetary Econ. 44 (2), 195–222.

Geyer, C.J., 1994. Estimating Normalizing Constants and Reweighting Mixtures in Markow Chain Monte Carlo. Technical Report 568. University of Minnesota, School of Statistics.

Giandrea, M.D., Sprague, S.A., 2017. Estimating the US labor share. Mon. Labor Rev. February 2017, 1–18.

Granville, B., Zeng, N., 2019. Time variation in inflation persistence: new evidence from modelling US inflation. Econ. Modell. 81, 30–39.

Harris, M.N., Le Bihan, H., Sevestre, P., 2020. Identifying price reviews by firms: an econometric approach. J. Money Credit Bank. 52 (2–3), 293–322.

Joe, H., 2014. Dependence Modeling with Copulas. CRC Press, Taylor & Francis Group, Boca Raton, FL.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90 (430), 773–795.

Keen, B.D., 2007. Sticky price and sticky information price-setting models: what is the difference? Econ. Inq. 45 (4), 770–786.

Kiley, M.T., 2007. A quantitative comparison of sticky-price and sticky-information models of price setting. J. Money Credit Bank. 39, 101–125.

Kim, I., Kim, Y.S., 2019. Inattentive agents and inflation forecast error dynamics: a Bayesian DSGE approach. J. Macroecon. 62, 103139.

Klenow, P.J., Willis, J.L., 2007. Sticky information and sticky prices. J. Monetary Econ. 54 (Suppl. 1), 79–99.

Knotek, E.S., 2010. A tale of two rigidities: sticky prices in a sticky-information environment. J. Money Credit Bank. 42 (8), 1543–1564.

Lawless, M., Whelan, K.T., 2011. Understanding the dynamics of labor shares and inflation. J. Macroecon. 33 (2), 121–136.

Leitch, G., Tanner, J.E., 1995. Professional economic forecasts: are they worth their costs? Journal of Jorecasting 14, 142–157.

Mallick, S.K., Sousa, R.M., 2013. Commodity prices, inflationary pressures, and monetary policy: evidence from BRICS economies. Open Econ. Rev. 24 (4), 677–694.

Mankiw, N.G., Reis, R., 2002. Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve. Q. J. Econ. 117 (4), 1295–1328.

Marshall, A.W., Olkin, I., 1985. A family of bivariate distributions generated by the bivariate Bernoulli distribution. J. Am. Stat. Assoc. 80 (390), 332–338.

McKnight, S., Mihailov, A., Rumler, F., 2020. Inflation forecasting using the new keynesian phillips curve with a time-varying trend. Econ. Modell. 87, 383–393.

Öller, L.E., Barot, B., 2000. The accuracy of European growth and inflation forecasts. Int. J. Forecast. 16, 293–315.

Reis, R., 2006. Inattentive producers. Rev. Econ. Stud. 73 (3), 793–821.

Rotemberg, J.J., Woodford, M., 1999. The cyclical behaviour of prices and costs. In: Taylor, J.B., Woodford, M., Uhlig, H. (Eds.), Handbook of Macroeconomics. Elsevier, Amsterdam, pp. 1051–1135.

Rudd, J.B., Whelan, K., 2005. Does labor's share drive inflation? J. Money Credit Bank. 37 (2), 297–312.

Samuelson, W., Zeckhauser, R., 1988. Status quo bias in decision making. J. Risk Uncertain. 1 (1), 7–59.

Sautua, S.I., 2017. Does uncertainty cause inertia in decision making? An experimental study of the role of regret aversion and indecisiveness. J. Econ. Behav. Organ. 136, 1–14.
Schorfheide, F., 2010. Bayesian methods in macroeconometrics. In: Durlauf, S.N., Blume, L.E. (Eds.), Macroeconometrics and Time Series Analysis. Palgrave Macmillan, London, pp. 28–34.
Sims, C.A., 2003. Implications of rational inattention. J. Monetary Econ. 50 (3), 665–690.
Sinclair, T., Stekler, H., Kitzinger, L., 2010. Directional forecasts of GDP and inflation: a joint evaluation with an application to Federal Reserve predictions. Appl. Econ. 42, 2289–2297.
Swanson, N.R., White, H., 1997a. Forecasting economic time series using flexible versus fixed specification and linear versus nonlinear econometric models. Int. J. Forecast. 13, 439–461.
Swanson, N.R., White, H., 1997b. A model selection approach to realtime macroeconomic forecasting using linear models and artificial neural networks. Review of Economics and Business Statistics 79, 540–550.
Wheelock, D.C., Wohar, M.E., 2009. Can the term spread predict output growth and recessions? A survey of the literature. Fed. Reserv. Bank St. Louis Rev. 91 (5, Part 1), 419–440 5, 91.
Zbaracki, M.J., Ritson, M., Levy, D., Dutta, S., Bergen, M., 2004. Managerial and customer costs of price adjustment: direct evidence from industrial markets. Rev. Econ. Stat. 86 (2), 514–533.