

Dynare Working Papers Series
<https://www.dynare.org/wp/>

Intrinsic persistence of wage inflation in New Keynesian models of the business cycles

Giovanni Di Bartolomeo
Marco Di Pietro

Working Paper no. 55

February 2020

CEPREMAP

CENTRE POUR LA RECHERCHE ECONOMIQUE ET SES APPLICATIONS

48, boulevard Jourdan — 75014 Paris — France

<https://www.cepremap.fr>

Intrinsic persistence of wage inflation in New Keynesian models of the business cycles*

Giovanni Di Bartolomeo

giovanni.dibartolomeo@uniroma1.it

Sapienza University of Rome

Marco Di Pietro

marco.dipietro@uniroma1.it

Sapienza University of Rome

Abstract

Our paper derives and estimates a New Keynesian wage Phillips curve that accounts for intrinsic inertia. Our approach considers a wage-setting model featuring an upward-sloping hazard function, that is based on the notion that the probability of resetting a wage depends on the time elapsed since the last reset. According to our specification, we obtain a wage Phillips curve that also includes backward-looking terms, which account for persistence. We test the slope of the hazard function using GMM estimation. Then, placing our equation in a small-scale New Keynesian model, we investigate its dynamic properties using Bayesian estimation. Model comparison shows that our model outperforms commonly used alternative methods to introduce persistence.

JEL classification: E24, E31, E32, C11.

Keywords: duration-dependent wage adjustments, intrinsic inflation persistence, DSGE models, hybrid Phillips curves, model comparison.

1 Introduction

Micro empirical evidence suggests that nominal wages are sticky and that wage inflation is persistent (Barattieri *et al.*, 2014). In aggregate models, these imperfections play an important role both in transmitting monetary policy and in our understanding of business cycle fluctuations (Christiano *et al.*, 2005; Rabanal and Rubio-Ramirez, 2005; Olivei and Tenreyro, 2010). If wage stickiness is ignored, macroeconomic models are unable to mimic the inertial dynamics of output that are observed in the data, unless implausibly high price stickiness is assumed (Christiano *et al.*,

*The authors are grateful to two anonymous referees and the Editor Pok-sang Lam for excellent comments and insightful suggestions. We would like to thank Klaus Adam, Barbara Annichiarico, Guido Ascari, Efrem Castelnuovo, Bob Chirinko, Jordi Galí, Michel Juillard, Francesco Lippi, Peter McAdam, Chiara Perricone, Ricardo Reis, Tiziano Ropele, Lorenza Rossi, Massimiliano Tancioni, Patrizio Tirelli. We have benefited from comments on various workshops and conferences including the MBF workshop (University of Milano Bicocca), CMC workshop (Sapienza University of Rome), ICMAIF (University of Crete), ZEW Conference on Recent Developments in Macroeconomics (Mannheim), CGBCR (University of Manchester) and Dynare Conference (National Bank of Belgium, Brussels). The authors also acknowledge financial support by Sapienza University of Rome.

2005). Moreover, the inertial structural component of wage inflation also affects the persistence of price inflation, which is present in the data (Fuhrer and Moore, 1995; Fuhrer, 2011). All else equal, price inflation depends on expected future real marginal cost, which in turn depends on wages.

In macroeconomic models, stickiness and the persistence of wage inflation are typically captured by assuming wage adjustment processes *à la* Calvo and some forms of indexation to past price inflation (e.g., Erceg *et al.*, 2000; Christiano *et al.*, 2005; Rabanal and Rubio-Ramirez, 2005; Smets and Wouters, 2007). The former accounts for wage stickiness, whereas the latter introduces intrinsic inflation persistence.¹ The successful implementation of these assumptions is mainly related to their simplicity and tractability (Tsoukis *et al.*, 2009). However, both assumptions seem to be rejected by the micro evidence on the wage-setting process.

In a recent study of the U.S., Barattieri *et al.* (2014) use quantitative data from the Survey of Income and Program Participation and find that wages are sticky but that the hazard function of a nominal wage change is not constant. This finding stands in stark contrast with the Calvo mechanism, whereby the probability of changing a wage is unrelated to the time elapsed from the last wage reset (i.e., the hazard function is flat). In the Barattieri *et al.* (2014) sample, the hazard function is initially increasing, with a peak at twelve months, which signals that the probability of observing a wage reset positively depends on the time elapsed from the last wage adjustment, i.e., newer wages are stickier than older. Moreover, the probability of a wage change does not vary across quarters, i.e., the wage reset process is unaffected by seasonality.

Barattieri *et al.* (2014) also reject wage indexation and instead find that only a fraction of wages are reset in every period, whereas the remaining wages are left unchanged, a finding that is at odds with the assumption of wage indexation, which entails that all wages are updated in every period. Moreover, the degree of wage indexation largely varies across time (Holland, 1988) and seems to be endogenously determined by business cycle fluctuations or other factors (Hofmann *et al.*, 2010; Acocella *et al.*, 2015). As a result, the degree of wage indexation might not be a structural parameter *à la* Lucas (1976). In general, inflation indexation can be considered an *ad hoc* assumption to introduce persistence because it is not supported by the survey evidence (see, e.g., Dhyne *et al.*, 2005; Fabiani *et al.*, 2005).

In light of the foregoing, we propose a different approach for modeling the wage adjustment process and introduce wage inflation persistence. Our starting point is Sheedy (2007, 2010),² who shows that a Phillips curve with any number of lags in past inflation rates can be obtained by assuming a positive hazard function in price setting.³ We borrow Sheedy’s mechanism and extend it to wage setting.

¹Following Fuhrer (2011), by “intrinsic persistence” we refer to the inertia that does not depend on real activity, but that is proper for the inflation process, i.e., the inertia that would be present even if the driving variable of the Phillips curve, (e.g., the output gap or real marginal cost) were not persistent (see also Angeloni *et al.*, 2006; Rudd and Whelan, 2006).

²Price adjustments with non-constant hazard functions are considered by many other papers— including Taylor (1980), Goodfriend and King (1997), Dotsey *et al.* (1999), Wolman (1999), Guerrieri (2001, 2006), and Mash (2004). These models are based on state or time-dependent assumptions and focus on price dynamics.

³Micro evidence on price setting regarding the slope of the hazard function is mixed. The results depend on the sample, countries, periods considered and methodologies used (see, e.g., Cecchetti, 1986; Nakamura and Steinsson, 2013). For a detailed discussion on the positive hazard function based on macroeconomic evidence, see instead Sheedy (2007, 2010) and Yao (2011).

We assume that the length of a wage spell directly influences the reset probability. If the slope of the hazard is positive (negative), the probability of posting a new wage increases (decreases) for the wages that have been left fixed for many periods. We derive a wage equation that accounts for stickiness and intrinsic wage inflation inertia, by considering a non-constant hazard function without assuming that all the wages are adjusted in every period (as with indexation)—in line with Barattieri *et al.* (2014). After deriving the wage Phillips curve, we perform a single equation estimation using generalized method of moments (GMM) and then embed the wage equation in a small-scale DSGE model which we estimate employing Bayesian techniques.⁴

Our paper makes both theoretical and empirical contributions to the literature on wage-setting behavior and inflation persistence. To our knowledge, ours is the first attempt to capture wage intrinsic persistence by assuming a positive hazard function and the first to estimate the resulting wage Phillips curve with macro data. Our main results can be summarized as follows.

We extend the Sheedy (2007, 2010) price adjustment mechanism to the wage-setting process. Assuming duration-dependent reset probabilities, we analytically derive a forward-looking wage Phillips curve that also embeds past terms for wage inflation rates. In so doing, we provide microfoundations and a theoretical justification for intrinsic wage inflation persistence that is not at odds with the micro evidence.

We estimate our wage Phillips curve as a single equation using GMM. We find that the estimated hazard parameters are positive and statistically significant, which offers evidence that a positive hazard function emerges for wage changes at the macro aggregate level. Our estimation also provides evidence for intrinsic—versus extrinsic—wage inflation persistence.

Including our equation in a DSGE macro model, we generalize Erceg *et al.* (2000, EHL henceforth) to account for possible duration-dependent wage adjustments. In estimating our DSGE macro model, we find that hazard gradients are positive for both prices and wages—confirming our GMM results and those of Sheedy (2007, 2010). Following Benati (2008, 2009), we then test the robustness of vintage-dependent adjustments to policy regime shifts. By considering sub-samples, we find that the parameters encoding intrinsic persistence also remain significantly different from zero also during the Great Moderation. Moreover, the parameters governing wage adjustments do not change significantly across regimes.

Finally, our model outperforms alternative specifications for price and wage adjustments. Following Rabanal and Rubio-Ramirez (2005), we evaluate the empirical performance of different models by comparing marginal likelihoods (via the Bayes factor). As alternatives, we consider flat hazard functions (price and wage Phillips curves *à la* Calvo) and past or dynamic indexation mechanisms (see Galí and Gertler, 1999; Christiano *et al.*, 2005).⁵

The remainder of the paper is organized as follows. In the next section, we introduce the hazard function and show the derivation and estimation of our duration-dependent wage Phillips curve. Section 3 presents a DSGE simple small-scale model characterized by a price and wage

⁴With respect to price adjustment in macro models with rational expectations, there has been a long debate regarding single equation estimation versus full model specification. See, e.g., Galí *et al.* (2005), Lindé (2005), Mavroeidis (2005), Rudd and Whelan (2005). We consider both approaches for the sake of robustness.

⁵Similar results are found by Laforte (2007) for price setting. In terms of the Bayes factor, he finds that the predictive ability of a model with positive hazard functions (Wolman, 1999) is substantially higher than that of models with indexation (Smets and Wouters, 2007) and sticky information (Mankiw and Reis, 2002).

Phillips curve that can account for intrinsic inflation persistence. Section 4 provides our model estimations and compares them to commonly used alternatives based on different types of inflation indexation. The final section concludes and offers some future lines of research.

2 Wage Phillips curve and intrinsic persistence

This section illustrates the main characteristics of a hazard function and shows how to derive a wage Phillips curve assuming that wages are reset following a vintage-dependent mechanism. Moreover, we perform the GMM estimation for our Phillips curve to test whether the sign of the hazard slope is positive and to assess the statistical significance of the Phillips curve coefficients.

2.1 Hazard function and duration-dependent adjustment

According to Sheedy (2007),⁶ the probability of a wage adjustment is not random (as in the Calvo specification) but instead depends on the time elapsed since the last wage reset. Therefore, the probability of posting a new wage is not equal among households, but is a positive function of duration. Formally, wage adjustments are defined using a hazard function, which expresses the relationship between the probability of updating a wage and the duration of wage stickiness. The hazard function is defined by the sequence of probabilities $\{\alpha_{w,l}\}_{l=1}^{\infty}$, where $\alpha_{w,l}$ represents the probability of resetting a wage that has remained unchanged for l periods. The hazard function is specified as follows:⁷

$$\alpha_{w,l} = \alpha_w + \varphi_w (1 - \alpha_{w,l-1})^{-1}, \quad \text{for } l > 1 \quad (1)$$

where α_w is the initial value of the hazard function (for $l = 1$) and φ_w is its slope. In what follows we assume that only one parameter controls the slope of the hazard,⁸ as described below:

$$\begin{cases} \varphi_w = 0, & \longrightarrow \text{the hazard is flat (Calvo case);} \\ \varphi_w > 0, & \longrightarrow \text{the hazard is upward-sloping;} \\ \varphi_w < 0, & \longrightarrow \text{the hazard is downward-sloping.} \end{cases} \quad (2)$$

Thus, the hazard is positive if $\varphi_w > 0$. As discussed above, a positive hazard function translates into a higher probability of updating a wage that last reset many periods ago. Equation (1) helps us to grasp the intuition about this point: Whenever $\varphi_w > 0$, $\alpha_{w,l}$ shifts upward, implying that $\alpha_{w,l+1} > \alpha_{w,l}$; then, older wages will more likely be reset than newer wages.

Each hazard function is related to a survival function, which expresses the probability that a wage remains fixed for l periods. As for the hazard, the survival function is defined by a sequence of probabilities: $\{\varsigma_{w,l}\}_{l=0}^{\infty}$, where $\varsigma_{w,l}$ denotes the probability that a wage fixed at time t will

⁶Notably, the hazard function modeled in Sheedy (2010) is equivalent to that modeled in Sheedy (2007). The differences are based only on parameterization choice. Both hazard functions lead to the same Phillips curve specification.

⁷Further details about the hazard function properties are provided in Appendix A.

⁸In Appendix A, we provide the general case in which the sequence $\{\varphi_{w,l}\}_{l=1}^n$ affects the hazard gradient.

remain in use at time $t + l$. Formally, the survival function is defined as follows:

$$\varsigma_{w,l} = \prod_{h=1}^l (1 - \alpha_{w,h}) \quad (3)$$

where $\varsigma_{w,0} = 1$.

By making use of (3), we can rewrite the non-linear recursion (1) for the wage adjustment probabilities as a linear recursion for the corresponding survival function:

$$\varsigma_{w,l} = (1 - \alpha_w)\varsigma_{w,l-1} - \varphi_w\varsigma_{w,l-2}, \quad \text{for } l > 1 \quad (4)$$

where $\varsigma_{w,1} = (1 - \alpha_w)$ for $l = 1$.⁹ Let $\theta_{w,t}$ denote the proportion of households at time t earning a wage posted at period $t - l$. The sequence $\{\theta_{w,t}\}_{l=0}^{\infty}$ indicates the distribution of the duration of wage stickiness at time t . If both the hazard function and the evolution over the time of the distribution of wage duration satisfy certain conditions,¹⁰ the following three relations are obtained:

$$\begin{cases} \theta_{w,l} = (\alpha_w + \varphi_w)\varsigma_{w,l} \\ \alpha_w^e = \alpha_w + \varphi_w \\ D_w^e = \frac{1 - \varphi_w}{\alpha_w + \varphi_w} \end{cases} \quad (5)$$

where $\theta_{w,l}$ represents the unique stationary distribution to which the economy always converges, α_w^e indicates the unconditional probability of a wage reset and D_w^e denotes the unconditional expected duration of wage stickiness.

2.2 Duration-dependent wage Phillips curve derivation

The supply side of the economy we consider is fairly standard (see EHL, 2000)¹¹ and consists of a continuum of monopolistically competitive firms indexed on the unit interval $\Omega \equiv [0, 1]$. The production function of the representative firm $i \in \Omega$ is described by a Cobb-Douglas without capital:

$$Y_t(i) = A_t N_t(i)^{1-\phi}, \quad (6)$$

where $Y_t(i)$ is the output of good i at time t , A_t represents the state of technology, $N_t(i)$ is the quantity of labor employed by i -firm and $1 - \phi$ measures the elasticity of output with respect to labor. The quantity of labor used by firm i is defined by:

$$N_t(i) = \left[\int_{\Omega} N_t(i, j)^{\frac{\varepsilon_w - 1}{\varepsilon_w}} dj \right]^{\frac{\varepsilon_w}{\varepsilon_w - 1}} \quad (7)$$

where $N_t(i, j)$ is the quantity of j -type labor employed by firm i in period t , and ε_w denotes the elasticity of substitution between workers. Cost minimization with respect to the quantity of labor

⁹It derives from (3).

¹⁰See Appendix A for all the restrictions that must be satisfied.

¹¹The notation we used here is similar to that in Galí's textbook (Chapter 6, 2015).

employed yields to the labor demand schedule:

$$N_t(i, j) = \left(\frac{W_t(j)}{W_t} \right)^{-\varepsilon_w} N_t(i) \quad (8)$$

where $W_t(j)$ is the nominal wage paid to j -type worker and W_t is the aggregate wage index defined as follows:

$$W_t = \left[\int_0^1 W_t(j)^{1-\varepsilon_w} dj \right]^{\frac{1}{1-\varepsilon_w}} \quad (9)$$

We consider a continuum of monopolistically competitive households indexed on the unit interval $\Theta \equiv [0, 1]$. Each household supplies a different type of labor, $N_t(j) = \int_0^1 N_t(i, j) di$ to all the firms. The representative household $j \in \Theta$ chooses the quantity of labor $N_t(j)$ to supply in order to maximize the following separable utility:

$$U(C_t(j), N_t(j)) = E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left[G_t \frac{(C_t(j) - hC_{t-1}(j))^{1-\sigma}}{1-\sigma} - \frac{N_t(j)^{1+\gamma}}{1+\gamma} \right] \right\} \quad (10)$$

where E_0 is the expectation operator conditional on time $t = 0$ information, β is the discount factor, σ denotes the relative risk aversion coefficient, γ is the inverse of Frisch labor supply elasticity and h is an internal habit on consumption. Finally, G_t is a preference shock that affects the marginal utility of consumption and is assumed to follow an $AR(1)$ stationary process. Assuming complete financial markets, the household faces a standard budget constraint specified in nominal terms as follows:

$$P_t C_t(j) + E_t [Q_{t+1,t} B_t(j)] \leq B_{t-1}(j) + W_t(j) N_t(j) + T_t(j) \quad (11)$$

where P_t is the price of the consumption good, $B_t(j)$ denotes the holdings of one-period nominally riskless state-contingent bonds purchased in period t and maturing in period $t + 1$, Q_t is the bond price, T_t represents a lump-sum government nominal transfer. Finally, $C_t(j)$ represents the consumption of household j and is described by a CES aggregator: $C_t(j) = \left[\int_{\Theta} C_t(i, j)^{\frac{\varepsilon_p - 1}{\varepsilon_p}} di \right]^{\frac{\varepsilon_p}{\varepsilon_p - 1}}$, where $C_t(i, j)$ denotes the quantity of i -type good consumed by household j , and ε_p is the elasticity of substitution between goods.

In our framework, households are wage setters. In setting wages, each maximizes (10) taking account of (11) and internalizes the effects of aggregate labor demand. Households are subject to a random probability of updating their wage, but, according to our duration-dependent mechanism, a wage change will be more likely to be observed when the last wage reset occurred far in the past. Formally, suppose that at time t , a household sets a new wage that is denoted by W_t^* ,¹² if the household still earns this wage at time $\tau \geq t$, its real wage will be W_t^*/P_{τ} . By considering the

¹²Because each household solves the same optimization problem, index j is henceforth omitted.

survival function, the household will then choose its optimal reset wage by solving:

$$\max_{W_t^*} \sum_{\tau=t}^{\infty} (\beta^{\tau-t} \varsigma_{w,\tau-t}) E_t \left[U_{c,\tau} \frac{W_t^*}{P_\tau} N_{\tau|t} - \frac{N_{\tau|t}^{1+\gamma}}{1+\gamma} \right] \quad (12)$$

where $U_{c,\tau}$ indicates the marginal utility of consumption and $N_{\tau|t}$ denotes the level of employment in period τ among workers whose last wage reset was in period t . This maximization is subject to the budget constraint (11) and the aggregate labor demand schedule (obtained by integrating (8) over all firms i). Equation (12) yields the following first-order condition:

$$\sum_{\tau=t}^{\infty} (\beta^{\tau-t} \varsigma_{w,\tau-t}) E_t \left\{ N_{\tau|t} \left[U_{c,\tau} \frac{W_t^*}{P_\tau} + \mu_w U_{n,\tau|t} \right] \right\} = 0 \quad (13)$$

where $U_{n,\tau|t}$ is the marginal utility of labor, and $\mu_w = \frac{\varepsilon_w}{\varepsilon_w - 1}$ represents the desired wage mark-up. Considering that $MRS_{\tau|t} = -\frac{U_{n,\tau|t}}{U_{c,\tau}}$ defines the marginal rate of substitution between consumption and labor in period τ for the workers posting a new wage in time t , equation (13) can be expressed as follows:

$$\sum_{\tau=t}^{\infty} (\beta^{\tau-t} \varsigma_{w,\tau-t}) E_t \left\{ N_{\tau|t} U_{c,\tau} \left[\frac{W_t^*}{P_\tau} - \mu_w MRS_{\tau|t} \right] \right\} = 0 \quad (14)$$

Assuming that the economy has converged to $\{\theta_{w,l}\}_{l=0}^{\infty}$, wage level (9) can be expressed as a weighted-average of past reset wages:

$$W_t = \left(\sum_{l=0}^{\infty} \theta_{w,l} W_{t-l}^{*1-\varepsilon_w} \right)^{\frac{1}{1-\varepsilon_w}} \quad (15)$$

As is common practice in DSGE models, we log-linearize (14) and (15) around a deterministic steady state. Specifically, there is no trend inflation, i.e., $\Pi^w = 1$ and $\Pi^p = 1$, where Π^w and Π^p represent the steady state of the wage and price inflations, respectively. This assumption implies that the steady state value for the relative reset wage is 1 and that the steady state of the real interest rate is equal to β^{-1} . In this manner, we thus obtain:¹³

$$w_t^* = \sum_{\tau=t}^{\infty} \left(\frac{\beta^{\tau-t} \varsigma_{w,\tau-t}}{\sum_{j=0}^{\infty} \beta^j \varsigma_{w,j}} \right) E_t [w_\tau - \Xi_w \mu_\tau^w] \quad (16)$$

where $\Xi_w = \frac{1}{1+\varepsilon_w \gamma}$, μ_τ^w denotes the deviations of the economy's average wage mark-up from its desired level, i.e., a mark-up on the marginal rate of substitution, and

$$w_t = \sum_{l=0}^{\infty} \theta_{w,l} w_{t-l}^* \quad (17)$$

Equations (16) and (17) describe the wage adjustment mechanism. The duration-dependent wage Phillips curve is derived by combining them with (4) and (5).

¹³Lower-case letters denote log-deviations from the steady state.

Specifically, inserting (4) into (16), we obtain the following:

$$w_t^* = \beta(1 - \alpha_w)E_t w_{t+1}^* - \beta^2 \varphi_w E_t w_{t+2}^* + [1 - \beta(1 - \alpha_w) + \beta^2 \varphi_w] (w_t - \Xi_w \mu_t^w) \quad (18)$$

By making use of (5), equation (17) can be recast as follows:

$$w_t = (1 - \alpha_w) w_{t-1} - \varphi_w w_{t-2} + (\alpha_w + \varphi_w) w_t^* \quad (19)$$

where we have used the fact that the stationary distribution of the wage duration (5) can be rewritten recursively as:

$$\theta_{w,l} = (1 - \alpha_w)\theta_{w,l-1} - \varphi_w \theta_{w,l-2} \quad \text{for } l > 1 \quad (20)$$

where $\theta_{w,0} = \alpha_w + \varphi_w$ and $\theta_{w,1} = (1 - \alpha_w)(\alpha_w + \varphi_w)$ because of (5) and (3).

The general expression for the wage Phillips curve is obtained from (18) and (19):

$$\pi_t^w = \psi_w \pi_{t-1}^w + \beta [1 + (1 - \beta) \psi_w] E_t \pi_{t+1}^w - \beta^2 \psi_w E_t \pi_{t+2}^w - k_w \mu_t^w, \quad (21)$$

where π_t^w is the wage inflation. Moreover:

$$\begin{cases} \psi_w = \frac{\varphi_w}{(1 - \alpha_w) - \varphi_w [1 - \beta(1 - \alpha_w)]} \\ k_w = \frac{(\alpha_w + \varphi_w) [1 - \beta(1 - \alpha_w) + \beta^2 \varphi_w]}{(1 - \alpha_w) - \varphi_w [1 - \beta(1 - \alpha_w)]} \Xi_w \end{cases}, \quad (22)$$

where ψ_w and k_w are coefficients depending on the hazard parameters.¹⁴ In particular, φ_w and α_w control the slope and the initial level of the hazard function, respectively.

As in the standard case, equation (21) describes a negative relation between current wage inflation and the wage mark-up: A negative μ_t^w involves the presence of average wage mark-up below the desired level, inducing households allowed to post a new wage to raise the latter, thus generating positive wage inflation. This relation is affected by expectations over future wage inflation¹⁵ due to the wage stickiness.

The novelty of our approach is the introduction of intrinsic wage inertia. Our wage Phillips curve has a “history dependent” dimension, as it establishes that current wage inflation also depends on past wage inflation. Unlike the case involving indexation, here the backward term is on wage inflation—and not on price inflation—which indicates a “purely” intrinsic inertia, i.e., wage inflation is driven by its own lags with positive coefficients. The presence of an endogenous lagged term in (21) is not obtained from an *ad hoc* assumption, but instead has a clear theoretical foundation that derives from a positive “selection effect” stemming from our pricing mechanism. By positive selection effect we mean that wage setters who have not made a wage change for a long time are more likely to post a new wage than wage setters who recently did it. The selection effect works through the duration of a wage spell and generates wage inflation persistence. The

¹⁴We assumed that only one parameter controls the slope of the hazard. The evolution of the wage Phillips curve coefficients for the general case in which n parameters affect the hazard gradient is shown in Appendix A.

¹⁵Both inflation at time $t + 1$ and $t + 2$ are relevant. Although the coefficient associated with the latter is negative, the overall effect of expected inflation is positive on its current rate. The second-order term in the difference equation thus captures the dynamics of the adjustment process. See Sheedy (2007) for a discussion.

intuition can be explained as follows.

The wage setters aim to adjust nominal wages to close the gap between the actual relative wage and the desired wage. Thus, they react to all the shocks that create a wedge between the two. In sticky wage models, after a temporary shock has vanished, incentives remain to adjust wages. Specifically, when a temporary wage inflationary shock hits the economy, some wage setters are not able to raise their wages, because of the presence of nominal rigidities, but the average wage level increases because others will do it. Once the shock has dissipated, some wage setters will meet a higher wage than desired (i.e., those wage setters who have been able to adjust previously), whereas others will meet lower wages because wage inflation occurred (i.e., those who have not been able to adjust their wages). It follows that the latter will attempt to raise their nominal wages to maintain the desired relative wage (“catch-up” effect) and the former to reduce them because now their relative wage is too high (“roll-back” effect). If both groups have the same probability to readjust wages, as in Calvo, the two effects offset one another, but this does occur under non-constant hazard.

In our case, wages that remained unvaried for longer periods have a larger probability to be changed than newly set wages (positive selection effect). Thus, it follows that, after a rise in the wage inflation, further increases are likely (the “catch-up” effect prevails over the “roll-back”). As a result, aggregate wage inflation remains positive after the shock has vanished, generating wage inflation persistence. This is captured by the presence of the past wage inflation rate in (21). The reason behind the positive selection effect is that the average gains from adjusting wages are expected to be higher the longer a wage has remained unchanged.

Our wage equation encompasses the Calvo model’s purely forward-looking specifications as a particular case when the hazard is flat (i.e., $\varphi_w = 0$), which implies that ψ_w drops to zero and, consequently, we return to the standard textbook case.

2.3 Hazard function estimation

As in Sheedy (2007, 2010), we estimate our wage Phillips curve via GMM, to inspect the shape of the hazard function and the statistical significance of the coefficients attached to the lead and lag components of the wage equation. For simplicity’s sake, we show the estimation of (21) when only one parameter affects the hazard slope.¹⁶ As it is not easy to find an observable proxy for the wage mark-up, the latter is expressed as a function of the unemployment rate, following Galí (2011):

$$\mu_t^w = \gamma u_t \tag{23}$$

where u_t indicates the unemployment rate. Therefore, (21) becomes:

$$\pi_t^w = \psi_w \pi_{t-1}^w + \beta [1 + (1 - \beta) \psi_w] E_t \pi_{t+1}^w - \beta^2 \psi_w E_t \pi_{t+2}^w - k_w \gamma u_t \tag{24}$$

To perform a GMM estimation of (24), we must use a set of instruments and impose an orthogonality condition. Let z_{t-1} represent a vector of observable variables known at time $t - 1$.

¹⁶For the more general case, we find that the extra leads and lags deriving from this specification are not statistically significant.

Under rational expectations the error forecast of π_t^w is uncorrelated with information contained in z_{t-1} ; thus, taking the unconditional expectation of (24) and multiplying it by the vector z_{t-1} yields the following orthogonality condition:

$$E \left\{ \left[\pi_t^w - \psi_w \pi_{t-1}^w - \beta(1 + (1 - \beta) \psi_w) E_t \pi_{t+1}^w + \beta^2 \psi_w E_t \pi_{t+2}^w + k_w \gamma u_t \right] z_{t-1} \right\} = 0 \quad (25)$$

When addressing small samples, the nonlinear estimation using GMM might be sensitive to the manner in which the moment conditions are normalized. Because of the non-linearity of (25) in the structural parameters, the orthogonality condition is normalized with a procedure that minimizes the non-linearities. This procedure, called bounded normalization, consists of multiplying the condition (25) by a function of the parameters characterizing the wage Phillips curve, to ensure that the wage Phillips curve coefficients are bounded. As noted by Galí and Gertler (1999), simulation studies show that bounded normalization has better small-sample properties than alternative normalization procedures. Thus,

$$E \left\{ \chi_w \left[\pi_t^w - \psi_w \pi_{t-1}^w - \beta(1 + (1 - \beta) \psi_w) E_t \pi_{t+1}^w + \beta^2 \psi_w E_t \pi_{t+2}^w + k_w \gamma u_t \right] z_{t-1} \right\} = 0 \quad (26)$$

where $\chi_w = (1 - \alpha_w) - \varphi_w [1 - \beta(1 - \alpha_w)]$, while ψ_w and k_w are also functions of α_w and φ_w (see (22)).

Our estimation is made using quarterly U.S. data from the FRED database ranging from 1960:1 to 2011:4. Wage inflation is measured by *compensation per hour*, whereas we use the *civilian unemployment rate* for the unemployment rate. The set of instruments consists of the lags of the following observable variables: wage inflation, unemployment, price inflation, consumer price index, output gap,¹⁷ labor share (nonfarm business sector), and the spread between ten-year Treasury Bond and three-month Treasury Bill yields. In particular, six lags in price inflation, wage inflation and CPI, four lags for the output gap and two lags for the remaining instruments are used.¹⁸

As only a subset of parameters can be identified from moment condition (26), some coefficients must be imposed. We aim to estimate the parameters of the hazard (α_w and φ_w); thus, no restrictions are imposed on them. Following common practice, we set $\beta = 0.99$ equal to the inverse of the real interest rate. We calibrate the inverse Frisch elasticity (γ) equal to 2; as result, following Galí (2011), $\varepsilon_w = 8.85$ since $\varepsilon_w = [1 - \exp(-\gamma u^n)]^{-1}$, where the natural unemployment rate u^n is equal to 6%, i.e., the average rate of the period considered.

Table 1 presents the results for the structural form estimation. We show the estimation for the structural parameters φ_w (hazard slope) and α_w (hazard initial value); moreover, we also report D_w^e and α_w^e (computed as in (5)) and the $J - stat$.¹⁹

¹⁷The output gap is obtained by Hodrick-Prescott filtering of the Real GDP series.

¹⁸The sample range, the data and the wage Phillips curve specification used for GMM estimation are different from those that will be used in the Bayesian estimation because we sought to avoid the possibility that the Bayesian comparison might unduly favor our model with respect to the alternatives considered. However, we chose to perform a “non-informative” estimation for the hazard slope parameters to test the robustness of our comparison (see Section 4).

¹⁹The standard errors of D_w^e and α_w^e are computed using the delta method (see Papke and Wooldridge, 2005).

Table 1 – Wage Phillips curve estimation (structural form)²⁰

α_w	φ_w	D_w^e	α_w^e	$J - stat$
0.318*	0.126*	1.964*	0.444*	19.527
(0.050)	(0.030)	(0.146)	(0.033)	[0.813]

Notes: a 6-lag Newey-West estimate of the covariance matrix is used.

Standard errors are shown in parentheses.

For the $J - stat$ the p-value is shown in brackets.

* denotes statistical significance at 5% level.

All the estimated coefficients are statistically significant, and the hazard function is estimated to be upward sloping. Wages are estimated to be rigid, as an adjustment comes every two quarters. The $J - stat$ is a test of over-identifying moment conditions: In our case, we accept the null hypothesis that the over-identifying restrictions are satisfied (the model is “valid”).

We now report the reduced form of (24) that is obtained by substituting the estimated values of α_w and φ_w into (22):

$$\pi_t^w = \underset{(0.038)}{0.197\pi_{t-1}^w} + \underset{(0.000)}{0.991E_t\pi_{t+1}^w} - \underset{(0.037)}{0.193E_t\pi_{t+2}^w} - \underset{(0.006)}{0.03u_t} \quad (27)$$

All the coefficients are statistically significant at the 5% level (standard errors computed using the delta method are reported in parentheses).

In line with underlying theory, our wage Phillips curve can capture the well-known negative relation between unemployment and wage inflation, as shown by the negative coefficient measuring the slope of the curve. Therefore, our estimation indicates that a 1% fall of unemployment below its steady state involve a 3% increase of wage inflation.

In Figure 1, we provide a graphical representation for the hazard and survival functions derived from our estimation and computed using (1) and (4), respectively.²¹ The hazard clearly shows a positive slope, which means that a duration-dependent mechanism for wage adjustment emerges.

²⁰The estimation has been performed using Cliff’s (2003) GMM package for MATLAB, which is available at <https://sites.google.com/site/mcliffweb/programs>.

²¹The crosses joined by the dotted lines are the point estimates. The thick and thin bars represent the one-standard-deviation and two-standard deviation bands around the point estimate, respectively. The standard deviation is computed using the delta method.

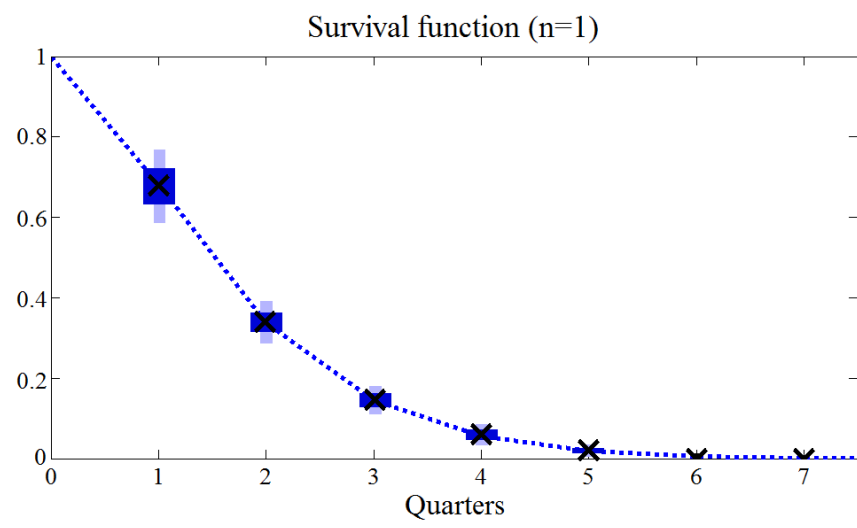
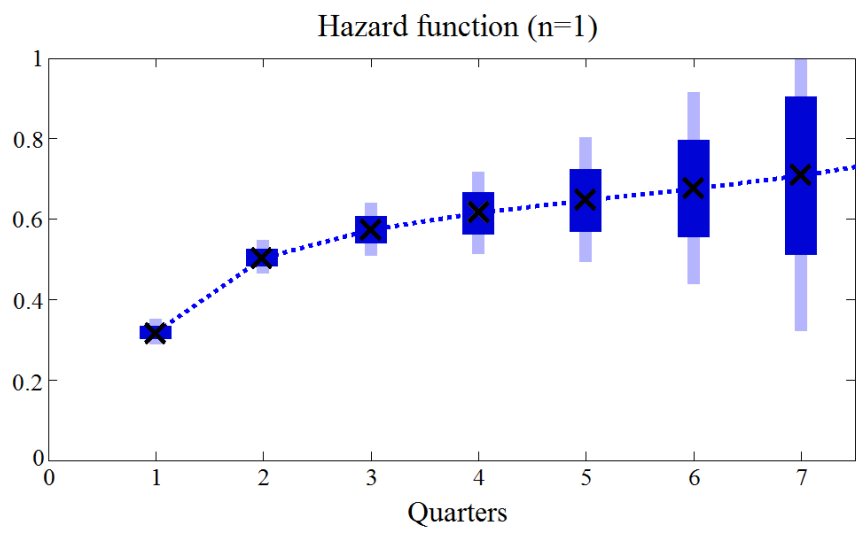


Figure 1 - Hazard and survival function deriving from GMM estimation.

3 The DSGE model

Our model generalizes EHL (2000) by assuming that price and wage adjustments are governed by a vintage-dependent mechanism. Moreover, to improve its empirical realism, we also consider habit formation in consumption, which implies persistence in the IS curve. Because the main difference with EHL (2000) is the derivation of the Phillips curves, as explained in Section 2, the description of the model derivation is not detailed. For a full derivation, we refer to EHL (2000). All the equations reported herein are expressed as log-linear deviations from the steady state.

3.1 The log-linearized economy

The demand side of the economy is described by a simple IS curve, which is obtained by log-linearizing the Euler equation around the steady state. As usual, the Euler equation is derived from maximizing the utility function (10), subject to the budget constraint (11). Formally,

$$y_t = \frac{1}{1+h} E_t y_{t+1} + \frac{h}{1+h} y_{t-1} - \frac{1-h}{\sigma(1+h)} (i_t - E_t \pi_{t+1}^p + E_t g_{t+1} - g_t), \quad (28)$$

where y_t is the output, π_t^p is the price inflation rate, i_t is the nominal interest rate set by the central bank, and g_t is a preference shock. The lagged term on output is attributable to the presence of internal consumption habits.

Price adjustment is described by a Phillips curve with a duration-dependent mechanism that is similar to that used for wages. As shown by Sheedy (2007), by using the same structure of (1) for prices, the price Phillips curve is:²²

$$\pi_t^p = \psi_p \pi_{t-1}^p + \beta [1 + (1-\beta)\psi_p] E_t \pi_{t+1}^p - \beta^2 \psi_p E_t \pi_{t+2}^p + k_p (mc_t + \zeta_t), \quad (29)$$

where mc_t is the real marginal cost and ζ_t is a price mark-up shock.²³ The coefficients ψ_p and k_p are a function of the parameters characterizing the hazard function for prices φ_p and α_p :

$$\begin{cases} \psi_p = \frac{\varphi_p}{(1-\alpha_p) - \varphi_p [1-\beta(1-\alpha_p)]} \\ k_p = \frac{(\alpha_p + \varphi_p) [1-\beta(1-\alpha_p) + \beta^2 \varphi_p]}{(1-\alpha_p) - \varphi_p [1-\beta(1-\alpha_p)]} \eta_{cx} \end{cases}$$

Parameter φ_p controls the gradient of the hazard function, and α_p is its starting level; the coefficient $\eta_{cx} = \frac{1-\phi}{1-\phi+\phi\varepsilon_p}$ is the elasticity of a firm's marginal cost with respect to average real marginal cost and depends on $1-\phi$ and the elasticity of substitution between goods (ε_p).

The log-linearized real marginal cost is:

$$mc_t = \omega_t + n_t - y_t, \quad (30)$$

where ω_t denotes the real wage and n_t is the number of hours worked. Equation (30) is derived

²²The curve is derived in a similar manner as the derivation explained in Section 2. For a detailed derivation, see Sheedy (2007).

²³For an interpretation of the source of this shock, see Smets and Wouters (2003) and Rabanal *et al.* (2005). See also Galí (2015: Appendix 5.2).

from cost minimization and is subject to the production function (6), which can be written in a log-linearized form as:

$$y_t = a_t + (1 - \phi) n_t, \quad (31)$$

where a_t is the technology shock.

The wage adjustment is described by the wage Phillips curve (21) previously derived, i.e.:

$$\pi_t^w = \psi_w \pi_{t-1}^w + \beta [1 + (1 - \beta) \psi_w] E_t \pi_{t+1}^w - \beta^2 \psi_w E_t \pi_{t+2}^w - k_w (\omega_t - mrs_t), \quad (32)$$

where the wage mark-up is expressed as the difference between the real wage (ω_t) and the marginal rate of substitution (mrs_t) because the labor market is characterized by imperfect competition.

The real wage, by definition, follows:

$$\omega_t = \pi_t^w - \pi_t^p + \omega_{t-1}. \quad (33)$$

The marginal rate of substitution between consumption and hours worked is obtained from the wage setter's problem and equals the ratio between the marginal utility of leisure and consumption. Formally, in log-linear terms, it is given by:

$$mrs_t = \frac{\sigma}{1 - h} (y_t - h y_{t-1}) + \gamma n_t - g_t, \quad (34)$$

Finally, monetary policy is assumed to follow a simple Taylor rule:

$$i_t = \rho_r i_{t-1} + (1 - \rho_r) (\delta_\pi \pi_t^p + \delta_x y_t) + v_t, \quad (35)$$

where ρ_r captures the degree of interest rate smoothing, δ_π and δ_x measure the response of the monetary authority to the deviation of inflation and output from their steady state values and v_t is a monetary policy shock.

Aside from the monetary disturbance,²⁴ all the shocks considered in the model follow an $AR(1)$ process:

$$\begin{cases} a_t = \rho_a a_{t-1} + \varepsilon_t^a, \\ g_t = \rho_g g_{t-1} + \varepsilon_t^g, \\ \zeta_t = \rho_\zeta \zeta_{t-1} + \varepsilon_t^\zeta, \\ v_t = \varepsilon_t^v, \end{cases} \quad (36)$$

where $\varepsilon_t^j \sim N(0, \sigma_j^2)$ are white noise shocks uncorrelated among them and ρ_j are the parameters measuring the degree of autocorrelation, for $j = \{a, g, \zeta\}$.

In summary, our model consists of eight equations, describing the following: the dynamic IS (28); the price Phillips curve (29); real marginal cost (30); the production function (31); the wage Phillips curve (32); real wage dynamics (33); the marginal rate of substitution (34); and the Taylor rule (35). The dynamics of the four shocks are described by (36).

²⁴Monetary policy persistence is already captured by the lagged term in (35). However, we have successfully checked the robustness of our results with regard to alternative assumptions. Following Woodford (2003), who shows that the optimal interest rate in New Keynesian models should have an $AR(2)$ interest rate smoothing, we have considered an $AR(2)$ process for the nominal rate in equation (35). Our results are available upon request. See also Coibion and Gorodnichenko (2012) for a wider discussion.

4 Empirical analysis

We estimate our model (28)-(36) for the U.S. economy using Bayesian techniques. Our choice is motivated by the fact that Bayesian methods outperform GMM and maximum likelihood in small samples.²⁵ After writing the model in state-space form, the likelihood function is evaluated using the Kalman filter, whereas prior distributions are used to introduce additional non-sample information into the parameters estimation: Once a prior distribution is elicited, the posterior density for the structural parameters can be obtained by reweighting the likelihood by a prior. The posterior is computed using numerical integration by employing the Metropolis-Hastings algorithm for Monte Carlo integration; for the sake of simplicity, all structural parameters are supposed to be independent of one another.

We use four observable macroeconomic variables: real GDP, price inflation, real wage, and nominal interest rate. The dynamics are driven by four orthogonal shocks, including monetary policy, productivity, preference and price mark-up. As the number of observable variables equals the number of exogenous shocks, the estimation does not present problems deriving from stochastic singularity.²⁶ The model estimation is performed using informative priors, and non-informative priors for those parameters characterizing the slope of the hazard function as a robustness check.

We aim to test whether the model exhibits a positive hazard function, i.e., whether the history-dependent price/wage adjustments hold. Following Benati (2008, 2009), we also test the robustness of our pricing mechanism to policy regime shifts. Considering only price rigidity and flexible wages, Benati (2009) analyses several models to build inflation persistence including Sheedy (2007)²⁷ and finds evidence of positive-sloping hazard functions; however, by considering the Great Moderation sub-sample, he also finds that the parameters encoding the hazard slope have dropped to zero over the most recent thirty years. He concludes that these parameters depend on the monetary regime, referring to the switch in the manner monetary policy is conducted, as discussed in Clarida *et al.* (2000). However, Benati (2009) focuses only on price inflation: We generalize his setup by considering staggered wages with a possible duration-dependent adjustment process in the labor markets. As discussed above, nominal rigidity and wage persistence may have important implications for both inflation persistence and monetary policy effects.

After estimating our model for the full sample (1960:1-2008:4), we also consider a smaller sample (1982:1-2008:4), representative of the Great Moderation, to investigate whether a positive hazard function still holds in a period characterized by low volatility in shocks and more aggressive tactics by central bankers in the fight against inflation.

Finally, we evaluate the empirical performance of our duration-dependent Phillips curves in relation to alternative specifications commonly used in the literature. We consider the traditional forward-looking Phillips curves derived in EHL (2000) extended by price and wage indexation, which is often a main assumption used to account for inflation persistence. Model comparison is based on log-marginal likelihood. To apply this methodology, we will show how the models

²⁵For an exhaustive analysis of Bayesian estimation methods, see Geweke (1999), An and Schorfheide (2007) and Fernández-Villaverde (2010).

²⁶The problems deriving from misspecification are widely discussed in Lubik and Schorfheide (2006) and Fernández-Villaverde (2010).

²⁷Specifically, Benati (2009) analyzed Fuhrer and Moore (1995), Galí and Gertler (1999), Blanchard and Galí (2007), Sheedy (2007), and Ascari and Ropele (2009).

compared here are nested.

The next subsection presents the data used and prior distributions. Subsection 4.2 provides the estimation for the baseline model. Subsection 4.3 evaluates our duration-dependent model against alternative specifications.

4.1 Data and prior distributions

We use U.S. quarterly data in our estimations, and all the time series used are from the FRED database maintained by the Federal Reserve Bank of St. Louis. We use *real gross domestic product* as the measure of the output, and the *effective Fed funds rate* is used for the nominal interest rate. Price inflation is measured using the *GDP implicit price deflator* taken in log-difference. The real wage is obtained dividing the nominal wage, measured by the *compensation per hour in nonfarm business sector*, by the *GDP implicit price deflator*. All the variables have been demeaned; moreover, output and real wage are detrended using Baxter and King’s bandpass filter.

Our choices regarding prior beliefs are as follows. The coefficients of the Taylor rule are centered on a prior mean of 1.5 for inflation and 0.125 for the output gap, which are Taylor’s (1999) estimates, and follow a Normal distribution. These values are standard in the literature. The smoothing parameter is assumed to follow a Beta distribution, with a mean of 0.6 and a standard deviation of 0.2. The same choice has been made for the consumption habit. We assume that the inverse of Frisch elasticity is based on a Gamma distribution, with a mean of 2 and a standard deviation 0.375. These priors are fairly diffuse and broadly consistent with those adopted in previous studies, including Del Negro *et al.* (2007), Smets and Wouters (2007), Justiniano and Primiceri (2008), and Justiniano *et al.* (2013).

For the hazard function coefficients, we perform an “informative estimation” using the estimated coefficients from a single equation GMM as priors;²⁸ we assign a Normal distribution to φ_p and φ_w with a standard deviation equal to 0.2, whereas α_p and α_w follow a Beta distribution with a standard deviation of 0.1. Following Benati (2009), we also perform a robustness check and estimate the model using non-informative priors for the parameters affecting the slope of the hazard function, instead of those derived from the GMM estimations. Unlike his approach, however, we employ a Uniform distribution with support $[-1, 1]$: We choose such a large interval because we want to investigate whether the hazard slope is positive, negative or zero.

We must calibrate some parameters to avoid identification problems.²⁹ Because we consider a production function without capital, it is difficult to estimate β and ϕ , which are set to 0.99 and 0.33, respectively. Similarly, we fix $\varepsilon_p = 6$ and $\varepsilon_w = 8.85$, implying a price and wage mark-up equal to 1.20 and 1.12, respectively. Following Sheedy (2007), price elasticity is calibrated to be coherent with the hazard priors derived from his GMM estimation. As explained in Section 2, wage elasticity is set equal to 8.85. Finally, all the autoregressive coefficients of the shocks follow

²⁸The values of φ_w and α_w estimated in Section 2 are used as priors. For the hazard characterizing price adjustment, we directly use as priors the GMM estimates of Sheedy (2007).

²⁹The identification procedure has been performed using the Identification toolbox for Dynare, which implements the identification condition developed by Iskrev (2010a, 2010b). For a review of identification issues arising in DSGE models, see Canova and Sala (2009).

a Beta distribution, with a mean of 0.5 and a standard deviation equal of 0.2. The prior for the shock standard deviations is an Inverse Gamma, with a mean of 0.01 and 2 degrees of freedom.

4.2 Estimation results

Our estimations are reported in Table 2, which also summarizes the 90% probability intervals and our beliefs about priors. The table describes the results for both the full sample and the Great Moderation. We report a posterior estimation of the shocks and structural parameters that are obtained by the Metropolis-Hastings algorithm, when informative priors for the hazard slope are used.

Table 2 – Prior and posterior distributions³⁰

	Prior distribution			Posterior distribution (full sample)			Posterior distribution (Great Moderation)		
	Density	Mean	St. Dev. ³¹	Mean	5%	95%	Mean	5%	95%
σ	Gamma	1.0	0.375	1.324	0.673	1.955	1.227	0.581	1.820
γ	Gamma	2.0	0.375	2.515	2.041	2.997	2.249	1.732	2.748
h	Beta	0.6	0.2	0.906	0.866	0.946	0.908	0.863	0.955
δ_π	Normal	1.5	0.25	1.423	1.197	1.650	1.851	1.524	2.158
δ_x	Normal	0.125	0.05	0.215	0.152	0.279	0.164	0.096	0.235
ρ_r	Beta	0.6	0.2	0.818	0.787	0.850	0.850	0.819	0.883
α_p	Beta	0.132	0.1	0.020	0.001	0.042	0.063	0.001	0.124
φ_p	Normal	0.222	0.2	0.195	0.157	0.233	0.128	0.048	0.213
α_w	Beta	0.318	0.1	0.126	0.073	0.179	0.151	0.073	0.228
φ_w	Normal	0.126	0.2	0.242	0.210	0.277	0.250	0.203	0.297
ρ_a	Beta	0.5	0.2	0.781	0.706	0.854	0.850	0.819	0.883
ρ_g	Beta	0.5	0.2	0.768	0.717	0.817	0.802	0.738	0.867
ρ_ζ	Beta	0.5	0.2	0.825	0.762	0.889	0.822	0.732	0.910
σ_a	Inv. Gamma	0.01	2	0.019	0.013	0.025	0.014	0.008	0.019
σ_g	Inv. Gamma	0.01	2	0.053	0.038	0.068	0.044	0.028	0.059
σ_v	Inv. Gamma	0.01	2	0.002	0.002	0.002	0.001	0.001	0.002
σ_ζ	Inv. Gamma	0.01	2	0.020	0.013	0.028	0.030	0.012	0.047

In the full sample case, the estimated hazard function is upward sloping because φ_p and φ_w are both positive. Thus, the duration-dependent mechanism appears able to account for inflation inertia for both prices and wages. The (unconditional) expected duration of price is 3.7 quarters, whereas wages appear to be less sticky, because their duration is 2.05 quarters.³² These durations

³⁰The posterior distributions are obtained using the Metropolis-Hastings algorithm; the procedure is implemented using the MATLAB-based Dynare package. The mean and posterior percentiles are from two chains of 250,000 draws each from the Metropolis-Hastings algorithm, for which we discarded the initial 30% of draws.

³¹For the Inverse Gamma distribution the degrees of freedom are indicated.

³²The durations (D_i^e) of price and wage stickiness are computed using the following relation: $D_i^e = \frac{1-\varphi_i}{\alpha_i+\varphi_i}$ for $i = \{p, w\}$ (see (5)).

are similar to those obtained in the literature, in which it is also common to find greater duration in price settings than wages (e.g., Rabanal and Rubio-Ramirez, 2005; Galí *et al.*, 2011).

The estimations for the parameters characterizing the utility function (i.e., habit, relative risk aversion and the inverse of Frisch elasticity) are coherent with the standard findings in the literature (see, e.g., Del Negro *et al.* 2007; Smets and Wouters, 2007; Justiniano and Primicieri, 2008; Justiniano *et al.*, 2013).

The response of the monetary authority to inflation and the output gap is consistent with the Taylor principle and the estimated coefficients of the monetary rule are in line with the literature. The estimated degree of interest rate smoothing is 0.82, and all the shocks exhibit a high degree of autocorrelation, i.e., approximately 0.8.

As expected, by considering the Great Moderation period, we find a more aggressive monetary policy stance (Clarida *et al.*, 2000). As opposed to Benati (2009), we find that the hazard functions still exhibit positive slopes in this sub-sample as well. This result offers evidence that a pricing mechanism based on a hazard function still holds also in a period characterized by a central bank more concerned with fighting inflation, as highlighted by the higher estimated coefficient for δ_π . As a result, intrinsic persistence also holds for the Great Moderation period.

The price duration increases to 4.5 quarters and is highlighted by the fact that the hazard function sloping remains positive, but smaller. This fact is consistent with macroeconomic theory: During the Great Moderation, inflation has dropped, and the cost of not adjusting a price is smaller compared to the previous period, which translates into a longer expected duration of prices.

By contrast, computed wage stickiness is rather stable, which reflects the fact that wage bargaining is more influenced by institutional factors related to the labor market than by monetary policy. Rabanal and Rubio-Ramirez (2005) also find the stability of wage duration over time.

In Figure 2, we plot the prior distribution, the posterior distribution and the posterior mode of the estimated parameters.

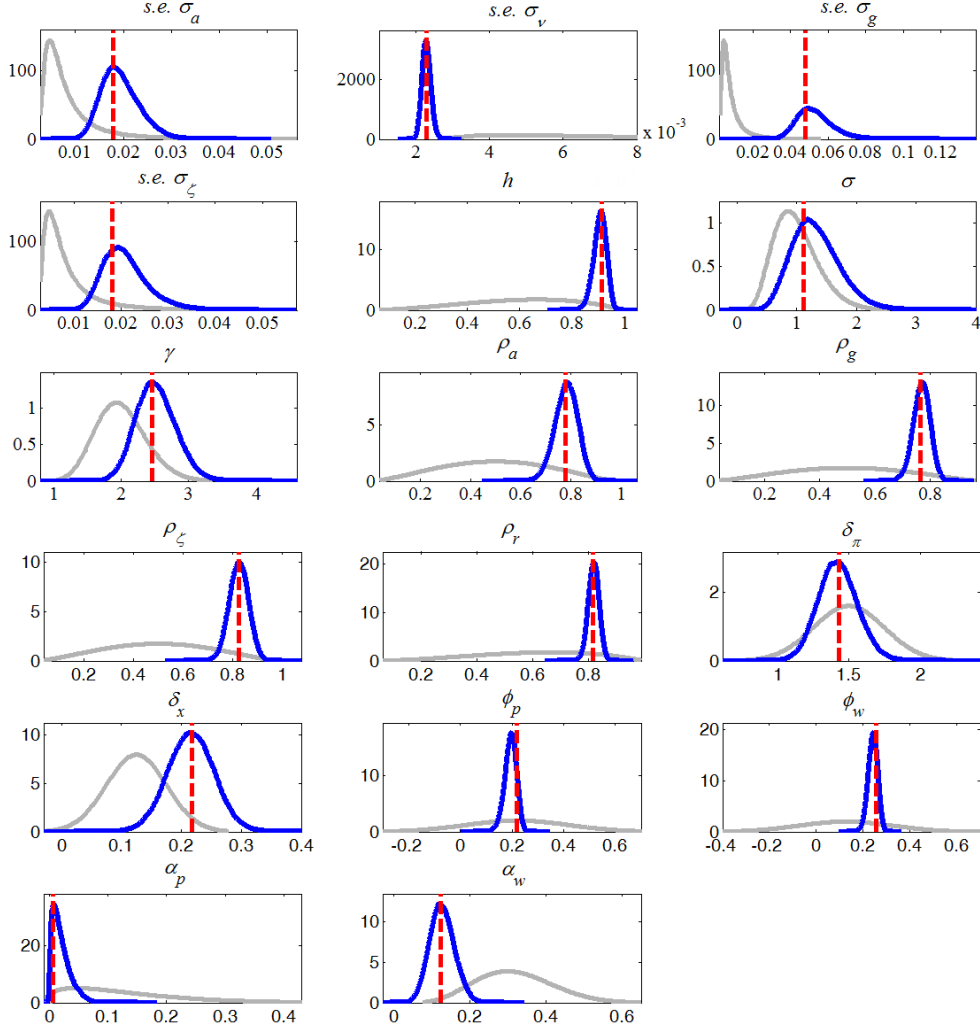


Figure 2 - Prior distribution (grey curve), Posterior distribution (blue curve) and Posterior mode (dotted line) of the estimated parameters.

Bayesian estimations of DSGE models can be quite sensitive to the choice of priors for model-specific parameters and other assumptions regarding, e.g., measures of the variables used and shock specifications. Thus, we have checked the robustness of our analysis by also considering uniform priors for the parameters φ_p and φ_w with support $[-1; 1]$,³³ whereas the prior distributions for the remaining parameters are the same as those used previously. The results are presented in Table 3.

³³Choosing this large support allows us to test whether the hazard slope is negative, positive or flat. The prior mean is centered on 0.

The “non-informative” estimation confirms our results regarding the hazard function, which remains characterized by a positive slope, both in the full sample and during the Great Moderation; the estimated parameters for the hazard slope are similar to those estimated under “informative” priors. This result shows that the hazard function mechanism is robust to a change of policy.

Table 3 - Prior and posterior distributions under non-informative priors

	Prior distribution			Posterior distribution (full sample)			Posterior distribution (Great Moderation)		
	Density	Mean	St. Dev.	Mean	5%	95%	Mean	5%	95%
σ	Gamma	1.0	0.375	1.321	0.670	1.933	1.227	0.595	1.834
γ	Gamma	2.0	0.375	2.511	2.021	2.974	2.251	1.738	2.753
h	Beta	0.6	0.2	0.906	0.868	0.948	0.909	0.865	0.957
δ_π	Normal	1.5	0.25	1.428	1.203	1.661	1.855	1.545	2.171
δ_x	Normal	0.125	0.05	0.215	0.151	0.277	0.165	0.096	0.234
ρ_r	Beta	0.6	0.2	0.818	0.787	0.850	0.851	0.820	0.882
α_p	Beta	0.132	0.1	0.020	0.001	0.041	0.067	0.001	0.133
φ_p	Uniform	0	0.57	0.195	0.158	0.236	0.125	0.042	0.213
α_w	Beta	0.318	0.1	0.126	0.073	0.177	0.151	0.072	0.225
φ_w	Uniform	0	0.57	0.243	0.209	0.276	0.252	0.207	0.298
ρ_a	Beta	0.5	0.2	0.780	0.704	0.854	0.832	0.755	0.910
ρ_g	Beta	0.5	0.2	0.768	0.719	0.817	0.800	0.738	0.866
ρ_ζ	Beta	0.5	0.2	0.824	0.760	0.887	0.824	0.737	0.914
σ_a	Inv. Gamma	0.01	2	0.019	0.013	0.025	0.014	0.008	0.019
σ_g	Inv. Gamma	0.01	2	0.052	0.038	0.067	0.044	0.029	0.061
σ_v	Inv. Gamma	0.01	2	0.002	0.002	0.002	0.001	0.001	0.002
σ_ζ	Inv. Gamma	0.01	2	0.020	0.013	0.028	0.029	0.013	0.044

We have also successfully checked the robustness of our results by considering different model specifications (i.e., a model without habit), various specifications for the Taylor rule (as previously discussed) and alternative series for observable variables.³⁴ Results are available upon request.

In the figures below, we plot the dynamic behavior of the model variables, described by the Bayesian impulse response functions, conditional on the price mark-up and monetary policy shocks.³⁵ The solid line represents the estimated response, with the shaded area capturing the corresponding 90% confidence interval. The horizontal axis measures the quarters after the initial shock. The monetary shock is illustrated in Figure 3, whereas the cost-push shock is described by Figure 4.

The monetary shock affects both real and nominal variables and has persistent effects on output. As expected, in response to the monetary tightening, GDP declines with a characteristic

³⁴In particular, we have considered a different measure for prices by using the *nonfarm business sector implicit price deflator*. With regard to wages, we considered alternative measures given by: *average hourly earnings of production; business sector compensation per hour*; and *hourly earnings for manufacturing sector*.

³⁵The figures show the Bayesian impulse response functions (IRFs) to a monetary and a cost-push shock. The sizes of the shocks are obtained from their estimated standard deviations (see Table 2).

hump-shaped pattern. It reaches a trough after four quarters and then slowly reverts back to its initial level. Both price inflation and the real wage also exhibit a hump-shaped behavior. Similarly, a positive shock to the price mark-up affects all the variables and has persistent effects on output. In both cases, the pattern of interest rate, real output, hours, real wage and inflation are consistent with the literature (see, e.g., Smets and Wouters, 2003; Christiano *et al.*, 2005; Güneş and Millard, 2012).

With regard to wage dynamic adjustments, which are the core of our investigation, wage inflation exhibits a pattern similar to that of price inflation. Figure 3 shows that, following a monetary shock, wage inflation has a moderate hump shape and progressively returns to its steady state with a little overshooting. In the case of a cost-push shock, which is shown in Figure 4, the dynamic response of wage inflation is strongly hump shaped with a peak after approximately ten quarters, which denotes a high level of wage persistence. We will discuss in more detail the response of nominal wage inflation in the next section, when we compare the responses of our model to alternatives in which inertia is introduced by wage indexation to past prices.

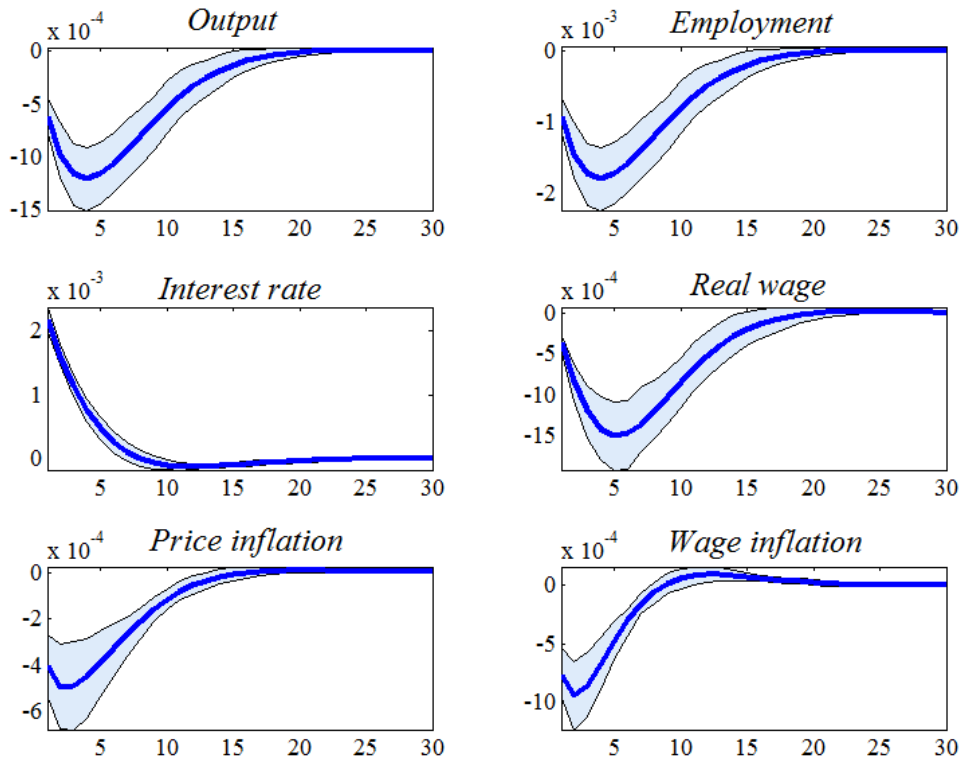


Figure 3 - Bayesian IRFs conditional to a monetary shock.

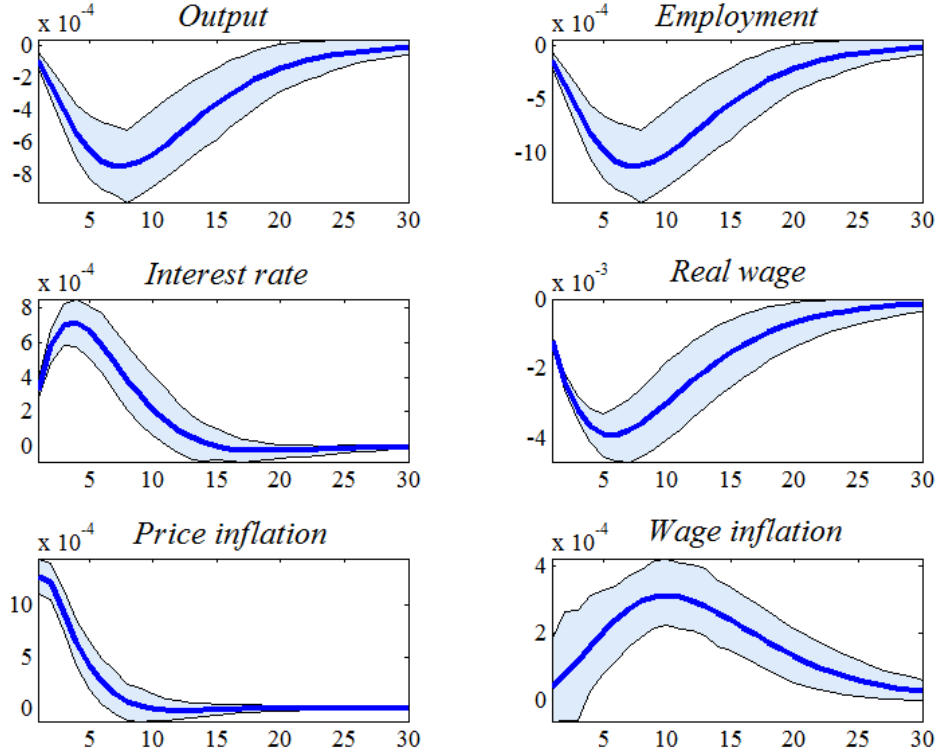


Figure 4 - Bayesian IRFs conditional to a price mark-up shock.

4.3 Duration-dependent Phillips curves vs. alternatives

In this section, we aim to compare the empirical performance of our duration-dependent Phillips curves to different specifications accounting for price and wage inflation inertia. Specifically, we focus on EHL (2000), augmenting it by indexation because, as discussed above, this is a common method of introducing price and wage persistence in New Keynesian DSGE models. We consider two different forms of indexation that are widely used, and we compare these alternatives to our baseline model in terms of log-marginal density.

4.3.1 Alternative price-setting mechanisms

Simply by setting $\varphi_p = 0$ and $\varphi_w = 0$ in (29) and (32), we obtain flat hazard functions and, as a result, price and wage Phillips curves *à la* Calvo as in EHL (2000), which is nested in our model. Formally, these two purely forward-looking curves are

$$\pi_t^p = \beta E_t \pi_{t+1}^p + \lambda_p (mc_t + \zeta_t) \quad (37)$$

$$\pi_t^w = \beta E_t \pi_{t+1}^w - \lambda_w \mu_t^w \quad (38)$$

where $\lambda_p = \frac{\alpha_p[1-\beta(1-\alpha_p)]}{1-\alpha_p}\eta_{cx}$ and $\lambda_w = \frac{\alpha_w[1-\beta(1-\alpha_w)]}{1-\alpha_w}\Xi_w$.

Equations (37) and (38) clearly do not exhibit any persistence. However, price and wage inertia can be easily introduced by considering indexation to past price inflation. Two popular ways to undertake such indexation have been proposed by Galí and Gertler (1999) and Christiano *et al.* (2005). The former obtains a lagged term in the aggregate Phillips curve by assuming the existence of a fraction of agents who set prices/wages according to a backward-looking rule of thumb. Alternatively, the latter considers dynamic indexation, i.e., all agents not able to reset their price/wage adjust according to past price inflation, and the past inflation rate consequently appears in the Phillips curve.

According to Galí and Gertler (1999), the Phillips curves can be rewritten as follows:

$$\pi_t^p = \frac{\xi_p}{\Lambda_p} \pi_{t-1}^p + \frac{\beta(1-\alpha_p)}{\Lambda_p} E_t \pi_{t+1}^p + \lambda_p^\xi (mc_t + \zeta_t) \quad (39)$$

$$\pi_t^w = \frac{\xi_w}{\Lambda_w} \pi_{t-1}^w + \frac{\beta(1-\alpha_w)}{\Lambda_w} E_t \pi_{t+1}^w - \lambda_w^\xi \mu_t^w \quad (40)$$

where ξ_p (ξ_w) measures the fraction of the rule-of-thumb agents, i.e., the degree of price (wage) indexation to past price inflation; $\Lambda_p = 1 - \alpha_p + \xi_p [\alpha_p + (1 - \alpha_p)\beta]$, $\Lambda_w = 1 - \alpha_w + \xi_w [\alpha_w + (1 - \alpha_w)\beta]$, $\lambda_p^\xi = \frac{(1-\xi_p)(1-\alpha_p)\lambda_p}{\Lambda_p}$, and $\lambda_w^\xi = \frac{(1-\xi_w)(1-\alpha_w)\lambda_w}{\Lambda_w}$.

Instead, assuming dynamic indexation as in Christiano *et al.* (2005), equations (37) and (38) can be written as:

$$\pi_t^p = \frac{\iota_p}{(1+\iota_p\beta)} \pi_{t-1}^p + \frac{\beta}{1+\iota_p\beta} E_t \pi_{t+1}^p + \lambda_p^\iota (mc_t + \zeta_t) \quad (41)$$

$$\pi_t^w = \iota_w \pi_{t-1}^w - \iota_w \beta \pi_t^w + \beta E_t \pi_{t+1}^w - \lambda_w \mu_t^w \quad (42)$$

where ι_p (ι_w) denotes the degree of price (wage) indexation to last period's price inflation and $\lambda_p^\iota = \frac{\lambda_p}{(1+\iota_p\beta)}$.

4.3.2 Model comparison

Our formalization nests different models of price and wage adjustment. The differences depend only on the Phillips curve parameterization. With different assumptions regarding φ_p , φ_w , ι_p , ι_w , ξ_p , ξ_w , we can consider positive or flat hazard functions augmented by the two different types of aforementioned indexation. We compare our baseline (BASE) to three alternative scenarios:³⁶

1. An EHL model with dynamic indexation (DYNind), by considering (41) and (42);
2. An EHL model with rule of thumb indexation *à la* Galí-Gertler (GG), by considering (39) and (40).
3. An EHL model with inertia only in the wage equation (MIXED), i.e., Phillips curves are (37) and (32).

³⁶We omit the comparison with a model characterized by simple forward-looking Phillips curves *à la* Calvo because this model does not have intrinsic persistence. However, Rabanal and Rubio-Ramirez (2005) showed that this model performs the same as a model with indexation.

The measure used to compare the models is the log-marginal likelihood, which is a measure of the fit of a model in explaining the data.³⁷ The aim is to evaluate whether the manner in which the price and wage adjustments are modeled affects the fit of a model. The model with the highest log-marginal likelihood better explains the data.³⁸ Table 4 reports our results.

Table 4 - Log-marginal data densities and Bayes factors for different models³⁹

Model	Log-marginal data density	Bayes factor vs. BASE
BASE	3615.6	
BASE (non-info)	3613.6	exp [-2.0]
MIXED	3598.1	exp [-17.5]
DYNind	3569.7	exp [-45.9]
GG	3564.8	exp [-50.8]

In terms of marginal likelihood, the difference between the Galí-Gertler specification and dynamic indexation is minimal. According to Jeffreys' scale of evidence,⁴⁰ this difference must be considered as "slight" evidence in favor of DYNind with respect to GG. However, our model clearly outperforms both the alternatives considered: In particular, the Bayes factor provides "very strong" evidence in favor of our specification. As the models considered differ only for the Phillips curves, this result indicates that the data prefer a pricing method that is based on positive hazard functions, which is not surprising as the micro evidence rejects both constant hazard and indexation. Under "non-informative" priors, we observe a slight decrease in the marginal likelihood due to an increase in model complexity under diffuse priors which penalizes the marginal data density (this effect dominates the improvement in model fit).

In comparing our duration-dependent adjustment model and those models with dynamic indexation or rule-of-thumb agents, the fit of the different models is judged by the log-marginal likelihoods. The differences are significant (as evidenced by the Bayes factors). However, while useful, the marginal likelihood is not very informative about the source of these gains. Thus, in order to inspect the nature behind the best fit of our setup we analyse its dynamic behavior.

The higher posterior odds of our specification comes from differences in the wage equations. Notably, all the models differ only in the wage and price Phillips curves. All price equations are similar (they include backwards- and forwards-looking terms for the main dependent variables) and, after estimation, lead to identical dynamics. By contrast, there are substantial differences between the wage dynamics of our model and those generated by models with indexation. These differences are illustrated in Figure 5 that depicts the IRFs of the wage inflation in the three models considered,⁴¹ conditional on 1% shocks of monetary policy (upper panel) and price mark-up (lower panel).

³⁷The estimated reduced forms of the Phillips curves considered are reported in Appendix B.

³⁸For details of model comparison technique, see Fernández-Villaverde and Rubio-Ramirez (2004), Rabanal and Rubio-Ramirez (2005), Lubik and Schorfheide (2006), and Riggi and Tancioni (2010).

³⁹We used the modified harmonic mean estimator based on Geweke (1999) for the computation of the marginal likelihood for different model specifications. The Bayes factor is the ratio of posterior odds to prior odds (see Kass and Raftery, 1995).

⁴⁰Jeffreys (1961) developed a scale to evaluate the Bayes factor indication. Odds ranging from 1:1 to 3:1 give "very slight evidence", odds from 3:1 to 10:1 are "slight evidence", odds from 10:1 to 100:1 give "strong to very strong evidence", and odds greater than 100:1 are "decisive evidence."

⁴¹The dynamics of all the other variables do not exhibit qualitative differences, so we do not report them. Clearly, these dynamics are similar to those reported in Figures 3 and 4.

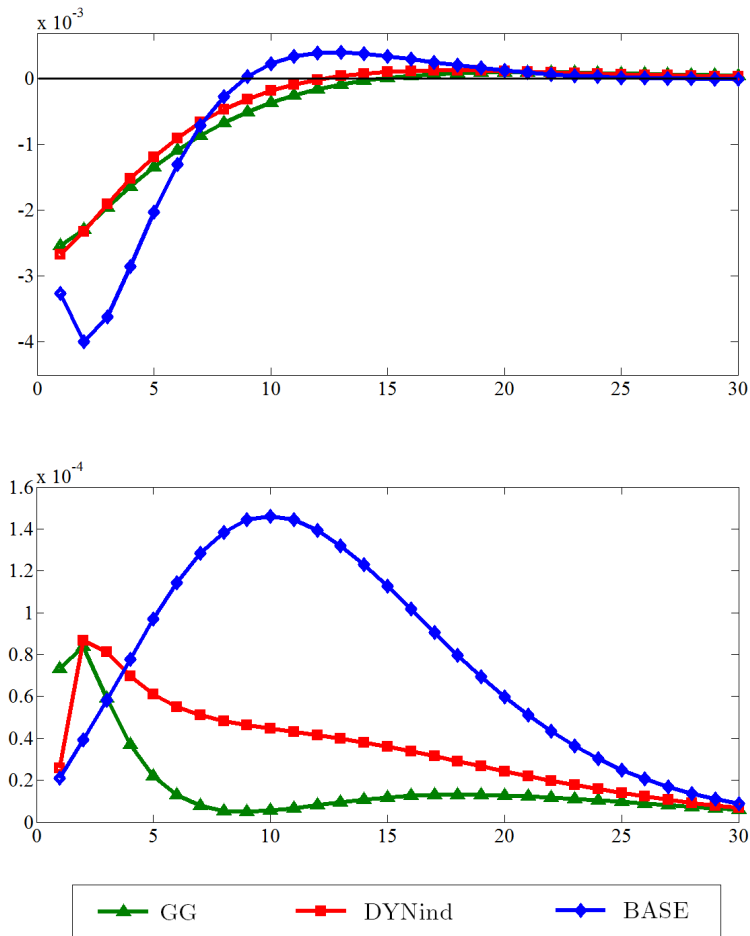


Figure 5 - IRFs of the wage inflation to 1% monetary policy (upper panel) and price mark-up (lower panel) shocks for several model specifications.

Considering monetary shock first, in both cases of indexation to past price, wage inflation follows a smooth path, whereas in our framework, beyond exhibiting a small hump immediately after the shock, it is more volatile. Under a cost-push shock, for both models with indexation, wage inflation responds with a spike in the early periods and then slowly returns to its stationary value. A model with a vintage-dependent wage adjustment instead generates a dynamic quite different than that associated with indexation models.

The higher predictive ability associated with our model by the Bayesian comparison emerges from the different dynamic behavior of nominal wages generated by our wage Phillips curve because the reduced forms of the three wage Phillips curves considered herein are different. Our mechanism embeds the wage Phillips curve of a lagged term for wage inflation (intrinsic inertia). In this manner, it differs from wage equations with indexation, for which the backward term is on past price inflation (inherited inertia). This is in line with the micro evidence on wages which shows that the indexation scheme is a poor expedient to explain the wage inflation persistence, in particular

since the Great Moderation. As the marginal likelihood is a measure of the ability of a model to fit the data, we can conclude that our wage-setting mechanism better captures the information contained in the data than models with some form of indexation.

In line with the above observations, the crucial role of wage adjustment is also highlighted by the fact that the log-marginal likelihood of MIXED is significantly higher than that of indexation models (see Table 4). It is worth recalling that MIXED refers to the model with duration-dependent adjustment only in wages, while prices follow the common Calvo scheme with only forward-looking term. The importance of this result is twofold: First, it confirms the key role of nominal wage rigidities to explain the data—as claimed by, e.g., Christiano *et al.* (2005), Rabanal and Rubio-Ramirez (2005), Olivei and Tenreyro (2010). Second, it shows that introducing wage persistence in a way consistent with the micro evidence can lead to a better fit of the model. Such a kind of model, with inertia only on the wage side generated by positive hazard rates, exhibits better predictive ability than models with indexation in both prices and wages. This improvement is consistent with the fact that micro evidence rejects the mechanism of indexation to explain the inflation persistence.

5 Conclusions

Our paper proposed a new approach to model wage inflation persistence and evaluated its empirical relevance. In line with micro evidence on wages, we assumed that wage adjustments are governed by a vintage-dependent mechanism, and showed how to derive a New Keynesian wage Phillips curve that also embeds backward terms for past wage inflation (intrinsic persistence). In our specification, the presence of endogenous-lagged terms does not rely on the unrealistic assumption of indexation to past price inflation rates but has a theoretical reason justified by the presence of a positive selection effect. Due to wage stickiness, wage setters continue to adjust wages after the shock has vanished. In the standard Calvo model, upward and downward adjustments compensate one another (no selection effect). In our case, due to a positive hazard function, a wage setter is more likely to adjust wages in the same direction of the past shock, inducing wage persistence.

We presented evidence about the relevance of intrinsic wage persistence from both single equation and full model estimation. Lagged terms for wage inflation are significantly different from zero in single equation GMM estimation. Placing our equation in a small-scale DSGE model, we confirmed using Bayesian estimation that an upward-sloping hazard function emerges for both prices and wages. By comparing log-marginal likelihoods, we found that our model outperforms alternatives based on popular mechanisms for modeling inflation persistence. We showed that the key rationale of this result is found in our wage adjustment mechanism. Our result confirms the crucial role of nominal wage rigidities to understanding economic fluctuations (see, e.g., Christiano *et al.*, 2005). Finally, we determine that the hazard function slope does not change with the policy regime.

It would be interesting to observe other features of the different modeling choices for price setting, such as the implications for welfare and optimal monetary policy. However, this subject is beyond the scope of the current paper, and we leave it for future research.

Appendix A – Hazard function properties

This appendix provides further details regarding hazard function properties. We refer to Sheedy (2007) for the proofs relative to the hazard function discussed herein, in particular his Appendices A.2 and A.5. Moreover, we show the evolution of the hazard and the resulting wage Phillips curve for the general case in which $n > 1$ parameters control the hazard gradient.

Assuming that $\Gamma_t \subset \Theta$ denotes the set of households that post a new wage at time t , the length of wage stickiness can be defined as:

$$D_t(j) \equiv \min \{l \geq 0 \mid j \in \Gamma_{t-l}\} \quad (43)$$

where $D_t(j)$ is the duration of wage stickiness for household j for which the last reset was l periods ago.

As explained in the text, the hazard function is defined by a sequence of probabilities: $\{\alpha_{w,l}\}_{l=1}^{\infty}$, where $\alpha_{w,l}$ represents the probability of resetting a wage that had remained unchanged for l periods. This probability is defined as: $\alpha_{w,l} \equiv \Pr(\Gamma_t \mid D_{t-1} = l - 1)$.

Each hazard function is related to a survival function, which expresses the probability that a wage remains fixed for l periods. The survival function for the hazard is defined by a sequence of probabilities: $\{\varsigma_{w,l}\}_{l=0}^{\infty}$, where $\varsigma_{w,l}$ denotes the probability that a wage fixed at time t will still be in use at time $t + l$.

The hazard function can be reparameterized by making use of a set of $n + 1$ parameters and rewritten as (1) if $n = 1$ and in the following way for the general case $n > 1$:

$$\alpha_{w,l} = \alpha_w + \sum_{j=1}^{\min(l-1,n)} \varphi_{w,j} \left[\prod_{k=l-j}^{l-1} (1 - \alpha_{w,k}) \right]^{-1}, \quad (44)$$

where α_w is the initial value of the hazard function and $\varphi_{w,j}$ is its slope; n is the number of parameters that control the slope. The sequence of parameters $\{\varphi_{w,l}\}_{l=1}^n$ affect the gradient of the hazard function in the following way:

$$\begin{cases} \varphi_{w,l} = 0, \forall l = 1, \dots, n & \longrightarrow \text{the hazard is flat (Calvo case);} \\ \varphi_{w,l} \geq 0, \forall l = 1, \dots, n & \longrightarrow \text{the hazard is upward-sloping;} \\ \varphi_{w,l} \leq 0, \forall l = 1, \dots, n & \longrightarrow \text{the hazard is downward-sloping.} \end{cases} \quad (45)$$

The survival function (4) in the general case is rewritten as:

$$\varsigma_{w,l} = (1 - \alpha_w)\varsigma_{w,l-1} - \sum_{h=1}^{\min(l-1,n)} \varphi_{w,h}\varsigma_{w,l-1-h} \quad (46)$$

Following Sheedy (2007), we assume that the hazard function satisfies two weak restrictions:

$$\begin{cases} \alpha_{w,1} < 1, \text{ meaning that there is a degree of wage stickiness;} \\ \alpha_{w,\infty} > 0, \text{ with } \alpha_{w,\infty} = \lim_{l \rightarrow \infty} \alpha_{w,l}. \end{cases} \quad (47)$$

We now introduce $\theta_{w,lt} \equiv \Pr(D_t = l)$ which denotes the proportion of households earning at time t a wage posted at period $t - l$. The sequence $\{\theta_{w,lt}\}_{l=0}^{\infty}$ indicates the distribution of the duration of wage stickiness at time t . This distribution evolves over time based on the following:

$$\begin{cases} \theta_{w,0t} = \sum_{l=1}^{\infty} \alpha_{w,l} \theta_{w,l-1,t-1} \\ \theta_{w,lt} = (1 - \alpha_{w,l}) \theta_{w,l-1,t-1} \end{cases} \quad (48)$$

If the hazard function satisfies the restrictions (47) and the evolution over the time of the distribution of wage length evolves as in (48), then *a*) from whatever starting point, the economy always converges to a unique stationary distribution $\{\theta_{w,l}\}_{l=0}^{\infty}$; hence $\theta_{w,lt} = \theta_{w,l} = \Pr(D_t = l)$, $\forall t$ and *b*) let us consider (1) and assume that the economy has converged to $\{\theta_{w,l}\}_{l=0}^{\infty}$; the relations expressed in (5) are obtained. For $n > 1$, the conditions in (5) become:

$$\begin{cases} \theta_{w,l} = \left(\alpha_w + \sum_{h=1}^n \varphi_{w,h} \right) \varsigma_{w,l} \\ \alpha_w^e = \alpha_w + \sum_{l=1}^n \varphi_{w,l} \\ D_w^e = \frac{1 - \sum_{l=1}^n l \varphi_{w,l}}{\alpha_w + \sum_{l=1}^n \varphi_{w,l}} \end{cases} \quad (49)$$

Then, inserting (46) in (16), we obtain:

$$w_t^* = \beta(1 - \alpha_w) E_t w_{t+1}^* - \sum_{l=1}^n \beta^{l+1} \varphi_{w,l} E_t w_{t+l+1}^* + \left[1 - \beta(1 - \alpha_w) + \sum_{l=1}^n \beta^{l+1} \varphi_{w,l} \right] (w_t - \Xi_w \mu_t^w) \quad (50)$$

By making use of (49), equation (17) can be recast as follows:

$$w_t = (1 - \alpha_w) w_{t-1} - \sum_{l=1}^n \varphi_{w,l} w_{t-1-l} + \left(\alpha_w + \sum_{h=1}^n \varphi_{w,h} \right) w_t^* \quad (51)$$

Finally, the wage Phillips curve in the case $n > 1$ is obtained by mixing (50) and (51):

$$\pi_t^w = \sum_{l=1}^n \psi_{w,l} \pi_{t-l}^w + \sum_{l=1}^{n+1} \delta_{w,l} E_t \pi_{t+l}^w - k_w \mu_t^w \quad (52)$$

where the coefficients $\psi_{w,l}$, $\delta_{w,l}$ and k_w have the following parameterization:

$$\begin{aligned}\psi_{w,l} &= \frac{\varphi_{w,l} + \sum_{h=l+1}^n \varphi_{w,h} \left[1 - \beta(1 - \alpha_w) + \sum_{k=1}^{h-1} \beta^{k+1} \varphi_{w,k} \right]}{\chi_w} && \text{for } l = 1, \dots, n \\ \delta_{w,1} &= \frac{\beta \left[(1 - \alpha_w) - \sum_{h=1}^n \beta^h \varphi_{w,h} \left(\alpha_w + \sum_{k=1}^{h-1} \varphi_{w,k} \right) \right]}{\chi_w} \\ \delta_{w,l+1} &= - \frac{\beta^{l+1} \left[\varphi_{w,l} + \sum_{h=l+1}^n \beta^{h-1} \varphi_{w,h} \left(\alpha_w + \sum_{k=1}^{h-1} \varphi_{w,k} \right) \right]}{\chi_w} && \text{for } l = 1, \dots, n \\ k_w &= \frac{\Xi_w \left[\left(\alpha_w + \sum_{h=1}^n \varphi_{w,h} \right) \left[1 - \beta(1 - \alpha_w) + \sum_{h=1}^n \beta^{h+1} \varphi_{w,h} \right] \right]}{\chi_w}\end{aligned}$$

where $\chi_w = (1 - \alpha_w) - \sum_{h=1}^n \varphi_{w,h} \left[1 - \beta(1 - \alpha_w) + \sum_{k=1}^{h-1} \beta^{k+1} \varphi_{w,k} \right]$.

It is easy to check that if we assume that only one parameter controls the slope of the hazard function (i.e., $n = 1$), the wage Phillips curve (52) becomes that reported in the paper, i.e., (21).

Appendix B – Reduced form Phillips curves from the Bayesian estimation

The reduced form from the Bayesian estimation for the price and wage Phillips curves in the EHL model with dynamic indexation (DYNind) are:

$$\pi_t^p = 0.138\pi_{t-1}^p + 0.853E_t\pi_{t+1}^p + 0.01mc_t \quad (53)$$

$$\pi_t^w = 0.153\pi_{t-1}^p - 0.151\pi_t^p + 0.99E_t\pi_{t+1}^w - 0.01\mu_t^w \quad (54)$$

The EHL model with rule-of-thumb indexation *à la* Galí-Gertler (GG) implies:

$$\pi_t^p = 0.144\pi_{t-1}^p + 0.847E_t\pi_{t+1}^p + 0.009mc_t \quad (55)$$

$$\pi_t^w = 0.106\pi_{t-1}^p + 0.884E_t\pi_{t+1}^w - 0.01\mu_t^w \quad (56)$$

Finally, our duration-dependent specification is associated with:

$$\pi_t^p = 0.2\pi_{t-1}^p + 0.99E_t\pi_{t+1}^p - 0.196E_t\pi_{t+2}^p + 0.012mc_t \quad (57)$$

$$\pi_t^w = 0.288\pi_{t-1}^w + 0.99E_t\pi_{t+1}^w - 0.283E_t\pi_{t+2}^w - 0.007\mu_t^w \quad (58)$$

As shown in the paper, the above reduced forms imply that models without duration-dependent adjustment capture persistence in the wage equation by past price inflation (i.e., inherited inertia), whereas (57)-(58) include a backward term for wage inflation. Both for price and wage inflation equations, our estimated Phillips curves exhibit a higher degree of inertia, as highlighted by the coefficients attached to backward inflation, with respect to models based on indexation. This is not surprising as, since the Great Moderation, indexation to past inflation has progressively vanished, and thus the parameter encoding it has progressively dropped to zero (see Benati, 2008).

References

- Acocella, N., G. Di Bartolomeo and P. Tirelli (2015), "U.S. Trend Inflation Reinterpreted: The Role of Fiscal Policies and Time-Varying Nominal Rigidities," *Macroeconomic Dynamics*, 19(6): 1294-1308.
- Adjemian, S., H. Bastani, M. Juillard, F. Karamé, F. Mihoubi, G. Perendia, J. Pfeifer, M. Ratto and S. Villemot (2011), "Dynare: Reference Manual, Version 4," *Dynare Working Papers*, 1, CEPREMAP.
- An, S. and F. Schorfheide (2007), "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26(2-4): 113-172.
- Angeloni, I., L. Aucremanne, M. Ehrmann, J. Galí, A. Levin, and F. Smets (2006), "New Evidence on Inflation Persistence and Price Stickiness in the Euro Area: Implications for Macro Modelling," *Journal of the European Economic Association*, 4(2-3): 562-574.
- Ascari, G. and T. Ropele (2009), "Trend Inflation, Taylor Principle, and Indeterminacy," *Journal of Money, Credit and Banking*, 41(8): 1557-1584.
- Barattieri, A., S. Basu and P. Gottschalk (2014), "Some Evidence on the Importance of Sticky Wages," *American Economic Journal: Macroeconomics*, 6(1): 70-101.
- Benati, L. (2008), "Investigating Inflation Persistence Across Monetary Regimes," *Quarterly Journal of Economics*, 123(3): 1005-1060.
- Benati, L. (2009), "Are 'Intrinsic Inflation Persistence' Models Structural in the Sense of Lucas (1976)?," *ECB Working Paper Series*, No. 1038.
- Bils, M. and P.J. Klenow (2004), "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112(5): 947-985.
- Blanchard, O.J. and J. Galí (2007), "Real Wage Rigidities and the New Keynesian Model," *Journal of Money, Credit and Banking*, 39(1): 35-65.
- Calvo, G.A. (1983), "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12(3): 383-398.
- Canova, F. and L. Sala (2009), "Back to Square One: Identification Issues in DSGE Models," *Journal of Monetary Economics*, 56(4): 431-449.
- Cecchetti, S.G. (1986), "The Frequency of Price Adjustment: A Study of the Newsstand Prices of Magazines," *Journal of Econometrics*, 31(3), 255-274.
- Christiano, L.J., M. Eichenbaum and C.L. Evans (2005), "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1): 1-45.
- Clarida, R., J. Galí and M. Gertler (2000), "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115(1): 147-180.
- Coibion, O. and Y. Gorodnichenko (2012), "Why Are Target Interest Rate Changes So Persistent?," *American Economic Journal: Macroeconomics*, 4(4): 126-162.
- Del Negro, M., F. Schorfheide, F. Smets, and R. Wouters (2007), "On the Fit and Forecasting Performance of New Keynesian Models," *Journal of Business and Economic Statistics*, 25(2), 123-162.

- Dhyne, E., L.J. Álvarez, H. Le Bihan, G. Veronese, D. Dias, J. Hoffmann, N. Jonker, P. Lünne-
mann, F. Rumler and J. Vilmunen (2005), “Price Setting in the Euro Area: Some Stylized
Facts from Individual Consumer Price Data,” *ECB Working Paper Series*, No. 524.
- Dotsey, M., R.G. King and A.L. Wolman (1999), “State Dependent Pricing and the General
Equilibrium Dynamics of Money and Output,” *Quarterly Journal of Economics*, 114(2): 655-
690.
- Erceg, C.J., Henderson, D.W. and A.T. Levin (2000), “Optimal Monetary Policy with Staggered
Wage and Price Contracts,” *Journal of Monetary Economics*, 46(2): 281-313.
- Fabiani, S., M. Druant, I. Hernando, C. Kwapil, B. Landau, C. Loupias, F. Martins, T.Y. Matha,
R. Sabbatini, H. Stahl, Ad C.J. Stokman, (2005), “The Pricing Behaviour of Firms in the
Euro Area: New Survey Evidence,” *ECB Working Paper Series*, No. 535.
- Fernández-Villaverde, J. and J.F. Rubio-Ramirez (2004), “Comparing Dynamic Equilibrium
Economies to Data: a Bayesian Approach,” *Journal of Econometrics*, 123(1): 153-187.
- Fernández-Villaverde, J. (2010), “The Econometrics of DSGE Models,” *SERIEs Spanish Economic
Association*, 1(1): 3-49.
- Fuhrer, J. and G. Moore (1995), “Inflation Persistence,” *Quarterly Journal of Economics*, 109(2):
127-159.
- Fuhrer, J. (2011), “Inflation Persistence,” *Handbook of Monetary Economics Volume 3A*, B.J.
Friedman and M. Woodford eds. Elsevier Science, North-Holland: 423-483.
- Galí, J. (2011), “The Return of Wage Phillips Curve,” *Journal of the European Economic Asso-
ciation*, 9(3): 436-461.
- Galí, J. (2015), “Monetary Policy, Inflation and the Business Cycle: An Introduction to the New
Keynesian Framework,” 2nd Edition, *Princeton University Press*.
- Galí, J. and M. Gertler (1999), “Inflation Dynamics: a Structural Econometric Analysis,” *Journal
of Monetary Economics*, 44(2): 195-222.
- Galí, J., M. Gertler and D. Lopez-Salido (2005), “Robustness of the estimates of the hybrid New
Keynesian Phillips curve,” *Journal of Monetary Economics*, 52(6): 1107-1118.
- Galí, J., F. Smets and R. Wouters (2011), “Unemployment in an estimated New Keynesian model,”
NBER Macroeconomics Annual 2011: 329-360.
- Geweke, J. (1999), “Using Simulation Methods for Bayesian Econometric Models: Inference, De-
velopment and Communication,” *Econometric Reviews*, 18(1): 1-73.
- Goodfriend, M. and R.G. King (1997), “The New Neoclassical synthesis and the Role of Monetary
Policy,” *NBER Macroeconomics Annual 1997*: 231-283.
- Guerrieri, L. (2001), “Inflation Dynamics,” *International Finance Discussion Paper*, No. 715.
- Guerrieri, L. (2006), “The Inflation Persistence of Staggered Contracts,” *Journal of Money, Credit
and Banking*, 38(2), 483-494.
- Günes, K. and S. Millard (2012), “Using Estimated Models to Assess Nominal and Real Rigidities
in the United Kingdom,” *International Journal of Central Banking*, 8(4): 97-119.
- Hofmann, B., G. Peersman and R. Straub, (2012), “Time variation in U.S. wage dynamics,”
Journal of Monetary Economics, 59(8): 769-783.

- Holland, S. (1988), "The Changing Responsiveness on Wages to Price-Level Shocks: Explicit and Implicit Indexation", *Economic Inquiry* 26: 265-279.
- Iskrev, N. (2010a), "Local Identification in DSGE Models," *Journal of Monetary Economics*, 57(2): 189-202.
- Iskrev, N. (2010b), "Evaluating the strength of identification in DSGE models. An a priori approach," *Working Papers w201032, Banco de Portugal*, 1-70.
- Jeffreys, H. (1961), "Theory of Probability," *Oxford University Press*, Oxford.
- Justiniano, A. and G. Primiceri (2008), "The Time-Varying Volatility of Macroeconomic Fluctuations," *American Economic Review*, 94(1): 190-217.
- Justiniano A., G. Primiceri, and A. Tambalotti (2013), "Is There a Trade-Off Between Inflation and Output Stabilization?," *American Economic Journal: Macroeconomics*, 5(2): 1-31.
- Kass, R.E. and A.E. Raftery (1995), "Bayes Factors," *Journal of the American Statistical Association*, 90(430): 773-795.
- Klenow, P. and B.A. Malin (2011), "Microeconomic Evidence on Price-Setting," *Handbook of Monetary Economics Volume 3A*, B.J. Friedman and M. Woodford eds. Elsevier Science, North-Holland: 231-284.
- Laforte, J.P. (2007), "Pricing Models: A Bayesian DSGE Approach for the U.S. Economy," *Journal of Money, Credit and Banking*, 39(s1): 127-154.
- Lindé J. (2005), "Estimating New-Keynesian Phillips curves: A Full Information Maximum Likelihood Approach," *Journal of Monetary Economics*, 52(6): 1135-1149.
- Lubik, T.A. and F. Schorfheide (2006), "A Bayesian Look at New Open Economy Macroeconomics," *NBER Macroeconomics Annual 2005*, M. Gertler and K. Rogoff eds. The MIT Press, Cambridge: 316-366.
- Lucas, R.E. (1976), "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy*, 1: 19-46.
- Mankiw, N.G. and R. Reis (2002), "Sticky Information Versus Sticky Prices: A Proposal To Replace The New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 117(4): 1295-1328.
- Mash, R. (2004), "Optimizing Microfoundations for Inflation Persistence," *Economics Series Working Papers*, University of Oxford, Department of Economics, No. 183.
- Mavroeidis, S. (2005), "Identification Issues in Forward-Looking Models Estimated by GMM, with an Application to the Phillips Curve," *Journal of Money, Credit and Banking*, 37(3): 421-448.
- Nakamura, E. and J. Steinsson, (2013), "Price Rigidity: Microeconomic Evidence and Macroeconomic Implications," *Annual Review of Economics*, 5(1): 133-163.
- Olivei, G. and S. Tenreyro (2010), "Wage-Setting Patterns and Monetary Policy: International Evidence," *Journal of Monetary Economics*, 57(7): 785-802.
- Papke, L.E. and J.M. Wooldridge (2005), "A Computational Trick for Delta-method Standard Errors," *Economics Letters*, 86(3): 413-417.
- Rabanal, P. and J.F. Rubio-Ramirez (2005), "Comparing New Keynesian Models of the Business Cycles: A Bayesian Approach," *Journal of Monetary Economics*, 52(6): 1151-1166.

- Riggi, M. and M. Tancioni (2010), “Nominal vs. Real Wage Rigidities in New Keynesian Models with Hiring Costs: A Bayesian Evaluation,” *Journal of Economic Dynamics and Control*, 34(7): 1305-1324.
- Rudd, J. and K. Whelan (2005), “New Tests of the New-Keynesian Phillips Curve,” *Journal of Monetary Economics*, 52(6): 1167-1181.
- Schorfheide, F. (2000), “Loss Function-Based Evaluation of DSGE Models,” *Journal of Applied Econometrics*, 15(6): 645-670.
- Sheedy, K.D. (2007), “Intrinsic Inflation Persistence,” *Centre for Economic Performance Discussion Paper*, No. 837.
- Sheedy, K.D. (2010), “Intrinsic Inflation Persistence,” *Journal of Monetary Economics*, 57(8): 1049-1061.
- Smets, F. and R. Wouters (2003), “An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area,” *Journal of the European Economic Association*, 1(5): 1123-1175.
- Smets, F. and R. Wouters (2007), “Shock and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach,” *American Economic Review*, 97(3): 586-606.
- Taylor, J.B. (1980), “Aggregate Dynamics and Staggered Contracts,” *Journal of Political Economy*, 88(1): 1-23.
- Taylor, J.B. (1993), “Discretion versus Policy Rules in Practice,” *Carnegie-Rochester Conference Series on Public Policy*, 39: 195-214.
- Tsoukis C., G. Kapetanios, and J. Pearlman (2009), “Elusive Persistence: Wage and Price Rigidities, the New Keynesian Phillips Curve and Inflation Dynamics,” *Journal of Economic Surveys*, 25(4): 737-768.
- Wolman, A.L. (1999), “Sticky Prices, Marginal Cost, and the Behaviour of Inflation,” *Federal Reserve Bank of Richmond Economic Quarterly*, 85(4): 29-48.
- Woodford, M. (2003), “Interest and prices: Foundations of a theory of monetary policy,” *Princeton University Press*.
- Yao, F. (2011), “Non-constant Hazard Function and Inflation Dynamics,” *SFB 649 Discussion Papers, 2009-030*, Humboldt University, Berlin, Germany.